Development of genetic tools for metabolic engineering of *Clostridium pasteurianum*

by

Michael Pyne

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Chemical Engineering

Waterloo, Ontario, Canada, 2014

Author’s declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Reducing the production cost of industrial biofuels will greatly facilitate their proliferation and co-integration with fossil fuels. The cost of feedstock is the largest cost in most fermentation bioprocesses and therefore represents an important target for cost reduction. Meanwhile, the biorefinery concept advocates revenue growth through complete utilization of by-products generated during biofuel production. Taken together, the production of biofuels from low-cost crude glycerol, available in oversupply as a by-product of bioethanol production, in the form of thin stillage, and biodiesel production, embodies a remarkable opportunity to advance affordable biofuel development. However, few bacterial species possess the natural capacity to convert glycerol as a sole source of carbon and energy into value-added bioproducts. Of particular interest is the anaerobe *Clostridium pasteurianum*, the only microorganism known to convert glycerol alone directly into butanol, which currently holds immense promise as a high-energy biofuel and bulk chemical. Unfortunately, genetic and metabolic engineering of *C. pasteurianum* has been fundamentally impeded due to a complete lack of genetic tools and techniques available for the manipulation of this promising bacterium. This thesis encompasses the development of fundamental genetic tools and techniques that will permit extensive genetic and metabolic engineering of *C. pasteurianum*.

We initiated our genetic work with the development of an electrotransformation protocol permitting high-level DNA transfer to *C. pasteurianum* together with accompanying selection markers and vector components. The CpaAI restriction-modification system was found to be a major barrier to DNA delivery into *C. pasteurianum* which we overcame by *in vivo* methylation of the recognition site (5'-CGCG-3') using the M.FnuDII methyltransferase. Systematic investigation of various parameters involved in the cell growth, washing and pulse delivery, and outgrowth phases of the electrotransformation procedure significantly elevated the electrotransformation efficiency up to 7.5×10^4 transformants μg^{-1} DNA, an increase of approximately three orders of magnitude. Key factors affecting the electrotransformation efficiency include cell-wall-weakening using glycine, ethanol-mediated membrane solubilization, field strength of the electric pulse, and sucrose osmoprotection.

Following development of a gene transfer methodology, we next aimed to sequence the entire genome of *C. pasteurianum*. Using a hybrid approach involving 454 pyrosequencing, Illumina dye sequencing, and single molecule real-time sequencing platforms, we obtained a

near-complete genome sequence comprised of 12 contigs, 4,420,100 bp, and 4,056 candidate protein-coding genes with a GC content of 30.0%. No extrachromosomal elements were detected. We provide an overview of the genes and pathways involved in the organism's central fermentative metabolism.

We used our developed electrotransformation procedure to investigate the use of established clostridial group II intron biology for constructing chromosomal gene knockout mutants of *C. pasteurianum*. Through methylome analysis of *C. pasteurianum* genome sequencing data and transformation assays of various vector deletion constructs, we identified a new Type I restriction-modification system that inhibits transfer of vectors harboring group II intron gene knockout machinery. We designated the new restriction system CpaAII and proposed a recognition sequence of 5'-AAGNNNNNCTCC-3'. Overcoming restriction by CpaAII, in addition to low intron retrohoming efficiency, allowed the isolation of a gene knockout mutant of *C. pasteurianum* with a disrupted CpaAI Type II restriction system. The resulting mutant strain should be efficienty transformed with plasmid DNA lacking M.FnuDII methylation.

Lastly, we investigated the use of plasmid-based gene overexpression and chromosomal gene downregulation to alter gene expression in *C. pasteurianum*. Using a β-galactosidase reporter gene, we characterized promoters corresponding to the ferredoxin and thiolase genes of *C. pasteurianum* and show that both promoters permitted high-level, constitutive gene expression. The thiolase promoter was then utilized to drive transcription of an antisense RNA molecule possessing complementarity to mRNA of our β-galactosidase reporter gene. Our antisense RNA system demonstrated 52-58% downregulation of plasmid encoded β-galactosidase activity throughout the duration of growth. In an attempt to perturb the central fermentative metabolism of *C. pasteurianum* and enhance butanol titers, we prepared several antisense RNA constructs for downregulation of 1,3-propanediol, butyrate, and hydrogen production pathways. The resulting downregulation strains are expected to exhibit drastically altered central fermentative metabolism and product distribution.

Taken together, we have demonstrated that *C. pasteurianum* is amendable to genetic manipulation through the development of methods for plasmid DNA transfer and gene overexpression, knockdown, and knockout. Further, our genome sequence should provide valuable nucleotide sequence information for the application of our genetic tools. Thus, the genome sequence, electrotransformation method, and associated genetic tools and techniques

reported here should promote extensive genetic manipulation and metabolic engineering of this biotechnologically important bacterium.

Keywords: Antisense RNA, biofuels, butanol, *Clostridium*, genetic engineering, glycerol, knockout, metabolic engineering, overexpression, restriction, transformation

Acknowledgements

I am infinitely grateful to have Professor C. Perry Chou, Professor Duane Chung, and Professor Murray Moo-Young as my PhD supervisors. They have pushed me immensely over the past 6 years, but I learned very early on that they want nothing more than for their students to succeed. It was both refreshing and valuable to have supervisors with such different scientific approaches and outlooks. Professor Chung – from you I learned that curiosity and skepticism are just as important as intellect. I have benefitted greatly from your entrepreneurial viewpoint. Professor Chou – you demonstrated that work ethic, attention to detail, and modesty are essential to a successful career in academia. If I inherited even a fraction of these traits from you, I will surely succeed in my career. Professor Moo-Young – you showed me first-hand of what hard work and ambition can achieve in science. I am honoured to have been your student.

I am greatly indebted to my committee members – Professor Frank Gu, Professor Bill Anderson, and the late Professor Jeno Scharer from the Department of Chemical Engineering and Professor Trevor Charles from the Department of Biology – for attending my comprehensive examination and defense, and providing valuable suggestions and comments. I am also honoured to have had Professor E. Terry Papoutsakis from the University of Delaware serve as the external examiner and share his expertise at my oral defense.

I owe sincere thanks to our collaborators, Professor Steven Brown and Sagar Utturkar at Oak Ridge National Laboratories in Oak Ridge, Tennessee, for their genome sequencing contributions. I would also like to thank Dr. Jiri Perutka (University of Texas at Austin), Scott Binford (Pacific Biosciences), and Dr. Bill Jack and Dr. Geoffrey Wilson (New England Biolabs) for their valuable input and assistance. This work would also not have been possible without plasmid and strain donations from Professor E. Terry Papoutsakis (University of Delaware), Professor Nigel Minton (University of Nottingham), Professor Sheng Yang (Shanghai Institutes for Biological Sciences), Professor George Bennett (Rice University), and New England Biolabs. I acknowledge financial support from Natural Sciences and Engineering Research Council (NSERC) of Canada and the Canada Research Chair program.

I would also like to acknowledge the hard-working staff and faculty members of the Department of Chemical Engineering, especially Rose Guderian, Judy Caron, Liz Bevan, Ingrid Sherrer, Camille Graham, Rick Hecktus, and Bert Habicher.

I cannot think of anyone better than Karan Sukhija to have started grad school with 6 years ago. This work would not have been possible without the help of Karan and all of my other past and present lab mates – Niju, Shreyas, Victoria, Weifang, Val, Kajan, Fred, Steve, Daryoush, Rana, Lu, Saideh, Bahram, Adam, Lamees, Yasmin, Will, Ziggy, Ian, Mark, Aaqil, and Nimhani. These people made the lab a fun, helpful, and cooperative environment. I also cannot envision a successful laboratory without the dedication of co-op students. Although there have been too many over the years to list, I am grateful for the hard work provided by all of our co-op students, especially Saad, Bensen, Nagma, Kelly, and Pullkit for donating two semesters of their time.

On a personal note, I would like to thank my friends and family for their love, encouragement, and laughs. I cannot say enough about my parents and their support. They instilled in me the value of hard work, dedication, and puns. I also owe a lot to my siblings, Jenn and Lonnie, for their support and friendly hazing. Lastly, I owe all of my grad school successes to Summer, whom I met only four months before I began at the University of Waterloo. She has been by my side for all of the quick trips to the lab and weekends spent writing at coffee shops. Thank you for understanding that “making potions” is not a 9 to 5 job and for lifting me up when nothing seemed to be going right.

To Summer, Hudson, and my amazing parents.

Contents

List of figures

List of tables

List of abbreviations

AB	Acetone-butanol
ABE	Acetone-butanol-ethanol
ACE	Allele-coupled exchange
Ap	Ampicillin
asRNA	Antisense RNA
BuOH	Butanol
Cm	Chloramphenicol
Dam	DNA adenine methylase
Dcm	DNA cytosine methylase
Em	Erythromycin
EPB	Electroporation buffer
ET	Electrotransformation
EtOH	Ethanol
FOA	5-Fluoroorotic acid
GC content	Guanine+cytosine content
Gly	Glycine
IEP	Intron-encoded protein
Ll.ltrB	Lactococcal group II intron
m6A	6-Methyladenine
m4C	4-Methylcytosine
m5C	5-Methylcytosine
MCS	Multiple cloning site
mRNA	Messenger RNA
NGS	Next generation sequencing
NR	Non-recombinant
ORF	Open reading frame
PCR	Polymerase chain reaction
1,3-PDO	1,3-Propanediol
RAM	Retrotransposition-activated marker
REN	Restriction endonuclease
RM	Restriction-modification (methylation)
RBS	Ribosome binding site
rpm	Revolutions per minute
rRNA	Ribosomal RNA
SMP	Sucrose-magnesium-phosphate buffer
SMRT sequencing	Single molecule real-time sequencing
SOE PCR	Splicing by overlap extension PCR
sRNA	Small non-coding RNA
Suc	Sucrose
Thr	Threonine
Tm	Thiamphenicol
UV	Ultraviolet
WT	Wild-type

CHAPTER 1

Introduction

Due to global energy demands, rising environmental concerns, and diminishing oil reserves, investigations into renewable sources of fuel have led to the development and exploitation of numerous bioprocesses that are environmentally viable, yet currently lack economic feasibility. Biofuels, including bioalcohols, biodiesel, and biohydrogen, are at the forefront of this global impetus and represent the most promising forms of renewable energy. Beginning in the 1970s with the onset of the global oil crisis, interest in the large-scale production of clean and renewable biofuels has been vastly increasing. Bioethanol, largely produced from the microbial fermentation of corn-based substrates, currently represents the largest biofuel market in the U.S., with production reaching 7.5 billion gallons in 2007 [1] of a total 12.5 billion gallons worldwide [2]. Recently, biodiesel, or fatty acid alkyl esters, has also emerged as a promising alternative to petroleum-based diesel, with global production totaling 5 billion gallons in 2007 [2]. While the bioethanol and biodiesel markets continue to expand, still other potential biofuels, although currently less developed than bioethanol and biodiesel, have received significant attention over the past three decades. Butanol, a high-energy four-carbon alcohol, is amongst these biofuels, as it possesses physicochemical properties that more closely resemble those of conventional gasoline.

Fermentative production of butanol represents one of the oldest industrial bioprocesses. Prior to the advent of the petrochemical industry, butanol was produced at an industrial scale through the acetone-butanol-ethanol fermentation (ABE fermentation) of substrates such as maize, wheat, and molasses [3,4]. However, fluctuating feedstock costs and the establishment of petrochemical processes during the 1950s led to the eventual downfall of bulk biobutanol production. Currently, the majority of the world's butanol supply is produced through chemical processes such as the oxo process, in which propylene is converted to butyraldehyde and subsequently hydrogenated to butanol [5]. A potential revival of the ABE fermentation process for industrial biobutanol production has been proposed in response to the mounting environmental and economic issues surrounding the use of petroleum-based fuels. However, many process shortcomings have yet to be resolved, including high feedstock and product recovery costs, microbial solvent toxicity, and low butanol titers [6]. While the revival of a large-

scale ABE fermentation process depends largely upon the resolution of such fundamental shortcomings, selection and optimization of the host organism is critical to the future success of butanol as a renewable fuel source.

In nature, butanol is produced solely by species from the bacterial genus *Clostridium* [7]. While most clostridia produce large amounts of butyrate, certain species are able to produce butanol as the major fermentation product, partially through reassimilation of acetate and butyrate. As natural solventogenic organisms, the prospect of increasing butanol production through genetic and metabolic engineering techniques has been an area of intense interest in the past decades [8]. Most efforts involving biobutanol to date have focused on the model solventogenic *Clostridium*, *C. acetobutylicum* ATCC 824. The development of a host-vector system [9], electrotransformation procedure [10], gene reporter system [11,12], antisense RNA strategy [13], chromosomal gene knockout system [14,15], and a complete genome sequence [16] have all been recently achieved with *C. acetobutylicum* ATCC 824. However, genetic engineering techniques are far less advanced in other solventogenic clostridia. Of particular significance, *C. pasteurianum* has been shown to be an ideal host for biobutanol production [17]], particularly due to its unique capacity to ferment waste biodiesel-derived glycerol [18], yet no genetic engineering attempts have been documented with this organism. It is widely held that genetic engineering and strain construction will play a key role in the development of an efficient biobutanol process and it would be interesting to determine if the genetic tools made available for *C. acetobutylicum* and similar organisms can be directly applied towards the genetic engineering of *C. pasteurianum*.

This thesis investigates the genetic tractability of *C. pasteurianum* with respect to application of genetic tools and methodologies established in *C. acetobutylicum* and related species. The successful development of genetic methods in closely-related clostridial species has demonstrated the utility of genetic and metabolic engineering of the solventogenic clostridia. The development of a host-vector system and transformation procedure will be investigated first and, if successful, will lead to a range of genetic techniques. These include gene overexpression, knockdown, and knockout. To further improve the potential of *C. pasteurianum* and provide necessary gene sequencing data, we aim to sequence and annotate the organism's genome using next-generation sequencing technologies. Collectively these tools should contribute to the

advancement of *C. pasteurianum* as a potential future industrial host organism for the production of biobutanol from low-value, crude glycerol.

1.1. Research objectives

The chief objectives of this research can be described as follows:

1. Develop a host-vector system and electroporation-mediated method of gene transfer for *C. pasteurianum*. This objective includes implementing a methylation strategy for the protection of plasmid DNA from restriction by *C. pasteurianum* and subsequent optimization of electrotransformation efficiency to a level equal to or above that of related clostridia (10^4-10^6 transformants μg^{-1} DNA).
2. Sequence, assemble, and annotate the entire genome sequence for *C. pasteurianum* to obtain sequence information for genes involved in the central fermentative metabolism of *C. pasteurianum*, which is required for the application of genetic and metabolic engineering techniques. If a complete genome cannot be attained, a draft genome assembly comprised of a small number of contigs is desirable.
3. Apply established clostridial group II intron technology to generate chromosomal gene knockout mutants of *C. pasteurianum*. This objective includes overcoming an uncharacterized restriction barrier, via methylome analysis, that hinders electrotransformation of plasmids harboring group II intron machinery.
4. Utilize a β-galactosidase gene reporter system to investigate plasmid-based gene overexpression and asRNA-mediated downregulation of chromosomal genes in *C. pasteurianum*. Apply the devised asRNA strategy to generate *C. pasteurianum* mutant strains with altered fermentative metabolism.

1.2. Thesis organization

This thesis is comprised of five manuscripts corresponding to Chapters 2-6. Original research manuscripts make up Chapters 3-6 and are preceded by a literature review manuscript in Chapter 2. Chapter 2 has been submitted for publication, Chapter 3 has been published, and Chapters 4-6 are in preparation.

Chapter 2 provides a comprehensive summary of the genus *Clostridium*, including fermentative butanol production and biotechnological progress made in clostridial genetic tool

development. *C. pasteurianum* is also discussed as a potential future host for the conversion of crude glycerol into butanol. Relevant strain engineering techniques, from genome sequencing and establishment of a gene transfer methodology through to deployment of advanced genome editing procedures, are discussed in detail to provide a blueprint for genetic manipulation of *C. pasteurianum*.

Chapter 3 encompasses the first step for genetic manipulation of *C. pasteurianum* through the development of a host-vector system and high-level transformation methodology. We initiated this work by investigating common antibiotics utilized in the clostridia for their effectiveness to inhibit growth of *C. pasteurianum*. We then constructed various *E. coli*-*C. pasteurianum* shuttle vectors and achieved low level transformation using electroporation and a methylation strategy to overcome host restriction. Systematic improvements were then made to the procedure to attain high level electrotransformation of *C. pasteurianum*.

Chapter 4 details the steps and collaborations undertaken to obtain a near-complete, fully-annotated genome sequence of *C. pasteurianum*. The resulting draft assembly is discussed with respect to relevant phenotypes and metabolic characteristics.

Chapter 5 provides a detailed description of the work involved in the development of a gene knockout system for *C. pasteurianum*. Methylome analysis from *C. pasteurianum* genome sequencing data illuminated a unique Type I restriction-modification system that inhibits electrotransformation of gene knockout vectors harboring group II intron machinery. Overcoming this restriction barrier and low gene knockout efficiency allowed, for the first time, isolation of gene disruption mutants of *C. pasteurianum*.

Chapter 6 investigates the utilization of our devised electrotransformation procedure for the application of gene overexpression and knockdown techniques. A thermophilic β-galactosidase enzyme was utilized as a reporter system to quantify the strength of native *C. pasteurianum* gene promoters. A strong, constitutive promoter was then selected to drive transcription of an antisense RNA molecule targeting the β-galactosidase reporter gene. After successful demonstration of β-galactosidase gene downregulation, we selected endogenous *C. pasteurianum* chromosomal genes involved in competing metabolic pathways for asRNA-mediated gene knockdown.

Lastly, Chapter 7 provides a summary of the genetic and metabolic work performed throughout the entirety of this thesis. Recommendations and future directions are provided as a

means of further developing *C. pasteurianum* as a potential host for the production of biobutanol from cheap and abundant crude glycerol. Emphasis is placed here on future targets for genetic and metabolic engineering using the results gleaned from this thesis.

CHAPTER 2

Literature review

2.1. Butanol as a prospective biofuel

As a higher alcohol, butanol possesses many physicochemical properties that make it an attractive alternative to petroleum fuel. Compared to ethanol, butanol is less corrosive, volatile, and hydroscopic, which, unlike ethanol, allows butanol to adapt well to current fuel transportation infrastructure [5]. Due to the increased carbon content of butanol in comparison to ethanol, butanol also possesses more desirable fuel characteristics, such as higher energy density and air-fuel ratio, in addition to research and motor octane numbers that more closely resemble those of gasoline [7]. Furthermore, the 4-carbon backbone of butanol closely mimics the 4- to 12-carbon hydrocarbons that make up petroleum gasoline [19]. Alternatively, ethanol is severely limited as an additive or blending agent, since most North American automobiles can only handle blends up to 10% (v/v) ethanol [5]. For these reasons, butanol is much more attractive as an additive and even as a replacement to conventional fuel, with most current automobile engines capable of utilizing blends up to 100% (v/v) butanol without modification to the engine or fuel system [20]. In fact, an unmodified 1992 Buick Park Avenue was driven across the United States in 2005 with 100% butanol as the sole fuel source (http://www.butanol.com/). Although fuel consumption increased by 9% compared to gasoline, combustion of butanol was much cleaner, with a 95% and 37% reduction in hydrocarbon and NO_x emissions, respectively. While the feasibility of butanol as a fuel source has clearly been demonstrated, butanol is utilized predominantly as a bulk solvent, rather than a fuel. Further, butanol is produced petrochemically and, therefore, is unsustainable. Although currently not economically feasible, butanol can be produced biologically (biobutanol) and sustainably solely by members of the genus *Clostridium*.

2.2. Clostridial fermentative butanol production

2.2.1. Overview of the genus Clostridium

Clostridium is one of the largest genera of prokaryotes and is comprised of approximately 200 distinct species of bacteria [21]. Prior to the advent of 16S rRNA sequencing, subscription to

the genus was traditionally granted based on the fulfillment of three simple criteria, namely the capacity to form endospores, an obligately anaerobic metabolism, and an inability to perform sulphate reduction [22]. In practice, the clostridia are more generally recognized by distinct rod-shaped cells, the presence of Gram-positive cell walls, and DNA of low guanine+cytosine content (GC content; typically 26-32%), although these characteristics are not strict taxonomic markers. As a result of such simple classification requirements, *Clostridium* is notorious for its extreme phylogenetic heterogeneity and possesses more genetic diversity than almost all other microbial genera [21]. In fact, *Clostridium* contains numerous uncharacteristic phenotypes, including species that are cocci-shaped, asporogenous, aerotolerant, intermediate in GC content, and Gram-negative or Gram-variable [23]. As such, 16S rRNA cataloguing and next generation genome sequencing have led to several massive proposed rearrangements of *Clostridium* [24,25,26,27]. Only 73 species were found to possess sufficient relatedness to the type species, *C. butyricum*, leading to the proposal of a new core genus, termed *Clostridium sensu strict*o [21,28]. The heterogeneity of the genus allows *Clostridium* as a whole to thrive within diverse habitats, as the clostridia are truly ubiquitous in nature. Owing to their ability to form endospores that are resistant to oxygen, heat, dessication, acid, and alcohol, dormant spores of *Clostridium* are commonly found within soil, aquatic sediments, intestinal tracts of mammals, and unpasteurized or spoiled foods [4]. From this rich diversity many species of *Clostridium* have arisen in recent years in both the medical and industrial sectors of biotechnology.

Clostridium can be divided into pathogenic and apathogenic species. Some of the most notorious and potent human toxins are produced by the pathogenic clostridia, which encompasses *C. botulinum*, *C. difficile*, *C. perfringens*, *C. septicum*, and *C. tetani* [29]. *C. botulinum* and *C. tetani* produce two of the most potent neurotoxins and are the causative agents of the well-known diseases botulism and tetanus, respectively. The botulism toxin is the most toxic agent currently known to humans, as type H toxin was recently found to have a lethal dose of 2-13 ng for an average-sized human [30]. The advent of proper food handling and pasteurization practices in the case of botulism, and access to effective vaccination programs for the prevention of tetanus, have drastically diminished the prevalence of these fatal diseases within the developed world. Still, an estimated 58,000 fatal cases of newborn tetanus were reported globally in 2010 and, as of 2012, 31 countries have yet to effectively eliminate maternal and neonatal tetanus (World Health Organization; http://www.who.int/). *C. difficile*, on the other

hand, produces two highly virulent exotoxins and is the leading cause of antibiotic-associated diarrhea and nosocomial infection [31]. Dormant spores of *C. difficile* often survive routine hospital sanitation practices, including exposure to alcohol-based hand rubs, and can be ingested through the fecal-oral route, allowing proliferation in the small intestine of patients with compromised gut flora, typically resulting from antibiotic treatment. Infection by *C. difficile* causes a range of intestinal maladies, including severe diarrhea, pseudomembranous colitis, septicemia, and death [29]. The occurrence of healthcare-associated *C. difficile* infection has increased drastically over the past decade and has risen to epidemic status [32,33]. Other pathogenic clostridia, particularly *C. perfringens* and *C. septicum*, are associated with food poisoning, gas gangrene, and meningitis, among other ailments [29]. It is estimated that *C. perfringens* accounts for more than 10% of the 9.4 million annual cases of foodborne illness in the United States [34].

2.2.2. Overview of fermentative metabolism

Within the apathogenic clostridia, several species have garnered immense interest in the field of industrial biotechnology as a result of their diversity of substrate utilization and unique metabolic capabilities [35]. While many clostridia produce standard fermentation products, including organic acids and carbon dioxide and hydrogen gases, a number of species produce varying amounts of alcohols and solvents, such as acetone, ethanol, 1,3-propanediol, isopropanol, and butanol, which have industrial potential as bulk solvents and prospective biofuels [22]. In fact, the exploitation of clostridia for large-scale production of commodity chemicals represents one of the first worldwide industrial bioprocesses [4]. Prior to the dominance of the current petrochemical industry, large-scale production of acetone, and later butanol, as solvents (AB fermentation) was carried out by species of *Clostridium*. Fluctuating costs of molasses and maize feedstocks, coupled to the establishment of more economical petrochemical processes during the 1950s, however, led to the eventual downfall of fermentative AB production. A potential revival of the industrial AB fermentation for the production of butanol, this time as a promising biofuel, is currently underway in response to the mounting environmental and political issues surrounding the production and consumption of petroleum-based fuels [3]. Unfortunately, many process shortcomings have yet to be resolved, including high feedstock costs, poor solvent yields, and product toxicity [6]. The revival of a competitive

AB fermentation process depends upon the resolution of these fundamental issues. However, it is the capacity to genetically manipulate the solventogenic clostridia that determines the future success of clostridial production of bulk solvents and biofuels [8].

Historically, the exploitation of *Clostridium* for the industrial production of acetone and butanol has centered around *C. acetobutylicum* and *C. beijerinckii* [4], and has led to the widespread investigation of these species with respect to genetics, physiology, and growth and metabolism. Numerous other *Clostridium* species, however, have garnered significant attention in recent years, largely the result of their unique and promising metabolic capabilities [35]. Examples include production of biofuels (ethanol or butanol) from diverse substrates, such as cellulose (e.g., *C. cellulolyticum* and *C. thermocellum* [36]), crude biodiesel-derived glycerol (*C. pasteurianum* [18]), and synthesis gas (e.g., *C. ljungdahlii* and *C. carboxidivorans* [37]). Solventogenic *Clostridium* species exhibit a unique biphasic fermentative metabolism. In the first stage, acidogenesis, exponential growth is coupled to vigorous acid (acetate and butyrate) and gas production (H_2 and CO_2) [38]. As the concentration of acids approaches deleterious levels, the culture pH drops significantly and the cells enter stationary phase. At this point, a metabolic shift occurs in which vigorous growth and acid production cease, as the cells enter the solvent-producing, or solventogenic, phase. Solventogenesis is characterized by inhibition of the acid and H_2 production pathways, lack of growth and cell division, and induction of solvent-producing pathways, stress response, and endospore formation [38]. Solvents such as acetone, ethanol, and butanol are partially produced through the uptake and reutilization of acetic and butyric acids accumulated during acidogenesis. The metabolic pathways associated with acidogenesis and solventogenesis are discussed below (Figure 2.1).

2.2.3. Acidogenic pathways

During acidogenesis, clostridial species metabolize hexose sugars through the Embden-Meyerhof-Parnas (EMP) pathway, generating two moles each of pyruvate, NADH, and ATP per mole of glucose [38]. The resulting pyruvate is further oxidized by pyruvate-ferredoxin oxidoreductase, rather than pyruvate dehydrogenase or pyruvate decarboxylase, through the phosphoroclastic reaction, yielding acetyl-CoA and CO_2 [39]. As with pyruvate dehydrogenase and pyruvate decarboxylase, thiamine pyrophosphate (TPP) is an essential coenzyme in the pyruvate-ferredoxin oxidoreductase reaction by transferring the C-2 hydroxyethyl moiety from

pyruvate to CoA [39]. Cleavage of pyruvate involves a redox reaction in which two hydrogen are transferred to ferredoxin, as opposed to NAD^+, as is the case with pyruvate dehydrogenase. Oxidized ferredoxin is then regenerated through the production of H_2, catalyzed by a hydrogenase. This series of reactions can be summarized as follows [39]:

$$\text{pyruvate} + \text{TPP-oxidoreductase} \rightleftharpoons \text{hydroxyethyl-TPP-oxidoreductase} + CO_2$$

$$\text{hydroxyethyl-TPP-oxidoreductase} + Fd_{ox} + \text{CoA} \rightleftharpoons \text{TPP-oxidoreductase} + \text{acetyl-CoA} + Fd_{red}$$

$$Fd_{red} \rightleftharpoons Fd_{ox} + H_2$$

Acetyl-CoA represents a major branch point in clostridial fermentations [4]. During acidogenesis, the fate of acetyl-CoA involves the production of acetate and butyrate. Since both acetate and acetyl-CoA are 2-C compounds, the acetate pathway involves a simple two-step phosphorylation-dependent conversion analogous to the *E. coli* acetate pathway [40]. Acetyl-CoA is first phosphorylated by acetate kinase (Ack) to generate acetyl-phosphate, which is then converted to acetate through the action of phosphotransacetylase (Pta) with the concomitant generation of ATP through substrate-level phosphorylation:

$$\text{acetyl-CoA} + P_i \xrightarrow{\text{Ack}} \text{acetyl-phosphate} + \text{CoA}$$

$$\text{acetyl-phosphate} + \text{ADP} \xrightarrow{\text{Pta}} \text{acetate} + \text{ATP}$$

As with many acetate-producing organisms, the genes encoding Ack and Pta are located in an operon and are transcribed from the same promoter in *C. acetobutylicum*, with *pta* upstream of *ack* [40]. The acetate pathway shown above generates 2ATP per mole of glucose and, as a result, acetate production represents a major energy-producing pathway in the solventogenic clostridia. However, as the acetate-producing pathway does not consume reducing equivalents (i.e. NADH), the NADH generated during glycolysis must be oxidized in order to restore electron

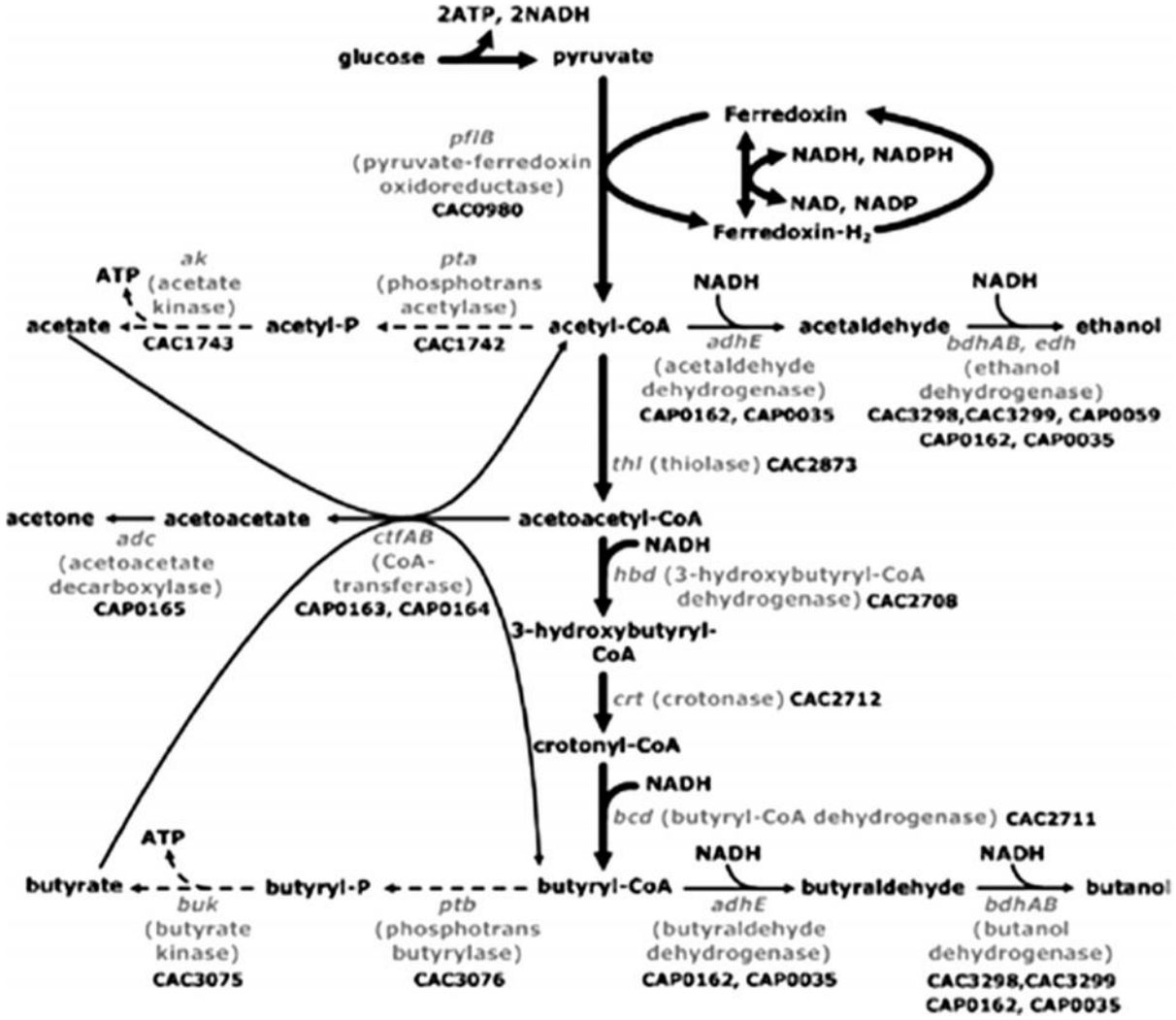


Figure 2.1. Central heterofermentative metabolism of glucose by *C. acetobutylicum* ATCC 824. Dotted lines represent pathways that are active only under acidogenic conditions. Adapted from [7].

balance. One way in which clostridia dispose of excess reducing power is through the generation of H_2 by the combined action of an NADH-ferredoxin oxidoreductase and hydrogenase. In this sense, the amount of acetate (and butyrate, see below) that can be produced is directly related to the ability of the cell to evolve H_2 [38]. However, unless the partial H_2 pressure is kept very low, only modest amounts of H_2 can be evolved. Consequently, free-living clostridia are not able to expose of their excess reducing equivalents through the production of H_2 [39]. In order to cope with excess reducing power following H_2 generation, solventogenic clostridia rely on the production of reduced end products such as butyrate, butanol, and ethanol.

While the acetate pathway generates 2ATP yet has no effect on the electron balance, the butyrate production pathway produces 1ATP and oxidizes 2NADH through the conversion of acetyl-CoA to butyryl-CoA [39]. Butyryl-CoA, a 4-C intermediate, is analogous to acetyl-CoA and is also a major branch point in butyrate production [4]. Butyryl-CoA is produced from acetyl-CoA through the successive action of four enzyme systems, which collectively oxidize 2NADH in the process:

$$2\text{acetyl-CoA} \xrightarrow{(1)} \text{acetoacetyl-CoA} + \text{CoA}$$

$$\text{acetoacetyl-CoA} + \text{NADH} \xrightarrow{(2)} \text{3-hydroxybutyryl-CoA} + \text{NAD}^+$$

$$\text{3-hydroxybutyryl-CoA} \xrightarrow{(3)} \text{crotonyl-CoA} + H_2O$$

$$\text{crotonyl-CoA} + \text{NADH} \xrightarrow{(4)} \text{butyryl-CoA} + \text{NAD}^+$$

where reactions (1) to (4) are catalyzed by thiolase (acetyl-CoA acetyltransferase), 3-hydroxybutyryl-CoA dehydrogenase, crotonase, and butyryl-CoA dehydrogenase, respectively. As is shown, two molecules of acetyl-CoA are condensed in order to convert the 2-C acetyl-CoA into the 4-C butyryl-CoA. In addition, 2NADH are consumed which aids the cell in disposing of excess reducing equivalents not evolved through H_2 production via NADH-ferredoxin oxidoreductase and hydrogenase. Butyryl-CoA represents the end point of the main trunk in the characteristic branched fermentation exhibited by solventogenic clostridia. Depending on many factors, the fate of butyryl-CoA ends with either butanol or butyrate production. As butyryl-CoA is analogous to acetyl-CoA, the conversion of butyryl-CoA follows the same two-step mechanism as the conversion of acetyl-CoA into acetate:

$$\text{butyryl-CoA} + P_i \xrightarrow{\text{Buk}} \text{butyryl-phosphate} + \text{CoA}$$

$$\text{butyryl-phosphate} + \text{ADP} \xrightarrow{\text{Ptb}} \text{butyrate} + \text{ATP}$$

where Buk and Ptb are butyrate kinase and phosphotransbutyrylase, respectively. Although NADH is not consumed directly in the production of butyrate, the conversion of acetyl-CoA to butyryl-CoA oxidizes 2NADH per mole of butyryl-CoA that is generated. As with *pta* and *ack*, the genes encoding Buk and Ptb (i.e. *buk* and *ptb*, respectively), form an operon in *C. acetobutylicum* with *ptb* preceding *buk* [41,42].

Although most solventogenic clostridia possess an inducible NADH-dependent lactate dehydrogenase gene, *ldhA*, lactate is generally only produced under certain limiting conditions [43]. Since lactate is produced from pyruvate rather than acetyl-CoA, this gives the cell a means of disposing excess reducing power if the phosphoroclastic splitting of pyruvate to acetyl-CoA, and therefore the H_2 pathway, is inhibited [38]. Pyruvate is converted to lactate in the same manner as most organisms through the action of lactate dehydrogenase (LdhA):

$$\text{pyruvate} + \text{NADH} \xrightarrow{\text{LdhA}} \text{lactate} + \text{NAD}^+$$

This reaction is unidirectional and only occurs with a buildup of pyruvate due to blockage of the pyruvate-ferredoxin oxidoreductase reaction. In addition, significant lactate production has been observed under stress conditions, such as iron and sulfur limitation [44] and inhibition of hydrogenase activity by carbon monoxide [45], potentially due to a lack of ferredoxin for use in the pyruvate cleavage reaction [38].

2.2.4. Solventogenic pathways

Although the metabolic shift that occurs from the acidogenic phase to the solventogenic phase has been an area of intense research in the past decades, the precise mechanism and factors involved in the trigger have yet to be completely worked out. It is widely accepted that the shift to the solventogenic phase occurs as an adaptive mechanism by the cells in order to cope with the high undissociated butyric acid concentration and low pH resulting from acidogenesis

[46,47,48]. During acidogenesis, the culture pH drops to approximately 3.8-5.0 largely due to extreme undissociated butyric acid concentration [4]. In order to mitigate the deleterious effects of this environment, cells begin to reassimilate the acetic and butyric acids into butanol, ethanol, and acetone. At this point, acid uptake and utilization causes the pH to partially increase to 6.8-7.0 [49]. However, factors other than high acid concentration and low pH have been found to trigger the metabolic shift from acidogenesis to solventogenesis. Through the generation and characterization of a *buk* knockout mutant of *C. acetobutylicum*, it has been proposed that butyryl-phosphate rather than butyrate acts as a global regulator of solventogenesis [50,51]. This is due to the fact that, although butyrate production was significantly decreased, solventogenesis was initiated at an earlier stage in the fermentation and resulted in increased acetone and butanol titers compared to the wild-type strain [50]. Since *ptb* is located upstream of *buk* within the butyrate-pathway operon, Ptb activity was unaffected by the *buk* mutation and butyryl-phosphate was able to accumulate. In addition to a high concentration of butyryl-phosphate, substrate surplus, high levels of ATP compared to ADP, and a large NADH/NAD^+ ratio have all been implicated as key factors involved in the switch from acidogenic to solventogenic metabolism [5,46,52]. As solvents, especially butanol, are more toxic to the cell than butyric and acetic acids, cells respond to solventogenesis through the formation of endospores in preparation for long-term survival [5]. In fact, the onset of solventogenesis is coupled to spore formation through the master transcriptional regulator, *spo0A* [53].

Solventogenic *Clostridium* species produce varying ratios of the solvents acetone, butanol, and ethanol. A typical acetone, butanol, ethanol ratio for a batch *C. acetobutylicum* fermentation is approximately 3:6:1, with butanol as the major fermentation end product [4]. As previously mentioned, both butanol and butyrate are produced from butyryl-CoA through the final branch point in fermentative clostridial metabolism. At the onset of solventogenesis, very high levels of undissociated butyric acid are present which can be taken up by the cell and assimilated into butyryl-CoA through the action of an acetoacetyl-CoA:acetate/butyrate:CoA transferase [54]. This enzyme system acts on both acetate and butyrate for the regeneration of acetyl-CoA and butyryl-CoA, respectively. Acetyl-CoA can then be condensed and siphoned towards acetone and butanol production via acetoacetyl-CoA and butyryl-CoA, respectively, or can be converted directly into ethanol. The most likely path via acid uptake for production of ethanol is:

$$\text{acetate} + \text{CoA} \xrightarrow{\text{CoA transferase}} \text{acetyl-CoA}$$

$$\text{acetyl-CoA} + \text{NADH} \xrightarrow{\text{acetaldehyde DH}} \text{acetaldehyde} + \text{NAD}^+$$

$$\text{acetaldehyde} + \text{NADH} \xrightarrow{\text{EtOH DH}} \text{ethanol} + \text{NAD}^+$$

and butanol production:

$$\text{butyrate} + \text{CoA} \xrightarrow{\text{CoA transferase}} \text{butyryl-CoA}$$

$$\text{butyryl-CoA} + \text{NADH} \xrightarrow{\text{butyraldehyde DH}} \text{butyraldehyde} + \text{NAD}^+$$

$$\text{butyraldehyde} + \text{NADH} \xrightarrow{\text{BuOH DH}} \text{butanol} + \text{NAD}^+$$

As with acetate and butyrate production, ethanol and butanol are produced in analogous pathways. As discussed above, acetic and butyric acids are assimilated into acetyl-CoA and butyryl-CoA, respectively, by acetoacetyl-CoA:acetate/butyrate:CoA transferase (abbreviated CoA transferase). Conversion of acetyl-CoA or butyryl-CoA to the corresponding aldehyde is catalyzed by the respective acetaldehyde or butyraldehyde dehydrogenase (DH) and ethanol (EtOH) or butanol (BuOH) dehydrogenase, respectively. However, the genome sequence of *C. acetobutylicum* ATCC 824 revealed eight putative aldehyde-alcohol dehydrogenases that could act on both 2- and 4-C aldehydes and alcohols [16]. The most versatile aldehyde-alcohol dehydrogenases produced by *C. acetobutylicum* are AdhE1 and AdhE2, encoded by *adhE* (locus tag: CAP0035) and *adhE2* (CAP0162), respectively [55,56,57]. However, these enzymes are active under different conditions, with AdhE1 acting under typical solventogenic conditions and Adhe2 active only under alcohologenic conditions in which acetone production is absent, such as fermentation of more reduced substrates than glucose [5,56]. AdhE1 and Adhe2 are capable of catalyzing both reactions involved in the conversion of either acetyl- or butyryl-CoA into ethanol and butanol, respectively [16]. As a result, these aldehyde-alcohol dehydrogenases are involved in both ethanol and butanol production under either solventogenic or alcohologenic conditions. However, AdhE1 and AdhE2 are not the enzymes responsible for high-level butanol production. Instead, a second set of butanol dehydrogenases, BdhA and BdhB, have been purified and characterized which are more specific to butyraldehyde than acetaldehyde and are unable to

catalyze the conversion of butyryl-CoA to butyraldehyde [58]. While *bdhA* is constitutively expressed and involved only in low-level butanol production, *bdhB* is induced later in solventogenesis than *adhE* and was found to be the predominant enzyme involved in high-level butanol production [59].

As acetyl-CoA and butyryl-CoA represent major branch points for the production of ethanol and butanol, respectively, acetoacetyl-CoA represents a third major branch point for the production of acetone [38]. Acetoacetyl-CoA:acetate/butyrate:CoA transferase, although previously discussed as the enzyme involved in uptake of acetic and butyric acids, also catalyzes the first step in acetone production from acetoacetyl-CoA to yield acetoacetate [54]. To generate acetone, acetoacetate is then decarboxylated by an acetoacetate decarboxylase (Adc) [60]:

$$\text{acetoacetyl-CoA} \xrightarrow{\text{CoA}} \text{acetoacetate} + \text{CoA}$$

$$\text{acetoacetate} \xrightarrow{\text{Adc}} \text{acetone} + CO_2$$

Although acetone is the end product of this pathway in *C. acetobutylicum*, some species of *Clostridium* are capable of reducing acetone further to isopropanol through the use of an isopropanol dehydrogenase [61]. Unlike the acetone pathway, the isopropanol dehydrogenase reaction is coupled to the oxidation of NADH which allows for additional disposal of excess reducing power generated during cell growth.

2.3. Clostridium pasteurianum *as a native butanol producer*

2.3.1. Characteristics of C. pasteurianum

C. pasteurianum is a saccharolytic, nitrogen-fixing species of *Clostridium* historically regarded as a butyric acid producer [39]. However, several early studies revealed the capability of *C. pasteurianum* to produce significant amounts of alcohols, such as butanol, ethanol, and 1,3-propanediol (13-PDO), through the fermentation of various substrates [62,63]. Although largely dictated by the specific culture conditions, the unique fermentative metabolism of *C. pasteurianum* appears to possess similarities to both classical ABE organisms, such as *C. acetobutylicum* and *C. beijerinckii*, and species that produce predominantly butyrate and 1,3-

PDO as major fermentation products, such as *C. butyricum* [17]. As a result, interest has been demonstrated in *C. pasteurianum* as a potential host for the biological production of butanol, 1,3-PDO, and even H_2 [64,65,66].

As most efforts regarding the biotechnological production of butanol have focused on *C. acetobutylicum* as the host organism, investigations into alternative clostridial species, such as *C. pasteurianum*, have been steadily increasing. Much of the interest surrounding this organism is related to the recent overabundance of crude glycerol derived as a byproduct of biodiesel production [18,67]. For every 10 kg of biodiesel produced, 1 kg of crude glycerol is generated. This has led to a drastic decrease in the price of raw glycerol and, consequently, the closure of numerous glycerol refineries throughout the US [68]. Glycerol, once considered a high-value commodity chemical valued at almost $0.60-0.90/lb, has now plummeted to approximately $0.05/lb and is considered an attractive no-cost substrate for the biological production of value-added bulk chemicals [67]. Unlike *C. acetobutylicum*, which is only able to metabolize glycerol when it is supplemented with glucose, and *C. butyricum*, which ferments glycerol but lacks a functional butanol pathway, *C. pasteurianum* is the only organism known to naturally ferment glycerol to butanol [66]. In fact, *C. pasteurianum* is not only able to ferment pure glycerol, but can effectively convert biodiesel-derived waste glycerol into predominantly butanol and 1,3-PDO [18].

2.3.2 Glycerol metabolism

Early studies involving the fermentation of simple hexose sugars by *C. pasteurianum* led to the conclusion that the organism is a classical butyric acid producer [39]. However, it was found that *C. pasteurianum* can produce appreciable amounts of acetone, butanol and ethanol when grown in media of high sugar content (≥6% w/v) [69]. In terms of butanol production, the optimum glucose concentration was found to be 12.5% w/v, yielding a butanol concentration of approximately 10 g/l. More importantly, fermentation of more reduced substrates than glucose, such as glycerol, was first demonstrated with the growth of *C. pasteurianum* on algal biomass supplemented with glycerol [63]. Butanol was found to increase with increasing glycerol concentrations and yielded a maximum concentration of 9.42 g/l with 4% glycerol and a 10-fold concentrate of algal biomass as the substrate. 1,3-PDO often competed with butanol as the main fermentation product. Since this initial finding, several studies have focused on the fermentation

of glycerol to solvents by *C. pasteurianum* [17,18,64,65,66]. The use of glycerol as sole carbon and energy source offers two predominant advantages. First, as previously stated, the widespread development of biodiesel technologies has caused glycerol to become a no-cost waste stream for use in microbial fermentations [67]. Second, due to its high reduction state, glycerol actually favors the production of butanol due to the additional reducing equivalents generated from its breakdown, as compared to less reduced substrates such as glucose [66].

Although fermentative butanol production by *C. pasteurianum* possesses many similarities to the well-characterized pathways of *C. acetobutylicum*, many prominent differences have been observed. As stated above, the most salient feature is the ability of *C. pasteurianum* to directly convert glycerol into butanol as the major fermentation product, in addition to appreciable amounts of 1,3-PDO. In fact, the presence of an active 1,3-PDO pathway is the sole characteristic that affords *C. pasteurianum* the ability to utilize substrates such as glycerol, sorbitol, and mannitol as the sole carbon and energy source. During fermentation on glycerol, most of the substrate metabolized by *C. pasteurianum* enters the glycolytic pathway through the action of glycerol dehydrogenase (glycerol DH) and dihydroxyacetone kinase (DHA kinase):

$$\text{glycerol} + NAD^+ \xrightarrow{\text{glycerol DH}} \text{dihydroxyacetone} + NADH$$

$$\text{dihydroxyacetone} + ATP \xrightarrow{\text{DHA}} \text{dihydroxyacetone phosphate} + ADP + P_i$$

The dihydroxyacetone phosphate produced is then able to proceed through glycolysis to generate pyruvate (Figure 2.2). Since the conversion of glucose to dihydroxyacetone phosphate does not produce any reducing equivalents, the metabolism of one mole of glycerol, equivalent to half a mole of glucose, generates an additional mole of NADH through the glycerol DH reaction. This excess reducing power must be disposed of through the formation of reduced fermentation products such as butanol, ethanol, and 1,3-PDO, the pathways of which all consume NADH. The remaining portion of glycerol that does not enter glycolysis must be converted into 1,3-PDO in order to maintain redox balance through the consumption of NADH:

$$\text{Glycerol} \xrightarrow{\text{glycerol dehydratase}} \text{3-hydroxypropionaldehyde} + H_2O$$

$$\text{3-hydroxypropionaldehyde} + NADH \xrightarrow{\text{1,3-PDO DH}} \text{1,3-propanediol} + NAD^+$$

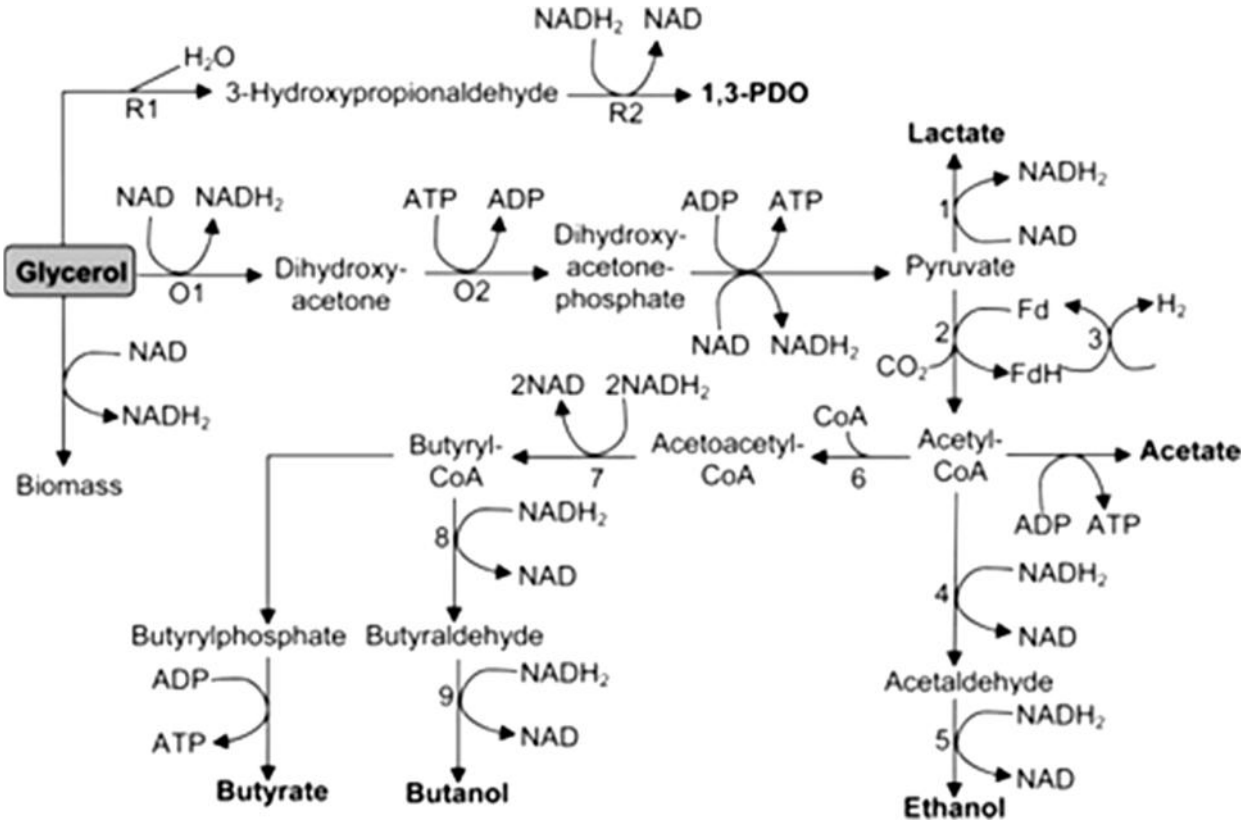


Figure 2.2. Fermentative metabolism of *C. pasteurianum* during growth on glycerol as the sole source of carbon and energy. Enzyme names are as follows: R1, glycerol dehydratase; R2, 1,3-PDO dehydrogenase; 01, glycerol dehydrogenase; 02, dihydroxyacetone kinase; 1, lactate dehydrogenase; 2, pyruvate-ferredoxin oxidoreductase; 3, hydrogenase; 4, acetaldehyde dehydrogenase; 5, ethanol dehydrogenase; 6, thiolase; 7, butyryl-CoA dehydrogenase; 8, butyraldehyde dehydrogenase; 9, butanol dehydrogenase. Adapted from [67].

Therefore, a given amount of 1,3-PDO must always be produced during growth on glycerol as a sole carbon source. Specifically, it has previously been shown that if butanol were to be the only product other than 1,3-PDO following fermentation of glycerol by *C. pasteurianum*, approximately 81.8 and 9.1% of glycerol would be converted to butanol and 1,3-PDO, respectively [66]. The remaining 9.1% of glycerol would be required for the formation of biomass. In contrast, butyrate and acetate each as a major product would only lead to 7.1 and 6.6%, respectively, for the formation of biomass from glycerol. In addition, this would require 50.0 and 64.5% of glycerol used for the production of 1,3-PDO in order to compensate for the lack of reducing equivalents consumed by the butyrate- and acetate-producing pathways, respectively [66]. Therefore, compared to the other reduced compounds produced by *C. pasteurianum* when grown on glycerol, butanol production is favoured by the cell in terms of energy production and biomass formation [67].

2.3.3. General glycerol fermentation characteristics and yields

Although studies involving solvent production from glycerol by *C. pasteurianum* are quite limited, several striking observations have been made. Fermentations by *C. pasteurianum* appear to be characterized by relatively weak regulation of the major metabolic pathways, especially compared to *C. acetobutylicum* in which the central fermentative metabolism and product distribution is under strict control [66]. Under seemingly identical fermentation conditions, product distribution can vary substantially. In addition, pH has repeatedly been found to have essentially no effect on growth rate and product formation over the range of 4.5-7.5 [17,18,66]. This result is in sharp contrast to *C. acetobutylicum*, in which pH drastically alters product distribution, with solvents being favored at low pH [4]. Phosphate limitation produced similar effects in both *C. pasteurianum* and *C. acetobutylicum* through the elevation of solvent production [17,70]. However, this effect was only observed when glycerol was used as a substrate for *C. pasteurianum*. Furthermore, iron limitation favoured lactate production in *C. acetobutylicum* in order to dispose of excess reducing power, whereas *C. pasteurianum* utilized a combination of the lactate and 1,3-PDO pathways in order to maintain redox balance under conditions of iron limitation in which the production of all CoA-dependent products was significantly reduced [17,44]. It was speculated that iron limitation significantly decreased butanol and ethanol production either through the inhibition of alcohol dehydrogenases that are

potentially iron-dependent and/or through decreased flux through the phosphoroclastic reaction due to a non-functional pyruvate-ferredoxin oxidoreductase, an iron-sulfur protein [17]. Perhaps the most efficient strategy for increasing butanol production by *C. pasteurianum* thus far involves the inhibition of hydrogenase by carbon monoxide [17]. It is widely known that clostridial hydrogenases are strongly inhibited by CO, as demonstrated with *C. acetobutylicum*, resulting in an increased flux to the butanol and ethanol pathways in order to dispose of excess reducing power that is typically lost as H_2 [45,71,72]. The inhibition of hydrogenase by either CO or methyl viologen has been shown in vitro using the well-characterized hydrogenase system from *C. pasteurianum* [73,74,75]. In addition, when *C. pasteurianum* was exposed to up to 15% CO during glucose fermentation, levels of lactate, ethanol, and butanol were drastically increased due to hydrogenase inhibition [17]. CO-derived hydrogenase inhibition has yet to be performed using glycerol as substrate, which would likely result in similar results with the exception that 1,3-PDO production would be potentially enhanced due to excessive NADH levels.

A typical ABE fermentation by *C. acetobutylicum* or *C. beijerinckii* utilizing glucose as the sole carbon source generates approximately 0.2-0.4 g solvents/g glucose, with a production distribution of 0.1 g acetone, 0.2 g butanol, and 0.03 g ethanol per gram of glucose [4]. While glucose is not a feasible substrate for fermentation by *C. pasteurianum* due to its high cost and low reduction state, glycerol represents a very attractive substrate for fermentation by *C. pasteurianum*. Although only a few studies have been conducted involving glycerol fermentation by *C. pasteurianum*, butanol yields of 0.24 [17], 0.31 [66], and 0.36 [18] g/g glycerol have been achieved. The average yield of butanol by *C. pasteurianum* during glycerol fermentation is in the range of 0.17-0.20, which corresponds precisely to that of glucose fermentation by *C. acetobutylicum* (0.15-0.20) [18], indicating that both organisms likely possess similar butanol tolerances. When grown on raw biodiesel-derived glycerol, butanol production was largely unaffected compared to pure glycerol, with an average yield of 0.22 g butanol/g raw glycerol [18]. In fact, the highest *C. pasteurianum* butanol yields of 0.31 and 0.36 g/g pure glycerol and 0.30 g/g raw glycerol are equal to or greater than the highest yields achieved to date by classical ABE fermentation of hexose sugars by *C. acetobutylicum* and *C. beijerinckii*. While additional studies are required concerning the key factors involved in high-level butanol production from glycerol by *C. pasteurianum*, the application of genetic and metabolic engineering strategies to

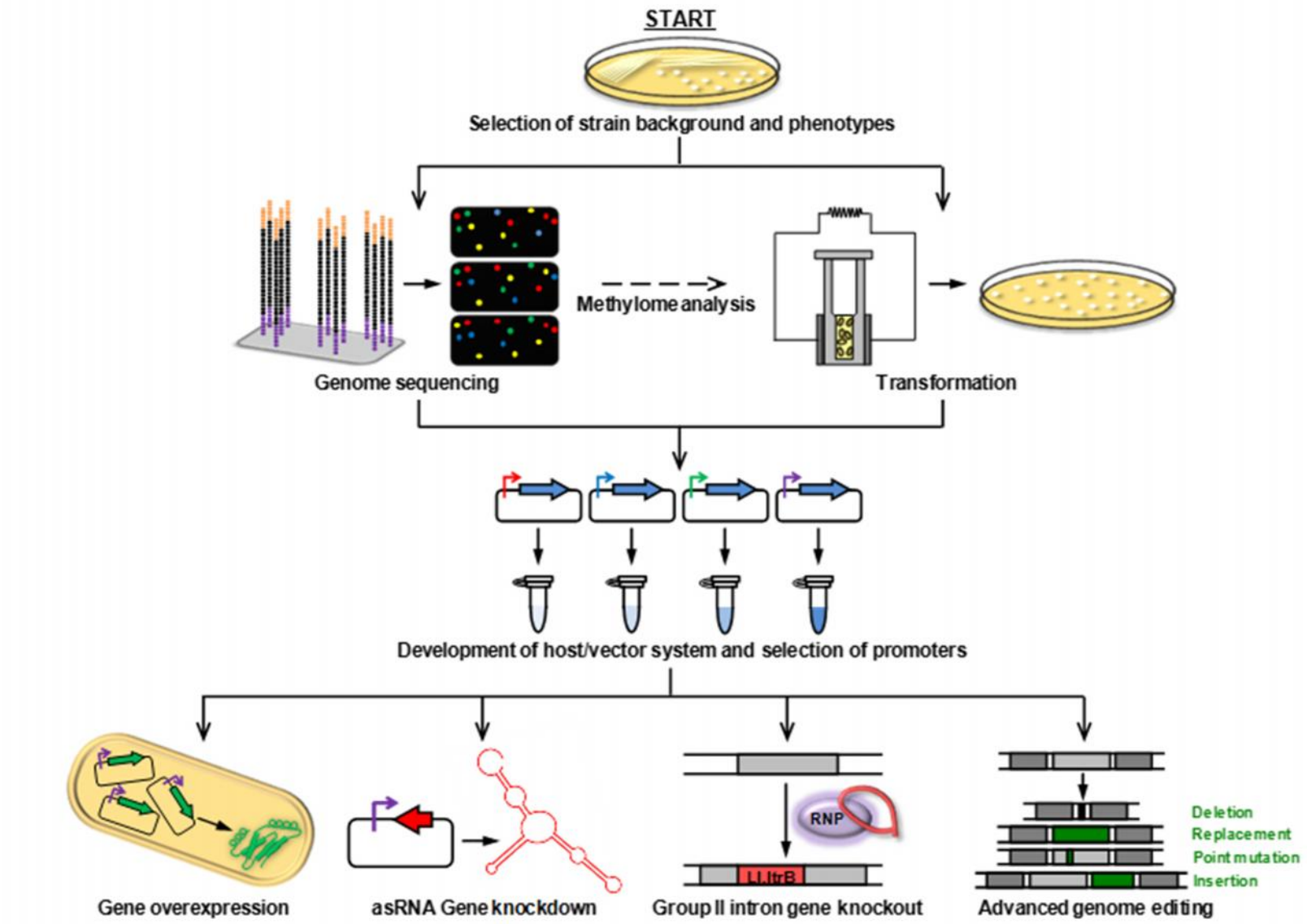
C. pasteurianum would likely lead to significant process enhancements such as decreased acid production and increased growth on crude glycerol.

2.4. Clostridial genetic engineering

The benchmark and model *Clostridium* for all strain construction endeavors is *C. acetobutylicum* ATCC 824. Over the past twenty years, work with this organism has led to the development of a host-vector system, electrotransformation protocol, gene reporter system, antisense-RNA gene knockdown and chromosomal gene knockout methodologies, a fully sequenced genome, and, most recently, chromosomal gene integration procedures [8]. The progression of *C. acetobutylicum* as the model *Clostridium* has followed an iterative, bottom-up approach (Figure 2.3), which has been broadly applied to the advancement of related clostridia in recent years. The major initial hurdle for genetic manipulation of *C. acetobutylicum* involved the identification and subsequent protection of plasmid DNA against a Type II restriction endonuclease, Cac824I [10,76]. With a high-efficiency electrotransformation procedure in place, efforts were then focused on the selection of both constitutive and inducible promoters for plasmid-based gene expression using a gene reporter system [12]. Promoters conferring high-level gene expression were then utilized for antisense-RNA-mediated gene knockdown of key pathways involved in the central metabolism of *C. acetobutylicum* [13,77]. The genome of *C. acetobutylicum* ATCC 824 was fully sequenced and annotated in 2001 [16] and signified a major milestone as the first clostridial genome sequence to be completed. A consistent and reliable method for chromosomal gene disruption didn't appear until 2007 in *Clostridium*, which employs Group II intron technology adapted from methods developed for *E. coli* [14,15]. Very recently, methodologies have arisen detailing the introduction of foreign DNA into clostridial chromosomes based on double-crossover homologous recombination [78,79], representing the current era of genetic engineering in *Clostridium*.

Owing to the ubiquity of clostridia in nature, there is a growing need for the ability to genetically manipulate non-type strains and environmental isolates in order to fully tap into the metabolic potential of the genus. Rather than engineering a heterologous pathway into a common, tractable laboratory organism, it is often advantageous to isolate an environmental

Figure 2.3. Schematic diagram depicting the general workflow involved in clostridial strain development using conventional genetic methods. Genetic techniques are organized corresponding to the chronological order of development for *C. acetobutylicum*. The host strain for manipulation is selected based on its background or phenotypic properties. If genome sequencing data is desired and not available, one or a combination of next generation sequencing platforms are utilized. Methylome analysis can be performed from single-molecule real-time sequencing reads to provide methylation data that can aid in elucidating the recognition sites of putative restriction-modification systems. Gene transfer is then performed based on electrotransformation and, to a lesser extent, conjugation using appropriately methylated DNA. Plasmid transformation opens the door to several genetic techniques, all of which require a host/vector system and access to various gene promoters. These applications are gene overexpression, knockdown, and knockout, as well as advanced genome editing, and are selected for use based on the specific genetic or metabolic engineering objectives.


START
Selection of strain background and phenotypes
Genome sequencing
Methylome analysis
Transformation
Development of host/vector system and selection of promoters
Gene overexpression
asRNA Gene knockdown
RNP
LI.ltrB
Group II intron gene knockout
Deletion
Replacement
Point mutation
Insertion
Advanced genome editing

strain that naturally expresses the target pathway or phenotype at a high level. Since the clostridia possess an extensive substrate-utilization range and produce an array of valuable metabolites, they provide far superior metabolic frameworks for strain development and metabolic engineering efforts [35]. Perhaps the most successful application of this approach involved the genetic development of the Q Microbe® (Qteros, Inc.; http://www.qteros.com/), an environmental isolate of *C. phytofermentans*. The strain was first isolated in 1996 in an attempt to identify superior biomass-degrading, ethanologenic bacteria [80]. In an approach analogous to that of *C. acetobutylicum*, *C. phytofermentans* has advanced as a promising industrial host strain for which an impressive repertoire of genetic methodologies have been developed, including conjugative gene transfer, intron-mediated gene knockout, plasmid-based overexpression, and chromosomal integration of non-replicative vectors [81]. Work with *C. phytofermentans* has thus paved the way for the genetic modification of other novel environmental isolates of *Clostridium*.

The overall state of genetic engineering within the *Clostridium* is sparse given the immense medical and industrial biotechnological potential surrounding many species of *Clostridium.* Of the approximately 200 traditionally-classified species, only a small subset of the clostridia have been probed with genome sequencing and manipulated using gene transfer methods (Table 2.1). Further, numerous landmark advancements have been made since the advent of recombinant DNA technology [8], yet even the most genetically advanced clostridia lag far behind other microbes, especially the model organisms *Escherichia coli*, *Saccharomyces cerevisiae*, and *Bacillus subtilis*, in terms of access to genetic tools and technologies. In fact, only a small collection of clostridia have experienced any significant degree of genetic advancement. Comprehensive genetic work within the genus, including genome sequencing, the development of gene transfer procedures, host-vector systems, gene overexpression, knockout, and knockdown tools, and advanced genome editing technologies, is required for the advancement of lesser known species, non-type strains, and environmental isolates of *Clostridium*. Therefore, in-depth experimental aspects of relevant genetic tools and techniques are discussed below. A schematic template depicting the workflow involved in the development of clostridial host strains, from establishing plasmid transformation and genome sequencing through to advanced genome editing technologies, is shown in Figure 2.3. Readers are also directed toward recent reviews pertaining to various aspects of clostridial genetics, metabolism, and pathogenicity discussed herein [7,8,35,82,83,84,85].

Table 2.1
Overview of the genus *Clostridium* with respect to available genome sequencing data and gene transfer procedures.

Clostridium species	Available genome sequencing data[a]	Published procedures for gene transfer[b]
Untapped clostridia		
59[c] or 184[d] species	41 species (63 strains)	2 species
Pathogenic clostridia		
C. botulinum	31 strains	+
C. difficile	218 strains	+
C. perfringens	11 strains	+
C. septicum	NA	+
C. tetani	1 strain	+
Medical clostridia		
C. acetobutylicum	3 strains	+
C. butyricum	6 strains	+
C. novyi	1 strain	–
C. sporogenes	2 strains	+
C. tetani	1 strain	+
C. tyrobutyricum	4 strains	+
Industrial clostridia		
C. acetobutylicum	3 strains	+
C. beijerinckii	2 strains	+
C. butyricum	6 strains	+
C. carboxidivorans	2 strains	–
C. cellulolyticum	1 strain	+
C. kluyveri	2 strains	–
C. ljungdahlii	1 strain	+
C. pasteurianum	3 strains	+
C. phytofermentans	1 strain	+
C. saccharoperbutylacetonicum	1 strain	+
C. thermocellum	7 strains	+
C. tyrobutyricum	4 strains	+

[a] Includes finished genomes and draft assemblies. NA: not available.
[b] Includes both transformation and conjugation methods. +: procedures for gene transfer have been developed. –: procedures for gene transfer have not yet been reported.
c Based on the core *Clostridium sensu stricto* genus (73 species).
d Based on the original *Clostridium* genus (approximately 203 species).

2.4.1. Genome sequencing

The first fully sequenced and annotated clostridial genome appeared in 2001 and was completed prior to the arrival of next generation sequencing (NGS) technologies, thus relying on classical Sanger chain-termination sequencing [16]. Since this time, the genomes of many species of *Clostridium* and hundreds of distinct strains have been sequenced and analyzed using NGS platforms. Currently, the genome sequences from at least 22 *Clostridium* species have been fully completed, whereas draft assemblies or raw sequencing data exists for at least 37 additional species, bringing the total to 59 clostridial species, encompassing roughly 374 strains (Table 2.2). Within some species, only a single strain has been sequenced, whereas up to 218 strains of *C. difficile* have been partially or fully sequenced. A similar situation exists for *C. botulinum*, which includes 31 whole genome sequencing projects, and embodies the intense medical interest in these notable human pathogens. Clostridial genomes vary drastically with respect to size (2.5 to 6.7 Mbp) and GC content (approximately 26 to 56%), illustrating the tremendous heterogeneity of the genus. Within a species, genome size and GC content can vary by greater than 1 Mbp [*C. acetobutylicum* [86], *C. botulinum* and *C. saccharolyticum*] and 5% (*C. saccharolyticum*), respectively. Several clostridia have been found to harbor extrachromosomal elements, as many as five in *C. botulinum* group II isolates [87], including small, cryptic plasmids and megaplasmids up to several hundred kb in size [88].

Although genome sequencing is often carried out in conjunction with establishing a means of gene transfer for new or undeveloped host strains (Figure 2.3), a fortunate and coincidental secondary outcome of SMRT sequencing can afford a substantial incentive to performing genome sequencing prior to carrying out transformation studies. As a result of inherent DNA polymerase kinetics, fluorescence measurements performed during SMRT sequencing differ reliably between unmodified and modified nucleotides, wherein different types of methylation can be distinguished with single base-pair resolution [89]. In a prokaryotic context, the relevant modified bases that can be detected and differentiated are N6-methyladenine (m6A) and N4- and N5-methylcytosine (m4C and m5C, respectively). Accordingly, the ability to probe an organism's unique methylation signature, or methylome, has

Table 2.2

Summary of completed and initiated whole genome sequencing projects. Data is adapted from the National Center for Biotechnology Information (NCBI) and Restriction Endonuclease database (REBASE).

Organism	Strains completed or initiated	Size of completed genomes (Mbp)	GC%	Extrachromosomal elements	M genes[a]
C. acetobutylicum	3 (completed)	4.13-4.15	30.9	1, 2	6
C. acidurici	1 (completed)	3.11	29.9	1	4
C. aldenense	1 (SRA or traces)	-	-	0	-
C. alkalicellulosi	1 (SRA or traces)	-	-	0	-
C. arbusti	1 (scaffolds or contigs)	-	29.8	0	4
C. asparagiforme	1 (scaffolds or contigs)	-	55.6	0	0
C. autoethanogenum	1 (scaffolds or contigs) 1 (SRA or traces)	-	31.0	0	4
C. bartletti	1 (scaffolds or contigs)	-	28.8	0	0
C. beijerinckii	1 (completed) 1 (scaffolds or contigs)	6.00	29.7-29.9	0	2, 4
C. bifermentans	2 (scaffolds or contigs) 1 (SRA or traces)	-	28.1-28.4	0	-
C. bolteae	6 (scaffolds or contigs) 2 (SRA or traces)	6.38	48.6-49.6	0	0
C. botulinum	13 (completed) 16 (scaffolds or contigs) 2 (SRA or traces)	3.21-4.26	27.0-29.0	0, 1, 2, 5	0-7, 9
C. butyricum	6 (scaffolds or contigs)	-	28.5-28.8	0	3
C. cadaveris	1 (scaffolds or contigs) 4 (SRA or traces)	-	31.1	0	-
C. carboxidivorans	2 (scaffolds or contigs)	-	29.7	0	7, 8
C. celatum	1 (scaffolds or contigs)	-	27.7	0	-
C. cellulolyticum	1 (completed)	4.07	37.4	0	10
C. cellulovorans	1 (completed) 1 (scaffolds or contigs)	5.26	31.1-31.2	0	13, 14
C. citroniae	1 (scaffolds or contigs) 1 (SRA or traces)	-	48.9	0	1
C. clariflavum	1 (completed)	4.90	35.7	0	15
C. clostridioforme	10 (scaffolds or contigs) 1 (SRA or traces)	-	48.7-49.2	0	-
C. colicanis	1 (scaffolds or contigs)	-	26.1	0	-
C. difficile	9 (completed) 209 (scaffolds or contigs)	4.05-4.46	28.1-52.8	0, 1	0-5
C. diolis	1 (scaffolds or contigs)	-	29.7	0	2
C. hathewayi	3 (scaffolds or contigs)	-	48.1-50.0	0	1
C. hiranonis	1 (scaffolds or contigs)	-	31.0	0	1
C. hylemonae	1 (scaffolds or contigs) 1 (SRA or traces)	-	48.9	0	-
C. intestinale	1 (scaffolds or contigs)	-	30.2	0	-
C. kluyveri	2 (completed)	3.96-4.02	32.0	1	15

C. lentocellum	1 (completed) 1 (SRA or traces)	4.71	34.3	0	4
C. leptum	1 (scaffolds or contigs) 1 (SRA or traces)	-	50.2	0	5
C. ljungdahlii	1 (completed)	4.63	31.1	0	7
C. methoxybenzovorans	1 (scaffolds or contigs)	-	44.5	0	-
C. methylpentosum	1 (scaffolds or contigs)	-	51.8	0	-
C. nexile	1 (scaffolds or contigs)	-	40.1	0	6
C. novyi	1 (completed)	2.55	28.9	0	3
C. papyrosolvens	2 (scaffolds or contigs)	-	37.0-37.1	0	3, 6
C. paraputrificum	1 (scaffolds or contigs) 1 (SRA or traces)	-	29.6	0	-
C. pasteurianum	1 (completed) 1 (scaffolds or contigs) 1 (SRA or traces)	5.04	29.8-30.6	1	5, 8
C. perfringens	3 (completed) 8 (scaffolds or contigs)	2.96-3.26	28.1-28.6	0, 1, 2, 4	1, 2, 7, 8
C. phytofermentans	1 (completed)	4.85	35.3	0	2
C. ramosum	1 (scaffolds or contigs)	-	31.4	0	0
C. saccharobutylicum	1 (completed)	5.11	28.7	0	-
C. saccharolyticum	2 (completed)	3.77-4.66	45.0-50.2	0	11, 16
C. saccharoperbutylacetonicum	1 (completed)	6.67	29.5	1	3
C. sartagoforms	1 (scaffolds or contigs)	3.98	27.9	0	-
C. scindens	1 (scaffolds or contigs)	-	46.4	0	-
C. spiroforme	1 (scaffolds or contigs)	-	28.6	0	0
C. sporogenes	2 (scaffolds or contigs)	-	27.8-28.0	0	1, 6
C. sporosphaeroides	1 (scaffolds or contigs)	-	53.5	0	-
C. stercorarium subsp. stercorarium	1 (completed)	2.97	42.2-42.3	0	6
C. sticklandii	1 (completed)	2.72	33.3	0	10
C. symbiosum	3 (scaffolds or contigs)	-	47.7-48.2	0	0, 3
C. termitidis	1 (scaffolds or contigs)	-	41.2	0	2
C. tetani	1 (completed)	2.87	28.6	1	3
C. thermocellum	2 (completed) 5 (scaffolds or contigs)	3.56-3.84	39.0-39.1	0	5-7, 12
C. tunisiense	1 (scaffolds or contigs)	-	31.2	0	4
C. tyrobutyricum	4 (scaffolds or contigs)	-	30.8	0	8, 9
C. ultunense	1 (scaffolds or contigs)	-	40.9	0	9
SUMMARY	**59 species** **374 strains**	**2.55-6.67**	**26.1-55.6**	**0-5**	**0-16**

SRA: Sequence Read Archive

[a] Genes annotated as methyltransferases.

valuable implications for strain manipulation. Often the chief barrier to genetic manipulation lies in the restriction-modification systems expressed by the host organism, as transfer of foreign DNA into many species is potently hindered by restriction (see section 2.4.2 below). Attaining a glimpse into an organism's methylome can allow identification of putative restriction-modification systems and help deduce DNA sequences recognized by such enzymes. Currently, methylome analysis is in its infancy and has only been performed on roughly 45 bacterial species, which includes *C. difficile*, *C. perfringens*, and *C. thermocellum* [90]. The utility of this strategy is highly dependent on the accurate detection of both m5C and m6A, respectively considered the fifth and sixth nucleotide bases of DNA. Within current detection limits, m6A detection requires only 25× coverage, which is attainable from most SMRT sequencing projects. Unfortunately, m5C demands at least 250× coverage for reliable determinations. Due to rapid advancement in NGS technology, these process shortcomings are expected to be resolved in the near future, leading to highly accurate *de novo* profiling of nucleotide modifications.

2.4.2. Establishing transformation

For genetic manipulation of bacterial strains, often the most imminent and difficult step involves the establishment of an efficient means of introducing foreign DNA into host cells, which paves the way to an array of valuable genetic applications (Figure 2.3). Thus, establishing a methodology for efficient transfer of plasmid DNA is paramount for the development of superior *Clostridium* strains. Two means of DNA transfer dominate within *Clostridium*: conjugation and transformation. Bacterial conjugation involves direct cell-to-cell transfer of plasmid DNA from one donor species to the target, or recipient, species, while transformation involves the uptake of DNA by competent cells. Transformation is more commonly employed in the clostridia due to its technical simplicity, greater reliability, independence from donor species, and high efficiency of plasmid transfer [91]. For these reasons, transformation is typically attempted first and, if unsuccessful, conjugative plasmid transfer is then explored. Many methods of transformation have been utilized for plasmid transfer to bacteria. Within the clostridia, however, transformation via electroporation, or electrotransformation, is the chief method that allows high-level transformation. Electrotransformation was developed in the 1980s and was first used in *Clostridium* for the transfer of plasmid DNA to *C. perfringens* [92] and *C. beijerinckii* [93]. This paved the way for numerous other clostridial electrotransformations, leading to the

genetic manipulation of approximately fifteen species of *Clostridium*, in addition to at least five other species that rely on conjugative plasmid transfer (*C. butyricum*, *C. phytofermentans*, *C. septicum*, *C. sordellii*, and *C. sporogenes*).

Gram-positive bacteria, especially members of *Clostridium*, are among the most difficult bacteria to transform. Rarely can commonalities in protocols be applied between different species or strains without rigorous optimization (Table 2.3). Further, some strains appear to be recalcitrant to electrotransformation. For example, the type species of *Clostridium*, *C. butyricum*, has proven refractory to transformation and has only been manipulated using interspecies conjugation with *Escherichia coli* [94]. Electrotransformation resistance in *Clostridium* is frequently the result of plasmid restriction by the host cell or from an inability of DNA to penetrate the thick peptidoglycan cell wall of Gram-positive bacteria [91]. Restriction can be overcome by appropriate methylation for protection of transforming plasmid DNA, whereas the Gram-positive cell wall can be weakened either through altering the physiological state of recipient cells or by modulating the electrical parameters of the electroporation pulse [95]. Strategies to overcome clostridial electrotransformation resistance are discussed in greater detail below. Further, a summary flowchart to achieve high-level electrotransformation when working with a new or untransformable strain of *Clostridium* is shown in Figure 2.4. Following optimization of electotransformation, efficiencies on the order of 10^3 to 10^6 transformants μg^{-1} should ideally be obtained. High-level efficiencies are required for most genetic engineering applications, such as gene knockout and chromosomal gene integration. If poor electrotransformation efficiency results following extensive electrotransformation troubleshooting, conjugative plasmid transfer should be employed. Conjugative plasmid transfer has been shown to succeed in some instances where electrotransformation failed, presumably due to the ability of plasmid DNA to evade host restriction [96]. Discussion of the experimental factors involved in interspecies conjugation is beyond the scope of this thesis, as protocols for conjugative plasmid transfer are often more standardized and broadly applicable across different species. Readers are directed toward published protocols for conjugative plasmid transfer to several species of *Clostridium* [96,97,98,99,100].

Table 2.3
Summary of established clostridial electrotransformation procedures.

Organism	Growth phase (OD_{600})	EPB[a]	kV	kV cm^{-1}	µF	Ω	Transformants μg^{-1}	Reference
C. acetobutylicum ATCC 824	Early stat. (1.2)	SP (7.4)	2.0	5.0	25	∞	5.0×10^4	[10,101]
C. acetobutylicum DSM 792	Mid log. (0.7)	SP (6.0)	1.8	4.5	50	600	6.0×10^2	[102]
C. beijerinckii NCIMB 8052	Mid log. (0.6)	SMP (7.4)	2.0	5.0	25	NR	2.9×10^3	[93]
C. beijerinckii NRRL B-592	Mid log.	SMH (7.4)	2.5	6.25	25	∞	NR	[103]
C. botulinum Hall A	Mid log. (0.8)	PEG	2.5	6.25	25	∞	3.4×10^3	[104]
C. botulinum ATCC 25765	(0.8)	SMP	2.0	10	25	400	0.8×10^4	[105]
C. cellulolyticum ATCC 35319	Late log. (0.5-1.0)	SMP (7.4)	1.5	7.5	25	100	2.0×10^2	[96]
C. difficile P-881	NR	SMP (7.4)	2.5	6.25	25	200	Integration[b]	[106]
C. ljungdahlii DSM 13528 (ATCC 55383)	Early log. (0.2-0.3)	SMP (6.0)	0.625	6.25	25	600	1.7×10^4	[107]
C. paraputrificum M-21	(0.4-0.6)	S	0.5	5.0	25	400	5.6×10^3	[108]
C. pasteurianum ATCC 6013	Mid log. (0.6-0.8)	SMP (6.0)	1.8	4.5	25	∞	7.5×10^4	[91]
C. perfringens 3624A	Late log./stat. (0.85-1.3)	SMP (7.4)	2.5	6.25	25	NR	9.2×10^4	[92,109,110]
C. saccharoperbutylacetonicum NI-4	(0.6)	MOPS (6.5)	NR	2.5	25	350	1.2×10^5	[111]
C. sporogenes	NR	PEG	1.25	6.25	25	100	$1.0\text{-}2.0 \times 10^2$	[112]
C. tetani CN655	Mid log.	SMP (7.4)	2.5	NR	25	200	NR	[113]
C. thermocellum DSM 1313	NR	CB	5.0[c]	25	NR	NR	2.2×10^5[d]	[114]
C. thermocellum ATCC 27405							5.0×10^4[d]	
C. thermocellum DSM 4150							1.0×10^3[d]	
C. thermocellum DSM 7072							1.0×10^3[d]	
C. thermosaccharolyticum ATCC 31960	(1.05)	G	2.0	10	25	800	5.2×10^1	[115]
C. tyrobutyricum ATCC 25755	(0.8)	SMP (7.4)	2.5	6.25	25	600	Integration[b]	[116]
C. tyrobutyricum JM1	Late log.	SMP (7.4)	2.5	6.25	NR	NR	$3.0\text{-}4.0 \times 10^0$	[117]
CONSENSUS	**Mid log. (0.5-0.8)**	**SMP (7.4)**	**2.0-2.5**	**5.0-6.25**	**25**	**200-600**	**10^3-10^4**	**[91]**

NR: not reported.

[a] S: 270-272 mM sucrose, M: 1 mM $MgCl_2$, P: 5-7 mM sodium phosphate, H: 7 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*'-2-ethanesulfonic acid), PEG: 10% polyethylene glycol 8000, MOPS: 65 mM 3-(*N*-morpholino)propanesulfonic acid, CB: 200 mM cellobiose, G: 20% glycerol.

[b] Non-replicating (i.e., integrative) plasmid was used.

[c] Square-wave pulse was administered.

[d] Induction of antibiotic selection prior to plating on selective media.

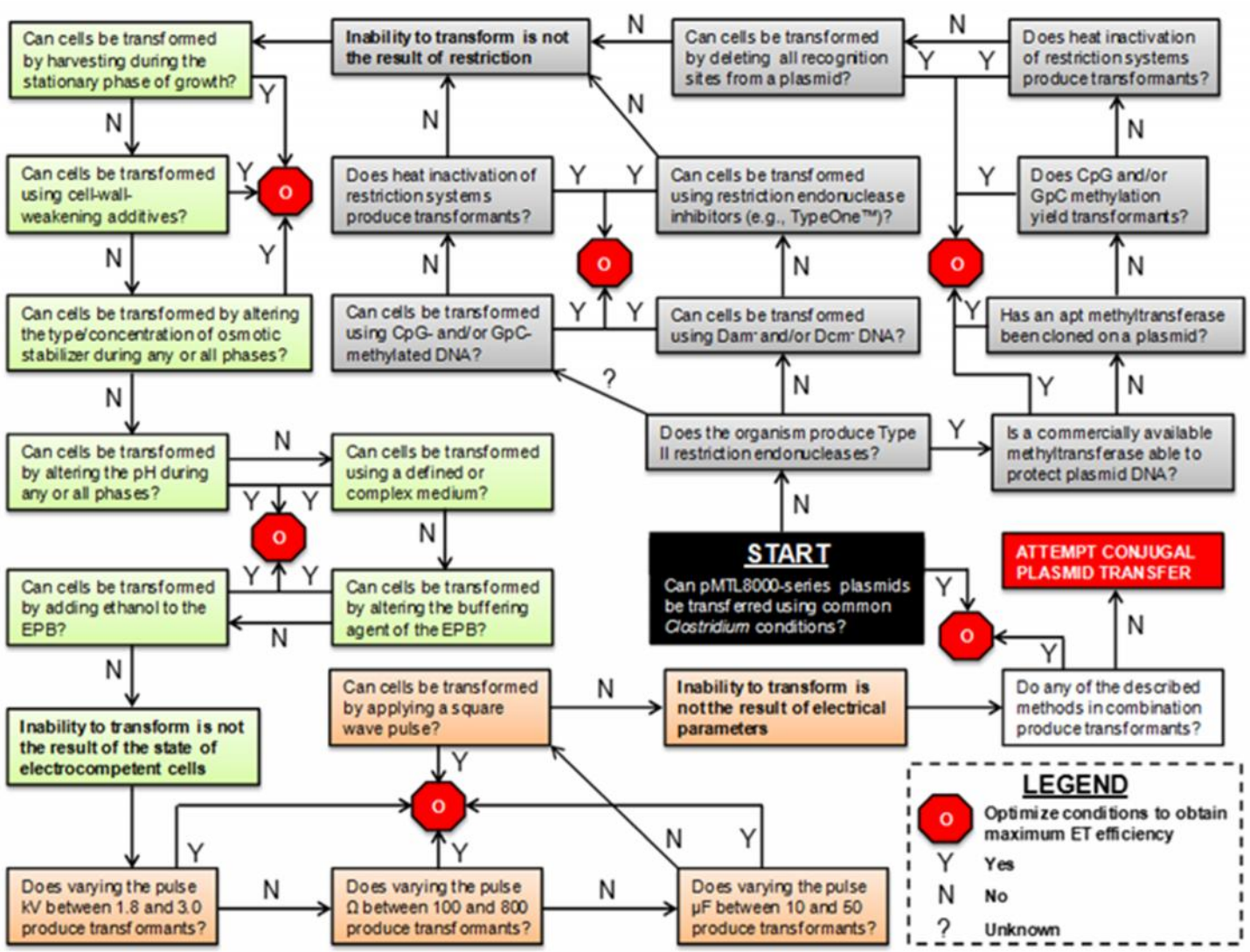


Figure 2.4. Flow chart for establishing an electrotransformation protocol in an untransformed or intractable *Clostridium* species. Factors affecting efficient transformation are grouped according to common electroporation barriers; namely, restriction (blue shading), competent cell physiology (green shading), and electrical parameters (red shading). Adapted from [118].

(i) Presence of active restriction-modification systems

The presence of highly active restriction-modification (RM) systems is the most probable factor responsible for hindering electrotransformation. If improperly methylated plasmid DNA is used to transform even highly competent cells, few or no transformants are obtained [10,91]. As such, an all-or-nothing effect on electrotransformation tends to be indicative of the presence of at least one active RM system. When attempting to transform a new or uncharacterized strain of *Clostridium*, often a key step entails assaying crude cell lysates for the presence of Type II restriction endonucleases by incubation of the active lysate in the presence of unmethylated plasmid DNA. Using this technique, Type II RM systems have been detected in *C. acetobutylicum* [10], *C. cellulolyticum* [96], *C. difficile* [98], *C. pasteurianum* [119], and *C. thermocellum* and *C. thermosaccharolyticum* [120]. Analysis of cell lysates for restriction activity can also be coupled to methylome analysis via SMRT sequencing to decipher the recognition sequence of putative restriction barriers.

Once a putative recognition sequence is identified, efforts can be made to protect, via methylation, the recognition sequences on transforming plasmids. For example, many *Clostridium* species produce isoschizomers of the *E. coli* Dam (DNA adenine methylase) RM system [90], and therefore restrict DNA prepared from Dam^- *E. coli* hosts. Most common laboratory *E. coli* strains are Dam^+ and, therefore, such barriers are easily overcome. Other identified Type II RM systems from *Clostridium* possess more unique recognition sequences and DNA protection requires methylation from a specific methyltransferase. DNA methylation can be performed in vivo by expressing the methyltransferase in *E. coli* cloning strains or, ideally, *in vitro* if the methyltransferase is available commercially. Most clostridial Type II RM systems possess recognition sequences rich in GC [e.g., 5'-GCNGC-3' for Cac824I from *C. acetobutylicum* [10] and 5'-CGCG-3' for CpaAI from *C. pasteurianum* [119]], owing to the low GC content that is a typical characteristic of the clostridia. In the case of *C. cellulolyticum*, a Type II RM system, CceI, was identified within cell lysates and found to cleave the sequence 5'-CCGG-3', which can be conveniently protected by the commercial methyltransferase, M.MspI (5'-m5CCGG-3'), but not M.HpaII (5'-Cm5CGG-3'), demonstrating the importance of methylation pattern within a given recognition sequence [96]. Unfortunately, few methyltransferases are commercially available and cloning of an appropriate methyltransferase

for *in vivo* methylation is necessary. This strategy is routinely employed for electrotransformation of *C. acetobutylicum* and *C. pasteurianum* [10,91].

In cases where neither commercial methyltransferases nor suitable methyltransferase genes are available, other methods have found utility in evading RM systems in a range of bacteria (Figure 2.4). For example, it has been shown that RM systems in some organisms can be temporarily heat-inactivated, allowing transformation to occur unimpeded by restriction. Inactivation regimens differ widely among bacteria with respect to duration of exposure, including 55 °C for 15 minutes for *C. acetobutylicum* (polyethylene-glycol-mediated protoplast transformation) [121] and 56 °C for 2 minutes for *Staphylococcus carnosus* [122]. This method has not been demonstrated in a clostridial electrotransformation context, but should prove effective in some mesophilic clostridia. An alternative method involves methylation of transforming DNA with the commercially available CpG (M.SssI) and/or GpC (M.CviPI) methyltransferases, which yield 5'-m5CG-3' and 5'-Gm5C-3' target sequences, respectively. Theoretically, methylation from one or both of these methyltransferases should confer protection against most known clostridial Type II restriction endonucleases as a result of the high occurrence of G and C bases in clostridial recognition sequences [REBASE [90]]. Many bacteria, however, including *E. coli*, potently restrict DNA possessing abundant m5C residues [123]. Consequently, extensively CpG- and GpC-methylated substrates failed to transform *C. pasteurianum*, despite adequate protection against the CpaAI (5'-CGCG-3') Type II endonuclease [91]. Although RM systems generally lead to an all-or-nothing effect on transformation, some transforming plasmids can escape restriction and become methylated by the endogenous methyltransferase, allowing propagation. Therefore, if at least one transformant can be obtained, plasmid DNA can then be isolated for retransformation into the host strain. If a Type II endonuclease is responsible for inhibiting electrotransformation, utilizing plasmid DNA isolated from the same host should transform several orders of magnitude better than *E. coli*-prepared plasmid DNA, as demonstrated in *C. septicum* [124] and *C. acetobutylicum* [10]. This strategy represents a classical approach for determining the presence of RM systems and, if practical, for overcoming the transformation barriers posed by such RM systems. Finally, all known restriction recognition sequences can be mutated from a transforming plasmid to evade host restriction. However, the recognition sites of many clostridial Type II endonucleases occur

frequently, especially within the *E. coli* regions of *E. coli-Clostridium* shuttle vectors, making this method less practical.

While most bacterial RM systems acting on transforming DNA substrates are Type II, Types I, III, and IV restriction endonucleases can also drastically reduce electrotransformation efficiencies. Type I and III restriction endonucleases are more complex than Type II systems, as they cleave non-specifically and at regions remote from their recognition sequences. Nevertheless, at least one mode of overcoming Type I restriction is commercially available. Named TypeOne™ Restriction Inhibitor (Epicentre; http://www.epibio.com/), the bacteriophage T7 Ocr protein has been shown to increase electrotransformation efficiency by several orders of magnitude in *E. coli*, *Salmonella typhimurium* , and *C. ljungdahlii* [125,126]. Alternatively, Type IV systems possess a reversed RM specificity by cleaving methylated recognition sequences, such as restriction of 5'-Sm5CNGS-3' (where S=G or C; N=A, C, G, or T) by SauUSI from *Staphylococcus aureus* subsp. *aureus* USA300 [127]. Compared to Type II enzymes, the respective mechanisms and recognition sequences of Type I, III, and IV RM systems remain largely unknown. Improvements in SMRT sequencing and subsequent methylome analysis are expected to decipher the recognition sites of many Type I, III, and IV RM systems and determine the relative contribution of such systems to various bacterial transformations.

(ii) Physiological state of electrocompetent cells

Inability to transfer plasmid DNA to a recipient cell can also be the result of the physical barrier imposed by the Gram-positive cell wall [95]. Since the cell wall is continuously remodeled throughout the course of bacterial growth, the growth phase of cells at the time of harvest is often critical. Whereas it is usually optimal to use mid-logarithmic-phase cells for electrotransformation (Table 2.3), some species of *Clostridium* undergo autolysis during stationary phase, which has been found to enhance permeabilization. Electrotransformation of *C. perfringens* is optimum using cells harvested from the late logarithmic or stationary phase, as mid-logarithmic-phase cells cannot be electrotransformed [110]. The structure and density of the cell wall can also be altered by the formulation of the growth medium and wash and electroporation buffer (EPB), including the pH and buffer type and strength, in addition to the presence of cell-wall-disrupting agents and associated osmotic stabilizers [95]. Work with *C. perfringens* has demonstrated the importance of maintaining a favorable and consistent pH

during all phases of electrotransformation, as adjusting the pH of the recovery and plating media to match that of the growth medium afforded a 2- to 6-fold enhancement in electrotransformation efficiency [109]. On the other hand, the majority of clostridia are best electrotransformed using an EPB consisting of 5-7 mM sodium phosphate buffer, pH 7.4, 1 mM magnesium chloride, and isotonic (270 mM) sucrose (Table 2.3). However, in addition to 10-20% glycerol or polyethylene glycol solutions, it is often beneficial to assess the effect of alternative EPBs, such as HEPES and MOPS buffering systems, which have been shown to affect permeabilization of the clostridia [103,104,111,115].

The most robust method for enhancing electrotransformation of *Clostridium* and other Gram-positive organisms entails the application of cell-wall-weakening additives, muralytic enzymes, and membrane-solubilizing agents. Various cell-wall-weakening compounds, including glycine, isonicotinic acid hydrazide (isoniacin), penicillin G and ampicillin, and DL-threonine, have all been investigated for stimulating clostridial electrotransformation [91,114,116,128]. Glycine [129] and DL-threonine [130] weaken the cell wall through incorporation and subsequent disruption of cross-linking, while isoniacin [131] and penicillin G and ampicillin [132] inhibit key enzymes involved in cell wall synthesis. Glycine was found to significantly stimulate electrotransformation of *C. pasteurianum* [91] and *C. cellulolyticum* [128], and DL-threonine is indispensable for plasmid DNA transfer to *C. tyrobutyricum* [116] and, to a lesser extent, *C. ljungdahlii* [125]. Further, muralytic enzymes, such as lysostaphin and lysozyme, have proven effective for the partial removal of cell walls prior to transformation in *C. perfringens* [133,134]. Still, many species of *Clostridium* do not require any special compound or enzyme treatments to achieve high-level electrotransformation and the effect of such additives on electrotransformation cannot be foreseen without experimentation. Since these additives are strongly detrimental to growth, effective cell wall disruption must be carefully balanced with cell viability [95]. Thus, cell-wall-weakening compounds and muralytic enzymes are added during cell growth, typically during logarithmic phase, or for a minimal span of time immediately prior to electrotransformation. Regardless of timing, exposure to cell-wall-weakening additives creates an increased requirement for osmotic stabilization, as cells with disrupted cell walls are highly susceptible to growth inhibition, cytoplasmic leakage, and death. In fact, the concentration and duration of exposure of such additives is dependent on the amount of osmotic stabilizer present in the growth medium, as typically a greater concentration of osmoprotectant permits more

extensive cell wall disruption [95]. Sucrose, ranging from 0.25 to 0.5 M, is the most common osmotic stabilizer and is supplied to yield an iso- or hypertonic environment. Osmotic stabilization has proven vital not just during exposure to cell-wall-weakening additives, but also during all stages of electrotransformation, including washing, pulse delivery, and recovery [91]. Finally, in an analogous manner to cell wall disruption, the cellular membrane of bacteria can be disturbed via partial solubilization in the presence of ethanol for *E. coli* [135], *Oenococcus oeni* [136], and *C. pasteurianum* [91], and Tween 80 for *Bacillus amyloliquefaciens* [137].

(iii) Electrical parameters

Since electrotransformation depends on the development of transient pores in cellular membranes through the delivery of an electric pulse, the electrical parameters of the administered pulse are critical for effective permeabilization and plasmid transfer. Electroporation pulses are typically described by the voltage (kV) or field strength (kV cm^{-1}), capacitance (µF), and resistance (Ω). In addition, the duration of the pulse is described using the time constant, which is the product of the total resistance and capacitance of the circuit, or the time it takes for a pulse to exponentially decrease to 37% of its initial peak voltage (Bio-Rad; http://www.bio-rad.com/). While eukaryotic cells require moderate field strengths for efficient electrotransformation, smaller bacterial cells necessitate high field strengths up to 18 kV cm^{-1}. The majority of the clostridia, however, are transformed optimally with field strengths between 5.0 and 6.25 kV cm^{-1}, generally corresponding to voltages of 2.0-2.5 kV (Table 2.3). For high-level electrotransformation to occur, it is often assumed that a significant proportion of the cell population might not survive pulse delivery. Therefore, the strength of the electric pulse must be balanced with cell viability. The strength of the electric pulse is also dependent on the physiological state of the recipient cells at the time of electroporation, as cells possessing weakened cell walls require lower voltages to achieve adequate electrotransformation compared to untreated cells [95]. Glycine-treated *C. pasteurianum* [91] and DL-threonine-treated *C. ljungdahlii* [125], therefore, utilize low voltages of 1.8 and 0.625 kV, respectively. In terms of capacitance, all clostridial electrotransformation protocols described to date, except one developed for *C. acetobutylicum* [102], utilize capacitors of 25 µF, although this parameter can be modulated to optimize the pulse time constant. In a similar manner, the pulse resistance can be varied to alter the duration of the electric pulse. Parallel resistors between 100 to 800 Ω, or ∞

Ω (absence of parallel resistors), are effective for transforming *Clostridium* species (Table 2.3), and should be investigated when developing an electrotransformation methodology for a new or untransformed strain. Finally, rather than traditional exponential-decay pulses, which are the most commonly employed electroporation pulses, square wave pulses have found utility within the *Clostridium*. Square wave pulses maintain a consistent voltage throughout the duration of the pulse and have led to increased electrotransformation efficiencies in *C. acetobutylicum* [138] and *C. thermocellum* [114].

2.4.3. Development of host/vector system and selection of promoters

With an efficient method of DNA transfer in place, development of a host-vector system, encompassing a broad selection of antibiotic-resistance determinants and functional replication origins, is paramount for effective strain manipulation in the clostridia. Recombinant strains of most species are selected using either erythromycin/clarithromycin (from *ermB*) or thiamphenicol/chloramphenicol (from *catP*), although determinants for spectinomycin- and tetracycline-resistance have been effective for some species [14]. Recent work has also led to the construction of a series of vectors containing four functional and distinct origins of replication and four common clostridial selectable markers [139]. Different replication origins can then be easily assessed, as the superior origin has been shown to differ between species. Segregational stabilities of the four modular clostridial origins were found to vary between 69.0 and 99.9% per generation for *C. acetobutylicum*, *C. botulinum*, *C. difficile*, and *C. tyrobutyricum* hosts [139,140]. Yu et al. [140] demonstrated the importance of proper replicon selection by heterologously expressing the *C. acetobutylicum adhE2* gene in *C. tyrobutyricum* from four shuttle vectors, each harboring a different modular clostridial origin. While segregational stability was greater than 95% for three of the four vectors, conjugative plasmid transfer efficiency varied by an order of magnitude between the vectors. More importantly, butanol titers differed by more than 7-fold (0.90 to 6.87 g L^{-1}), suggesting profound implications for plasmid-based metabolic engineering strategies. Since almost all genetic applications require selection and stable replication of clostridial shuttle plasmids (Figure 2.3), optimal design of host/vector systems is critical.

In addition to selectable markers and replication origins, evaluation and selection of strong gene promoters is needed. The genetic techniques depicted in Figure 2.3, specifically

plasmid-based gene overexpression, antisense-RNA-mediated gene knockdown, chromosomal gene knockout, and advanced genome editing technologies, demand access to a repertoire of both constitutive and inducible promoter systems. In order to assess promoter function and strength, numerous gene reporter systems have been adapted or developed for use in *Clostridium*, including chloramphenicol acetyltransferase (*catP*) from *C. perfringens* [141], β-galactosidase (*lacZ*) from *Thermoanaerobacterium thermosulfurogenes* [12], β-1,4-endoglucanase (*eglA*) from *C. acetobutylicum* [142], luciferase (*luxAB*) from *Vibrio fischeri* [143], β-glucuronidase (*gusA*) from *E. coli* [11], and an oxygen-independent green fluorescent protein from *Pseudomonas putida* [128]. Among these, *lacZ* and *gusA* have been most widely used for genetic studies. In each case, the selected promoterless reporter gene is cloned into a shuttle vector for transcriptional fusion with various promoters in order to evaluate relative promoter strength. Promoters from genes involved in the central carbon metabolism of the clostridia are presumed to be strong and are commonly chosen for investigation. Acetoacetate decarboxylase (*adc*), phosphotransbutyrylase (*ptb*), and thiolase (*thl*) gene promoters [12,128,144] have been extensively applied for gene expression in *C. acetobutylicum* and related clostridia. Promoters from other central fermentative genes leading to high-level gene expression include ones from ferredoxin [*fdx*; *C. pasteurianum* and *C. sporogenes* [14,145]], butanol dehydrogenase [*bdh*; *C. saccharoperbutylacetonicum* [146]] and pyruvate:ferredoxin oxidoreductase [*pfo*; *C. phytofermentans* [99]].

Certain genetic applications require the use of inducible promoter systems, resulting from either the toxicity or the desire for transient expression of the encoded protein. As a result, inducible gene expression systems have recently been adapted for use in the clostridia, with induction arising from either biotic or abiotic inducers (Table 2.4). These systems include a xylose-inducible promoter [11], a *lac* operator-containing ferredoxin promoter (*fac* promoter) induced by isopropyl β-D-1-thiogalactopyranoside [14], a UV-inducible *recA* promoter [147,148], a *tetO1* operator-containing promoter induced by anhydrotetracycline [149], and, most recently, a lactose-inducible promoter [78,150]. Clostridial promoter-repressor systems developed to date vary drastically with respect to stringency, as expression in the presence of inducer has been found to increase from 1.4- up to approximately 120-fold. For an optimum inducible expression system, promoter activity should be efficiently repressed in the absence of inducer and activated to a high level upon induction. Of all clostridial inducible promoters, it

Table 2.4
Summary of clostridial inducible promoter-repressor systems developed to date.

Organism	Basal promoter	Repressor	Operator	Inducer	Maximum inducibility[a]	Reference
C. acetobutylicum ATCC 824	*recA* (*C. acetobutylicum*)	*dinR* (endogenous)	Cheo box	2 Gy (radiation)	1.4- to 4.1-fold	[147,148]
C. acetobutylicum ATCC 824	*xylA* (*Staphylococcus xylosus*)	*xylR* (*S. xylosus*)	xylO	10 g/l xylose (in absence of glucose)	18-fold	[11]
C. botulinum ATCC 3502 *C. difficile* 630 *C. sporogenes* NCIMB 10696	*fdx* (*C. pasteurianum*)	*lacI* (*E. coli*)	lacO	1 mM IPTG	NR	[14]
C. acetobutylicum ATCC 824	*catP* (*C. perfringens*)	*tetR* (*E. coli*)	2× tetO1	100 ng/ml anhydrotetracycline	41- to 119-fold	[149]
C. acetobutylicum ATCC 824 *C. perfringens* SM101 *C. perfringens* JGS4143 *C. perfringens* 13	*bgaL* (*C. perfringens*)	*bgaR* (*C. perfringens*)	NR	10 mM lactose	10- to 15-fold (*C. acetobutylicum*) 80-fold (*C. perfringens*)	[78] [150]

[a] Calculated by dividing reporter enzyme activity under induction by the corresponding activity under uninduced conditions.

appears the anhydrotetracycline-inducible Pcm-2tetO1 promoter [149] is the most stringent system currently available (Table 2.4). Also, another recent inducible gene expression system, based on lactose induction in both *C. acetobutylicum* [78] and *C. perfringens* [150], also exhibits marked repression in the absence of inducer and has been utilized successfully for the propagation of a plasmid harboring a functional *mazF* gene, the expression of which is lethal. The clostridial inducible promoters should prove indispensable for future genetic applications requiring transient gene expression or strict control of intracellular protein levels, particularly for processes involving toxic or lethal gene products.

2.4.4. Plasmid-based gene overexpression

Gene overexpression applications encompass both homologous and heterologous expression strategies. While homologous overexpression aims to amplify the activity of a resident chromosomal gene through increasing gene dosage, heterologous strategies are commonly utilized to impart a foreign or non-native activity to the host strain. Many clostridial plasmid-based gene overexpression strategies have been reported in the past twenty years, most of which involve *C. acetobutylicum* (Table 2.5). Several genes involved in the central fermentative metabolism of *C. acetobutylicum* have been overexpressed individually and in combination, including the aldehyde/alcohol dehydrogenases, *aad*/*adhE1* [57,151] and *adhE2* [152], thiolase, *thl* [151], and acetone/solvent formation operons, *adc-ctfAB* [153] and *aad-ctfAB* [154]. These strategies have led to both positive and negative outcomes in terms of butanol and total solvent titers, demonstrating the limitations of rational metabolic engineering approaches in clostridia with highly complex and branched fermentations. Non-solventogenic pathway genes that have been overexpressed in *C. acetobutylicum* include the master sporulation transcriptional regulator, *spo0A* [101,155], and the class I heat shock response operon, *groESL* [156], yielding important mutant strains with enhanced solvent production and tolerance. Homologous gene expression has also been performed in related clostridia. In analogous studies, regulatory proteins controlling botulism [*botR*/*A*; [157]] and tetanus [*tetR*; [113]] toxin production were overexpressed in their respective hosts, *C. botulinum* and *C. tetani*, allowing characterization of their previously unknown roles in toxin production. Clostridial hydrogenases, encoded by *hydA*, have also been common targets for homologous gene overexpression due to their central roles in redox balance during fermentative growth. Hydrogenases from *C. acetobutylicum* [158,159], *C.*

Table 2.5
Summary of notable clostridial homologous and heterologous gene overexpression strategies.

Host organism	Gene(s)/operon	Phenotype relative to control strain(s)	Reference
Homologous overexpression			
Clostridium acetobutylicum ATCC 824/DSM 792	Acetone operon; *adc-ctfAB*	Higher yield of solvents	[153]
	Sporulation transcriptional regulator; *spo0A*	Enhanced solvent formation and tolerance	[101,155]
	Class I heat shock genes; *groESL*	Increased solvent titers and tolerance	[156]
	Hydrogenase; *hydA*	NR	[158]
		Unaltered hydrogen yield and productivity	[159]
	Aldehyde/alcohol dehydrogenase; *aad*	Increased ethanol and acetate production	[151,160]
	Thiolase; *thl*	Increased acid and decreased solvent production	[151]
	Aldehyde/alcohol dehydrogenase; *adhE2*	Partially rescued butanol production[a]	[152]
	Solvent operon; *aad-ctfAB*	Fully rescued butanol production with acetate accumulation[a]	[154]
C. botulinum A strain 62	Botulism neurotoxin regulatory protein; *botR/A*	Increase in botulism neurotoxin production	[157]
C. paraputrificum M-21	Hydrogenase; *hydA*	Increased hydrogen gas productivity and acetic acid production	[161]
C. tetani CN655	Tetanus toxin regulatory protein; *tetR*	Increase in tetanus toxin production	[113]
C. tyrobutyricum JM1	Hydrogenase; *hydA*	Increased hydrogen yield	[117]
Heterologous overexpression			
Clostridium acetobutylicum ATCC 824/DSM 792	Minicellulosome components (*C. cellulolyticum* + *C. thermocellum*)	Secretion of active, heterologous and chimeric cellulases and scaffoldin proteins	[162,163,164,165,166]
	1,3-Propanediol operon; *dhaB1-dhaB2-dhaT* (*C. butyricum*)	1,3-propanediol as major fermentation product[a]	[167]
	Hydrogenase; *hydA* (algal) (*C. butyricum*)	NR	[158]
		Unaltered hydrogen yield and productivity	[159]

	Transaldolase; *talA* (*E. coli*)	Enhanced xylose utilization and solvent production	[168]
	Acetoin reductase; *acr* (*C. beijerinckii*)	Conversion of acetone to 2,3-butanediol	[169]
	Primary/secondary alcohol dehydrogenase; adh_{B-593} (*C. beijerinckii*)	Conversion of acetone to isopropanol	[170]
C. cellulolyticum H10	Synthetic ethanol formation operon; *pdc-adhII* (*Zymomonas mobilis*)	Improved growth and ethanol production	[171]
	Chimeric isobutanol operon; *kivD-yqhD-alsS-ilvC-ilvD* (*E. coli*, *B. subtilis*, and *L. lactis*)	Small amounts of isobutanol produced	[172]
C. ljungdahlii DSM 13528	Synthetic butanol formation operons; *bcd-hbd-crt* and *thlA-bdhA-adhE* (*C. acetobutylicum*)	Small amounts of butanol produced	[125]
C. perfringens SM101	Alternative RNA polymerase sigma factor; *txeR* (*C. difficile*)	Activation of *toxA* and *toxB* promoters	[173,174]
	Negative regulator of toxin synthesis; *tcdC* (*C. difficile*)	Decreased toxin gene expression	[175]
C. tyrobutyricum ATCC 25755	Aldehyde/alcohol dehydrogenase; *aad* (*C. acetobutylicum*)	Significant production of butanol and ethanol	[144]
	Aldehyde/alcohol dehydrogenase; *adhE2* (*C. acetobutylicum*)		[140]

NR: not reported.

[a] Host strain was a solvent-defective, degenerate mutant of *C. acetobutylicum*, either strain M5 or DG1

paraputrificum [161], and *C. tyrobutyricum* [117] have been homologously overexpressed in their respective hosts. As shown with other central fermentative genes, overexpression of hydrogenases in clostridial species often yields unpredictable outcomes, ranging from drastic effects on redox balance and product distribution [117,161] to no detectable changes in growth and metabolism [159].

In contrast to homologous overexpression strategies, expression of foreign or non-native activities in *Clostridium* is a relatively new avenue of research. Nevertheless, work from the past decade has led to the engineering of *C. acetobutylicum* for the expression of a range of heterologous activities, including most notably components of a functional minicellulosome from *C. cellulolyticum* and *C. thermocellum* [162,163,164,165,166], a 1,3-propanediol production pathway from *C. butyricum* [94,167], a *C. beijerinckii* primary/secondary alcohol dehydrogenase [170] and acetoin reductase [169] for conversion of acetone to isopropanol or 2,3-butanediol, respectively, and transaldolase from *E. coli* [168] for improved xylose utilization. A number of prokaryotic and eukaryotic hydrogenases have also been heterologously expressed in *C. acetobutylicum*, including ones of clostridial [159] and algal [158] origin. In other instances, a genetically-tractable species is utilized to study the genetics of a related strain for which genetic tools are not available, as has been shown through the investigation of *C. difficile* pathogenicity regulatory elements in *C. perfringens* [e.g., [173,174,175]]. Conversely, solventogenic genes from *C. acetobutylicum*, the most tractable *Clostridium*, have been heterologously expressed in hopeful alternative hosts for which genetic tools have been recently developed, such as *C. ljungdahlii* [125] and *C. tyrobutyricum* [140,144]. While efforts with *C. ljungdahlii* have led to only modest titers of butanol (approximately 0.15 g L^{-1}) and demand more extensive metabolic engineering strategies, mutants of *C. tyrobutyricum* expressing a heterologous *adhE2* gene from *C. acetobutylicum* have been reported to produce butanol with titers of 10-20 g L^{-1}. Finally, *C. cellulolyticum* has shown promise as a host for the production of ethanol and isobutanol from cellulose through construction and expression of synthetic, chimeric operons harboring heterologous genes from *E. coli*, *B. subtilis*, *L. lactis*, and *Zymomonas mobilis* [171,172].

Owing in part to the low GC content of most industrially- and medically-significant clostridia (roughly 27-32%; Table 2.2), codon bias is an important consideration for heterologous gene expression. *C. perfringens* exhibits one of the strongest codon biases of all bacteria

analyzed to date [176,177]. Of the six codons encoding arginine, *C. perfringens* favors AGA by a factor of approximately 10-260× relative to four of the other degenerate codons, while the remaining CGG codon is essentially unused [176]. Genes that are highly expressed have been found to exhibit a greater degree of codon bias, whereas poorly expressed genes tend to follow a more random selection of codons [178]. Consequently, codon bias has a marked effect on heterologous gene expression approaches involving the clostridia. However, codon optimization is commonly avoided in clostridial heterologous gene expression approaches, especially those involving large operons and multi-subunit gene products. Instead, genes or operons of interest are often amplified from the native organism's DNA and cloned directly to the host strain for heterologous expression. Although this strategy is often adequate when cloning genes between related clostridia, which are expected to possess similar codon biases, poor expression can arise when expressing genes from non-clostridial origins. Therefore, codon optimization via continuously-improving commercial gene synthesis will allow the generation of superior clostridial strains through the improved expression of heterologous genes, operons, and metabolic pathways.

2.4.5. asRNA gene knockdown

In contrast to plasmid-based overexpression strategies, antisense RNA (asRNA) molecules possessing complementarity to the mRNA transcript of a target gene can be utilized to decrease, yet not entirely abolish, protein expression. Accumulating evidence from transcriptome studies suggests that asRNA transcription occurs naturally in the bacterial kingdom. Approximately 1,000 different asRNAs have been identified in *E. coli* [179] and the transcriptome of *B. subtilis* encompasses asRNAs for 18% of the 506 genes analyzed [180]. The basic premise of asRNA techniques lies in the interaction of the asRNA with its complementary mRNA target, forming duplex RNA with altered the secondary structure and potentially reduced stability and half-life. Antisense RNA, therefore, allows inhibition of gene expression by either initiating degradation of the mRNA transcript or by restricting ribosome-mediated translation [13,181,182].

Prior to the major advancement of gene knockout technology in the clostridia, asRNA gene knockdown was the most common tool for reducing expression of native chromosomal genes for metabolic engineering of *C. acetobutylicum*. Desai and Papoutsakis [13] separately

targeted asRNAs to butyrate kinase (*buk*) and phosphotransbutyrylase (*ptb*) genes, resulting in enhanced production of acetone and butanol for the *buk* knockdown mutant, but deteriorated solvent production for the *ptb* downregulated strain. Similarly, Tummala et al. [183] demonstrated effective knockdown of coenzyme A-transferase (*ctfAB*) involved in acetone production, resulting in significant reduction in acetone titer and mildly lower levels of butanol. Downregulation of *ctfB* has also been combined with overexpression of an alcohol-aldehyde dehydrogenase (*aad*) gene to return butanol yield to control levels with a 23-fold increase in ethanol titer, while maintaining low acetone selectivity [183]. In addition, Sillers et al. [151] overexpressed the thiolase gene of *C. acetobutylicum* along with *ctfAB* knockdown and *aad* overexpression to further increase butanol production. Thus, asRNA approaches can work effectively with overexpression approaches for optimal metabolic engineering of *Clostridium*.

Antisense-RNA-mediated gene knockdown has also been employed in *Clostridium* beyond *C. acetobutylicum*. In *C. saccharoperbutylacetonicum*, asRNA was applied effectively to downregulate expression of the hydrogenase gene cluster, *hupCBA*, and alter solvent selectivity, leading to a butanol/acetone ratio of 1.3, compared to 2.9 for the control strain [146]. Antisense RNA was also employed successfully in *C. beijerinckii* to reduce the expression of a glycerol dehydrogenase gene, *gldA*, and increase butanol tolerance [184], and in *C. cellulolyticum* to downregulate *cel48F*, a major component of the organism's cellulosome [185]. Finally, within the medical clostridia, asRNA-mediated gene knockdown has been utilized for downregulation of alpha- and beta-type small acid-soluble spore proteins to investigate the relationship between these proteins and their roles in resistance of *C. perfringens* spores to moist heat and UV radiation [186].

For clostridial species with limited availability of genetic tools, asRNA can be the simplest genetic technique to carry out. Gene knockdown requires only access to a functional and stable shuttle vector, a constitutive or inducible promoter, a means of plasmid transfer, and sequence data for the target gene to be downregulated. Being less drastic than gene knockouts, asRNA gene knockdown strategies are less likely to be lethal and often allow isolation of a viable mutant where harsher knockout methods fail. Additionally, careful promoter selection can facilitate asRNA knockdown experiments through strict control over the degree of downregulation by fine tuning of inducible promoters or selection of appropriate constitutive promoters [13]. Antisense RNA strategies may be more conducive to a high-throughput

screening approach than more complicated approaches, such as group II intron gene knockout (refer to section 2.4.6).

Effective downregulation of a target gene by asRNA can be unpredictable due to many complex factors involved in the design of the asRNA. Care should be taken in terms of size of the asRNA molecule, degree of homology to the target mRNA, region of mRNA to be targeted (e.g. coding sequence, untranslated region, ribosome binding site, etc.), and asRNA secondary structure which can be predicted using RNA structural prediction software [77]. Many successful asRNA approaches have included the ribosome binding site, in addition to a small 5' portion of the target coding sequence [77,185]. Unfortunately, few studies have systematically investigated asRNA design for the effectiveness of gene downregulation in the clostridia. Tummala et al. [77] evaluated the design of asRNA constructs in an attempt to illuminate design parameters fundamental to the success of asRNA downregulation in *C. acetobutylicum*. This study employed RNA prediction software to design several different asRNAs targeting the acetoacetate decarboxylase (*adc*) and coenzyme A-transferase (*ctfAB*) mRNAs. Antisense RNAs were structurally analyzed and the normalized metrics of these structural features were plotted against the degree of downregulation. Two structural features, referred as components (regions of intramolecular complementarity within an asRNA molecule) and free nucleotides (which do not reside within a component but are available for interaction with the target mRNA), were analyzed for correlation with *in vivo* asRNA downregulation [77,187]. Experimental data suggest that the normalized number of components correlates well with the degree of asRNA downregulation. Specifically, the ratio of the number of components to the number of total free nucleotides can be used as a predictor of asRNA effectiveness, whereby maximal downregulation is achieved using the asRNA construct having the lowest component/nucleotide ratio [77].

Finally, in terms of the broader role of non-coding RNAs as tools for modulating gene expression, RNA deep sequencing has revealed an abundance of small non-coding RNAs (sRNAs) in *C. acetobutylicum* [188], at up to 7.5% of the total sequencing reads. As the *Clostridium* sRNome continues to be deciphered, sRNAs may become an important part of the clostridial genetic engineering toolkit.

2.4.6. Group II intron gene knockout

It took approximately 15 years after the first reports of plasmid transfer for the first broadly-applicable gene knockout methodology to reach widespread use in the *Clostridium* community. The current established method of gene knockout in clostridia was derived from group II intron technology, termed TargeTron gene knockouts, originally developed and commercialized for use in *E. coli* and related bacteria (Sigma-Aldrich; http://www.sigmaaldrich.com/). The clostridial TargeTron system, or ClosTron, was developed independently by two groups in 2007 [14,15] and employs the mobile *Lactococcus lactis* Ll.ltrB group II intron and its cognate intron-encoded protein (IEP), LtrA, which is essential for intron function. According to the Ll.ltrB intron splicing mechanism, the bacterial intron can be programmed to insert into precise regions of bacterial genomes [189]. The predominant factor dictating site-specific intron insertion was found to be minor base-pairing (generally 11-16 bp for group II introns) between the intron RNA and host target DNA [190,191]. Once target site recognition has occurred, the IEP nicks one strand of the host's DNA and integrates the Ll.ltrB intron into the genome through reverse transcriptase activity [192]. The LtrA IEP thus exhibits a tripartite functionality, acting as a maturase, endonuclease, and reverse transcriptase [193]. Since the intron homing mechanism utilizes an RNA intermediate to generate a chromosomal DNA disruption via reverse transcription, the intron mechanism has been termed retrohoming. Further, by identifying the regions of the intron RNA involved in base-pairing and generating a pool of introns randomly mutated in such regions, the rules governing intron insertion were systematically elucidated in *E. coli* [194]. From these efforts, a computer algorithm was then designed to utilize the rules for Ll.ltrB insertion site recognition and splicing to predict putative insertion sites within a target gene query [195]. PCR primers are then generated to allow mutation of the plasmid-encoded Ll.ltrB intron, via splicing by overlap extension PCR (SOE PCR), for retargeting the intron to the selected chromosomal gene. Since the targeting portion of the intron is less than 400 bp, it is generally more cost-effective to have retargeted introns synthesized commercially in order to minimize labor involved in generating gene knockout mutants [145].

Perhaps the most promising aspect of the Ll.ltrB retrohoming mechanism is that targeted insertion occurs essentially independent of host functions. This has allowed the technology to be broadly applied to at least 11 species of *Clostridium* (Table 2.6), including *C. acetobutylicum*, *C. beijerinckii*, *C. botulinum*, *C. butyricum*, *C. cellulolyticum*, *C. difficile*, *C. perfringens*, *C.*

phytofermentans, *C. sordellii*, *C. sporogenes*, and *C. thermocellum* [14,15,99,128,145,196,197,198]. The ClosTron technology has also been significantly advanced in recent years [http://clostron.com/; [145]]. The entire intron retargeting procedure, from selection of putative insertion sites through to intron mutation and vector construction, has been automated and streamlined. ClosTrons have also been revamped to better suit the modular clostridial shuttle vector repertoire. Still, targeting frequencies have varied drastically, ranging from less than a fraction of a percent up to approximately 60% [14,15], and this can make screening become a major barrier to isolating group II intron gene knockouts. Dong et al. [149] screened approximately 900 transformants using colony PCR in order to identify one knockout colony. To overcome this inefficiency, ClosTrons have been modified to contain an optional retrotransposition-activated marker [RAM; [14]], based on antibiotic selection that is activated only upon intron insertion, but not when carried on the intron-containing plasmid [194]. With the improved RAM functionality, targeting frequency is elevated substantially, generally ranging from 20-100% [14]. However, certain predicted insertion sites still experience inefficient targeting (with targeting frequencies of 2.5-20%) and demand relatively extensive colony screening despite the presence of a selectable RAM. Further, although the lactococcal Ll.ltrB intron represents the model bacterial group II intron for conducting gene knockouts, superior alternatives have been unearthed in recent years. Of particular relevance is the EcI5 intron, which is isolated from a virulent *E. coli* plasmid and has been shown to retrohome several-fold better than Ll.ltrB (up to 98% targeting efficiency without selection) [199]. However, this superior intron has not yet been applied to the clostridia. On the other hand, a thermophilic cyanobacterial group II intron has recently been developed for gene targeting in thermophiles, with *C. thermocellum* acting as the model host species [198]. Thermotargetron disruption efficiencies were found to be similar to the superior EcI5 intron (67-100% targeting efficiency without selection) when assayed using six different chromosomal gene insertions [198]. Higher level integration is presumed to arise from the increased temperatures used for gene targeting (48 °C compared to 37 °C for the Ll.ltrB and EcI5 introns), which enhances DNA melting for more effective base pairing between the intron RNA and host DNA.

Despite its proven utility within *Clostridium*, group II intron-mediated gene knockouts suffer from certain inherent limitations. It is estimated that a 1 kb gene is expected to possess at least five predicted group II intron insertion sites. Disruption of particularly small genes becomes

Table 2.6
Summary of available ClosTron shuttle vectors for performing chromosomal gene knockouts in various *Clostridium* species. Data is adapted and updated from [145], among others.

Organism	Plasmid transfer method	Intron-harboring shuttle vector	Intron promoter[a]	Reference
C. acetobutylicum	Electrotransformation	pSY6	*ptb*	[15]
		pMTL007C-E2	*fdx* (*C. sporogenes*)	[145]
C. beijerinckii	Conjugation or electrotransformation	pMTL007S-E2	*fdx* (*C. sporogenes*)	[145]
		pWJ1	NR	[200]
		pYW1	*ptb*	[201]
C. botulinum	Conjugation or electrotransformation	pMTL007C-E2	*fdx* (*C. sporogenes*)	[145]
		pMTL007C-E5	*fdx* (*C. sporogenes*)	
C. butyricum	Conjugation	pMTL007	*fac*/*fdx* (*C. pasteurianum*)	[196]
		pMTL007C-E2	*fdx* (*C. sporogenes*)	[202]
C. cellulolyticum	Electrotransformation	pWH199	*fdx* (*C. pasteurianum*)	[197]
		pSY6	*ptb* (*C. acetobutylicum*)	[128]
C. difficile	Conjugation	pMTL007C-E2	*fdx* (*C. sporogenes*)	[203]
		pMTL007C-E5	*fdx* (*C. sporogenes*)	[145]
C. perfringens	Electrotransformation	pJIR750ai	*cpb2*	[204]
		pMTL007C-E5	*fdx* (*C. sporogenes*)	[145]
C. phytofermentans	Conjugation	pQint3367	*pfo*	[99]
C. sordellii	Conjugation	pMTL007C-E5	*fdx* (*C. sporogenes*)	[145]
		pDLL2	*gdh*	[205]
C. sporogenes	Conjugation	pMTL007C-E2	*fdx* (*C. sporogenes*)	[145]
C. thermocellum	Electrotransformation	pHK-TT1A	*groEL*	[198]

NR: not reported.
[a] Unless stated otherwise, the promoter controlling intron and IEP transcription is native to the host organism.

difficult, as a 400 bp gene, for example, gives a minimum of only two potential intron insertion sites. Sites that target the antisense strand of a coding gene are generally preferred, as sense orientation insertions can yield conditional knockouts if the LtrA intron-encoded protein remains present [190], thus reducing the number of usable insertion sites. Occasionally sites with high scores, based on the developed computer algorithm [195], still do not result in a stable knockout mutant. Ectopic intron insertions, which occur concomitantly with the desired disruption due to sequence similarity, have also been shown to be a relatively common event in group II intron technology [206]. Such events can be partially circumvented or avoided altogether by optimal intron design and employing an algorithm designed to scan the genome for prospective unwanted insertion sites (TargeTronics, LLC; http://www.targetrons.com/). Still, current clostridial group II intron knockout protocols advocate the use of both gene complementation and Southern blot analyses to rule out the presence of multiple intron insertions in isolated mutants [207], greatly increasing the labor and time for obtaining desired knockout mutants. Lastly, the ClosTron system has been investigated for its capacity to insert heterologous DNA into bacterial genomes, since cargo DNA can be inserted into the Ll.ltrB intron and delivered during intron retrohoming. Targeting efficiency was found to decrease markedly with cargo sequences greater than 1 kb [145], impeding the usefulness of this strategy, especially for introns with a selectable RAM.

Based on the extensive applicability with a broad host range, group II intron technology has served as a powerful tool for clostridial genetic and metabolic engineering. Nonetheless, certain applications, particularly for exogenous DNA delivery, demand more sophisticated chromosomal engineering tools for use in the clostridia.

2.4.7. Advanced and anticipated genome editing technologies

Development of advanced genome editing technologies allowing precise manipulation of clostridial chromosomes, encompassing chromosomal deletions, replacements, point mutations, and insertions, have yet to be realized due to inefficient host homologous recombination and inefficient transformation. The successful isolation of chromosomal integration mutants is also a laborious task and only a few mutant strains have been isolated to date. Such methods have employed both replicative and non-replicative vectors for chromosomal integration using the natural recombination systems of the host organisms [208,209]. Unstable single crossover events tend to dominate, however, leading to unwanted insertion of the entire integrative vector into the

host chromosome [209,210,211]. Such barriers have been partially alleviated in recent years by utilizing negative selection to select for double crossover integration events. Numerous counter-selection markers have been exploited for this purpose in a range of clostridia, including *mazF* from *Bacillus subtilis* [78] and *codA* from *E. coli* [212], in addition to native clostridial *pyrE*, *pyrF*, *galK*, *tdk*, and *hpt* genes [14,79,213,214,215]. Nevertheless, many native clostridial markers are limited in their use, as the chromosomal copy must first be inactivated prior to the introduction of plasmid-encoded counter-selection [212].

To develop more powerful homologous-recombination-based genetic tools, Heap et al. [79] conceived Allele-Coupled Exchange (ACE) for effective selection of double crossover mutants without the use of plasmid-encoded counter-selection. Here, the desired double-crossover recombination event is coupled to either expression or silencing of a selectable marker. This approach was demonstrated by recombination at the *pyrE* locus of *C. acetobutylicum* through positive selection on growth medium lacking uracil and negative selection by supplementing the growth medium with 5-fluoroorotic acid (FOA). By carefully controlling the sequence of integration events using the length of flanking homology regions, single crossover integration of the donor plasmid was selected by *pyrE* positive selection, whereas the double crossover event, resulting in truncation of *pyrE* and plasmid excision, could be subsequently selected via FOA supplementation. The same approach was also used to fuse a promoterless, vector-borne copy of *pyrE* with a strong, native chromosomal promoter, demonstrating the utility of ACE for genome-wide recombination at any desired locus. This method has been employed for the construction of elaborate synthetic chromosomal operons expressing functional cellulosome elements in *C. acetobutylicum* [163].

ACE, together with similar methods developed by Cartman et al. [212] and Al-Hinai et al. [78], exemplifies the current state of the art in clostridial genome editing and proves indispensable for the genetic engineering of a range of medical and industrial clostridia. However, these methods are still considered laborious. The path to simpler and more efficient genome editing techniques will be achieved by closely mimicking the pioneering genetic work done in *E. coli* and *Bacillus subtilis*. The benchmark for performing highly precise and efficient genome editing in *E. coli* over the past decade has been phage-mediated recombinogenic engineering, or recombineering [216]. Rather than relying on natural host recombination, which is inefficient in *E. coli*, or the use of tedious counter-selection markers, recombineering utilizes

efficient phage recombination systems comprised of an endonuclease gene paired with its cognate recombinase, either *recET* from the *E. coli* Rac prophage [216] or *exo* and *bet* genes from phage lambda [217]. These gene products can be employed for recombination of linear PCR cassettes [218]. Additionally, the recombinase (*recT* or *bet*) can be expressed alone for recombination of short single-stranded oligonucleotides possessing homology to host chromosomes [219]. It is expected that clostridial recombineering systems are on the verge of development, as a report detailing site-specific recombination of short synthetic nucleotides appeared recently [220]. Rather than employing the proven *E. coli* recombineering machinery, which function poorly in Gram-positive hosts [221], this method relies on a *recT* homolog from *C. perfringens*. With these proof-of-principle results, it is expected that more practical phage recombination-based demonstrations will emerge in the near future.

CHAPTER 3

Development of an electrotransformation protocol for *C. pasteurianum*

3.1. Introduction

Unlike the past industrial workhorses, *C. acetobutylicum* and *C. beijerinckii*, *C. pasteurianum* has garnered nominal attention as a potential host for the production of butanol. This is largely due to the current inability to transfer DNA to *C. pasteurianum*. Based on early genetic studies, it appears efforts were in place to conduct genetic manipulation of *C. pasteurianum*, since a method for producing and regenerating protoplasts was developed [222] and a Type-II restriction endonuclease was identified as a potential barrier to gene transfer [119]. Successful conjugation-based plasmid transfer to *C. pasteurianum* has also been documented [119], yet no protocol has been described, nor have any genetic mutants arisen from this work. Accordingly, no genetic tools are currently available for the manipulation of *C. pasteurianum*. In order to construct superior *C. pasteurianum* strains through rational genetic and metabolic engineering strategies, it is pertinent to develop methods to transfer foreign DNA to this microorganism.

In contrast to Gram-negative bacteria, Gram-positive cells possess an extensive exterior network of peptidoglycan which physically restricts passage of exogenous DNA into the cell. For this reason, electrotransformation of Gram-positive species is generally less efficient than Gram-negative strains [223]. To overcome the thick Gram-positive cell wall, mild or brief pretreatment using cell-wall-weakening agents, such as lysozyme, glycine, DL-threonine, or penicillin G, is commonly required to achieve optimal electrotransformation while maintaining sufficient cell viability [95,223]. Whereas some species of *Clostridium* can be electrotransformed without the use of cell-wall-weakening agents [10,96], others are electrotransformed at elevated levels when treated with such additives [107,116,134]. Poor electrotransformation efficiency of Gram-positive bacteria is further compounded within the clostridia due to the unusually high production of non-specific cell-wall-associated nucleases [10]. A number of highly-specific clostridial Type-II restriction endonucleases have also been identified [10,96,120], including

CpaAI from *C. pasteurianum* ATCC 6013 [119], highlighting the importance of DNA protection via methylation of the transforming DNA. Unidentified restriction-modification systems are likely the underlying cause of electrotransformation recalcitrance that has been observed with certain species, such as *C. butyricum* [167]. In summary, development of electrotransformation within the clostridia entails investigation of cell-wall-weakening additives, inactivation or evasion of non-specific nucleases, and protection of foreign DNA against highly specific restriction endonucleases, in addition to examination of other common parameters involved in the cell growth, washing and pulse delivery, and recovery phases of the standard bacterial electrotransformation procedure [95,223].

As an entry point to allow genetic manipulation of *C. pasteurianum*, here we report the development of an electroporation-mediated transformation system for *C. pasteurianum* ATCC 6013. CpaAI was validated as a major restriction endonuclease attacking foreign DNA delivered into *C. pasteurianum* and this mechanistic limitation was resolved by *in vivo* methylation of the recognition site (5'-CGCG-3') prior to electroporation. Methylation alone, however, did not result in high-level transfer of DNA into *C. pasteurianum*. Instead, through systematic investigation, we developed an efficient electrotransformation method that is dependent on weakening of the cell wall using glycine, ethanol-mediated membrane solubilization, a low electric field, and osmotic stabilization afforded by sucrose. In addition to the electroporation protocol, we also identified effective antibiotic selection systems and origins of replication capable of sustaining plasmids in *C. pasteurianum*. To our knowledge, this study presents the first demonstration of DNA transfer into *C. pasteurianum* with a high efficiency and opens an avenue for extensive genetic and metabolic engineering of *C. pasteurianum*.

3.2. Methods

3.2.1. Bacterial strains, plasmids, and olgionucleotides

The bacterial strains, plasmids, and oligonucleotides utilized in this work are listed in Table 3.1. *E. coli* DH5α was utilized for routine vector construction and propagation, and *E. coli* ER1821 for maintenance of M.FnuDII-methylated *E. coli-C. pasteurianum* shuttle vectors. *C. pasteurianum* ATCC™ 6013 (Winogradsky 5; W5) was acquired from the American Type Culture Collection (Manassas, VA). Modular pMTL-series shuttle vectors [139] were kindly

provided by Prof. Nigel Minton (University of Nottingham, Nottingham, UK). Plasmids pFnuDIIM [224], pSC12 [225], and pSY6 [15] were respectively provided by Dr. Geoffrey Wilson (New England Biolabs, Inc. (NEB), Ipswich, MA), Prof. George Bennett (Rice University, Houston, TX), and Prof. Sheng Yang (Shanghai Institutes for Biological Sciences, Shanghai, China). Plasmids pHT3 [12] and pIMP1 [10] were provided by Prof. Terry Papoutsakis (University of Delaware, Newark, DE). Oligonucleotide primers were synthesized and purified by Integrated DNA Technologies (IDT; Iowa City, IA) using standard desalting.

3.2.2. Bacterial growth and maintenance conditions

Unless stated otherwise, all chemicals were purchased from Sigma-Aldrich (St. Louis, MO) and stock solutions were prepared according to the manufacturer's recommendations. *E. coli* strains were grown aerobically at 37 °C in lysogeny broth (LB; 10 g/l NaCl, 5 g/l Bacto yeast extract, and 10 g/l Bacto tryptone). Solid and liquid cultures of recombinant *E. coli* were supplemented with 100, 34, or 30 µg/ml of ampicillin, chloramphenicol, and kanamycin, respectively. For selection of strains harboring two compatible plasmids, antibiotic concentrations were reduced by 50%. Recombinant *E. coli* stocks were stored at -80 °C in 15% glycerol. Unless specified otherwise, growth and manipulation of *C. pasteurianum* was performed in a controlled anaerobic atmosphere (85% N_2, 10% H_2, and 5% CO_2) within an anaerobic chamber (Plas-Labs, Inc.; Lansing, MI). Oxygen was purged from growth medium by autoclaving and trace O_2 was reduced using a palladium catalyst fixed to the heating unit of the anaerobic chamber. Agar-solidified medium was prepared aerobically and allowed to equilibrate within the anaerobic chamber for 36 hours prior to use. Anaerobic conditions were monitored by addition of 1 mg/l resazurin to both solid and liquid media. Solid and liquid cultures of recombinant *C. pasteurianum* were supplemented with 15 µg/ml thiamphenicol. Cells were maintained as spores on solidified 2×YTG (16 g/l Bacto tryptone, 10 g/l Bacto yeast extract, 5 g/l glucose, 4 g/l NaCl, and 12 g/l agar) plates. Sporulated agar plate stocks were prepared by streaking colonies from an exponential-phase culture (OD_{600} of 0.4-0.6) and cultivating for more than seven days under anaerobic conditions, followed by exposure and storage in air at 4 °C for up to two months [3]. For long-term storage, vegetative stock cultures (OD_{600} of 0.4-0.6) were prepared and stored at -80 °C in 10% glycerol by inoculating a single sporulated plate colony into 10 ml 2×YTG and heat shocking at 80 °C for 10 minutes to induce germination.

Table 3.1
Strains, plasmids, and oligonucleotides.

Strain	Relevant characteristics	Source or reference
Escherichia coli DH5α	F$^-$ *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG φ80dlacZ*a*M15* a*(lacZYA-argF)U169, hsdR17($r_K^- m_K^+$),* a$^-$	Lab stock
Escherichia coli ER1821	F$^-$ *endA1 glnV44 thi-1 relA1? e14$^-$(mcrA$^-$) rfbD1? spoT1?* a*(mcrC-mrr)114::IS10*	Lab stock; New England Biolabs
Clostridium pasteurianum ATCC 6013	Wild-type	American Type Culture Collection
Plasmid	**Relevant characteristics**	**Source or reference**
pET-20b(+)	*E. coli* pET-series expression vector (Ap^R; ColE1 ori)	Novagen
pETKnFRT	Derived by inserting the FRT-flanked kan gene of pKD4 into the MCS of pET-20b(+) (Ap^R; ColE1 ori; FRT-Kn^R-FRT)	This study
pFnuDIIM	The M.FnuDII methyltransferase gene of *Fusobacterium nucleatum* inserted into the *tet* gene of pACYC184 (p15A ori; Cm^R)	[224]
pFnuDIIMKn	Derived by inserting the FRT-flanked *kan* gene of pETKnFRT into the *cat* gene of pFnuDIIM (p15A ori; Kn^R)	This study
pHT3	*E. coli-C. pasteurianum* shuttle vector containing *lacZ* from *Thermoanaerobacterium thermosulfurogenes* EM1 (Ap^R; ColE1 ori; Erm^R; pIM13 ori)	[12]
pIMP1	*E. coli-C. pasteurianum* shuttle vector (Ap^R; ColE1 ori; Erm^R; pIM13 ori)	[10]
pKD4	Template vector (Ap^R; pR6K ori; FRT-Kn^R-FRT)	[218]
pMTL82151	*E. coli-C. pasteurianum* shuttle vector (Cm^R; ColE1 ori; pBP1 ori)	[139]
pMTL83151	*E. coli-C. pasteurianum* shuttle vector (Cm^R; ColE1 ori; pCB102 ori)	[139]
pMTL84151	*E. coli-C. pasteurianum* shuttle vector (Cm^R; ColE1 ori; pCD6 ori)	[139]
pMTL85141	*E. coli-C. pasteurianum* shuttle vector (Cm^R; ColE1 ori; pIM13 ori)	[139]
pMTL85141ermB	Derived by insertion of the *ermB* gene of pIMP1 into pMTL85141	This study
pSC12	*E. coli-C. pasteurianum* shuttle vector (Cm^R; ColE1 ori; pIM13 ori)	[225]
pSY6	*E. coli-C. pasteurianum* expression vector carrying the *L. lactis ltrB* group II intron under control of the *C. acetobutylicum ptb* promoter, and *ltrA* ORF (Ap^R; ColE1 ori; Erm^R; pIM13 ori)	[15]
pSY6catP	Derived by replacing the *ermB* gene of pSY6 with the *catP* gene from pSC12	This study

Oligonucleotide	Sequence (5'-3')*
KnFRT.BlpI.S	ACACGTGCTCAGCGATTGTGTAGGCTGGAGCTGCTTCG
KnFRT.XhoI.AS	GCCATGCTCGAGATGAATATCCTCCTTAGTTCCTATTCC
ermB.NdeI.S	ATTACGCATATGTTTGGCTAACACACACGCCATTCC
ermB.PvuI.AS	CTTTTTCGATCGTTTCCGACGCTTATTCGCTTCGCT
catP.BclI.S	GTTTGATCATGGTCTTTGTACTAACCTGTGG
pSC12.SOE.AS	tacagcatgaccgttaaagtgg
pSC12.SOE.S	ccactttaacggtcatgctgtaAGTGCAAGGTACACTTGCAAAGTAGTGG
catP.ClaI.AS	GGATCGATCCAACTTAATCGCCTTGCAGCACA
pMTL.seq.S	GGGAGGTCAATCTATGAAATGCG
pMTL.seq.AS	CGGAGCATTTGGCTTTCCTTCCAT

* Lower case: overlap sequences used in SOE PCR; Underline: restriction recognition sequences

3.2.3. Preparation of protoplasts and assay of CpaAI activity

Protoplasts of *C. pasteurianum* were prepared by suspension of cells from a 100 ml culture (OD_{600} of 0.4-0.6) in 25 ml of protoplast buffer (25 mM potassium phosphate, pH 7.0, 6 mM $MgSO_4$, and 15% lactose) containing 200 µg/ml lysozyme, followed by incubation for 45 minutes anaerobically at 37 °C, as described previously [222,226]. For preparation of crude protoplast lysates, 25 ml of protoplasts were collected by centrifugation at 8,500×g for 20 minutes and lysed by resuspension in 20 ml of TEMK buffer (4 mM Tris-HCl, pH 8.0, 10 mM EDTA, 6.6 mM 2-mercaptoethanol, and 25 mM KCl) [10]. After incubation at 37 °C for 1 hour, cell debris was cleared by centrifugation at 20,000×g for 15 minutes and supernatants containing protoplast extracts were stored at -80 °C. CpaAI activity was assayed as previously described [119]. Reaction mixtures contained 1.0 µg plasmid DNA and 25% crude protoplast lysate in a total volume of 20 µl of 1× CpaAI reaction buffer (6 mM Tris-HCl, pH 7.4, 6 mM $MgCl_2$, and 6 mM 2-mercaptoethanol). Optimal digestion occurred at 37 °C for 2-4 hours.

3.2.4. DNA Isolation and manipulation

Plasmid DNA was extracted and purified from *E. coli* DH5α and ER1821 using an EZ-10 Spin Column Plasmid DNA Miniprep Kit from Bio Basic, Inc. (Markham, ON). Recombinant DNA manipulations were performed according to standard procedures [227]. *Taq* DNA polymerase, restriction endonucleases, CpG (M.SssI) and GpC (M.CviPI) methyltransferases, Quick Ligation Kit, and 1 kb DNA ladder were purchased from NEB (Ipswich, MA). *Pfu* DNA polymerase and RNase A were purchased from Bio Basic, Inc. (Markham, ON). All commercial enzymes and kits were used according to the manufacturer's instructions.

Plasmid DNA was extracted and purified from *C. pasteurianum* using a previously described method [12]. Briefly, 3-9 ml of late-exponential phase cells were collected by centrifugation and washed twice in KET buffer (0.5 M KCl , 0.1 M EDTA, and 0.05 M Tris-HCl, pH 8.0) and once in SET buffer (25% sucrose, 0.05 M EDTA, and 0.05 M Tris-HCl, pH 8.0). Cells were then suspended in 200 µl of SET buffer containing 5 mg/ml lysozyme and incubated anaerobically at 37 °C for 20 minutes. RNase A was added to a final concentration of 100 µg/ml and cell lysis and plasmid purification were carried out using the protocol for Purification of Low-Copy Number Plasmid and an EZ-10 Spin Column Plasmid DNA Miniprep

Kit from Bio Basic, Inc. (Markham, ON) beginning with addition of 400 µl of alkaline SDS solution II.

Colony PCR of wild-type and recombinant *C. pasteurianum* was performed by suspending single colonies in 50 µl colony lysis buffer (20 mM Tris-HCl, pH 8.0, containing 2 mM EDTA and 1% Triton X-100), heating in a microwave for 2 minutes at maximum power setting, and adding 1 µl of the resulting cell suspension to a 9 µl PCR containing Standard Taq DNA Polymerase (NEB; Ipswich, MA). An initial denaturation of 5 minutes at 95 °C was employed to further cell lysis. Colonies screened in this manner by suspension in deionized H_2O failed to yield appreciable amplification.

3.2.5. Vector construction

Plasmid pFnuDIIMKn was derived from pFnuDIIM to allow methylation of *E. coli-C. pasteurianum* shuttle vectors and possesses a kanamycin-resistance determinant, as both pFnuDIIM [224] and the *E. coli-C. pasteurianum* shuttle vectors used in this study carry the same chloramphenicol-resistance marker. First, an FRT-*kan*-FRT PCR cassette was amplified from plasmid pKD4 [218] using primers KnFRT.BlpI.S and KnFRT.XhoI.AS and inserted into the MCS of BlpI/XhoI-digested pET-20b(+) (Novagen; Madison, WI) to generate pETKnFRT. Next, the FRT-*kan*-FRT cassette was digested out of pETKnFRT using ScaI and EcoRI and subcloned into the corresponding restriction sites within the *catP* gene of pFnuDIIM to yield pFnuDIIMKn.

Plasmid pSY6catP was derived from pSY6 [15] by swapping the *ermB* marker with the *catP* determinant from pSC12 [225]. The internal BsrGI recognition site within the coding sequence of *catP* was mutated by introducing two silent mutations using splicing by overlap extension (SOE) PCR to prevent interference with future group II intron retargeting, which requires use of BsrGI. The *catP* gene was amplified in two parts from template pSC12 using primer sets catP.BclI.S/pSC12.SOE.AS and pSC12.SOE.S/catP.ClaI.AS with 22 bp of overlap between products. The resulting overlapping PCR products were separated on a 2.0% agarose gel, pierced three times with a P10 micropipette tip, and used as template in a SOE PCR by cycling for 10 cycles prior to adding primers catP.BclI.S and catP.ClaI.As and cycling for 25 additional cycles. The mutated PCR product was purified using a EZ-10 Spin Column PCR

Products Purification Kit (Bio Basic, Markham, ON), digested with BclI/ClaI, and inserted into the corresponding sites of pSY6 to generate pSY6catP.

Plasmid pMTL85141ermB was derived from pMTL85141 via insertion of the *ermB* marker from pIMP1 into pMTL85141. The *ermB* gene and associated promoter was PCR-amplified from template pIMP1 using primers ermB.NdeI.S and ermB.PvuI.AS. The resulting 1,238 bp PCR product was purified using an EZ-10 Spin Column PCR Products Purification Kit (Bio Basic, Markham, ON), digested with NdeI/PvuI, and inserted into the corresponding sites of pMTL85141 to generate pMTL85141ermB.

3.2.6. Preparation of electrocompetent cells and electrotransformation

For preparation of electrocompetent cells of *C. pasteurianum* using the high-level protocol, a seed culture was first prepared by inoculating 20 ml of reduced 2×YTG with 0.2 ml of a thawed glycerol stock. The culture was then 20^{-2}-diluted and, following overnight growth at 37 °C, 1 ml of the seed culture was transferred to a 125 ml Erlenmeyer flask containing 20 ml of reduced 2×YTG. Cells were grown to early exponential phase (OD_{600} of 0.3-0.4), at which time filter-sterilized stock solutions of 2 M sucrose and 18.77% glycine were added to respective concentrations of 0.4 M and 1.25%. Growth was resumed until the culture attained an OD_{600} of 0.6-0.8 (approximately 2-3 h) and 20 ml culture was transferred to a 50 ml pre-chilled, screw-cap centrifuge tube. At this point, all manipulations were performed at 4 °C using an ice-bath and pre-chilled reagents. Cells were removed from the anaerobic chamber and collected by centrifugation at 8,500×g and 4 °C for 20 minutes. The resulting cell pellet was returned to the anaerobic chamber and washed once in 5 ml of filter-sterilized SMP buffer (270 mM sucrose, 1 mM $MgCl_2$, and 5 mM sodium phosphate, pH 6.5). Following centrifugation, the final cell pellet was resuspended in 0.6 ml SMP buffer.

For transfer of plasmids to *C. pasteurianum*, *E. coli-C. pasteurianum* shuttle vectors were first co-transformed with pFnuDIIMKn into *E. coli* ER1821 to methylate the external cytosine residue within 5'-CGCG-3' tetranucleotide recognition sites of CpaAI. Plasmid mixtures were then isolated and 0.5 µg, suspended in 20 µl of 2 mM Tris-HCl, pH 8.0, was added to 580 µl of *C. pasteurianum* electrocompetent cells. The cell-DNA mixture was transferred to a pre-chilled electroporation cuvette with 0.4 cm gap (Bio-Rad; Richmond, CA), 30 µl of cold 96% ethanol was added, and the suspension was incubated on ice for 5 minutes. A single exponential decay

pulse was applied using a Gene Pulser (Bio-Rad, Richmond, CA) set at 1.8 kV, 25 µF, and ∞ Ω, generating a time constant of 12-14 ms. Immediately following pulse delivery, the cuvette was flooded with 1 ml 2×YTG medium containing 0.2 M sucrose and the entire suspension was transferred to 9 ml of the same medium. Recovery cultures were incubated for 4-6 hours prior to plating 50-250 µl aliquots onto 2×YTG agar plates containing 15 µg/ml thiamphenicol, 4 µg/ml clarithromycin, or 20 µg/ml erythromycin. Plates were incubated for 2-4 days under secondary containment within 3.4 L Anaerobic Jars each equipped with a 3.5 L Anaerobic Gas Generating sachet (Oxoid Thermo Fisher; Nepean, ON).

3.3. Results

3.3.1. Protection of plasmid DNA from CpaAI restriction

To develop a *C. pasteurianum* transformation protocol, we first assayed crude cell lysates for the presence of restriction-modification systems, which potently inhibit plasmid DNA transfer to bacteria. At least one Type-II restriction endonuclease, designated CpaAI with 5'-CGCG-3' recognition and an isoschizomer of ThaI and FnuDII, has been previously identified in cell-free lysates of *C. pasteurianum* ATCC 6013 [119]. We initially prepared crude cell lysates through sonication of whole cells. As found in other species, such as *C. acetobutylicum* [10,228], lysates generated in this manner potently degraded all plasmid DNA substrates, presumably due to non-specific cell-wall-associated nucleases (data not shown). To overcome non-specific nuclease activity, we then aimed to assay CpaAI restriction activity using protoplast extracts, which allowed clear detection of CpaAI activity. Optimal digestion occurred between 2-4 hours incubation at 37 °C and produced a restriction pattern identical to that of BstUI, a commercial isoschizomer of CpaAI (Figure 3.1A). Since all known BstUI isoschizomers catalogued in REBASE [90] are sensitive to methylation of both external cytosine residues within the 5'-CGCG-3' recognition sequence, we next assessed the effect of external cytosine methylation by expression of the M.FnuDII methyltransferase (with 5'-^{m}CGCG-3' methylation site of both DNA strands) from plasmid pFnuDIIMKn. M.FnuDII methylation protected pMTL85141, an *E. coli-Clostridium* shuttle vector [139], from degradation by CpaAI and BstUI (Figure 3.1B). While unmethylated substrates were significantly restricted after 2 hours incubation at 37 °C, M.FnuDII-methylated substrates were completely resistant to cleavage, even after 8 h. Note that

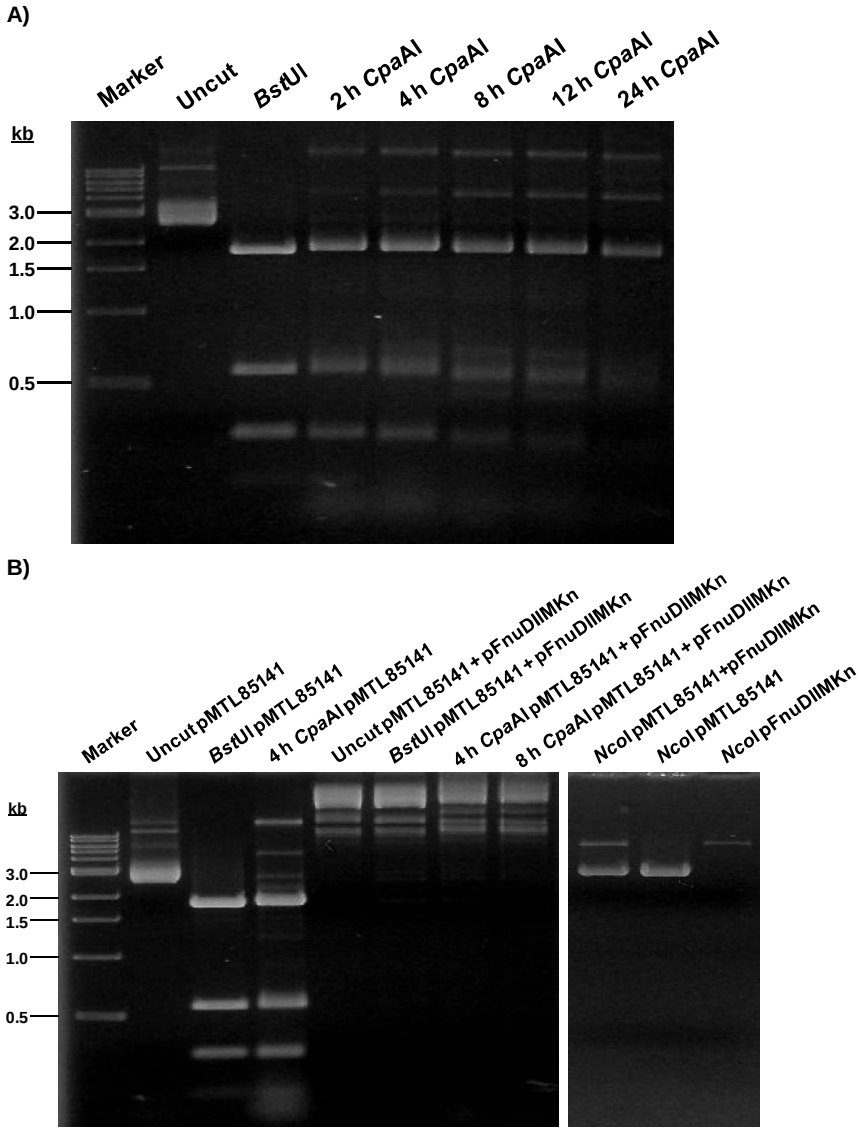


Figure 3.1. M.FnuDII methyltransferase-mediated protection of pMTL85141 against CpaAI endonuclease. A) Time course digestion of pMTL85141 using crude protoplast extracts possessing CpaAI restriction activity, resolved on a 2% agarose gel. Digestion reactions contained 1.0 µg pMTL85141 and 25% protoplast extract in a total volume of 1× CpaAI custom buffer. For comparison, pMTL85141 is shown undigested and digested with BstUI, a commercial isoschizomer of CpaAI. Expected digestion products are 1785, 581, 270, 252, and 75 bp. **B)** M.FnuDII-mediated protection of pMTL85141 from CpaAI digestion (left panel). Protoplast extract digestions contained 1.0 µg pMTL85141 or pMTL85141+pFnuDIIMKn, the vector harboring the M.FnuDII methyltransferase gene, and 25% protoplast extract in a total volume of 1× CpaAI custom buffer. pMTL85141 preparation in the presence of plasmid pFnuDIIMKn afforded protection of pMTL85141 from both BstUI and CpaAI restriction, as no digestion products could be detected. Methylation treatment resulted in the presence of high-molecular weight bands. Linearization of the high molecular weight bands by NcoI digestion (right panel) confirmed the presence of pMTL85141 (2,963 bp) and the methylating plasmid, pFnuDIIMKn (6,449 bp), at the correct sizes of the individual linearized vectors.

methylated pMTL85141 plasmid preparations, which also contains the pFnuDIIMKn methylating plasmid, migrated at a different molecular weight than unmethylated plasmid preparations. However, when we linearized the double-plasmid preparation, in addition to preparations of the two individual plasmids, we observed no detectable changes in plasmid size or unexpected products (Figure 3.1B, right panel). *In vitro* methylation with commercial M.SssI (5'-^{m}CG-3' methylation site) and M.CviPI (5'-GmC-3' methylation site) methyltransferases also protected plasmids from digestion by CpaAI in protoplast extracts and commercial BstUI (not shown).

3.3.2. Initial electrotransformation of C. pasteurianum

To electrotransform *C. pasteurianum*, we employed a series of *E. coli-Clostridium* shuttle vectors which differ only in their Gram-positive origins of replication: pMTL82151 (pBP1 ori from *C. botulinum*); pMTL83151 (pCB102 ori from *C. butyricum*); pMTL84151 (pCD6 ori from *C. difficile*); and pMTL85141 (pIM13 ori from *Bacillus subtilis*) [139]. We utilized conditions common to clostridial electrotransformation procedures (Table 3.2) and M.FnuDII-methylated DNA. Of the four vectors tested, pMTL83151, pMTL84151, and pMTL85141 yielded colonies using thiamphenicol selection, corresponding to electrotransformation efficiencies of 0.7×10^1, 0.3×10^1, and 2.4×10^1 transformants μg^{-1} DNA, respectively. Accordingly, pMTL85141 was selected as the vector used for all subsequent electrotransformation work. Importantly, no transformants were obtained with unmethylated plasmid, validating the necessity to protect transforming DNA against the endogenous CpaAI restriction endonuclease. Interestingly, while *in vivo* methylation was essential for transformation, we did not obtain transformants when pMTL85141 was methylated *in vitro* with M.SssI or M.CviPI methyltransferases, although both enzymes protect pMTL85141 from digestion by CpaAI. This result is unexpected, but is speculated upon later in this report.

To confirm the presence of pMTL85141 in transformed colonies, we screened thiamphenicol-resistant colonies for the presence of the *catP* resistance marker within pMTL85141 using colony PCR (Figure 3.2A). All of the colonies screened generated a single expected product of 518 bp. To further confirm the presence of plasmid and determine if rearrangements or recombinations occurred upon transfer to *C. pasteurianum*, plasmid pMTL85141was isolated and purified from thiamphenicol-resistant colonies and digested with

Table 3.2
Consensus clostridial electrotransformation conditions leading to initial low-level transformation of *C. pasteurianum*

Electrotransformation parameter	Consensus for *Clostridium* species	Low-level transformation of *C. pasteurianum*	Selected references
Cell growth			
Growth medium	YTG or 2×YTG	2×YTG	[92,93,113]
Growth phase and OD_{600} at time of harvest	mid to late exponential phase (OD_{600} 0.5–0.9)	OD_{600} 0.6-0.8	[96,229]
Washing and pulse delivery			
Wash and electroporation buffer	5–7 mM sodium phosphate, pH 6.5-7.4, containing 270 mM sucrose and 1 mM $MgCl_2$	5 mM sodium phosphate, pH 6.5, containing 270 mM sucrose and 1 mM $MgCl_2$	[10,96,116]
Number of wash steps	1	1	[92,102,113]
Cuvette gap width	0.4 cm	0.4 cm	[102,104,116]
Volume of cells	600 µl	600 µl	[10]
Pulse parameters	2.0–2.5 kV; 25 µF; 200–800 Ω; 4–8 ms	2.0–2.5 kV; 25 µF; 200–800 Ω; 6-9 ms	[92,104,115]
Outgrowth			
Recovery and plating medium	YTG or 2×YTG	2×YTG	[93,113]
Transformation efficiency	Up to 10^6 transformants μg^{-1} DNA	2.4×10^1 transformants μg^{-1} pMTL85141	[138]

XhoI. XhoI digestion of all plasmid preparations from *C. pasteurianum* yielded a single band on a 1.0% agarose gel, similar to the digestion of pMTL85141 prepared from *E. coli* DH5α (Figure 3.2B). The presence of the methyltransferase vector, pFnuDIIMKn, could not be detected in *C. pasteurianum* plasmid preparations.

3.3.3. High-level electrotransformation of C. pasteurianum

An electrotransformation efficiency of 2.4×10^1 transformants μg^{-1} DNA is very low relative to the efficiency of other clostridia (i.e., up to 10^6 transformants μg^{-1} DNA [138]) and did not permit transfer of certain gene knockout vectors, as described later in this report. Therefore, we were prompted to develop a high-level electrotransformation protocol by systematically investigating the effects on electrotransformation efficiency of parameters throughout all phases of the electrotransformation procedure. Electrotransformation efficiencies reported below represent the average of at least two electrotransformation experiments using the same preparation of electrocompetent cells.

(i) Cell-wall-weakening

We first investigated the use of cell-wall-weakening agents due to their potential to greatly enhance electrotransformation by weakening of the Gram-positive cell wall [95]. A screening experiment was conducted to identify potential additives capable of enhancing electrotransformation of *C. pasteurianum*, including glycine, DL-threonine, lysozyme, and penicillin G (Figure 3.3A). Individually, we screened the effect of glycine and DL-threonine by supplying the additives in the presence of 0.25 M sucrose at the first signs of growth (OD_{600} of 0.3-0.4) because cultures failed to grow to sufficient cell densities if glycine or DL-threonine were present without sucrose supplementation or if the additives were present at the time of inoculation. Cell growth rate was slightly reduced in the presence of both glycine and DL-threonine. On the other hand, lysozyme and penicillin G were screened by addition at the wash stage in the wash and electroporation buffer, followed by incubation at 37 °C for 30 minutes prior to electroporation. Additive concentrations were chosen based on previous electrotransformation studies with various species of Gram-positive bacteria [116,137,230,231,232]. Of the four additives screened, only glycine and DL-threonine improved the electrotransformation efficiency. The samples treated with 40 µg/ml lysozyme and 30 µg/ml

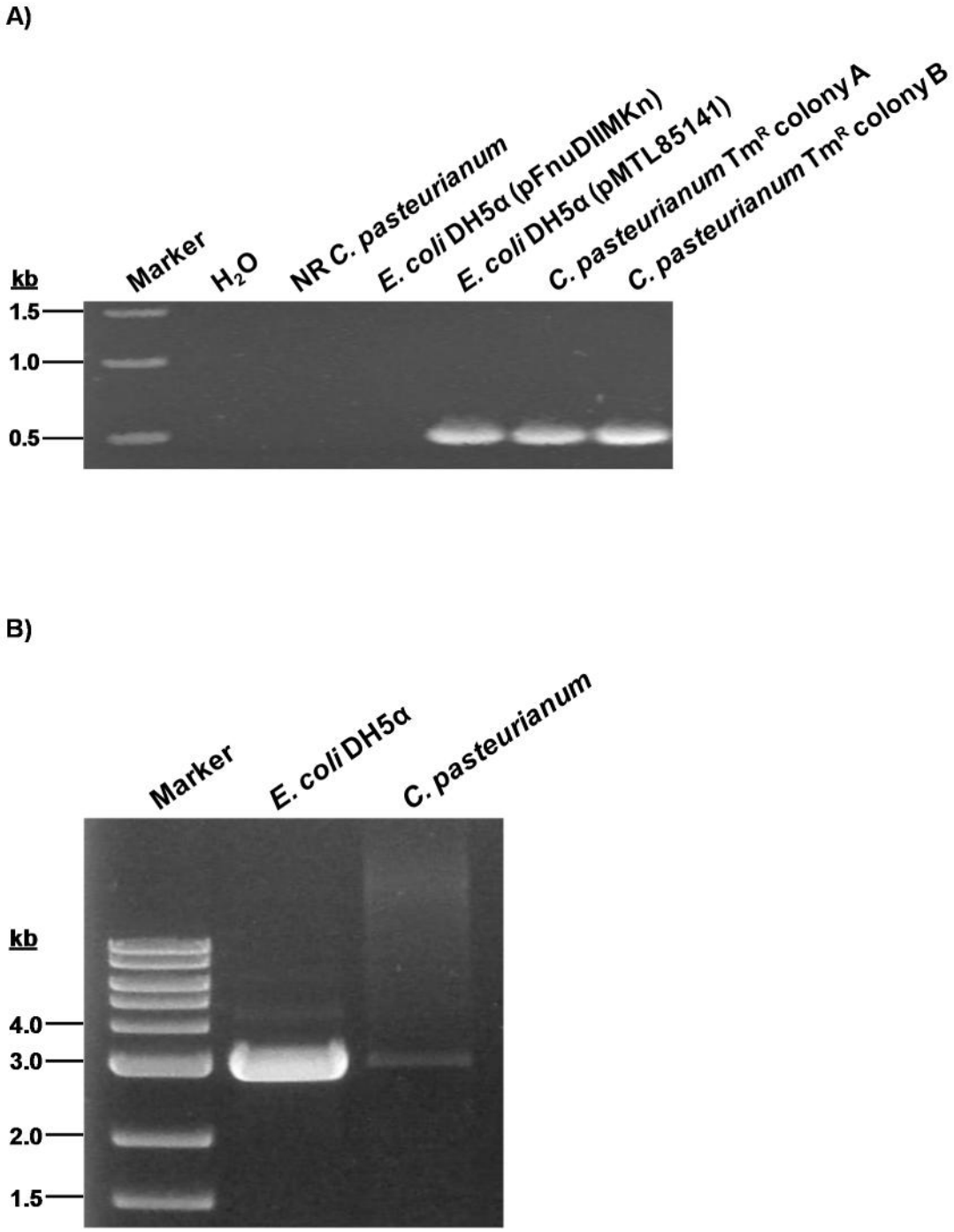
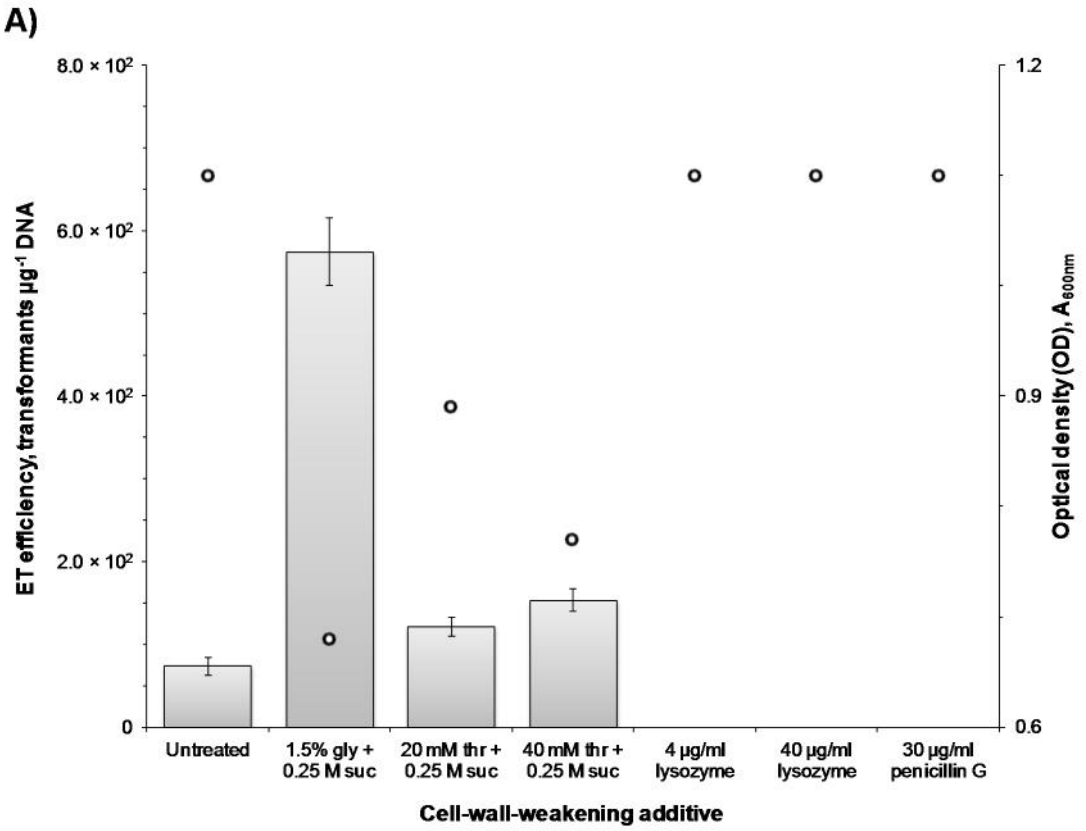


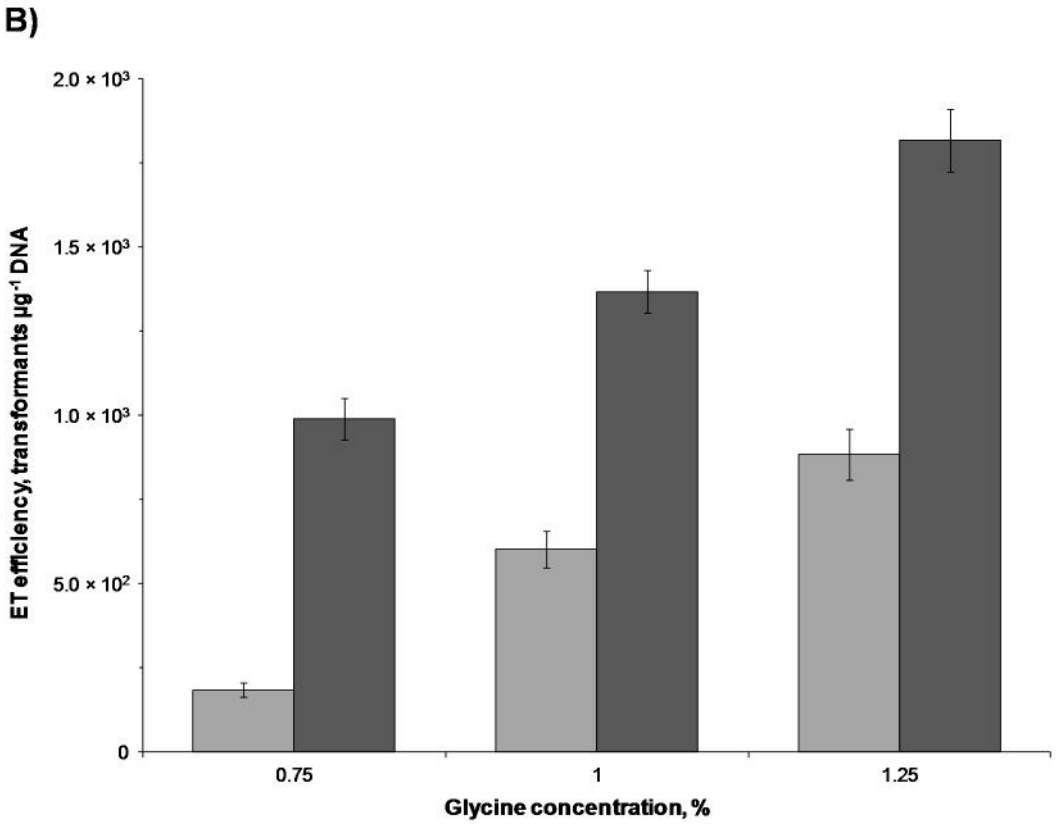
Figure 3.2. Low-level electrotransformation of *C. pasteurianum*. A) Colony PCR confirmation of pMTL85141 presence in *C. pasteurianum* transformants using primers pMTL.seq.S and pMTL.seq.AS. The expected product of 518 bp could only be amplified from *E. coli* transformed with pMTL85141 and from thiamphenicol-resistant *C. pasteurianum* colonies, and not from non-recombinant *C. pasteurianum*. **B)** XhoI-linearized pMTL85141 plasmid prepared from *E. coli* DH5α and from a representative transformant of *C. pasteurianum* showing the expected plasmid size of 2,963 bp. Some undigested vector remains visible in the *E. coli* preparation.

penicillin G even failed to grow during the outgrowth period following electroporation, potentially due to cell lysis. Despite a slight inhibition on cell growth, more than 7-fold enhancement of electrotransformation efficiency was attained using 1.5% glycine, compared to the control experiment with no cell-wall-weakening agent. Supplementation of 20 and 40 mM DL-threonine provided approximately 1.6- and 2.1-fold increases, respectively, in electrotransformation efficiency. Although glycine and DL-threonine have different mechanisms of cell wall disruption, combining glycine and DL-threonine treatments did not lead to a synergistic increase in electrotransformation efficiency.

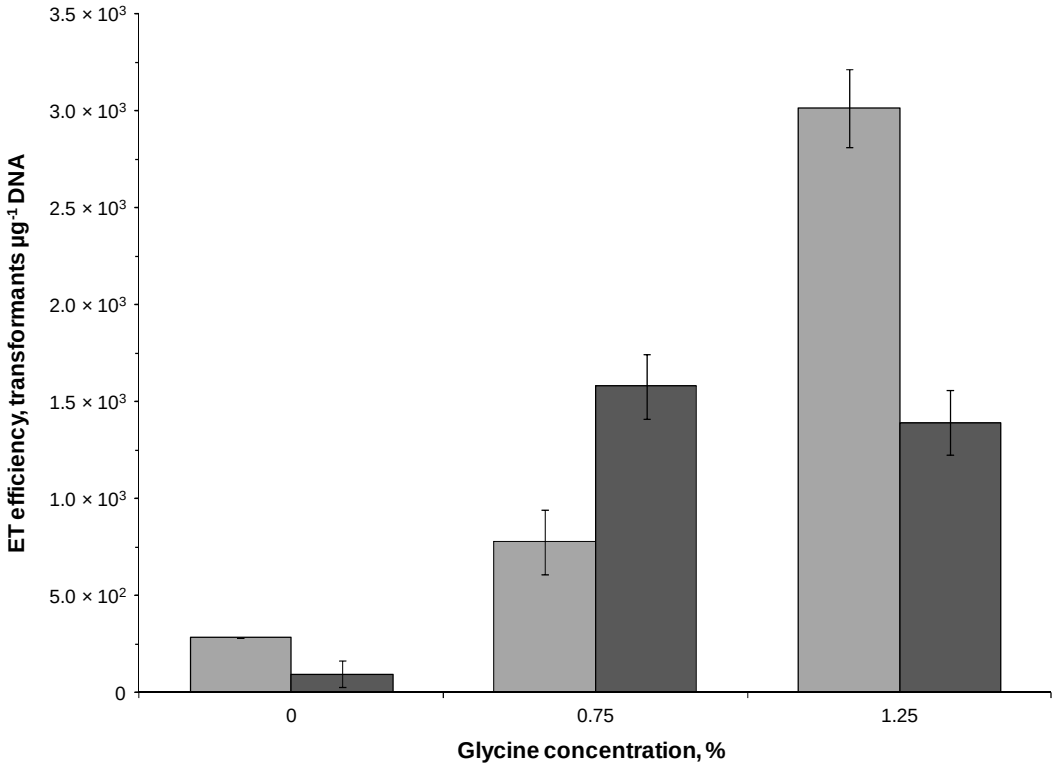
As a result of the clear benefit of glycine on the electrotransformation efficiency, we set out to determine the optimum glycine regimen with respect to concentration and duration of exposure. This investigation was done concomitant with investigating the effect of sucrose on electrotransformation efficiency by providing osmoprotection during the various cell-wall-weakening glycine treatments. We tested glycine at 0.75, 1.0, and 1.25% in the presence of either 0.25 or 0.4 M sucrose, corresponding to nearly isotonic and hypertonic extracellular environments, respectively. The highest glycine concentration was selected as 1.25% to minimize growth inhibition, which becomes significant at concentrations equal to or greater than 1.5%. Increasing the sucrose concentration from 0.25 to 0.4 M led to a significant increase in electrotransformation efficiency under all glycine concentrations tested (Figure 3.3B). To examine the effect of the duration of glycine exposure on electrotransformation efficiency, cultures were incubated with 0, 0.75, or 1.25% glycine in the presence of 0.4 M sucrose starting at an OD_{600} of 0.4 for either 2.5 or 4.5 hours prior to washing and pulse delivery (Figure 3.3C). Maximum electrotransformation efficiency was attained by exposing cells to 1.25% glycine for 2.5 hours in the presence of 0.4 M sucrose, a 10.7-fold increase compared to the untreated control culture. Interestingly, lower glycine concentrations could be compensated for by increasing the duration of exposure. When using a glycine concentration of 0.75% in the growth medium, 4.5 hours rather than 2.5 hours of exposure generated a greater electrotransformation efficiency at this lower glycine concentration, although the absolute gain in electrotransformation efficiency was still lower than with 1.25% glycine.

Figure 3.3. Investigation of cell-wall-weakening and osmoprotection on electrotransformation of *C. pasteurianum*. A) Investigation of cell-wall-weakening agents. Cells were grown to early exponential phase (OD_{600} 0.3-0.4) and glycine (gly) or DL-threonine (thr) was added along with 0.25 M sucrose (suc). For lysozyme and penicillin G treatments, additives were supplemented to buffer SMP prior to electroporation and incubated anaerobically at 37 °C for 30 minutes. An untreated culture was included as a control. The OD_{600} of each culture at time of harvest is shown (O). Pulse duration was unaffected between samples. **B)** Investigation of glycine and sucrose concentrations. Six cultures were grown to an OD_{600} of 0.4 and glycine was added to a final concentration of 0.75, 1.0, or 1.25% together with sucrose at 0.25 (light shading) or 0.4 M (dark shading). Growth was minimally affected between samples, as all cultures attained a final OD_{600} of 1.2-1.5. Pulse duration was unaffected between samples. **C)** Investigation of glycine concentration and duration of exposure. Two cultures were grown to an OD_{600} of 0.4 and glycine was added to a final concentration of 0.75 or 1.25% together with 0.4 M sucrose. An additional control culture was prepared without either glycine or sucrose supplementation. Cells were harvested, washed, and electroporated at either 2.5 (light shading) or 4 hours (dark shading) following supplementation with glycine and sucrose. Pulse duration was unaffected between samples. **D)** Effect of sucrose concentration within the wash and electroporation buffer. Cultures were washed and electroporated in SMP buffer containing either isotonic (0.27 M) or hypertonic sucrose (0.5 M). Pulse duration was unaffected between samples. **E)** Effect of sucrose concentration within the outgrowth medium. Cultures were electroporated and resuspended and grown in 2×YTG medium containing either 0.2 or 0.4 M sucrose.

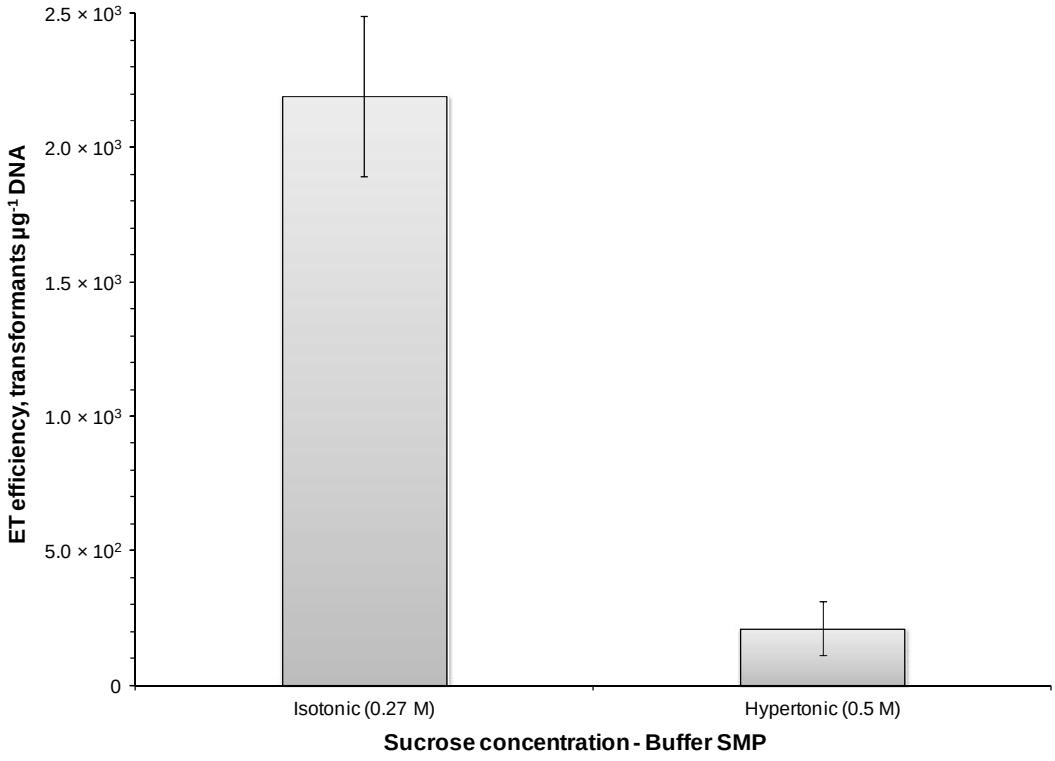

A)
8.0 × 10²
6.0 × 10²
4.0 × 10²
2.0 × 10²
0
ET efficiency, transformants µg⁻¹ DNA
1.2
0.9
0.6
Optical density (OD), A600nm
Untreated
1.5% gly + 0.25 M suc
20 mM thr + 0.25 M suc
40 mM thr + 0.25 M suc
4 µg/ml lysozyme
40 µg/ml lysozyme
30 µg/ml penicillin G
Cell-wall-weakening additive


B)
2.0 × 10³
1.5 × 10³
1.0 × 10³
5.0 × 10²
0
ET efficiency, transformants µg⁻¹ DNA
0.75
1
1.25
Glycine concentration, %

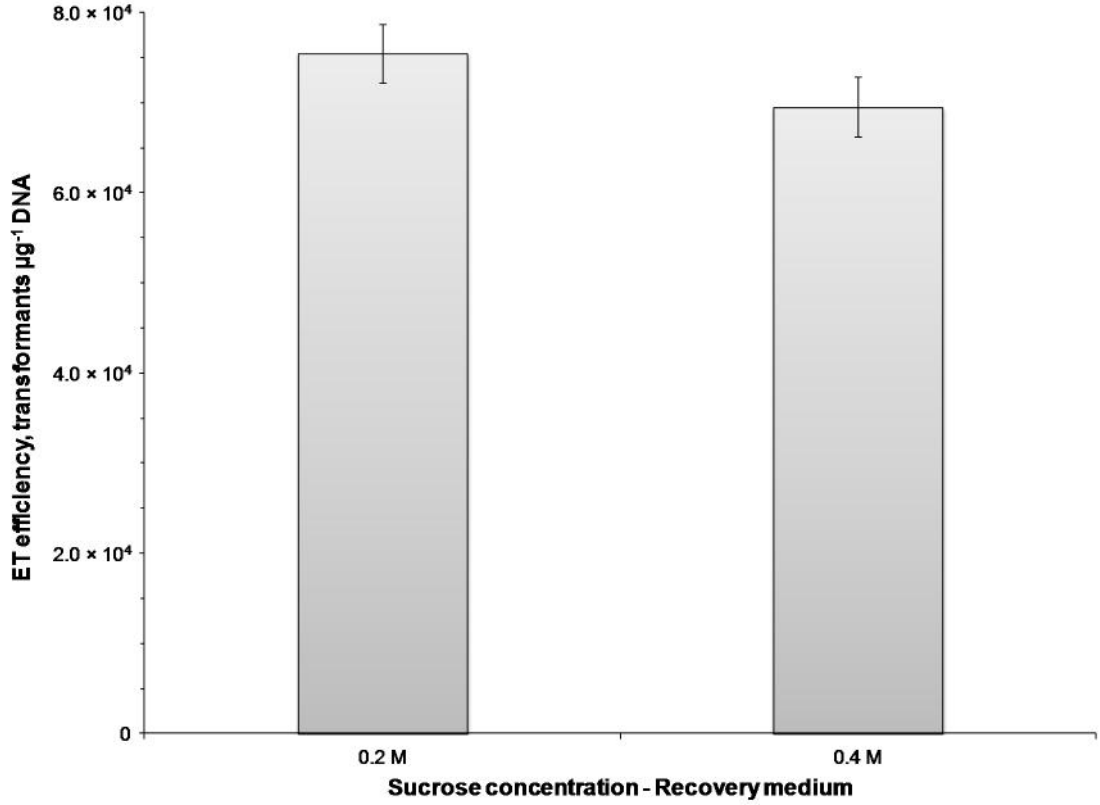
C)


3.5 × 10³
3.0 × 10³
2.5 × 10³
2.0 × 10³
1.5 × 10³
1.0 × 10³
5.0 × 10²
0
ET efficiency, transformants µg⁻¹ DNA
0
0.75
1.25
Glycine concentration, %

D)


2.5 × 10³
2.0 × 10³
1.5 × 10³
1.0 × 10³
5.0 × 10²
0
ET efficiency, transformants µg⁻¹ DNA
Isotonic (0.27 M)
Hypertonic (0.5 M)
Sucrose concentration - Buffer SMP

E)


8.0 × 10⁴
6.0 × 10⁴
4.0 × 10⁴
2.0 × 10⁴
0
ET efficiency, transformants µg⁻¹ DNA
0.2 M
0.4 M
Sucrose concentration - Recovery medium

(ii) Osmoprotection

We continued to investigate the effect of the osmoprotectant concentration on electrotransformation efficiency during the subsequent washing and electroporation phase and the outgrowth phase following electroporation. Cells grown in the presence of 1.25% glycine and 0.4 M sucrose were washed and electroporated in the common clostridial SMP buffer containing either 0.27 M (isotonic) or 0.5 M (hypertonic) sucrose (Figure 3.3D). SMP buffer outperformed other buffers tested, such as 10% PEG 8000, 15% glycerol, protoplast buffer with lysozyme omitted, and SMP buffer supplemented with 15% glycerol (data not shown). Hypertonic sucrose, which improved the electrotransformation efficiency when included during the growth phase, reduced electrotransformation efficiency by a factor of 10.3 when included at the washing and electroporation phase. Thus, 0.27 M sucrose was adopted as the optimum sucrose concentration in the wash and electroporation buffer.

To assess the effect of sucrose osmoprotection during cell recovery immediately following delivery of the electric pulse, cells were grown, made electrocompetent, pulsed, and resuspended in 10 ml 2×YTG containing either 0.2 or 0.4 M sucrose (Figure 3.3E). Similar to the washing and electroporation phase, hypertonic sucrose again reduced electrotransformation efficiency, although the effect was modest (a 1.1-fold decrease), and thus, 0.2 M was adopted as the optimum sucrose concentration in the outgrowth medium.

(iii) Cell membrane solubilization

After developing a regimen to weaken the exterior cell wall while supporting cell viability with sucrose osmoprotection, we next sought to enhance transfer of plasmid DNA to *C. pasteurianum* with the use of ethanol to solubilize the cell membrane, a strategy which has proved effective with some species of Gram-negative bacteria [135,136]. We also extended this approach to butanol, which elicits a more pronounced toxic effect on cells. To achieve maximum membrane solubilization without adversely affecting cell viability, we utilized concentrations near the toxicity threshold for many species of *Clostridium*, which were up to 15% (v/v) for ethanol and 2% (v/v) for butanol [233,234]. Five minutes prior to electroporation, ethanol or butanol was added directly to the cell-DNA suspension. Ethanol added at 5 and 10% provided a 1.6- and 1.3-fold respective increase in electrotransformation efficiency, compared to the control experiment with no ethanol treatment (Figure 3.4). Butanol, and ethanol at an elevated

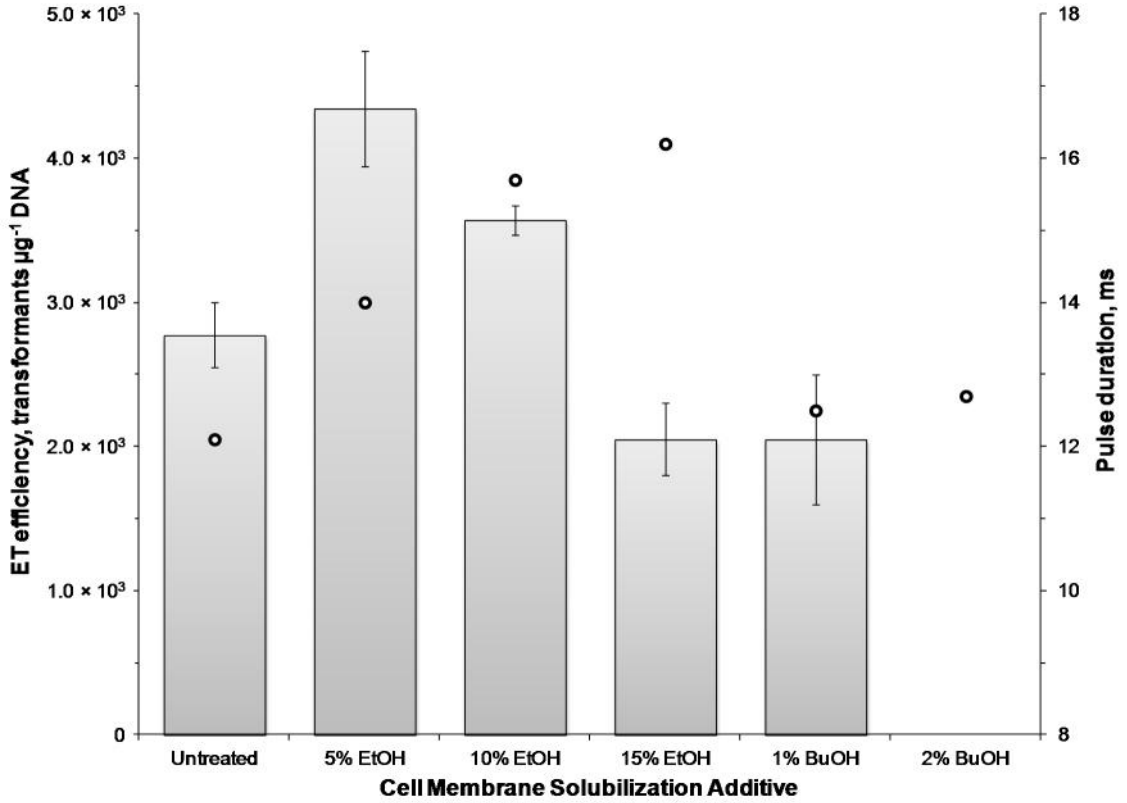


Figure 3.4. Investigation of membrane permeabilization on the electrotransformation of *C. pasteurianum*. Immediately prior to pulse delivery, cell-DNA suspensions were supplemented with 5, 10, or 15% ethanol (EtOH) or 1 or 2% butanol (BuOH). An untreated sample was included as a control. The time constant of each pulse is shown (O). The sample treated with 2% butanol failed to grow during the allotted 16-hour recovery period following electroporation.

concentration of 15%, proved to be detrimental to electrotransformation. The 2% butanol sample grew extremely slowly during the outgrowth period following electroporation. The addition of ethanol increased the pulse time constant, which may have influenced electrotransformation efficiency (Figure 3.4). Butanol did not significantly affect the pulse time constant.

(iv) Electric pulse parameters

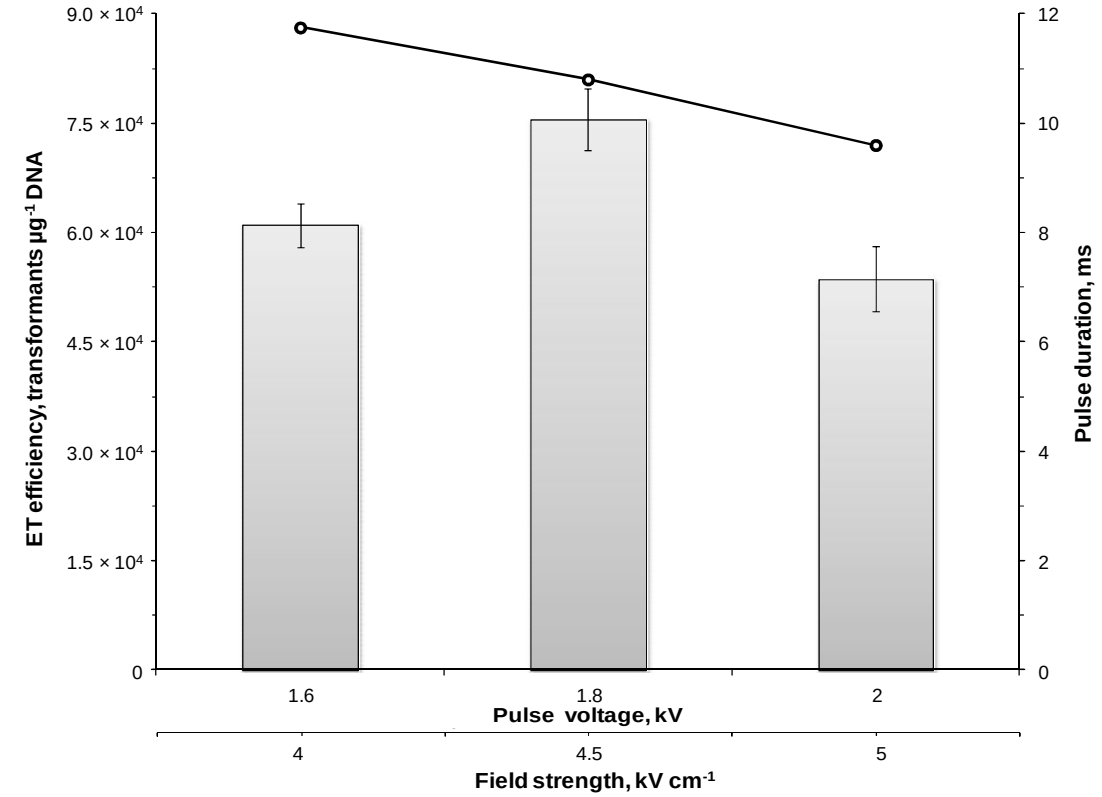
We investigated the effects of the electrical pulse with respect to voltage (i.e., field strength), capacitance, and resistance (Figure 3.5A-C). In an initial screening experiment, low voltages in the range of 1.8-2.0 kV generated significantly more transformants than voltages of 2.0-2.5 kV (data not shown), which are representative of most electrotransformation protocols using species of *Clostridium* [92,102,104,116]. Hence, pulses of 1.6, 1.8, and 2.0 kV were administered, corresponding to field strengths of 4.0, 4.5, and 5.0 kV cm^{-1} (Figure 3.5A), using a capacitance of 25 µF and a resistance of ∞ Ω (i.e., without the use of Pulse Controller module). A voltage of 1.8 kV was found to produce the greatest electrotransformation efficiency, although pulses of 1.6 and 2.0 kV only slightly reduced the electrotransformation efficiency. Pulse duration decreased by approximately 1 ms when increasing pulse voltage from 1.6 to 1.8 kV and from 1.8 to 2.0 kV. Next, capacitances of 25 and 50 µF were assessed at voltages of 1.8 and 2.25 kV, and ∞ Ω (Figure 3.5B). At both voltages, increasing the capacitance from 25 to 50 µF reduced electrotransformation efficiency by a factor of 2.7 (1.8 kV) and 15.6 (2.25 kV), respectively. Similarly, decreasing resistance from ∞ Ω to 200 and 600 Ω, at 2.25 kV and 25 µF was unproductive and resulted in a 3.3- and 2.3-fold decrease in electrotransformation efficiency, respectively (Figure 3.5C). Pulse duration changes were not predictive of the effects on electrotransformation efficiency, as increases in the time constant accompanying changes in capacitance and decreases in the time constant accompanying changes in resistance both correlated with decreased electrotransformation efficiency.

(v) DNA quantity and outgrowth duration

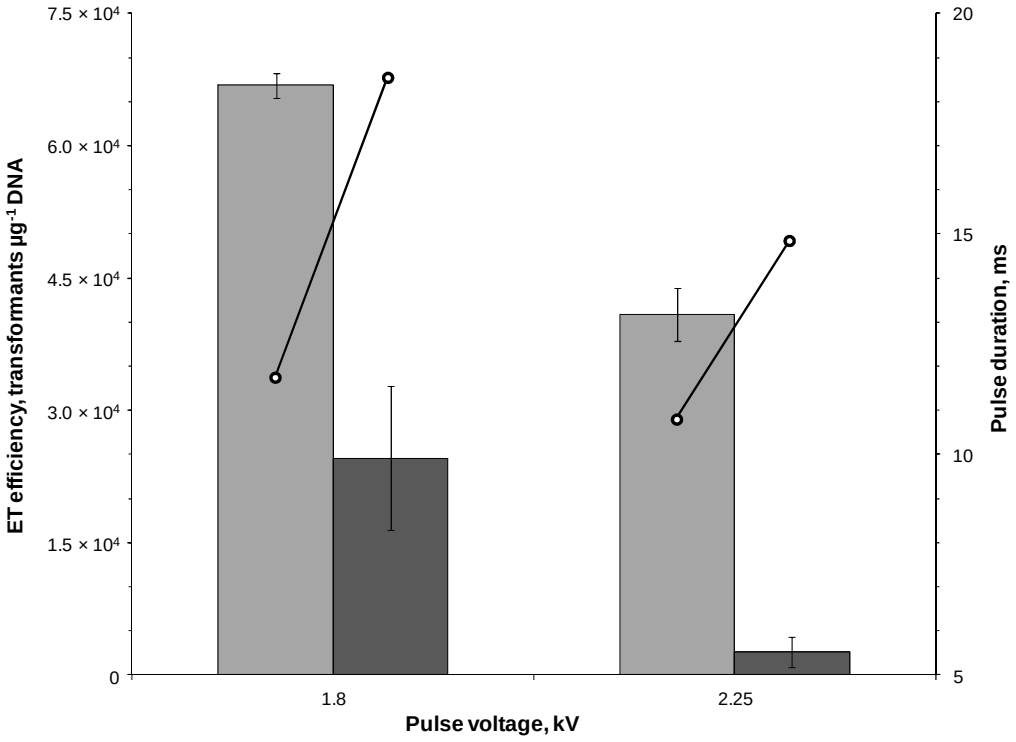
Finally, we evaluated the effect of DNA amount on both number of transformants and electrotransformation efficiency (Figure 3.6A) and the effect of the duration of outgrowth following electroporation. Although the total number of transformants was found to increase linearly between 0.5 and 5.0 µg of pMTL85141, the greatest electrotransformation efficiency

Figure 3.5. Investigation of electric pulse parameters on the electrotransformation of *C. pasteurianum*. A) Effect of pulse voltage (field strength). Electrotransformation efficiency was measured using electric pulses of 1.6, 1.8, or 2.0 kV, corresponding to field strengths of 4.0, 4.5, and 5.0 kV cm^{-1}. The time constant of each pulse is shown (⭘). **B)** Effect of pulse capacitance. Electrotransformation efficiency was measured at 25 (light shading) and 50 µF (dark shading) under voltages of 1.8 and 2.25 kV. The time constant of each pulse is shown (⭘). **C)** Effect of pulse resistance. Electrotransformation efficiency was measured at 200, 600, and ∞ Ω at a voltage of 2.25 kV. The time constant resulting from each administered pulse is shown (⭘).

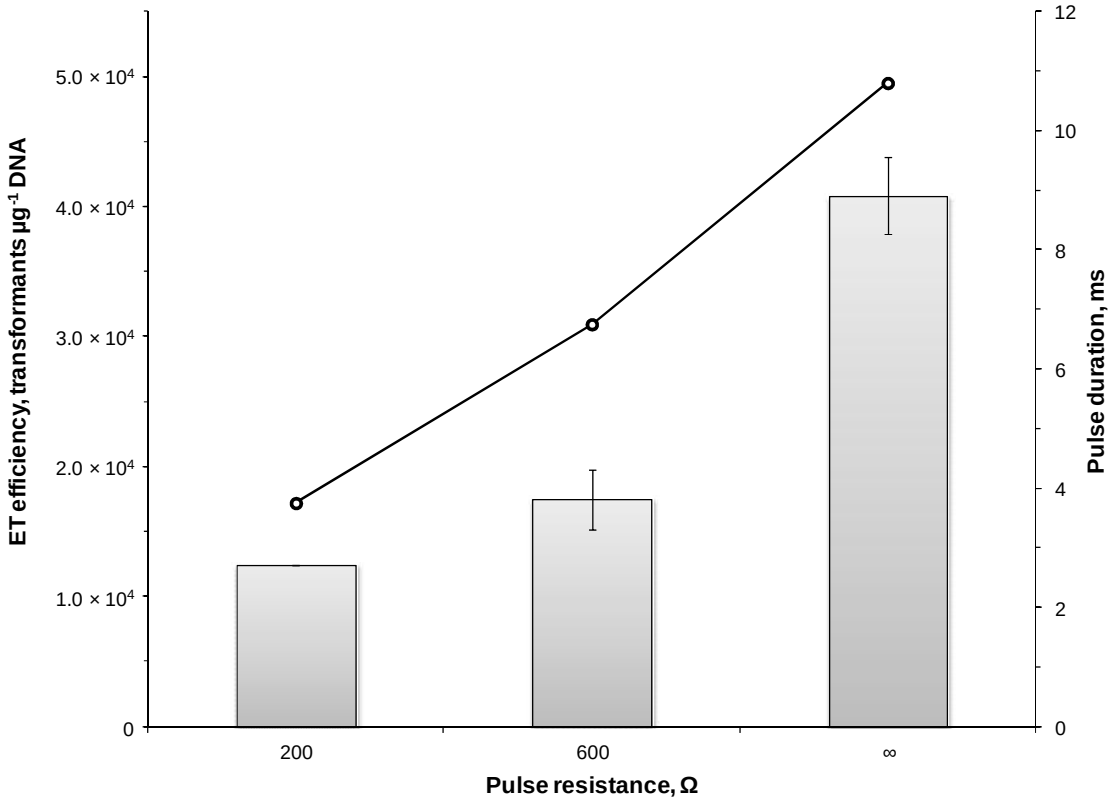
A)


9.0 × 10⁴
7.5 × 10⁴
6.0 × 10⁴
4.5 × 10⁴
3.0 × 10⁴
1.5 × 10⁴
0
ET efficiency, transformants µg⁻¹ DNA
12
10
8
6
4
2
0
Pulse duration, ms
1.6
1.8
2
Pulse voltage, kV
4
4.5
5
Field strength, kV cm⁻¹

B)


7.5 × 10⁴
6.0 × 10⁴
4.5 × 10⁴
3.0 × 10⁴
1.5 × 10⁴
0
ET efficiency, transformants µg⁻¹ DNA
20
15
10
5
Pulse duration, ms
1.8
2.25
Pulse voltage, kV

C)


ET efficiency, transformants µg-1 DNA
5.0 × 10⁴
4.0 × 10⁴
3.0 × 10⁴
2.0 × 10⁴
1.0 × 10⁴
0
12
10
8
6
4
2
0
Pulse duration, ms
200
600
∞
Pulse resistance, Ω

occurred using 0.5 µg of plasmid DNA. Transformants could be detected at the lowest quantity of DNA tested, 0.25 µg, and saturation with pMTL85141 was not observed up to 5.0 µg, the highest quantity of DNA tested.

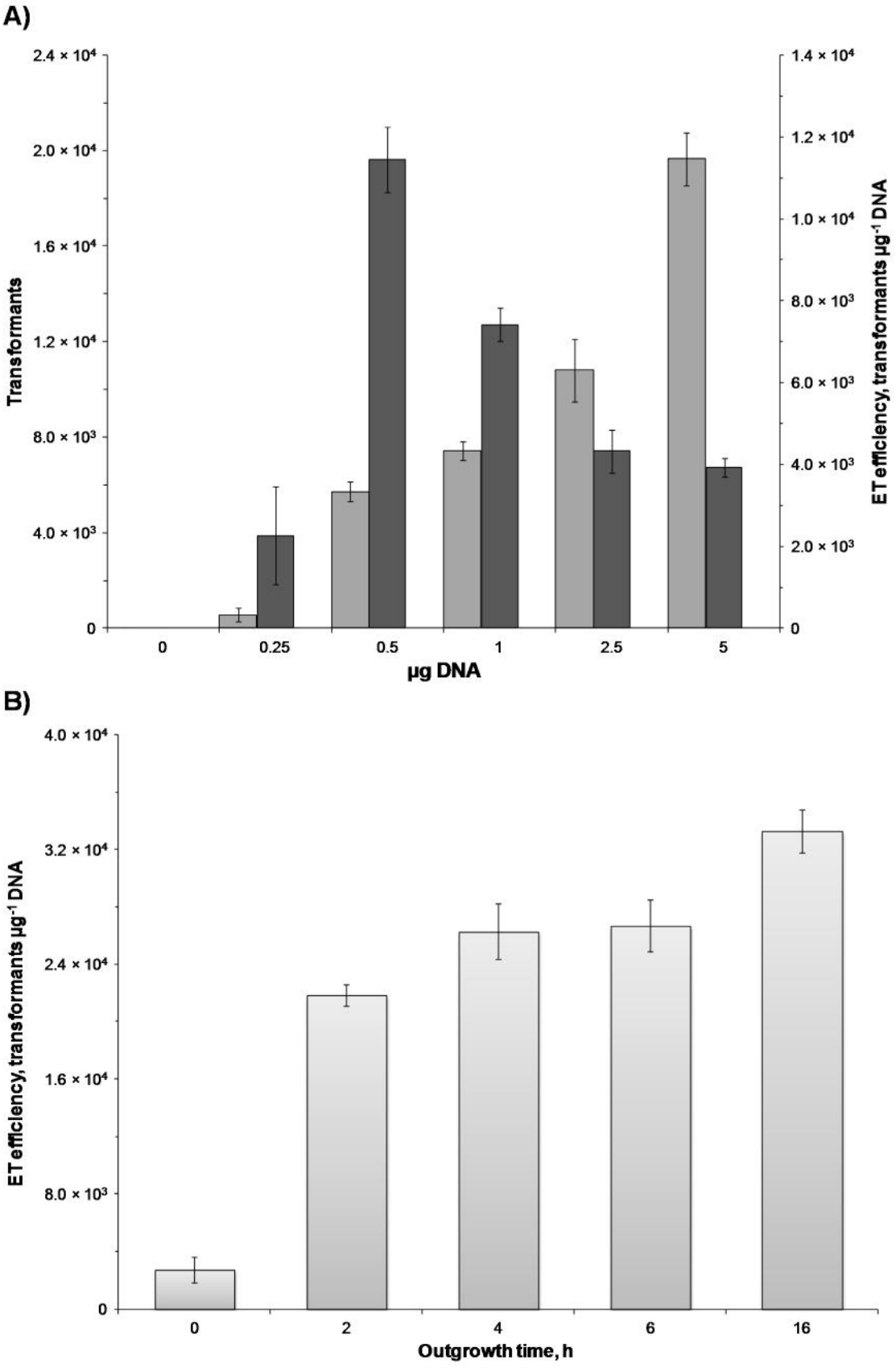
For assessing outgrowth duration, we incubated electroporated cells for 0, 2, 4, 6, or 16 hours prior to plating on selective medium. Growth in the form of gas formation and increased culture turbidity could be detected as early as 2 hours following transfer to recovery medium. Transformants could be obtained without recovery (i.e., 0 hours incubation), although at a significantly reduced efficiency (7.9- to 12.1-fold reduction compared to 2-16 hours incubation) (Figure 3.6B). As expected, the greatest electrotransformation efficiency was attained using the longest recovery time tested (i.e., 16 hours), which was approximately 1.3-fold greater than at 4-6 hours outgrowth, during which time the electrotransformation efficiency was unchanged. While 16 hours of outgrowth is a convenient duration due to the length of the pre-growth and washing and electroporation phases, electrotransformation efficiency for clostridia is typically reported following 4-6 hours of outgrowth. Thus, the electrotransformation efficiencies reported here, all of which involved 16 hour outgrowth experiments, can be divided by 1.3 for comparison to other clostridial electroporation efficiencies.

3.3.4. Application of the electrotransformation protocol to other vectors

Since many clostridial vectors favor the *ermB* determinant for erythromycin or clarithromycin selection, rather than *catP*-based thiamphenicol selection used in this study, we constructed pMTL85141ermB, a dual *catP* and *ermB* selectable plasmid. Comparable, high-level electrotransformation efficiencies (1.0-1.4 × 10^4 transformants μg^{-1} DNA) were obtained by selection of pMTL85141ermB using 15 µg/ml thiamphenicol, 4 µg/ml clarithromycin, or 20 µg/ml erythromycin. Control plasmid transformations lacking the *ermB* determinant failed to generate clarithromycin- or erythromycin-resistant colonies. Therefore, *ermB*-based clarithromycin or erythromycin selection is effective using *C. pasteurianum*.

To determine the generality of our high-efficiency electrotransformation protocol for other vectors, we also attempted electrotransfer of pSY6catP into *C. pasteurianum*. pSY6catP is a modified form of pSY6 [15] whereby the *ermB* erythromycin-resistance determinant is replaced with *catP* from pMTL85141. pSY6 is one of several *E. coli-Clostridium* shuttle vectors (in addition to, e.g., the ClosTron system of vectors [145]), which harbours the Ll.ltrB group II

Figure 3.6. Investigation of amount of DNA and outgrowth duration on the electrotransformation of *C. pasteurianum*. A) Effect of plasmid DNA amount on total number of transformants and electrotransformation efficiency of *C. pasteurianum*. Zero, 0.25, 0.5, 1.0, 2.5, or 5.0 µg of pMTL85141 were added to electrocompetent cells of *C. pasteurianum* and electroporated. Total number of thiamphenicol-resistant transformants (light shading) and electrotransformation efficiency (dark shading) were enumerated. Pulse duration was unaffected between samples. Zero µg pMTL85141 failed to generate transformants. **B)** Effect of post-electroporation incubation time. Cells were electroporated, transferred to outgrowth medium containing 0.2 M sucrose, and incubated for 0, 2, 4, 6, or 16 hours prior to selective plating.


A)
Transformants
ET efficiency, transformants µg-1 DNA
2.4 × 10⁴
2.0 × 10⁴
1.6 × 10⁴
1.2 × 10⁴
8.0 × 10³
4.0 × 10³
0
1.4 × 10⁴
1.2 × 10⁴
1.0 × 10⁴
8.0 × 10³
6.0 × 10³
4.0 × 10³
2.0 × 10³
0
0
0.25
0.5
1
2.5
5
µg DNA
B)
ET efficiency, transformants µg-1 DNA
4.0 × 10⁴
3.2 × 10⁴
2.4 × 10⁴
1.6 × 10⁴
8.0 × 10³
0
0
2
4
6
16
Outgrowth time, h

intron machinery necessary for performing intron-mediated gene knockouts in clostridia. A pSY6-based vector was chosen because it possesses the same pIM13 replicon as pMTL85141, thereby eliminating potential variation in efficiency due to differences in the origin of replication. Unexpectedly, pSY6catP transformed *C. pasteurianum* at a significantly decreased efficiency of 1.1×10^1 transformants μg^{-1} DNA, an efficiency approximately 1,000-fold lower than achieved with pMTL85141. To rule out a vector size effect on the reduction in electrotransformation efficiency (pSY6catP is 8,498 bp, whereas pMTL85141 is 2,963 bp), we also attempted to transform pHT3, a 7,377 bp vector with the same fundamental vector components as pMTL85141ermB, in addition to a heterologous *lacZ* gene from *Thermoanaerobacterium thermosulfurogenes* EM1 [12] (Table 3.1). Unlike pSY6catP, pHT3 transformed at a high efficiency of 1.8×10^4 transformants μg^{-1} DNA, which is comparable to pMTL85141ermB. Therefore, the dramatic reduction in electrotransformation efficiency is likely not due to differences in plasmid size. At this point, we hypothesize the presence of an additional unidentified restriction system which targets certain common site(s) of pSY6catP, but not pMTL85141, pMTL85141ermB, or pHT3, much like the *dcm*-methylation-dependent restriction systems recently addressed in *C. thermocellum* and *C. ljungdahlii* [107,235]. Our observation of the transformability of *in-vivo*-methylated plasmids, but not *in-vitro*-methylated plasmids, may also be the result of an unidentified methylation-dependent restriction system, which may or may not be the same one affecting pSY6catP. The *in vitro* M.SssI and M.CviPI methyltransferases we utilized, with 5'-^{m}CG-3' and 5'-G^{m}C-3' methylation sites, respectively, each methylate a larger number of sites than the M.FnuDII *in vivo* methyltransferase (5'-^{m}CGCG-3' methylation site). Since certain DNA substrates containing 5-methylcytosine are restricted, it is likely that over-methylated pMTL85141 could be unexpectedly targeted by a Type II or Type IV 5-methylcytosine-specific restriction endonuclease in *C. pasteurianum*. Experiments are underway to probe for an additional restriction system in *C. pasteurianum*. Nonetheless, even with the reduced electrotransformation efficiency of pSY6catP, we have used it to successfully introduce type II introns into the *C. pasteurianum* genome in preliminary experiments (data not shown), whereas this was not possible with the low-level electrotransformation efficiency we initially obtained prior to our electrotransformation studies detailed in this report.

3.4. Discussion

Lack of a gene transfer system has fundamentally hindered genetic manipulation of *C. pasteurianum*. We report here the first development of an electroporation-mediated procedure for high-level gene transfer to *C. pasteurianum*. The first step in overcoming this barrier involved confirming the previously identified *C. pasteurianum* restriction endonuclease, CpaAI, which cleaves 5'-CGCG-3' nucleotide sequences. *In vivo* methylation using the M.FnuDII methyltransferase allowed protection of *E. coli-C. pasteurianum* shuttle vectors. Active protoplast extracts of *C. pasteurianum* completely degraded pMTL85141 prepared from *E. coli* DH5α, but not when prepared from *E. coli* ER1821 harboring both pMTL85141 and pFnuDIIMKn (Figure 3.1A, B). This finding allowed us to demonstrate low-level electrotransformation of *C. pasteurianum* by using conditions commonly utilized for clostridial electroporation (Table 3.2). Despite repeated attempts, unmethylated pMTL85141 (i.e., prepared from *E. coli* DH5α in the absence of plasmid pFnuDIIMKn) was unable to transform *C. pasteurianum*, confirming the requirement for proper methylation of plasmid DNA prior to electrotransformation. Plasmid transfer was confirmed through colony-PCR-amplification, isolation and purification, and DNA sequencing of plasmid pMTL85141 from thiamphenicol-resistant colonies (Figure 3.2A, B; data not shown). The resulting plasmid preparation from *C. pasteurianum* was devoid of rearrangements and recombination and could only transform restriction-deficient strains of *E. coli*, such as ER1821. The initial efficiency of approximately 2.4×10^1 transformants μg^{-1} DNA paled in comparison to electrotransformation levels achieved with other species of *Clostridium* (up to 10^6 transformants μg^{-1} DNA [138]) and prevented transfer of low-level-transforming vectors, such as pSY6catP (discussed below). Thus, we investigated the effects of several parameters critical for the electrotransformation of Gram-positive bacteria. With the modified protocol, we were able to elevate the electrotransformation efficiency to a maximum of 7.5×10^4 transformants μg^{-1} DNA, an increase of more than three orders of magnitude from the original low-level electrotransformation efficiency (Table 3.3). The electroporation parameters determined to be most influential on the electrotransformation of *C. pasteurianum* are discussed in detail below.

Gram-positive bacteria are electrotransformed at efficiencies several orders of magnitude lower than Gram-negative cells, highlighting the importance of effective weakening of the Gram-positive cell wall prior to electroporation. The most widely used additives for this purpose

Table 3.3
Summary of protocol for high-level electrotransformation of *C. pasteurianum* and comparison to initial low-level protocol

Electrotransformation parameter	Low-level protocol	High-level protocol
Cell growth		
Growth additive	None	1.25% glycine (at OD_{600} 0.3-0.4)
Osmotic stabilizer	None	Hypertonic sucrose (0.4 M; at OD_{600} 0.3-0.4)
OD_{600} at time of harvest	OD_{600} 0.6-0.8	OD_{600} 0.6-0.8
Washing and pulse delivery		
Osmotic stabilizer	Isotonic sucrose (0.27 M)	Isotonic sucrose (0.27 M)
Cell membrane solubilizer	None	5% (v/v) ethanol 5 min prior to pulse
DNA amount	5 µg	0.5 µg
Pulse parameters	2.0-2.5 kV; 25 µF; 200-800 Ω; 6-9 ms	1.8 kV; 25 µF; ∞ Ω; 12-14 ms
Outgrowth		
Osmotic stabilizer	None	Hypotonic sucrose (0.2 M)
Recovery time	16 h	4-6 h
Transformation efficiency	2.4×10^1 transformants μg^{-1} pMTL85141	Up to 7.5×10^4 transformants μg^{-1} pMTL85141

are glycine, DL-threonine, lysozyme, and penicillin G, all of which were screened in this study (Figure 3.3A). Under the conditions tested, only glycine and DL-threonine afforded enhanced levels of electrotransformation, with glycine rendering the cell most susceptible to electrotransformation. Glycine has been widely used for increasing electrotransformation efficiency of Gram-positive bacteria [137,230,231,236,237,238,239,240,241] and its mode of action has been extensively studied [129]. Specifically, glycine is incorporated into the peptidoglycan network through replacement of critical D- and L-alanine residues, generating a weakened cell wall due to a lesser degree of cross-linking [129]. To balance glycine supplementation with cell viability, we determined that optimum electroporation is achieved using brief exposure to the highest glycine concentration that still permits sufficient cell growth. The optimized glycine regimen for *C. pasteurianum* involves exposure of early exponential phase cells to 1.25% glycine for 2-3 hours (Figure 3.3B, C), which is most similar to protocols developed for *Mycobacterium avium* [231], *Bacillus subtilis* [239], and *B. thuringiensis* [240]. Of particular importance is the timing of glycine addition. Specifically, glycine should be added only when the cells enter the exponential growth phase, as glycine supplementation at the time of inoculation resulted in significant growth inhibition. While cell-wall-weakening is not an indispensable requirement for electrotransformation of *C. pasteurianum*, a significant increase in electrotransformation efficiency was obtained with glycine treatment. To our knowledge this study represents the first use of glycine as a cell-wall-weakening and electroporation-enhancing agent within the *Clostridium* genus.

Upon treatment with glycine, cells possess a compromised cell wall, and therefore, it is paramount that cells be stabilized osmotically. Our results clearly demonstrate the importance of sucrose osmoprotection during glycine treatment, as cultures without sucrose supplementation failed to grow in the presence of glycine. Although 0.25 M sucrose in the growth medium was sufficient to promote growth, increasing the concentration to 0.4 M afforded a significant enhancement to electrotransformation (Figure 3.3B). In opposition to the growth medium, however, hypertonic sucrose in the recovery medium (0.4 M) was found to have a slightly negative effect on electrotransformation efficiency compared to a slightly hypotonic environment (0.2 M) (Figure 3.3E). A greater degree of osmotic stabilization is likely required in the growth medium compared to the recovery medium as a result of cell-wall-weakening that occurs in the presence of glycine. In a similar manner, hypertonic sucrose (0.5 M) in the wash and

electroporation buffer greatly reduced electrotransformation efficiency compared to an isotonic buffer (0.27 M) (Figure 3.3D). This outcome is in agreement with other clostridial electrotransformation protocols that employed phosphate-buffered sucrose as the wash and electroporation buffer, as all such protocols utilize isotonic rather than hypertonic sucrose [10,92,93,96,102,104,113]. Alternatively, a hypertonic wash and electroporation buffer has been shown to enhance electrotransformation of *Streptococcus cremoris* and *S. lactis* [242], *Listeria monocytogenes* [243], and *Lactobacillus sake* [244]. Based on these studies and our data presented here it is evident that the degree of osmotic protection required to achieve optimum electrotransformation must be carefully determined for each phase of the electroporation process. While it is commonly assumed that the role of osmotic stabilizer on cell-wall-weakened and electrotransformed cells is strictly protective, it cannot be ruled out that higher electrotransformation efficiencies arise, at least in part, from faster growth rates exhibited by cultures that actively utilize the stabilizer as a carbon and energy source, as observed with the use of sucrose for *C. pasteurianum* (this study) and *Lactococcus lactis* subsp. c*remoris* [230].

In addition to the use of cell-wall-weakening additives, the Gram-positive cell wall can be destabilized through the application of high-voltage pulses during electroporation, as stronger electric fields are typically required for Gram-positive compared to Gram-negative bacteria [95,223]. Hence, optimization of the electrical parameters must be performed to achieve efficient electroporation without compromising cell viability. We aimed to tailor our electric pulse, defined by the pulse voltage (kV) and corresponding field strength (kV cm^{-1}), capacitance (μF), and resistance (Ω), based on our developed glycine and sucrose regimen. In contrast to most clostridial electroporation procedures, which commonly utilize voltages of 2.0-2.5 kV (5.0-6.25 kV cm^{-1}), optimum electrotransformation of glycine-treated *C. pasteurianum* was found to occur under a lower voltage of 1.8 kV (4.5 kV cm^{-1}; Figure 3.5A). Within the clostridia, this low field strength has only been matched by the described protocol for *C. acetobutylicum* DSM 792 [102]. Increasing the capacitance at voltages of either 1.8 or 2.25 kV generated significantly reduced electrotransformation efficiencies (Figure 3.5B). In addition, using a higher voltage of 2.25 kV and attempting to truncate the duration of the pulse by using lower resistances of 200 and 600 Ω also produced a reduced number of transformants (Figure 3.5C). Therefore, optimum electrotransformation of glycine-treated *C. pasteurianum* was found to occur under a relatively low electric field of 4.5 kV cm^{-1} at 25 μF and ∞ Ω, generating a time constant of 12-14 ms. We

postulate that a relatively weak electric field is best for enhancing electroporation of *C. pasteurianum* as a result of the compromised cell wall associated with the application of our glycine regimen. A similar effect has been observed using glycine-treated cells of *Bacillus cereus*, in which glycine had no effect under a high field strength (20 kV cm^{-1}), yet a pronounced positive effect under a low field strength (12 kV cm^{-1}) [241].

To a lesser extent than the Gram-positive cell wall, the cell membrane also acts as a physical barrier to transfer of plasmid DNA into the cell. Aside from the presumed pore formation that occurs immediately following delivery of the electric pulse, little work has been done to enhance electroporation through increasing the extent of membrane permeabilization. Recently, two reports have detailed the use of ethanol as a membrane-solubilizing agent to enhance electroporation of *Escherichia coli* [135] and *Oenococcus oeni* [136]. The effect of ethanol, a fermentation end product of *C. pasteurianum*, on growth of species of *Clostridium* has been extensively studied and it has been shown that ethanol toxicity occurs through direct interaction with the cell membrane resulting in a decreased extent of lipid organization and increased membrane fluidity and cytoplasmic leakage [233,245]. It has been proposed that leakage occurs through an ethanol-induced increase in membrane pore size [246], which has clear implications to electroporation. In this report we assessed the electroporation-enhancing effect of ethanol and also extended our approach to butanol. Whereas butanol inhibited electrotransformation and cell growth of *C. pasteurianum* under the conditions tested, we found that ethanol supplemented at an appropriate concentration (5 or 10%) had a clear positive effect on electrotransformation (Figure 3.4). To our knowledge, this study represents the first use of ethanol to enhance electrotransformation within the *Clostridium* genus.

The investigation of various electrotransformation parameters enabled us to generate a maximum of 7.5×10^4 transformants μg^{-1} DNA, a more than 3,000-fold increase compared to our initial attempts using common clostridial electroporation conditions (Table 3.3).The results reported in this study demonstrate that *C. pasteurianum* ATCC 6013 is amendable to genetic manipulation. Our hope is that our developed gene transfer protocol will allow genetic and metabolic engineering of *C. pasteurianum* and promote further development of this biotechnologically important microorganism, as the maximum electrotransformation efficiency attained for *C. pasteurianum* is among the highest reported in the *Clostridium* genus. However, significant barriers remain to be resolved; namely the low electrotransformation efficiency for

plasmids which carry group II intron machinery, such as pSY6catP, necessary for constructing gene knockout mutants in clostridia. Since we have observed a drop in electrotransformation efficiency for pSY6catP (8,498 bp), yet not pHT3 (7,377 bp), a plasmid of comparable size, we suspect the decrease in efficiency is not related to plasmid size. Instead, we speculate that *C. pasteurianum* possesses at least one additional restriction-modification system, in addition to CpaAI, that is specifically active on pSY6catP, but not on pMTL85141. It is likely that additional uncharacterized restriction activities are also responsible for our inability to electrotransform substrates methylated *in vitro* using CpG and GpC methyltransferases, despite proper protection against the previously identified CpaAI restriction endonuclease. We are currently conducting genome sequencing of *C. pasteurianum* which will enable us to identify potential candidate restriction-modification genes that may be responsible for the low electrotransformation efficiency of certain non-pMTL85141 and *in-vitro*-methylated vectors. Also, it should be mentioned that, even with the reduced electrotransformation efficiency of pSY6catP, our electrotransformation protocol developed herein allows the introduction of group II introns into the *C. pasteurianum* genome in preliminary experiments (data not shown). Thus, should genome sequencing reveal additional restriction-modification genes, our electrotransformation method should enable investigators to knockout such genes using group II introns. Finally, the work here also demonstrates that *catP* (using thiamphenicol) and *ermB* (using clarithromycin or erythromycin) comprise effective selection marker systems for performing genetic engineering in *C. pasteurianum*. The pIM13 origin of replication from *Bacillus subtilis* and the pCB102 and pCD6 origins of replication from *C. butyricum* and *C. difficile*, respectively, were shown to support plasmid maintenance in *C. pasteurianum* and they round out the vector toolkit now available for genetic engineering in this important bacterium. Taken together, the high-level electrotransformation protocol and the vector and selection tools described herein set the stage for biotechnological exploitation of *C. pasteurianum* for the first time, thereby opening an important avenue for the production of biofuels from low-value and abundant crude glycerol.

3.5. Conclusions

In this work, M.FnuDII methylation, together with cell-wall-weakening, partial membrane solubilization, a low electric field, and osmoprotection enabled the

electrotransformation of *C. pasteurianum* ATCC 6013 at an efficiency of up to 7.5 x 10^4 transformants ug^{-1} DNA. The work here also demonstrates the development of a *C. pasteurianum* genetic toolkit currently comprised of two selectable markers (*catP*-based thiamphenicol selection and *ermB*-based clarithromycin or erythromycin selection) and three Gram-positive origins of replication (pIM13 from *Bacillus subtilis*, pCD6 from *C. difficile*, and pCB102 from *C. butyricum*). This is the first report of a genetic transformation procedure for *C. pasteurianum* and represents a key advancement for this industrially-important bacterium with important implications for low-cost biofuel production.

CHAPTER 4

Draft genome sequence of *Clostridium pasteurianum*

4.1. Introduction

The development of a transformation procedure allowing high-level gene transfer to *C. pasteurianum* [91] paves the way to an array of powerful genetic tools, including gene overexpression, knockout, and knockdown techniques. However, none of these techniques are possible without access to a genome sequence for *C. pasteurianum*. At the time of this report, and despite significant interest in the solventogenic clostridia, only one completed genome sequence is available for *C. pasteurianum* and is from an environmental isolate (strain BC1). It is unclear how similar strain BC1 is to the type strain of the species, strain ATCC 6013 (DSM 525).

In the current era of genomics and bioinformatics, where genome sequencing, assembly, and annotation can be outsourced for ever-decreasing costs, it is often unjustifiable to undertake genetic studies using an organism lacking genome sequencing data. As depicted in Figure 2.3, a critical step when attempting to manipulate a clostridial host strain entails sequencing the target organism's genome. Recent bacterial genome sequencing projects have involved at least one of three major NGS technologies; namely, Illumina dye sequencing, 454 pyrosequencing, and Pacific Biosciences single molecule real time sequencing (SMRT), all examples of DNA-polymerase-mediated sequencing-by-synthesis [247]. Illumina dye sequencing and 454 pyrosequencing were developed in the 1990s, whereas SMRT sequencing arrived in 2009. As of 2013, an updated version of the SMRT technology has become available for commercial use. Although all three technologies differ in the length of reads and depth of coverage obtained, generally a single run of any of the three NGS technologies is insufficient to capture a full-length bacterial genome, resulting in a large number of noncontiguous sequences [248]. In order to move from a fragmented draft assembly to a finished genome, often a hybrid approach comprising more than one NGS platform is utilized, which drives up costs and demands accurate reads for effective error correction [249]. A recent exception to this principle, which is expected

to exemplify the near future of NGS technology, involves a nonhybrid approach wherein a finished bacterial genome can be generated from a single DNA library using SMRT sequencing [250]. An important advantage of this technique lies in its ability to resolve long repeat regions of DNA, which largely comprise the noncontiguous gaps found in most bacterial draft genome assemblies.

In this study, we detail the steps taken to obtain a near-complete genome sequence for *C. pasteurianum* ATCC 6013. A hybrid sequencing approach combining all three major NGS platforms was employed by outsourcing sequencing, assembly, and annotation to various genome sequencing centres in North America. A collaboration was initiated with Oak Ridge National Laboratories to expedite genome assembly and annotation.

4.2. Methods

4.2.1. Growth and maintenance conditions

C. pasteurianum ATCC 6013 was grown and maintained as described previously (refer to section 3.2.2).

4.2.2. Total genomic DNA isolation and purification

Intact, high molecular weight *C. pasteurianum* genomic DNA was extracted from a 60 ml culture (OD 0.5-0.7) by first washing cells in 40 ml of a buffer containing 25 mM potassium phosphate, pH 7.0, and 6 mM $MgSO_4$, followed by resuspension in 15 ml of the same buffer supplemented with 50% sucrose and 200 µg/ml lysozyme [86,226]. After anaerobic incubation at 37 °C for 45 minutes, genomic DNA was extracted from 1.0-2.0 ml samples of protoplast suspension using a DNeasy Blood and Tissue Kit from Qiagen (Valencia, CA). Due to the high nuclease content of clostridia, hypertonic sucrose (50% w/v) was added to buffer ATL during cell lysis [86]. Eluted genomic DNA was treated with 100 µg/ml RNase A prior to additional purification using a Genomic DNA Clean &Concentrator kit from Zymo Research (Irvine, CA).

4.2.3. DNA sequencing, assembly, and annotation

454 pyrosequencing was performed using the Genome Sequencer FLX and Titanium reagents. Sequencing was performed by Genome Canada via the Génome Québec and McGill

University Innovation Centre (Montréal, QC). Two separate 1/8 runs were performed using a paired-end *C. pasteurianum* genomic library to allow *de novo* construction of raw sequence reads. An 8 kb library insert size was used.

Illumina sequencing was performed using the MiSeq sequencer at Oak Ridge National Laboratories (Oak Ridge, TN). A paired-end *C. pasteurianum* genomic library was prepared to allow *de novo* construction of raw sequence reads.

SMRT sequencing using the RS I analyzer and a 5 kb insert *C. pasteurianum* genomic library was performed at the DNA Sequencing Core at the University of Michigan (Ann Arbor, MI). Two parallel SMRT flow cells were utilized from the same genomic library. SMRT sequencing using the RS II analyzer and a long insert *C. pasteurianum* genomic library (11-20 kb) was performed at the Genomic Resource Center at the Institute for Genome Sciences (University of Maryland School of Medicine; Baltimore, MD).

Assembly and annotation of raw sequencing reads following each round of sequencing was performed Oak Ridge National Laboratories (Oak Ridge, TN). Newbler (454/Roche; Basel, Switzerland) and CLC Genomics Workbench & CLC Assembly Cell (CLC bio; Aarhus, Denmark) software packages were utilized for assembly.

4.3. Results and discussion

To obtain a draft genome assembly for *C. pasteurianum*, a hybrid NGS approach was employed utilizing all three major NGS platforms. Two 1/8 runs of 454 pyrosequencing resulted in 245,433 reads assembled into 1,231 contigs and 43 scaffolds. Following Illumina MiSeq sequencing, a 454 and Illumina hybrid draft genome reduced the number of contigs to 214. Contig size ranged from 512 to 162,006 bp, with a mean of 19,697 bp. SMRT sequencing from an RS I sequencer led to an improved draft assembly possessing 4.4 Mbp in 30 contigs. Finally, to further drive down the number of gaps in our draft assembly, we supplemented our RS I data with two additional SMRT flow cells using the updated RS II sequencer. The final draft assembly following our hybrid NGS approach yielded 12 contigs containing 4.4 Mbp. Remaining contigs range in size from 8,251 to 1,860,121 bp, with five contigs under 12 kb. We are currently employing traditional PCR, primer walking, and Sanger sequencing using the *C. pasteurianum* BC1 genome as a reference for gap closure of the remaining 12 contigs in our

latest draft genome assembly. By mapping the 12 contigs from our draft sequence to the finished BC1 genome, it is expected that all gaps within our sequence are less than 2 kb.

Our *C. pasteurianum* genome sequence contains approximately 4,420,100 and 4,056 candidate protein-coding genes, and has a GC content of 30%. No extrachromosomal elements were identified or suspected, although the presence of such elements cannot be ruled out. Contigs 8 and 1 possess uncharacteristic GC contents of 37.5 and 38.2%, respectively. Approximately 74 tRNA genes were identified in the genome. A total of 16 rRNA clusters comprised of 16S and 23S rRNA genes were detected, many of which flank the remaining contigs, demonstrating the limitations of NGS technologies for sequencing highly repetitive rRNA-encoding DNA. In contrast, the genome sequence of *C. pasteurianum* strain BC1 is 4.99 Mbp with a 53,393 bp plasmid, with GC contents of 30.6 and 28.9%, respectively. Strain BC1 possesses 9 rRNA clusters composed of 16S and 23S genes, in addition to 10 5S subunit genes.

All putative acido- and solventogenic pathway genes were annotated within the *C. pasteurianum* genome, in addition to a full complement of glycolytic pathway genes. Locus tags corresponding to the genes discussed below are given in parentheses for the *C. pasteurianum* DSM 525 (ATCC 6013) draft genome sequence (accession NZ_ANZB00000000) that is available through the National Center for Biotechnology Information (NCBI; http//www.ncbi.nlm.nih.gov). Two apparent mechanisms are utilized by the organism for transitioning from glycolysis to fermentative metabolism (conversion of pyruvate to acetyl-CoA), namely pyruvate formate lyase (*pfl*) and pyruvate:ferredoxin oxidoreductase (*pfo*). *C. pasteurianum* possesses three copies of each of these genes (*pfl*: F502_15710, F502_15690; F502_19556; *pfo*: F502_01955, F502_07648, F502_07643), with two *pfo* genes arranged adjacently within contig 9 (F502_07648 and F502_07643). As is the case with *C. acetobutylicum*, many central fermentative pathway genes were found within operon structures in the genome sequence of *C. pasteurianum*. These include genes for acetate (*pta-ack*; F502_04382 and F502_04387) and butyrate (*ptb-buk*; F502_11866 and F502_11871) production, in addition to the multistep conversion of acetyl-CoA to butyryl-CoA (*crt-bcd-etfB-etfA-hbd*; F502_06297 to F502_06277). A second copy of *hbd* (F502_07923) is located elsewhere in the chromosome. The genome sequence contains 17 genes that are annotated as iron-type alcohol dehydrogenases, 10 of which could play roles in production of ethanol and butanol from acetaldehyde/acetyl-CoA (*adh*) and butyraldehyde/butyryl-CoA (*bdh*), respectively.

Interestingly, *C. pasteurianum* appears to possess all of the genes involved in acetone production, yet formation of acetone by *C. pasteurianum* has never been reported. Three CoA transferase gene clusters have been annotated for potential conversion of acetoacetyl-CoA to acetoacetate with concomitant production of acetyl-CoA and butyryl-CoA directly from acetate and butyrate, respectively. A complete acetone-formation locus has been identified in the genome of *C. pasteurianum*, which possesses the same orientation as that from *C. acetobutylicum*. This cluster is comprised of both subunits of CoA transferase (*ctfAB*; F502_16510 and F502_16505), preceded by a bifunctional acetaldehyde/alcohol dehydrogenase (*aad*/*adhE1*; F502_16515), and followed by a reverse orientation acetoacetate decarboxylase gene (*adc*; F502_16500) for conversion of acetoacetate to acetone. The amino acid sequence of the *C. pasteurianum adc* gene possesses significant identity (84%) to the *C. acetobutylicum* homolog. Despite these similarities, *C. pasteurianum* fails to produce acetone, indicating a lack of activity or absence of a suitable transcriptional promoter for any of the *ctf* or *adc* genes involved in the acetone formation pathway. An additional copy of subunit A of CoA transferase is arranged in an operon with additional copies of both crotonase and butyryl-CoA dehydrogenase (*crt-ctfA-bcd*; F502_09038 to F502_09028), while the other CoA transferase cluster (*ctfAB*; F502_16080 and F502_16075) does not appear to be a part of an operon with other genes. *C. pasteurianum* also differs from *C. acetobutylicum* in its ability to form 1,3-propanediol when grown on reduced substrates, such as mannitol and glycerol. The genes for glycerol dehydratase and 1,3-propanediol dehydrogenase have been cloned and sequenced in *E. coli* and are found within a large 1,3-propanediol production locus of contig 5. This 7.1 kb, 8-gene cluster contains the large, medium, and small subunits of glycerol dehydratase (*dhaBCE*; F502_03402 to F502_03412), two subunits of a glycerol dehydratase reactivation factor (F502_03417 and F502_03422), two cobalamin adenosyltransferase genes conferring vitamin B_{12} dependence (F502_03427 and F502_03432), and 1,3-propanediol dehydrogenase (*dhaT*; F502_03437).

In the time that we undertook our genome sequencing studies, a second draft genome for the type strain of *C. pasteurianum* (ATCC 6013/DSM 525) has been made publicly available [251]. The second genome assembly currently possesses 37 contigs encompassing 4,285,687 bp (Table 4.1). It is expected that both genomes possess many of the same gap regions between contigs, which often represent highly-repetitive DNA encoding ribosomal RNA. Indeed, our

draft assembly contains 16 copies of the 16S rRNA gene, while the second assembly contains only two copies of the same gene. Finally, an additional draft genome for a third strain of *C. pasteurianum*, NRRL B-598, was published online in December, 2013. The NRRL B-598 assembly currently contains 138 contigs, yet is very large (6.0 Mbp) (Table 4.1), demonstrating significant variation in clostridial genomes, even between strains of the same species.

4.4. Conclusions

In this study, we have obtained a near-complete genome sequence of the type strain of *C. pasteurianum*, which currently contains 12 contigs. By mapping our 12 contigs to the finished genome sequence for *C. pasteurianum* BC1, we suspect that we have sequenced 4.4 Mbp of a predicted 4.5 Mbp for the genome of strain ATCC 6013. We are currently finishing the genome sequence by closing the gap regions between contigs using standard PCR and Sanger sequencing. We have also succeeded in capturing the nucleotide sequences of all relevant central fermentative genes of *C. pasteurianum*, which are required for genetic manipulation and metabolic engineering studies involving gene overexpression, knockout, and knockdown.

Table 4.1
Summary of completed and in progress *C. pasteurianum* genome sequencing efforts

Strain	Progress	Size	GC content	Candidate protein coding genes	Institute
ATCC 6013/DSM 525 (type strain)	37 contigs	4,285,687 bp	29.8 %	3,887	Hamburg University of Technology (Hamburg, Germany)
ATCC 6013/DSM 525 (type strain)	12 contigs	4,420,100 bp	30.0%	4,056	University of Waterloo (Waterloo, ON)
BC1	completed	4,990,707 bp (+53,393 bp plasmid)	30.6 %	4,512	Brookhaven National Laboratory (Upton, NY)
NRRL B-598	138	6,041,878 bp	29.6%	5,367	Institute of Chemical Technology (Prague, Czech Republic)

CHAPTER 5

Overcoming host restriction and low retrohoming efficiency permits group-II-intron-mediated chromosomal gene knockout in *Clostridium pasteurianum*

5.1. Introduction

An overall lack of genetic tools has impeded metabolic engineering of *C. pasteurianum*, allowing only tedious random chemical mutagenesis techniques [252,253]. It is expected that efficient plasmid transfer will allow rational metabolic engineering approaches, including the application of gene knockout, knockdown, and overexpression techniques, none of which have been employed yet in *C. pasteurianum*. Of these indispensable techniques, gene knockouts are the most robust and widely employed method for decreasing or entirely abolishing the expression of a native chromosomal gene or metabolic pathway. In *Clostridium*, the preferred tool for conducting gene knockouts is the ClosTron [14], which was adapted from TargeTron technology and exploits the unique retrohoming mechanism of bacterial group II introns [189]. Owing to the broad host range of group II intron machinery, ClosTron-mediated gene knockouts have been performed in at least 11 species of *Clostridium* and have proven indispensable in genetic studies involving both the medical and industrial clostridia [145].

Following our initial report of gene transfer to *C. pasteurianum*, we have observed that electrotransformation efficiency varies drastically between certain shuttle vectors. Poor electrotransformation efficiency has been determined to be specific to the group II intron machinery within such vectors [91]. Restriction systems are the most common cause of observed transformation recalcitrance in bacteria and often potently inhibit plasmid transfer [95]. *C. pasteurianum* ATCC 6013 produces at least two active RM systems, CpaI (5’-GATC-3’) [254] and CpaAI (5’-CGCG-3’) [119], the recognition sites of which are blocked using Dam- (CpaI) and M.FnuDII– (CpaAI) methylated plasmid DNA, respectively, prior to electrotransformation [91]. It is possible that *C. pasteurianum* expresses a third RM system that recognizes a specific nucleotide sequence that is found within the group II intron or its cognate intron-encoded protein vector region. On the other hand, transformation resistance could potentially be the result of

toxicity of a plasmid-encoded gene product or plasmid instability due to selection of poor replication origins. In this study we examined the nature of this electrotransformation variation and aimed to resolve the issue to allow more efficient plasmid transfer. We demonstrate that host restriction of intron-harbouring vectors and low intron retrohoming efficiency hinder chromosomal gene knockout techniques in *C. pasteurianum*. The resolution of these barriers permits successful utilization of group II intron technology and will lead to improved metabolic engineering outcomes with *C. pasteurianum*.

5.2. Methods

5.2.1. Bacterial strains, plasmids, and oligonucleotides

Bacterial strains and plasmids employed in this work are listed in Table 5.1 and oligonucleotides sequences are given in Table 5.2. *E. coli* DH5α was utilized for vector construction and cloning purposes and ER1821 [pFnuDIIMKn] for methylation of *E. coli-C. pasteurianum* shuttle vectors destined for *C. pasteurianum* [91]. All oligonucleotides and gBlocks were synthesized by Integrated DNA Technologies (IDT; Coralville, IA). Oligonucleotides were prepared at the 25 nm scale using standard desalting. Custom gene synthesis was performed by Bio Basic Inc. (Markham, ON).

5.2.2. Growth and maintenance conditions

All recombinant and non-recombinant *E. coli* and *C. pasteurianum* strains were cultivated and maintained according to section 3.2.2. Recombinant *C. pasteurianum* strains were selected, when necessary, with 10 µg ml^{-1} thiamphenicol or 20 µg ml^{-1} erythromycin.

5.2.3. DNA isolation and manipulation and electrotransformation

Plasmid DNA was extracted from *E. coli* and purified using an EZ-10 Spin Column Plasmid DNA Miniprep Kit (Bio Basic, Inc., Markham, ON). Intact, high molecular weight *C. pasteurianum* genomic DNA was extracted as outlined in section 4.2.2. DNA restriction fragments and PCR products were purified directly or from agarose gels using an EZ-10 Spin Column DNA Gel Kit (Bio Basic, Inc.; Markham, ON). Vector construction was carried out according to standard procedures [227]. Restriction enzymes, Standard *Taq* DNA Polymerase,

Phusion High-Fidelity DNA Polymerase, and Quick Ligation Kit were purchased from New England Biolabs (Whitby, ON). All commercial enzymes and kits were used according to the manufacturer's instructions. Electrotransformation of *C. pasteurianum* was performed according to section 3.2.6.

5.2.4. Vector construction

Plasmid pMTL007C-E6 was constructed from pMTL007C-E2 by ligation of a 0.9 kb AscI + FseI digestion fragment of pMTL85141 containing the pIM13 replication module with a 7.4 kb product of pMTL007C-E2 resulting from digestion with the same restriction enzymes.

pSY6catP-Ll.ltrB was constructed by digesting pSY6catP with NheI + BstAPI, extracting the resulting 6.1 kb fragment, and ligating it with complementary oligos pSY6catP-ltrB.NheI.S + pSY6catP-ltrB.BstAPI.AS that had been annealed to generate compatible NheI and BstAPI restriction ends. Complementary oligonucleotides were mixed in equimolar amounts, heated to 95 °C in a thermocycler, and allowed to anneal overnight by disconnecting the power source from the thermocycler. For construction of pSY6catP-ltrA, the entire 3.4 kb Ll.ltrB-ltrA intron region was removed using XhoI + BstAPI digestion and replaced with a 2.4 kb PCR product containing only the *ltrA* coding sequence, generated using primers pltrA.XhoI.S + pltrA.BstAPI.AS. The resulting product was digested with XhoI + BstAPI and ligated with the 5.1 kb vector backbone to place *ltrA* under transcriptional control of the *ptb* promoter. Plasmid pSY6catP-delPptb was derived from pSY6catP by replacement of the 130 bp SmaI + XhoI digestion product containing the *C. acetobutylicum ptb* promoter with a 56 bp stuffer fragment lacking -35 and -10 promoter signals derived by annealing oligonucleotides pSY6.del-Pptb.S + pSY6.del-Pptb.AS. Ligation-proficient SmaI and XhoI restriction ends were generated upon successful annealing of complementary oligonucleotides. pSY6catP-del2dcm was derived from pSY6catP by replacing a 0.93 kb SacII + BstAPI restriction fragment with the same 0.93 kb sequence in which the two *dcm* sites were deleted by two single-base-pair mutations. The *dcm* deletion fragment was synthesized by Bio Basic, Inc. (Markham, ON), digested with SacII + BstAPI, and ligated into the corresponding sites of pSY6catP. The three restriction fragment

Table 5.1
Strain and plasmids employed in this study.

Strains or plasmids	Relevant characteristics	Source or reference
Strains		
Escherichia coli DH5α	F$^-$ *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG φ80dlacZ*a*M15* a*(lacZYA-argF)U169, hsdR17(*$r_K^-m_K^+$*),* a-	Lab stock
Escherichia coli ER1821	F$^-$ *endA1 glnV44 thi-1 relA1? e14$^-$(mcrA$^-$) rfbD1? spoT1?* a*(mcrC-mrr)114::IS10*	Lab stock; New England Biolabs
Clostridium pasteurianum ATCC 6013	Wild-type	American Type Culture Collection
Clostridium pasteurianum ΔREN	Disruption mutant generated by inserting the Ll.ltrB intron into position 176a of the CpaAI gene	This study
Plasmids		
pFnuDIIMKn	M.FnuDII methyltransferase plasmid for methylation of *E. coli-C. pasteurianum* shuttle vector s (Kn^R ; p15A ori)	[91]
pIMP1	*E. coli-Clostridium* shuttle vector (Ap^R; ColE1 ori; Em^R; pIM13 ori)	[10]
pMTL85141	*E. coli-Clostridium* shuttle vector (Cm^R/Tm^R; ColE1 ori; pIM13 ori)	[139]
pSY6catP	*E. coli-Clostridium* shuttle vector expressing the Ll.ltrB intron and *ltrA* IEP from the *C. acetobutylicum ptb* promoter (Ap^R ; Cm^R/Tm^R; ColE1 ori; pIM13 ori)	[91]
pMTL007C-E2	ClosTron vector expressing the Ll.ltrB intron with RAM and *ltrA* IEP from the *C. sporogenes fdx* promoter (Cm^R/Tm^R; ColE1 ori; pCB102 ori; Em^R RAM)	[145]
pMTL007C-E6	pIM13 derivative of pMTL007C-E2	This study
pSY6catP-Ll.ltrB	*ltrA*-deletion derivative of pSY6catP	This study

pSY6catP-ltrA	Ll.ltrB-deletion derivative of pSY6catP	This study
pSY6catP-delPptb	Derived by deleting the -35 and -10 signals of the *C. acetobutylicum ptb* promoter from plasmid pSY6catP	This study
pSY6catP-del2dcm	Derived by mutating the two *E. coli* Dcm restriction recognition sites downstream of the *ltrA* coding sequence within plasmid pSY6catP	This study
pSY6catP-MB	Derived by replacing a 1,661 bp MfeI + BstAPI restriction fragment of pSY6catP with a 48 bp stuffer fragment	This study
pSY6catP-NM	Derived by replacing a 688 bp NheI + MfeI restriction fragment of pSY6catP with a 48 bp stuffer fragment	This study
pSY6catP-NS	Derived by replacing a 1,434 bp NheI + SacII restriction fragment of pSY6catP with a 48 bp stuffer fragment	This study
pSY6catP-frag1	Derived by replacing a 1,332 bp BglII + EcoO109I restriction fragment of pSY6catP with a 589 bp *ltrA* region	This study
pSY6catP-frag2	Derived by replacing a 1,332 bp BglII + EcoO109I restriction fragment of pSY6catP with a 363 bp *ltrA* region	This study
pSY6catP-frag3	Derived by replacing a 1,332 bp BglII + EcoO109I restriction fragment of pSY6catP with a 574 bp *ltrA* region	This study
pMTLSY334	Derived by subcloning a 334 bp SacII + AatII fragment of the *ltrA* coding sequence into plasmid pMTL85141	This study
pSY6catP-mut98	pSY6catP derivative possessing 98 silent mutations in the *ltrA* coding sequence	This study
pMTLCP-E2	Derived by subcloning a 1,427 bp MscI + AclI restriction fragment of pSY6catP-mut98 into plasmid pMTL007C-E2	This study
pMTLCP-E6	Derived by subcloning a 1,427 bp MscI + AclI restriction fragment of pSY6catP-mut98 into plasmid pMTL007C-E6	This study
pSYCP-REN	Targeting construct of plasmid pSYCP for disruption of the CpaAI REN gene at position 176a	This study
pMTLCP-E2-REN	Targeting construct of plasmid pMTLCP-E2 for disruption of the CpaAI REN gene at position 176a	This study
pMTLCP-E6-REN	Targeting construct of plasmid pMTLCP-E6 for disruption of the CpaAI REN gene at position 176a	This study

Table 5.2
Oligonucleotides employed in this study

Oligonucleotide	Sequence (5'-3')*
pSY6catP-ltrB.NheI.S	CTAGCGCTATATGCGTTGATGCAATTTCTATGCACTCGTAGTAGTCTGAGAAGGCATATG
pSY6catP-ltrB.BstAPI.S	ATGCCTTCTCAGACTACTACGAGTGCATAGAAATTGCATCAACGCATATAGCG
pltrA.XhoI.S	TTTCTACTCGAGGCGTTGATGCAATTTCTATGCACTC
pltrA.BstAPI.AS	GGCATCAGAGCAGATTGTACTGAG
pSY6.del-Pptb.S	GGGGTTAATCATTTAACATAGATAATTAAATAGTAAAAGGGAGTGTCGAGATATCC
pSY6.del-Pptb.AS	TCGAGGATATCTCGACACTCCCTTTTACTATTTAATTATCTATGTTAAATGATTAACCCC
pSY6.MfeI/BstAPI.S	AATTGATTTAGTAATTTCTATAAGCAGGTTAGCTGTAAAACTAGCAGTAGCACGCATATG
pSY6.MfeI/BstAPI.AS	ATGCGTGCTACTGCTAGTTTTACAGCTAACCTGCTTATAGAAATTACTAAATC
pSY6.NheI/MfeI.S	CTAGCATTTAGTAATTTCTATAAGCAGGTTAGCTGTAAAACTAGCAGTAGCACC
pSY6.NheI/MfeI.AS	AATTGGTGCTACTGCTAGTTTTACAGCTAACCTGCTTATAGAAATTACTAAATG
pSY6.NheI/SacII.S	CTAGCATTTAGTAATTTCTATAAGCAGGTTAGCTGTAAAACTAGCAGTAGCACCCGC
pSY6.NheI/SacII.AS	GGGTGCTACTGCTAGTTTTACAGCTAACCTGCTTATAGAAATTACTAAATG
pSY6-frag1.BglII.S	GGGATATGATATACGAGTAAGGAGATCTGG
pSY6-frag1.EcoO109I.AS	AGTATTAGGCCCTGACGTCCCACATAATTCACAACATTTAGC
pSY6-frag2.BglII.S	AACAGGAGATCTGCTAAATGTTGTGAATTATGTGGGACGTC
pSY6-frag2.EcoO109I.AS	TACTCTAGGCCCTGGAGACCCCACACTACCATCG
pSY6-frag3.BglII.S	TCGCCAAGATCTCGATGGTAGTGTGGGGTCTCC
pSY6-frag3.EcoO109I.AS	GTGCCACCTGACGTCTAAGAAACC
pSY6.3'SOE.S	TGGGAAATGGCAATGATAGCGAAAC
pSY6-SOE.EcoO109I.AS	ATAGGCGTATCACGAGGCCCTTTC
pSY6-gBlock.BglII.S	CGAGTAAGGAGATCTGGAACGATAAAACG
pMTL.CpaAII-anneal.S	CTAGAGTCGACGTCACGCGTCCAAGGAGATCTCCAGGCCTGCAGACATGCA
pMTL.CpaAII-anneal.AS	AGCTTGCATGTCTGCAGGCCTGGAGATCTCCTTGGACGCGTGACGTCGACT
pSYCP.gBlock.S	GGAGGTCAATCTATGAAAATGCGATTAAGC
pSYCP.gBlock.AS	CTTTCGTTTCGTTCCCATAGGTTCTCC
pMTLCP.REN-HindIII.S	GTATTTAAAGCTTATAATTATCCTTAAATTTCTTAAAAGTGCGCC
Cpa-REN.seq.Fw	CTACTTGAGGTCTAGGACTTCTATCT
Cpa-REN.seq.Rv	ACAGATAGGCCATTAAAGGTATTCA
Ll.ltrB.seq-int.asFw	CCAACGCGTCGCCACGTAATAAAT
Ll.ltrB.seq-int.asRv	ATGGGAACGAAACGAAAGCGATGC

* Underline: restriction recognition sequences and overhangs.

deletion constructs, pSY6catP-MB, pSY6catP-NM, and pSY6catP-NS were prepared by digesting pSY6catP with MfeI + BstAPI, NheI + MfeI, and NheI + SacII, respectively, and annealing the resulting vector backbones with the respective annealed oligonucleotide pairs, pSY6.MfeI/BstAPI.S + pSY6.MfeI/BstAPI.AS (pSY6catP-MB), pSY6.NheI/MfeI.S + pSY6.NheI/MfeI.AS (pSY6catP-NM), and pSY6.NheI/SacII.S + pSY6.NheI/SacII.AS (pSY6catP-NS).

To construct pSY6catP-frag1, pSY6catP-frag2, and pSY6catP-frag3, a 1.3 kb BglII+EcoO109I restriction fragment was removed from pSY6catP and replaced with a 589 (primers pSY6-frag1.BglII.S + pSY6-frag1.EcoO109I.AS), 363 (primers pSY6-frag2.BglII.S + pSY6-frag2.EcoO109I.AS), and 574 (primers pSY6-frag3.BglII.S + pSY6-frag3.EcoO109I.AS) bp PCR product, respectively, corresponding to various products of the *ltrA* coding region of pSY6catP. To construct pMTLSY334, a 334 bp SacII + AatII fragment of the *ltrA* coding sequence was subcloned into the corresponding sites of pMTL85141. To mutate the unknown restriction recognition sequence within a 334 bp region of the *ltrA* coding sequence, a 655 bp gBlock was synthesized possessing 98 silent mutations in which 83 codons were altered. A 731 bp PCR product containing the 3' *ltrA* coding sequence and downstream region and possessing 25 bp overlap with the mutated gBlock was amplified using primers pSY6.3'SOE.S + pSY6-SOE.EcoO109I.AS. The PCR product was loaded on a 1.0% agarose gel, stabbed with a micropipette tip, and used as template along with 5 ng of the purified gBlock in a SOE PCR by cycling for 10 cycles prior to adding primers pSY6-gBlock.BglII.S + pSY6-SOE.EcoO109I.AS and cycling for 25 additional cycles. The resulting product was digested with BglII + EcoO109I and ligated with pSY6catP that had been digested with the same restriction endonucleases to generate pSY6catP-mut98. Plasmids pMTLCP-E2 and pMTLCP-E6 were constructed in a similar manner by subcloning a 1,427 bp MscI + AclI restriction fragment of pSY6catP-mut98 into the corresponding sites of pMTL007C-E2 and pMTL007C-E6, respectively.

Ll.LtrB intron design was performed using the computer algorithm developed by TargeTronics, LLC (Austin, TX; http://www.targetrons.com). The insertion site with the highest predicted insertion score splicing into the antisense strand was selected corresponding to nucleotide position 176 (score of 7.3) of the CpaAI restriction endonuclease (REN). Plasmid pSY6catP-mut98 was used as the basis for a *C. pasteurianum* TargeTron gene knockout vector. For retargeting pSY6catP-mut98, a 572 bp gBlock fragment was synthesized possessing

mutations in the IBS, EBS2, and EBS1d intron regions corresponding to position 176 of the REN insertion site. The retargeted gBlock was designed with a constitutive *C. pasteurianum* thiolase promoter controlling transcription of the Ll.ltrB intron and *ltrA* gene. The gBlock fragment was PCR-amplified using primers pSYCP.gBlock.S + pSYCP.gBlock.AS, digested with BamHI + BsrGI, and ligated into the corresponding sites of pSY6catP-mut98 to generate pSYCP-REN. To retarget the ClosTron vectors pMTLCP-E2 and pMTLCP-E6 to the CpaAI gene of *C. pasteurianum*, primers pMTLCP.REN-HindIII.S and pSYCP.gBlock.AS were used to amplify the gBlock targeted to the CpaAI endonuclease. The resulting 384 bp PCR product was digested with HindIII + BsrGI and ligated into the corresponding sites of pMTLCP-E2 and pMTLCP-E6 to give pMTLCP-E2-REN and pMTLCP-E6-REN, respectively.

5.2.5. Single molecule real-time genome sequencing and methylome analysis

Single molecule real-time genome sequencing data from an RS-II sequencer (Pacific Biosciences; Menlo Park, CA) from section 4.2.3 was utilized for methylome analysis. Our 12-contig draft genome assembly was used as a reference genome for methylome analysis, which was carried out by Pacific Biosciences (Menlo Park, CA).

5.2.6. Group-II-intron-mediated gene knockout, enrichment, and screening

To isolate a chromosomal gene knockout, plasmid pSYCP targeted to the gene encoding CpaAI was electrotransformed to *C. pasteurianum* and transformants were selected using 10 µg ml^{-1} thiamphenicol. Colonies containing the intron insertion were identified using a gene-specific and intron-specific primer pair. Both gene-intron junctions were screened using primers Ll.ltrB.seq-int.asFw + Cpa-REN.seq.Fw for one junction and primers Ll.ltrB.seq-int.asRv + Cpa-REN.seq.Rv for the adjacent junction. Positive colonies were then pooled for enrichment of the gene knockout genotype by repeated subculturing in selective growth medium. Briefly, sporulated colonies were heat-shocked at 80 ºC for 10 minutes in 10 ml of 2×YTG medium, cooled on ice, and supplemented with 10 µg ml^{-1} thiamphenicol. Following approximately 24 h of growth, 0.5 ml was used to inoculate a fresh tube of 10 ml 2×YTG containing 10 µg ml^{-1} thiamphenicol. This process was repeated approximately every 10-12 hours for a total of 10 transfers, at which point serial dilutions were plated onto non-selective 2×YTG agar, as previously reported [78]. To identify a homogenous gene knockout colony, colony PCR was

performed using two gene-specific primers flanking the intron insertion site (Cpa-REN.seq.Fw + Cpa-REN.seq.Rv).

5.3. Results

5.3.1. Shuttle vectors harboring Ll.ltrB intron machinery hinder electrotransformation of C. pasteurianum

To attempt gene knockouts in *C. pasteurianum*, we first transferred plasmid pSY6catP, which was developed for use in *C. acetobutylicum* [15]. This vector harbors the Ll.ltrB intron and its cognate *ltrA* IEP, both transcribed from the same *ptb* promoter from *C. acetobutylicum* within a pIMP1 vector backbone. Electrotransformation efficiencies of 3.7×10^4 and 3.7×10^0 transformants μg^{-1} were obtained for pIMP1 and pSY6catP, respectively, indicating an inability of pSY6catP to transform *C. pasteurianum* (Figure 5.1). As the only difference between pIMP1 and pSY6catP is the presence of the intron machinery, we also attempted to transfer the ClosTron plasmid pMTL007C-E2, which expresses the same Ll.ltrB intron elements, but in a different vector backbone. Like pSY6catP, pMTL007C-E2 also yielded a poor electrotransformation efficiency (1.9×10^1 transformants μg^{-1}) (Figure 5.1). Since pMTL007C-E2 possesses a different replication origin from pSY6catP (pCB102 from *C. butyricum* and pIM13 from *Bacillus subtilis*, respectively), we constructed a pIM13 derivative of pMTL007C-E2, named pMTL007C-E6, in order to allow direct comparison with pSY6catP. Like pSY6catP, pMTL007C-E6 suffered from the same electrotransformation inhibition, generating only 3.2×10^0 transformants μg^{-1}, whereas the empty control vector, pMTL85141, gave 1.5×10^4 transformants μg^{-1} (Figure 5.1). Taken together, these outcomes suggest that shuttle vectors carrying the Ll.ltrB intron machinery are inhibitory to electrotransformation of *C. pasteurianum*.

5.3.2. Inability to transform Ll.ltrB-containing vectors is due to presence of the ltrA *IEP gene and is not the result of toxicity or* E. coli *methylation*

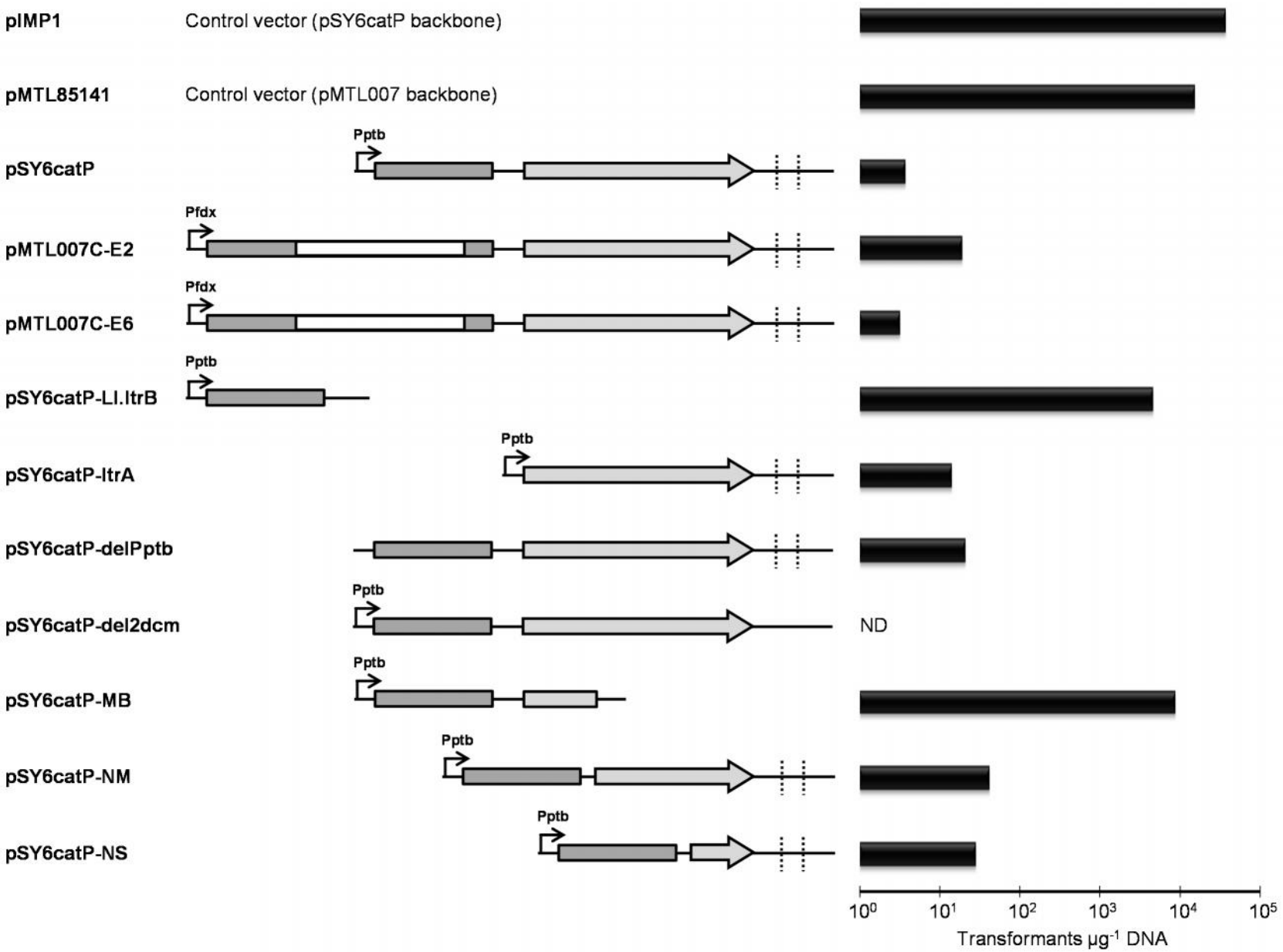
Since the Ll.ltrB intron and its cognate IEP gene, *ltrA*, are expressed from the same *ptb* promoter, we aimed to express the intron components separately in order to determine which element is responsible for the reduction in electrotransformation efficiency. We constructed plasmids pSY6catP-Ll.ltrB and pSY6catP-ltrA, which individually express the Ll.ltrB intron and

ltrA gene, respectively, both from the same constitutive *C. acetobutylicum ptb* promoter. Upon transfer to *C. pasteurianum*, only pSY6catP-ltrA hindered electrotransformation (efficiency of 1.4×10^1 transformants μg^{-1}), whereas pSY6catP-Ll.ltrB generated 4.6×10^3 transformants μg^{-1} (Fiure 5.1), indicating that the transformation inhibition is specific to the *ltrA* region of pSY6catP.

As plasmid pSY6catP-ltrA expresses a functional LtrA IEP, the low electrotransformation efficiency obtained could be the result of toxicity of the lactococcal *ltrA* gene product. To test this hypothesis, we constructed pSY6catP-delPptb, in which the -35 and -10 signals of the *ptb* promoter within pSY6catP were deleted. The resulting plasmid should not express any Ll.ltrB intron RNA or functional IEP. Despite deletion of the promoter controlling transcription of the intron and IEP, pSY6catP-delPptb failed to improve the electrotransformation efficiency and generated 2.1×10^1 transformants μg^{-1}, a comparable efficiency to unmodified pSY6catP (Figure 5.1).

If toxicity of the IEP is not responsible for the drastic reduction in electrotransformation efficiency, we reasoned that a specific nucleotide sequence within the IEP must be cleaved by an uncharacterized restriction-modification system in *C. pasteurianum*. Plasmid pSY6catP-ltrA contains two *E. coli* Dcm (5'-CCWGG-3') restriction recognition sites immediately downstream of the *ltrA* coding sequence, which are not found in the parental vector, pIMP1. Since *E. coli-C. pasteurianum* shuttle vector preparations destined for *C. pasteurianum* are first methylated in a Dcm^+ *E. coli* host strain (ER 1821), all such plasmids are methylated at both Dcm sites (5'-CmCWGG-3'). Certain methylated Dcm sites have been shown to potently inhibit electrotransformation of *C. thermocellum* [235] and *C. ljungdahlii* [107]. Therefore, we questioned if one or both of the Dcm methylation sites located downstream of the *ltrA* coding region were responsible for the decline in electrotransformation efficiency. We constructed plasmid pSY6catP-del2dcm, in which both Dcm recognition sites were mutated. Similar to unmodified pSY6catP, plasmid pSY6catP-del2dcm failed to generate any detectable transformants (electrotransformation efficiency of 0 transformants μg^{-1}) (Figure 5.1), indicating that Dcm methylation is not detrimental to electrotransformation of *C. pasteurianum*.

Figure 5.1. Electrotransformation data demonstrating that *C. pasteurianum* restricts a 932 bp SacII-BstAPI fragment within *ltrA* of pSY6catP. Only relevant vector regions corresponding to the Ll.ltrB intron elements are shown and are depicted to scale. Shaded box: Ll.ltrB intron; unshaded box: *ermB* RAM; shaded arrow: *ltrA*; dashed line: Dcm recognition site; Pptb: *ptb* promoter (*C. acetobutylicum*); Pfdx: *fdx* promoter (*C. sporogenes*); ND: not detected.


pIMP1
Control vector (pSY6catP backbone)
pMTL85141
Control vector (pMTL007 backbone)
Pptb
pSY6catP
Pfdx
pMTL007C-E2
Pfdx
pMTL007C-E6
Pptb
pSY6catP-Ll.ltrB
Pptb
pSY6catP-ltrA
pSY6catP-delPptb
Pptb
pSY6catP-del2dcm
ND
Pptb
pSY6catP-MB
Pptb
pSY6catP-NM
Pptb
pSY6catP-NS
10^0
10^1
10^2
10^3
10^4
10^5
Transformants µg-1 DNA

5.3.3. C. pasteurianum *restricts a 339 bp region of the* ltrA *ORF, which can be overcome by extensive codon modification*

To attempt to identify the unknown restriction recognition sequence within plasmid pSY6catP-ltrA, three constructs were prepared in which various-sized restriction fragments of the *ltrA* gene and the downstream region of the coding sequence were deleted and replaced with a 48 bp stuffer fragment. The sizes of these restriction fragment deletions were 1,661 bp (pSY6catP-MB), 1,434 bp (pSY6catP-NS), and 688 bp (pSY6catP-NM). Upon electrotransformation of the three vectors, only pSY6catP-MB gave an improved electrotransformation efficiency (8.8×10^3 transformants μg^{-1}), whereas pSY6catP-NS and pSY6catP-NM yielded efficiencies of 2.8×10^1 and 4.2×10^1 transformants μg^{-1}, respectively, which are comparable to pSY6catP (Figure 5.1). This result indicates that the putative restriction endonuclease recognition sequence is contained within a 932 bp SacII-BstAPI restriction fragment of pSY6catP corresponding to 493 bp of the 3' region and 439 bp downstream of the *ltrA* coding sequence.

In further attempts to reduce the size of the vector region responsible for electrotransformation inhibition, we constructed three vectors in which a 1,332 bp BglII-EcoO109I restriction fragment was removed from pSY6catP and replaced with one of three different regions of the 932 bp SacII-BstAPI restriction fragment that was found to inhibit electrotransformation of *C. pasteurianum*. Of the initial 932 bp region, plasmids pSY6catP-frag1, pSY6catP-frag2, and pSY6catP-frag3 possess portions of sizes 339, 338, and 305 bp, respectively. Approximately 20-30 bp overlap was contained between PCR fragments to ensure the unknown restriction recognition site would be represented in its entirety. Of these three constructs, only pSY6catP-frag1 showed a reduction in electotransformation efficiency (7.1×10^1 transformants μg^{-1}) compared to the control vector, as pSY6catP-frag2 and pSY6catP-frag3 yielded efficiencies of 3.5×10^4 and 9.7×10^3 transformants μg^{-1}, respectively (Figure 5.2). This result allowed the inhibitory region of pSY6catP to be reduced from 932 to 339 bp. To further demonstrate that the 339 bp region of the *ltrA* coding sequence is responsible for inhibition of electrotransformation of *C. pasteurianum*, we aimed to insert the detrimental sequence into a control vector that is otherwise able to electrotransform *C. pasteurianum* at a high efficiency. Thus, we constructed plasmid pMTLSY334 by subcloning the 339 bp SacII-AatII restriction fragment from pSY6catP into pMTL85141. Plasmid pMTL85141 consistently transforms *C.*

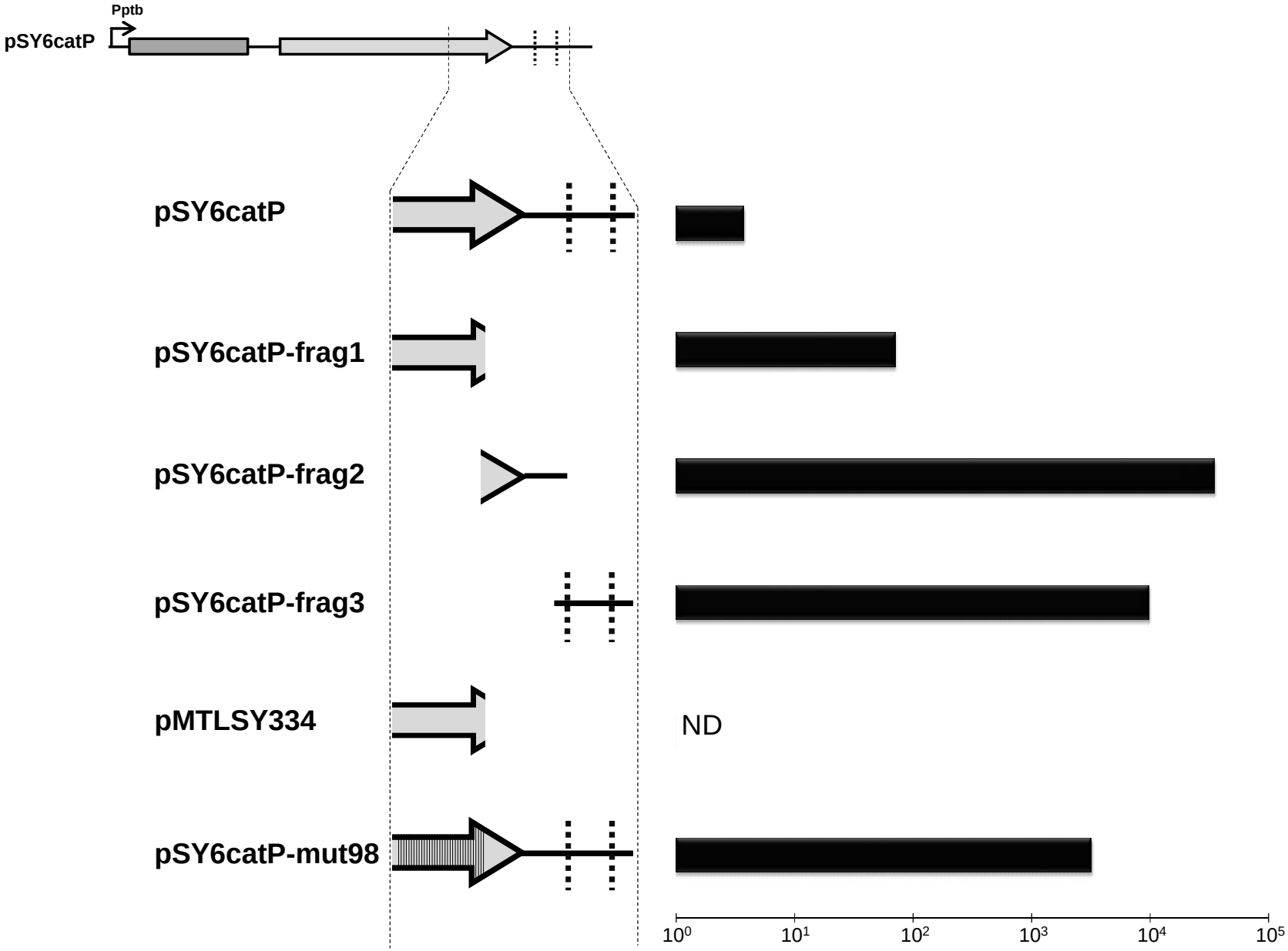
pasteurianum with efficiencies on the order of 10^3-10^4 transformants μg^{-1}. As expected, pMTLSY334 failed to generate any detectable transformants upon two attempts at electrotransformation (electrotransformation efficiency of 0 transformants μg^{-1}) (Figure 5.2), thus demonstrating the potent activity of the uncharacterized restriction-modification system expressed by *C. pasteurianum*.

In an initial attempt to overcome the uncharacterized restriction barrier posed by *C. pasteurianum*, we attempted to mutate the unknown restriction recognition sequence by mutating the entire 339 bp SacII-AatII restriction region of pSY6catP. Since this region is contained within the *ltrA* open reading frame, our efforts were limited to silent mutations which would preserve the amino acid sequence and yield a functional *ltrA* gene product. Consequently, we mutated 83 codons within the *ltrA* coding sequence, in which 98 bp of the 339 bp region were effectively mutated, to generate plasmid pSY6catP-mut98. The greatest number of consecutive nucleotides left unaltered following mutagenesis was 5 bp. Since most restriction-modification recognition sequences are ≥6 bp, we hypothesized that the unknown restriction recognition site would be effectively mutated within pSY6catP-mut98. Accordingly, pSY6catP-mut98 transformed *C. pasteurianum* with an efficiency of 3.2×10^3 transformants μg^{-1}, an increase of approximately three orders of magnitude compared to unmodified pSY6catP (Figure 5.2). We then subcloned the mutated 339 bp *ltrA* coding region into plasmids pMTL007C-E2 and pMTL007C-E6 to give pMTLCP-E2 and pMTLCP-E6, respectively, to determine if the same codon modification strategy could also enhance electrotransformation of ClosTron vectors. Similar to pSY6catP-mut98, codon modification led to significantly increased electrotransformation efficiencies for both pMTLCP-E2 (1.9×10^3 transformants μg^{-1}) and pMTLCP-E6 (2.7×10^3 transformants μg^{-1}), compared to 1.9×10^1 transformants μg^{-1} for pMTL007C-E2 and 3.2×10^0 transformants μg^{-1} for pMTL007C-E6.

5.3.4. SMRT methylome analysis unveils a unique Type I restriction-modification system in C. pasteurianum *that restricts a 5'-AAGNNNNNCTCC-3' site within pSY6catP*

As a result of polymerization kinetics exhibited by DNA polymerase in single molecule real-time sequencing applications, it is possible to obtain insight into an organism's unique DNA methylation profile, or methylome [89]. To determine a putative recognition sequence of the

Figure 5.2. Electrotransformation data demonstrating that *C. pasteurianum* restricts a 339 bp region of *ltrA* within pSY6catP. The 932 bp SacII-BstAPI region of pSY6catP (see Figure 5.1) is enlarged to better show relevant vector components. Point mutations within the *ltrA* coding sequence are depicted as vertical bands. Enlarged vector components are shown to scale. Shaded box: Ll.ltrB intron; shaded arrow: *ltrA*; thick dashed line: Dcm recognition site; Pptb: *ptb* promoter (*C. acetobutylicum*); ND: not detected.


Pptb
pSY6catP
pSY6catP
pSY6catP-frag1
pSY6catP-frag2
pSY6catP-frag3
pMTLSY334
ND
pSY6catP-mut98
10^0
10^1
10^2
10^3
10^4
10^5

uncharacterized RM system expressed by *C. pasteurianum*, we first obtained SMRT sequencing data using the upgraded RS II analyzer and total genomic DNA from *C. pasteurianum*. A coverage of 30-60× was obtained. As a result of current technical restraints in SMRT methylome analysis, only m6A residues could be detected, as m5C methylation requires at least 250× coverage [89]. Methylome analysis revealed three previously unidentified m6A RM systems in *C. pasteurianum*, in addition to the documented CpaI enzyme, an isoschizomer of the *E. coli* Dam system (Table 5.3).

The DNA sequence of the 339 bp SacII-AatII fragment of pSY6catP was scanned for the presence of the predicted RM recognition sequences unveiled following methylome analysis. One such system, which recognizes the sequence 5'-AAGNNNNNCTCC-3', was found once within the 339 bp region shown to inhibit electrotransformation of *C. pasteurianum*. The recognition sequence of the new restriction activity is characteristic of a Type I system, which possess a defined 3 bp 5' region and a defined 3-4 bp 3' region separated by 5-8 degenerate N residues [255]. Moreover, a single Type I RM system composed of restriction (*hsdR*), methylation (*hsdM*), and specificity (*hsdS*) host specificity domains (*hsd*) was identified in the draft genome assembly of *C. pasteurianum* ATCC 6013. Therefore, the newly identified Type I RM system in *C. pasteurianum* was named CpaAII.

We scanned all other *E. coli-C. pasteurianum* shuttle vectors transformed to date and, as expected, only pMTL007C-E2 and pMTL007C-E6 were found to possess a CpaAII recognition site. The CpaAII site within these vectors is identical to the one in pSY6catP, as it is located within the *ltrA* gene sequence of the Ll.ltrB intron machinery. Unexpectedly, however, pMTL007C-E2 and pMTL007C-E6 were found to possess an additional CpaAII site within the erythromycin resistance RAM. To assess the effect of the additional CpaAII site on electrotransformation of *C. pasteurianum*, we electrotransformed plasmids pMTLCP-E2 and pMTLCP-E6, in which the detrimental CpaAII site within the *ltrA* coding sequence is mutated, yet the additional site within the *ermB* RAM is left unaltered. As stated above (section 5.3.3), pMTLCP-E2 and pMTLCP-E6 generated electrotransformation efficiencies of 1.9×10^3 and 2.7×10^3 transformants μg^{-1}, respectively, signifying that the additional CpaAII site within these vectors may not be subject to restriction by CpaAII. Sanger sequencing of the *ermB* RAM

Table 5.3
Methylation motifs identified from methylome analysis of *C. pasteurianum* SMRT sequencing data

Putative recognition sequence (5'-3')[a]	Modification type[b]	Number of sites in genome	Number of methylated sites detected (percentage)
A<u>A</u>GNNNNNC**T**CC	m6A	524	487 (92.9%)
GRTAA<u>A</u>G	m6A	2,612	987 (37.8%)
G<u>A</u>**T**C	m6A	13,630	5,105 (37.4%)
CAAAA<u>A</u>R	m6A	3,945	945 (24.0%)

[a] Methylated residues are underlined and bolded in the sense and antisense strands, respectively.
[b] m6A was the only modification that could be detected in this experiment; coverage was not great enough to allow detection of m5C and m4C residues.

vector region in pMTL007C-E2, pMTLCP-E2, pMTL007C-E6, and pMTLCP-E6 confirmed the presence of an additional CpaAII recognition sequence in all four vectors.

5.3.5. Generation of an intron-mediated gene knockout mutant of C. pasteurianum

Since plasmid pSY6catP-mut98 afforded a substantial improvement in electrotransformation efficiency compared to pSY6catP (approximately three orders of magnitude) by mutation of the unique CpaAII restriction recognition sequence, we used it as the basis for our *C. pasteurianum* gene knockout vector. To further improve our knockout system, we replaced the *C. acetobutylicum ptb* promoter controlling Ll.ltrB + *ltrA* transcription with a strong, constitutive promoter from the *C. pasteurianum* thiolase gene. The resulting vector, pSYCP, was then targeted to nucleotide position 176 within the antisense strand (176a) of the CpaAI gene coding sequence. We selected the gene encoding the CpaAI restriction endonuclease for disruption since the 176a insertion site generated a high predicted intron insertion score of 7.3, allowed for optimal δ-δ' base-pairing between the intron RNA (δ=G) and the 176a insertion site (δ'=C) [194], is unlikely to lead to a lethal phenotype, and should produce a restriction-deficient mutant strain that is useful for future strain construction endeavors by abolishing the requirement for plasmid methylation prior to electrotransformation [228]. The retargeted pSYCP-REN plasmid was electroporated to *C. pasteurianum* and transformants were selected using thiamphenicol. Transformant colonies were screened for insertion of the Ll.ltrB intron within the CpaAI coding sequence, resulting in an insertion of approximately 0.9 kb, using two gene-specific primers flanking the 176a insertion site. Of 28 screened colonies, all 28 possessed only the wild-type PCR product, as the knockout band could not be detected (data not shown). However, both adjacent gene-intron junction PCR products could be detected using one gene-specific and one intron-specific primer in approximately half of the colonies screened, signifying successful intron insertion, yet poor retrohoming efficiency. The resulting colonies were comprised of a mixture of wild-type and knockout cells. To separate the mosaic colonies and enrich the gene knockout genotype, we pooled and subcultured mosaic colonies in selective liquid medium every 12 hours for a total of 5 days (10 transfers) and screened the resulting colonies for the intron insertion using two gene-specific primers. Of 12 colonies screened, four possessed the desired intron insertion, three possessed the wild-type genotype, and five failed to

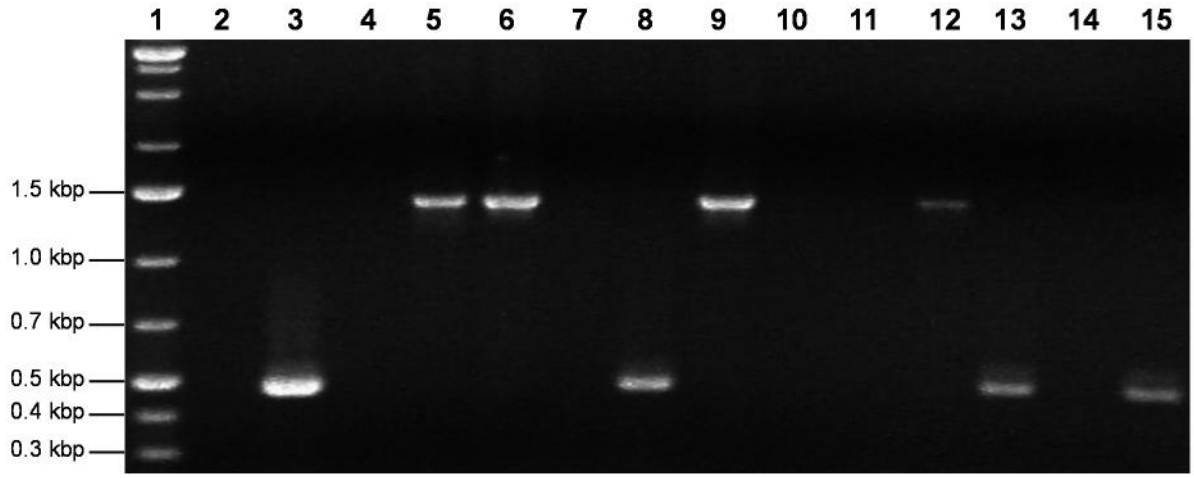


Figure 5.3. Identification and verification of the *C. pasteurianum* CpaAI disruption mutant. Colony PCR screening of non-selective colonies resulting from the gene knockout enrichment procedure was performed with gene-specific primers flanking the predicted 176a CpaAI intron insertion site. Lane 1: marker; lane 2: no template control; lane 3: wild-type *C. pasteurianum* colony; lanes 4-15: non-selective enrichment colonies; lanes 5, 6, 9, and 12: positive clones; lanes 8, 13, and 15: negative clones; lanes 4, 7, 10, 11, and 14: no amplification.

generate a detectable PCR product (Figure 5.3). Therefore, our gene knockout enrichment procedure generated a population in which more than half of the cells harbored the desired intron insertion. We are currently characterizing the resulting CpaAI REN disruption mutant for the ability to be transformed with unmethylated plasmid substrates.

5.4. Discussion

Genetic tools, particularly for chromosomal gene knockout, have been developed and applied for metabolic engineering of many species of solventogenic clostridia [8]. To date, however, no techniques have been reported for disrupting native chromosomal genes in *C. pasteurianum*. Here we show that the ability to carry out group-II-intron-mediated gene knockouts in *C. pasteurianum* is limited by host restriction and low-level intron retrohoming, both of which have been overcome in this report.

During development of a host-vector system and electrotransformation methodology for *C. pasteurianum*, we observed that all vectors harboring lactococcal group II intron machinery transformed at two to three orders of magnitude less than vectors lacking the Ll.ltrB intron and its cognate IEP gene [91]. By assessing transformability of various deletion derivatives of pSY6catP and performing methylome analysis from *C. pasteurianum* SMRT sequencing data, we concluded that the inability of intron-containing vectors to efficiently electrotransform *C. pasteurianum* is due to restriction by a new RM system that we have designated CpaAII. Based on methylome analysis data, M.CpaAII methylates the recognition sequence 5'-A<u>A</u>GNNNNNCTCC-3' and its complement 5'-GG<u>A</u>GNNNNNCTT-3', (methylated nucleotide residues are underlined). The *C. pasteurianum* DSM 525 (ATCC 6013) draft genome sequence possesses seven DNA methyltransferase genes that could play a role in host restriction, six of which are Type II, while the remaining is Type I. Two of these enzymes correspond to the previously characterized adenine-specific M.CpaI (5'-G<u>A</u>TC-3') and cytosine-specific M.CpaAI (5'-<u>C</u>GCG-3') Type II RM systems [90]. This leaves five methyltransferases that could be responsible for the M.CpaAII activity identified in this study. However, only two such enzymes, one Type I and one Type II, are associated with a corresponding restriction endonuclease, one of which is presumed to be CpaAII. We hypothesize that the CpaAII system is of the Type I family. Type I RM systems are encoded by three host specificity domains (*hsd*) corresponding to restriction (*hsdR*), modification (*hsdM*), and specificity (*hsdS*) [255]. A single Type I RM system

is annotated in our draft genome sequence for *C. pasteurianum*, including all three *hsd* genes. All *hsdS* proteins consist of two highly conserved target recognition domains which are responsible for recognizing the bipartite arrangement of Type I recognition sites [255]. Type I restriction recognition sequences possess a defined 3 bp sequence at the 5' end and a defined 3-4 bp sequence at the 3' end, which are separated by 5-8 degenerate N residues [90]. The putative CpaAII recognition sequence conforms to this bipartite structure. Further, all Type I systems uncovered to date possess adenine-specific M.CpaAII proteins, as is the case for *C. pasteurianum*. Therefore, we propose that the newly identified CpaAII system is part of the Type I RM family.

All vectors that demonstrated a significantly reduced ability to electrotransform *C. pasteurianum* were found to possess at least one CpaAII recognition sequence. In the case of pSY6catP, mutation of the single CpaAII recognition site within the *ltrA* coding sequence rescued high-level electrotransformation. Conversely, only one of the two CpaAII restriction recognition sequences within ClosTron plasmids pMTL007C-E2 and pMTL007C-E6, however, was found to negatively affect electrotransformation efficiency of *C. pasteurianum*. A possible explanation for this is that the actual CpaAII recognition sequence may differ slightly from the sequence predicted from methylome analysis. The actual CpaAII recognition sequence could be longer than the predicted 5'-AAGNNNNNCTCC-3' sequence. For example, the Type I RM system from *E. coli* ECL 394, 5'-GACNNNNNRTAAY-3' [90], contains a 5 bp element at the 3' end, rather than the typical 3-4 bp, and also incorporates degenerate nucleotides (R=A or G; Y=C or T). Increasing the coverage of our SMRT sequencing reads from the RS II sequencer and repeating the methylome analysis would improve the accuracy of our predicted recognition sequence and resolve the discrepancy in our experimental data.

Having resolved the restriction issue pertaining to plasmid pSY6catP, we used pSY6catP-mut98 as the basis for our gene knockout vector, in which the single CpaAII site was deleted and transcription of the Ll.ltrB intron and its cognate IEP is controlled from the constitutive *C. pasteurianum thl* promoter. To demonstrate chromosomal gene knockout in *C. pasteurianum* we targeted our resulting pSYCP-REN vector to the Type II CpaAI restriction endonuclease gene. Retrohoming efficiency proved to be too low to isolate a gene knockout mutant directly from transformant colonies, as most colonies were found to represent a heterogeneous mixture of wild-type and knockout cells. The use of a selectable RAM inserted into domain IV of the intron

also did not permit isolation of a gene knockout mutant, likely due to a further decrease in retrohoming efficiency due to the presence of excess cargo DNA within the intron [145]. To isolate a true knockout colony, we developed an enrichment procedure in which pSYCP-REN transformant colonies are subjected to repeated subculturing in selective liquid medium to encourage intron insertion. Following our devised enrichment protocol, approximately half of the colonies possessed the desired intron insertion. The resulting insertional mutant is expected to be deficient in expression of the CpaAI Type II restriction endonuclease, and therefore, should be transformable using unmethylated plasmid DNA [228]. We are currently characterizing our gene knockout mutant for loss of CpaAI restriction activity. The CpaAI mutant should prove useful for future genetic applications, as electrotransformation will be greatly simplified without the requirement for plasmid methylation from the FnuDIIM methyltransferase.

5.5. Conclusions

In this work, a previously uncharacterized Type I RM system, designated CpaAII, was identified and shown to inhibit electrotransformation of plasmids harboring group II intron gene knockout machinery. Overcoming CpaAII restriction, coupled to a gene knockout enrichment procedure, allowed for the first time disruption of a chromosomal gene in *C. pasteurianum*. Our resulting CpaAI-deficient mutant, in addition to our optimized gene knockout vector and enrichment protocol, should prove invaluable for future genetic and metabolic engineering approaches involving *C. pasteurianum*.

CHAPTER 6

Plasmid-based gene overexpression and antisense-RNA-mediated gene knockdown in *Clostridium pasteurianum*

6.1. Introduction

For organisms that carry out highly branched, complex fermentations, like that of the *C. pasteurianum* glycerol fermentation, gene knockout and knockdown techniques have proven vital for engineering strains with improved metabolic phenotypes [83,256]. Although *C. pasteurianum* naturally produces considerable quantities of butanol (up to 17 g l^{-1}) when grown on highly reduced substrates, competing products, particularly 1,3-propanediol and butyrate, accumulate in the culture medium, thus limiting flux through the butanol biosynthetic pathway [17,66]. Therefore, metabolic engineering is a viable option for increasing butanol yields and productivities from the fermentation of glycerol by *C. pasteurianum*. Unfortunately, few genetic tools are currently available to perform metabolic engineering of this important industrial bacterium.

A fundamental facet of metabolic engineering is the ability to decrease the expression of genes involved in competing metabolic pathways, which is accomplished through the use of gene knockdown or knockout techniques [8]. Gene knockouts lead to complete inhibition of gene expression and, therefore, often produce lethal phenotypes, particularly for genes involved in central fermentative metabolism. Antisense-RNA-mediated gene downregulation, on the other hand, is a more subtle approach for decreasing the activity of native chromosomal genes and metabolic pathways, as production of the target protein is still permitted, though at a significantly reduced level [13]. Antisense RNAs are designed to possess partial or complete complementarity to the target mRNA molecule, leading to base-pairing and RNase-mediated degradation of the resulting RNA duplex before protein translation is allowed to occur. This strategy has been employed in *C. acetobutylicum*, *C. beijerinckii*, and *C. saccharoperbutylacetonicum*, among other species, for downregulation of competing metabolic pathways, including butyrate and acetone formation pathways, to enhance butanol titers, selectivity, and tolerance [13,77,146,184]. The degree of downregulation achieved in these reports has ranged from 44 to 92% downregulation [77], which is commonly assessed by measuring enzyme activity or target

protein levels via Western blot analysis. Downregulation can be finely tuned by selection of the gene promoter driving production of the asRNA transcript, where stronger promoters generate more asRNA product, leading to a greater degree of downregulation. Promoters utilized in successful asRNA approaches include ones from genes involved in central carbon metabolism, such as phosphotransbutyrylase (*ptb*), thiolase (*thl*), and butanol dehydrogenase (*bdh*).

In this report, we investigate the efficacy of asRNA strategies to perturb the carbon and electron flow of the glycerol fermentation carried out by *C. pasteurianum*. We first employed a common clostridial gene reporter system based on a thermophilic β-galactosidase gene (*lacZ*) [12] to assess the relative strength of various gene promoters involved in the central fermentative metabolism of *C. pasteurianum*. Then, we utilized our chosen promoter to drive transcription of an asRNA molecule targeted to our *lacZ* reporter gene to assess the ability of asRNA to downregulate genes in *C. pasteurianum*. Finally, we applied our asRNA strategy to attempt downregulation of several central fermentative genes that potentially limit carbon and electron flow to the butanol biosynthetic pathway.

6.2. Methods

6.2.1. Bacterial strains, cultivation conditions, and electrotransformation

Bacterial strains utilized in this study are listed in Table 6.1. For routine cultivation and maintenance, recombinant and non-recombinant *E. coli* and *C. pasteurianum* strains were grown according to section 3.2.2. Recombinant *C. pasteurianum* strains were selected, when necessary, with 10 μg ml^{-1} thiamphenicol or 20 μg ml^{-1} erythromycin.

Batch shaker flask glycerol fermentations were performed by first growing a single heat-shocked sporulated colony in 10 ml 2×YTG containing an appropriate antibiotic. Cultures grown to exponential phase were used to inoculate plastic screw-cap flasks containing 50 ml of Biebl medium (per liter: 50 g glycerol, 0.5 g KH_2PO_4, 0.5 g K_2HPO_4, 0.2 g $MgSO_4 \cdot 7H_2O$, 0.02 g $CaCl_2 \cdot 2H_2O$, and 2 ml trace element solution SL7) [66], with concentrations of the $(NH_4)_2SO_4$ (7.35 g l^{-1}), yeast extract (5.08 g l^{-1}), $CaCO_3$ (3.15 g l^{-1}), and $FeSO_4 \cdot 7H_2O$ (60 mg l^{-1}) components optimized for butanol production [257,258]. Vitamin B_{12} (cyanocobalamin), cysteine-HCl, and resazurin were also added to final concentrations of 0.2 μM, 0.5 g l^{-1}, and 1 mg l^{-1}, respectively. Following inoculation, cultures were grown anaerobically without shaking

for 12-18 hours until gas production could be observed, at which time the flasks were sealed and incubated outside of the anaerobic chamber at 37 ºC and 200 rpm. A single end-point sample for growth and product analyses was taken after 24-48 h of incubation with shaking, corresponding to a change in color of the culture medium to pink due to cessation of growth and gas production. All strains were cultivated in duplicate, unless stated otherwise.

6.2.2. DNA isolation, manipulation, and plasmid construction

Plasmids and oligonucleotides utilized in this work are shown in Table 6.1. Oligonucleotides were synthesized by Integrated DNA Technologies (IDT; Coralville, IA) at the 25 nm scale using standard desalting. Restriction endonucleases, Quick Ligation Kit, Standard *Taq* DNA Polymerase, and Phusion DNA Polymerase were purchased from New England Biolabs (Ipswich, MA). Molecular biology kits for purification of plasmid DNA, PCR products, and gel extracted DNA, and custom gene synthesis products, were purchased from Bio Basic Inc. (Markham, ON). Recombinant DNA procedures were performed according to standard procedures [227]. Total genomic DNA was isolated from *C. pasteurianum* as described in section 4.2.2.

A thiamphenicol derivative of promoterless pHT3, designated pHT3catP, was constructed by subcloning a 1.3 kbp ClaI + SacI restriction fragment of pSY6catP containing a thiamphenicol-resistance cassette into the corresponding restriction sites of the erythromycin-resistance determinant of pHT3. Transcriptional promoter fusions with *lacZ* in pHT3catP were prepared by amplifying putative *fdx* and *thl* promoter regions from *C. pasteurianum* total genomic DNA using primer pairs Pfdx.NarI.S + Pfdx.XmaI.AS and Pthl.NarI.S + Pthl.XmaI.AS, respectively. The resulting 565 bp (*fdx*) and 512 bp (*thl*) PCR products were digested with NarI + XmaI and ligated into the corresponding sites of pHT3catP to generate pHT3catP-fdx and pHT3catP-thl. To construct plasmid pHTaslacZ110, a *lacZ* asRNA expression cassette containing the *C. pasteurianum thl* promoter, *lacZ* asRNA, and *C. pasteurianum fdx* transcriptional terminator was obtained as a custom AhdI- and BstAPI-flanked gene synthesis fragment. The resulting 517 bp fragment was digested with AhdI + BstAPI and inserted into the corresponding restriction sites of pHT3catP-thl. Plasmids pHTasbuk123 and pHTasptb123 were constructed by first obtaining 224 bp BamHI- and BstXI-flanked custom gene synthesis products possessing a 123 nt *buk* or *ptb* asRNA upstream of the *C. pasteurianum fdx* transcriptional

Table 6.1
Strain, plasmids, and oligonucleotides.

Strains or plasmids	Relevant characteristics	Source or reference
Strains		
Escherichia coli DH5α	F- *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG φ80dlacZ*a*M15* a*(lacZYA-argF)U169, hsdR17($r_K^-m_K^+$),* a-	Lab stock
Escherichia coli ER1821	F- *endA1 glnV44 thi-1 relA1? e14-(mcrA-) rfbD1? spoT1?* a*(mcrC-mrr)114::IS10*	Lab stock; New England Biolabs
Clostridium pasteurianum ATCC 6013	Wild-type	American Type Culture Collection
Plasmids		
pFnuDIIMKn	*E. coli* vector expressing the M.FnuDII methyltransferase for plasmid methylation prior to electrotransformation of *C. pasteurianum*	[91]
pMTL85141	*E. coli-C. pasteurianum* shuttle vector (Cm^R/Tm^R; ColE1 ori; pIM13 ori)	[139]
pSY6catP	*E. coli-Clostridium* shuttle vector expressing the Ll.ltrB intron and *ltrA* (Ap^R ; Cm^R/Tm^R; ColE1 ori; pIM13 ori)	[91]
pHT3	Promoterless *E. coli-C. pasteurianum* β-galactosidase gene expression reporter vector (Ap^R; ColE1 ori; Em^R; pIM13 ori)	[12]
pIMP1	*E. coli-Clostridium* shuttle vector (Ap^R; ColE1 ori; Em^R; pIM13 ori)	[10]
pHT3catP	Thiamphenicol-resistant derivative of pHT3	This study
pHT3catP-fdx	Derived by transcriptionally fusing the *C. pasteurianum fdx* promoter with the *lacZ* gene of pHT3catP	This study
pHT3catP-thl	Derived by transcriptionally fusing the *C. pasteurianum thl* promoter with the *lacZ* gene of pHT3catP	This study
pHTaslacZ110	Vector expressing a 110 nt *lacZ* asRNA from the *C. pasteurianum thl* promoter	This study
pHTasbuk123	Vector expressing a 123 nt *buk* asRNA from the *C. pasteurianum thl* promoter	This study
pHTasptb123	Vector expressing a 123 nt *ptb* asRNA from the *C. pasteurianum thl* promoter	This study
pHTasdhaB139	Vector expressing a 139 nt *dhaB* asRNA from the *C. pasteurianum thl* promoter	This study
pHTasdhaT146	Vector expressing a 146 nt *dhaT* asRNA from the *C. pasteurianum thl* promoter	This study
pHTashydA175	Vector expressing a 175 nt *hydA* asRNA from the *C. pasteurianum thl* promoter	This study

Oligonucleotide	**Sequence (5'-3')**
Pfdx.NarI.S	CAGAACGGCGCCGAAGATATAAGAAAAAGACTCCCAAAGG
Pfdx.XmaI.AS	ATCATACCCGGGCCATAACTTATTGTATCATGTTTTTAAAC
Pthl.NarI.S	CTGTAGGCGCCGATATAGTCTATAAGCATTTAGATGGAGTTAG
Pthl.XmaI.AS	AAATACCCGGGTCGATTGTTATTTAATTCACAATTTAATTATAACC
asdhaB139.BamHI.S	CTTTAAGGATCCTACTAGGCTTTGGATCATTGGGA
asdhaB139.XhoI.AS	CAACAGCTCGAGGATAAAGGAGGATTATATATGAAGTCAAAACGAT
asdhaT146.BamHI.S	ATCTTCGGATCCTCTCAAGAATTTATCTGTAACTATTAAAGCTTTC
asdhaT146.XhoI.AS	TAATTTCTCGAGCTAGGAGGAATTATAAAATGAGAATGTATGATTT
ashydA175.BamHI.S	CTAATCGGATCCCCTCTACCTCTACAGTACATATTTCACAC
ashydA175.XhoI.AS	AAAATTCTCGAGAGGAGGCTAGATATGAAAACAATAATTATAAATGG

* Underline: restriction recognition sequences.

terminator. Antisense RNAs were designed to target the RBS and 5' region of the respective *ptb* and *buk* coding sequences. A unique XhoI site was inserted between the asRNA and *fdx* terminator regions to allow simplified cloning of other asRNA constructs. The gene synthesis products were digested with BamHI + BstXI and ligated into the corresponding sites of pHT3catP-thl for transcriptional fusion of the asRNAs with the *C. pasteurianum thl* promoter. To construct pHTasdhaB139, pHTasdhaT146, and pHTashydA175, asRNAs targeting the RBS and 5' coding region of the *dhaB* (139 nt), *dhaT* (146 nt), and *hydA* (175 nt) genes were amplified from *C. pasteurianum* genomic DNA using primer pairs asdhaB139.BamHI.S + asdhaB139.XhoI.AS, asdhaT146.BamHI.S + asdhaT146.XhoI.AS, and ashydA175.BamHI.S + ashydA175.XhoI.AS, respectively. The resulting PCR products were digested with BamHI + XhoI and ligated with pHTasbuk123 that had been digested with the same enzymes.

6.2.3. β-Galactosidase time course studies and enzymatic assay

β-Galactosidase time course studies and enzymatic assay were performed as previously described [12], with slight modifications. For time course studies of pHT3catP, pHT3catP-fdx, pHT3catP-thl, and pHTaslacZ110, 45 ml samples were removed from each flask in duplicate every 3-5 hours starting at an A_{600} of ~0.5. Cells were harvested from cultures by centrifugation at 10,000 × *g* for 10 min at 4 °C, and the resulting cell pellets were stored at -80 °C for up to 7 days. Cell pellets were thawed on ice, resuspended in Z buffer to an A_{600} of 20 in a minimum volume of 1 ml, and lysed by incubation in the presence of 1 mg/ml lysozyme for 1 hour at 37 °C, unless specified otherwise. One ml of the resulting lysates was cleared of cell debris by centrifugation at 16,000 × *g* for 15 min at 4 °C and the supernatant was heated in a 60 °C water bath for 30 min. Denatured heat-labile proteins were cleared by centrifugation at 16,000 × *g* for 30 min at 4 °C. The resulting supernatants were stored at 4 °C prior to enzymatic assays. β-Galactosidase activity was assayed as described [12] at 60 °C by measuring A_{420} following development of yellow color resulting from cleavage of ortho-nitrophenyl-β-galactoside (ONPG) to orthonitrophenol.

6.3. Results

6.3.1. Constitutive expression of β-galactosidase

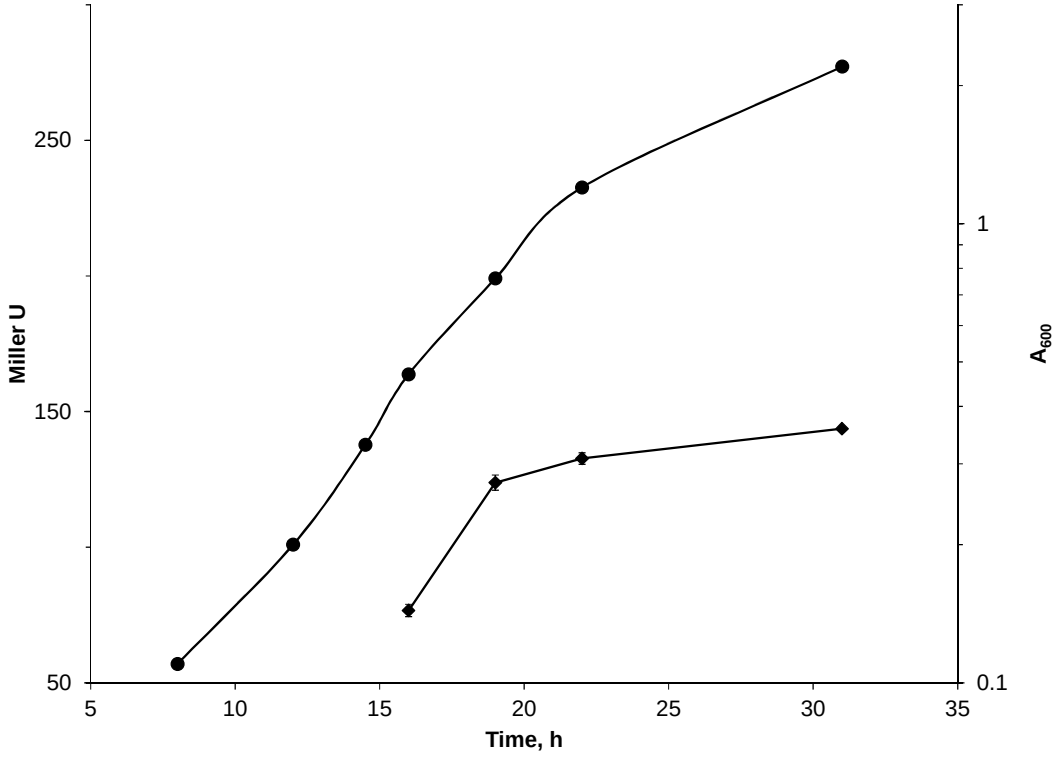
A modified thiamphenicol-selectable reporter system, pHT3catP, containing a thermophilic β-galactosidase gene (*lacZ*) was utilized to assess the function of two promoters from genes involved in the central metabolism of *C. pasteurianum*. Putative promoter regions from the native ferredoxin (*fdx*) and thiolase (*thl*) genes of *C. pasteurianum* were cloned upstream of *lacZ* in pHT3catP to generate transcriptional fusions. The resulting plasmids, pHT3catP-fdx and pHT3catP-thl, in addition to the promoterless pHT3catP control vector, were electrotransformed to *C. pasteurianum* and used in time course studies to assess constitutive expression of *lacZ*. Time course β-galactosidase activity profiles from static flask cultures are shown in Figure 6.1. Promoterless plasmid pHT3catP failed to yield appreciable β-galactosidase activity (6-7 MU) throughout the duration of growth. Conversely, both pHT3catP-fdx and pHT3catP-thl generated significant β-galactosidase activity. The *thl* promoter yielded greater maximal β-galactosidase activity (up to 440 MU; Figure 6.1B), compared to the *fdx* promoter (up to 144 MU; Figure 6.1A). Although the *fdx* promoter has been used for gene expression applications in the solventogenic clostridia, we opted for the stronger *thl* promoter to achieve high level expression of asRNA molecules.

6.3.2. Antisense-RNA-mediated downregulation of plasmid-encoded lacZ

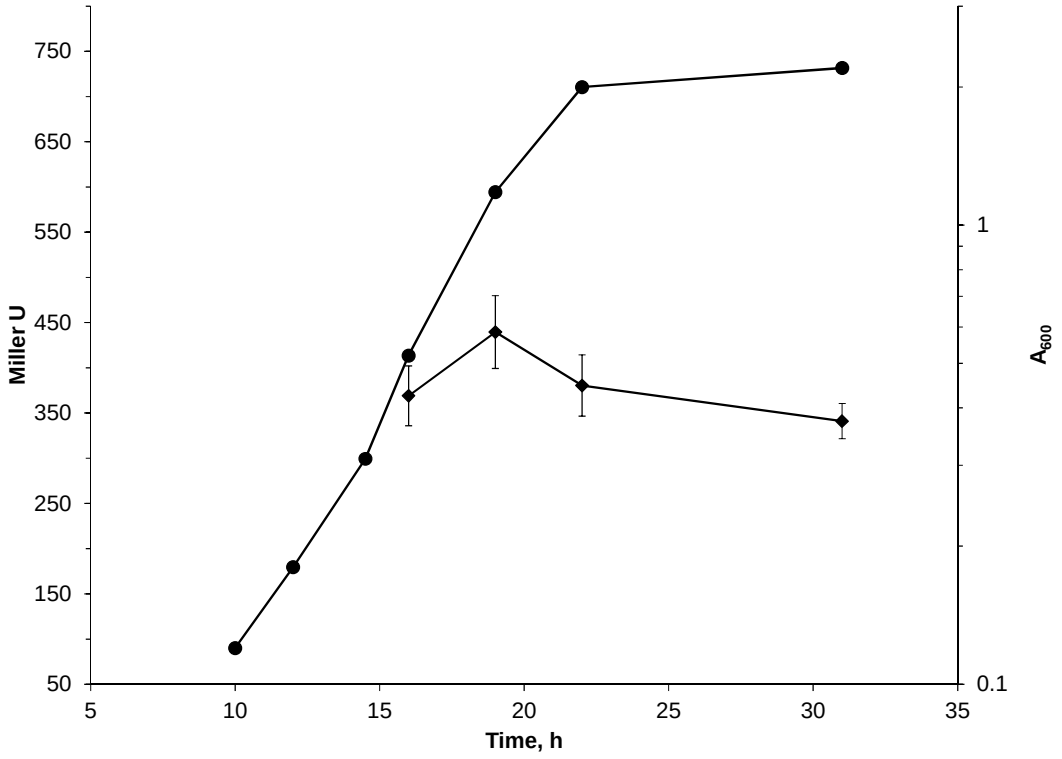
To assess the efficacy of asRNA-mediated gene downregulation in *C. pasteurianum*, we utilized the same *lacZ* gene as a reporter. An asRNA cassette was first constructed by placing a 110 bp *lacZ* asRNA molecule under transcriptional control of the strong *C. pasteurianum thl* promoter. A strong transcriptional terminator from the *C. pasteurianum fdx* gene was placed immediately downstream of the asRNA expression cassette. The asRNA molecule was designed to target 66 bp upstream of the *lacZ* initiation codon, including the RBS, and 44 bp of the 5' region of the *lacZ* coding sequence. The resulting *lacZ* asRNA cassette was inserted into plasmid pHT3catP-thl. In this configuration, both the full-length *lacZ* ORF and the *lacZ* asRNA are expressed from the same plasmid with different copies of the *thl* promoter. Strains harboring pHT3catP, pHT3catP-thl, and pHTaslacZ110 were cultivated and cells were harvested and assayed for β-galactosidase activity at various time points. At all three time points tested, the *lacZ* downregulation strain generated only 52-58% of the β-galactosidase activity produced from the control strain harboring pHT3catP-thl (Figure 6.2), demonstrating significant downregulation

Figure 6.1. Time course β-galactosidase activities of static flask cultures harboring *lacZ* transcriptional fusion constructs. A) ATCC 6013 P_{fdx}-*lacZ* transcriptional fusion (pHT3catP-fdx). B) ATCC 6013 P_{thl}-*lacZ* transcriptional fusion (pHT3catP-thl). Symbols: ●, A_{600}; ♦, β-galactosidase activity (Miller Units).

A)


250
150
50
Miller U
1
0.1
A_{600}
5
10
15
20
25
30
35
Time, h

B)


750
650
550
450
350
250
150
50
Miller U
1
0.1
A_{600}
5
10
15
20
25
30
35
Time, h

of *lacZ*. This effect was not due to promoter titration as a result of the presence of two copies of the *thl* promoter contained within pHTaslacZ110, as an alternative *lacZ* asRNA construct lacking a transcriptional terminator failed to downregulate *lacZ* and produced approximately the same levels of β-galactosidase activity as the strain possessing pHT3catP-thl (data not shown). Taken together, these data suggest that both a strong promoter and transcriptional terminator are essential for effective asRNA-mediated gene downregulation in *C. pasteurianum*.

6.3.3. Antisense-RNA-mediated downregulation of chromosomal genes in C. pasteurianum

Since our devised asRNA approach demonstrated effective downregulation of plasmid-encoded *lacZ*, we aimed to knockdown key fermentative pathway genes in the chromosome of *C. pasteurianum*. In an attempt to perturb the central carbon and electron flow in *C. pasteurianum*, we separately expressed various asRNA molecules from the *C. pasteurianum thl* promoter. Antisense RNAs targeting production of 1,3-propanediol, butyrate, and hydrogen were separately assessed for their effect on growth and product distribution. Strains were cultivated in semi-defined growth medium optimized for butanol production and buffered with solid calcium carbonate. Strains were incubated statically and anaerobically until exponential phase, at which time they were capped and moved to a shaker in the presence of air. Cultures were stopped and sampled when the culture broth turned pink, signifying the presence of oxygen that was able to enter the flasks due to termination of gas production from the cultures. Growth data from one such experiment is shown in Table 6.2. Strains harboring plasmids pMTL85141 and pIMP1 were included as vector controls. All strains grew to a final OD of 7.2 to 10.2. The *hydA* asRNA strain reached the endpoint the fastest and attained the greatest final OD. This result is consistent with hydrogenase downregulation, as decreased hydrogen production would lead to less total gas production and a greater susceptibility to oxygen exposure. All remaining asRNA strains appeared similar to the pMTL85141 and pIMP1 control cultures with respect to growth and OD. Endpoint samples from all strains are currently being assessed for metabolite production using HPLC analysis.

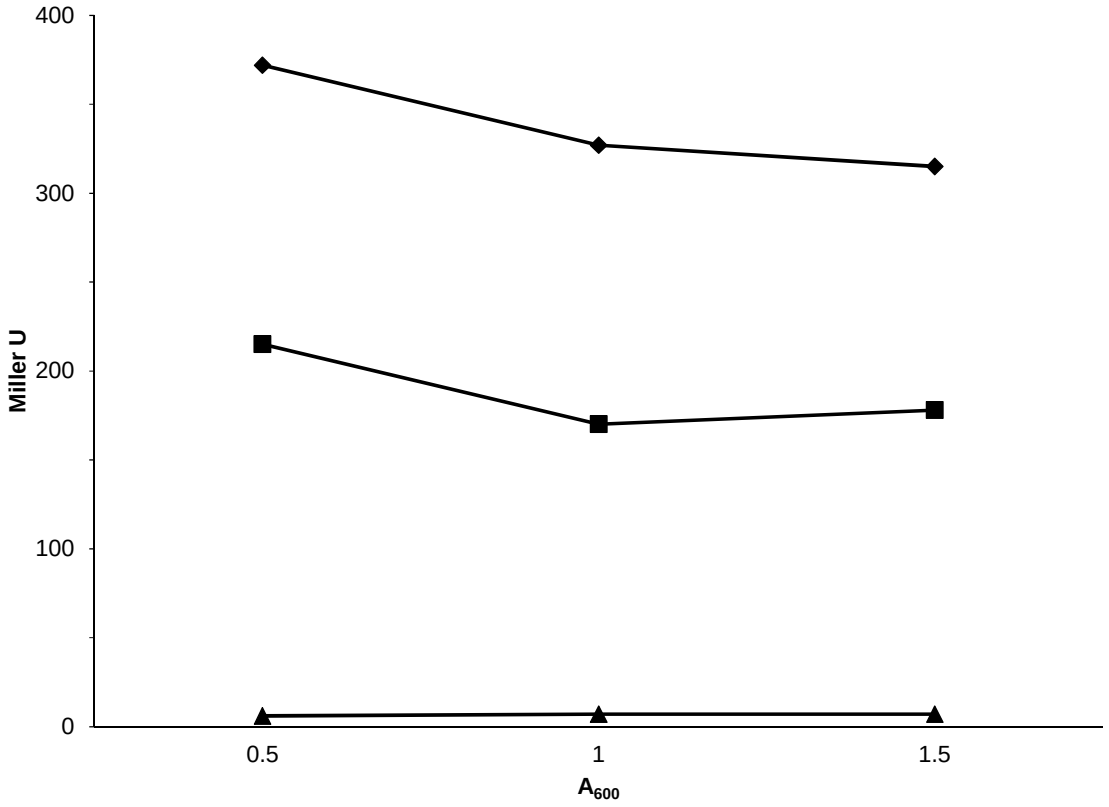


Figure 6.2. Time course β-galactosidase activities of static flask cultures harboring promoterless construct pHT3catP (▲), P_{thl}-*lacZ* transcriptional fusion plasmid pHT3catP-thl (♦), and *lacZ* asRNA construct pHTaslacZ110 (■).

Table 6.2
Growth data of control and asRNA downregulation strains grown on glycerol in semi-defined medium

Plasmid	Relevant characteristics	Static incubation[a]	Shaking incubation[b]	OD_{600}
pIMP1	Empty plasmid control (Em^R)	18.5 h	54 h	7.6
pMTL85141	Empty plasmid control (Tm^R)	18.5 h	34 h	7.6
pHTasbuk123	Expression of a 123 nt *buk* asRNA	10.5 h	45 h	8.0
pHTasptb123	Expression of a 123 nt *ptb* asRNA	10.5 h	45 h	8.8
pHTasdhaB139	Expression of a 139 nt *dhaB* asRNA	18.5 h	40.5 h	7.4
pHTasdhaT146	Expression of a 146 nt *dhaT* asRNA	18.5 h	49 h	7.2
pHTashydA175	Expression of a 175 nt *hydA* asRNA	18.5 h	24.5 h	10.2

[a] Strains were cultivated anaerobically without shaking until exponential-phase growth could be observed.
[b] Once the cultures reached exponential phase, flasks were sealed and transferred to a shaking incubator in the presence of air until the cultures turned pink due to oxygen.

6.4. Discussion

The ability to both upregulate and downregulate gene expression is fundamental to metabolic engineering [8]. These tools have found widespread use within the clostridia for various strain construction applications [83]. Since a method of gene transfer has only been reported for *C. pasteurianum* recently [91], however, these methods are lacking for genetic manipulation of this important bacterium. Here we demonstrate the utility of plasmid-based gene overexpression and asRNA-mediated downregulation of chromosomal genes in *C. pasteurianum*.

Essential to the development of fundamental genetic techniques is access to a repertoire of functional gene promoters. To compare promoters and provide a quantitative means of assessing relative promoter strength, gene reporter systems are commonly employed. In this study we selected a thermophilic β-galactosidase gene encoded by *lacZ* as the reporter system since it has found widespread use in the clostridia [12], the corresponding activity assay is straightforward, and *C. pasteurianum* has negligible endogenous β-galactosidase activity. We compared promoters from the *C. pasteurianum* ferredoxin and thiolase genes, where the thiolase gene promoter produced three-fold higher maximal β-galactosidase activity. A similar result was obtained with *C. acetobutylicum*, as the *thl* promoter out-competed promoters from the phosphotransbutyrylase and acetoacetate decarboxylase genes, although the *ptb* promoter has been used in many clostridial gene expression applications. Thiolase, which catalyzes the step in moving from 2C to 4C metabolic intermediates, would be expected to be expressed at a high level in both of these organisms, since the predominant products during glucose fermentation are butyrate in the case of *C. pasteurianum* and butyrate and butanol in the case of *C. acetobutylicum*.

Having identified the *thl* promoter as a strong, constitutive promoter in *C. pasteurianum*, we selected it to drive transcription of our asRNA constructs. We constructed a single plasmid system for expression of both our *lacZ* reporter gene and *lacZ* asRNA. Using this system, we achieved greater than 50% downregulation at all time points measured, demonstrating that asRNA downregulation represents a viable option for decreasing gene expression in *C. pasteurianum*. Using our strategy, the functional *lacZ* gene and the *lacZ* asRNA were separately expressed from two copies of the same *thl* promoter within plasmid pHTaslacZ110. Assuming transcription levels are comparable for RNAs of different size, it can be assumed that the cytoplasmic level of functional *lacZ* mRNA would be equal to that of *lacZ* asRNA. Under these

conditions, downregulation would be expected to reach 100% since the *lacZ* mRNA would be completely titrated with *lacZ* asRNA. It is clear from this study, however, that approximately half of the *lacZ* mRNA transcripts are translated into functional proteins, suggesting that translational machinery is able to out-compete a proportion of the asRNA molecules for binding to the mRNA transcript. An alternative outcome is expected to arise when employing our plasmid-based asRNA system for downregulation of chromosomal genes in *C. pasteurianum*. Antisense RNAs expressed from the strong *thl* promoter from our multicopy plasmid should greatly outnumber mRNA transcripts produced from single-copy chromosomal genes. In this case, downregulation would be expected to greatly exceed 50%. Indeed, the extent of downregulation reported in studies involving plasmid-encoded asRNAs targeted to native chromosomal genes commonly reach up to 92% [77]. Thus, under optimal conditions, it is nearly possible to entirely abolish expression of chromosomal genes using asRNA.

Next we investigated the ability of our asRNA system to downregulate chromosomal genes encoding key metabolic pathway enzymes in *C. pasteurianum*. Genes encoding phosphotransbutyrylase, butyrate kinase, glycerol dehydratase (large subunit), 1,3-propanediol dehydrogenase, and hydrogenase were individually targeted using our asRNA approach in an attempt to perturb carbon and electron flow. Further, as a result of their central roles in production of competing metabolic pathways, downregulation of the aforementioned genes should, theoretically, enhance production of butanol. We are currently characterizing the resulting asRNA mutants with respect to growth and product formation. Previous studies have reported that during fermentation of glycerol by *C. pasteurianum*, the predominant products formed are butanol and 1,3-propanediol, with only small amounts of acetate, butyrate, and ethanol [17,18,66]. Growth experiments in our laboratory under similar or seemingly identical conditions, however, typically generate only small amounts of 1,3-propanediol, with butyrate and butanol representing the main fermentative products. It is widely held that the ability of *C. pasteurianum* to ferment glycerol is solely due to presence of an active 1,3-propanediol production pathway [66]. Our preliminary data contradict this notion, as some experiments using wild-type *C. pasteurianum* and glycerol as a sole carbon source have generated less than 0.1 g l^{-1} of 1,3-propanediol. Supplementation with vitamin B_{12} ranging from 0.025 to 0.2 μM had no effect on 1,3-propanediol production (data not shown). Therefore, it is expected that our asRNA strains targeting the 1,3-propanediol pathway (*dhaB* and *dhaT*) will have only minor effects on

growth and product distribution of *C. pasteurianum*, whereas ones targeted to *ptb*, *buk*, and *hydA* are expected to lead to significant alterations in the central fermentative metabolism of *C. pasteurianum*.

6.5. Conclusions

In this work, we show that plasmid-based gene overexpression and asRNA-mediated gene downregulation are effective tools for altering expression of target genes in *C. pasteurianum*. We identified the *thl* promoter as a strong, constitutive promoter, which should find use in future gene expression applications. Our devised asRNA system should also lead to extensive metabolic engineering of *C. pasteurianum*.

CHAPTER 7

Original contributions and recommendations

7.1. Original contributions

7.1.1. High-level gene transfer to C. pasteurianum

Efficient electroporation-mediated plasmid transfer to *C. pasteurianum* was demonstrated for the first time. Electrotransformation was shown to be dependent on methylation and protection of CpaAI sites (5'-CGCG-3') within transforming plasmid DNA by the M.FnuDII methyltransferase. Partial disruption of the cell wall and cell membrane via glycine and ethanol, respectively, were shown to enhance electrotransformation. A final electrotransformation efficiency on the order of 10^4 transformants μg^{-1} was achieved, which is among the highest reported in the clostridia. Since our initial report, we have disseminated plasmids, strains, and protocols pertaining to electrotransformation of *C. pasteurianum* to research groups in the U.S., U.K., and Germany.

7.1.2. Draft genome sequence of C. pasteurianum

We sequenced, assembled, and annotated a nearly complete genome sequence for the type strain of *C. pasteurianum*. Although another research group reported the genome sequence of *C. pasteurianum* in parallel to our efforts, our draft is currently comprised of 12 contigs, compared to 37 for the alternate assembly.

7.1.3. Development of a group-II-intron-mediated chromosomal gene knockout system for C. pasteurianum

Established clostridial group II intron gene knockout vectors and procedures were not successful in disrupting chromosomal genes in *C. pasteurianum* due to host restriction and low intron retrohoming efficiency. Using deletion vector constructs and methylome analysis from SMRT sequencing data, we showed that Type I host restriction inhibits electrotransformation of

shuttle vectors harboring the Ll.ltrB IEP gene (*ltrA*). We designated the new Type I RM system CpaAII and propose a putative recognition sequence of 5'-AAGNNNNNCTCC-3'. By mutating the unique CpaAII site within pSY6catP, we resolved the inability to electrotransform shuttle vectors encoding Ll.ltrB intron elements and generated a vector for performing chromosomal gene knockouts in *C. pasteurianum*. Since retrohoming efficiency remained too low to directly isolate intron insertional mutants, we devised an enrichment procedure in which >50% of the resulting colonies possessed the desired intron insertion.

7.1.4. Development of plasmid-based gene overexpression and asRNA-mediated chromosomal gene downregulation methodologies in C. pasteurianum

We applied a well-known clostridial gene reporter system based on β-galactosidase to generate transcriptional fusions with promoters from genes involved in the central fermentative metabolism of *C. pasteurianum*. Of the promoters tested, the promoter from the *C. pasteurianum thl* promoter conferred greatest β-galactosidase activity throughout the course of growth, thus demonstrating successful overexpression of a heterologous gene in *C. pasteurianum*. We applied our strong, constitutive *thl* promoter to drive transcription of a synthesized asRNA molecule targeted to plasmid-encoded *lacZ*. β-Galactosidase activity in the asRNA strain was reduced to approximately 50% of the level of activity exhibited by the control strain, thus demonstrating asRNA-mediated gene downregulation as a viable genetic tool in *C. pasteurianum*. We constructed asRNA constructs for downregulation of 1,3-propanediol (*dhaB* and *dhaT*), butyrate (*buk* and *ptb*), and hydrogen production pathways (*hydA*), which are currently being characterized with respect to growth and product distribution.

7.2. Recommendations

1. The methods developed herein for chromosomal gene knockdown and knockout should find widespread use in *C. pasteurianum* metabolic engineering applications. The 1,3-propanediol pathway represents an interesting target for knockdown and, if permissible, gene knockout, since it is currently unclear if *C. pasteurianum* is able to ferment glycerol as a sole carbon source without production of 1,3-propanediol. If knockout of the 1,3-propanediol pathway proves lethal, knockdown techniques should still permit growth on glycerol. Based on our fermentation data, the butyrate pathway also represents a promising

target for knockout or downregulation. In *C. beijerinckii*, knockout of the *buk* gene encoding butyrate kinase leads to enhanced solvent production [201]. Disurption of the analogous *ack* gene encoding acetate kinase in the acetate pathway has also produced strains with altered fermentative metabolism [259] and it would be interesting to determine what effect such knockouts or knockdowns would have on the glycerol fermentation of *C. pasteurianum*.

2. Plasmid-based gene overexpression has also been investigated in this thesis, yet attempts have not yet been made to utilize this tool to alter the fermentative metabolism of *C. pasteurianum*. As work with *C. acetobutylicum* has demonstrated, homologous overexpression of key aldehyde/alcohol dehydrogenases is a promising strategy for enhancing solvent production [151]. Numerous iron-containing aldehyde/alcohol dehydrogenases have been annotated in our genome sequence of *C. pasteurianum*, including ones with predicted activity on 4C substrates, which could be overexpressed from a plasmid to enhance butanol titers. A similar approach could involve overexpression of genes involved in the central trunk pathway to increase flux from acetyl-CoA to butyryl-CoA. Of these, thiolase, which catalyzes the condensation of two 2C acetyl-CoA into a 4C acetoacetyl-CoA, has been a common target for overexpression [151]. Conversely, heterologous gene expression strategies can be employed to impart new metabolic capabilities to *C. pasteurianum*. Successful heterologous gene expression strategies with other species of *Clostridium* are summarized in Table 2.5 and a few brief approaches are discussed below for *C. pasteurianum*.
3. It is clear that *C. pasteurianum* is naturally an excellent producer of butanol when grown on reduced substrates, such as glycerol. It has been proposed that product toxicitity is likely the factor limiting butanol production. Therefore, overexpression of any genes able to enhance alcohol tolerance could potentially lead to increased butanol production. Homologous overexpression of the *groESL* operon from *C. pasteurianum*, which has been annotated to contig 5 in our genome sequence, or heterologous expression of *groESL* homologs in *C. acetobutylicum* [156] is a promising avenue for further research.
4. A significant obstacle to efficient conversion of crude, biodiesel-derived glycerol to butanol by *C. pasteurianum* appears to be toxicity and growth inhibition resulting from contaminants present in the unpurified glycerol feedstock [18]. A major potential growth-

inhibiting impurity is methanol, which is the most widely used alcohol for production of biodiesel. Alleviation of methanol-mediated growth inhibition of *C. pasteurianum* on biodiesel-derived glycerol could therefore theoretically be achieved by heterologous expression of methylotrophic enzymes from organisms able to ferment methanol. For example, significant methanol, formaldehyde, and formate dehydrogenase activities, which collectively oxidize methanol to carbon dioxide, have been assayed from *Paracoccus denitrificans* grown anaerobically on methanol [260]. Therefore, heterologous expression of *P. denitrificans* genes encoding methanol dehydrogenase or both methanol and formaldehyde dehydrogenase in *C. pasteurianum* could lead to improved growth on crude glycerol via conversion of methanol into formaldehyde or formate, respectively, both of which are produced by wild-type *C. pasteurianum* [261].

5. Finally, recombinogenic engineering (recombineering) of *Clostridium* represents one of the newest and most exciting avenues of clostridial genetic engineering. The chief prerequisite for recombineering is a highly efficient electrotransformation procedure, which we have developed in this thesis. Further, a *recT* gene from *C. perfringens* has been shown to permit recombination of single-stranded oligonucleotides in *C. acetobutylicum* [220]. The *C. perfringens recT* could therefore be expressed in *C. pasteurianum* to assess the utility of oligonucleotide recombination. If successful, this system could then serve as the basis for the development of highly precise genome editing techniques in *C. pasteurianum*.

References

1. Sivakumar G, Vail DR, Xu JF, Burner DM, Lay JO, et al. (2010) Bioethanol and biodiesel: Alternative liquid fuels for future generations. Engineering in Life Sciences 10: 8-18.
2. Jegannathan KR, Chan ES, Ravindra P (2009) Harnessing biofuels: A global Renaissance in energy production? Renewable & Sustainable Energy Reviews 13: 2163-2168.
3. Awang GM, Jones GA, Ingledew WM (1988) The actone-butanol-ethanol fermentation. Crit Rev Microbiol 15: S33-S67.
4. Jones DT, Woods DR (1986) Acetone-butanol fermentation revisited. Microbiological Reviews 50: 484-524.
5. Durre P (2008) Fermentative butanol production - Bulk chemical and biofuel. Incredible Anaerobes: From Physiology to Genomics to Fuels. Oxford: BLACKWELL PUBLISHING. pp. 353-362.
6. Zheng YN, Li LZ, Xian M, Ma YJ, Yang JM, et al. (2009) Problems with the microbial production of butanol. Journal of Industrial Microbiology & Biotechnology 36: 1127-1138.
7. Lee SY, Park JH, Jang SH, Nielsen LK, Kim J, et al. (2008) Fermentative butanol production by clostridia. Biotechnology and Bioengineering 101: 209-228.
8. Papoutsakis ET (2008) Engineering solventogenic clostridia. Current Opinion in Biotechnology 19: 420-429.
9. Lee SY, Bennett GN, Papoutsakis ET (1992) Construction of *Escherichia coli Clostridium acetobutylicum* shuttle vectors and transformation of *Clostridium acetobutylicum* strains. Biotechnology Letters 14: 427-432.
10. Mermelstein LD, Welker NE, Bennett GN, Papoutsakis ET (1992) Expression of cloned homologous fermentative genes in *Clostridium acetobutylicum* ATCC 824. Bio-Technology 10: 190-195.
11. Girbal L, Mortier-Barriere I, Raynaud F, Rouanet C, Croux C, et al. (2003) Development of a sensitive gene expression reporter system and an inducible promoter-repressor system for *Clostridium acetobutylicum*. Applied and Environmental Microbiology 69: 4985-4988.
12. Tummala SB, Welker NE, Papoutsakis ET (1999) Development and characterization of a gene expression reporter system for *Clostridium acetobutylicum* ATCC 824. Applied and Environmental Microbiology 65: 3793-3799.
13. Desai RP, Papoutsakis ET (1999) Antisense RNA strategies for metabolic engineering of *Clostridium acetobutylicum*. Applied and Environmental Microbiology 65: 936-945.
14. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP (2007) The ClosTron: A universal gene knock-out system for the genus *Clostridium*. Journal of Microbiological Methods 70: 452-464.
15. Shao L, Hu S, Yang Y, Gu Y, Chen J, et al. (2007) Targeted gene disruption by use of a group II intron (targetron) vector in *Clostridium acetobutylicum*. Cell Research 17: 963-965.
16. Nolling J, Breton G, Omelchenko MV, Makarova KS, Zeng QD, et al. (2001) Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. Journal of Bacteriology 183: 4823-4838.
17. Dabrock B, Bahl H, Gottschalk G (1992) Parameters affecting solvent production by *Clostridium pasteurianum*. Applied and Environmental Microbiology 58: 1233-1239.
18. Taconi KA, Venkataramanan KP, Johnson DT (2009) Growth and solvent production by *Clostridium pasteurianum* ATCC (R) 6013 (TM) utilizing biodiesel-derived crude glycerol as the sole carbon source. Environmental Progress & Sustainable Energy 28: 100-110.

19. Keasling JD, Chou H (2008) Metabolic engineering delivers next-generation biofuels. Nature Biotechnology 26: 298-299.
20. Dernotte J, Mounaim-Rousselle C, Halter F, Seers P (2010) Evaluation of butanol-gasoline blends in a port fuel-injection, spark-ignition engine. Oil & Gas Science and Technology-Revue De L Institut Francais Du Petrole 65: 345-351.
21. Wiegel J, Tanner R, Rainey FA (2006) An introduction to the family Clostridiaceae. In: Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E, editors. The Prokaryotes. 3 ed. New York: Springer. pp. 654-678.
22. Gottschalk G, Andreesen JR, Hippe H (1981) The genus *Clostridium* (non-medical aspects). In: Starr MP, Stolp H, Truper HC, Balows A, Schlegel HG, editors. The prokaryotes, a handbook on habitats, isolation and identification of bacteria. New York: Springer-Verlag. pp. 1767-1803.
23. Finegold SM, Song YL, Liu CX (2002) Taxonomy - General comments and update on taxonomy of clostridia and anaerobic cocci. Anaerobe 8: 283-285.
24. Gupta RS, Gao B (2009) Phylogenomic analyses of clostridia and identification of novel protein signatures that are specific to the genus *Clostridium sensu stricto* (cluster I). International Journal of Systematic and Evolutionary Microbiology 59: 285-294.
25. Ludwig W, Schleifer KH, Whitman WB (2009) Revised road map to the phylum Firmicutes. In: De Vos P, Garrity GM, Jones D, Krieg NR, Ludwig W et al., editors. Bergey's Manual of Systematic Bacteriology. 2 ed. New York: Springer. pp. 1-8.
26. Rainey FA, Hollen BJ, Small A (2009) Genus I. *Clostridium*. In: De Vos P, Garrity GM, Jones D, Krieg NR, Ludwig W et al., editors. Bergey's Manual of Systematic Bacteriology. 2 ed. New York: Springer. pp. 738-830.
27. Yutin N, Galperin MY (2013) A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. Environmental Microbiology 15: 2631-2641.
28. Lawson PA, Llopperez P, Hutson RA, Hippe H, Collins MD (1993) Towards a phylogeny of the clostridia based on 16S ribosomal-RNA sequences. Fems Microbiology Letters 113: 87-92.
29. Shone CC, Hambleton P (1989) Toxigenic clostridia. In: Minton NP, Clarke DJ, editors. Clostridia. New York: Plenum Press. pp. 265-292.
30. Barash JR, Arnon SS (2013) A novel strain of *Clostridium botulinum* that produces Type B and Type H botulinum toxins. J Infect Dis.
31. McFarland LV, Mulligan ME, Kwok RYY, Stamm WE (1989) Nosocomial acquisition of *Clostridium difficile* infection. New England Journal of Medicine 320: 204-210.
32. Chalmers JD, Al-Khairalla M, Short PM, Fardon TC, Winter JH (2010) Proposed changes to management of lower respiratory tract infections in response to the *Clostridium difficile* epidemic. Journal of Antimicrobial Chemotherapy 65: 608-618.
33. Kelly CP, LaMont JT (2008) *Clostridium difficile*-more difficult than ever. New Engl J Med 359: 1932-1940.
34. Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, et al. (2011) Foodborne illness acquired in the United States-Major pathogens. Emerging Infectious Diseases 17: 7-15.
35. Tracy BP, Jones SW, Fast AG, Indurthi DC, Papoutsakis ET (2012) Clostridia: The importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. Current Opinion in Biotechnology 23: 364-381.

36. Carere CR, Sparling R, Cicek N, Levin DB (2008) Third generation biofuels via direct cellulose fermentation. International Journal of Molecular Sciences 9: 1342-1360.
37. Munasinghe PC, Khanal SK (2010) Biomass-derived syngas fermentation into biofuels: Opportunities and challenges. Bioresource Technology 101: 5013-5022.
38. Jones DT, Woods DR (1989) Solvent production. In: Minton NP, Clarke DJ, editors. Clostridia. New York: Plenum Press. pp. 105-144.
39. Gottschalk G (1986) Bacterial metabolism. New York: Springer-Verlag.
40. Boynton ZL, Bennett GN, Rudolph FB (1996) Cloning, sequencing, and expression of genes encoding phosphotransacetylase and acetate kinase from *Clostridium acetobutylicum* ATCC 824. Applied and Environmental Microbiology 62: 2758-2766.
41. Cary JW, Petersen DJ, Papoutsakis ET, Bennett GN (1988) Cloning and expression of *Clostridium acetobutylicum* phosphotransbutyrylase and butyrate kinase genes in *Escherichia coli*. Journal of Bacteriology 170: 4613-4618.
42. Walter KA, Nair RV, Cary JW, Bennett GN, Papoutsakis ET (1993) Sequence and arrangement of 2 genes of the butyrate-synthesis pathway of *Clostridium acetobutylicum* ATCC 824. Gene 134: 107-111.
43. Freier D, Gottschalk G (1987) L(+)-lactate dehydrogenase of *Clostridium acetobutylicum* is activated by fructose-1,6-bisphosphate. Fems Microbiology Letters 43: 229-233.
44. Bahl H, Gottwald M, Kuhn A, Rale V, Andersch W, et al. (1986) Nutritional factors affecting the ratio of solvents produced by *Clostridium acetobutylicum*. Applied and Environmental Microbiology 52: 169-172.
45. Kim BH, Bellows P, Datta R, Zeikus JG (1984) Control of carbon and electron flow in *Clostridium acetobutylicum* fermentations - Utilization of carbon monoxide to inhibit hydrogen production and to enhance butanol yields. Applied and Environmental Microbiology 48: 764-770.
46. Girbal L, Croux C, Vasconcelos I, Soucaille P (1995) Regulation of metabolic shifts in *Clostridium acetobutylicum* ATCC 824. Fems Microbiology Reviews 17: 287-297.
47. Husemann MHW, Papoutsakis ET (1988) Solventogenesis in *Clostridium acetobutylicum* fermentations related to carboxylic acid and proton concentrations. Biotechnology and Bioengineering 32: 843-852.
48. Monot F, Engasser JM, Petitdemange H (1984) Influence of pH and undissociated butyric acid on the production of acetone and butanol in batch cultures of *Clostridium acetobutylicum*. Applied Microbiology and Biotechnology 19: 422-426.
49. Qureshi N (2009) Solvent production. In: Schaechter M, editor. Encyclopedia of microbiology. Oxford: Elsevier. pp. 512-528.
50. Harris LM, Desai RP, Welker NE, Papoutsakis ET (2000) Characterization of recombinant strains of the *Clostridium acetobutylicum* butyrate kinase inactivation mutant: Need for new phenomenological models for solventogenesis and butanol inhibition? Biotechnology and Bioengineering 67: 1-11.
51. Zhao YS, Tomas CA, Rudolph FB, Papoutsakis ET, Bennett GN (2005) Intracellular butyryl phosphate and acetyl phosphate concentrations in *Clostridium acetobutylicum* and their implications for solvent formation. Applied and Environmental Microbiology 71: 530-537.
52. Grupe H, Gottschalk G (1992) Physiological events in *Clostridium acetobutylicum* during the shift from acidogenesis to solventogenesis in continuous culture and presentation of a model for shift induction. Applied and Environmental Microbiology 58: 3896-3902.

53. Ravagnani A, Jennert KCB, Steiner E, Grunberg R, Jefferies JR, et al. (2000) Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. Molecular Microbiology 37: 1172-1185.
54. Cary JW, Petersen DJ, Papoutsakis ET, Bennett GN (1990) Cloning and expression of *Clostridium acetobutylicum* ATCC 824 acetoacetyl-coenzyme A acetate butyrate coenzyme A transferase in *Escherichia coli*. Applied and Environmental Microbiology 56: 1576-1583.
55. Fischer RJ, Helms J, Durre P (1993) Cloning, sequencing, and molecular analysis of the *sol* operon of *Clostridium acetobutylicum*, a chromosomal locus Involved in solventogenesis. Journal of Bacteriology 175: 6959-6969.
56. Fontaine L, Meynial-Salles I, Girbal L, Yang XH, Croux C, et al. (2002) Molecular characterization and transcriptional analysis of *adhE2*, the gene encoding the NADH-dependent aldehyde/alcohol dehydrogenase responsible for butanol production in alcohologenic cultures of *Clostridium acetobutylicum* ATCC 824. Journal of Bacteriology 184: 821-830.
57. Nair RV, Bennett GN, Papoutsakis ET (1994) Molecular characterization of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. Journal of Bacteriology 176: 871-885.
58. Welch RW, Rudolph FB, Papoutsakis ET (1989) Purification and characterization of the NADH-dependent butanol dehydrogenase from *Clostridium acetobutylicum* (ATCC 824). Archives of Biochemistry and Biophysics 273: 309-318.
59. Sauer U, Durre P (1995) Differential induction of genes related to solvent formation during the shift from acidogenesis to solventogenesis in continuous culture of *Clostridium acetobutylicum*. Fems Microbiology Letters 125: 115-120.
60. Gerischer U, Durre P (1990) Cloning, sequencing, and molecular analysis of the acetoacetate decarboxylase gene region from *Clostridium acetobutylicum*. Journal of Bacteriology 172: 6907-6918.
61. George HA, Johnson JL, Moore WEC, Holdeman LV, Chen JS (1983) Acetone, isopropanol, and butanol production by *Clostridium beijerinckii* (syn *Clostridium butylicum*) and *Clostridium aurantibutyricum*. Applied and Environmental Microbiology 45: 1160-1163.
62. Bahadur K, Saroj KK (1960) A study of the influence of hydrogen ion concentration of the culture media on the formation of acetone and butanol by *Clostridium pasteurianum*, *Clostridium butylicus* and *Clostridium acetobutylicum* using sucrose as substrate in a synthetic media. Jpn J Microbiol 4: 341-349.
63. Nakas JP, Schaedle M, Parkinson CM, Coonley CE, Tanenbaum SW (1983) System development for linked fermentation production of solvents from algal biomass. Applied and Environmental Microbiology 46: 1017-1023.
64. Heyndrickx M, Devos P, Vancanneyt M, Deley J (1991) The fermentation of glycerol by *Clostridium butyricum* LMG $1212t_2$ and LMG $1213t_1$ and *C. pasteurianum* LMG 3285. Applied Microbiology and Biotechnology 34: 637-642.
65. Heyndrickx M, Devos P, Deley J (1991) Fermentation characteristics of *Clostridium pasteurianum* LMG 3285 grown on glucose and mannitol. Journal of Applied Bacteriology 70: 52-58.
66. Biebl H (2001) Fermentation of glycerol by *Clostridium pasteurianum* - Batch and continuous culture studies. Journal of Industrial Microbiology & Biotechnology 27: 18-26.

67. Johnson DT, Taconi KA (2007) The glycerin glut: Options for the value-added conversion of crude glycerol resulting from biodiesel production. Environmental Progress 26: 338-348.
68. Yazdani SS, Gonzalez R (2007) Anaerobic fermentation of glycerol: A path to economic viability for the biofuels industry. Current Opinion in Biotechnology 18: 213-219.
69. Harris J, Mulder R, Kell DB, Walter RP, Morris JG (1986) Solvent production by *Clostridium pasteurianum* in media of high sugar content. Biotechnology Letters 8: 889-892.
70. Bahl H, Andersch W, Gottschalk G (1982) Continuous production of acetone and butanol by *Clostridium acetobutylicum* in a 2-stage phosphate limited chemostat. European Journal of Applied Microbiology and Biotechnology 15: 201-205.
71. Meyer CL, McLaughlin JK, Papoutsakis ET (1985) The effect of CO on growth and product formation in batch cultures of *Clostridium acetobutylicum*. Biotechnology Letters 7: 37-42.
72. Datta R, Zeikus JG (1985) Modulation of acetone-butanol-ethanol fermentation by carbon monoxide and organic acids. Applied and Environmental Microbiology 49: 522-529.
73. Adams MWW, Mortenson LE (1984) The physical and catalytic properties of hydrogenase II of *Clostridium pasteurianum* - A comparison with hydrogenase I. Journal of Biological Chemistry 259: 7045-7055.
74. Erbes DL, Burris RH (1978) Kinetics of methyl viologen oxidation and reduction by hydrogenase from *Clostridium pasteurianum*. Biochimica Et Biophysica Acta 525: 45-54.
75. Erbes DL, Burris RH, Ormejohnson WH (1975) Iron-sulfur cluster in hydrogenase from *Clostridium pasteurianum* W5. Proceedings of the National Academy of Sciences of the United States of America 72: 4795-4799.
76. Mermelstein LD, Papoutsakis ET (1993) In vivo methylation in *Escherichia coli* by the *Bacillus subtilis* phage phi 3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. Applied and Environmental Microbiology 59: 1077-1081.
77. Tummala SB, Welker NE, Papoutsakis ET (2003) Design of antisense RNA constructs for downregulation of the acetone formation pathway of *Clostridium acetobutylicum*. Journal of Bacteriology 185: 1923-1934.
78. Al-Hinai MA, Fast AG, Papoutsakis ET (2012) Novel system for efficient isolation of *Clostridium* double-crossover allelic exchange mutants enabling markerless chromosomal gene deletions and DNA integration. Applied and Environmental Microbiology 78: 8112-8121.
79. Heap JT, Ehsaan M, Cooksley CM, Ng YK, Cartman ST, et al. (2012) Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. Nucleic Acids Research 40.
80. Warnick TA, Methe BA, Leschine SB (2002) *Clostridium phytofermentans* sp nov., a cellulolytic mesophile from forest soil. International Journal of Systematic and Evolutionary Microbiology 52: 1155-1160.
81. Tolonen AC, Petit E, Blanchard JL, Warnick T, Leschine SB (2013) Technologies to study plant biomass fermentation using the model bacterium *Clostridium phytofermentans*. In: Sun J, Ding SY, Peterson JD, editors. Biological conversion of biomass for fuels and chemicals. Cambridge: Royal Society of Chemistry. pp. 114-139.
82. Gheshlaghi R, Scharer JM, Moo-Young M, Chou CP (2009) Metabolic pathways of clostridia for producing butanol. Biotechnology Advances 27: 764-781.

83. Lutke-Eversloh T, Bahl H (2011) Metabolic engineering of *Clostridium acetobutylicum*: recent advances to improve butanol production. Current Opinion in Biotechnology 22: 634-647.
84. Minton NP (2003) Clostridia in cancer therapy. Nature Reviews Microbiology 1: 237-242.
85. Van Mellaert L, Barbe S, Anne J (2006) *Clostridium* spores as anti-tumour agents. Trends in Microbiology 14: 190-196.
86. Wilkinson SR, Young M (1993) Wide diversity of genome size among different strains of *Clostridium acetobutylicum*. Journal of General Microbiology 139: 1069-1076.
87. Skarin H, Hafstrom T, Westerberg J, Segerman B (2011) *Clostridium botulinum* group III: a group with dual identity shaped by plasmids, phages and mobile elements. Bmc Genomics 12.
88. Iacobino A, Scalfaro C, Franciosa G (2013) Structure and genetic content of the megaplasmids of neurotoxigenic *Clostridium butyricum* Type E strains from Italy. Plos One 8.
89. Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, et al. (2010) Direct detection of DNA methylation during single-molecule, real-time sequencing. Nature Methods 7: 461-U472.
90. Roberts RJ, Vincze T, Posfai J, Macelis D (2010) REBASE-a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Research 38: D234-D236.
91. Pyne ME, Moo-Young M, Chung DA, Chou CP (2013) Development of an electrotransformation protocol for genetic manipulation of *Clostridium pasteurianum*. Biotechnology for Biofuels 6.
92. Allen SP, Blaschek HP (1988) Electroporation-induced transformation of intact cells of *Clostridium perfringens*. Applied and Environmental Microbiology 54: 2322-2324.
93. Oultram JD, Loughlin M, Swinfield TJ, Brehm JK, Thompson DE, et al. (1988) Introduction of plasmids into whole cells of *Clostridium acetobutylicum* by electroporation. Fems Microbiology Letters 56: 83-88.
94. Gonzalez-Pajuelo M, Meynial-Salles I, Mendes F, Soucaille P, Vasconcelos I (2006) Microbial conversion of glycerol to 1,3-propanediol: Physiological comparison of a natural producer, *Clostridium butyricum* VPI 3266, and an engineered strain, *Clostridium acetobutylicum* DG1(pSPD5). Applied and Environmental Microbiology 72: 96-101.
95. Aune TEV, Aachmann FL (2010) Methodologies to increase the transformation efficiencies and the range of bacteria that can be transformed. Applied Microbiology and Biotechnology 85: 1301-1313.
96. Jennert KCB, Tardif C, Young DI, Young M (2000) Gene transfer to *Clostridium cellulolyticum* ATCC 35319. Microbiology 146: 3071-3080.
97. Lyras D, Rood JI (1998) Conjugative transfer of RP4-*oriT* shuttle vectors from *Escherichia coli* to *Clostridium perfringens*. Plasmid 39: 160-164.
98. Purdy D, O'Keeffe TAT, Elmore M, Herbert M, McLeod A, et al. (2002) Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. Molecular Microbiology 46: 439-452.
99. Tolonen AC, Chilaka AC, Church GM (2009) Targeted gene inactivation in *Clostridium phytofermentans* shows that cellulose degradation requires the family 9 hydrolase Cphy3367. Molecular Microbiology 74: 1300-1313.
100. Williams DR, Young DI, Young M (1990) Conjugative plasmid transfer from *Escherichia coli* to *Clostridium acetobutylicum*. Journal of General Microbiology 136: 819-826.

101. Harris LM, Welker NE, Papoutsakis ET (2002) Northern, morphological, and fermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. Journal of Bacteriology 184: 3586-3597.
102. Nakotte S, Schaffer S, Bohringer M, Durre P (1998) Electroporation of, plasmid isolation from and plasmid conservation in *Clostridium acetobutylicum* DSM 792. Applied Microbiology and Biotechnology 50: 564-567.
103. Birrer GA, Chesbro WR, Zsigray RM (1994) Electrotransformation of *Clostridium beijerinckii* NRRL B-592 with shuttle plasmid pHR106 and recombinant derivatives. Applied Microbiology and Biotechnology 41: 32-38.
104. Zhou YT, Johnson EA (1993) Genetic transformation of *Clostridium botulinum* Hall A by electroporation. Biotechnology Letters 15: 121-126.
105. Davis TO, Henderson I, Brehm JK, Minton NP (2000) Development of a transformation and gene reporter system for group II, non-proteolytic *Clostridium botulinum* type B strains. Journal of Molecular Microbiology and Biotechnology 2: 59-69.
106. Ackermann G, Tang YJ, Henderson JP, Rodloff AC, Silva J, et al. (2001) Electroporation of DNA sequences from the pathogenicity locus (PaLoc) of toxigenic *Clostridium difficile* into a non-toxigenic strain. Molecular and Cellular Probes 15: 301-306.
107. Leang C, Ueki T, Nevin KP, Lovley DR (2013) A genetic system for *Clostridium ljungdahlii*: A chassis for autotrophic production of biocommodities and a model homoacetogen. Applied and Environmental Microbiology 79: 1102-1109.
108. Sakka K, Kawase M, Baba D, Morimoto K, Karita S, et al. (2003) Electrotransformation of *Clostridium paraputrificum* M-21 with some plasmids. Journal of Bioscience and Bioengineering 96: 304-306.
109. Allen SP, Blaschek HP (1990) Factors involved in the electroporation-induced transformation of *Clostridium perfringens*. Fems Microbiology Letters 70: 217-220.
110. Kim AY, Blaschek HP (1989) Construction of an *Escherichia coli-Clostridium perfringens* shuttle vector and plasmid transformation of *Clostridium perfringens* Applied and Environmental Microbiology 55: 360-365.
111. Nakayama S, Irie R, Kosaka T, Matsuura K, Yoshino S, et al. (2007) New host-vector system in solvent-producing *Clostridium saccharoperbutylacetonicum* strain N1-4. Journal of General and Applied Microbiology 53: 53-56.
112. Liu SC, Minton NP, Giaccia AJ, Brown JM (2002) Anticancer efficacy of systemically delivered anaerobic bacteria as gene therapy vectors targeting tumor hypoxia/necrosis. Gene Therapy 9: 291-296.
113. Marvaud JC, Eisel U, Binz T, Niemann H, Popoff MR (1998) TetR is a positive regulator of the tetanus toxin gene in *Clostridium tetani* and is homologous to BotR. Infection and Immunity 66: 5698-5702.
114. Tyurin MV, Desai SG, Lynd LR (2004) Electrotransformation of *Clostridium thermocellum*. Applied and Environmental Microbiology 70: 883-890.
115. Klapatch TR, Guerinot ML, Lynd LR (1996) Electrotransformation of *Clostridium thermosaccharolyticum*. Journal of Industrial Microbiology 16: 342-347.
116. Zhu Y, Liu XG, Yang ST (2005) Construction and characterization of *pta* gene-deleted mutant of *Clostridium tyrobutyricum* for enhanced butyric acid fermentation. Biotechnology and Bioengineering 90: 154-166.

117. Jo JH, Jeon CO, Lee SY, Lee DS, Park JM (2010) Molecular characterization and homologous overexpression of FeFe -hydrogenase in *Clostridium tyrobutyricum* JM1. International Journal of Hydrogen Energy 35: 1065-1073.
118. Saunders JR, Docherty A, Humphreys GO (1984) Transformation of bacteria by plasmid DNA. Methods Microbiol 17: 61-95.
119. Richards DF, Linnett PE, Oultram JD, Young M (1988) Restriction endonucleases in *Clostridium pasteurianum* ATCC 6013 and *C. thermohydrosulfuricum* DSM 568. Journal of General Microbiology 134: 3151-3157.
120. Klapatch TR, Demain AL, Lynd LR (1996) Restriction endonuclease activity in *Clostridium thermocellum* and *Clostridium thermosaccharolyticum*. Applied Microbiology and Biotechnology 45: 127-131.
121. Lin YL, Blaschek HP (1984) Transformation of heat-treated *Clostridium acetobutylicum* protoplasts with pUB110 plasmid DNA. Applied and Environmental Microbiology 48: 737-742.
122. Lofblom J, Kronqvist N, Uhlen M, Stahl S, Wernerus H (2007) Optimization of electroporation-mediated transformation: *Staphylococcus carnosus* as model organism. Journal of Applied Microbiology 102: 736-747.
123. Raleigh EA, Wilson G (1986) *Escherichia coli* K-12 restricts DNA containing 5-methylcytosine. Proceedings of the National Academy of Sciences of the United States of America 83: 9070-9074.
124. Kennedy CL, Krejany EO, Young LF, O'Connor JR, Awad MM, et al. (2005) The alpha-toxin of *Clostridium septicum* is essential for virulence. Molecular Microbiology 57: 1357-1366.
125. Kopke M, Held C, Hujer S, Liesegang H, Wiezer A, et al. (2010) *Clostridium ljungdahlii* represents a microbial production platform based on syngas. Proceedings of the National Academy of Sciences of the United States of America 107: 13087-13092.
126. Walkinshaw MD, Taylor P, Sturrock SS, Atanasiu C, Berge T, et al. (2002) Structure of Ocr from bacteriophage T7, a protein that mimics B-form DNA. Molecular Cell 9: 187-194.
127. Xu SY, Corvaglia AR, Chan SH, Zheng Y, Linder P (2011) A type IV modification-dependent restriction enzyme SauUSI from *Staphylococcus aureus* subsp. *aureus* USA300. Nucleic Acids Research 39: 5597-5610.
128. Cui GZ, Hong W, Zhang J, Li WL, Feng YG, et al. (2012) Targeted gene engineering in *Clostridium cellulolyticum* H10 without methylation. Journal of Microbiological Methods 89: 201-208.
129. Hammes W, Schleife.Kh, Kandler O (1973) Mode of action of glycine on biosynthesis of peptidoglycan. Journal of Bacteriology 116: 1029-1053.
130. Chassy BM (1976) Gentle method for lysis of oral streptococci. Biochemical and Biophysical Research Communications 68: 603-608.
131. Hermans J, Boschloo JG, Debont JAM (1990) Transformation of *Mycobacterium aurum* by electroporation - the use of glycine, lysozyme and isonicotinic acid hydrazide in enhancing transformation efficiency. Fems Microbiology Letters 72: 221-224.
132. Blumberg PM, Stroming Jl (1974) Interaction of penicillin with the bacterial cell: Pencillin-binding proteins and penicillin-sensitive enzymes. Bacteriol Rev 38: 291-335.
133. Reysset G (1993) Transformation and electrotransformation in clostridia. In: Sebald M, editor. Genetics and molecular biology of anaerobic bacteria. New York: Springer-Verlag. pp. 111-119.

134. Scott PT, Rood JI (1989) Electroporation-mediated transformation of lysostaphin-treated *Clostridium perfringens*. Gene 82: 327-333.
135. Sharma AD, Singh J, Gill PK (2007) Ethanol mediated enhancement in bacterial transformation. Electronic Journal of Biotechnology 10: 166-168.
136. Assad-Garcia JS, Bonnin-Jusserand M, Garmyn D, Guzzo J, Alexandre H, et al. (2008) An improved protocol for electroporation of *Oenococcus oeni* ATCC BAA-1163 using ethanol as immediate membrane fluidizing agent. Letters in Applied Microbiology 47: 333-338.
137. Zhang HH, Li YT, Chen XM, Sheng HM, An LZ (2011) Optimization of electroporation conditions for *Arthrobacter* with plasmid pART2. Journal of Microbiological Methods 84: 114-120.
138. Tyurin M, Padda R, Huang KX, Wardwell S, Caprette D, et al. (2000) Electrotransformation of *Clostridium acetobutylicum* ATCC 824 using high-voltage radio frequency modulated square pulses. Journal of Applied Microbiology 88: 220-227.
139. Heap JT, Pennington OJ, Cartman ST, Minton NP (2009) A modular system for *Clostridium* shuttle plasmids. Journal of Microbiological Methods 78: 79-85.
140. Yu MR, Du YM, Jiang WY, Chang WL, Yang ST, et al. (2012) Effects of different replicons in conjugative plasmids on transformation efficiency, plasmid stability, gene expression and *n*-butanol biosynthesis in *Clostridium tyrobutyricum*. Applied Microbiology and Biotechnology 93: 881-889.
141. Matsushita C, Matsushita O, Koyama M, Okabe A (1994) A *Clostridium perfringens* vector for the selection of promoters. Plasmid 31: 317-319.
142. Quixley KWM, Reid SJ (2000) Construction of a reporter gene vector for *Clostridium beijerinckii* using a *Clostridium* endoglucanase gene. Journal of Molecular Microbiology and Biotechnology 2: 53-57.
143. Phillips-Jones MK (2000) Use of a *lux* reporter system for monitoring rapid changes in alpha-toxin gene expression in *Clostridium perfringens* during growth. Fems Microbiology Letters 188: 29-33.
144. Yu MR, Zhang YL, Tang IC, Yang ST (2011) Metabolic engineering of *Clostridium tyrobutyricum* for *n*-butanol production. Metabolic Engineering 13: 373-382.
145. Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, et al. (2010) The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. Journal of Microbiological Methods 80: 49-55.
146. Nakayama SI, Kosaka T, Hirakawa H, Matsuura K, Yoshino S, et al. (2008) Metabolic engineering for solvent productivity by downregulation of the hydrogenase gene cluster *hupCBA* in *Clostridium saccharoperbutylacetonicum* strain N1-4. Applied Microbiology and Biotechnology 78: 483-493.
147. Nuyts S, Theys J, Landuyt W, Van Mellaert L, Lambin P, et al. (2001) Increasing specificity of anti-tumor therapy: Cytotoxic protein delivery by non-pathogenic clostridia under regulation of radio-induced promoters. Anticancer Research 21: 857-861.
148. Nuyts S, Van Mellaert L, Theys J, Landuyt W, Bosmans E, et al. (2001) Radio-responsive *recA* promoter significantly increases TNF alpha production in recombinant clostridia after 2 Gy irradiation. Gene Therapy 8: 1197-1201.
149. Dong HJ, Tao WW, Zhang YP, Li Y (2012) Development of an anhydrotetracycline-inducible gene expression system for solvent-producing *Clostridium acetobutylicum*: A useful tool for strain engineering. Metabolic Engineering 14: 59-67.

150. Hartman AH, Liu HL, Melville SB (2011) Construction and characterization of a lactose-inducible promoter system for controlled gene expression in *Clostridium perfringens*. Applied and Environmental Microbiology 77: 471-478.
151. Sillers R, A-Hinai MA, Papoutsakis ET (2009) Aldehyde-alcohol dehydrogenase and/or thiolase overexpression coupled with CoA transferase downregulation lead to higher alcohol titers and selectivity in *Clostridium acetobutylicum* fermentations. Biotechnology and Bioengineering 102: 38-49.
152. Sillers R, Chow A, Tracy B, Papoutsakis ET (2008) Metabolic engineering of the non-sporulating, non-solventogenic *Clostridium acetobutylicum* strain M5 to produce butanol without acetone demonstrate the robustness of the acid-formation pathways and the importance of the electron balance. Metabolic Engineering 10: 321-332.
153. Mermelstein LD, Papoutsakis ET, Petersen DJ, Bennett GN (1993) Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for increased solvent production by enhancement of acetone formation enzyme activities using a synthetic acetone operon Biotechnology and Bioengineering 42: 1053-1060.
154. Lee JY, Jang Y-S, Lee J, Papoutsakis ET, Lee SY (2009) Metabolic engineering of *Clostridium acetobutylicum* M5 for highly selective butanol production. Biotechnology journal 4: 1432-1440.
155. Alsaker KV, Spitzer TR, Papoutsakis ET (2004) Transcriptional analysis of *spo0A* overexpression in *Clostridium acetobutylicum* and its effect on the cell's response to butanol stress. Journal of Bacteriology 186: 1959-1971.
156. Tomas CA, Welker NE, Papoutsakis ET (2003) Overexpression of *groESL* in *Clostridium acetobutylicum* results in increased solvent production and tolerance, prolonged metabolism, and changes in the cell's transcriptional program. Applied and Environmental Microbiology 69: 4951-4965.
157. Marvaud JC, Gibert M, Inoue K, Fujinaga Y, Oguma K, et al. (1998) botR/A is a positive regulator of botulinum neurotoxin and associated non-toxin protein genes in *Clostridium botulinum* A. Molecular Microbiology 29: 1009-1018.
158. Girbal L, von Abendroth G, Winkler M, Benton PMC, Meynial-Salles I, et al. (2005) Homologous and heterologous overexpression in *Clostridium acetobutylicum* and characterization of purified clostridial and algal Fe-only hydrogenases with high specific activities. Applied and Environmental Microbiology 71: 2777-2781.
159. Klein M, Ansorge-Schumacher MB, Fritsch M, Hartmeier W (2010) Influence of hydrogenase overexpression on hydrogen production of *Clostridium acetobutylicum* DSM 792. Enzyme and Microbial Technology 46: 384-390.
160. Nair RV, Papoutsakis ET (1994) Expression of plasmid-encoded *aad* in *Clostridium acetobutylicum* M5 restores vigorous butanol production Journal of Bacteriology 176: 5843-5846.
161. Morimoto K, Kimura T, Sakka K, Ohmiya K (2005) Overexpression of a hydrogenase gene in *Clostridium paraputrificum* to enhance hydrogen gas production. Fems Microbiology Letters 246: 229-234.
162. Chanal A, Mingardon F, Bauzan M, Tardif C, Fierobe HP (2011) Scaffoldin modules serving as "cargo" domains to promote the secretion of heterologous cellulosomal cellulases by *Clostridium acetobutylicum*. Applied and Environmental Microbiology 77: 6277-6280.

163. Kovacs K, Willson BJ, Schwarz K, Heap JT, Jackson A, et al. (2013) Secretion and assembly of functional mini-cellulosomes from synthetic chromosomal operons in *Clostridium acetobutylicum* ATCC 824. Biotechnology for Biofuels 6.
164. Mingardon F, Chanal A, Tardif C, Fierobe HP (2011) The issue of secretion in heterologous expression of *Clostridium cellulolyticum* cellulase-encoding genes in *Clostridium acetobutylicum* ATCC 824. Applied and Environmental Microbiology 77: 2831-2838.
165. Mingardon F, Perret S, Belaich A, Tardif C, Belaich JP, et al. (2005) Heterologous production, assembly, and secretion of a minicellulosome by *Clostridium acetobutylicum* ATCC 824. Applied and Environmental Microbiology 71: 1215-1222.
166. Perret S, Casalot L, Fierobe HP, Tardif C, Sabathe F, et al. (2004) Production of heterologous and chimeric scaffoldins by *Clostridium acetobutylicum* ATCC 824. Journal of Bacteriology 186: 253-257.
167. Gonzalez-Pajuelo M, Meynial-Salles I, Mendes F, Andrade JC, Vasconcelos I, et al. (2005) Metabolic engineering of *Clostridium acetobutylicum* for the industrial production of 1,3-propanediol from glycerol. Metabolic Engineering 7: 329-336.
168. Gu Y, Li J, Zhang L, Chen J, Niu L, et al. (2009) Improvement of xylose utilization in *Clostridium acetobutylicum* via expression of the *talA* gene encoding transaldolase from *Escherichia coli*. J Biotechnol 143: 284-287.
169. Siemerink MAJ, Kuit W, Contreras AML, Eggink G, van der Oost J, et al. (2011) d-2,3-Butanediol production due to heterologous expression of an acetoin reductase in *Clostridium acetobutylicum*. Applied and Environmental Microbiology 77: 2582-2588.
170. Lee J, Jang YS, Choi SJ, Im JA, Song H, et al. (2012) Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for isopropanol-butanol-ethanol fermentation. Applied and Environmental Microbiology 78: 1416-1423.
171. Guedon E, Desvaux M, Petitdemange H (2002) Improvement of cellulolytic properties of *Clostridium cellulolyticum* by metabolic engineering. Applied and Environmental Microbiology 68: 53-58.
172. Higashide W, Li YC, Yang YF, Liao JC (2011) Metabolic engineering of *Clostridium cellulolyticum* for production of isobutanol from cellulose. Applied and Environmental Microbiology 77: 2727-2733.
173. Mani N, Dupuy B (2001) Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. Proceedings of the National Academy of Sciences of the United States of America 98: 5844-5849.
174. Mani N, Lyras D, Barroso L, Howarth P, Wilkins T, et al. (2002) Environmental response and autoregulation of *Clostridium difficile* TxeR, a sigma factor for toxin gene expression. Journal of Bacteriology 184: 5971-5978.
175. Matamouros S, England P, Dupuy B (2007) *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. Molecular Microbiology 64: 1274-1288.
176. Musto H, Romero H, Zavala A (2003) Translational selection is operative for synonymous codon usage in *Clostridium perfringens* and *Clostridium acetobutylicum*. Microbiology 149: 855-863.
177. Sharp PM, Bailes E, Grocock RJ, Peden JF, Sockett RE (2005) Variation in the strength of selected codon usage bias among bacteria. Nucleic Acids Research 33: 1141-1153.
178. Gouy M, Gautier C (1982) Codon usage in bacteria - correlation with gene expressivity. Nucleic Acids Research 10: 7055-7074.

179. Dornenburg JE, DeVita AM, Palumbo MJ, Wade JT (2010) Widespread antisense transcription in *Escherichia coli*. Mbio 1: 1-4.
180. Lee JM, Zhang SH, Saha S, Anna SS, Jiang C, et al. (2001) RNA expression analysis using an antisense *Bacillus subtilis* genome array. Journal of Bacteriology 183: 7371-7380.
181. Georg J, Hess WR (2011) *cis*-Antisense RNA, another level of gene regulation in bacteria. Microbiology and Molecular Biology Reviews 75: 286-300.
182. Wagner EGH, Simons RW (1994) Antisense RNA control in bacteria, phages, and plasmids. Annual Review of Microbiology 48: 713-742.
183. Tummala SB, Junne SG, Papoutsakis ET (2003) Antisense RNA downregulation of coenzyme A transferase combined with alcohol-aldehyde dehydrogenase overexpression leads to predominantly alcohologenic *Clostridium acetobutylicum* fermentations. Journal of Bacteriology 185: 3644-3653.
184. Liyanage H, Young M, Kashket ER (2000) Butanol tolerance of *Clostridium beijerinckii* NCIMB 8052 associated with down-regulation of *gldA* by antisense RNA. Journal of Molecular Microbiology and Biotechnology 2: 87-93.
185. Perret S, Maamar H, Belaich JP, Tardif C (2004) Use of antisense RNA to modify the composition of cellulosomes produced by *Clostridium cellulolyticum*. Molecular Microbiology 51: 599-607.
186. Raju D, Setlow P, Sarker MR (2007) Antisense-RNA-mediated decreased synthesis of small, acid-soluble spore proteins leads to decreased resistance of *Clostridium perfringens* spores to moist heat and UV radiation. Applied and Environmental Microbiology 73: 2048-2053.
187. Patzel V, Sczakiel G (1998) Theoretical design of antisense RNA structures substantially improves annealing kinetics and efficacy in human cells. Nature Biotechnology 16: 64-68.
188. Venkataramanan KP, Jones SW, McCormick KP, Kunjeti SG, Ralston MT, et al. (2013) The *Clostridium* small RNome that responds to stress: the paradigm and importance of toxic metabolite stress in *C. acetobutylicum*. Bmc Genomics 14: 849.
189. Mohr G, Smith D, Belfort M, Lambowitz AM (2000) Rules for DNA target-site recognition by a lactococcal group II intron enable retargeting of the intron to specific DNA sequences. Genes & Development 14: 559-573.
190. Guo HT, Karberg M, Long M, Jones JP, Sullenger B, et al. (2000) Group II introns designed to insert into therapeutically relevant DNA target sites in human cells. Science 289: 452-457.
191. Karberg M, Guo HT, Zhong J, Coon R, Perutka J, et al. (2001) Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. Nature Biotechnology 19: 1162-1167.
192. Matsuura M, Noah JW, Lambowitz AM (2001) Mechanism of maturase-promoted group II intron splicing. Embo Journal 20: 7259-7270.
193. Matsuura M, Saldanha R, Ma HW, Wank H, Yang J, et al. (1997) A bacterial group II intron encoding reverse transcriptase, maturase, DNA endonuclease activities: Biochemical demonstration of maturase activity and insertion of new genetic information within the intron. Genes & Development 11: 2910-2924.
194. Zhong J, Karberg M, Lambowitz AM (2003) Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. Nucleic Acids Research 31: 1656-1664.

195. Perutka J, Wang WJ, Goerlitz D, Lambowitz AM (2004) Use of computer-designed group II introns to disrupt *Escherichia coli* DExH/D-box protein and DNA helicase genes. Journal of Molecular Biology 336: 421-439.
196. Cai GQ, Jin B, Saint C, Monis P (2011) Genetic manipulation of butyrate formation pathways in *Clostridium butyricum*. J Biotechnol 155: 269-274.
197. Li YC, Tschaplinski TJ, Engle NL, Hamilton CY, Rodriguez M, et al. (2012) Combined inactivation of the *Clostridium cellulolyticum* lactate and malate dehydrogenase genes substantially increases ethanol yield from cellulose and switchgrass fermentations. Biotechnology for Biofuels 5.
198. Mohr G, Hong W, Zhang J, Cui GZ, Yang YF, et al. (2013) A targetron system for gene targeting in thermophiles and its application in *Clostridium thermocellum*. Plos One 8.
199. Zhuang FL, Karberg M, Perutka J, Lambowitz AM (2009) EcI5, a group IIB intron with high retrohoming frequency: DNA target site recognition and use in gene targeting. RNA 15: 432-449.
200. Xiao H, Li ZL, Jiang Y, Yang YL, Jiang WH, et al. (2012) Metabolic engineering of d-xylose pathway in *Clostridium beijerinckii* to optimize solvent production from xylose mother liquid. Metabolic Engineering 14: 569-578.
201. Wang Y, Li XZ, Milne CB, Janssen H, Lin WY, et al. (2013) Development of a gene knockout system using mobile group II introns (Targetron) and genetic disruption of acid production pathways in *Clostridium beijerinckii*. Applied and Environmental Microbiology 79: 5853-5863.
202. Cai GQ, Jin B, Monis P, Saint C (2013) A genetic and metabolic approach to redirection of biochemical pathways of *Clostridium butyricum* for enhancing hydrogen production. Biotechnology and Bioengineering 110: 338-342.
203. Kuehne SA, Cartman ST, Heap JT, Kelly ML, Cockayne A, et al. (2010) The role of toxin A and toxin B in *Clostridium difficile* infection. Nature 467: 711-713.
204. Chen Y, McClane BA, Fisher DJ, Rood JI, Gupta P (2005) Construction of an alpha toxin gene knockout mutant of *Clostridium perfringens* type A by use of a mobile group II intron. Applied and Environmental Microbiology 71: 7542-7547.
205. Carter GP, Awad MM, Hao YB, Thelen T, Bergin IL, et al. (2011) TcsL ss an essential virulence factor in *Clostridium sordellii* ATCC 9714. Infection and Immunity 79: 1025-1032.
206. Lehmann K, Schmidt U (2003) Group II introns: Structure and catalytic versatility of large natural ribozymes. Critical Reviews in Biochemistry and Molecular Biology 38: 249-303.
207. Cooksley CM, Zhang Y, Wang HZ, Redl S, Winzer K, et al. (2012) Targeted mutagenesis of the *Clostridium acetobutylicum* acetone-butanol-ethanol fermentation pathway. Metabolic Engineering 14: 630-641.
208. Brown DP, Ganovaraeva L, Green BD, Wilkinson SR, Young M, et al. (1994) Characterization of *spo0A* homologs in diverse *Bacillus* and *Clostridium* species identifies a probable DNA-binding domain. Molecular Microbiology 14: 411-426.
209. Green EM, Boynton ZL, Harris LM, Rudolph FB, Papoutsakis ET, et al. (1996) Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. Microbiology 142: 2079-2086.
210. Coleman JP, Hudson LL (1995) Cloning and characterization of a conjugated bile-acid hydrolase gene from *Clostridium perfringens*. Applied and Environmental Microbiology 61: 2514-2520.

211. Nair RV, Green EM, Watson DE, Bennett GN, Papoutsakis ET (1999) Regulation of the *sol* locus genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 by a putative transcriptional repressor. Journal of Bacteriology 181: 319-330.
212. Cartman ST, Kelly ML, Heeg D, Heap JT, Minton NP (2012) Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between the *tcdC* genotype and toxin production. Applied and Environmental Microbiology 78: 4683-4690.
213. Argyros DA, Tripathi SA, Barrett TF, Rogers SR, Feinberg LF, et al. (2011) High ethanol titers from cellulose by using metabolically engineered thermophilic, anaerobic microbes. Applied and Environmental Microbiology 77: 8288-8294.
214. Nariya H, Miyata S, Suzuki M, Tamai E, Okabe A (2011) Development and application of a method for counterselectable in-frame deletion in *Clostridium perfringens*. Applied and Environmental Microbiology 77: 1375-1382.
215. Tripathi SA, Olson DG, Argyros DA, Miller BB, Barrett TF, et al. (2010) Development of *pyrF*-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a *pta* mutant. Applied and Environmental Microbiology 76: 6591-6599.
216. Zhang YM, Buchholz F, Muyrers JPP, Stewart AF (1998) A new logic for DNA engineering using recombination in *Escherichia coli*. Nature Genetics 20: 123-128.
217. Murphy KC, Campellone KG (2003) Lambda Red-mediated recombinogenic engineering of enterohemorrhagic and enteropathogenic *E. coli*. Bmc Molecular Biology 4.
218. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of Sciences of the United States of America 97: 6640-6645.
219. Ellis HM, Yu DG, DiTizio T, Court DL (2001) High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. Proceedings of the National Academy of Sciences of the United States of America 98: 6742-6746.
220. Dong HJ, Tao WW, Gong F, Li Y, Zhang Y (2014) A functional *recT* gene for recombineering of *Clostridium*. J Biotechnol http://dx.doi.org/10.1016/j.jbiotec.2013.12.011.
221. Datta S, Costantino N, Zhou XM, Court DL (2008) Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages. Proceedings of the National Academy of Sciences of the United States of America 105: 1626-1631.
222. Clarke DJ, Fuller FM, Morris JG (1979) Proton-translocating adenosine triphosphatase of the obligately anaerobic bacterium *Clostridium pasteurianum* .1. ATP phosphohydrolase activity. European Journal of Biochemistry 98: 597-612.
223. Trevors JT, Chassy BM, Dower WJ, Blaschek HP (1992) Electrotransformation of bacteria by plasmid DNA. In: Chang DC, Chassy BM, Saunders JA, Sowers AE, editors. Guide to electroporation and electrofusion. New York: Academic Press Inc. pp. 265-283.
224. Lunnen KD, Barsomian JM, Camp RR, Card CO, Chen SZ, et al. (1988) Cloning Type II restriction and modification genes. Gene 74: 25-32.
225. Zhao YS, Hindorff LA, Chuang A, Monroe-Augustus M, Lyristis M, et al. (2003) Expression of a cloned cyclopropane fatty acid synthase gene reduces solvent formation in *Clostridium acetobutylicum* ATCC 824. Applied and Environmental Microbiology 69: 2831-2841.
226. Kell DB, Peck MW, Rodger G, Morris JG (1981) On the permeability to weak acids and bases of the cytoplasmic membrane of *Clostridium pasteurianum*. Biochemical and Biophysical Research Communications 99: 81-88.

227. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A laboratory manual. 2 ed. Cold Spring Harbor: Cold Spring Harbor Press.
228. Dong HJ, Zhang YP, Dai ZJ, Li Y (2010) Engineering *Clostridium* strain to accept unmethylated DNA. Plos One 5: e9038.
229. Tardif C, Maamar H, Balfin M, Belaich JP (2001) Electrotransformation studies in *Clostridium cellulolyticum*. Journal of Industrial Microbiology & Biotechnology 27: 271-274.
230. Holo H, Nes IF (1989) High-frequency transformation, by electroporation, of *Lactococcus lactis* subsp *cremoris* grown with glycine in osmotically stabilized media. Applied and Environmental Microbiology 55: 3119-3123.
231. Lee SH, Cheung M, Irani V, Carroll JD, Inamine JM, et al. (2002) Optimization of electroporation conditions for *Mycobacterium avium*. Tuberculosis 82: 167-174.
232. Rodriguez MC, Alegre MT, Mesas JM (2007) Optimization of technical conditions for the transformation of *Pediococcus acidilactici* P60 by electroporation. Plasmid 58: 44-50.
233. Baskaran S, Ahn HJ, Lynd LR (1995) Investigation of the ethanol tolerance of *Clostridium thermosaccharolyticum* in continuous culture. Biotechnology Progress 11: 276-281.
234. Tomas CA, Beamish J, Papoutsakis ET (2004) Transcriptional analysis of butanol stress and tolerance in *Clostridium acetobutylicum*. Journal of Bacteriology 186: 2006-2018.
235. Guss AM, Olson DG, Caiazza NC, Lynd LR (2012) Dcm methylation is detrimental to plasmid transformation in *Clostridium thermocellum*. Biotechnology for Biofuels 5: 1-6.
236. Buckley ND, Vadeboncoeur C, LeBlanc DJ, Lee LN, Frenette M (1999) An effective strategy, applicable to *Streptococcus salivarius* and related bacteria, to enhance or confer electroporation competence. Applied and Environmental Microbiology 65: 3800-3804.
237. Framson PE, Nittayajarn A, Merry J, Youngman P, Rubens CE (1997) New genetic techniques for group B streptococci: High-efficiency transformation, maintenance of temperature-sensitive pWV01 plasmids, and mutagenesis with Tn917. Applied and Environmental Microbiology 63: 3539-3547.
238. Helmark S, Hansen ME, Jelle B, Sorensen KI, Jensen PR (2004) Transformation of *Leuconostoc carnosum* 4010 and evidence for natural competence of the organism. Applied and Environmental Microbiology 70: 3695-3699.
239. McDonald IR, Riley PW, Sharp RJ, McCarthy AJ (1995) Factors affecting the electroporation of *Bacillus subtilis*. Journal of Applied Bacteriology 79: 213-218.
240. Peng D, Luo Y, Guo S, Zeng H, Ju S, et al. (2009) Elaboration of an electroporation protocol for large plasmids and wild-type strains of *Bacillus thuringiensis*. Journal of Applied Microbiology 106: 1849-1858.
241. Turgeon N, Laflamme C, Ho J, Duchaine C (2006) Elaboration of an electroporation protocol for *Bacillus cereus* ATCC 14579. Journal of Microbiological Methods 67: 543-548.
242. Powell IB, Achen MG, Hillier AJ, Davidson BE (1988) A simple and rapid method for genetic transformation of lactic streptococci by electroporation. Applied and Environmental Microbiology 54: 655-660.
243. Park SF, Stewart G (1990) High-efficiency transformation of *Listeria monocytogenes* by electroporation of penicillin-treated cells. Gene 94: 129-132.
244. Berthier F, Zagorec M, ChampomierVerges M, Ehrlich SD, MorelDeville F (1996) Efficient transformation of *Lactobacillus sake* by electroporation. Microbiology-UK 142: 1273-1279.

245. Ingram LO (1986) Microbial tolerance to alcohols - Role of the cell membrane. Trends in Biotechnology 4: 40-44.
246. Kundu S, Ghose TK, Mukhopadhyay SN (1983) Bioconversion of cellulose into ethanol by *Clostridium thermocellum* - Product inhibition. Biotechnology and Bioengineering 25: 1109-1126.
247. Liu L, Li YH, Li SL, Hu N, He YM, et al. (2012) Comparison of next-generation sequencing systems. Journal of Biomedicine and Biotechnology.
248. Mavromatis K, Land ML, Brettin TS, Quest DJ, Copeland A, et al. (2012) The fast changing landscape of sequencing technologies and their impact on microbial genome assemblies and annotation. Plos One 7.
249. Bashir A, Klammer AA, Robins WP, Chin CS, Webster D, et al. (2012) A hybrid approach for the automated finishing of bacterial genomes. Nature Biotechnology 30: 701-707.
250. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, et al. (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nature Methods 10: 563-569.
251. Rappert S, Song L, Sabra W, Wang W, Zeng A-P (2013) Draft genome sequence of type strain *Clostridium pasteurianum* DSM 525 (ATCC 6013), a promising producer of chemicals and fuels. Genome announcements 1.
252. Jensen TO, Kvist T, Mikkelsen MJ, Westermann P (2012) Production of 1,3-PDO and butanol by a mutant strain of *Clostridium pasteurianum* with increased tolerance towards crude glycerol. *AMB Express* 2: 44.
253. Malaviya A, Jang YS, Lee SY (2012) Continuous butanol production with reduced byproducts formation from glycerol by a hyper producing mutant of *Clostridium pasteurianum*. Applied Microbiology and Biotechnology 93: 1485-1494.
254. Roberts RJ (1987) Restriction enzymes and their isoschizomers. Nucleic Acids Research 15: R189-R215.
255. Kneale GG (1994) Symmetrical model for the domain structure of Type I DNA methyltransferases. Journal of Molecular Biology 243: 1-5.
256. Papoutsakis ET, Bennett GN (1999) Metabolic engineering of *Clostridium acetobutylicum*. In: Lee SY, Papoutsakis ET, editors. Metabolic engineering. New York: Marcel Dekker Inc. pp. 253-279.
257. Du T-f, He A-y, Wu H, Chen J-n, Kong X-p, et al. (2013) Butanol production from acid hydrolyzed corn fiber with *Clostridium beijerinckii* mutant. Bioresource Technology 135: 254-261.
258. Moon C, Lee CH, Sang BI, Um Y (2011) Optimization of medium compositions favoring butanol and 1,3-propanediol production from glycerol by *Clostridium pasteurianum*. Bioresource Technology 102: 10561-10568.
259. Kuit W, Minton NP, Lopez-Contreras AM, Eggink G (2012) Disruption of the acetate kinase (*ack*) gene of *Clostridium acetobutylicum* results in delayed acetate production. Applied Microbiology and Biotechnology 94: 729-741.
260. Bamforth CW, Quayle JR (1978) Aerobic and anaerobic growth of *Paracoccus denitrificans* on methanol. Archives of Microbiology 119: 91-97.
261. Bagaeva TV, Zinurova EE (2004) Comparative characterization of extracellular and intracellular hydrocarbons of *Clostridium pasteurianum*. Biochemistry-Moscow 69: 427-428.