NORTHWESTERN UNIVERSITY

METABOLIC ENGINEERING OF *CLOSTRIDIUM ACETOBUTYLICUM* FOR ENHANCED BUTANOL PRODUCTION AND SELECTIVITY

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Department of Chemical and Biological Engineering

By

RYAN SILLERS

EVANSTON, ILLINOIS

DECEMBER 2008

ABSTRACT

Increasing butanol titers and selectivity in *Clostridium acetobutylicum*

Ryan Sillers

To increase the butanol titers and selectivity in *Clostridium acetobutylicum* we replaced the promoter of the alcohol/aldehyde dehydrogenase (*aad*) gene with the phosphotranbutyrylase (*ptb*) promoter and combined this with CoAT downregulation to minimize acetone production. This led to early production of high alcohol (butanol plus ethanol) titers and overall solvent titers of 30 g/L. To increase then the carbon flux towards butanol, we examined thiolase (*thl*) overexpression. The combined *thl* overexpression with *aad* overexpression decreased, as expected, acetate and ethanol production while increasing acetone and butyrate formation. *thl* overexpression in strains with CoAT downregulation did not significantly alter product formation thus suggesting that a more complex metabolic engineering strategy is necessary to achieve improved butanol titers and selectivity. The *aad* with the *ptb* promoter was also used in the M5 strain and restored butanol production to wildtype levels. Thiolase overexpression was combined with *aad* overexpression aiming to enhance butanol formation. While acetate was reduced, butanol titers were not improved. We generated acetate kinase (AK) and butyrate kinase (BK) knockout (KO) mutants of M5, reducing the respective acid formation. We could not successfully transform the BKKO M5 strain. The AKKO M5 strain overexpressing *aad* produced less acetate, but also less butanol compared to the M5 *aad* overexpression strain. The difficulty in generating high butanol producers without acetone and acid production is likely

hindered by the inability to control the electron flow, which may be affected by unknown pSOL1 genes**.**

Using antisense RNA, we investigated the downregulation of the histidine kinase coded by CAC0654 on solventogenesis and sporulation. Our findings suggest the two-component system coded by CAC0654 and CAC0653 (a kinase and its cognate response regulator) is a putative negative regulator of sporulation. pH-controlled fermentations of the CAC0654 asRNA strain construct targeting show reduced glucose uptake and an 80% reduction in butanol production compared to the plasmid control strain. Transcriptional profiling of 824(pHK654) using microarrays shows profound upregulation of the sporulation cascade downstream of Spo0A, the master regulator of sporulation, through the expression of *sigK*. This two-component system may act affect the phosphorylation state of Spo0A repressing *sigF*, *spoIIE*, and *sigE*.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

1.1 General clostridia characteristics

The clostridia genus contains diverse set of members ranging from pathogens (*C. difficile*, *C. botulinum*, *C. tetani*, and *C. perfringens*) to soil members that produce a variety of solvents and organic acids (*C. acetobutylicum and C. beijerinkii*). A number of reviews are available that present this material in greater detail (Mitchell 2001; Dürre 2005). Common attributes of clostridia members include anaerobic metabolism, Gram-positive staining, and endospore formation. Actively growing cells exhibit a rod-shaped morphology that under appropriate conditions can give rise to endospores. Although clostridia do not grow in the presence of oxygen, the ability of individual species to tolerate oxygen varies. The mechanism of oxygen inhibition is not completely understood, but it is suggested that clostridia have limited ability to tolerate the peroxides and free hydroxyls formed under aerobic growth. Under anaerobic growth, clostridia tolerate a wide range of temperatures and pH. Most clostridia grow well at neutral pH, but the acid-producing members typically lower their extracellular pH environment to pH 4-5. Optimal temperatures for the majority of clostridia occur at 30 – 40 ^{o}C, but thermopilic clostridia (e.g. *C. thermocellum*) grow well in excess of 60 ^{o}C. Many of these themophilic clostridia are of interest for their thermo-stable enzymes and in particular the degradation of cellulose. During growth clostridia are able to metabolize many different carbon sources, from simple hexose and pentose sugars, to more complex starches and cellulose. This wide substrate usage in clostridia makes them industrially relevant as they hold the potential to ferment inexpensive and renewable sources into a variety of compounds.

1.2 *Clostridium acetobutylicum* characteristics and fermentation

C. acetobutylicum and related strains were used extensively in the acetone-butanol fermentation and remain targets of basic and applied research. *C. acetobutylicum* ATCC 824 was isolated in Connecticut in the 1920's and is considered the type strain. This genome has been sequenced and annotated (Nölling, Breton et al. 2001), and methods for genetic deletions (Harris, Welker et al. 2002; Heap, Pennington et al. 2007; Shao, Hu et al. 2007) and gene overexpression (Mermelstein and Papoutsakis 1993) have been developed, making it an attractive organism for development. The sequencing data revealed a 3.94 Mb chromosome and a 192 kb megaplasmid (pSOL1) containing 3,740 open reading frames (ORFs). The genome is very AT-base pair rich (72%). The megaplasmid contains genes necessary for sporulation (still unknown) and all essential genes for solvent production (known). The loss of (pSOL1) is common in ATCC 824 under continuous culture conditions or repeated vegetative transfers and leads to degeneracy and the inability to form spores or produce solvents (Cornillot, Nair et al. 1997).

The fermentation of *C. acetobutylicum* is characterized by two phases, the acidogenic and the solventogenic phase. The acidogenic phase occurs during vegetative growth, when the cultures produce the organic acids acetate and butyrate, which lower the culture pH. In the solventogenic phase, the culture produces butanol, acetone, and ethanol as the culture enters into the stationary phase. Butyrate and acetate are typically re-assimilated to produce solvents, thus raising the pH of the culture. This shift towards solvent production is thought to be a way for the culture to detoxify the culture environment of the harmful acids towards the less damaging solvents (Jones and Woods 1986). In both batch and continuous cultures of the wildtype strain, peak solvent production occurs at low pH, between 4.5 and 5.0, but as extracellular pH is

increased solvent production decreases and production of acetate and butyrate is elevated (Bahl 1982; Monot 1984). The understanding of this pH dependence is still not fully understood, but the onset of solvent production has been associated with the level of undisassociated butyric acid and butyryl phosphate (Huang, Forsberg et al. 1986; Harris, Desai et al. 2000; Zhao, Tomas et al. 2005).

1.3 ABE fermentation and biofuel production

The production of acetone and butanol by Clostridium strains was one of the earliest industrial fermentations. In the early 1900's, the industrial acetone-butanol-ethanol (ABE) fermentation was second in size only to ethanol fermentation. Early interest in the ABE fermentation was based on acetone production for gunpowder in World War One. The ABE fermentation continued through the Second World War and into the early 1960's until the process was unable to complete with the petrochemical process. The major factors contributing to the decline of the ABE fermentation are high substrate costs, low solvent titers, and high product recovery costs. Molasses were typically used as a substrate and were responsible for over half of the total costs associated with the process. The increased use of molasses in livestock feed raised prices and contributed to the decline of the process.

With concerns over crude oil supplies and rising costs, there is significant interest in reducing dependence on fossil fuels and increasing the use of renewable energy sources. The recent advances in molecular biology and ME techniques offer an opportunity to re-establish the AB fermentation as an economically viable process. *Clostridium acetobutylicum* is a model host for the production of acetone and butanol. It has a naturally evolved metabolism for these solvents and can metabolize both hexose and pentose sugars. Mixed carbon sources such as

hemicellulose are inexpensive and readily available from biomass hydrolysis. Active research is also underway to use cellulose directly as a carbon source (Perret, Casalot et al. 2004), without the need for acid hydrolysis pretreatment and this may greatly reduce process costs. Newly engineered strains have increased butanol titers significantly (from 13 to 18 g/l) (Green, Boynton et al. 1996; Tomas, Welker et al. 2003). Further increasing titers to 22-28 g/L would reduce separation costs and could make the AB fermention once again economically viable (Woods 1995).

Although initial efforts in biofuels have focused on ethanol, butanol is widely considered a superior energy carrier. This is based on the similar energy content per volume of the butanol and traditional gasoline, the ability for butanol to be used either directly as a transportation fuel or at high blend rates with gasoline without modifications to current vehicles, and the low vapor pressure and water solubility of butanol which allow it to be shipped using current infrastructure. Ethanol, in contrast, has lower energy content than butanol or fossil fuels per volume, requires engine modifications for high blend rates, and its water miscibility and vapor pressure make its transport via existing infrastructure difficult.

1.4 Toxicity

Butanol toxicity is a major limiting factor that prevents high solvent titers necessary for economical solvent production. Butanol has a chaotropic effect on the cell membrane disrupting membrane fluidity and function (Bowles and Ellefson 1985). Experimental results from *E. coli* suggest that butanol partitions into the cell membrane disrupting the packing of fatty acyl chains (Ingram 1976). The cell then responds to this stress by altering the cell membrane to increase the ratio of saturated to unsaturated fatty acids (Sullivan, Hegeman et al. 1979). In *C.*

acetobutylicum, butanol disrupts the pH and transmembrane electrical gradients, lowers intracellular ATP concentration, and inhibits glucose uptake (Bowles and Ellefson 1985; Terracciano and Kashket 1986). Butanol levels of 14-15 g/L have been found to completely inhibit growth and maintenance of the pH gradient (Bowles and Ellefson 1985; Ounine, Petitdemange et al. 1985).

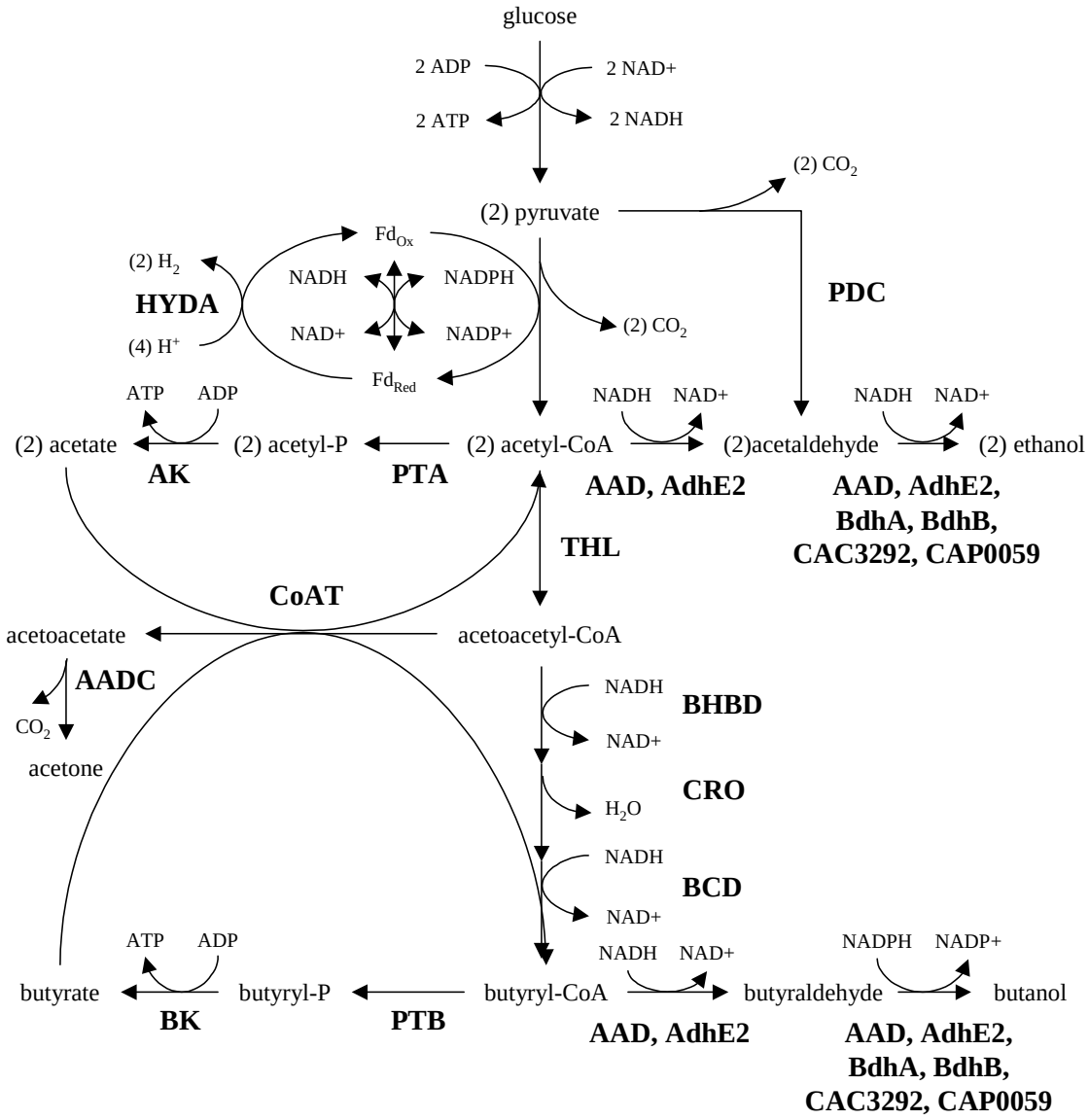
Due to this inhibitory effect of butanol, there is significant interest in developing strains that have increased butanol tolerance and formation capability. In the past, mutant strains have been screened based on their ability to tolerate high levels of butanol in hopes of creating strains that will produce butanol in excess of the normal toxicity levels. Unfortunately, while some mutant strains tolerated high levels of butanol, they did not produce higher butanol titers (Jones and Woods 1986). Recombinant strains have been developed that produce butanol titers in excess of normal inhibitory levels. The heat shock overexpression strain 824(pGROE1) and the butyrate kinase mutant strain PJC4BK produce 17-18 g/L butanol (Green, Boynton et al. 1996; Tomas, Welker et al. 2003). The ability of these strains to increase butanol tolerance and titers without membrane alterations suggests further improvements can be made using ME.

1.5 Primary metabolism

The metabolism of glucose and other hexose sugars occurs through the Embden-Meyerhof pathway (Figure 1.1) and produces 2 mol of pyruvate, 2 mol of adenosine triphosphate (ATP), and 2 mol of reduced nicotinamide adenine dinucleotide (NADH) for every mol of glucose. Pentose sugars are utilized through the pentose phosphate pathway and yield 5 mol of ATP and NADH for every 3 mol of sugar. Pyruvate is cleaved to form carbon dioxide, acetyl-CoA, and reduced ferredoxin. Under rare conditions, pyruvate may also form lactate to

Figure 1.1 Primary metabolism of *C. acetobutylicum*.

Enzymes are abbreviated as follows: phosphotransacetylase (PTA); acetate kinase (AK); thiolase (THL); β-hydroxybutyryl dehydrogenase (BHBD); crotonase (CRO); butyryl-CoA dehydrogenase (BCD); CoA Transferase (CoAT); acetoacetate decarboxylase (AADC); butyrate kinase (BK); phophotransbutyrylase (PTB); alcohol/aldehyde dehydrogenase (AAD); hydrogenase (HYDA); pyruvate decarboxylase (PDC). Note: AAD is believed to be the primary enzyme for butanol and ethanol formation but additional genes exist that code for alcohol forming enzymes (*adhe2*, *bdhA*, *bdhB*, CAC3292, CAP0059).


glucose
2 ADP
2 ATP
2 NAD+
2 NADH
(2) pyruvate
(2) CO_2
PDC
Fd_{Ox}
Fd_{Red}
(2) H_2
HYDA
(4) H^+
NADH
NADPH
NAD+
NADP+
(2) CO_2
ATP
ADP
(2) acetate
AK
(2) acetyl-P
PTA
(2) acetyl-CoA
AAD, AdhE2
(2)acetaldehyde
AAD, AdhE2,
BdhA, BdhB,
CAC3292, CAP0059
(2) ethanol
THL
CoAT
acetoacetate
acetoacetyl-CoA
AADC
CO_2
acetone
BHBD
CRO
H_2O
BCD
butyrate
BK
butyryl-P
PTB
butyryl-CoA
AAD, AdhE2
butyraldehyde
AAD, AdhE2,
BdhA, BdhB,
CAC3292, CAP0059
butanol

provide another source of NADH oxidation when the generation of molecular hydrogen is disrupted (Green 1998). Reduced ferredoxin can use protons as an electron acceptor producing molecular hydrogen, which re-oxidizes ferredoxin. Hydrogen production is highest during the acidogenic phase, when hydrogenase (HYDA) is primarily responsible for the oxidation of ferredoxin. During solventogenesis, reduced ferredoxin is oxidized producing NADH and NADPH, which are consumed in the production of butanol and ethanol. Acetyl-CoA serves as the first major branch point for ethanol, acetate, and acetoacetyl-CoA formation.

Under acid-formation conditions, acetyl-CoA is converted to acetate by phosphate acetytransferse (PTA) and acetate kinase (AK) producing 1 mol ATP for every mol acetate formed. Acetyl-CoA can also be converted butyryl-CoA, which is the precursor to butyrate. During the conversion of acetyl-CoA to butyryl-CoA, acetyl-CoA is converted to acetoacetyl-CoA by thiolase (THL). Acetoacetyl-CoA is converted to butyryl-CoA by three enzymes, β-hydroxybutryl-coenzyme A dehydrogenase (BHBD), butyryl-CoA dehydrogenase (BCD), and crotonase (CRO). Once butyryl-CoA is formed, it can be converted in a manner analogous to acetate by phosphotransbutyrlyase (PTB) and butyrate kinase (BK) producing 1 mol ATP for every mol butyrate formed. During the conversion of acetyl-CoA to butyryl-CoA, 2 mol NADH is consumed regenerating NAD^+ used in glycolysis. Butyrate production is thus a NADH neutral pathway, whereas acetate production results in the net production of 1 mol NADH for every mol acetate produced.

During solventogenesis, ethanol is produced from acetyl-CoA via a two-step process by the bi-fuctional alcohol aldehyde dehydrogenase (AAD). Butyryl-CoA is then converted to butanol by the same AAD enzyme that produces ethanol. AAD had been shown to be the primary enzyme responsible for butanol and ethanol production, but other genes in the genome

may contribute to alcohol formation under certain conditions. There exists a second bifuntional alcohol aldehyde dehydrogenase in the chromosome. This second enzyme, AdhE2, is specifically induced in high pH conditions combined with increased NADH availability (Fontaine, Meynial-Salles et al. 2002). This enzyme is not highly transcribed in typically batch fermentations and does not contribute to butanol formation under glucose fed batch cultures. In addition to the two alcohol aldehyde dehydrogenases, there are four other alcohol dehydrogenases on the chromosome that may aid the conversion of the respective aldehyde into ethanol or butanol (Nölling, Breton et al. 2001). Ethanol may also be formed from pyruvate independent of acetyl-CoA by pyruvate decarboxylase (PDC). Once acetaldehyde is formed, any of the alcohol dehydrogenases may convert it into ethanol.

Both acetate and butyrate can be re-assimilated by the CoA transferase (CoAT) enzyme thus replenishing acetyl-CoA and butyryl-CoA, respectively. This reaction is coupled to the conversion of acetoacetyl-CoA to acetoacetate, where the CoA species is transferred from acetoacetyl-CoA to acetate or butyrate. Acetoacetate is then converted to acetone by acetoacetate decarboxylase (AADC). This coupled reaction dictates that the molar amount of acids reassimilated is equal to the number of moles of acetone formed.

1.6 Metabolic engineering

Metabolic engineering (ME) has been defined as the “improvement of cellular activities by manipulation of enzymatic, transport, and regulatory functions of the cell with the use of recombinant DNA technology” (Bailey 1991). Later definitions emphasize the “purposeful” and “directed” approach ME applies identifying genetic targets (Cameron and Tong 1993; Stephanopoulos 1999). These definitions contrast ME’s rational and logical approach compared

to earlier efforts of genetic modifications that used random mutagenesis and selection to improve cellular properties. The rapid advancements in bioinformatics and genomic tools along with molecular biology techniques provide many opportunities to improve the bioprocessing traits of microorganisms in a directed manner. ME applications range from heterologous gene transfer to develop novel compounds or utilize alternate substrates within a host, to the optimization of product pathways naturally occurring in a cell. These applications may require multiple genetic modifications to develop an industrially relevant strain. Each modification must be followed by detailed analysis to determine the cellular response and identify further genetic targets. This iterative nature is essential to the success of ME strategies.

Strain development uses a variety of techniques to modify the metabolism of a host. Heterologous gene transfer into a host can expand natural pathways to introduce new cellular properties beneficial to bioprocessing. For example, the expression of an endoglucanase gene from *Clostridum thermocellum* in a *Klebsiella oxytoca* strain previously engineered for ethanol production allows the use of cellulose as the substrate (Wood and Ingram 1992). Heterologous enzymes may also show higher specificity for a substrate or increased reaction rates compared to native enzymes. Naturally existing pathways are not optimized for industrial processes and are often modified to increase product titers and yields. Branch points where intermediate metabolites may either form a desired product or an unwanted byproduct are common targets of ME applications. Eliminating competing pathways using genetic knockouts may raise product yields while overexpression of rate-limiting enzymes can increase carbon flux to the desired product. Genes not directly involved in the product metabolism, such as regulatory enzymes or transcriptional factors, are also important genetic targets that can produce beneficial effects.

The iterative process of strain development includes a genetic modification, determination of the cellular response and identification of future targets (Bailey 1991). The cellular response to simple genetic changes may be dramatic and may become more challenging to interpret as further modifications are made. Even the insertion of an empty cloning vector containing only antibiotic resistance into a strain can burden a cell (Ricci and Hernandez 2000) and produce widespread consequences throughout the organism altering growth and metabolism (Tomas, Welker et al. 2003). Advanced analytical and computational methods are needed to fully understand the effects of genetic modifications. Detailed analysis can determine successful modifications useful in the development of engineered strains.

1.7 Metabolic Flux Analysis

The emergence of genomic techniques allows for the rational design of recombinant strains, but the simple identification of the genes involved does not always provide enough information to increase and redirect the metabolism to the desired products. Metabolic flux analysis (MFA) is an important analytical tool to quantitate the intracellular fluxes through metabolic pathways. MFA calculates individual fluxes of cellular reactions using stoichiometric models and mass balances together with pseudo-steady state assumptions of metabolic intermediates (Vallino and Stephanopoulos 1993; Edwards 1999). The calculated fluxes through the intermediate metabolites provide a much-improved understanding of cellular kinetics and can identify prevailing flux bottlenecks in the biological system. These bottlenecks can then be used as targets for future genetic modifications to increase carbon flow to a specific product.

The basis for the stoichiometric model used for *C. acetobutylicum* is the fermentation equation (Papoutsakis 1984). However, this metabolic system is underdetermined due the presence of a

singularity concerning the acetone production via the uptake of either acetate or butyrate. Although butyrate is observed to be the primary acid reassimilated during growth, *in vitro* experiments suggest that the CoAT enzyme has higher substrate specificity for acetate than butyrate (Wiesenborn, Rudolph et al. 1989). This is resolved with the observed data suggesting that acetate production is maintained, combined with the increased acetate uptake, resulting in a observed smaller net acetate uptake. In resolving the singularity, the results from the *in vitro* experiments are used.

1.8 Sporulation

As clostridia enter into conditions that are unfavorable to growth, the cells differentiate and begin the process to form mature spores that are protective against a number of harsh environmental conditions (Fig 1.2). The conditions that start the sporulation process are not fully understood, and different species sporulate at different frequencies under different conditions. The clostridia cell cycle begins in vegetative growth, where the cells undergo normal cell replication and division. As the cells continue to grow they begin to differentiate into swollen clostridia cells begin granulose formation and storage. These swollen clostridia forms are commonly implicated in the formation of solvents and toxins. Later in development a forespore is formed. Following the maturation of the forespore, the mother cell undergoes lysis and the mature spore is released. Once the spores find a hospitable environment they will germinate and once again become vegetative cells.

1.9 Genetics of sporulation and solvent formation

As the cell begins its early vegetative growth the primary metabolic products are the organic acids, acetate and butyrate. In response to the accumulation of the acid products, genes

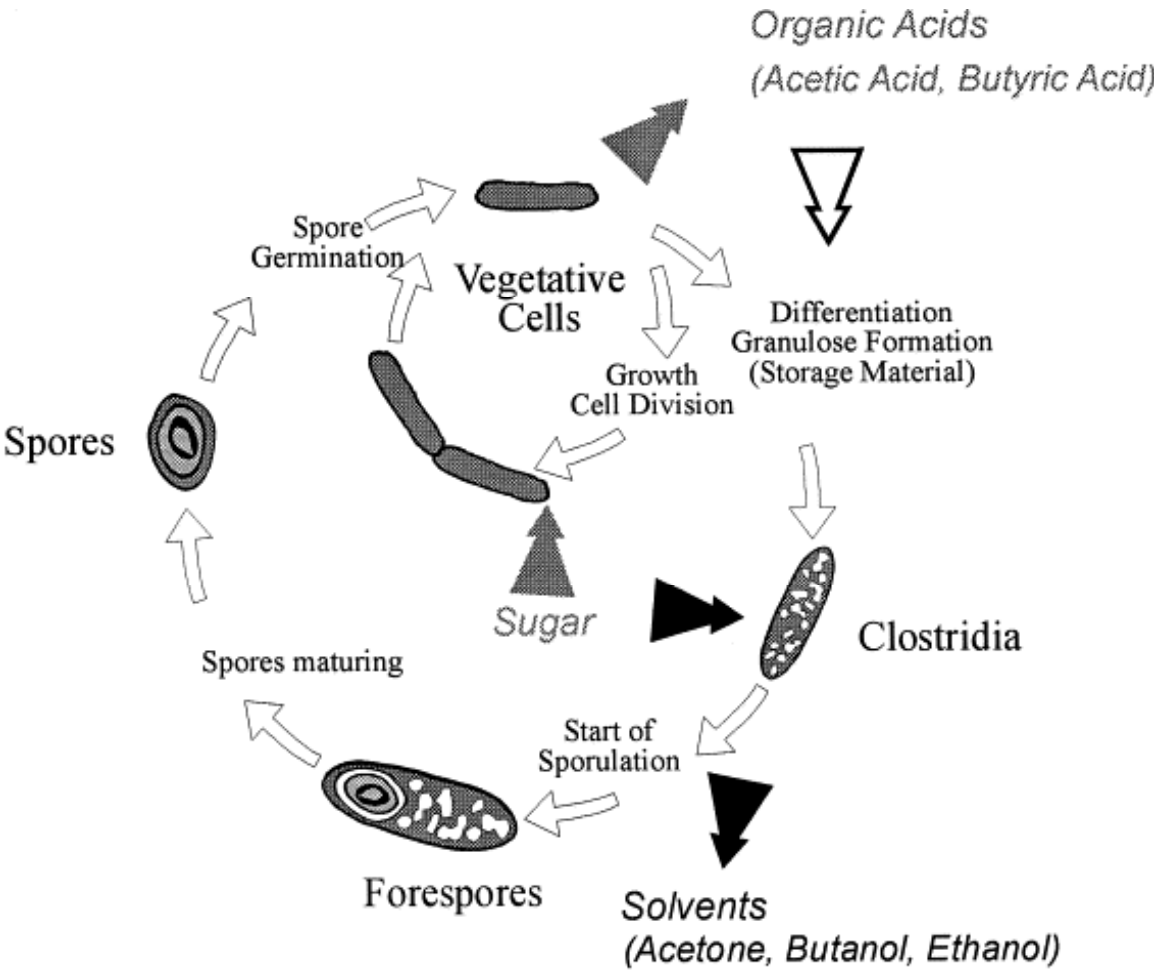


Figure 1.2 Diagram of the clostridial cell cycle and sporulation process

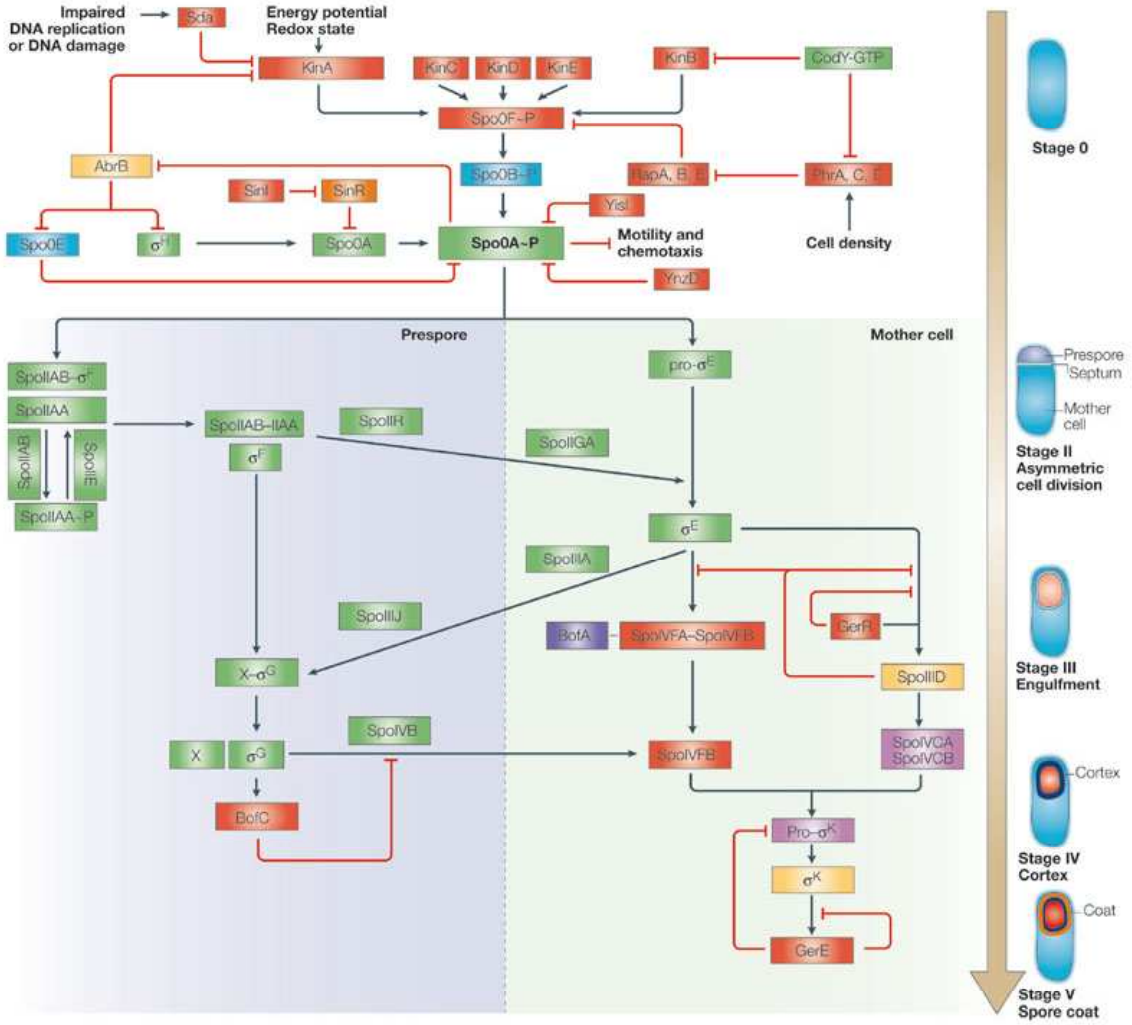
From (Mitchell 2001).

are induced to convert the acids into less harmful solvent products. This conversion to solvent products presumably allows the cell an opportunity for completion of the sporulation process. This provides the underlying inter-commitment between the sporulation process and solvent formation. At the molecular level both sporulation and solvent formation are regulated by the master transcriptional regulator, Spo0A. Once activated (phosphorylated), Spo0A binds its target genes causing this switch from acid to solvent production and sporulation. This has been demonstrated in both *C. acetobutylicum* and *C. beijerinckii spo0A* knockout mutants, which are unable to sporulate and only very low levels of solvents are produced (Ravagnani, Jennert et al. 2000; Harris, Welker et al. 2002). The direct binding of Spo0A causing the induction of solvent genes and repression of butyrate formation genes has been shown in *C. beijerinckii* (Ravagnani, Jennert et al. 2000). However, the mechanism of Spo0A activation remains unknown. Unlike the related and widely-studied spore forming *B. subtilis*, carbon and nitrogen starvation does not induce sporulation (Long, Jones et al. 1984) and many of the genetic elements proceeding Spo0A activation in *B. subtilis* are missing in clostridia (Nölling, Breton et al. 2001; Paredes, Alsaker et al. 2005) (Fig 1.3).

Proposed mechanisms for Spo0A activation include phosphorylation by a sensor kinase(s), acetyl-phosphate or butyryl-phosphate regulation, or a modified phosphorelay system (Paredes, Alsaker et al. 2005). In *B. subtilis*, Spo0A is phosphorylated by five different sensor kinases (Jiang, Shao et al. 2000). All five kinases are “orphan” kinases that are not directly associated with a response regulator in the same operon (Fabret, Feher et al. 1999). Within *C. acetobutylicum* there are five “orphan kinases” and two of these exhibit a significant induction just prior to the onset of sporulation (when gene targets of Spo0A are induced). Butyryl-phosphate has also been proposed as a signal molecule inducing solvent formation and

Figure 1.3 Sporulation cascade of *B. subtilis* and clostridial homologs.

From (Paredes, Alsaker et al. 2005). Proteins in red have not been identified in any clostridia species. Proteins in green are reported to exist or best-best BLASTP match found in all five clostridia. Proteins in blue may exist in *C. tetani* with low similarity. Proteins in yellow exist in all clostridia except *C. tetani.* Proteins in orange are found only in *C. acetobutylicum* and *C. tetani* (Note: AbrB existence in *C. perfringens* is also doubtful). Proteins in purple are found only in *C. acetobutylicum* and *C. botulinum*. In *B. subtilis* and *C. difficile,* σ^k is encoded as two fragments that are post-transcriptionally processed into a functional protein (pink).


Impaired DNA replication or DNA damage
Sda
Energy potential Redox state
KinA
KinC
KinD
KinE
KinB
CodY-GTP
Spo0F~P
AbrB
Spo0B~P
SinI
SinR
YisI
Spo0E
σH
Spo0A
Spo0A~P
Motility and chemotaxis
Cell density
YnzD
Prespore
Mother cell
SpoIIAB–σF
SpoIIAA
SpoIIAB–IIAA
SpoIIR
σF
SpoIIGA
SpoIIAA~P
pro-σE
σE
SpoIIA
SpoIIJ
BofA
SpoIVFA–SpoIVFB
GerR
SpoIID
X–σG
X
σG
SpoIVB
SpoIVFB
SpoIVCA SpoIVCB
BofC
Pro–σK
σK
GerE
Stage 0
Prespore
Septum
Mother cell
Stage II Asymmetric cell division
Stage III Engulfment
Cortex
Stage IV Cortex
Coat
Stage V Spore coat

sporulation (Harris, Desai et al. 2000). A butyrate kinase deletion mutant was found to produce lower levels of butyrate and resulting in increased solvent production and earlier sporulation (Green, Boynton et al. 1996; Harris, Desai et al. 2000). Later studies confirmed the increased concentration and earlier accumulation of butyryl-phosphate in the butyrate kinase mutant suggesting a possible role as a phosphate donor in cellular signaling (Zhao, Tomas et al. 2005).

During sporulation, Spo0A begins the signal cascade that results in the expression of other sigma factors that are critical for proper sporulation. Each factor in the cascade is necessary for the expression or activation of the next. The cascade in expressed in the following order: σ^H, Spo0A, σ^F, σ^E, σ^G, σ^K (Dürre and Hollergschwandner 2004; Paredes, Alsaker et al. 2005). *spo0A* is transcribed from a σ^H dependent promoter and once phosphorylated drives the expression of the *sigF*, *sigE* and *spoIIE* and represses *abrB*, a repressor of sporulation (Scotcher, Rudolph et al. 2005). SpoIIE is responsible for the release of σ^F from its anti-sigma factor. Once released, σ^F activates the transcription of *spoIIR*, which interacts with SpoIIGA to process the pro-σ^E factor. The functional σ^E then drives the expression of *spoIIID* and the *spoIIIAA-spoIIIAH* operon that interacts with *spoIIIJ* to activate σ^G. The expression of the final sporulation sigma factor, σ^K, requires *spoIIID* and σ^G.

The genes responsible for solvent production in *C. acetobutylicum* include *ctfA/B* and *adc* (acetone formation), *bdhA/B* and two *adhE* genes (primarily butanol but also ethanol formation). All but the *bdhA/B* genes are located on a large megaplasmid pSOL1 whose loss leads not only to the virtual elimination of solvent production but also abolishes sporulation (Cornillot, Nair et al. 1997). The five genes on pSOL1 are organized into three operons (Paredes, Rigoutsos et al. 2004). The *sol* operon contains *aad-ctfA-ctfB* while the *adc* and *adhE2* are organized in separate

monocistronic operons. The *sol* operon and *adc* are located adjacent to each other, but are transcribed from opposite strands. The *adhE2* is separated by about 50 kbp. The chromosomal *bdhA* and *bdhB* are organized into two consecutive monocistronic operons (Paredes, Rigoutsos et al. 2004).

As described earlier *adc* is directly controlled by Spo0A (Ravagnani, Jennert et al. 2000). Spo0A is also likely to directly control the expression of *sol* operon based on the Spo0A mutant and promoter studies (Thormann, Feustel et al. 2002). The disruption of the Spo0A motif upstream of the *sol* operon resulted in the elimination of promoter activity. The study also suggests the involvement of another transcription factor for the *sol* operon expression as overexpression of the promoter element alone causes a decrease in solvent production, presumably due to the titration of the unknown activator.

1.10 Microarrays

Microarrays are a high-throughput technology for the measurement of large amounts of genomic information. The earliest microarrays first reported in 1995 allowed for the analysis of fewer than 100 genes (Schena, Shalon et al. 1995), but in a relatively short time technology companies such as Agilent are producing DNA arrays that can measure nearly a quarter million different spots on a single 1 x 3 inch microscope slide. Microarrays consist of a probe that is bound to the slide that can be generated from PCR products that are printed on the slide, or can be synthesized directly on the slide. Initial uses of microarrays focused on the relative measurements of RNA transcription, but since then the uses have expanded greatly to include detection of DNA binding sites, SNP analysis, and various chromosomal modification analysis.

In a typical microarray experiment, two samples are compared for relative transcription levels. These samples may be different points in a timecourse, from organisms with different genetic backgrounds, or testing the transcriptional differences of an organism grow in different media. Once the RNA is isolated from a sample (or DNA, depending on the individual application) it is copied using reverse transcription or PCR and is labeled with an appropriate, normally fluorescent dye, such as Cy3 or Cy5. After the preparation of the individual samples, they are hybridized on the microarray for a number of hours to allow the binding of a target sequence to the probe on the microarray. After hybridization the microarrays are washed to eliminate any non-specific hybridization. The microarrays are then scanned to detect the levels of fluorescent intensity at each spot. Once this intensity data is collected the data is normalized and processed to account for sample variation and dye differences such as incorporation differences or differences in excitation strength. Once processed the data can be analyzed in many ways using a variety of clustering algorithms to group common expression patterns or identification of dramatically different expression profiles.

Microarrays have been used in a variety of organisms from eukaryotes to prokaryotes. They have been used to identify the transcriptional program differentiating different lymphoma cells (Alizadeh, Eisen et al. 2000) to the pattern of sporulation in budding yeast (Chu, DeRisi et al. 1998). In prokaryotes, there has been significant work in the identification of regulatory networks. This has used transcriptional data comparing gene induction between wildtype and mutant strains, as well as DNA binding data to identify the sites on the chromosome where a particular transcription factor will bind. This has allowed the construction of a transcription network of many of the transcription factors and sigma factors regulating sporulation in *B.*

subtilis (Britton, Eichenberger et al. 2002; Eichenberger, Jensen et al. 2003; Molle, Fujita et al. 2003; Molle, Nakaura et al. 2003; Eichenberger, Fujita et al. 2004; Wang, Setlow et al. 2006).

The application of microarrays to clostridia has focused on the development and sporulation program of *C. acetobutylicum* (Tomas, Alsaker et al. 2003; Alsaker and Papoutsakis 2005). Additional studies have investigated the cells response to butanol in a number of genetic backgrounds (Tomas, Beamish et al. 2003; Alsaker, Spitzer et al. 2004). The cellular response to a wide variety of metabolites has also been determined and a general transcriptional stress response has been proposed. Microarrays have also been used in the identification of enriched populations within a mixed plasmid overexpression library in order to identify genes that confer tolerance to butanol (Borden and Papoutsakis 2007).

CHAPTER 2 : MATERIALS AND METHODS

2.1 Culture conditions and maintenance of strains

E. coli strains were grown aerobically at 37°C and 200 rpm in liquid LB media or solid LB with agar (1.5%) media supplemented with the appropriate antibiotics (ampicillin at 50 μg/mL or chloramphenicol at 35 μg/mL). Frozen stocks were made from 1 mL overnight culture resuspended in LB containing 15% glycerol and stored at -85°C. *C. acetobutylicum* strains were grown anaerobically at 37°C in an anaerobic chamber (Thermo Forma, Waltham, MA). Cultures were grown in liquid CGM (containing 0.75 g KH_2PO_4, 0.982 g K_2HPO_4, 1.0 g NaCl, 0.01 g $MnSO_4$, 0.004 g PABA, 0.348 g $MgSO_4$, 0.01 g $FeSO_4$, 2.0 g asparagine, 5.0 g yeast extract, 2.0 g $(NH_4)_2SO_4$, and 80 g glucose, all per liter) media or solid 2xYTG pH 5.8 (containing 16 g Bacto tryptone, 10 g yeast extract, 4 g NaCl, and 5 g glucose, all per liter) plus agar (1.5%) supplemented with antibiotics as necessary (erythromycin at 100 μg/mL in liquid media and 40 μg/mL in solid media, clarithromycin at 75 μg/mL). Cultures were heat shocked at 70-80°C for 10 minutes prior to enhance solvent production and prevent strain degeneration (Cornillot, Nair et al. 1997). Frozen stocks were made from 10 mL of A_{600} 1.0 culture resuspended in 1 mL CGM containing 15% glycerol and stored at -85°C.

2.2 Bioreactor experiments

Fermentations were carried out using a BioFlo 110 or BioFlo II (New Brunswick Scientific Co., Edison, NJ) bioreactor with 4.0 L working volumes. Fermentations used a 10% v/v inoculum of a pre-culture with A_{600} equal to 0.2. CGM media were supplemented with

0.10% (v/v) antifoam and 75 μg/mL clarithromycin. Fermentations were maintained at constant pH using 6 M NH_4OH. Anaerobic conditions were maintained through nitrogen sparging. Temperature was maintained at 37°C and agitation was set at 200 rpm. Glucose was restored to the initial concentration (440 mM) in fermentations if glucose levels fell below 200 mM.

2.3 Analytical techniques

Cell density was measured at A_{600} using a Biomate3 spectrophotometer (Thermo Spectronic, Waltham, MA). Samples were diluted as necessary to keep absorbance below 0.40. Supernatant concentrations of glucose, acetone, acetate, acetoin, butyrate, butanol, and ethanol were determined using a high-pressure liquid chromatography system (HPLC) (Waters Corp., Milford, MA) (Buday, Linden et al. 1990). Mobile phase of 0.15 mM H_2SO_4 at 0.50 mL/min was used with an Aminex HPLC Organic Acid Analysis Column (Biorad, Hercules, CA). The column was cooled to 15°C and samples were run for 55 minutes.

2.4 RNA sampling and isolation

Cell pellets from 3 to 10 mL of culture were incubated at 37°C for 4 minutes in 200 μL of SET buffer (25% sucrose, 50 mM EDTA pH 8.0, 50 mM Tris-HCl pH 8.0) with 20 mg/mL lysozyme. 1 mL trizol was added to each sample and stored at -85°C until purification. 0.5 mL Trizol and 0.2 mL chloroform was added to 0.5 mL RNA sample and centrifuged at 12,000 rpm for 15 minutes. The aqueous phase was collected and added to an equal volume of isoproponal and RNA was precipitated at 12,000 rpm for 10 minutes. 1 mL ethanol was added to wash the pellet and centrifuged at 9,500 rpm for 4 minutes. Samples were dried and resuspended in 20-100 μL of RNase free water and stored at -85°C.

2.5 Quantitative (Q)-RT-PCR

Reverse transcription of RNA was carried out using random hexamer primers with 500 μM dNTPs, 2.0 μg RNA, 2 μL RNase inhibitor, 2.5 μL reverse transcriptase, and 2.5 μM random hexamers in a total volume of 100 μL (Applied Biosystems). The reaction was incubated at 25°C for 10 minutes, 48°C for 30 minutes, followed by inactivation of the enzymes by a five-minute incubation at 95°C. The SYBR green master mix kit (Applied Biosystems) was used for RT-PCR. Each PCR contained 1 μL cDNA and 1 μM gene specific primers (Table 2) in a total volume of 25 μL. Samples were performed in triplicate on a BioRad iCycler with the following parameters: 10 minutes at 95°C, forty cycles of 15 sec at 95°C and 1 minute at 60°C. All genes were normalized to the pullulanase gene (Tomas, Welker et al. 2003).

2.6 Metabolic flux analysis

Metabolic Flux analysis calculations were performed using a program developed by Desai et al (Desai, Nielsen et al. 1999). Product concentrations from bioreactor experiments were used to generate metabolic fluxes. Error associated with the calculated fluxes is typically less than 10 percent.

2.7 Genomic DNA isolation

50 mL of A_{600} 1.0 cells were collected by centrifugation at 4^{o}C, washed with lysis buffer and frozen at -20^{o}C. Cells were incubated with lysozyme and RNase A for 10 minutes at 37^{o}C. Cell debris was removed by centrifugation and the supernatant was extracted twice with phenol:chloroform:isoamyl alcohol, once with chloroform:isoamyl alcohol. DNA was precipitated with isopropanol and washed with 70% ethanol. DNA was then air dried and resuspended in TE buffer.

2.8 PCR

The PCR amplification of specific DNA sequences was performed according to manufacturer's instructions (Applied Biosystems). In a total reaction volume of 100 μL, the following components were added: 10 μL PCR buffer II, 10 μL $MgCl_2$ (10mM), 6.6 μL dNTPs (10mM), 1 μL each primer (100 mM), 500 ng chromosomal DNA template, 0.5 μL Amplitaq Polymerase Gold, and the remaining volume H_2O. For colony PCR colonies grown for 24-48 hours were resuspended in 10 μL H20 and 2 μL of suspension was used as the template.

CHAPTER 3 : DIFFERENT COMBINATIONS OF HIGHER ALDEHYDE-ALCOHOL DEHYDROGENASE AND/OR THIOLASE EXPRESSION WITH COA TRANSFERASE DOWNREGULATION LEAD TO HIGHER ALCOHOL TITERS AND SELECTIVITY IN CLOSTRIDIUM ACETOBUTYLICUM FERMENTATIONS: FOCUS ON THE ACETYL-COA AND BUTYRYL-COA FLUX NODES

3.1 Introduction

Recent advances in molecular biology and metabolic engineering (ME) techniques of butyric-acid clostridia offer an opportunity to re-establish the acetone, butanol and ethanol (ABE) fermentation as an economically viable process. *Clostridium acetobutylicum* is a model and prototypical organism for the production of these commodity chemicals, and especially that of butanol, which has now emerged as an important new biofuel. The genome of *C. acetobutylicum* has been sequenced and annotated (Nölling, Breton et al. 2001), and methods for genetic deletions (Harris, Welker et al. 2002; Heap, Pennington et al. 2007; Shao, Hu et al. 2007) and gene overexpression (Mermelstein and Papoutsakis 1993) have been developed. Furthermore, genome-scale microarray-based transcriptional analyses (Tomas, Alsaker et al. 2003; Tomas, Beamish et al. 2003; Alsaker, Spitzer et al. 2004; Alsaker, Paredes et al. 2005; Alsaker and Papoutsakis 2005) have illuminated its complex metabolism, thus allowing the development of more precise ME strategies, by combining genetic modifications. Here we address ME issues related to flux determinism in core primary-metabolism pathways.

High butanol selectivity and titers in the ABE fermentation are current obstacles for an economical industrial process. Butanol is the most valuable product, and thus the production of all other products must be minimized. Ethanol is the only other product that may be desirable as a co-product in the context of biofuel production. The ABE batch fermentation is characterized by two distinct phases, the acidogenic and the solventogenic phase. Initially, the cultures produce the organic acids butyrate and acetate, which lower the culture pH. In the solventogenic phase, the culture produces butanol, acetone, and ethanol. Butyrate and acetate are partially re-assimilated to produce solvents, thus raising the pH of the culture. The trigger responsible for the switch from acid to solvent formation (known as solventogenesis) has been extensively studied, but the exact mechanism for this change remains unknown. The external pH is known to affect solventogenesis and product formation (Husemann and Papoutsakis 1988). Recent evidence correlates increases of butyryl-phosphate (BuP) concentration with the onset of solvent formation and suggests that BuP may play a role in the regulation of solvent initiation (Zhao, Tomas et al. 2005).

In WT *C. acetobutylicum* fermentations, final acetone concentrations are typically one-half the final levels of butanol. Initial efforts to increase the selectivity of butanol to acetone used antisense RNA (asRNA) technology targeting the transcripts of enzymes in the acetone formation pathway (Fig. 3.1). The *ctfB* asRNA successfully reduced acetone production when designed to downregulate a subunit of the first enzyme in the acetone formation pathway, CoA transferase (CoAT) (Tummala, Welker et al. 2003). However, butanol titers were also significantly reduced in the *ctfB* asRNA strain. The *ctfB* gene is part of a tricistronic operon (*aad-ctfA-ctfB*) also containing the *aad* gene, whose product, the bi-functional AAD (aldehyde-alcohol dehydrogenase) protein, catalyzes the two-step conversion of butyryl-CoA to butanol or

of acetyl-CoA to ethanol (Nair, Bennett et al. 1994). Because the *ctfB* and *aad* genes reside on the same mRNA transcript, the *ctfB* asRNA resulted in a downregulation of both the *ctfB* and *aad* genes thus resulting in lower butanol production (Tummala, Welker et al. 2003). Follow-up studies were able to restore WT butanol titer levels while maintaining low acetone production by combining, in strain 824(pAADB1) the *ctfB* asRNA with the overexpression of the *aad* gene alone off plasmid pAADB1 using its own autologous promoter (Tummala, Junne et al. 2003). Significantly, this strain produced ca. 200 mM ethanol, the highest ever in *C. acetobutylicum*. The high ethanol production is due to the dual functionality of the AAD enzyme, which catalyzes both the formation of ethanol and butanol. In the WT strain, butanol is produced nearly six-fold higher than ethanol. The logical and accepted interpretation is that AAD has a much higher affinity for butyryl-CoA than for acetyl-CoA (Fig. 3.1). The high ethanol production by strain 824(pAADB1) suggests that the ratio of acetyl-CoA to butyryl-CoA is much higher in this strain than in the WT strain. In this study we aimed to explore ME strategies to enhance butanol formation and selectivity and, significantly, to be able to accelerate butanol production. We focused on exploring the regulation of fluxes around the two critical nodes of butyryl-CoA and acetyl-CoA (Fig. 3.1). First, we sought to overexpress *aad* (AAD) early by changing the temporal expression of this gene using the *ptb* promoter (of the *ptb-buk* operon, coding the two enzymes responsible for butyrate production from butyryl-CoA; Fig. 3.1) which is expressed early in the acidogenic growth phase (Tummala, Welker et al. 1999) when the *aad* natural expression is normally absent (Nair, Bennett et al. 1994). This early expression of *aad* sought to direct more of the carbon flux towards butanol production while limiting the formation of butyrate by competing early for butyryl-CoA. The possibility of reducing ethanol and acetate

production by altering the fluxes around the acetyl-CoA node focusing on the overxpression of the thiolase gene (Fig. 3.1) was also investigated.

3.2 Materials and methods

3.2.1 Bacterial strains and plasmids

The list of bacterial strains and plasmids are in Table 3.1.

3.2.2 Plasmid and strain construction

The *aad* gene (CAP0162) responsible for butanol formation was PCR amplified from *C. acetobutylicum* genomic DNA using primers aad_fwd and aad_rev to exclude the natural promoter. All primers used in plasmid construction are listed in Table 3.2. The pSOS94 vector was digested with *Bam*HI and *Ehe*I and blunt ended to remove the acetone formation genes while leaving the *ptb* promoter region and the *adc* terminator. The *aad* PCR product and the linearized pSOS94 vector were ligated to create p94AAD3. Both pCTFB1AS, containing the *ctfB* asRNA, and p94AAD3 were digested with *Sal*I to linearize pCTFB1AS and isolate the *aad* gene with the *ptb* promoter and *adc* terminator from p94AAD3. These fragments were ligated together to generate pCASAAD.The thiolase (*thl*) gene including the endogenous promoter and terminator regions was amplified from *C. acetobutylicum* genomic DNA using primers thl_fwd and thl_rev. Following purification, the PCR product was digested with *Sal*I and *Eco*RI as was the shuttle vector pIMP1. The digested PCR product was ligated into the pIMP1 shuttle vector to form the plasmid pTHL. The *aad* gene cassette from p94AAD3 was isolated using a *Sal*I digestion and purified. Plasmid pTHL was *Sal*I digested and ligated with the purified *aad* gene cassette to generate plasmid pTHLAAD.

Strain or Plasmid	Relevant Characteristics[a]	Source or Reference[b]
Bacterial Strains		
C. acetobutylicum		
ATCC 824		ATCC
M5		(Clark, Bennett et al. 1989)
E. coli		
Top10		Invitrogen
ER2275		New England Biolabs
Plasmids		
pAN1	Cm^r, Φ3T I gene, p15A origin	(Mermelstein and Papoutsakis 1993)
pSOS94[c]	acetone operon (*ptb* promoter)	Soucaille and Papoutsakis, unpublished
p94AAD3[c]	*aad* (*ptb* promoter)	This study
pCTFB1AS[c]	*ctfB* asRNA (*thl* promoter)	(Tummala, Welker et al. 2003)
pCASAAD[c]	*aad* (*ptb* promoter), *ctfB* asRNA (*thl* promoter)	This study
pAADB1[c]	*aad* (*aad* promoter), *ctfB* asRNA (*thl* promoter)	(Tummala, Junne et al. 2003)
pTHL[c]	*thl*	This study
pTHLAAD[c]	*thl*, *aad* (*ptb* promoter)	This study
pPTBAAD[c]	*aad* (*ptb* promoter)	This study
pCAS[c]	*ctfB* asRNA (*adc* promoter)	This study
pSOS95del[c]	*thl* promoter	(Tummala, Junne et al. 2003)
pSS2[c]	*aad* (*ptb* promoter), *ctfB* asRNA (*adc* promoter), *thl*	This study

[a]Cm^r, chloramphenicol resistance gene; *ptb*, phosphotransbutyrylase gene; *aad*, alcohol/aldehyde dehydrogenase gene; *ctfB*, CoA transferase subunit B gene; *thl*, thiolase gene; *adc*, acetoacetate decarboxylase gene.

[b]ATCC, American Tissue Culture Collection, Rockville, MD

[c]contains the following: ampicillin resistance gene; macrolide, lincosimide, and streptogramin B resistance gene; *repL*, pIM13 Gram-positive origin of replication; ColE1 origin of replication

Table 3.1 Bacterial strains and plasmids used in this study.

Primer Name	Sequence (5'-3')	Description
aad_fwd	TTAGAAAGAAGTGTATATTTAT	*aad* forward primer
aad_rev	AAACGACGGCCAGTGAAT	*aad* reverse primer
thl_fwd	CCATATGTCGACGGAAAGGCTTCA	*thl* forward primer
thl_rev	ACGCCTAGTACTGAATTCGCCTCA	*thl* reverse primer
p_adc_top	TCGACTAAAAATTTACTTAAAAAAACATATGTGTTATAAT GAAATATAAATAAATAGGACTAGAGGCGATTTATAATGTG AAGATAAAGTATGTTAG	*adc* promoter top oligonucleotide
p_adc_bot	AATTCTAACATACTTTATCTTCACATTATAAATCGCCTCT AGTCCTATTTATTTATATTTACATTATAACACATATTGTT TTTTTAAGTAAATTTTTAG	*adc* promoter bottom oligonucleotide
ctfBas_top	AATTCTTAATTCTCTTGCAACTCTTTTGGCTATTATTTCT TTCGCTAGGTTTTTATCATTAATCATTTTATGCAGGCTCC TTAAAAGTAATTACATTACA	*ctfB* asRNA top oligonucleotide
ctfBas_bot	TATGATATGTAATTACTTTTAAGGAGCCTGCATAAAATGA TTAATGATAAAAACCTAGCGAAAGAAATAATAGCCAAAAG AGTTGCAAGAGAATTAAG	*ctfB* asRNA bottom oligonucleotide
cas_fwd	TCGACTAAAAATTTACTTAAAAAAAC	*ctfB* asRNA forward primer
cas_rev	TATGATATGTAATTACTTTTAAGG	*ctfB* asRNA reverse primer
aad_rt_fwd	AGAAAATGGCTCACGCTTCA	*aad* RT-PCR forward primer
aad_rt_rev	GCAATGCCAACTAGGAATATTGTG	*aad* RT-PCR reverse primer
pul_rt_fwd	TTCTCCACTGTGGCGTAGAGTT	*thl* RT-PCR forward primer
pul_rt_rev	TCTCTAAGATCCCAATCTATCCAATTT	*thl* RT-PCR reverse primer

Table 3.2 List of primers and oligonucleotides used in this study.

A revised *ctfB* asRNA cassette was generated by first inserting a 100 bp oligonucelotide into the pIMP1 shuttle vector following digestion with *Sal*I and *Eco*RI. This oligonucleotide includes the sequence for the *adc* promoter element with compatible nucleotide overhangs for ligation. The complimentary oligonucleotides p_adc_top andp_adc_bot were first annealed together before ligating into the pIMP1 vector, creating pPADC, which was then digested with *Eco*RI and *Nde*I. A second set of complementary oligonucleotides, ctfBas_top and ctfBas_bot, were annealed and ligated to the digested pPADC to form pCAS. The new *ctfB* asRNA cassette was PCR amplified from this plasmid using primers cas_fwd and cas_rev and ligated into the pTHLAAD plasmid to generate plasmid pSS2.

All plasmids were transformed into Top 10 chemically competent *E. coli* (Invitrogen, Carlsbad, CA). Plasmids were confirmed using sequencing reactions. The plasmids were methylated using *E. coli* ER2275 (pAN1) cells to avoid the natural restriction system of *C. acetobutylicum* (Mermelstein and Papoutsakis 1993). Once methylated, the plasmids were transformed by electroporating *C. acetobutylicum* wildtype (WT) or mutant M5 strains as described (Mermelstein, Welker et al. 1992).

3.3 Results

3.3.1 Early and elevated expression of *aad* using the *ptb* promoter

We aimed to explore ME strategies to enhance butanol formation and selectivity and accelerate butanol production. We focused on exploring the regulation of fluxes around the two critical nodes of butyryl-CoA and acetyl-CoA (Fig. 3.1). First, we sought to enhance and accelerate butanol and ethanol production by increasing the expression of the enzyme responsible for butanol and ethanol formation, alcohol/aldehyde dehydrogenase (AAD) (Fig.

3.1). This was accomplished by changing the temporal expression of this gene using the *ptb* promoter, p_{ptb} (of the *ptb-buk* operon, coding the two enzymes responsible for butyrate production from butyryl-CoA), which is active early in the acidogenic growth phase (Tummala, Welker et al. 1999) when the *aad* natural expression is normally absent (Nair, Bennett et al. 1994). This early expression of *aad* sought to direct more of the carbon flux towards butanol production while limiting the formation of butyrate by competing early for butyryl-CoA. The combination of the early-plasmid expressed *aad* with the later chromosomal expressed *aad* from its natural promoter should allow for the sustained butanol production throughout both the acidogenic and solventogenic growth phases. Plasmid p94AAD3 was created to express *aad* from the p_{ptb}. p94AAD3 was first transformed into the degenerate strain M5 (which has lost the pSOL1 megaplasmid and thus the ability to express the *sol* operon and form butanol or acetone (Cornillot, Nair et al. 1997)) to confirm the proper expression of *aad* and the production of a functional protein. Production of butanol in M5(p94AAD3) was observed (results not shown) confirming the proper expression and translation of the *aad* gene from p_{ptb}. Then, *aad* expressed from the p_{ptb} was isolated from p94AAD3 and combined into a plasmid containing the *ctfB* asRNA, creating plasmid pCASAAD. Following the transformation of pCASAAD into the WT strain, controlled pH 5.0 fermentations were performed in duplicate to fully characterize the 824(pCASAAD) compared to the strain containing the *ctfB* asRNA and *aad* overexpression from its endogenous promoter, 824(pAADB1), and the plasmid control strain 824(pSOS95del) (Fig. 3.2).

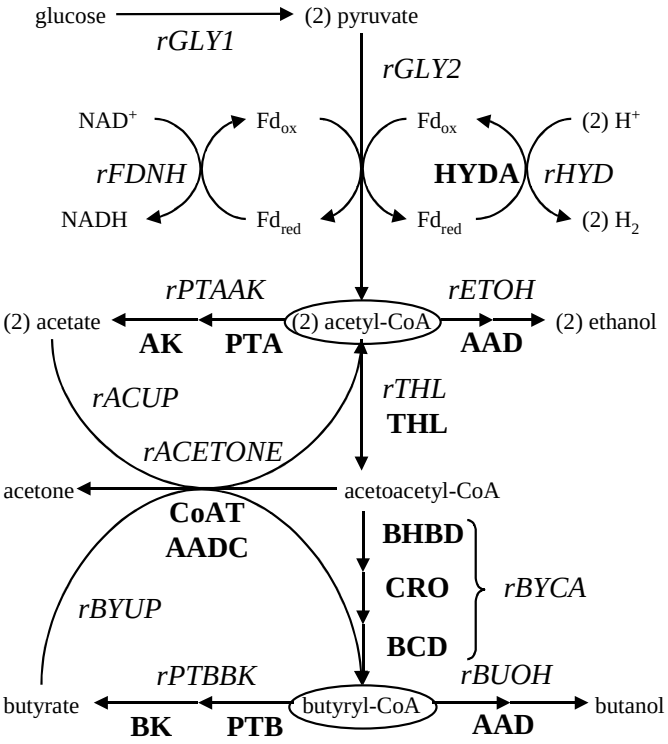


Figure 3.1 Metabolic pathways in *C. acetobutylicum* and associated calculated *in vivo* fluxes.

Selected enzymes are shown in bold and associated intracellular fluxes are shown in italics. The metabolic intermediates acetyl-CoA and butyryl-CoA are in ovals to highlight their importance in final product formation. Enzymes are abbreviated as follows: hydrogenase (HYDA); phosphotransacetylase (PTA); acetate kinase (AK); thiolase (THL); β-hydroxybutyryl dehydrogenase (BHBD); crotonase (CRO); butyryl-CoA dehydrogenase (BCD); CoA Transferase (CoAT); acetoacetate decarboxylase (AADC); butyrate kinase (BK); phophotransbutyrylase (PTB); alcohol/aldehyde dehydrogenase (AAD). Note: AAD is believed to be the primary enzyme for butanol and ethanol formation but additional genes exist that code for alcohol forming enzymes (*adhe2*, *bdhA*, *bdhB*, CAC3292, CAP0059).

RNA samples were collected during the fermentations and analyzed for the level of *aad* expression using Q-RT PCR. Comparing the *aad* expression between the strains, there exists a nearly ten-fold higher expression of *aad* in 824(pCASAAD) than in 824(pAADB1) during the first four timepoints (Fig. 3.3). These timepoints correspond to the exponential growth phase and the early transitional phase when the p_{ptb} is expected to have the highest activity. During the later timepoints *aad* expression continues to be higher in 824(pCASAAD), but at lower levels than initially observed. The expression of *aad* within each strain was also examined. In 824(pCASAAD), *aad* expression is highest during the first four timepoints after which the expression level decreases. This pattern is the opposite of the WT strain where *aad* expression is absent early, but is later induced in stationary phase (Alsaker and Papoutsakis 2005). This shows that the p_{ptb} was successful in enhancing the early expression of *aad*. The pattern of *aad* expression in 824(pAADB1) is more complex. There exists a distinct peak in expression of *aad* that corresponds to the entry into stationary phase, when *aad* is induced in the WT strain. After this point the *aad* expression begins to decrease.

3.3.2 *ptb*-promoter-driven *aad* expression leads to higher cell densities and increased, earlier butanol formation

Although the growth rate was similar between all strains, 824(pCASAAD) reached higher cell densities (Table 3.3) than either 824(pAADB1) or 824(pSOS95del). We attribute these higher cell densities to the lower butyrate concentrations observed in the 824(pCASAAD) strain (Fig. 3.2): butyrate was completely re-assimilated by both strains with the *ctfB* asRNA, but peak butyrate levels were reduced by two-thirds in 824(pCASAAD) compared to 824(pAADB1).

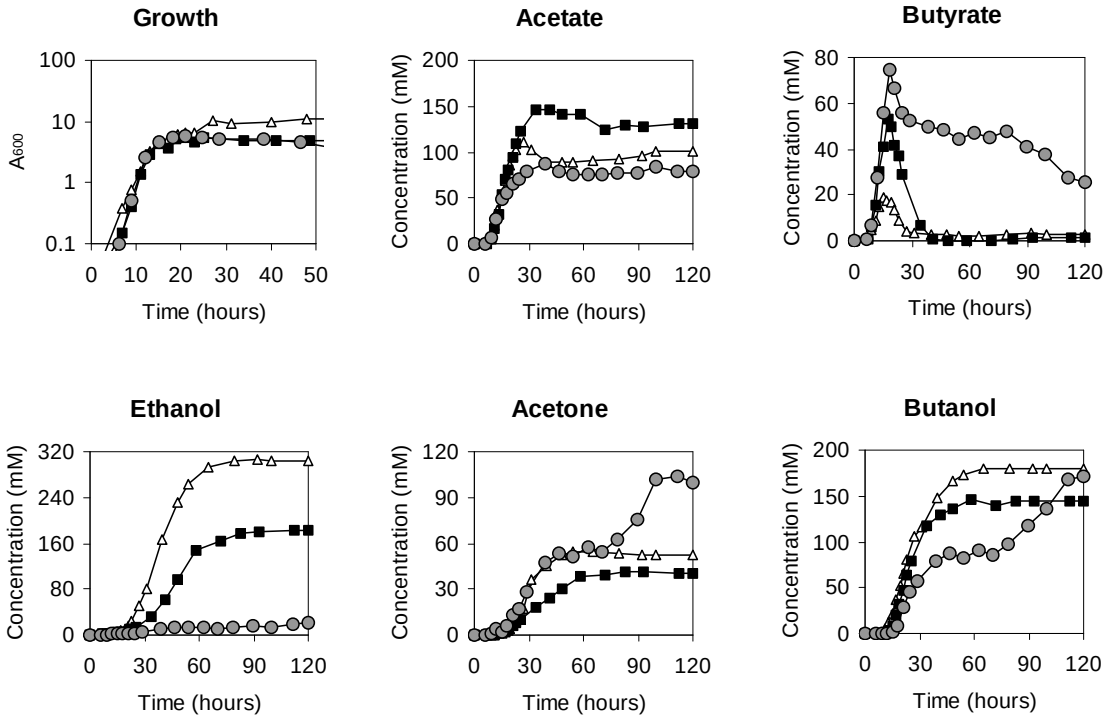


Figure 3.2 Growth and product concentrations of 824(pCASAAD), 824(pAADB1) and 824(pSOS95del) pH 5.0 fermentations.

Fermentations were performed in duplicate, while results are shown from one fermentation. Differences in product formation between duplicate fermentations are less than 5%. Lag times were standardized between fermentations by normalizing an A_{600} of 1.0 at hour 10 of the fermentation. 824(pCASAAD) results are shown as open triangles, 824(pAADB1) results are shown as closed squares, and 824(pSOS95del) results are shown as gray circles.

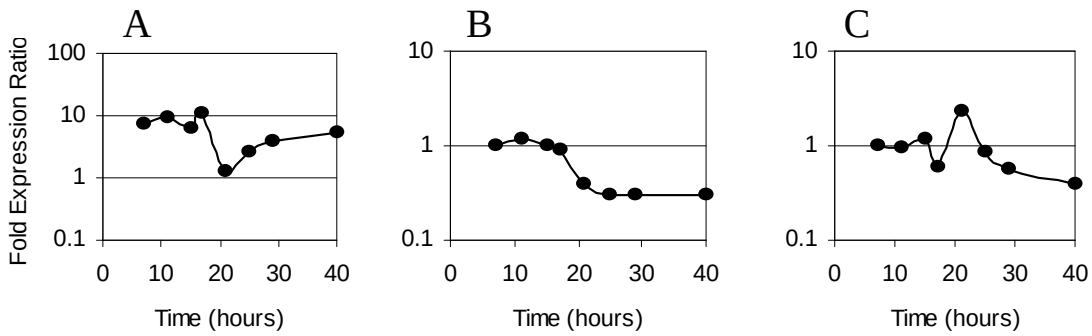


Figure 3.3 Q RT-PCR analysis of *aad* expression.

Samples were taken from bioreactor experiments shown in Figure 1. A. The ratio of *aad* expression in 824(pCASAAD) relative to 824(pAADB1) comparing similar timepoints. B. The ratio of *aad* expression in 824(pCASAAD) relative to the first timepoint sampled. C. The ratio of *aad* expression in 824(pAADB1) relative to the first timepoint sampled.

Furthermore, both the peak and final acetate levels were reduced in 824(pCASAAD) compared to 824(pAADB1): final acetate concentrations were 129 mM in 824(pAADB1), 85 mM in 824(pCASAAD) and 77 mM in the plasmid control strain 824(pSOS95del). The solvent formation profiles also show significant differences between strains. In the control 824(pSOS95del) strain, acetone and butanol are the primary solvents produced, 109 mM and 176 mM respectively, while ethanol formation is relatively minor, at ca. 20 mM. The acetone production of 824(pCASAAD) is slightly higher than in 824(pAADB1), but 824(pSOS95del) produces twice the acetone of 824(pCASAAD). Butanol production was higher in 824(pCASAAD) (178 mM) and 824(pSOS95del) than in 824(pAADB1) (146 mM). Significantly, butanol was produced earlier and reached its final levels in half the time in 824(pCASAAD) (ca. 60 hrs) than in 824(pSOS95del) (ca. 120 hrs) (Fig. 3.2). Earlier butanol formation is better demonstrated in the plots showing the specific intracellular fluxes (Fig. 3.4). Ethanol production was dramatically higher in 824(pCASAAD) and 824(pAADB1) than in 824(pSOS95del). 824(pCASAAD) produced 305 mM ethanol, 15 times higher than the control strain, while 824(pAADB1) produces 184 mM. This is the highest ethanol production reported by any solventogenic clostridium. These data show that butyryl-CoA and acetyl-CoA are important determinants of solvent yields and selectivities, and this is further illuminated by metabolic flux analysis, which is presented next.

3.3.3 Metabolic flux analysis of the three strains supports the limiting role of butyryl-CoA and acetyl-CoA for butanol vs. ethanol production, respectively

Using a previously developed model (Desai and Papoutsakis 1999), the fluxes of 824(pAADB1) and 824(pCASAAD) were calculated and normalized both for differences in lag

Strains	Fermentation characteristics[a]							
	Max A_{600}	Butanol	Ethanol	Acetone	$Acetate_{peak}$	$Acetate_{final}$	$Butyrate_{peak}$	$Butyrate_{final}$
824(pSOS95del)	5.79	176	19	109	80	77	73	37
824(pAADB1)	5.60	146	184	42	147	129	53	1
824(pCASAAD)	11.80	178	300	61	105	85	20	2
824(pTHL)	10.67	54	6	34	68	68	76	71
824(pTHLAAD)	9.75	153	28	98	68	67	39	17
824(pPTBAAD)	11.90	160	76	59	124	124	62	2
824(pSS2)	10.70	137	288	29	120	106	16	2
824(pCAS)[b]	3.59	53	11	29	38	38	45	34
824(pSOS95del)[b]	4.35	152	21	91	22	12	33	22

[a]All results shown are average mM concentration from duplicate experiments
[b]Results are from static flask experiments without pH control

Table 3.3 Product formation in pH 5.0 fermentation experiments

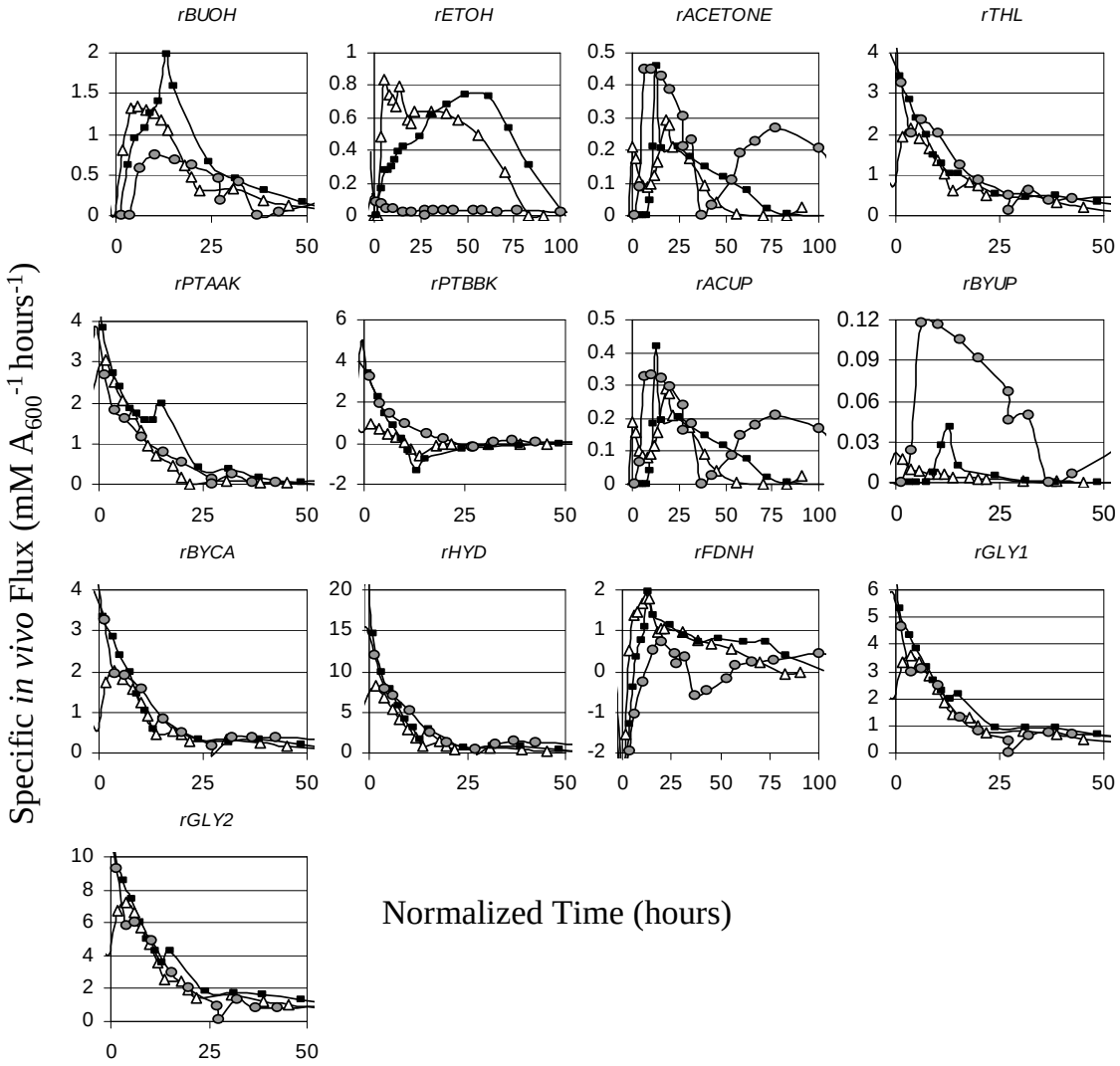


Figure 3.4 Metabolic flux analysis of 824(pCASAAD), 824(pAADB1) and 824(pSOS95del).

824(pCASAAD) results are shown as open triangles, 824(pAADB1) results are shown as closed squares, and 824(pSOS95del) results are shown as gray circles. Lag times were standardized between fermentations by normalizing an A_{600} of 1.0 at hour 10 of the fermentation.

times of growth and cell density. First, the core carbon fluxes GLY 1, GLY 2, thiolase and BYCA, and the H_2 formation flux were largely similar among the three strains (except for the first 3-4 normalized hours in strain 824(pCASAAD), which were lower, likely due to the metabolic burden of the early AAD overexpression), and thus unaffected by the genetic modifications, which is the theoretically expected and a desirable finding. The butanol and ethanol formation fluxes show significantly higher values early in 824(pCASAAD) than in 824(pAADB1) or the plasmid control. This is consistent with the observation that the FDNH fluxes ($NADH_2$ production from reduced ferredoxin coupled to the GLY 2 flux (Fig. 3.1)) show higher values earlier in the order (high to low) of 824(pCASAAD), 824(pAADB1) and 824(pSOSdel). In strain 824(pCASAAD) the butanol formation flux dropped to less than 25% its maximum at 21 hours, while the ethanol formation flux is maintained at over 50% its maximal value for nearly 60 hours. Although the peak formation values occur later in 824(pAADB1), the same trends of butanol and ethanol production are evident. In 824(pAADB1), the ethanol formation flux only reaches its maximum value after the butanol formation flux sharply decreases at 25 hours. Butanol formation precedes ethanol formation in all strains, but ethanol formation is sustained for longer time periods than butanol formation, and especially so after 25 hours when the flux (BYCA) to butyryl-CoA as well as the acetate and butyrate fluxes have largely been reduced to zero, while the GLY 2 flux (Fig. 3.1) remains still at reasonable levels (ca. 1 mM A_{600}^{-1} h^{-1}; Fig. 3.4). This suggests that butyryl-CoA availability limits butanol formation, while the substantial flux to acetyl-CoA combined with high levels of AAD expression feeds and sustains the flux to ethanol at high levels thus leading to the very high ethanol titers. The nearly zero flux of acetate formation after 25 hours and the sustained high ethanol fluxes late past 50 to 60 hours suggest that the high AAD activity combined with the

likely lower activities of the acetate formation enzymes (based on transcriptional information, data not shown) are responsible for channeling most of the carbon to ethanol by utilizing all the available reducing paper, and thus the zero H_2 formation flux at that time period.

The butyrate formation flux is particularly low in 824(pCASAAD), thus demonstrating that the strategy for channeling butyryl-CoA from butyrate to butanol formation by the early and strong *aad* overexpression has worked as anticipated. Due to the low butyrate formation, butyrate uptake is much lower in 824(pCASAAD). Acetate formation is also sustained better and longer in 824(pAADB1) than in 824(pCASAAD) and the plasmid-control strain, and this is consistent with the deduced longer sustained acetyl-CoA pool that sustains much longer a high ethanol flux. Comparing the acid uptake fluxes with the acetone formation flux, it is evident that acetone is produced almost solely from the uptake of acetate, as the acetate uptake flux is 10-fold higher than the butyrate uptake flux in both 824(pAADB1) and 824(pCASAAD). Acetone formation is also sustained longer in 824(pAADB1) than in 824(pCASAAD), but both strains show the anticipated lower acetone fluxes compared to the plasmid-control strain as a result of the asRNA dowregulation of the acetone-formation enzyme CoAT (Figs. 3.1 and 3.4).

3.3.4 Role of thiolase promoter and thiolase expression on the acetyl-CoA to butyryl-CoA flux, and its impact on product formation

The data discussed above (Figs. 3.2 and 3.4) suggest there exists a bottleneck for the formation of butyryl-CoA from acetyl-CoA. Four enzymes catalyze the conversion of butyryl-CoA from acetyl-CoA (Fig. 3.1). These are organized into two operons on the chromosome. The first enzyme in the pathway, thiolase (coded by the monocistronic *thl*), converts acetyl-CoA to acetoacetyl-CoA. The other three enzymes, β-hydroxybutryl-coenzyme A dehydrogenase

(BHBD), crotonase (CRO), and butyryl-CoA dehydrogenase (BCD), convert acetoacetyl-CoA to butyryl-CoA and are co-transcribed as a single, large operon (Boynton, Bennett et al. 1996). *thl* is expressed at high levels. Its constitutive-like expression (Tummala, Welker et al. 1999) makes it an ideal promoter (p_{thl}) for high expression studies in clostridia, and was thus used to drive the expression of the *ctfB* asRNA in these studies. It is possible that the use of p_{thl} for the *ctfB* asRNA could have lowered the expression of the endogenous *thl* gene, thereby lowering THL activity and creating the acetyl-CoA buildup. In order to test this hypothesis, the *ctfB* asRNA was expressed from another promoter, namely the promoter of the acetoacetate decarboxylase (Fig. 3.1) gene *adc* (p_{adc}), which is also highly expressed but a little later than p_{thl} (Tummala, Welker et al. 1999), and since it is used in the formation of acetone, any promoter titration effects would not negatively impact butanol formation. A 100-basepair oligonucleotide was designed to include the integral portions of the *adc* promoter (Gerischer and Dürre 1990). Following its ligation into the pIMP1 shuttle vector, another oligonucleotide was designed to include the Shine-Delgarno sequence of the *ctfB* gene and ca. 50 basepairs of the downstream coding sequence. Following the *ctfB* sequence was the *glnA* hairpin terminator (Desai and Papoutsakis 1999). This new *ctfB* asRNA was ligated downstream of p_{adc} to create the plasmid pCAS (Table 3.1). This plasmid was transformed into the WT strain to confirm the functionality of the new *ctfB* asRNA. This new strain (824(pCAS)) was characterized in static fermentations and compared to the original *ctfB* asRNA. Acetone formation was much lower in 824(pCAS) than in the 824(pSOS95del) control strain (Table 3.3). 824(pCAS) also has low overall solvent formation and higher acid formation with limited acid reassimilation compared to 824(pSOS95del). These results are consistent with the previous *ctfB* asRNA strain (Tummala, Junne et al. 2003; Tummala, Welker et al. 2003).

To determine if low THL levels were limiting the conversion of acetyl-CoA to butyryl-CoA in the WT strain without AAD overexpression, the *thl* gene including its endogenous promoter was amplified from genomic DNA and ligated into the pIMP1 shuttle vector to create plasmid pTHL. Following the transformation of this plasmid into the WT strain, pH controlled bioreactors were used to characterize the strain. The metabolism of the 824(pTHL) is characterized by initial levels of high acid production, typical in clostridial fermentations, but there is only very limited acid reassimilation (Table 3.3). Along with the elevated levels of acid production, there is a dramatic decrease in the levels of solvents produced. Additionally, there is a sharp decrease in the cell density of the culture and a plateau of the glucose uptake just a few hours following the peak butyrate production. This may indicate that the cells cannot reassimilate butyrate promptly and the solvent genes cannot be induced to respond to the butyrate production, which leads to growth inhibition. We hypothesized that *aad* overexpression using p_{ptb} would promote early butanol production and a means for preventing the accumulation of inhibitory butyrate concentrations.

Overexpression of AAD using p_{ptb} was analyzed with (strain 824(pTHLAAD)) and without (strain 824(pPTBAAD)) *thl* overexpression, and the fermentation data from the two strains are summarized in Table 3.3. As a result of AAD overexpression, ethanol levels in 824(pPTBAAD) increased to 76 mM, more than three times the WT production. Additionally butyrate was nearly completely re-assimilated by this strain, while the final butanol titer was 160 mM. Acetate production in 824(pPTBAAD) was also very high reaching final levels of 124 mM. With the addition of THL overexpression, 824(pTHLAAD) shows a significant shift in product formation compared to 824(pPTBAAD). Ethanol production is reduced from 76 mM in 824(pPTBAAD) to 28 mM in 824(pTHLAAD). Acetate formation in 824(pTHLAAD) is also

reduced to nearly half the level of 824(pPTBAAD). Butanol is produced at similar levels in both strains while THL overexpression causes a small increase in butyrate formation. Acetone levels were ca. 40% higher in 824(pTHLAAD) compared to 824(pPTBAAD).

Comparing the profiles of the different fluxes (Fig. 3.5) provides additional insight into the role of THL overexpression. Consistent with the THL overexpression, there is an increase in the thiolase flux in 824(pTHLAAD) from about 5 hours to 30 hours (notice that the time scale in the flux analysis is in normalized hours) compared to 824(pPTBAAD). The higher butanol and BYCA fluxes early (in normalized hours) in the fermentation show that butanol is produced earlier in 824(pTHLAAD) than in 824(pPTBAAD) apparently because THL overexpression can enhance the butyryl-CoA rate of formation. The ethanol formation flux is similar between the two strains until about 25 hours into the fermentation when the flux is sharply reduced to zero at 40 hours in 824(pTHLAAD), while the ethanol formation flux is sustained at a high level in 824(pPTBAAD) after 50 hours. The HYD and FDNH fluxes are not affected by *thl* overexpression. Comparing the acid formation fluxes there appears to be little difference, but the acid uptake fluxes are significantly increased in 824(pTHLAAD) compared with 824(pPTBAAD): the acetate uptake flux is nearly twice as high in 824(pTHLAAD) and is sustained longer than in 824(pPTBAAD), while the butyrate uptake flux has a similar magnitude, but is sustained longer in 824(pTHLAAD). The acetone formation flux follows a similar pattern as the acetate uptake flux showing that acetone formation is mostly due to acetate uptake. Significantly, except for the first few hours, the BYCA flux is identical between the two strains. These flux analysis data then show that THL overexpression enhances acetoacetyl-CoA formation, which enhances acetone formation and acetate uptake. Except for very early in the

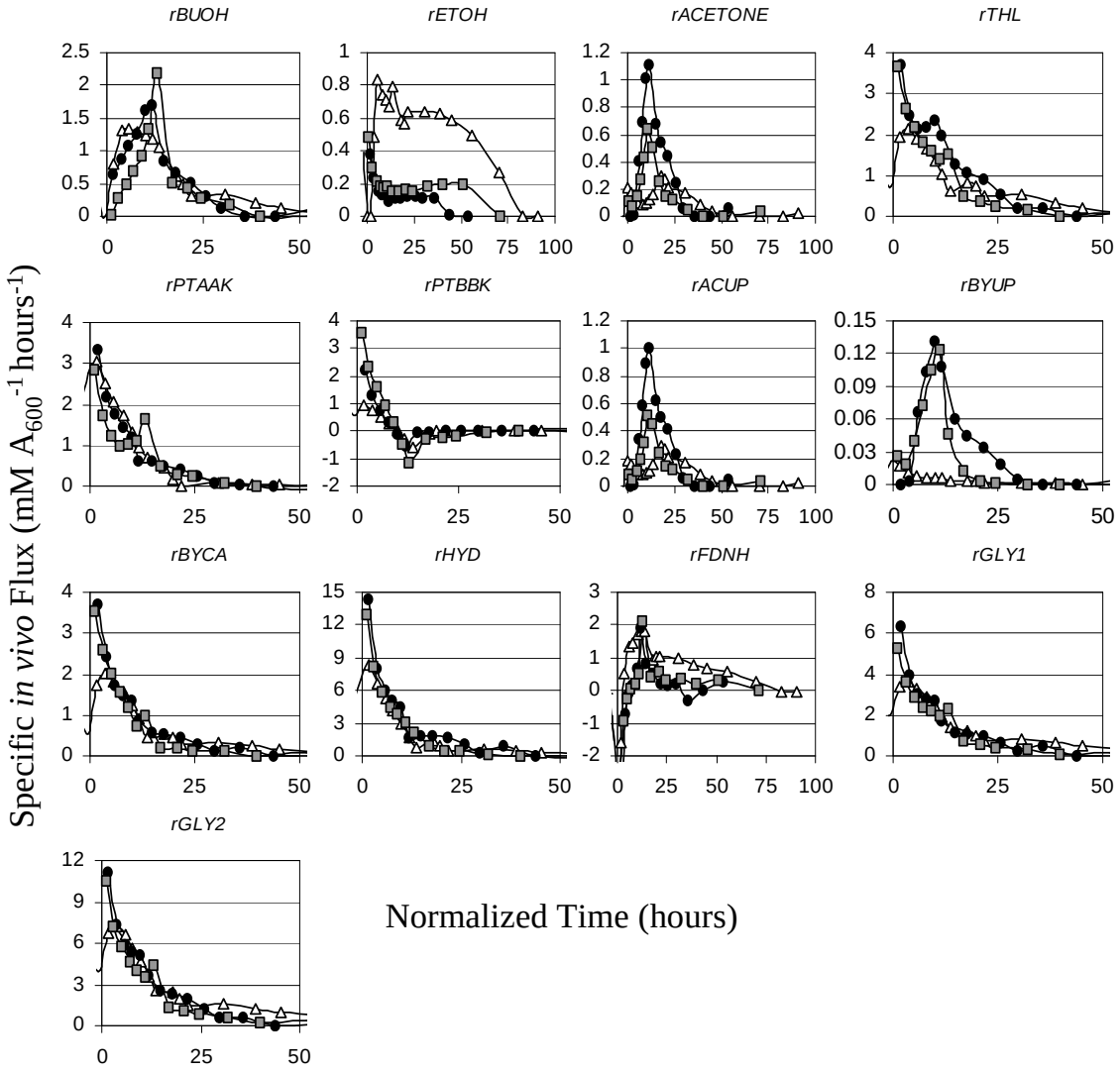


Figure 3.5 Metabolic Flux Analysis of 824(pTHLAAD), 824(pPTBAAD) and 824(pSOS95del).

824(pTHLAAD) results are shown as closed circles, 824(pPTBAAD) results are shown as grey squares, and 824(pCASAAD) results are shown as open triangles. Lag times were standardized between fermentations by normalizing an A_{600} of 1.0 at hour 10 of the fermentation.

fermentation (in normalized hours), the lack of a major impact on the BYCA flux suggests that that flux is limited by one of the HBD, CRO or BCD enzymes (Fig. 3.1). Thus, *thl* overexpression achieved in principle the goal of reducing the acetyl-CoA pool and thus reduce the formation of ethanol and acetate. Indeed, in 824(pPTBAAD) the ratio of the concentrations of the two carbon products (ethanol and acetate) to the four carbon products (butanol and butyrate) was 0.81. When THL was overexpressed with AAD in 824(pTHLAAD), this ratio more than doubled to 1.79.

A comparison of the fermentation data (Table 3.3) from strains 824(pPTBAAD) and 824(pCASAAD) illuminates the impact of the asRNA CoAT (Fig. 3.1) downregulation. This is analyzed in detail below.

3.3.5 Combined effect of *thl* and *aad* overexpression with CoAT dowregulation

Plasmid pSS2 (Table 3.1) was constructed to combine THL, AAD (from the p_{ptb}) overexpression, and CoAT downregulation by asRNA, but for the latter using the p_{adc} instead of the p_{thl} used in the pCASAAD and pAADB1 plasmids. pH controlled fermentations of strain 824(pSS2) were once again used to characterize the strain in order to compare to the 824(pCASAAD) and 824(pTHLAAD) strains (Fig. 3.6). Strain 824(pSS2) grew a little slower than either 824(pCASAAD) or 824(pTHLAAD) and product formation was delayed even when normalized for differences in lag times; this is probably due to a general metabolic burden by the larger plasmid. Peak acetate production in 824(pSS2) was similar to 824(pCASAAD), but final acetate concentrations were higher. Butyrate formation was nearly identical in 824(pSS2) compared with 824(pCASAAD), which has lower peak and final butyrate levels than 824(pTHLAAD). Ethanol formation at 288 mM was very high in 824(pSS2), nearly as high as

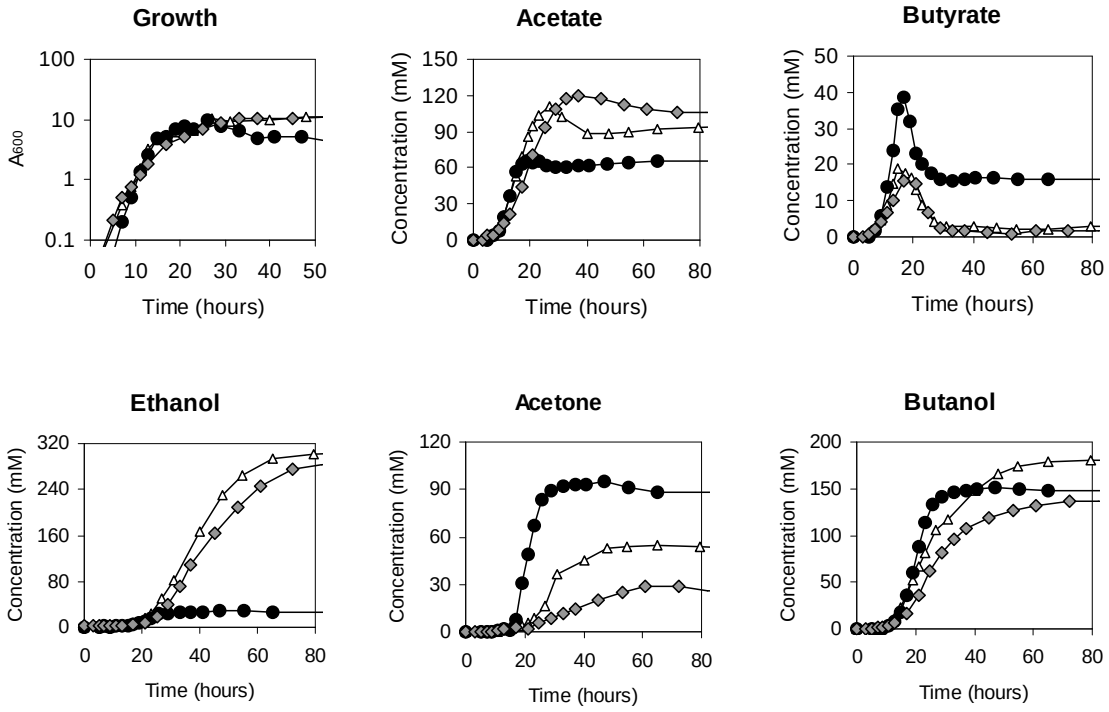


Figure 3.6 Growth and product concentrations of 824(pCASAAD), 824(pTHLAAD) and 824(pSS2) pH 5.0 fermentations.

Fermentations were performed in duplicate, while results are shown from one fermentation. Differences in product formation between duplicate fermentations are less than 5%. Lag times were standardized between fermentations by normalizing an A_{600} of 1.0 at hour 10 of the fermentation. 824(pCASAAD) results are shown as open triangles, 824(pTHLAAD) results are shown as closed circles, and 824(pSS2) results are shown as gray diamonds.

in 824(pCASAAD), which is much greater than ethanol production in 824(pTHLAAD). Acetone levels are much lower in the two strains harboring the *ctfB* asRNA, while butanol levels were fairly similar across all strains the lowest levels are achieved in 824(pSS2).

824(pSS2) shows a more similar profile to 824(pCASAAD), which does not overexpress *thl* (but produces somewhat higher butanol and acetone levels) than 824(pTHLAAD), which does overexpress *thl*. These results indicate that the *ctfB* asRNA combined with *aad* overexpression provide the dominant phenotype (high butanol and ethanol formation with suppressed acetone formation) that additional *thl* expression is unable to modulate in terms of enhancing butanol formation. It can be also argued that the p_{adc} driven asRNA CoAT downregulation has the desirable outcome, namely in producing a large suppression of acetone formation (which is fractionally larger than the suppression of either butanol or ethanol formation; compare the profiles of strain 824(pSS2) and 824(pCASAAD) in Fig. 3.6).

If one compares strains 824(pSS2) and 824(pTHLAAD), the impact of CoAT downregulation in the former is expected in that it reduces acetone formation, but unexpected in that it dramatically enhances ethanol and acetate formation apparently due to an increased acetyl-CoA pool. The same conclusion is drawn when one compares strains 824(pPTBAAD) and 824(pCASAAD) (Table 3.3): CoAT downregulation enhances dramatically ethanol formation but is accompanied by a lower final acetate production. An explanation of these phenotypic difference is a bit more complex thus necessitating a comparison of the fluxes between these two strains is shown in Fig. 3.5 pCASAAD has much higher ethanol and butanol formation fluxes, lower rTHL fluxes, dramatically lower acetate (rACUP) and butyrate (rBYUP) uptake fluxes, altered rFDNH, and altered acetate formation fluxes (higher early, lower later), all of which point to altered regulation around the acetyl-CoA node.

3.4 Discussion

We altered the pattern of *aad* expression by replacing the endogenous promoter with that of *ptb*, which is responsible for butyrate formation. This caused both earlier and higher expression of *aad* and had dramatic effects on the fermentation products (Figs. 3.2, 3.3 & 3.4). All the solvents (acetone, butanol, and ethanol) were produced at higher levels in 824(pCASAAD) with p_{ptb} driven *aad* expression than in 824(pAADB1), which uses the native *aad* promoter. The use of the *ctfB* asRNA kept acetone concentrations low, while ethanol concentrations reached the highest levels observed with this organism. The total solvent produced of 824(pCASAAD) is over 30 g/L with 13-14 g/L each of butanol and ethanol. WT fermentations only produce about 20 g/L solvents, but acetone and butanol are the primary products. Additionally, butyrate is not totally reassimilated by the WT strain as it is in 824(pCASAAD). Final acid levels can be 15-25% of the total products in WT *C. acetobutylicum* fermentations, but in 824(pCASAAD) fermentations acids are only 5-10% of the total products. Other high solvent producing clostridia strains have been engineered that produced between 25-29 g/L total solvents in batch cultures, but again, the primary products are butanol and acetone (Harris, Desai et al. 2000; Qureshi and Blaschek 2001; Tomas, Welker et al. 2003). With the emergence of biofuels, strains producing ethanol may be preferred over those producing acetone as other significant products.

Metabolic flux analysis showed that the earlier expression of *aad* resulted in earlier formation of both butanol and ethanol. It also appears that butyryl-CoA depletion may lead to the high ethanol yields. Ethanol production only becomes significant as butanol production decreases due to reduced availability of butyryl-CoA. As the same enzyme (AAD; Fig. 3.1) catalyzes butanol and ethanol formation, genomic manipulations to directly decrease ethanol

formation cannot be achieved by a simple metabolic engineering strategy. To further increase the butanol titers one must divert more of the acetyl-CoA (the precursor of ethanol) to butyryl-CoA. Thiolase (THL) is the first enzyme in the conversion of acetyl-CoA to butyryl-CoA and its role in solvent production was investigated. THL overexpression combined with AAD overexpression lowered production of acetate and ethanol, while increasing acetone and butyrate levels. The *ctfB* asRNA used in the earlier studies was also redesigned to eliminate the use of the *thl* promoter to alleviate concerns of transcription factor titration effects. Although the combined THL and AAD overexpression does produce a substantial shift in the fermentation products, the combination of THL and AAD overexpression with CoAT asRNA downregulation does not significantly alter product formation compared to the strain without THL ovexpression. This indicates that with fewer genetic manipulations *thl* gene dosage may be an important contributor to the final end products, but the combination of *ctfb* asRNA and AAD overexpression may result in internal metabolite concentrations too constrained for THL overexpression to impart any benefit. In particular, the BYCA flux appears resistant to change by THL overexpression (Fig. 3.6; compare strain 824(pSS2) to strain 824(pCASAAD); flux data are not shown), thus possibly necessitating the overexpression of one of the BHBD, CRO or BCD enzymes (Fig. 3.1). This is a rather complex undertaking for two reasons: first, the size of the plasmid required for overexpressing some or all of these genes together and the genes or RNAs already on plasmids pSS2 or pCASAAD will be very large and will impose a large metabolic burden on the cells. Second, it is not clear which of the three genes of the BYCA linear pathway is rate limiting and especially because additional electron transport proteins may be necessary (Boynton, Bennett et al. 1996). Nevertheless, this is a problem we plan to pursue in our lab in the quest for better

strains to produce butanol. At the same time, a fresh approach is necessary in order to be able to exercise control over the ratio of butanol to ethanol produced.

CHAPTER 4 : METABOLIC ENGINEERING OF THE NON-SPORULATING, NON-SOLVENTOGENIC CLOSTRIDIUM ACETOBUTYLICUM STRAIN M5 TO PRODUCE BUTANOL WITHOUT ACETONE DEMONSTRATES THE ROBUSTNESS OF THE ACID-FORMATION PATHWAYS AND THE IMPORTANCE OF THE ELECTRON BALANCE

4.1 Introduction

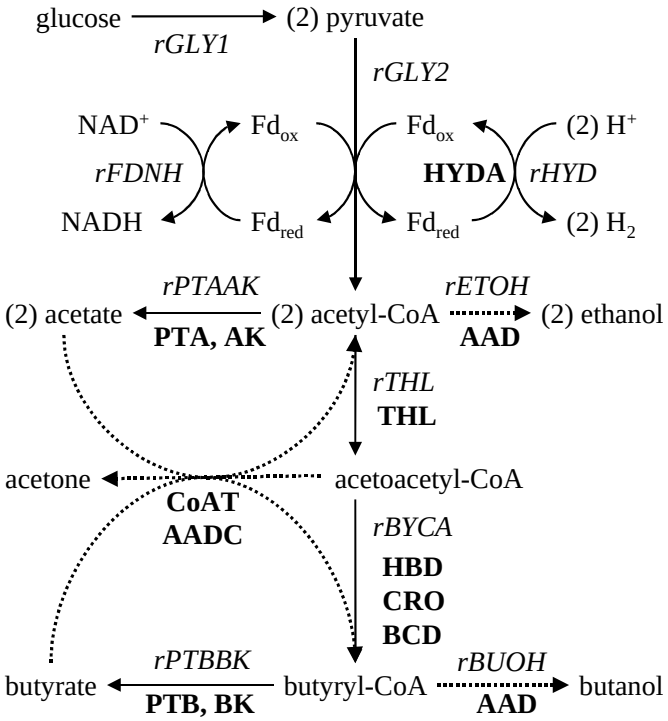
Clostridium acetobutylicum is an anaerobic, spore-forming prokaryote that produces the solvents butanol, acetone, and ethanol. Increased interest in the development of biologically based chemicals and fuels has generated renewed attention in this industrially important strain. The desired product of the *C. acetobutylicum* fermentation is butanol, which has superior fuel characteristics to ethanol, such as higher energy content and lower water miscibility. The *C. acetobutylicum* genome has been sequenced and annotated (Nölling, Breton et al. 2001), and methods for genetic deletions (Harris, Welker et al. 2002; Heap, Pennington et al. 2007; Shao, Hu et al. 2007) and gene overexpression (Mermelstein and Papoutsakis 1993) have been developed, making it an attractive organism for further strain development. Clostridia can also grow on a variety of substrates, from simple pentoses and hexoses to complex polysaccharides (Jones and Woods 1986).

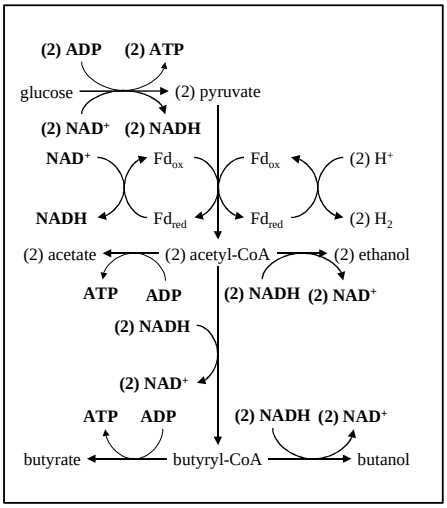
The metabolism of *C. acetobutylicum* is typically biphasic in batch culture: the cells first produce acetate and butyrate and later butanol, acetone, and ethanol. During growth, the production of acids lowers the pH of the culture, which combined with butyrate accumulation

(Husemann and Papoutsakis 1988) causes a shift in metabolism towards solvent production. As solvents are produced, the acids are typically reassimilated and converted into solvents. With initiation of solvent formation, the cells commit to their sporulation program. In continuous culture or upon consecutive vegetative transfers, cells may degenerate whereby they become asporogenous and lose the capability to produce solvents. In this organism, the degeneration process is due to the loss of the pSOL1 megaplasmid (Cornillot, Nair et al. 1997) which carries the key solvent formation genes in the so called *sol* locus made up of the *sol* operon (*aad-ctfA-ctfB*) (coding for the enzymes AAD, and CoAT; Fig. 4.1) and the *adc* gene (coding for the enzyme AADC; Fig. 4.1). From the practical point of view, a solvent-producing, non-sporulating strain is most desirable because it is known that solventogenic clostridia produce solvents only during a rather narrow window of their sporulation program, namely during and/or the near stage of the characteristic, cigar-shaped, clostridial-cell form (Jones and Woods 1986). The rest of the mixed-cell population does not likely contribute to solvent production, and this limits the specific cell productivity. Significantly, sporulating cells are not suitable for continuous or semicontinuous (fed-batch) fermentations due to the commitment to sporulation. Solvent producing non-sporulating strains for other clostridia have been derived by random mutagenesis (Lemme and Frankiewiez 1985; Jain 1993). For *C. acetobutylicum*, chemical mutagenesis might not be an option, because it has been reported (Clark, Bennett et al. 1989) to lead to strain degeneration (Cornillot, Nair et al. 1997); significantly, strains generated by random mutagenesis cannot be easily improved further by targeted metabolic engineering strategies. It has been shown however, that asporogenous strains (M5 and DG1) of *C. acetobutylicum* which have lost the pSOL1 megaplasmid can be complemented by plasmids carrying the *aad* gene (expressed

Figure 4.1 Metabolic pathways in *C. acetobutylicum* and associated calculated *in vivo* fluxes.

Selected enzymes are shown in bold and associated intracellular fluxes are shown in italics. Enzymes are abbreviated as follows: hydrogenase (HYDA); phosphotransacetylase (PTA); acetate kinase (AK); thiolase (THL); β-hydroxybutyryl dehydrogenase (BHBD); crotonase (CRO); butyryl-CoA dehydrogenase (BCD); CoA Transferase (CoAT); acetoacetate decarboxylase (AADC); butyrate kinase (BK); phophotransbutyrylase (PTB); alcohol/aldehyde dehydrogenase (AAD). Note: AAD is believed to be the primary enzyme for butanol and ethanol formation but additional genes exist that code for alcohol forming enzymes (*adhe2, bdhA, bdhB*, CAC3292, CAP0059). The pathways whose genes reside on the pSOL1 megaplasmid and are absent in M5 are shown as dotted lines. The boxed pathway shows the ATP generation and NADH production occurring during metabolism.


glucose
(2) pyruvate
rGLY1
rGLY2
NAD+
Fdox
Fdox
(2) H+
rFDNH
HYDA
rHYD
NADH
Fdred
Fdred
(2) H2
rPTAAK
rETOH
(2) acetate
(2) acetyl-CoA
(2) ethanol
PTA, AK
AAD
rTHL
THL
acetone
acetoacetyl-CoA
CoAT
AADC
rBYCA
HBD
CRO
BCD
rPTBBK
rBUOH
butyrate
butyryl-CoA
butanol
PTB, BK
AAD


(2) ADP
(2) ATP
glucose
(2) pyruvate
(2) NAD+
(2) NADH
NAD+
Fdox
Fdox
(2) H+
NADH
Fdred
Fdred
(2) H2
(2) acetate
(2) acetyl-CoA
(2) ethanol
ATP
ADP
(2) NADH
(2) NAD+
(2) NADH
(2) NAD+
ATP
ADP
(2) NADH
(2) NAD+
butyrate
butyryl-CoA
butanol

from its autologous promoter) and this leads to butanol formation, albeit at relatively low levels (Nair and Papoutsakis 1994; Cornillot, Nair et al. 1997), but no acetone production. Here we examined whether such strains can be made to produce higher levels of butanol similar to those of the WT strain. Such strains would be very desirable for an additional important reason: they would not produce acetone, which is not a desirable product, thus significantly raising the selectivity of the process for butanol, and thus increasing its economic appeal.

In this study we have used the non-sporulating, non-solventogenic strain M5 (Clark, Bennett et al. 1989) as a platform in order to generate strains that produce only butanol. First, the strain was transformed with a plasmid to overexpress the *aad* gene from an alternate promoter that bypasses the endogenous solvent gene regulation. Efforts to further enhance butanol production were investigated using thiolase (*thl*; Fig. 4.1) overexpression, and inactivation of the acid formation pathways, individually, by gene knockout (KO). From the fundamental point of view, this study also addresses the key question as to whether other pSOL1 genes are necessary for efficient cell growth and solvent production, and also if elimination of one of the acetate or butyrate formation pathways (Fig. 4.1) can be practically achieved in the M5 genetic background. A key tool for pursuing these goals is the detailed metabolic flux analysis of the cell's primary metabolism (Fig. 4.1), as has been developed and extensively validated by our group over the years (Papoutsakis 1984; Desai, Nielsen et al. 1999; Harris, Desai et al. 2000; Harris, Blank et al. 2001).

4.2 Materials and Methods

4.2.1 Bacterial strains and plasmids

The list of bacterial strains and plasmids are in Table 4.1.

4.2.2 Strain and plasmid construction

The *aad* gene (CAP0162) was PCR amplified from *C. acetobutylicum* genomic DNA using primers (aad_fwd and aad_rev) to exclude the natural promoter. All primers and oligonucleotides used are shown in Table 4.2. The pSOS94 vector was digested with *Bam*HI and *Ehe*I and blunt ended, leaving the *ptb* promoter region (*ptb* codes for the first enzyme, PTB, of the butyrate formation from butyryl-CoA; Fig. 4.1). The *aad* PCR product and the linearized pSOS94 vector were ligated to create p94AAD3. The thiolase gene (*thl*) with its natural promoter was PCR amplified from genomic DNA (thl_fwd and thl_rev) with *Sal*I and *Eco*RI sites designed into the primers. The pIMP1 plasmid and the thiolase gene were digested with *Sal*I and *Eco*RI and ligated to create pTHL. Both pTHL and p94AAD3 were digested with *Sal*I to linearize pTHL and isolate the *aad* gene with the *ptb* promoter from p94AAD3. These fragments were ligated together to produce pTHLAAD.

Partial acetate kinase (*ack*; 341 bp) and butyrate kinase (*buk*: 446 bp) gene fragments were PCR amplified using the akko_fwd and akko_rev primers or the bkko_fwd and bkko_rev primers, respectively. Following PCR, the fragments were cloned into the pTOPO-GW8 cloning vector (Invitrogen, Carlsbad, CA) per manufacturer's directions. Following confirmation of positive clones, the vectors containing the *ack* or *buk* fragment were digested with *Dra*I or *Nsi*I, respectively. A modified chloramphenicol marker was generated by PCR amplifying a 687 bp region pLHKO with mod-CM/SDG-F and mod-CM/SDG-R primers. This fragment was then ligated into *Bam*HI and *Kas*I digested pSOS94. The new chloroamphenicol-thiamphenicol resistance gene was PCR amplified using primers pSOS94_fwd and pSOS94_rev. The fragment was ligated into the *Dra*I or *Nsi*I digested vectors described above to create pAKKO and pBKKO, respectively.

Strain or Plasmid	Relevant Characteristics[a]	Source or Reference[b]
Bacterial Strains		
C. acetobutylicum		
M5	pSOL-	(Clark, Bennett et al. 1989)
M5 AKKO	pSOL-, *ack*-	This study
M5 BKKO	pSOL-, *buk*-	This study
E. coli		
Top10		Invitrogen
ER2275		New England Biolabs
Plasmids		
pAN1	Cm^r, Φ3T I gene, p15A origin	(Mermelstein and Papoutsakis 1993)
pSOS94[c]	acetone operon (*ptb* promoter)	Soucaille and Papoutsakis, unpublished
p94AAD3[c]	*aad*, (*ptb* promoter)	This study
pTHLAAD[c]	*thl*, *aad*, (*ptb* promoter)	This study
pLHKO[c]	Cm^r	(Harris 2001)
pGW8-TOPO		Invitrogen
pSOS94-CMptb[c]	Cm^r (*ptb* promoter)	This study
pAKKO	*ack* gene fragment, Cm^r (*ptb* promoter)	This study
pBKKO	*buk* gene fragment, Cm^r (*ptb* promoter)	This study

[a]Cm^r, chloramphenicol/thiamphenicol resistance gene; *ptb*, phosphotransbutyrylase gene; *aad*, alcohol/aldehyde dehydrogenase gene; *thl*, thiolase gene; *adc*, acetoacetate decarboxylase gene; *ack*, acetate kinase gene; *buk*, butyrate kinase gene.

[b]ATCC, American Tissue Culture Collection, Rockville, MD

[c]contains the following: ampicillin resistance gene; macrolide, lincosimide, and streptogramin B resistance gene; *repL*, pIM13 Gram-positive origin of replication; ColE1 origin of replication

Table 4.1 Bacterial strains and plasmids used in this study.

Primer Name	Sequence (5'-3')	Description
aad_fwd	TTAGAAAGAAGTGTATATTTAT	*aad* forward primer
aad_rev	AAACGACGGCCAGTGAAT	*aad* reverse primer
thl_fwd	CCATATGTCGACGGAAAGGCTTCA	*thl* forward primer
thl_rev	ACGCCTAGTACTGAATTCGCCTCA	*thl* reverse primer
mod-CM/SDG-F	CCGGATCCACTTGAATTTAAAAGGAGGGAACTTAGATGGT ATTTGAAAAAATTGAT	Cm forward primer with modifed Shine Delgarno
mod-CM/SDG-R	CGGCGCCAGTTACAGACAAACCTGAAGT	Cm reverse primer
pSOS94_fwd	GGAATGGCGTGTGTGTTAGCCAAA	Cm and primer forward primer
pSOS94_rev	TCACACAGGAAACAGCTATGACCA	Cm and primer reverse primer
akko_fwd	ATATATGGCGCGCCAGGACACAGAATCGTTCATGGTGG	*ack* fragment forward primer
akko_rev	ATATATGGCGCGCCCCCTAACTTTAGAACCTTCTGCAC	*ack* fragment reverse primer
bkko_fwd	ATATATGGCGCGCCATCAATCCTGGCTCGACCTCAACT	*buk* fragment forward primer
bkko_rev	ATATATGGCGCGCCCGTTGTACGCAATTCCGCCTGTTA	*buk* fragment reverse primer

Table 4.2 Primers used in plasmid construction.

All plasmids were transformed into Top 10 chemically competent *E. coli* (Invitrogen). Plasmids were confirmed by sequencing. The plasmids were methylated using *E. coli* ER2275 (pAN1) to avoid the natural restriction system of *C. acetobutylicum* (Mermelstein and Papoutsakis 1993). Once methylated, the plasmids were transformed by electroporating the *C. acetobutylicum* mutant M5 as described (Mermelstein, Welker et al. 1992).

4.3 Results

4.3.1 Expression of *aad* using the *ptb* promoter for increased butanol production in strain M5.

The asporogenous, non-solventogenic strain M5 strain (Clark, Bennett et al. 1989) has lost the megaplasmid pSOL1 that contains all the genes required for solvent production (Cornillot, Nair et al. 1997). The megaplasmid contains two independent bifunctional aldehyde/alcohol dehydrogenases (AADHs) capable of forming butanol from butyryl-CoA (Nölling, Breton et al. 2001). The primary AADH gene, *aad*, is induced during batch cultures at low pH (Nair, Bennett et al. 1994), while the second AADH gene, *adhE2*, is induced in alcohologenic cultures at near neutral pH (Fontaine, Meynial-Salles et al. 2002). Both genes were expressed in pSOL1-deficient backgrounds (Nair and Papoutsakis 1994; Fontaine, Meynial-Salles et al. 2002): both strains were only able to produce about half the levels of butanol that is typically produced by the WT strain. *aad* expression in the pSOL1 deficient background relied on the expression from the endogenous promoter. Expression of *adhE2* utilized the *thl* promoter, which has high, constitutive-like expression (Tummala, Welker et al. 1999), but was unable to achieve high butanol titers. As solvent formation and the sporulation process are co-regulated in the WT strain by a common transcriptional regulator, Spo0A (Harris,

Welker et al. 2002), it was thought that the natural promoter of *aad* may not be sufficiently induced in the asporogenous strain M5. Thus, the natural promoter of *aad* was replaced with the *ptb* promoter of the *ptb-buk* operon (Fig. 4.1), which is expressed highly especially in the early part of the fermentation (Tummala, Welker et al. 1999). Additionally, butyrate is the primary product of the pSOL1 deficient strain M5 and any promoter titration effects would be beneficial for the reduction of butyrate formation, whose accumulation leads to premature growth inhibition of the M5 strain. Strain M5(p94AAD3) was constructed to overexpress *aad* from the *ptb* promoter. Butanol formation in test tube cultures confirmed the proper expression and function of the AAD protein (M5 forms no butanol at all). Controlled pH 5.0 fermentations were carried out to characterize the strain (Table 4.3). At this pH, M5(p94AAD3) produced 92 mM butanol which is less than typical levels of ca. 160 mM produced in WT fermentations, but higher than the 84 mM produced by M5(pCAAD), which expresses *aad* from its endogenous promoter. Acid production was also altered. Butyrate formation was reduced from 99 mM in M5(pCAAD) to 72 mM in M5(p94AAD3), while acetate levels increased to 159 mM from 101 mM. The decreased butyrate formation could be due to either the titration effect of addition plasmid copies of the *ptb* promoter or from increased competition for butyryl-CoA from AAD. The increased acetate formation may reflect an effort by the cells to compensate for the loss of ATP generation from the decreased butyrate formation. In view of the fact that the *ptb* promoter and generally both butyrate and acetate formation are affected by pH levels (Bahl 1982; Monot 1984), we next tested the impact of culture pH on product formation.

Strains	Fermentation characteristics[a]						
	pH	Max A_{600}	Doubling Time	Butanol	Ethanol	Acetate	Butyrate
M5[b]	5.0	-	1.15	0	6	107	169
M5(pCAAD)[b]	5.0	-	1.59	84	8	101	99
M5(p94AAD3)	5.0	11.64	2.19	92	20	159	72
M5(p94AAD3)	5.5	12.81	1.82	140	36	222	73
M5(p94AAD3)	5.75	10.81	2.08	150	44	248	87
M5(p94AAD3)	6.0	9.23	1.98	138	34	256	144
M5(pTHLAAD)	5.5	8.90	1.86	108	19	172	84
M5(pTHLAAD)	6.0	8.25	1.88	77	17	148	196
M5 AKKO(p94AAD3)	5.5	7.58	2.25	92	22	180	75
AKKO(p94AAD3)	6.0	9.00	2.65	51	21	179	287
M5[c]	-	5.29	1.20	0	6	23	73
M5 AKKO[c]	-	3.81	1.48	0	7	9	75
M5 BKKO[c]	-	1.42	2.93	0	3	17	26

M5

[a]Doubling time (hours), product concentrations are shown in mM.
[b]Results from M5 and M5(pCAAD) are from a previous study (Nair and Papoutsakis 1994)
[c]Results are from static flask experiments without pH control

Table 4.3 Final product concentrations in pH controlled batch fermentations of *C. acetobutylicum* M5 and recombinant strains.

4.3.2 Culture pH profoundly impacts solvent production in M5(p94AAD3) with maximal butanol titers at pH 5.75.

We investigated the impact of culture pH in controlled pH fermentations carried out at pH 5.5, 5.75, and 6.0 (Fig. 4.2). Butanol titers at pH 5.5, 5.75, and 6.0 were 140, 150, and 137 mM, respectively, significantly higher than the 92 mM reached at pH 5.0. Interestingly, butanol production was slightly decreased at pH 6.0, combined with dramatic increases in the levels of acids produced. Acetate levels increased with increasing pH, reaching a maximum value of 256 mM at pH 6.0, compared to the 159 mM produced at pH 5.0. Butyrate levels were fairly similar between cultures maintained at pH 5.0 to 5.75, but nearly doubled to 144 mM at pH 6.0. In order to obtain insight into how pH had such profound effects on cell metabolism, we carried out metabolic flux analysis, which we discuss next.

4.3.3 Metabolic flux analysis of M5(p94AAD3) to assess the impact of culture pH

At pH 5.0, glucose utilization, rGLY1, and the conversion of pyruvate to acetyl-CoA, rGLY2, in M5(p94AAD3) began very high, but quickly dropped to near zero by hour 10 (Fig. 4.3). With increased pH, M5(p94AAD3) exhibited a bimodal glucose utilization pattern, with high utilization early and then again hour 10. This second period of glucose utilization corresponds with increased fluxes of butanol and acetate formation. The butanol formation flux reached its maximal value at about hour 10 at pH 5.5 – 6.0. Acetate formation began high, but showed a second large peak in its flux that overlaps with the peak of the butanol flux. The butyrate formation flux was also sustained longer with increasing pH. The ethanol formation flux followed similar trends as the butyrate formation flux displaying a high initial flux, and a more

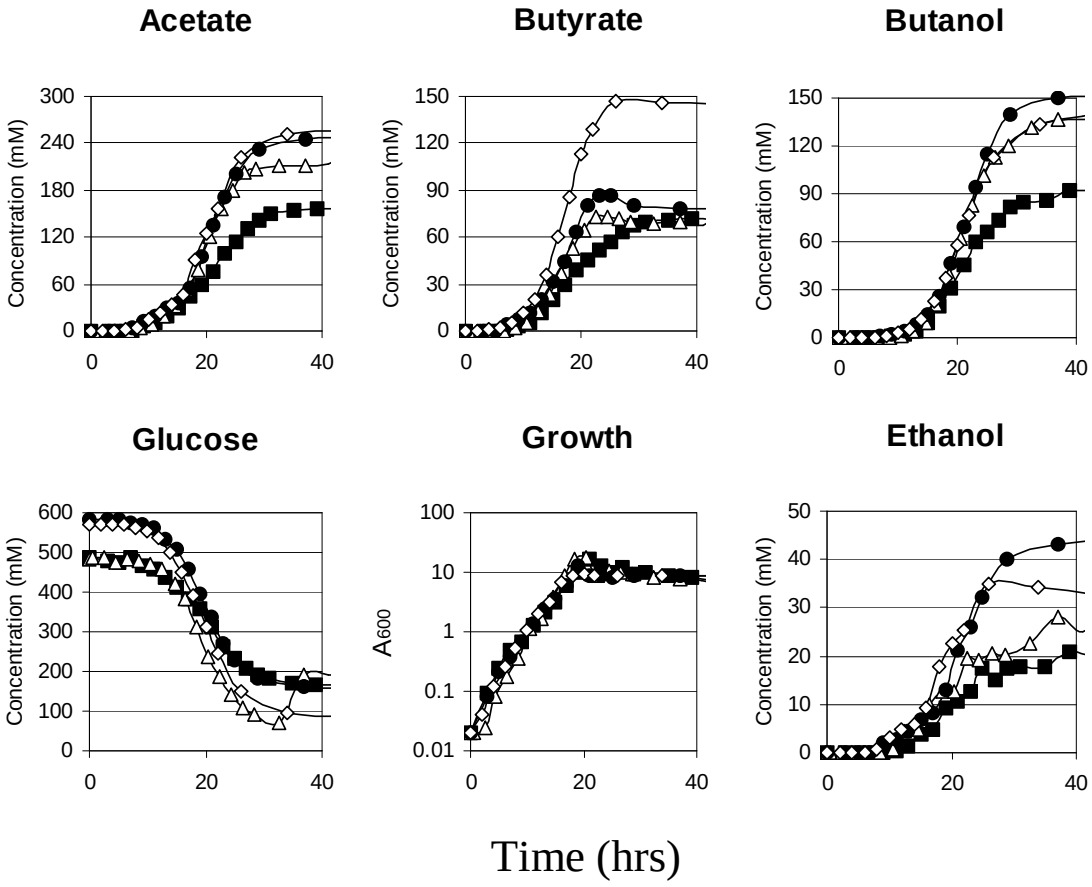


Figure 4.2 pH-controlled fermentations of M5(p94AAD3).

Product concentrations for the major metabolites and growth characteristics of the M5(p94AAD3) fermentations at various fermentation pH values. pH 5.0 (closed squares), pH 5.5 (open triangles), pH 5.75 (closed circles), and pH 6.0 (open diamonds) are shown. Time has been normalized based on culture absorbance to account for variations in lag time of the cultures (A_{600} of 1.0 occurs at hour 10).

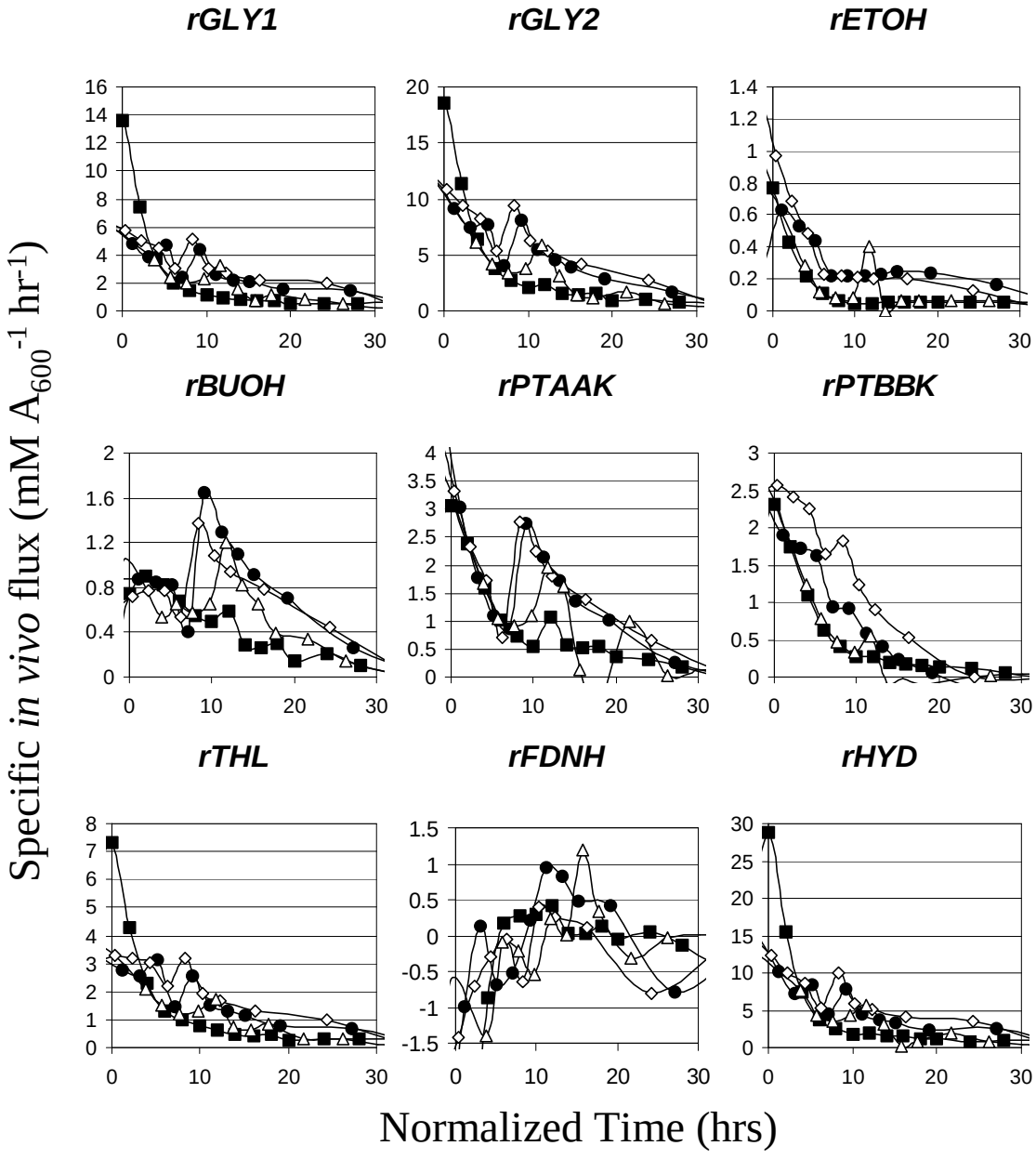


Figure 4.3 Metabolic flux analysis of M5(p94AAD3) at various fermentation pH settings

pH 5.0 (closed squares), pH 5.5 (open triangles), pH 5.75 (closed circles), and pH 6.0 (open diamonds) are shown. Time has been normalized based on culture absorbance to account for variations in lag time of the cultures (A_{600} of 1.0 occurs at hour 0).

prolonged flux with increasing pH. The rTHL and rHYD fluxes showed very similar patterns to the glucose utilization fluxes, exhibiting high initial values with a second peak apparent at pH 5.5 and above. The NADH from ferredoxin formation flux (rFDNH) displayed more complex patterns, mostly bimodal but also some trimodal patterns. These patterns were evident for all pH values and only approximately correspond to the peaks in the butanol or ethanol-formation fluxes, or the rHYD fluxes, thus suggesting unusual complexity in the regulation of electron flow.

4.3.4 Comparative analysis of M5(p94AAD3) and M5(pTHLAAD) fermentations shows that thiolase overexpression benefits butyrate plus butanol at the expense of acetate plus ethanol formation but suppresses butanol production

At pH values below 6.0, butyrate concentration is fairly low in M5(p94AAD3) cultures and only about half the butanol concentration. Acetate concentrations however are the highest of all metabolites in all pH values tested. We thus examined if *thl* overexpression might reduce acetate formation. THL is the first enzyme in the pathway converting acetyl-CoA to butyryl-CoA. Acetyl-CoA is the metabolite at the critical branchpoint where it can be converted by phosphotransacetylase (PTA) and acetate kinase (AK) to acetate, or by AAD to ethanol if not converted to butyryl-CoA. Butyryl-CoA is converted to either butyrate by phosphotransbutyrylase (PTB) and butyrate kinase (BK) or butanol by AAD. Increased expression of *thl* might increase the carbon flux from acetyl-CoA to butyryl-CoA reducing the formation of acetate and ethanol and increasing butyrate and butanol formation. *thl* with its natural promoter and the *aad* gene with the *ptb* promoter were combined on a single construct to create the plasmid pTHLAAD.

Fermentations of M5(pTHLAAD) were performed at pH 5.5 and 6.0 based on the performance of the M5(p94AAD3) strains. At pH 5.5, acetate levels were reduced from 222 mM in M5(p94AAD3) to 172 mM in strain M5(pTHLAAD) and ethanol production was reduced from 36 mM to 19 mM (Table 4.3). However, the lowered production of acetate and ethanol was not accompanied by increased butanol formation. Butanol production decreased from 140 mM to 108 mM, and butyrate increased from 73 mM to 84 mM. At pH 6.0 acetate was reduced from 256 mM in M5(p94AAD3) to 148 mM in M5(pTHLAAD) and ethanol was reduced from 34 mM to 17 mM. Butanol levels also decreased from 138 mM to 77 mM while butyrate production increased from 144 mM to 196 mM.

Metabolic flux analysis shows glucose utilization (rGLY1), rGLY2 and rTHL are all significantly higher (by 2-3 fold) early in the fermentation of M5(pTHLAAD) compared to M5(p94AAD3) at both pH values (Fig. 4.4). Most of the increased glucose utilization and thiolase flux is manifested in butyrate formation indicated by the higher initial fluxes of rPTBBK. The butyrate formation flux in M5(pTHLAAD) was roughly twice the initial flux in M5(p94AAD3). This difference is most evident at pH 6.0 when butyrate formation was sustained the longest. rBUOH was induced much earlier at pH 5.5 in M5(pTHLAAD) than M5(p94AAD3), but this was reversed at pH 6.0. This suggests that even under the *ptb* promoter, AAD expression and/or activity are more pronounced at lower pH, and thus able to compete more effectively for butyryl-CoA for butanol vs. butyrate formation. The butanol flux was trimodal in strain M5(pTHLAAD) at pH 6, but less pronounced so at pH 5.5. The acetate fluxes (rPTAAK) were overall lower in strain M5(pTHLAAD), and did not exhibit as pronounced a multimodal pattern as the fluxes in the M5(p94AAD3) strain. As was the case in the M5(p94AAD3) strain, the fluxes for acetate and butanol in M5(pTHLAAD) closely aligned with

each other. The ethanol fluxes were consistently lower in strain M5(pTHLAAD) than in strain M5(p94AAD3), thus suggesting that lower pools of acetyl-CoA due to *thl* overexpression (Fig. 1) penalize ethanol formation the most while enhancing butanol and butyrate formation. The rFDNH fluxes show distinct but difficult to interpret patterns: by carefully examining the areas under the curves, the most significant observation is that this flux remains mostly negative, especially for M5(pTHLAAD) (compare against the fluxes in Fig. 4.3). This means that NADH is used to produce reduced ferredoxin and eventually H_2. We conclude that *thl* overexpression benefits 4-C compound (butyrate + butanol) formation at the expense of 2-C compound (acetate + ethanol) formation (Table 4), but results in the suppression of butanol formation.

4.3.5 Can disruption of the *ack* and *buk* genes reduce acetate and butyrate formation?

As efforts thus far were unable to reduce the formation of acids by *aad* and/or *thl* overexpression, gene knockouts (KOs) were used to inactivate the acetate and butyrate formation pathways separately. A non-replicative plasmid method was used, utilizing a cloning vector lacking an origin of replication for *C. acetobutylicum.* As previous efforts utilizing antibiotic resistance markers had limited success in generating KO mutants, it was thought that by increasing the expression of the antibiotic resistance genes might result in greater efficiency of KO strain generation. Although antibiotic markers for erythromycin and thiamphenicol resistance have been used successfully with replicating cloning vectors, these plasmid-borne markers have multiple gene copies in each cell, while chromosomal integrants would have only one gene copy. To increase expression and translation of the thiamphenicol resistance gene, an optimized Shine-Delgarno sequence (AGGAGG) replaced the endogenous sequence and the gene was put under the control of the high expression *ptb* promoter. The thiamphenicol

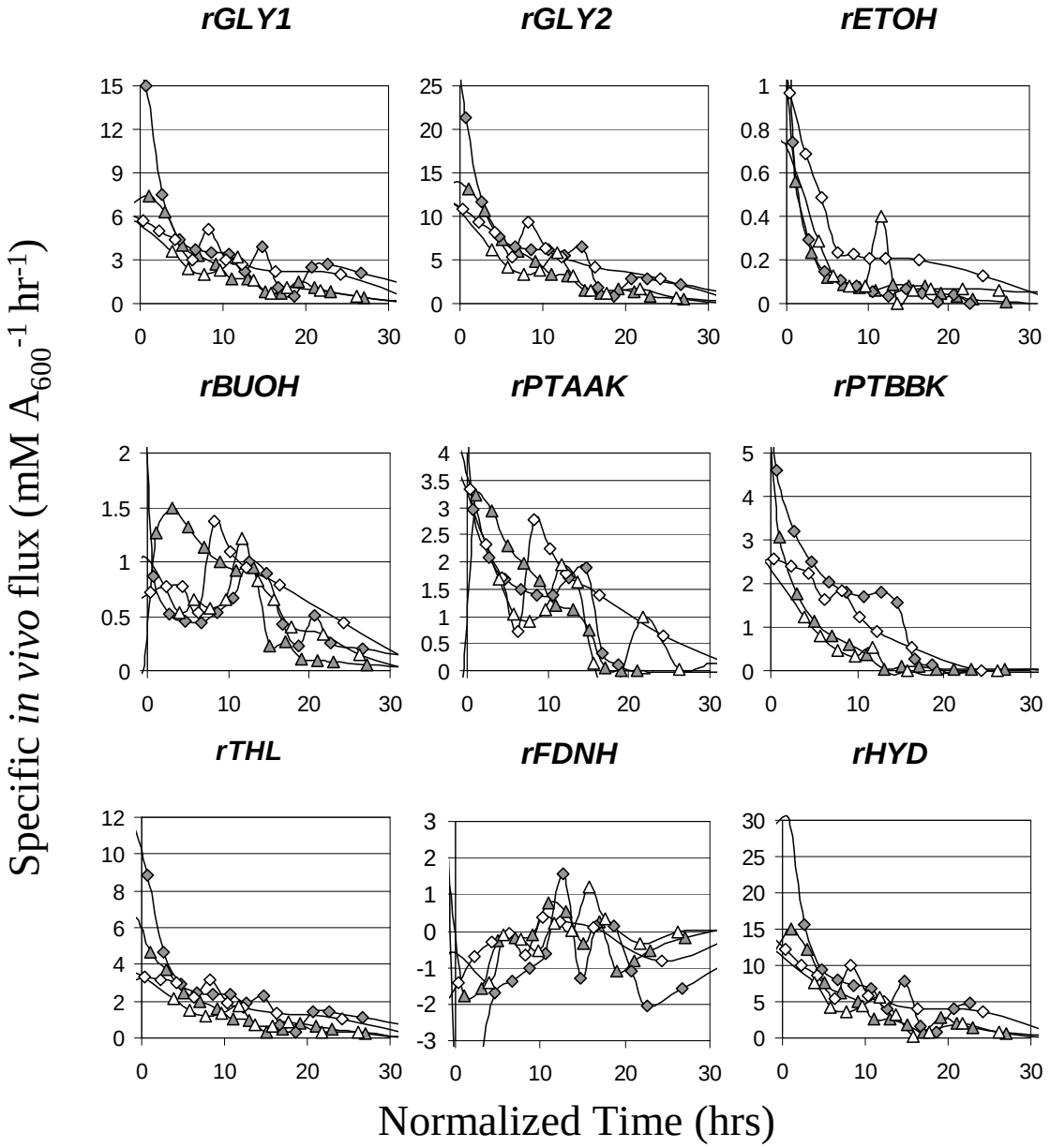


Figure 4.4 Metabolic flux analysis of M5(p94AAD3) and M5(pTHLAAD).

M5(p94AAD3) fluxes are shown as open symbols at pH 5.5 (triangles) and 6.0 (diamonds). M5(pTHLAAD) fluxes are shown as gray symbols at pH 5.5 (triangles) and 6.0 (diamonds). Time has been normalized based on culture absorbance to account for variations in lag time of the cultures (A_{600} of 1.0 occurs at hour 0).

resistance marker was used for compatibility with the other plasmids used in this study. The transformants were initially screened using colony PCR. After the confirmation of the genomic integration, one clone (M5 AKKO) for the *ack* disruption and one (M5 BKKO) for the *buk* disruption were chosen for further study. Genomic DNA was isolated from each strain and the chromosomal locus spanning the genomic integration plus upstream and downstream sequences were PCR amplified and sequenced. The sequencing data confirmed the expected single crossover event that has previously been observed with similar methods (Green, Boynton et al. 1996). The schematic of the genomic crossover and the sequencing data are shown in Appendix A.

The growth and metabolic profiles of M5 AKKO and M5 BKKO strains were compared against the parental M5 strain in 200 mL flask cultures without pH control (Fig. 4.5). The growth rates of the M5 and the M5 AKKO strains were fairly similar with doubling times of 1.20 and 1.48 hours, respectively. M5 BKKO displayed a dramatic growth inhibition, with a doubling time of 2.93 hours, which makes it nearly 2.5 times slower than the M5 and twice slower than the M5 AKKO strain. Product formation was altered in both recombinant strains compared to the parental strain, largely as would be expected. Acetate production was reduced by > 60% from 23 mM in M5 to just 9 mM in M5 AKKO. Butyrate production in the M5 and M5 AKKO strains was virtually the same at 73 mM and 76 mM, respectively. Butyrate was produced at very low levels in the M5 BKKO strain, reaching a maximum of 26 mM after 75 hours of growth. This is a reduction of 64% compared with M5. Acetate production was also reduced in M5 BKKO, reaching a final value of 17 mM. This level is more than the M5 AKKO could produce suggesting that the acetate pathway is less inhibited than the butyrate pathway in M5 BKKO.

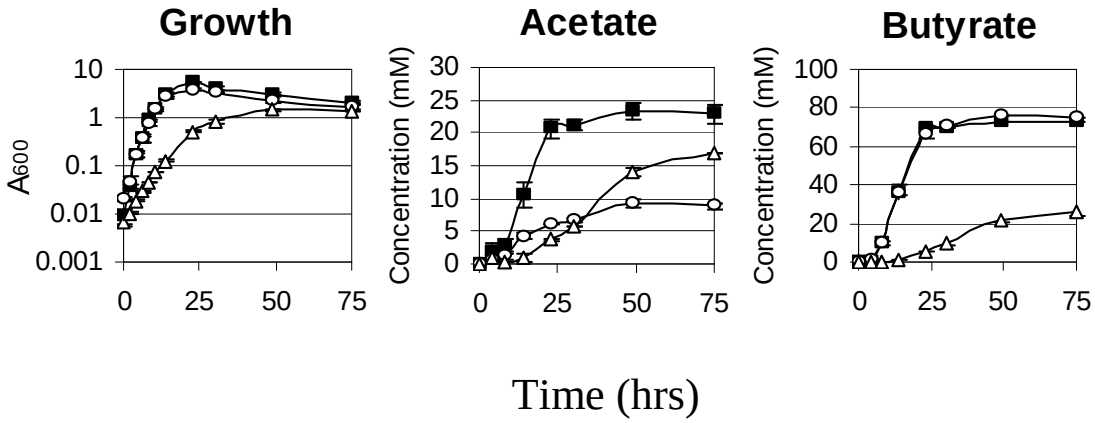


Figure 4.5 Growth profiles and product analysis of M5 AKKO, M5 BKKO, and M5

Growth profiles and acetate and butyrate production in static flask cultures without pH control. M5 (closed squares), M5 AKKO (open circles), and M5 BKKO (open triangles). Results shown are averages of two experiments. Error bars show the standard deviation between experiments. Error bars not visible are smaller than the symbols used.

Both the M5 AKKO and M5 BKKO could still produce low levels of the acids targeted by the genomic disruptions. This has been observed in other clostridia mutants involving the inactivation of the acid formation pathways where parallel and analogous pathways (e.g., acetate and butyrate) exist (Green, Boynton et al. 1996; Zhu, Liu et al. 2005; Liu, Zhu et al. 2006). The formation of the product in the inactivated pathway is attributed to the cross reactivity of the enzymes in the alternate pathway.

4.3.6 Overexpression of *aad* in the M5 AKKO strain: antibiotics and pH impact product formation, but butanol formation remains severely inhibited

Following the successful inactivation and characterization of the *ack* and *buk* genes in the M5 strain, we aimed to transform the two KO strains with the *aad* overexpression plasmid p94AAD3. The goal was to examine the combined effect of AAD overexpression with the downregulation of either the acetate or butyrate formation pathways. Despite repeated attempts to transform the M5 BKKO strain with p94AAD3, no transformants could be isolated. This is likely due to the very low growth rate achieved by these cells; it is also possible that the *buk* inactivation may impact the competence of the cell. Correct transformants containing p94AAD3 were however isolated from the M5 AKKO strain.

Initial results from tube cultures without pH control that used both erythromycin and thiamphenicol for strain maintenance showed that glucose uptake was inhibited and suggested that the cells could not grow well when using multiple antibiotics. In an earlier study (Harris, Blank et al. 2001), it had been determined that the use of tetracycline drastically inhibits solvent formation in *C. acetobutylicum*, but the effect of simultaneous usage of thiamphenicol and erythromycin has not been examined. In order to assess the impact of antibiotic usage on

butanol production, cultures were first inoculated in the presence of both thiamphenicol and erythromycin and then diluted 100-fold into media containing either both thiamphenicol and erythromycin, thiamphenicol alone, erythromycin alone, or no antibiotic. The thiamphenicol resistance gene is integrated into the genome disrupting the *ack* gene, while the erythromycin is used for the plasmid maintenance for *aad* overexpression.

Cultures produced the highest level of butanol when grown with erythromycin and nearly as much was produced when both antibiotics were used (Fig. 4.6). When no antibiotic or thiamphenicol alone was used, the cells produced low butanol levels and increased levels of butyrate. These results indicate that erythromycin is needed to maintain the plasmid and ensure high butanol production, while the dual antibiotic usage has a small inhibitory effect on butanol formation. In subsequent pH-controlled fermentation experiments, clarithromycin only was used for plasmid maintenance (clarithromycin is a pH stable derivative of erythromycin). Genomic integrations of this type have been previously studied and have been shown to be stable (Green, Boynton et al. 1996).

Product formation by M5 AKKO(p94AAD3) in pH 5.5 and 6.0 fermentations is summarized in Table 4.3. Corresponding data for strain M5(p94AAD3) are shown in order to comparatively assess the impact of *aad* overexpression with *ack* inactivation. At pH 5.5, acetate production was reduced from 222 mM in M5 (p94AAD3) to 180 mM in M5 AKKO(p94AAD3), but unexpectedly butanol production was also significantly reduced from 140 mM to 92 mM. Butyrate production was only slightly increased from 73 mM in M5 (p94AAD3) to 75 mM in M5 AKKO(p94AAD3). Changes in product formation were more dramatic at pH 6.0. Final acetate levels were reduced from 256 mM to 179 mM. Butanol production was reduced from

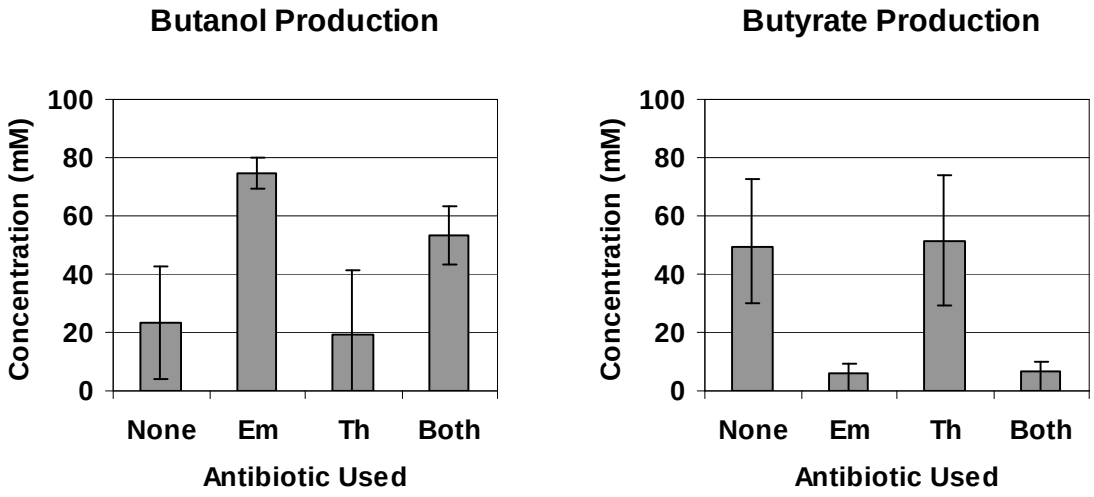


Figure 4.6 Erythromycin and thiamphenicol effects on product formation in M5 AKKO (p94AAD3)

Cultures were performed without pH control varying the addition of erythromycin (Em) and thiamphenicol (Th). Results shown are the average of 5 experiments with error bars showing the standard deviation.

138 mM to just 51 mM. The most extreme product shift was the final butyrate concentrations. In M5(p94AAD3), butyrate is produced at 144 mM while in M5 AKKO(p94AAD3) butyrate wass doubled to 287 mM, the highest levels ever observed in this organism. The shift in acid products can be seen clearly by comparing the ratio of acetate versus butyrate produced. At pH 5.5, M5(p94AAD3) had an acetate to butyrate ratio of 3.05, whereas strain M5 AKKO(p94AAD3) exhibited a reduced ratio of 2.40. At pH 6.0, M5(p94AAD3) displayed an acetate to butyrate ratio of 1.78, but this was reduced by two-thirds to 0.62 in M5 AKKO(p94AAD3).

Metabolic flux analysis was also performed to determine differences between M5 AKKO(p94AAD3) and M5(p94AAD3). Glucose utilization was increased early in the M5 AKKO(p94AAD3) compared to M5(p94AAD3) (Fig 4.7). Peak butanol formation occurred earlier in M5 AKKO(p94AAD3) than in M5(p94AAD3), reflecting the impact of *ack* deactivation presumably leading. Corresponding to the low butanol flux at pH 6.0 in M5 AKKO(p94AAD3), the butyrate formation flux was increased dramatically and was sustained at high levels much longer than in M5(p94AAD3). The thiolase flux showed a small induction in M5 AKKO(p94AAD3) at pH 6.0, primarily due to the large increase in butyrate formation flux compared to M5(p94AAD3). Once again, the acetate formation flux aligned with the butanol formation flux. Ethanol formation was significantly reduced in strain M5 AKKO(p94AAD3) despite the reduction in acetate formation, which might have anticipated a compensatory ethanol formation. There are also important differences in electron flow: for example, at pH 6, M5 AKKO(p94AAD3) displayed much higher levels of the rHYD flux and all negative values for the rFDNH flux (formation of reduced ferredoxin from NADH), which means all electrons were

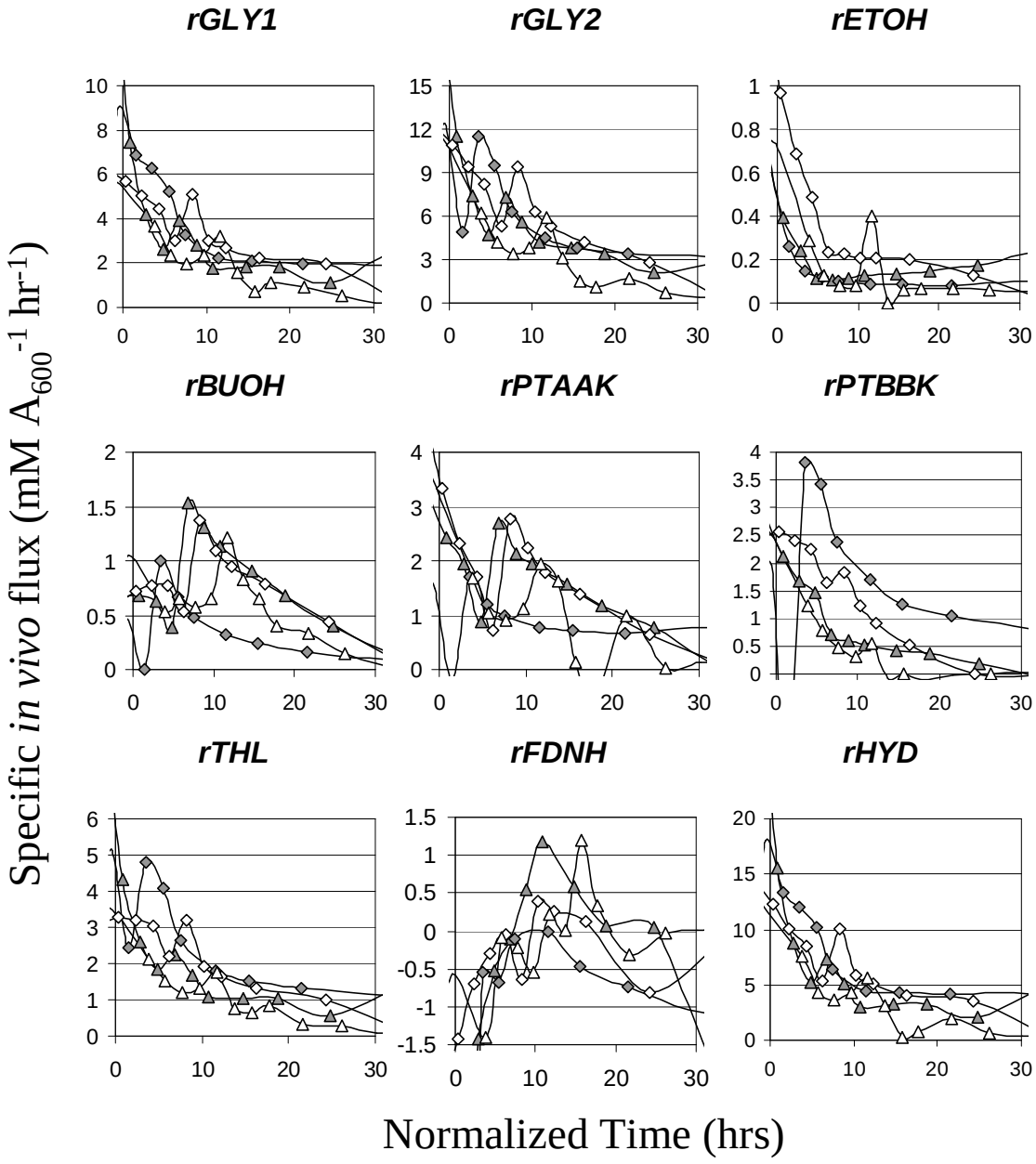


Figure 4.7 Metabolic flux analysis of M5(p94AAD3) and M5 AKKO(p94AAD3)

M5(p94AAD3) fluxes are shown as open symbols at pH 5.5 (triangles) and 6.0 (diamonds). M5 AKKO(p94AAD3) fluxes are shown as gray symbols at pH 5.5 (triangles) and 6.0 (diamonds). Time has been normalized based on culture absorbance to account for variations in lag time of the cultures (A_{600} of 1.0 occurs at hour 0).

shuffled to H_2 production rather than butanol/ethanol formation even late in culture, and this is an extreme pattern never previously observed, and was accompanied by extremely high levels of butyrate formation. It appears that M5 AKKO(p94AAD3) cannot effectively utilize NADH for alcohol formation, it thus uses butyrate formation to minimize NADH generation, and still produces substantial levels of H_2 to balance the electron flow. It appears as if other plasmid-coded proteins are necessary for making possible the beneficial use of reduction energy to produce butanol and ethanol.

4.4 Discussion

In this study we were able to achieve WT levels of butanol production without acetone production in a non-sporulating culture by expressing *aad* from the *ptb* promoter. It has been suggested that this level of butanol production in the pSOL1 deficient background was unlikely due to the inability of the strains to reassimilate acid products (Fontaine, Meynial-Salles et al. 2002). Reassimilation of acid products would replenish butyryl-CoA pools within the cell for butanol production. We have shown here that by manipulating the genetic regulation of *aad* by replacing the natural promoter with the *ptb* promoter that WT levels can be achieved without acid re-assimilation. It has also been thought that the pSOL1 deficient strains may be more sensitive to butanol than the WT sporulating strain. A previous study has shown that the *spo0A* overexpressing *C. acetobutylicum* strain enters sporulation earlier than the WT strain and that this early sporulation may explain its enhanced butanol tolerance (Alsaker, Spitzer et al. 2004). Sporulating cultures were thought to be more tolerant to solvents than non-sporulating cultures perhaps due to unknown genes on the pSOL1 plasmid (Fontaine, Meynial-Salles et al. 2002). We have shown here that M5 derivative strains were not only able to withstand WT levels of

butanol, but were also subjected to elevated levels of butyrate suggesting that tolerance of these strain is comparable to their sporulating counterparts.

As these strains are unable to reassimilate acids, high levels of acids were produced. In one strategy, high acetate production was countered using *thl* overexpression combined with the *aad* overexpression. The strain overexpressing both *thl* and *aad* did produce less acetate than the strain overexpression *aad* alone and produced a higher ratio of four-carbon products (butanol and butyrate) to two-carbon products (ethanol and acetate), but less butanol was produced and higher concentrations of butyrate were formed. In a second strategy, the acetate kinase and butyrate kinase genes were inactivated using homologous recombination via a non-replicating plasmid. Both of these strains successfully reduced the product formation of the targeted pathway, but the inactivation of the butyrate kinase resulted in a dramatic inhibition of growth that has blocked further genetic manipulations. Earlier studies were unable to achieve a *buk* disruption in the pSOL1 deficient strain DG1, and it was thus suggested that a mutant producing only acetate would not be viable (Gonzalez-Pajuelo, Meynial-Salles et al. 2005). Here we have created a *buk* disruption mutant, but the strain produced low butyrate levels and had a very low growth rate.

The acetate kinase mutant M5 AKKO was successfully complemented with *aad* overexpression and acetate levels were reduced in this strain, although this was accompanied by a reduction in butanol formation as well. This strain had a significant shift in the ratio of acetate to butyrate produced, particularly at higher pH. Overall, *thl* overexpression and inactivation of acetate kinase resulted in the production of higher levels of butyrate and lower levels of acetate. This shift in acid production has also been accompanied by lower butanol titers.

The slow growth rate of the butyrate kinase mutant suggests that M5 primarily balances NADH formation/use by producing butyrate (Fig. 4.1). The alternate route for NADH oxidation

is the reduction of ferredoxin ultimately leading to hydrogen production (Fig. 4.1). This alternate route of NADH oxidation has been suggested as thermodynamically unfavorable (Gonzalez-Pajuelo, Meynial-Salles et al. 2005). Here, while the rFDNH and rHYD appear overall operational (Figs. 4.3, 4.4 and 4.7), the patterns of the rFDNH flux are very different (mostly negative) than in strains based on the WT pSOL1-containing genetic background (Chapter 3). This suggests that this pathway for NADH utilization to form H_2 cannot be increased to compensate for the lack of butyrate formation.

The introduction of the *aad* gene into the M5 strain adds still another route for NADH oxidation. During cellular metabolism, the conversion of pyruvate to acetyl-CoA results in the reduction of ferredoxin, which must then be oxidized. If during metabolism ferredoxin is oxidized to generate H_2, acetate formation would result in an excess amount of NADH, while butanol production would result in an NADH deficit. In the WT strain, H_2 production decreases relatively early in the fermentation and thus subsequently NADH can be produced from the oxidation of reduced ferredoxin (Fig. 4.1). This reaction can provide the NADH necessary for solventogenesis. Here, it was observed that butanol production is correlated with acetate production in several recombinant strains in the M5 genetic background. It seems the FDNH flux is tightly regulated and the NADH necessary for butanol production must come from additional acetate production. This once again highlights the difficulty the M5 strain has in manipulating NADH generation/consumption to achieve the necessary electron balance. Unlike the parent, WT 824 strain, the reactions necessary to achieve the electron balance in the cells appear to be stiff, resisting the changes necessary to achieve an electron balance as necessitated by genetically imposed changes on the metabolic pathway, such as the elimination of butyrate formation.

This suggests that the pSOL1 megaplasmid carries some gene or genes of importance to making the electron flow flexible and easy as it occurs in the WT 824 strain, which contains the pSOL1 megaplasmid. There are several oxidoreductase genes residing on pSOL1, and these may be involved in this reaction. Alternatively, a large number of genes encoding transcription factors and DNA binding proteins also reside on pSOL1, and these may be involved in the regulation of the electron flow. Another noteworthy observation that supports the important role of other pSOL1 genes in providing product-formation and electron-balance flexibility is that when the pPTBAAD plasmid (which is identical to the p94AAD3 plasmid) is used to overexpressed *aad* in the WT 824 strain, large quantities of ethanol are formed in addition to butanol resulting in abnormally high ethanol/butanol ratios (Sillers, Al-Hanai et al.). This does not happen however in the M5 strain upon *aad* overexpression from the same plasmid. Further efforts to increase butanol production should focus on the identification of the gene or genes involved in this complex regulation of electron flow in the M5 genetic background.

CHAPTER 5 : THE *CLOSTRIDIUM ACETOBUTYLICUM* TWO-COMPONENT SYSTEM CODED BY CAC0653 AND CAC0654 REGULATES THE PROGRESSION OF SPORULATION

5.1 Introduction

Bacteria thrive in a variety of environments and have developed complex signaling pathways to coordinate adaptive responses to changing conditions. The most common manner for signal recognition and cellular response is through a two-component system (Hoch 2000), comprised of a sensor histidine kinase and a response regulator. Histidine kinases (HKs) contain an input domain and a transmitter domain (Rodrigue et al. 2000). The input domain is responsible for recognizing and binding the environmental stimulus and varies considerably among HKs. The transmitter domain, however, is highly conserved and this region contains the phosphorylatable histidine residue making the identification of HKs possible (Fabret et al. 1999). Once the signal is received the phosphate is transferred from the transmitter domain of the HK to the receiver domain of the response regulator (RR). This phosphate addition presumably alters the conformation of the protein allowing DNA binding and the induction or repression of its target genes.

Bacterial two-component systems have been reviewed in several organisms (Fabret et al. 1999; Mizuno 1997; Mizuno et al. 1996; Rodrigue et al. 2000; Wang et al. 2002). More complex sensory systems involve multiple proteins acting in a phosphorelay, such as the initiation of

sporulation in bacilli (Burbulys et al. 1991). Like bacilli, clostridia are also Gram-positive sporulating organisms, but are anaerobic and occupy a different environmental niche than bacilli. While many components of sporulation downstream of the phosphorelay are conserved in both clostridia and bacilli, none of the components of the bacilli phosphorelay are recognizable in clostridia (Paredes et al. 2005). Recent transcriptional profiling experiments of *C. acetobutylicum* using microarrays (Alsaker and Papoutsakis 2005; Alsaker et al. 2004; Jones et al. 2008; Tomas et al. 2003a) have produced significant information regarding its differentiation programs, yet little is known regarding the entry into and the regulation of sporulation and the characteristic stationary phase phenomenon of solvent production. Genomic analysis of *C. acetobutylicum* has identified 28 two-component systems ((Dürre 2005)), but only three (*kpdDE*, *phoPR*, and *agrAC*) have been annotated and two systems have been studied, including the *kpdDE* system (Behrens and Dürre 2000; Treuner-Lange et al. 1997) and the PhoP/R system (Fiedler et al. 2008).

Two recent DNA-microarray-based analyses have generated transcriptional data for all histidine kinases and two-component systems in *C. acetobutylicum* (Alsaker and Papoutsakis 2005; Jones et al. 2008), and such data are useful in identifying potential candidate regulators of sporulation (Paredes et al. 2005). Among the non-orphan histidine kinases (Paredes et al. 2005), the expression profile of CAC0653/CAC0654 coincides with the pattern of several other regulatory genes of early sporulation, and suggests that this two-component system may be involved in the regulation of either the process of solvent formation or sporulation (Alsaker and Papoutsakis 2005). In this study, we have examined the potential role of this two-component system coded by CAC0653(RR) and CAC0654(HK) on sporulation and solventogenesis.

5.2 Materials and methods

5.2.1 Bacterial strains

Bacterial strains and plasmids used in this study are listed in Table 5.1.

5.2.2 Strain construction

To produce the asRNA cassette for the downregulation of CAC0654, complementary 80 base pair oligonucleotides (IDT, Coralville, IA) were designed to include the ribosome-binding site of CAC0654 and 60 base pairs of downstream coding sequence of the CAC0654 gene with *BamH*I and *Kas*I overhangs according to a previously described method (Desai and Papoutsakis 1999). The nucleotide sequences are shown in Table 5.2. One microgram of each oligonucleotide was added to a 1X SSC solution in 20 μL and incubated at 95°C for 10 minutes and then allowed to cool to room temperature for one hour to anneal. pSOS95, a high expression clostridial vector utilizing the thiolase promoter (Tummala et al. 1999), was digested with *Bam*HI and *Kas*I and the large fragment was gel-band isolated. The annealed oligonucleotide was ligated in an antisense orientation into the large pSOS95 fragment to produce p654as. PT6534 was constructed by amplifying the two gene open reading frame using primers 653_fwd and 654_rev. Following digestion with *Bam*HI and *Kas*I the fragment was ligated into the digested pSOS95del. The plasmids were confirmed using sequence analysis before methylation using *E. coli* (pAN1) (Mermelstein and Papoutsakis 1993) to overcome the restriction of this organism. *C. acetobutylicum* was transformed with the methylated plasmid by electroporation as previously described (Mermelstein et al. 1992).

Strain or Plasmid	Relevant Characteristics[a]	Source or Reference[b]
Bacterial Strains		
C. acetobutylicum		ATCC
E. coli		
Top10		Invitrogen
ER2275		New England Biolabs
Plasmids		
pAN1	Cm^r, Φ3T I gene, p15A origin	(Mermelstein and Papoutsakis 1993)
pSOS95[c]	acetone operon (*thl* promoter)	Soucaille and Papoutsakis, unpublished
pSOS95del[c]	*thl* promoter	(Tummala et al. 2003a)
p654as[c]	654 asRNA (*thl* promoter)	This study
pT6534[c]	CAC0653 and CAC0654 (*thl* promoter)	This study

[a]Cm^r, chloramphenicol/thiamphenicol resistance gene; *ptb*, phosphotransbutyrylase gene; *aad*, alcohol/aldehyde dehydrogenase gene; *thl*, thiolase gene; *adc*, acetoacetate decarboxylase gene; *ack*, acetate kinase gene; *buk*, butyrate kinase gene.

[b]ATCC, American Tissue Culture Collection, Rockville, MD

[c]contains the following: ampicillin resistance gene; macrolide, lincosimide, and streptogramin B resistance gene; *repL*, pIM13 Gram-positive origin of replication; ColE1 origin of replication

Table 5.1 Bacterial strains and plasmids used in this study.

Primer Name	Sequence (5'-3')	Description
654astop	GATCTAATTAAAGCTATAGTCAACATACTATAAAGTAGAGCG	654 asRNA top
	ATCCTACTCTTTATCTTCAGCATAGCCTATACCCCTTAGC	oligonucleotide
654asbot	GCTAAGGGGTATAGGCTATGCTGAAGATAAAGAGTAGGAT	654 asRNA top
	CGCTCTACTTTATAGTATGTTGTCTATAGCTTTAATTAGATC	oligonucleotide
653_fwd	GATCCAGGATGGTGTAAAATGCATTTGC	CAC0653 forward primer
654_rev	CTAATATTTTCTCGAAGATAAGGCGCC	CAC0654 reverse primer
sigF_fwd	AGGTTGTATGGGACTCGTAAAGGC	*sigF* RTPCR forward primer
sigF_rev	CCCATAATCATAGGAACAGCATAGGT	*sigF* RTPCR reverse primer
sigE_fwd	TAGACTGGGATGGAAACAAGCTGC	*sigE* RTPCR forward primer
sigE_rev	AGCTGTTTATCAACCTCTCCTTCT	*sigE* RTPCR reverse primer
sigG_fwd	GGAATGGAAACACAACTGCCAGAG	*sigG* RTPCR forward primer
sigG_rev	AGCCCTATACACCCAACCTGAA	*sigG* RTPCR reverse primer
sigK_fwd	CCATAGGCACCATAGGGCTTATAAAGG	*sigK* RTPCR forward primer
sigK_rev	GCAGCATAGGTTGCAAGTCTAGTACC	*sigK* RTPCR reverse primer
abrB_fwd	CGCTTTAGAAATTTATGTAGATGGTGAGC	*abrB* RTPCR forward primer
abrB_rev	TGATAACATCGCTTGCATCTCCGC	*abrB* RTPCR reverse primer

Table 5.2 Primers used in this study

5.2.3 Crude cell extracts

Cell pellets of 2-10 mL were collected and washed with sterile water and stored at -85°C until purification. For purification, cell pellets were resuspended in 150 μL of SDS gel loading buffer (50 mM Tris pH 6.8, 2% SDS, 10% glycerol, 5% β-mercaptoethanol) and boiled for three minutes. Samples are spun at 16,000 rpm for three minutes and the supernatant collected and stored at -85°C until use. Protein samples are quantitated using a RC DC protein assay kit (Biorad).

5.2.4 Western blot analysis

25 μg of total protein was separated per lane in a SDS-PAGE (12% polyacrylamide, 4% stacking) Ready Gel (Biorad). MagicMark XP (Invitrogen, Carlsbad, CA) was used as a protein standard. Proteins were transferred to a PVDF membrane overnight. Membranes were blocked with Tris-buffered saline plus Tween 20 (TBST) and 4% ECL blocking agent (Amersham Biosciences, Piscataway, NJ) before hybridization. The blocked membrane was incubated with rabbit anti-CAC0654 polyclonal antiserum at 1:1000 dilution. Goat anti-rabbit immunoglobulin conjugated with horseradish peroxidase (1:10,000 dilution) was used as a secondary antibody (Sigma, St. Louis, MO) in TBST-4% ECL blocking agent. The conjugated antibodies were fluorescently reacted with an ECL+ kit (Amersham Biosciences) and detected with a Storm 860 Imager (Molecular Dynamics, Sunnyvale, CA). ImageQuant version 5.0 (Molecular Dynamics) was used to view and quantify Western blots.

5.2.5 Antibody Generation

The full coding sequence of CAC0653 and a truncated portion excluding the transmembrane domain of CAC0654 (amino acids 252-513) were PCR amplified. The PCR products were sent to Protein Tech Inc. (Chicago, IL) for protein expression and raising of the rabbit antibodies. Final production bleeds were used in all Western blots.

5.2.6 Microscopy

For phase contrast microscopy, cells were collected by centrifugation and resuspended 10-100 fold in a 1% NaCl solution. Cells were stored at 4^{o}C until imaged. 1-2 μL was placed on a microscope slide and imaged using a Leica widefield microscope with phase contrast lens at 1000X magnification. For electron microscopy, samples were fixed by addition of 16% paraformaldehyde and 8% glutaraldehyde to the culture medium for a final concentration of 2% paraformaldehyde and 2% glutaraldehyde. Cultures were fixed for 1 h at room temperature, pelleted and resuspended in buffer. Samples were embedded and analyzed as previously described (Jones et al. 2008).

5.2.7 cDNA probe synthesis and labeling for microarray analysis

Random hexamer primed reverse transcription reactions using Superscript II (Invitrogen) reverse transcriptase with aminoallyl dUTP were used as previously described (Alsaker, Paredes et al. 2005). Reactions were purified using Millipore YM-30 columns and dried in a rotary SpeedVac. cDNA samples were resuspended in 4.5 μL of 0.1 M sodium bicarbonate buffer pH 9.0 and incubated at room temperature in the dark for one hour with Cy3 or Cy5 dye (Amersham). Samples were purified on a Qiagen Qiaquick PCR purification column. Samples were stored at -20°C.

5.2.8 Microarray hybridizations and analysis

cDNA probes were mixed together, dried in a rotary SpeedVac and resuspended in Pronto Universal Hybridization Solution. Samples were hybridized to the microarray and incubated for 16 hours at 42°C as previously described (Alsaker, Paredes et al. 2005). Microarrays were scanned using a G2656BA Agilent Scanner and quantitated using GenePix Pro software version 5.0 (Molecular Devices, Sunnyvale, CA). Microarrays were normalized using a segmental nearest neighbor logarithmic expression ratio of the mean (Yang, Haddad et al. 2003) using 200 genes per interval, 25 intervals, and a gamma factor of 2.0. Hierarchical clustering is performed and visualized using MEV (Saeed, Sharov et al. 2003).

5.2.9 Assignment of *C. acetobutylicum* homologs to *B. subtilis* genes

The putative set of homologous genes between *B. subtilis* and *C. acetobutylicum* was determined using a best-best BLASTP search (Brüggemann, Baumer et al. 2003). Briefly, for each known *B. subtilis* gene a BLASTP (Altschul, Madden et al. 1997) search against the *Clostridium acetobutylicum* genome was carried out and the best scoring *C. acetobutylicum* gene was used as query for a 2nd BLASTP search against the *B. subtilis* genome. If the best scoring *B. subtilis* gene from this 2nd search and the original *B. subtilis* gene were the same and they share more than 30% amino acid identity over more than 60% of the length of both genes (Brüggemann, Baumer et al. 2003), we assumed that the *B. subtilis* and the *C. acetobutylicum* genes were homologs. The complete list of homolog genes is available in Appendix B.

5.2.10 Identification of putative regulons for Spo0A, σ^H, σ^F, σ^E, σ^G, and σ^K

The regulons of the sporulation specific sigma factors and Spo0A have been reported for *B. subtilis* (Britton, Eichenberger et al. 2002; Eichenberger, Jensen et al. 2003; Molle, Fujita et

al. 2003; Eichenberger, Fujita et al. 2004; Steil, Serrano et al. 2005; Wang, Setlow et al. 2006) however, similar in depth studies are not available for any clostridia. Following our previous mapping of the sporulation cascade in *C. acetobutylicum* (Paredes et al. Nat Rev Micr. 2005) we have combined the *B. subtilis* regulons of the sporulation specific sigma factors and Spo0A, with the *C. acetobutylicum*'s annotation paper (Nölling, Breton et al. 2001) to infer the putative sporulation regulons in *C. acetobutylicum*. For those genes not identified as *B. subtilis* homologs in the *C. acetobutylicum*'s annotation, homologs were identified using the double BLAST explained before. The inferred sporulation regulons are available in Appendix C.

5.2.11 Similarity searches for the CAC0653/CAC0654 pair using the KEGG database (http://www.genome.jp/kegg).

The KEGG SSDB (Sequence Similarity DataBase) contains information about amino acid sequence similarity among all protein-coding genes in completed genomes. All possible pairwise genome comparisons are performed by the SSEARCH program (http://www.biology.wustl.edu/gcg/ssearch.html), and gene pairs with a Smith-Waterman similarity score of 100 or more are entered in SSDB, together with information about best hits and bidirectional best hits (best-best hits). We first searched the KEGG SSDB for either CAC0653 or CAC0654 to identify potentially conserved gene clusters http://www.genome.jp/kegg/ssdb/. CAC0653 (or CAC0654) and its best-best hits were considered as an initial cluster. Neighboring genes on both sides of the chromosome were included in the cluster as long as they were also best-best hits. Gapped genes were included in the cluster if they were forward best hits.

5.3 Results

5.3.1 Expression profile of CAC0653 and CAC0654 in the wildtype (WT) *C. acetobutylicum* ATCC 824 strain

Computational predictions indicate that CAC0653 and CAC0654 are organized in a putative bicistronic operon flanked by rho-independent terminators upstream of CAC0653 and downstream of CAC0654 (Paredes et al. 2004). Examination of the expression profile of CAC0653 and CAC0654 in the WT strain shows that the induction of this two-component system is closely associated with induction of genes required for sporulation and solvent production (Fig 5.1). The timing of CAC0653 and CAC0654 expression suggests that the two-component system may be involved in sensing an environmental signal that regulates entry into the sporulation program and/or the production of solvents. The histidine kinase CAC0654 is a putative membrane-bound protein containing two transmembrane domains (Doss et al. 2005).

5.3.2 Functional analysis of CAC0653/654 system

We first aimed to construct deletion mutants for the two genes using a method which successfully inactivated genes in *the C. acetobutylicum* M5 strain (Sillers et al. 2008). Briefly, a new antibiotic selection marker was constructed that utilizes the high expression of the *ptb* promoter to generate higher antibiotic resistance gene expression to aid the viability and selection of integrational mutants. Despite the success of this method in the M5 strain (Sillers et al. 2008), no mutants could be isolated for either the CAC0653 or the CAC0654 gene using this approach. A replicative plasmid approach that was successful in the generation of the *spo0A* knockout strain (Harris et al. 2002) was also employed, but still no mutants could be isolated. A third new method that proved successful in inactivating several sigma factors in the WT strain

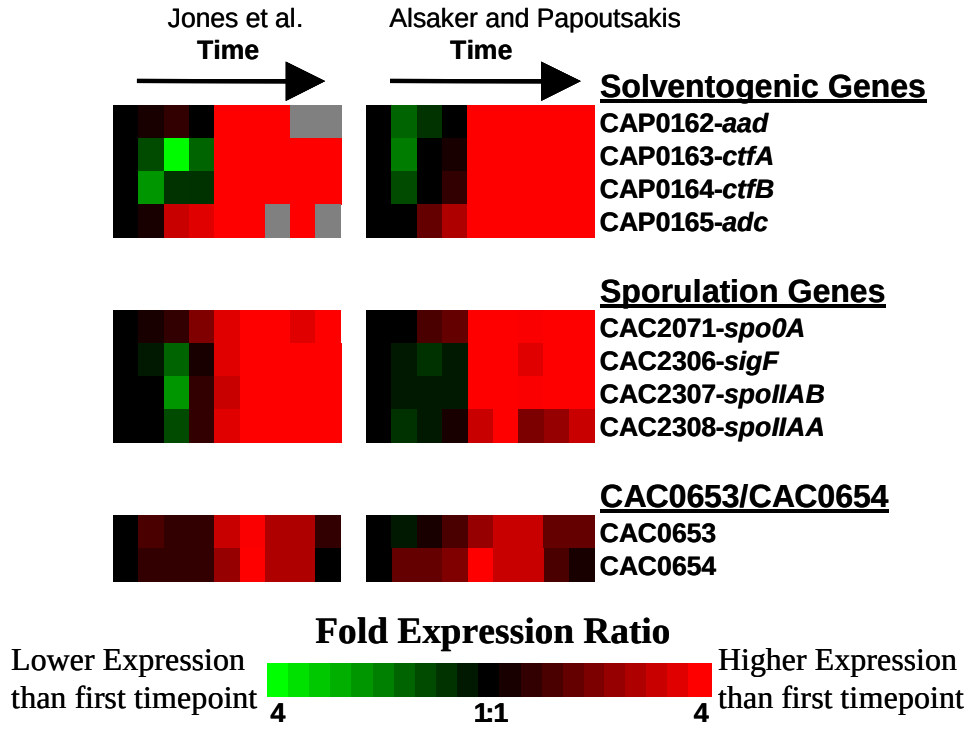


Figure 5.1 Expression profiles of solventogenic genes, sporulation genes, and the CAC0653/CAC0654 two component system in the *C. acetobutylicum* wild-type strain.

Colorimetric expression profiles are shown from an earlier microarray study (Alsaker and Papoutsakis 2005) and a recent microarray study (Jones et al. 2008) . Expression ratios are shown relative to the first timepoint of each study at hour 6. The induction of genes required for solvent formation and those related to the early sporulation program are shown to correlate with the induction of the CAC0653/CAC0654 two component system. The timepoints in each study are from hour 6, 7, 8, 9, 10, 12, 14, 16, and 18 of each set of experiments.

(data not shown) also proved unsuccessful. The lack of success using these methods points to two possibilities: one, that the CAC0653 and CAC0654 two component system is necessary for cell survival, or two, that the inactivation methods developed thus far are not effective in inactivating all gene targets. As deletion mutants could not be isolated, we utilized an asRNA approach that although is not as definitive as a knockout, has shown high utility in the downregulation of several target genes in *C. acetobutylicum* (Desai and Papoutsakis 1999; Scotcher and Bennett 2005; Tummala et al. 2003b) as well as in other clostridia (Carter et al. 2005; Raju et al. 2007). Gene knockdown (KD) using RNAi is also widely used in mammalian systems where gene deletions are very difficult. The histidine kinase CAC0654 was targeted using asRNA as response regulator domains tend to be highly conserved, while sensory domains of histidine kinases show greater variation (Fabret et al. 1999). Furthermore, the CAC0653 and CAC0654 operon was overexpressed to test the impact of increased levels of the two-component system on the cells.

The 824(pT6534) strain overexpressing CAC0653 and CAC0654 was utilized to confirm the overexpression and size of the respective proteins by Western analysis (Fig 5.2). Antibodies against a truncated version of the CAC0654 protein, excluding the transmembrane domain, were raised and used for Western analysis. Antibodies were generated against the full protein of the response regulator coded by CAC0653. When the levels of each protein in the overexpression strain were compared to the plasmid control strain, 824(pSOS95del), the protein levels in the overexpression strain were readily identifiable, but not always so in the 824(pSOS95del) control strain (Fig. 5.2A). Four biological replicates were analyzed for expression of the CAC0654 and two for the CAC0653 comparing the overexpression, control and asRNA strains. In only one experiment comparing the control strain to the asRNA strain was the protein level corresponding

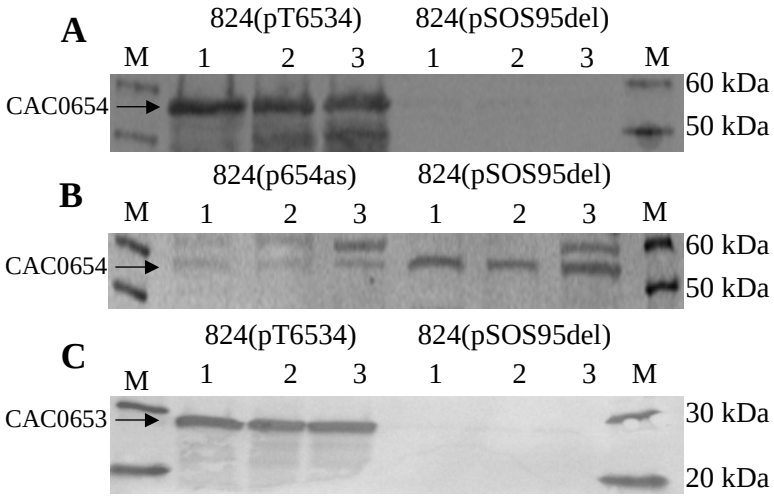


Figure 5.2 Western analysis of strains 824(p654as), 824(pSOS95del), and 824(pT6534).

25 μg of crude protein was loaded per sample from three timepoints, 1, 2 and 3, corresponding to exponential, transitional and stationary phase, respectively, of the cultures, A. Western blot of CAC0654 comparing protein abundance in crude protein extracts from 824(pT6534) and 824(pSOS95del). The arrow indicates the CAC0654 band. The CAC0654 bands are very faint in strain 824(pSOS95del). B. Western blot comparison of CAC0654 protein levels in 824(p654as) compared to 824(pSOS95del). C. Western blot comparison of CAC0653 protein levels in 824(pT6534) compared to 824(pSOS95del). The CAC0653 bands are very faint in strain 824(pSOS95del).

to CAC0654 detectible (Fig. 5.2B). Although no loading control system is available in this organism (there exist very few antibodies against *C. acetobutylicum* proteins), assuming accurate protein quantification, in the case of Fig. 5.2B, the KD extent in the levels of the CAC0654 protein was between 50 and 80%; such KD levels are consistent with previously reported asRNA KD of various proteins (Tummala et al. 2003b). Although gene expression of this two-component system could be reproducibly quantified by microarray analysis (Fig. 5.1), when Northern blots were utilized to determine the level of CAC0654 downregulation, only very faint bands could be observed above background and could not be reliably and reproducibly quantitated with such low intensities (generally, Northern blot analysis is not very sensitive for low abundance transcripts in this organism). These and the Western-analysis data suggest that this two-component system is expressed at low levels and that the corresponding proteins are in low abundance in the cells.

Despite the lack of definitive validation of the asRNA strain, a very reliable phenotype was always observed that resulted in a dramatic decrease of solvent production, as is detailed below.

5.3.3 Knockdown of CAC0654 alters cellular morphology and accelerates sporulation while overexpression of CAC0653 and CAC0654 appears to delay sporulation

Static-flask cultures of the 824(p654as), 824(pT6534) and 824(pSOS95del) were compared using phase contrast microscopy over a 4-day period to assess the impact of upregulation and downregulation of CAC0653 and/or CAC0654 proteins on the sporulation program. During the early stages of growth of the cells (less than two days), dramatic differences between the strains became apparent (Fig. 5.3). In the CAC0654 asRNA strain,

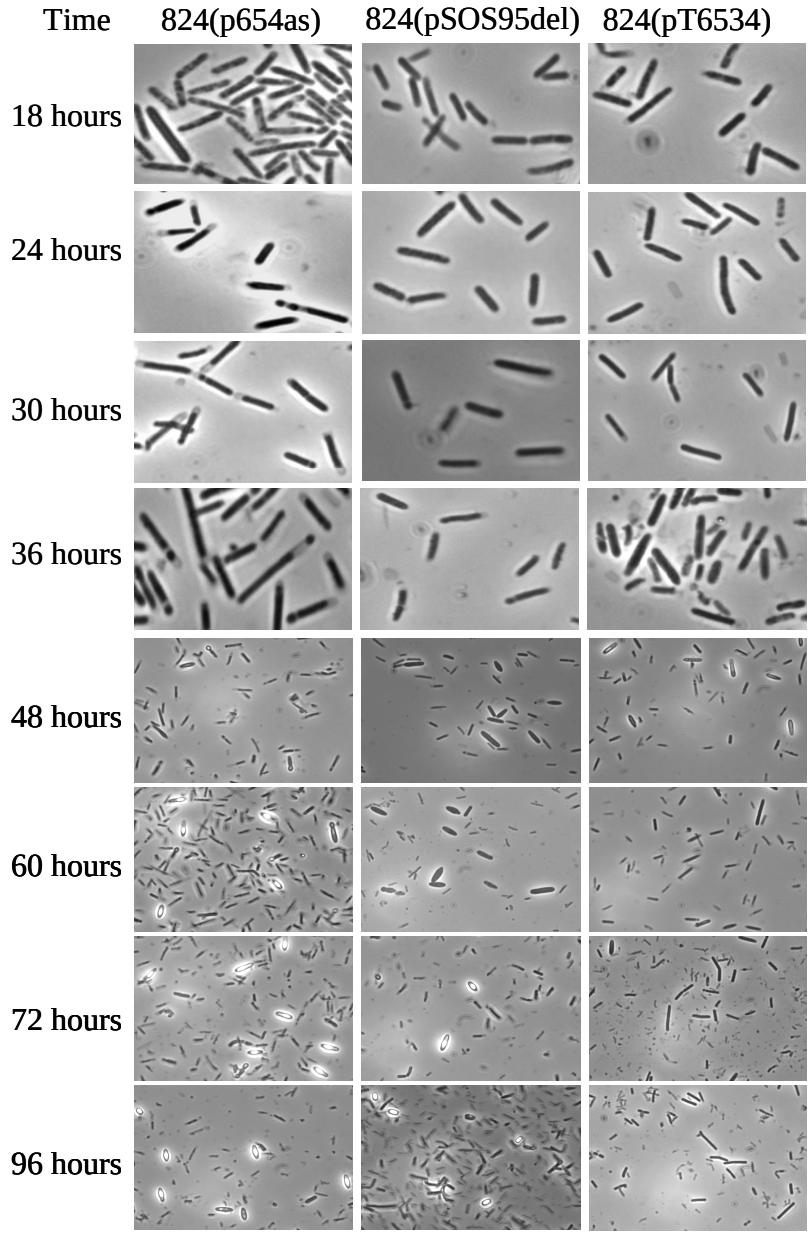


Figure 5.3 Phase-contrast microscopy analysis of strains 824(p654as) , 824(pSOS95del), and 824(pT6534).

Samples from 824(p654as), 824(pSOS95del), and 824(pT6534) were taken throughout a 4-day period of the cultures and visualized using phase contrast microscopy at 1000X magnification. Samples earlier than 48 hours are enlarged to show differences at the cellular level.

824(p654as), transparent polar bodies became visible in the majority of the largely elongated cells beginning at 18 hours and continuing until ca. hours 36. Such morphology was observed much less frequently in the control and overexpression strains. At ca. 60 hours, in the 824(p654as) strain, we observed a good frequency of phase-bright cells indicative of cells containing mature endospores and by 60 hours free spores were observed, 12-24 hours earlier than in the control strain. The overexpression 824(pT6534) strain exhibited morphology with somewhat delayed sporulation compared to the plasmid control and certainly the 824(p654as) strain. All three strains apparently completed sporulation, at least to some extent, as judged by the fact that they survived the standard heat shocking..

We sought to further support these findings by examining the cells from the 3 strains using TEM and SEM analysis from a separate set of two biological replicates. SEM images support the presence of swollen cells at hour 32-35 in the 824(pT6534) strain but not in the 824(pSOS95del) and 824(pT6534) strains (data not shown). Otherwise SEM images could not clearly distinguish cellular differences between the 3 strains. TEM images provided a better resolution of cellular differences. For example, at hour 18, the 824(p654as) strain displays some polar DNA condensation while the other 2 strains do not (Fig. 5.4). By hour 32-35, most cells of the 824(pSOS95del) strain display the electron-translucent granules characteristic of clostridial development (Jones et al. 2008) as well as a well formed prespore, while that happens at a much lower frequency in the other 2 strains (Fig. 5.4). By hour 82-87, 824(pSOS95del) cells display a high frequency of mature spores. Many of the mature spores of 824(p654as) show an accumulation of extra spore coat proteins, while several sporulating cells of the 824(pT6534) strain display discontinuous spore coat fragments indicative of an earlier stage of sporulation

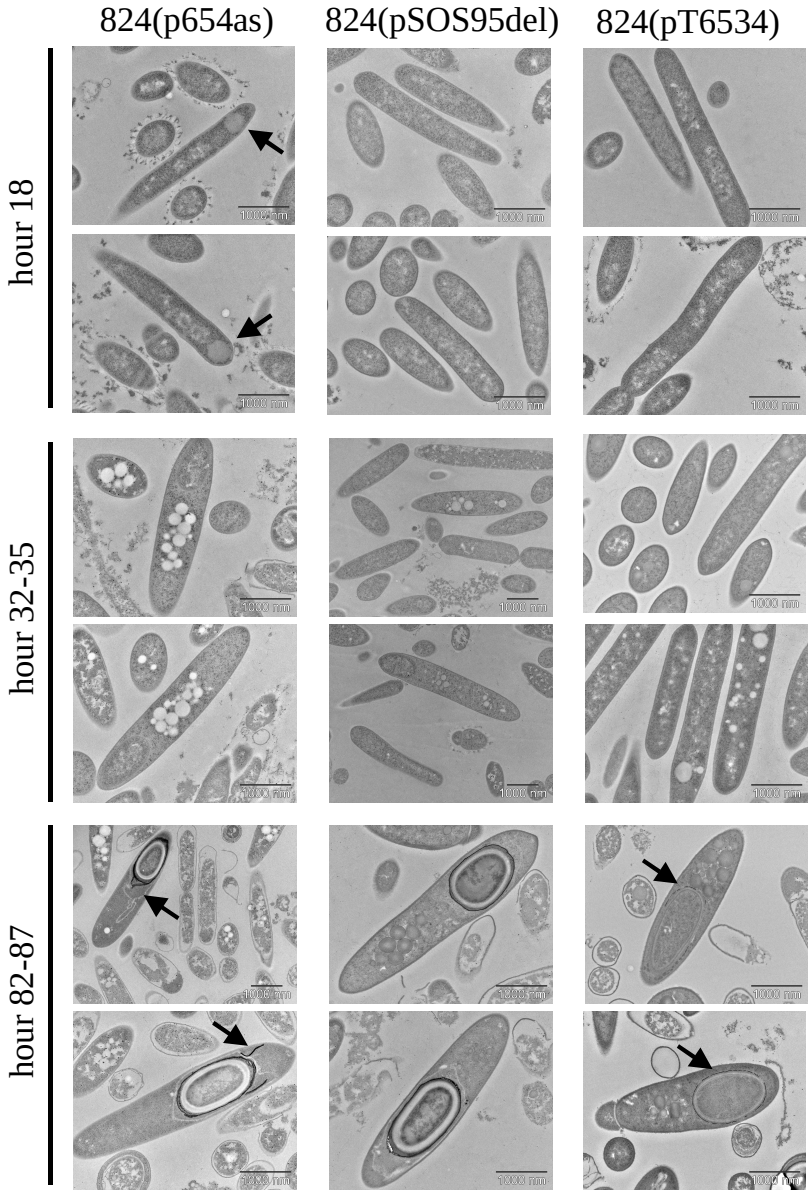


Figure 5.4 TEM images of 824(p654as), 824(pSOS95del), and 824(pT6534).

Arrows at hour 18 show the polar body formation exhibited in 824(p654as). Arrows at hour 82-87 show the extra spore coat protein produced by 824(p654as) and the discontinuous spore coat formation of 824(pT6534).

(Labbe 2005). These data support the hypothesis that the proteins coded by CAC654/653 directly or indirectly affect sporulation.

5.3.4 AsRNA knockdown of the CAC0654 protein results in dramatically reduced solvent formation and acid re-assimilation, but overexpresssion of the two-component proteins resulted in relatively small reduction in solvent formation

In flask batch cultures without pH control, strain 824(p654as) produced significantly lower (Table 5.3) levels of solvents than the plasmid control strain 824(pSOS95del), while strain 824(pT6534) displays a ca. 20% reduction (which we do not considersignificantly different) in solvent formation compared to the control strain (Table 5.3).

Next, for better controlled culture conditions, we used replicate pH-controlled fermentations to determine the effects of CAC0654 asRNA KD on the metabolism and transcriptional program of the organism compared to the plasmid control strain, 824(pSOS95del). The growth rate and peak optical density (A_{600}) measurements were very similar between the two strains, but the A_{600} of 824(p654as), after reaching its peak, continued to drop over the course of the culture, while 824(pSOS95del) maintained a higher A_{600} (Fig. 5.5). Furthermore, CAC0654 KD resulted in a dramatic reduction in the formation of solvents, reduced acid re-assimilation, and lower glucose consumption than the control strain (data not shown). Butanol and acetone titers were reduced by over 80% in 824(p654as) compared to the control strain. Butyrate, which the cell typically re-assimilates into solvents, remained elevated in 824(p654as), and was twice the final butyrate concentration of 824(pSOS95del).

Strains	End Product Concentrations (mM)				
	Butanol	Ethanol	Acetone	Acetate	Butyrate
824(pSOS95del)[a]	160	21	102	12	23
824(p654as)[a]	41	6	31	25	34
824(pT6534)[b]	127	16	80	10	21

[a] Cultures were performed in triplicate. The standard deviation of the replicates is less than 10%

[b] The values from 824(pT6534) are from a single experiment.

Table 5.3 Product formation in static flask fermentation experiments: effect of CAC0653/654 knockdown and overexpression.

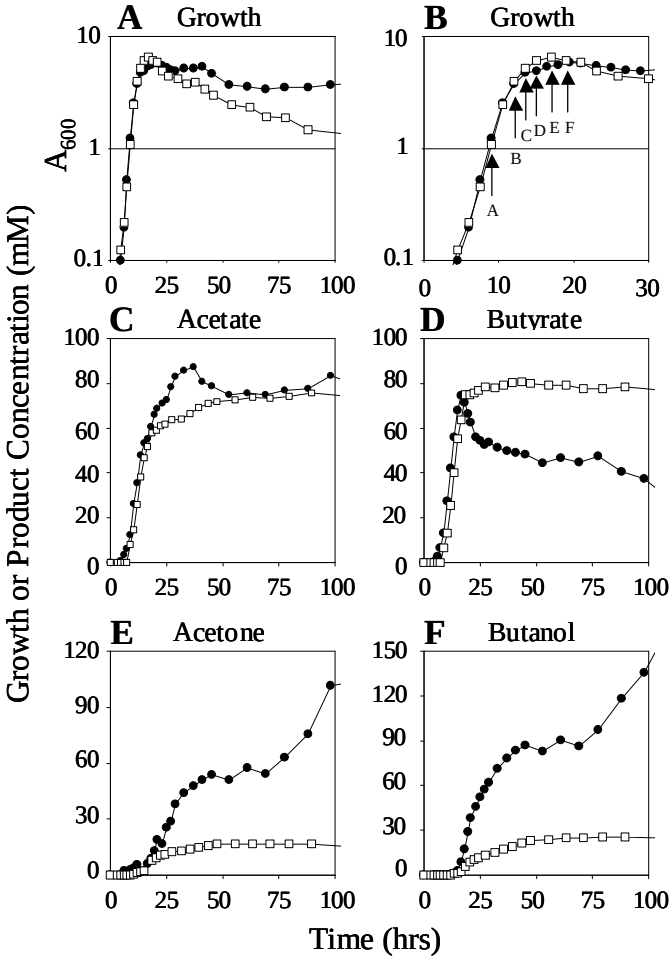


Figure 5.5 Growth curves and product formation kinetics from pH-controlled fermentation experiments.

Symbols: 824(p654as) (□); 824 (pSOS95del) (●). A: Growth curves for the whole 100-hour experiments. B: growth curve for the first 30 hours for more precisely identifying the sampling points (A-F, corresponding to hours 9, 12, 13.5, 15, 17, and 19, respectively) for microarray analysis. C-F: Carboxylic acid and solvent product formation kinetics. Results shown are averaged from replicate experiments. Concentration differences between experiments were typically less than 5 mM and error bars are not shown. Cell density (A600) differences were below 10%.

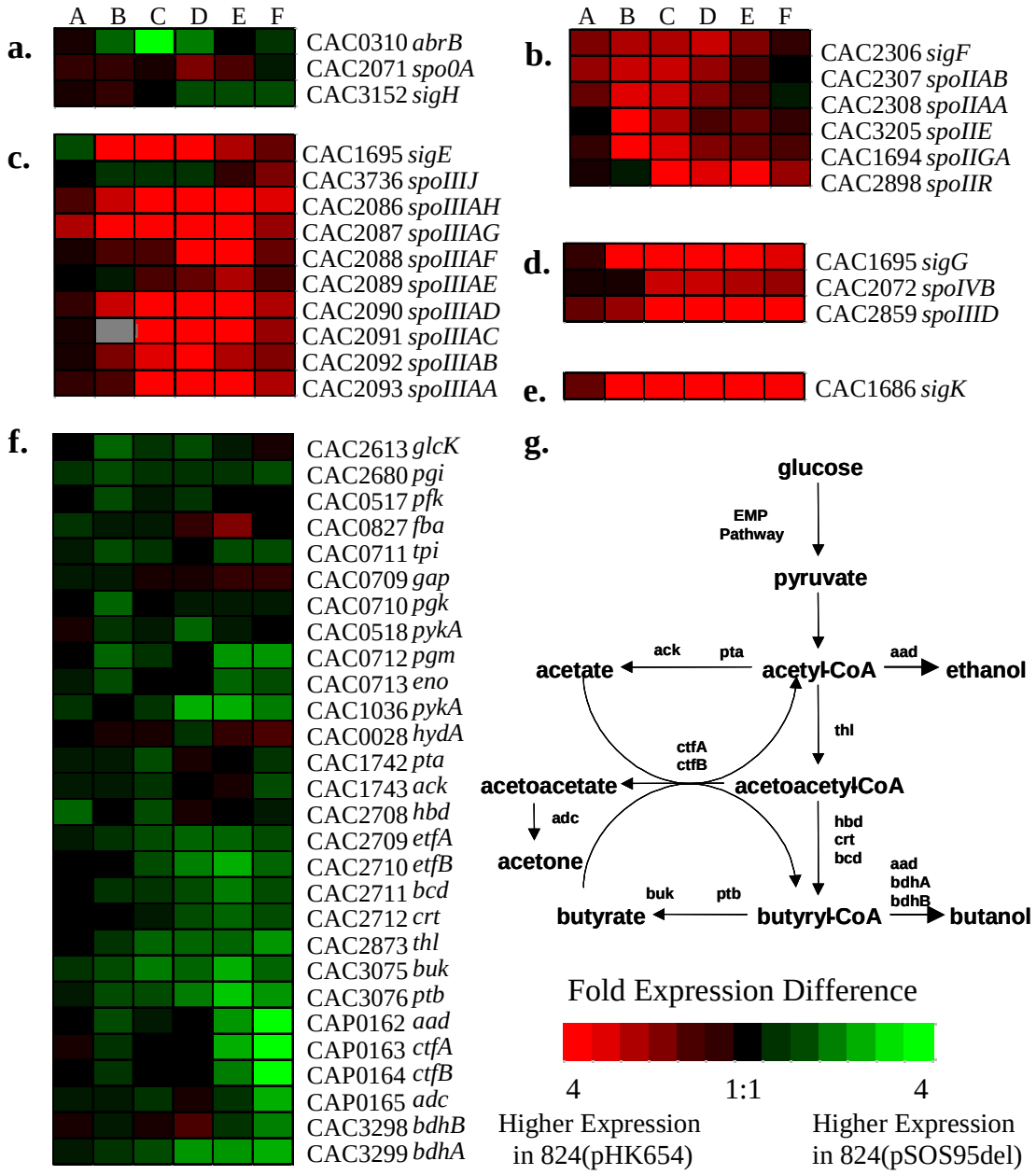
5.3.5 Microarray analysis of strain 824(p654as) shows that the CAC0654 asRNA knockdown results in an accelerated transcriptional sporulation program

Transcriptional analysis was performed on the CAC0654 asRNA strain using whole genome microarrays (Alsaker et al. 2005). Samples were taken from the pH-controlled fermentations (Fig. 5.5B) during exponential and stationary phase to capture the transcriptional effects following peak expression of CAC0654 (Fig. 5.1). Labeled cDNA from RNA isolated from samples of the 824(p654as) strain were hybridized against the corresponding cDNA from the plasmid control (824(pSOS95del)) strain at the same stage of culture for six timepoints (A through F; Fig. 5.5B).

Sporulation related transcription and sigma factors. Examination of the transcriptional profiles of key sporulation related genes revealed a dramatic upregulation of this gene subset. Initiation of sporulation in *C. acetobutylicum* is controlled by the master transcriptional regulator, Spo0A (Harris et al. 2002). This begins the signal cascade that results in the expression of other sigma factors that are necessary for sporulation. Each factor in the cascade is necessary for the expression or activation of the next. The cascade in expressed in the following order: σ^H, Spo0A, σ^F, σ^E, σ^G, σ^K (Dürre and Hollergschwandner 2004; Jones et al. 2008; Paredes et al. 2005). *spo0A* is transcribed from a σ^H dependent promoter and, once phosphorylated, drives the expression of the *sigF*, *sigE* and *spoIIE* and represses *abrB*, a repressor of sporulation (Scotcher et al. 2005). Neither *spo0A* nor *sigH* showed any significant difference in expression between the two strains. *abrB* showed overall downregulation in strain 824(p654as). Almost all other genes after *spo0A* in the sporulation cascade were highly upregulated in the 824(p654as) strain compared to the plasmid control strain (Fig. 5.6). This upregulation of sporulation-related genes occurred at timepoints A or B (Fig. 5.5), closely following the initiation of butanol formation

Figure 5.6 Temporal expression patterns of sporulation related sigma factors and associated genes (a-e) and of primary metabolism genes (g).

Ratios compare expression in 824(p654as) to 824(pSOS95del). A-F time points at the top of each gene group corresponds to the timepoints shown in Fig. 5.2. (a) Genes involved in early sporulation signaling. (b) *sigF* and genes required for σ^E processing. (c) *sigE* and genes required for σ^G processing. (d) *sigG* and genes required for *sigK* expression. (e) *sigK* expression. (f) Primary metabolic genes and abbreviated diagram of metabolism (g). Red indicates higher expression in the CAC0654 asRNA strain 824(pHK654) and green indicates higher expression in the plasmid control strain, 824(pSOS95del). Grey boxes indicate missing data on the microarrays.


A B C D E F
a.
CAC0310 abrB
CAC2071 spo0A
CAC3152 sigH
b.
A B C D E F
CAC2306 sigF
CAC2307 spoIIAB
CAC2308 spoIIAA
CAC3205 spoIIE
CAC1694 spoIIGA
CAC2898 spoIIR
c.
CAC1695 sigE
CAC3736 spoIIIJ
CAC2086 spoIIIAH
CAC2087 spoIIIAG
CAC2088 spoIIIAF
CAC2089 spoIIIAE
CAC2090 spoIIIAD
CAC2091 spoIIIAC
CAC2092 spoIIIAB
CAC2093 spoIIIAA
d.
CAC1695 sigG
CAC2072 spoIVB
CAC2859 spoIIID
e.
CAC1686 sigK
f.
CAC2613 glcK
CAC2680 pgi
CAC0517 pfk
CAC0827 fba
CAC0711 tpi
CAC0709 gap
CAC0710 pgk
CAC0518 pykA
CAC0712 pgm
CAC0713 eno
CAC1036 pykA
CAC0028 hydA
CAC1742 pta
CAC1743 ack
CAC2708 hbd
CAC2709 etfA
CAC2710 etfB
CAC2711 bcd
CAC2712 crt
CAC2873 thl
CAC3075 buk
CAC3076 ptb
CAP0162 aad
CAP0163 ctfA
CAP0164 ctfB
CAP0165 adc
CAC3298 bdhB
CAC3299 bdhA
g.
glucose
EMP Pathway
pyruvate
acetate
ack
pta
acetyl-CoA
aad
ethanol
thl
ctfA
ctfB
acetoacetate
acetoacetyl-CoA
adc
acetone
hbd
crt
bcd
aad
bdhA
bdhB
buk
ptb
butyrate
butyryl-CoA
butanol
Fold Expression Difference
4
1:1
4
Higher Expression in 824(pHK654)
Higher Expression in 824(pSOS95del)

(first detected at timepoint B). Notably, *sigF* and its regulon appear to be upregulated in strain 824(p654as) compared to plasmid control at point A (hour 12), which appears to be early than in the WT strain (Fig. 5.1), and this is consistent with the observation that sporulation is accelerated in strain 824(p654as) (Figs. 5.3 and 5.4). This upregulation of sporulation genes also correlated closely with the timing of peak expression of the CAC0653/CAC0654 operon in the wildtype strain (Fig. 5.1). Q-RT-PCR analysis of several important sporulation related sigma factors were used to validate the results from the microarray analysis. To do so, we carried out separate experiments and under slightly different culture conditions (no pH control) in order to ascertain the broader validity of our microarray findings, although we anticipated that some of the gene expression patterns would be qualitatively different given that both solvent formation (and thus sporulation) are pH dependent (e.g.,(Husemann and Papoutsakis 1988; Roos et al. 1985)). The transcriptional kinetics of the sporulation-related transcription and sigma factors *abrB*, *sigF*, *sigE*, *sigG* and *sigK* were analyzed by Q-RT-PCR in such a culture (Fig. 5.7), and the results were found to be generally similar to the corresponding microarray expression profiles, with some qualitative differences, and notably in *sigF* and *sigK* expression. Taken together, these data suggest that the 2-component system encoded by CAC0654/0653 acts as a transcriptional repressor of sporulation.

Genes of primary metabolism, motility, and chemotaxis. In agreement with the product-formation profiles (Fig. 5.5), the genes involved in solvent formation (and many of the primary metabolic genes) displayed a significant downregulation in 824(p654as) compared to 824(pSOS95del) (Fig. 5.6f and g). For solvent-formation genes (the *sol* operon, *aad-ctafA-ctfB* and *adc*), the differences became especially pronounced late, namely after about hour 17, which corresponds to point E of Fig. 5.5B). These data show that the accelerated sporulation of

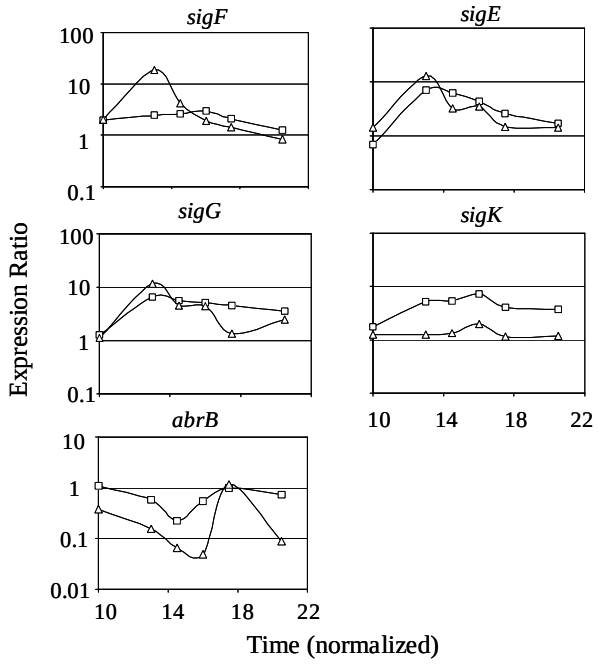


Figure 5.7 Comparison of expression data from microarray analysis with data from Q-RT-PCR.

The expression ratios shown are either from microarray data (Δ) for the experiment of Fig. 5.5, or from Q-RT-PCR data (□) from static flask cultures (no pH control). The ratios were calculated from the expression of each gene in 824(p654as) divided by the expression of the gene in 824(pSOS95del). The time scale for each culture was normalized to 10 hours at an A_{600} of 1.0 to correctly compare different cultures in the same stage of growth.

824(p654as) (Fig. 5.3) results in premature dormancy of metabolism by lowering the expression levels of the solvent-formation (but also other primary-metabolism) genes and thus resulting in reduced solvent production.

As sporulation advances, chemotaxis and motility are progressively lost, and thus transcriptional analyses of the asporogenous M5 (Alsaker and Papoutsakis 2005) and the Spo0A knockout (Tomas et al. 2003b) strains displayed increased expression of many chemotaxis and flagellar genes when compared to the WT strain. Here then, as would be expected from the upregulation of many sporulation related genes, and thus accelerated sporulation, in 824(p654as) compared to 824(pSOS95del), the large flagellar and chemotaxis operons were largely downregulated in 824(p654as) (Fig. 5.8).

Deduced Spo0A and sporulation sigma-factor activity. Spo0A must be phosphorylated to become active and other sigma factors also require post-translational modifications to bind DNA (e.g., σ^F requires SpoIIE for release from its anti-sigma factor and σ^E must be proteolytically cleaved to gain functionality). Thus, transcriptional upregulation of these important sporulation genes does not imply upregulation of that the encoded functional protein(Jones et al. 2008). However, activity of a sigma or transcription factors can be inferred by inspecting the expression profiles of their putative regulons, i.e., of the genes directly regulated by the sigma/transcription factor (Jones et al. 2008; Paredes et al. 2005). Information regarding the regulons of the sporulation specific sigma factors and Spo0A was drawn from recent genomic studies in *B. subtilis* ((Britton et al. 2002; Eichenberger et al. 2004; Eichenberger et al. 2003; Molle et al. 2003; Steil et al. 2005; Wang et al. 2006) and used to construct putative regulons in *C. acetobutylicum*. We erred on the side of including more genes, based on *B. subtilis* homologs, in these regulons than may actually be the case so that we would miss fewer possible interesting

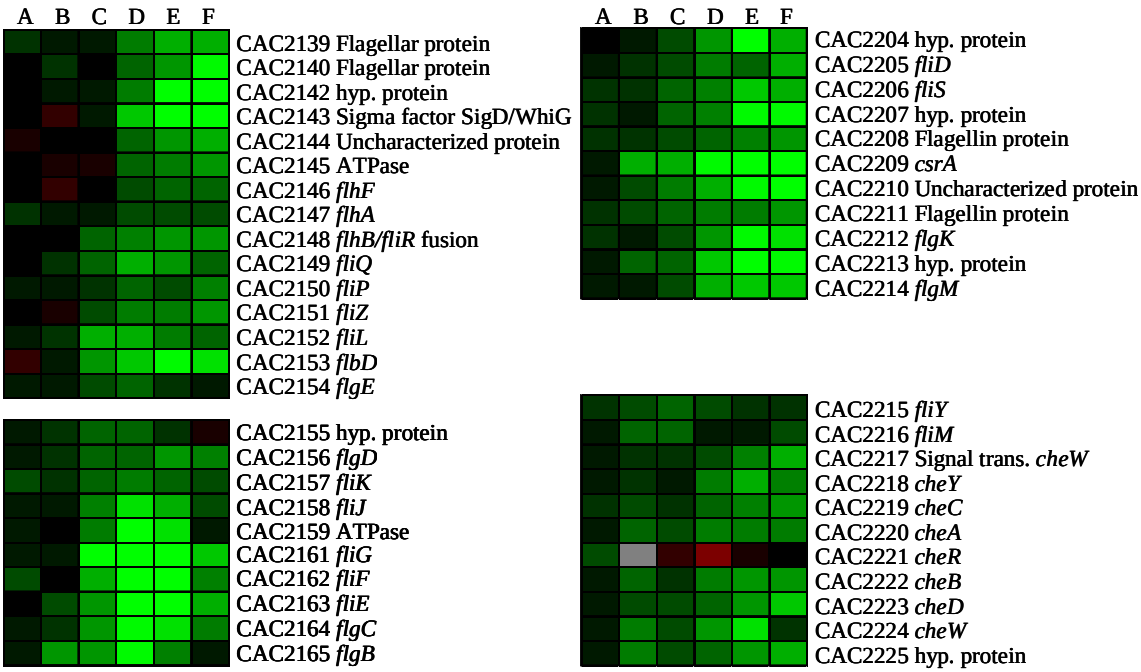


Figure 5.8 Temporal expression profiles of selected flagellar and chemotaxis genes.

Genes are organized in putative operons. Ratios compare expression in 824(p654as) to 824(pSOS95del). A-F at the top of each group correspond to the timepoints shown in Fig. 5.5B. The expression scales are the same as in Fig. 5.6. Grey boxes indicate missing data on the microarrays.

expression patterns, and let our microarray data dictate at the end which of these might belong to the true regulon of each factor. The expression profiles of these putative regulons are shown Fig. 5.9. σ^H is the first sigma factor of the sporulation cascade and showed no change in expression between 824(p654as) and 824(pSOS95del). The putative σ^H regulon is comprised of the sporulation related genes *spoVS*, *spoVG*, *ftsAZ*, and *spo0A*. These genes show very little change in expression between the two strains during the first three timepoints when σ^H should be active (i.e., before *spo0A* expression and solvent formation), indicating that CAC0654 asRNA does not affect σ^H activity. *spo0A* expression is little changed between the two strains, but the genes under control of Spo0A exhibit dramatic differences between the two strains. Based on *B. subtilis* studies (Molle et al. 2003), Spo0A positively regulates *sigF*, *sigE*, and the σ^F processing enzyme *spoIIE*, while negatively regulating the large flagellar operon (CAC2155-CAC2165) (Fig. 5.8). The *sigF* operon (CAC2306-CAC2308) was expressed higher in 824(p654as) than in 824(pSOS95del) beginning at the first timepoint of the microarray analysis, while *sigE* and *spoIIE* were upregulated at timepoint B, and finally the flagellar operon was downregulated at timepoints C and D. These data suggest that Spo0A may be phosphorylated earlier and possibly at a higher level in the CAC0654 asRNA strain than in the plasmid control strain. An earlier and stronger Spo0A phosphorylation would be consistent with the accelerated and enhanced sporulation of the 824(p654as) strain (Figs. 5.3 and 5.4). While the expression of *spo0A* and the Spo0A~P dependent genes have different expression profiles, the sigma factors downstream of Spo0A in the sporulation cascade have expression profiles largely similar to the genes in their regulons, which is consistent with our recent findings (Jones et al. 2008).

The putative σ^F regulon is the smallest of the investigated regulons, but contains several well-characterized genes such as *spoIIR*, *sigG*, *spoIIQ*, and the *gpr-spoIIP* operon, all of which

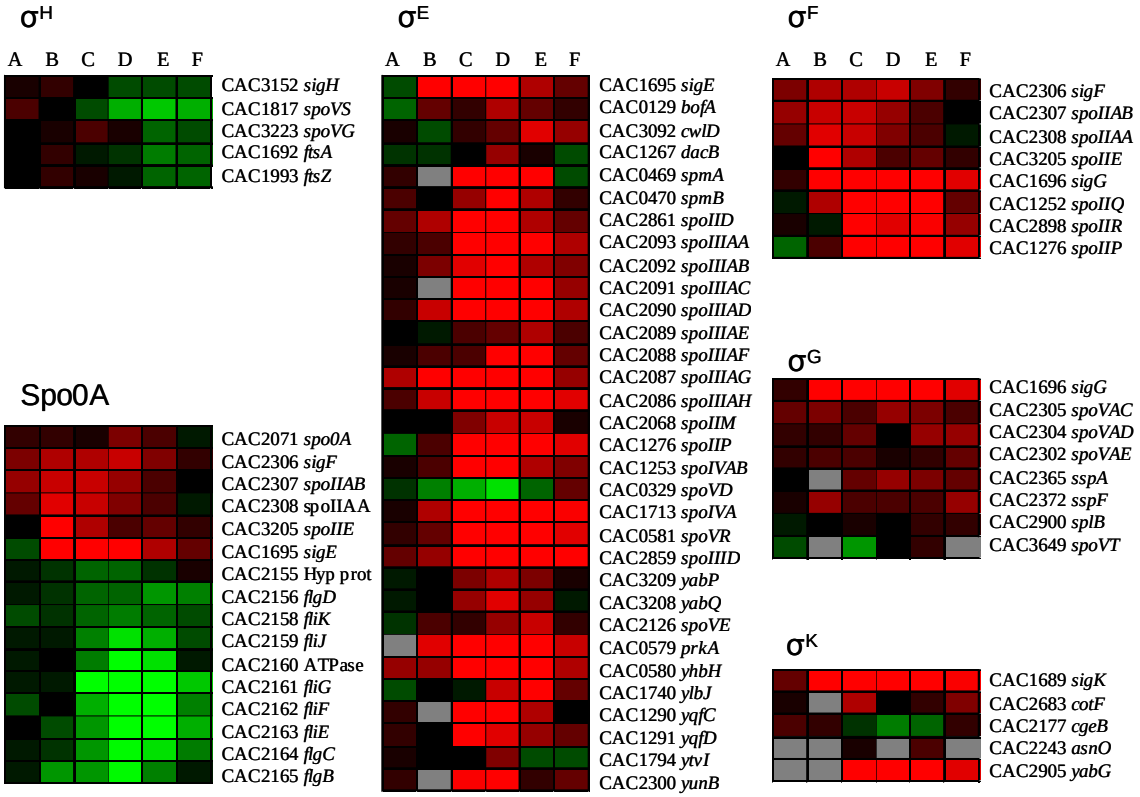


Figure 5.9 Expression patterns from microarray analysis for Spo0A and sporulation-related sigma factors with selected genes from their putative regulons.

Ratios compare expression in 824(p654as) to 824(pSOS95del). A-F at the top of each group correspond to the timepoints shown in Fig. 5.5B. The expression scales are the same as in Fig. 5.6. Grey boxes indicate missing data on the microarrays.

were, as expected, differentially upregulated. Those genes were transiently upregulated starting at timepoints B or C, closely following the upregulation of the *spoIIE*, rather than the expression pattern of *sigF*, which is upregulated starting at the first timepoint (A). SpoIIE is needed for proper σ^F functionality and this timing suggests that the genes comprising the putative σ^F regulon are upregulated due to increased *spoIIE* expression rather than increased *sigF* expression.

The σ^E putative regulon is the largest of the sigma factor regulons and contains many upregulated genes. Key differentially expressed genes belonging to this set include *spoIID*, the *spoIIIAA-spoIIIAH* operon, *spoIIID*, and *sigK*. A recent study by Eichenberger et al. (Eichenberger et al. 2003) identified both known and newly identified σ^E dependent genes required for proper sporulation in *B. subtilis*. Among those required for proper sporulation, 31 had identifiable homologs in *C. acetobutylicum*, and almost all 31 genes were upregulated in 824(p654as).

Although the transcripts of both the late sigma factors σ^G and σ^K show upregulation early in the timecourse, their putative regulons do not show the large upregulation like those of the σ^F and σ^E putative regulons. Our recent study (Jones et al. 2008) suggested that the annotated *sigK* gene may not be expressed, and that the late stage sporulation genes may be different in *C. acetobutylicum* compared to *B. subtilis*.

5.3.6 Homology searches for possible functional assignment.

A phylogenetic analysis of clostridial signal transduction kinases based on sequence similarity generated a phylogenic tree (Sabrina Doβ 2005), which groups the CAC0654 histidine kinase with several clostridial orphan kinases (i.e., kinases without a response regulator coded by

a gene adjacent to the kinase gene) comprising an orphan kinase cluster. This cluster of orphan kinases includes three (CAC0903, CAC0437 and CAC0323) of the five identifiable (Paredes et al. 2005) *C. acetobutylicum* orphan kinases. In *B. subtilis*, it is known that orphans kinases are responsible for the phosphorylation (activation) of Spo0A, and this is also likely the case in clostridia (Jones et al. 2008; Paredes et al. 2005). Although the CAC0654 kinase likely targets the CAC0653 response regulator, this analysis suggests that there may exist some promiscuity in the targeting of the CAC0654 kinase, and that this targeting may include direct Spo0A activation. The timing of the transcriptional activation of the CAC0654 gene (Fig. 5.1) is consistent with this conjecture.

We analyzed the possible conservation of the two component system coded by CAC0653/0654 using two related but distinct approaches. First, we used the KEGG SSDB (Sequence Similarity DataBase) as explained in the Methods section. Searching for either CAC0653 or CAC0654 gave the same results, namely similarity to three (one each from three different strains, B Eklund 17B, B1 Okra, and F Langeland) of *C. botulinum* two component systems, and one two-component system in each of *Sodalis glossinidius* and *Thermotoga lettingae*. The homologous gene pairs are shown in Figure 5.10. Alignment of the corresponding CAC0653 homologs is shown in Figure 5.11. None of these two-component systems has been functionally characterized, but KEGG suggests that the conserved two-component pair in *T. lettingae* is involved in copper-ion sensing (http://www.genome.jp/dbget-bin/show_pathway?tle02020+Tlet_0962), while that in *S. glossinidius* (http://www.genome.jp/dbget-bin/show_pathway?sgl02020+SG0960) is involved in iron sensing. In *C. acetobuylicum* and related clostridia, iron levels have been shown to have an important effect on solvent formation (Bahl et al. 1986; Dabrock et al. 1992).

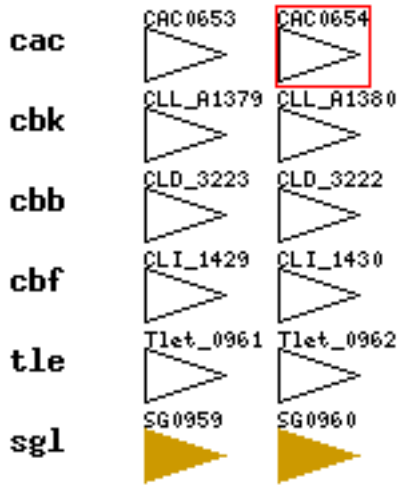


Figure 5.10 Conserved gene clusters of CAC0653/CAC0654 per the KEGG database.

Organism abbreviations are as follows: *C. acetobutylicum* (cac), *C. botulinum* B Eklund 17B (cbk), *C. botulinum* B1 Okra (cbb), *C. botulinum* F Langeland (cbf), *Thermotoga lettingae* (tle) and *Sodalis glossini*dius (sgl).

```
cbb_CLD_3223_DNA-binding_respo      -----MRLLIIEDNIELANSMKMGLEKMGFHIDVSNTGSDGEEKASINEY
cbf_CLI_1429_DNA-binding_respo      -----MRLLIIEDNIELANSMKMGLEKMGFHIDVSNTGSDGEEKASINEY
cbk_CLL_A1379_DNA-binding_resp      -----MRLIIIEDDKILANIVKLGLEEFGFSSDVANTAGDGEEKIFINEY
cac_CAC0653_response_regulator      -----MHLLIVEDDEELSILLKKGLKQYNCTCDISHDGEDGLYNLEINSY
tle_Tlet_0961_two_component_tr      MGLKKEKVLIVEDNEDLLNSIKEFLLKEGYDVSCAVDAEQAFDMVIEDHY
sgl_SG0959_DNA-binding_transcr      -----MKILVIEDDKLLREGLQQALEQESFACECASNLGQSWPLLTSTTY
                                          ::*::**:  *   ::  * : .   . :    :.       *

cbb_CLD_3223_DNA-binding_respo      DVILLDLNLPDIDGIEILNYLRSESIETPVIIVTARDEVEQLAFGLDNGA
cbf_CLI_1429_DNA-binding_respo      DVILLDLNLPDIDGIEILNYLRSESIETPVIIVTARDEVEQLAFGLDNGA
cbk_CLL_A1379_DNA-binding_resp      DTILLDLNLPDKDGLEVLKDIRESGITTPIIIITARIEVKERAIGLDLGS
cac_CAC0653_response_regulator      DAVILDINLPKMDGITVCRKMRGKGISTPVLMLTARTDTDDRVLGLDSGA
tle_Tlet_0961_two_component_tr      DCLVVDVLLPGINGFEFVEHLRESLINVPILILTALDAINDKVRGLSCGA
sgl_SG0959_DNA-binding_transcr      SLIILDLGLPDGDGMTLLEKLRRQGNGVPVLILTARDAIDERVRGLDFGA
                                    . :::*: **  :*: . . :* .   .*::::**    .: . **. *:

cbb_CLD_3223_DNA-binding_respo      DDYITKPFKLLELRARIHAIIRRFHGRTNPIINIGKLQLNPITRTVEIEN
cbf_CLI_1429_DNA-binding_respo      DDYITKPFKLLELRARIHAIIRRFHGRTNPIINIGNFQLNPITRTVEIEN
cbk_CLL_A1379_DNA-binding_resp      DDYIVKPFDLIELRARINAVIRRFYGRTNAEIKIGNLLINPKTRVVTYSL
cac_CAC0653_response_regulator      DDYLGKPFEFKELRARLHALIRRNYNKPSNEIKIKDLNIDTKAKSVKVGE
tle_Tlet_0961_two_component_tr      DDYLTKPFDFRELLARIKSLIRRSNITKGEEINFKGLKLNSRSRQVTVED
sgl_SG0959_DNA-binding_transcr      DDYLIKPFALSELLARSRAIIRRHQGVSNNLLWVDDISLDLTHHDVTVDG
                                    ***: *** : ** ** .::***     .  : .  : ::   : *

cbb_CLD_3223_DNA-binding_respo      KPVALASKEFDILEYICYRHPAVVSSEEIAEHIYDENFDPFSSVLRVHIA
cbf_CLI_1429_DNA-binding_respo      KPVALASKEFDILEYICYRHPAVVSSEEIAEHIYDENFDPFSSVLRVHIA
cbk_CLL_A1379_DNA-binding_resp      KPVKLSAKEFDVLEYIANNHPHVVSSESIAEHVYDEFFDPFSSVLRVHIS
cac_CAC0653_response_regulator      KLITLTAREYDILELLCYNYPNIVSAEEIIEHVWGDNDNQFSNVIRVHIA
tle_Tlet_0961_two_component_tr      TVLKMSKREFALLELFLRNPGIVFSREEIMEKVWENEKIIKSNVIDVYIL
sgl_SG0959_DNA-binding_transcr      KAVILTPKEFAILSRLMLRAGHRIHREILHQDLYSWDDDPSSNSLEVHIH
                                    . : :: :*: :*. :  .    .  * : :.::       *. : *:*

cbb_CLD_3223_DNA-binding_respo      RLKKKLSNASGKEILINIRGKGYVLCIE---
cbf_CLI_1429_DNA-binding_respo      RLKKKLSNASGKEILINIRGKGYVLCIE---
cbk_CLL_A1379_DNA-binding_resp      NLKKKLKLASGQELLLTLRGKGYRIWEID--
cac_CAC0653_response_regulator      NLRRKIKCSGGETLIETLKGKGYRLC-----
tle_Tlet_0961_two_component_tr      YLRNKLKPYGYDRYIQTVPGYGYRFCEE---
sgl_SG0959_DNA-binding_transcr      HLRQKIGRGR----IQTLRGFGYLLTTAEKS
                                     *:.*:        : .: * ** :
```

Figure 5.11 Alignment of CAC0653 orthologs identified by KEGG.

Organism abbreviations are as follows: *C. acetobutylicum* (cac), *C. botulinum* B Eklund 17B (cbk), *C. botulinum* B1 Okra (cbb), *C. botulinum* F Langeland (cbf), *Thermotoga lettingae* (tle) and *Sodalis glossini*dius (sgl). Symbols: (*) indicates positions which have a single, fully conserved residue. (:) indicates that one of the 'strong' groups (http://align.genome.jp/clustalw/clustalw_readme.html) is fully conserved. (.) indicates that one of the 'weaker' groups (http://align.genome.jp/clustalw/clustalw_readme.html) is fully conserved.

To further analyze the conservation of this two-component system across other species, a phylogenetic tree was constructed as follows. To identify homologs to CAC0653, the DNA binding domain of the response regulator was used, while for CAC0654, the ATP binding and phosphoacceptor domains were used in the histidine kinase BLAST. The proteins identified from each independent BLAST were then compared to identify homologous protein pairs likely to be in an operon. Once the orthologs were identified, alignments using ClustalW were used to construct phylogenetic dendrograms with branch lengths built from the alignment results, and the results are shown as Figures 5.10 and 5.11 for CAC0653 and CAC0654, respectively. These analyses show similarity of the CAC0653/0654 system to two-component systems in several other clostridia, bacilli and related organisms. The RR shows similarity to PhoP (which responds to phosphate and magnesium (Smith and Maguire 1998)), OmpR, and PmrA (iron, pH and magnesium sensing (Kox et al. 2000)) all of which are part of the lower branch of the CAC0653 dendrogram (Figure 5.10).

5.4 Discussion

Our data (Figs. 5.2-4 and Table 5.3) suggest that the altered sporulation characteristics due to CAC0654 KD are accompanied by reduced solvent formation, glucose utilization and acid re-assimilation apparently due to accelerated sporulation. CAC0653/654 overexpression has only a small impact on product formation, which is consistent with the less dramatic impact of the overexpression on the two proteins on sporulation.

As the microscopy shows, 824(p645as) has accelerated sporulation, while the overexpression strain, 824(pT6534), displays some delay in sporulation (Fig. 5.3 and 5.4). This accelerated sporulation program is consistent with the finding that most sporulation genes were

upregulated in the asRNA strain, 824(p654as), compared to the plasmid control strain, 824(pSOS95del). The expression profiles of the earliest transcriptional regulators of the sporulation cascade (σ^H and Spo0A) were largely unaffected by the asRNA KD of the CAC0654 protein, and the same is true for the putative σ^H regulon (Fig. 5.9). There are several important genes (*sigF*, *sigE*, *spoIIE*, and the flagellar operon CAC2155-2165) under putative Spo0A control that are differentially expressed between 824(p654as) and 824(pSOS95del), suggesting that the Spo0A was activated earlier and/or stronger in the CAC0654 asRNA strain. Additionally, 824(p654as) showed a premature termination of solvent formation and overall metabolism likely due to the accelerated sporulation (Fig. 5.6). Collectively, our data suggest that the CAC0653/CAC0654 response regulator and histidine kinase pair act as repressors of key sporulation genes, and most likely affecting Spo0A activation or sporulation signaling starting at SigF.

In *B. subtilis*, there a number of phosphotases that result in lower Spo0A activation (Perego 2001; Perego and Hoch 1991), but homologs of these genes either remain unidentified or are absent in clostridia (Nölling et al. 2001; Paredes et al. 2005). However, two-component systems have been also identified that affect the sporulation process in *B. subtilis*, and specifically the ComAP system, which controls many competence related genes and also the RapA phosphatase that can decrease Spo0A phosphorylation (Grossman 1995). The CAC0653 /CAC0654 two-component system may have a role in the regulation of similar, but yet unknown, proteins in *C. acetobutylicum* that may regulate the level of Spo0A phosphorylation.

The initial phenotype observed of 824(p654as) was that of reduced acid reassimilation and the concomitant decreased solvent production. The events of solvent production and sporulation are typically intertwined in this organism and based on the low solvent formation it

was initially expected that 824(p654as) would have decreased expression of the key sporulation genes. This occurs in the spo0A deletion strain where both solvent formation and sporulation are blocked (Harris et al. 2002). However, here we see an enhancement and acceleration in the sporulation program and this accelerated sporulation likely causes early cessation of cellular metabolism.

Comparisons of CAC0654 with other histidine kinases within the *C. acetobutylicum* genome show a tight clustering with orphan kinases, which are likely responsible for Spo0A phosphorylation. Homology comparisons of this two-component system to two-component systems in other organisms shows a good preservation among clostridia and bacilli, and indicate a potential role in iron or related metal sensing/regulation.

CHAPTER 6 CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

Metabolic engineering holds the potential to develop vastly improved microorganisms for various commercially relevant applications. Here we have developed a number of engineering strains for their potential use as butanol producers. In the wildtype strain, we have found that the replacement of the natural promoter of *aad* with the acidogenesis associated promoter of *ptb* can elevate transcription of *aad* 10-fold when overexpressed compared to the overexpression of the *aad* with its endogenous promoter. When this high expression level of *aad* was combined with *ctfB* asRNA used to limit acetone formation, solvent production reached record levels in this organism. In addition to butanol reaching 180 mM, ethanol production was elevated to ca. 300 mM, the highest ethanol titers reached in this organism. The addition of the *ctfB* asRNA kept acetone concentrations relatively low, about one-half the levels of the control strain. Metabolic flux analysis helped to identify a likely bottleneck between the conversion of acetyl-CoA and butyryl-CoA. To identify if butyryl-CoA limitation was prematurely ending butanol formation, thiolase overexpression was examined. Thiolase is the first enzyme in the pathway converting acetyl-CoA to butyryl-CoA. Thiolase overexpression solely resulted in an early termination of the fermentation with primarily acids produced as the products. It was thought the early termination of the fermentation was due to the early accumulation of butyrate. To overcome the early butyrate accumulation, the *aad* gene expressed from the *ptb* promoter was used in combination with thiolase overexpression. This early *aad* expression would provide another route for butyryl-CoA formation other than butyrate. When thiolase overexpression was

examined using the *ptb* driven *aad* expression, it was found to significantly reduce the side-products produced from acetyl-CoA. Acetate was reduced by 50% when thiolase and *aad* were overexpressed, while ethanol was reduced by two-thirds when compared to a strain only overexpressing *aad*. Although the higher *aad* expression and thiolase overexpression did have beneficial effects on product formation, the addition of thiolase overexpression to the strain overexpressing *aad* and *ctfB* asRNA, had little impact on product formation. This indicates that additional genes involved in the pathway converting acetyl-CoA to butyryl-CoA may be necessary to increase butanol production or that by constraining the acid re-assimilation the cell additional regulation at the genetic or metabolic intermediate level.

Following the successful utilization of the *ptb* expressed *aad* in the wildtype background we investigated its utility in a non-sporulating, non-solvent producing strain, M5. Previous studies expressing either of the two alcohol/aldehyde dehydrogenases (AADHs) in this background has lead to butanol production without acetone production, but at levels just one-half what is produced in the wildtype strain. Complementation of the M5 strain with the *aad* expressed from the *ptb* promoter resulted in increased butanol production compared to previous strains, but less than the wildtype. By increasing the pH of the culture, butanol production was increased to 150 mM, near wildtype levels. This increased butanol production is likely due to the increased expression of the *aad* at higher pH from the *ptb* promoter and perhaps in part due to the reduce toxicity of the acid products at increasing pH. Despite the increased butanol production in this strain, the acid products, particularly acetate, were still major by-products. Thiolase overexpression was again employed to limit the by-products from acetyl-CoA (acetate and butyrate) and increase butyrate and butanol production. Thiolase overexpression successfully redirected the carbon flux to the four-carbon products (butanol and butyrate) from

the two-carbon products (ethanol and acetate), but butyrate became the primary product at the expense of butanol formation. To reduce the production of the acid end-products, gene deletions strains were constructed to disrupt the acetate kinase and butyrate kinase genes individually. Genetic disruptions are particularly difficult in clostridia species and the successful mutant generation required the optimization of the expression and translation of the antibiotic marker. The new antibiotic marker was used successfully in generating both the acetate and butyrate kinase mutant strains, where previous methods had been unsuccessful. The success in the mutant generation is partly due to the use of the M5 strain, which has a higher transformation efficiency than does the wildtype strain. Analysis of the mutant strains revealed greater than 60% reduction in the respective acid production pathway compared to the parent strain. This residual acid formation is common in clostridial strains where one pathway has been downregulated or inactivated the analogous pathway exhibits a compensatory increase of activity in the other pathway. Although disruption of the acetate kinase gene had only a minimal effect on the growth rate of the strain, the butyrate kinase disruption severely retarded its growth rate. This was likely due to the elimination of the pathway that allows the oxidation of NADH in the cell. Acetate production produces an excess of NADH in the cell, while butyrate production oxidizes the NADH produced during glycolysis. This reduction in growth rate is more dramatic when considering that acetate production produces more ATP than butyrate per mol glucose, but the reduction of acetate formation is much more easily adapted to than butyrate depletion. Both the acetate kinase and butyrate kinase mutant strains were transformed with the *aad* overexpression plasmid used earlier, but only the acetate kinase mutant could be complemented. Likely the reduced growth rate of the butyrate kinase mutant also results in reduced transformation efficiency. The acetate kinase mutant strain with *aad* overexpression did produce

less acetate than the strain with just the *aad* overexpression, but butanol production was also reduced. This same occurrence of decreased butanol with decreased acetate production had also been observed in the strain overexpressing both thiolase and *aad*. Using metabolic flux analysis it was shown that the intracellular production rates of both acetate and butyrate correlated very closely regardless of the genetic background indicating a strong link between the two. This can be explained by once again looking at the oxidation of NADH within the cell. Acetate production results in excess NADH, while butanol production oxidized more NADH than is produced during glycolysis, however, the co-production would result in a balanced redox state within the cell. This indicates a similar electron balance as is observed in the butyrate kinase mutant whose growth rate is reduced when the NADH balanced pathway is disrupted. As our metabolic engineering efforts aimed to decrease acetate production and were successful, the side effect was reduced butanol formation.

Despite the efforts to increase expression and bypass the endogenous solvent gene regulation, relatively little is known regarding the entry into solvent formation and sporulation. In an effort to better understand the genetic regulation of these cellular processes, we investigated the expression profiles of response regulators and histidine kinases that allow the cell to adapt to changing conditions. A common mechanism for signal recognition and response is through a two-component system, which involves the phosphate transfer from a histidine kinase to a response regulator, which then alters gene expression. Using an antisense RNA (asRNA) construct, we have investigated the downregulation of the histidine kinase coded by CAC0654 on solventogenesis, the characteristic stationary-phase phenomenon of *C. acetobutylicum*, and key signaling pathways of sporulation. Our findings suggest the two-component system coded by CAC0654 and CAC0653 (a kinase and its cognate response

regulator) is a putative negative regulator of sporulation. pH-controlled fermentations of strain 824(pHK654) carrying the asRNA construct targeting CAC0654 show reduced glucose uptake and an 80% reduction in butanol production compared to the plasmid control strain. Transcriptional profiling of 824(pHK654) using microarrays shows profound upregulation of the sporulation cascade downstream of Spo0A, the master regulator of sporulation, through the expression of *sigK*. These results were confirmed using quantitative real-time PCR. This suggests the two-component system acts as a repressor of sporulation. BLASTP searches of this two-component system exhibit highest homology to unknown and uncharacterized two-component systems indicating that this is a novel two-component system that may act affect the phosphorylation state of Spo0A repressing *sigF*, *spoIIE*, and *sigE*.

6.2 Recommendations

Much of the work here has focused on the metabolic engineering of the primary metabolism and solvent formation pathways. Analyses of these strains have relied heavily on metabolic flux analysis to identify bottlenecks within the cell and to suggest new targets for genetic modification. Additional work focused on the identification of transcriptional regulators to gain further insight into the induction of solvent formation and sporulation. DNA microarrays were used to investigate the global changes in the transcriptional pattern between strains, and to identify likely mechanisms for gene targets. Identifying the common regulators of both these pathways and where the shared components diverge will greatly aid the development of future strains.

Along this same route our group recently performed an analysis of the full transcriptional pattern of *C. acetobutylicum* throughout its entire growth phase until fully developed spores were

formed. This provides a large data set and a wealth of information to probe for future experiments. Transcriptional data provides a wonderful starting point for experimental design and hypothesis, but the combination of multiple techniques together, a systematic "omics" approach, where transcriptional, protein, and metabolic tools are combined together should allow for more sophisticated analysis and even greater insight into how the cell performs.

As a first step, after the full transcriptional program was determined from the DNA microarrays, one can combine that data with the production characteristics and metabolic flux analysis to elucidate where the cellular metabolism is likely regulated. As genes need to be induced and expressed to produce protein, this regulation is relatively easy to measure and using recombinant techniques as used earlier, this regulation can be bypassed and changed to produce a variety of desired expression patterns for the desired gene. This is of course not the whole story, as the transcripts must be translated into proteins, and this production is another point at which additional regulation may come into play. Even once the protein is produced, if the substrates and other associated proteins and cofactors are not present, the protein will be unable to perform its desired function. This was observed with *aad* overexpression when the butyryl-CoA substrate was likely limited and AAD utilized acetyl-CoA as a substrate instead. So one can argue the transcriptional analysis as a starting point for the final function, and in the case of the product formation pathways, metabolic flux analysis can measure the final output of the system, the actual rate of formation.

Applying this analysis to the wildtype strain of *C. acetobutylicum*, one must identify the genes in the respective pathways. For the primary metabolic genes in product formation (e.g *aad* for butanol formation), these are relatively well studied and known, but for many of the glycolytic genes these genes must be identified based on genome annotation in the absence of

functional genetic information. In the current metabolic flux analysis model there are 15 genes annotated that are involved in the rGLY1 flux, the conversion of glucose to pyruvate. Interestingly, only three of these genes (CAC2613, CAC2680, and CAC0827) show a significant downregulation prior to the reduction in rGLY1 (Fig 6.1). These three genes are the only ones of the 15 whose expression drops to less than one-third of its initial expression level and maintains that low expression for the next 10 hours. Overexpression of any or a combination of the three genes could determine if increased expression could increase the overall metabolism of the cell.

Continuing in the primary pathway of the metabolism, pyruvate is converted to acetyl-CoA by a pyruvate-ferredoxin oxidoreductase. This gene has not been studied, but two genes are annotated as such. Looking at the transcriptional profiles only one (CAC2229) is expressed at all timepoints during the first twenty hours (Fig 6.2). CAC2499 meets the cutoff for expression in just a few of the early timepoints, suggesting that this gene is lowly expressed, if expressed at all. If the pattern of CAC2229 expression is evaluated, it is very stably expressed showing no significant induction or repression during the time of high rGLY2 flux. This would indicate that expression of CAC2229 does not contribute to the flux pathway. A similar pattern of flux profiles and transcriptional profiles is observed for the pathway converting acetyl-CoA to butyryl-CoA. For thiolase, two genes are annotated for this function, but the chromosomal copy CAC2873 has been shown to be the primary thiolase, expressed at significantly higher levels than the copy from the megaplasmid (Winzer, Lorenz et al. 2000). The five genes associated with the conversion of acetoacetyl-CoA to butyryl-CoA all show a very similar expression pattern. As these genes are predicted to be part of a single operon, this shared expression pattern is expected, but does not seem to be linked to the decrease in the associated flux profile.

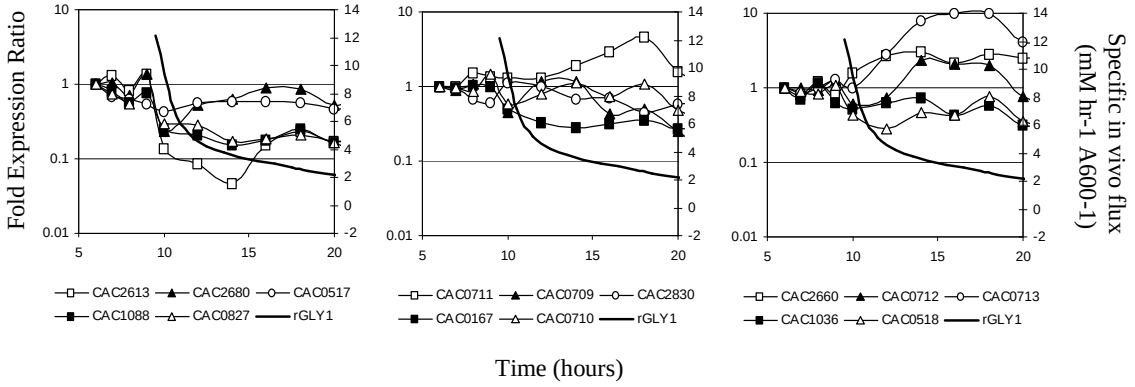


Figure 6.1 Transcriptional profiles of the genes involved in the conversion of glucose to pyruvate and the associated *in vivo* flux

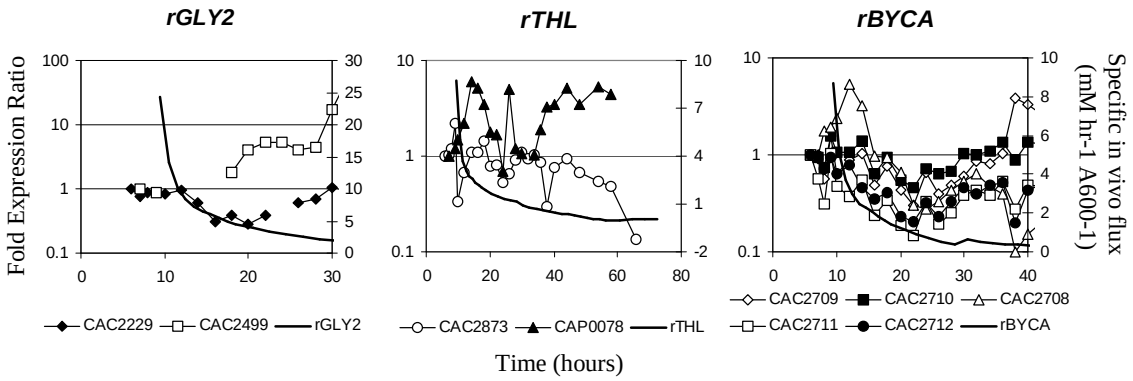


Figure 6.2 Transcriptional profiles of the genes involved in the conversion of pyruvate to butyryl-CoA and the associated *in vivo* flux

For most of the primary metabolic genes involved in the production of the end product acids and solvents, they have been functionally analyzed and studied. For the case of ethanol and butanol production this is partially true. The primary genes have been identified, but several other genes catalyze the same reaction and their involvement (or not) has not been fully determined. The acid formation fluxes, rPTAAK and rPTBBK, show very high values early in the fermentation and then quickly drop off to near zero values by 20 hours (Fig 6.3). The transcriptional profiles however, do not match this trend. The expression of *pta* and *ack* are continued at nearly the same level throughout the timecourse, while the expression of *ptb* and *buk* does decrease, this occurs only after the flux decreases suggesting that there is additional regulation at the protein or translational level diverting flux from these pathways. The acetone and butanol pathways show the main genes involved in their production induced at about hour 10, much earlier than the observed increase in flux. You would expect a flux controlled by the protein itself to have an increase in expression before the increase in flux, but there is a 10-20 hour lag between the gene induction and the peak in flux for each pathway, which is far greater than what a reasonable expectation would be of translation time. The ethanol formation flux exhibits a large increase just after 10 hours and then slowly decreases. For most of the enzymes annotated as part of this pathway there is little induction, but the primary gene for butanol production, and a butanol dehydrogenase are induced before the flux increase. Interestingly, the increase in expression and flux seem to be strongly correlated for the secondary alcohol/aldehyde dehydrogenase. This gene is normally strongly induced at neutral pH with excess NADH, but it may contribute to ethanol formation in the conditions here. High ethanol formation was not observed when this

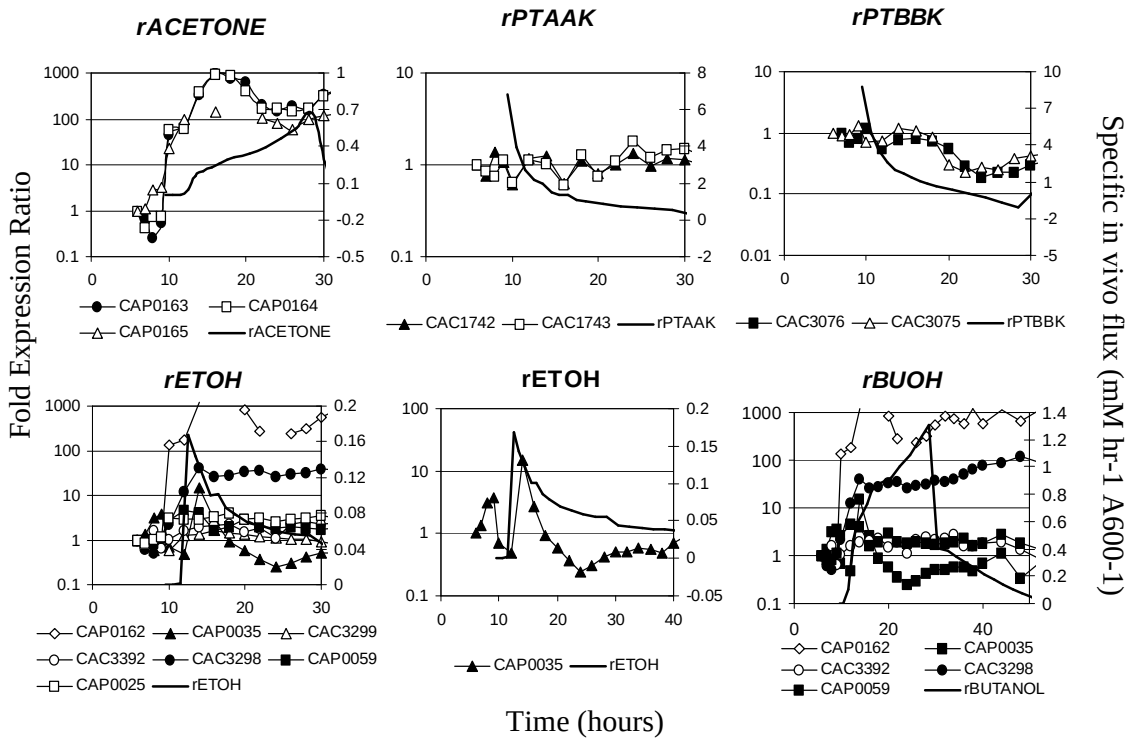


Figure 6.3 Transcriptional profiles of the genes involved in the formation of acid and solvent end products and the associated *in vivo* flux

gene was expressed in a degenerate strain without the megaplasmid (Fontaine, Meynial-Salles et al. 2002), but that study used different culture conditions and the impact of other genes on the megaplasmid may also alter its role within the cell.

The gene for hydrogen production has been identified and comparing its transcriptional profile along with the intracellular flux, a very close correlation is observed. The hydrogen flux begins very high and then decreased sharply to near zero flux by about 20 hours, while the transcriptional profile is expressed at its peak value initially and is decreased by 10-fold 15 hours into the fermentation (Fig 6.4). Both the hydrogenase and the NADH oxidoreductase share a similar role in the oxidation of ferredoxin. The hydrogenase utilizes hydrogen as an electron acceptor producing molecular hydrogen, while the NADH oxidoreductase uses NAD+ as an electron acceptor, producing NADH. It is this shift between hydrogen production and NADH production that is observed during the shift between acidogenesis and solventogenesis. There are five annotated NADH oxidoreductases in the *C. acetobutylicum* genome and their transcriptional profiles are shown in Figure 4 plotted against the NADH flux. It appears that only one of the five genes is significantly induced prior to the increase in NADH flux, indicating that this may be the primary NADH oxidoreductase in the cell. It is also important to note that the flux in increased in the NADH pathway just a couple of hours after gene induction. This combined with the fact the hydrogenase expression decreases at a similar time suggest that these two genes may dominate the redox state within the cell and contribute greatly to the timing and determination of the fermentation products. Efforts to inactivate the hydrogenase gene should be tested as well as the overexpression of the NADH oxidoreductase to confirm its role in NADH production.

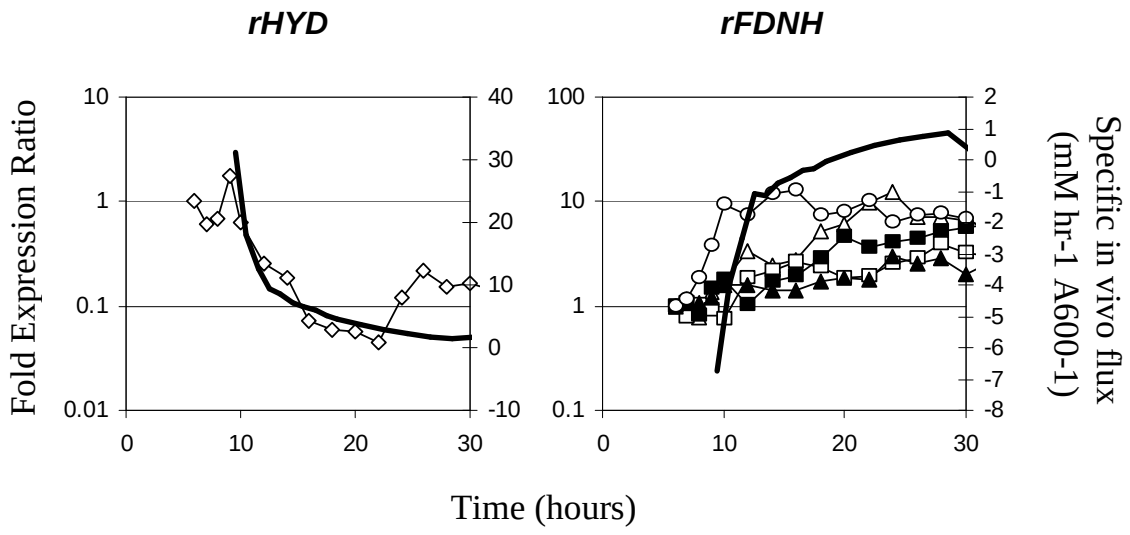


Figure 6.4 Transcriptional profiles of the genes involved in the formation of hydrogen and NADH and the associated *in vivo* flux

This gene should be overexpressed in the future utilizing a highly expressed acidogenesis associated promoter to increase NADH production and facilitate the production of increased butanol.

REFERENCES

Alizadeh, A. A., M. B. Eisen, et al. (2000). "Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling." Nature **403**(6769): 503-11.

Alsaker, K., C. J. Paredes, et al. (2005). "Design, Optimization, and Validation of Genomic DNA Microarrays for Examining the *Clostridium acetobutylicum* Transcriptome." Biotechnology and Bioprocess Engineering **10**(5): 432-443.

Alsaker, K. V. and E. T. Papoutsakis (2005). "Transcriptional program of early sporulation and stationary-phase events in Clostridium acetobutylicum." J Bacteriol **187**(20): 7103-18.

Alsaker, K. V., T. R. Spitzer, et al. (2004). "Transcriptional analysis of spo0A overexpression in Clostridium acetobutylicum and its effect on the cell's response to butanol stress." Journal of Bacteriology **186**(7): 1959-1971.

Altschul, S. F., T. L. Madden, et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search program." Nuc. Acids Res. **25**: 3389-3402.

Bahl, H., Andersch, W., Braun, K., and Gottschalk, G. (1982). "Effect of pH and butyrate concentration on the production of acetone and butanol by *Clostridium acetobutylicum* grown in continuous culture." Eur. J. Appl. Microbiol. Biotechnol. **14**: 17-20.

Bailey, J. E. (1991). "Toward a science of metabolic engineering." Science **252**: 1668-75.

Behrens, M. and P. Dürre (2000). "KdpE of Clostridium acetobutylicum is a highly specific response regulator controlling only the expression of the kdp operon." J Mol Microbiol Biotechnol **2**(1): 45-52.

Borden, J. R. and E. T. Papoutsakis (2007). "Dynamics of genomic-library enrichment and identification of solvent tolerance genes for Clostridium acetobutylicum." Appl Environ Microbiol **73**(9): 3061-8.

Bowles, L. K. and W. L. Ellefson (1985). "Effects of butanol on *Clostridium acetobutylicum*." Applied and Environmental Microbiology **50**(5): 1165-1170.

Boynton, Z. L., G. N. Bennett, et al. (1996). "Cloning, sequencing, and expression of clustered genes encoding β-hydroxybutyryl-coenzyme A (CoA) dehydrogenase, crotonase, and butyryl-CoA dehydrogenase from *Clostridium acetobutylicum* ATCC 824." J. Bacteriol. **178**(11): 3015-3024.

Britton, R. A., P. Eichenberger, et al. (2002). "Genome-wide analysis of the stationary-phase sigma factor (sigma-H) regulon of *Bacillus subtilis*." Journal of Bacteriology **184**(17): 4881-4890.

Brown, T. D., M. C. Jones-Mortimer, et al. (1977). "The enzymic interconversion of acetate and acetyl-coenzyme A in Escherichia coli." J Gen Microbiol **102**(2): 327-36.

Brüggemann, H., S. Baumer, et al. (2003). "The genome sequence of *Clostridium tetani*, the causative agent of tetanus disease." Proc. Natl Acad. Sci. USA **100**(3): 1316-1321.

Brüggemann, H. and G. Gottschalk (2004). "Insights in metabolism and toxin production from the complete genome sequence of *Clostridium tetani*." Anaerobe **10**(2): 53-68.

Buday, Z., J. C. Linden, et al. (1990). "Improved acetone butanol fermentation analysis using subambient HPLC column temperature." Enzyme and Microbial Technology **12**(1): 24-27.

Burbulys, D., K. A. Trach, et al. (1991). "Initiation of sporulation in B. subtilis is controlled by a multicomponent phosphorelay." Cell **64**(3): 545-52.

Cameron, D. C. and I. T. Tong (1993). "Cellular and metabolic engineering. An overview. [Review]." Applied Biochemistry and Biotechnology **38**(1-2): 105-140.

Chu, S., J. DeRisi, et al. (1998). "The transcriptional program of sporulation in budding yeast." Science **282**(5389): 699-705.

Clark, S. W., G. N. Bennett, et al. (1989). "Isolation and Characterization of Mutants of Clostridium acetobutylicum ATCC 824 Deficient in Acetoacetyl-Coenzyme A:Acetate/Butyrate:Coenzyme A-Transferase (EC 2.8.3.9) and in Other Solvent Pathway Enzymes." Appl Environ Microbiol **55**(4): 970-976.

Cornillot, E., R. V. Nair, et al. (1997). "The genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 reside on a large plasmid whose loss leads to degeneration of the strain." J Bacteriol **179**(17): 5442-5447.

Desai, R. P., L. K. Nielsen, et al. (1999). "Stoichiometric modeling of *Clostridium acetobutylicum* fermentations with non-linear constraints." Journal of Biotechnology **71**(1-3): 191-205.

Desai, R. P., L. N. Nielsen, et al. (1999). "Metabolic flux analysis of *C. acetobutylicum* fermentations using nonlinear constraints." Journal of Biotechnology **71**: 191-205.

Desai, R. P. and E. T. Papoutsakis (1999). "Antisense RNA strategies for metabolic engineering of Clostridium acetobutylicum." Appl Environ Microbiol **65**(3): 936-45.

Dürre, P., Ed. (2005). Handbook on Clostridia. Boca Raton, FL, Taylot & Francis Group.

Dürre, P. and C. Hollergschwandner (2004). "Initiation of endospore formation in *Clostridium acetobutylicum*." Anaerobe **10**(2): 69-74.

Edwards, J. S., Ramakrishna, R., Schilling, C., Palsson, B. O. (1999). Metabolic Engineering. New York, NY, Marcel Dekker.

Eichenberger, P., M. Fujita, et al. (2004). "The program of gene transcription for a single differentiating cell type during sporulation in *Bacillus subtilis*." PLoS Biol. **2**(10): 1664-1683.

Eichenberger, P., S. T. Jensen, et al. (2003). "The sigma(E) regulon and the identification of additional sporulation genes in *Bacillus subtilis*." J. Mol. Biol. **327**(5): 945-972.

Fabret, C., V. A. Feher, et al. (1999). "Two-component signal transduction in Bacillus subtilis: how one organism sees its world." J Bacteriol **181**(7): 1975-83.

Fontaine, L., I. Meynial-Salles, 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." J Bacteriol **184**(3): 821-830.

Gerischer, U. and P. Dürre (1990). "Cloning, sequencing, and molecular analysis of the acetoacetate decarboxylase gene region from *Clostridium acetobutylicum*." J Bacteriol **172**(12): 6907-6918.

Gonzalez-Pajuelo, M., I. Meynial-Salles, et al. (2005). "Metabolic engineering of Clostridium acetobutylicum for the industrial production of 1,3-propanediol from glycerol." Metab Eng **7**(5-6): 329-36.

Green, E. M. and G. N. Bennett (1996). "Inactivation of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824." Appl Biochem Biotechnol **57-58**: 213-221.

Green, E. M., Bennett, G. N. (1998). "Genetic manipulation of acid and solvent formation in *Clostridium acetobutylicum* ATCC 824." Biotechnology and Bioengineering **58**: 217-221.

Green, E. M., Z. L. Boynton, et al. (1996). "Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824." Microbiology **142**(Pt 8): 2079-2086.

Grossman, A. D. (1995). "Genetic networks controlling the initiation of sporulation and the development of genetic competence in Bacillus subtilis." Annu Rev Genet **29**: 477-508.

Harris, L. (2001). Cloning and characterization of the *Clostridium acetobutylicum* ATCC 824 gene encoding the Spo0A transcriptional regulator and its role in controlling solvent formation and sporulation-specific gene expression. Chemical Engineering. Evanston, IL, Northwestern University.

Harris, L. M., L. Blank, et al. (2001). "Fermentation characterization and flux analysis of recombinant strains of *Clostridium acetobutylicum* with an inactivated *solR* gene." Journal of Industrial Microbiology & Biotechnology **27**(5): 322-328.

Harris, L. M., R. P. Desai, et al. (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): 1-11.

Harris, L. M., N. E. Welker, et al. (2002). "Northern, morphological, and fermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824." Journal of Bacteriology **184**(13): 3586-3597.

Heap, J. T., O. J. Pennington, et al. (2007). "The ClosTron: a universal gene knock-out system for the genus Clostridium." J Microbiol Methods **70**(3): 452-64.

Hoch, J. A. (2000). "Two-component and phosphorelay signal transduction." Curr Opin Microbiol **3**(2): 165-70.

Huang, L., C. W. Forsberg, et al. (1986). "Influence of external pH and fermentation products on *Clostridium acetobutylicum* intracellular pH and cellular distribution of fermentation products." Applied and Environmental Microbiology **51**: 1230-1234.

Husemann, M. H. W. and E. T. Papoutsakis (1988). "Solventogenesis in Clostridium-Acetobutylicum Fermentations Related to Carboxylic-Acid and Proton Concentrations." Biotechnology and Bioengineering **32**(7): 843-852.

Ingram, L. O. (1976). "Adaptation of membrane lipids to alcohols." J. Bacteriol. **125**: 670-678.

Jain, M. K., Beacom, D and Datta, R (1993). Mutant strain of *C. acetobutylicum* and process for making butanol USA.

Jiang, M., W. Shao, et al. (2000). "Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*." Mol Microbiol **38**(3): 535-542.

Jones, D. T. and D. R. Woods (1986). "Acetone-butanol fermentation revisited." Microbiological Reviews **50**: 484-524.

Jones, S. W., C. J. Paredes, et al. (2008). "The transcriptional program underlying the physiology of clostridial sporulation." Gen Biol **Submitted**.

Lemme, C. J. and J. R. Frankiewiez (1985). Strain of Clostridium acetobutylicum and process for its preparation. USA.

Liu, X., Y. Zhu, et al. (2006). "Construction and characterization of ack deleted mutant of Clostridium tyrobutyricum for enhanced butyric acid and hydrogen production." Biotechnol Prog **22**(5): 1265-75.

Long, S., D. T. Jones, et al. (1984). "Initiation of solvent production, clostridial stage and endospore formation in *Clostridium acetobutylicum* P262." Appl. Microbiol. Biot. **20**(4): 256-261.

Mermelstein, L. D. and E. T. Papoutsakis (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." Appl Environ Microbiol **59**(4): 107710-81.

Mermelstein, L. D., N. E. Welker, et al. (1992). "Expression of cloned homologous fermentative genes in Clostridium acetobutylicum ATCC 824." Biotechnology (N Y) **10**(2): 190-5.

Mitchell, W. J. (2001). Biology and Physiology. Clostridia Biotechnology and Medical Applications. P. Durre**:** 49-104.

Mizuno, T. (1997). "Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of Escherichia coli." DNA Res **4**(2): 161-8.

Mizuno, T., T. Kaneko, et al. (1996). "Compilation of all genes encoding bacterial two-component signal transducers in the genome of the cyanobacterium, Synechocystis sp. strain PCC 6803." DNA Res **3**(6): 407-14.

Molle, V., M. Fujita, et al. (2003). "The Spo0A regulon of *Bacillus subtilis*." Molecular Microbiology **50**(5): 1683-1701.

Molle, V., Y. Nakaura, et al. (2003). "Additional targets of the *Bacillus subtilis* global regulator CodY identified by chromatin immunoprecipitation and genome-wide transcript analysis." Journal of Bacteriology **185**(6): 1911-1922.

Monot, F., Engasser, J., and Petitdemange, H. (1984). "Influence of pH and undissociated butyric acid on the production of acetone and butanol in batch cultures of *Clostridium acetobutylicum*." Appl Microbiol Biotechnol **19**: 422-426.

Nair, R. V., G. N. Bennett, et al. (1994). "Molecular Characterization of an Aldehyde/Alcohol Dehydrogenase Gene from Clostridium-Acetobutylicum Atcc-824." Journal of Bacteriology **176**(3): 871-885.

Nair, R. V. and E. T. Papoutsakis (1994). "Expression of plasmid-encoded aad in *Clostridium acetobutylicum* M5 restores vigorous butanol production." J Bacteriol **176**(18): 5843-5846.

Nölling, J., G. Breton, et al. (2001). "Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*." Journal of Bacteriology **183**(6): 4823-4838.

Ounine, K., H. Petitdemange, et al. (1985). "Regulation and butanol inhibition of D-xylose and D-glucose uptake in *Clostridium acetobutylicum*." Appl. Environ. Microbiol. **49**: 874-878.

Papoutsakis, E. T. (1984). "Equations and calculations for fermentations of butyric acid bacteria." Biotechnology and Bioengineering **26**: 174-187.

Paredes, C. J., K. Alsaker, et al. (2005). "A comparative genomic view of clostridial sporulation and physiology." Nature Reviews Microbiology **3**(12): 969-978.

Paredes, C. J., I. Rigoutsos, et al. (2004). "Transcriptional Organization Of The Clostridium acetobutylicum Genome." Nuc. Acids Res. **32**(6): 1973-1981.

Perego, M. (2001). "A new family of aspartyl phosphate phosphatases targeting the sporulation transcription factor Spo0A of Bacillus subtilis." Mol Microbiol **42**(1): 133-43.

Perego, M. and J. A. Hoch (1991). "Negative regulation of Bacillus subtilis sporulation by the spo0E gene product." J Bacteriol **173**(8): 2514-20.

Perret, S., L. Casalot, et al. (2004). "Production of heterologous and chimeric scaffoldins by Clostridium acetobutylicum ATCC 824." J Bacteriol **186**(1): 253-7.

Qureshi, N. and H. P. Blaschek (2001). "Recent advances in ABE fermentation: hyper-butanol producing Clostridium beijerinckii BA101." J Ind Microbiol Biotechnol **27**(5): 287-91.

Ravagnani, A., K. C. Jennert, et al. (2000). "Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia." Molecular Microbiology **37**(5): 1172-1185.

Ricci, J. C. D. and M. E. Hernandez (2000). "Plasmid effects on *Escherichia coli* metabolism." Critical Reviews in Biotechnology **20**(2): 79-108.

Rodrigue, A., Y. Quentin, et al. (2000). "Two-component systems in Pseudomonas aeruginosa: why so many?" Trends Microbiol **8**(11): 498-504.

Roos, J. W., J. K. McLaughlin, et al. (1985). "The Effect of Ph on Nitrogen Supply, Cell-Lysis, and Solvent Production in Fermentations of Clostridium-Acetobutylicum." Biotechnology and Bioengineering **27**(5): 681-694.

Rose, I. A. (1955). "Acetate Kinase of Bacteria." Methods in Enzymology **1**: 591-595.

Saeed, A. I., V. Sharov, et al. (2003). "TM4: a free, open-source system for microarray data management and analysis." Biotechniques **34**(2): 374-8.

Schena, M., D. Shalon, et al. (1995). "Quantitative monitoring of gene expression patterns with a complementary DNA microarray." Science **270**: 467-470.

Scotcher, M. C., F. B. Rudolph, et al. (2005). "Expression of abrB310 and sinR, and effects of decreased abrB310 expression on the transition from acidogenesis to solventogenesis, in Clostridium acetobutylicum ATCC 824." Applied and Environmental Microbiology **71**(4): 1987-1995.

Shao, L., S. Hu, et al. (2007). "Targeted gene disruption by use of a group II intron (targetron) vector in Clostridium acetobutylicum." Cell Res **17**(11): 963-5.

Shimizu, T., K. Ohtani, et al. (2002). "Complete genome sequence of *Clostridium perfringens*, an anaerobic flesh-eater." Proc. Natl Acad. Sci. USA **99**(2): 996-1001.

Sillers, R., M. Al-Hanai, et al. "Different combinations of higher aldehyde-alcohol dehydrogenase and/or thiolase expression with CoA transferase downregulation lead to higher alcohol titers and selectivity in *Clostridium acetobutylicum* fermentations: focus on the acetyl-CoA and butyryl-CoA flux nodes." Submitted.

Steil, L., M. Serrano, et al. (2005). "Genome-wide analysis of temporally regulated and compartment-specific gene expression in sporulating cells of *Bacillus subtilis*." Microbiol.-Sgm **151**: 399-420.

Stephanopoulos, G. (1999). "Metabolic fluxes and metabolic engineering." Metab Eng **1**(1): 1-11.

Sullivan, K. H., G. D. Hegeman, et al. (1979). "Alteration of the fatty acid composition of Escherichia coli by growth in the presence of normal alcohols." J Bacteriol **138**(1): 133-8.

Terracciano, J. S. and E. R. Kashket (1986). "Intracellular conditions required for the initiation of solvent production by *Clostridium acetobutylicum*." Applied and Environmental Microbiology **52**: 86-91.

Thormann, K., L. Feustel, et al. (2002). "Control of butanol formation in *Clostridium acetobutylicum* by transcriptional activation." J Bacteriol **184**(7): 1966-1973.

Tomas, C. A., K. V. Alsaker, et al. (2003). "DNA-array based transcriptional analysis of asporogenous, non-solventogenic *Clostridium acetobutylicum* strains SKO1 and M5." Journal of Bacteriology **185**(15): 4539-4547.

Tomas, C. A., J. A. Beamish, et al. (2003). "Transcriptional analysis of butanol stress and tolerance in *Clostridium acetobutylicum*." Journal of Bacteriology **186**(7): 2006-2018.

Tomas, C. A., N. E. Welker, et al. (2003). "Overexpression of *groESL* in *Clostridium acetobutylicum* results in increased solvent production and tolerance, prolonged metabolism, and large changes in the cell's transcriptional program." Applied and Environmental Microbiology **69**(8): 4951-4965.

Treuner-Lange, A., A. Kuhn, et al. (1997). "The kdp system of Clostridium acetobutylicum: cloning, sequencing, and transcriptional regulation in response to potassium concentration." J Bacteriol **179**(14): 4501-12.

Tummala, S. B., S. G. Junne, et al. (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**(12): 3644-3653.

Tummala, S. B., N. E. Welker, et al. (1999). "Development and characterization of a gene expression reporter system for Clostridium acetobutylicum ATCC 824." Applied and Environmental Microbiology **65**(9): 3793-3799.

Tummala, S. B., N. E. Welker, et al. (2003). "Design of antisense RNA constructs for downregulation of the acetone formation pathway of *Clostridium acetobutylicum*." Journal of Bacteriology **185**(6): 1923-1934.

Vallino, J. J. and G. Stephanopoulos (1993). "Metabolic flux distributions in *Corynebacterium glutamicum* during growth and lysine production." Biotechnol. Bioeng. **41**: 633-646.

Wang, L., Y. P. Sun, et al. (2002). "Genomic analysis of protein kinases, protein phosphatases and two-component regulatory systems of the cyanobacterium Anabaena sp. strain PCC 7120." FEMS Microbiol Lett **217**(2): 155-65.

Wang, S. T., B. Setlow, et al. (2006). "The Forespore Line of Gene Expression in Bacillus subtilis." Journal of Molecular Biology **358**(1): 16-37.

Wiesenborn, D. P., F. B. Rudolph, et al. (1989). "Coenzyme A transferase from *Clostridium acetobutylicum* ATCC 824 and its role in the uptake of acids." Appl Environ Microbiol **55**(2): 323-329.

Wiesenborn, D. P., F. B. Rudolph, et al. (1989). "Phosphotransbutyrylase from *Clostridium acetobutylicum* ATCC 824 and its role in acidogenesis." Appl Environ Microbiol **55**(2): 317-322.

Winzer, K., K. Lorenz, et al. (2000). "Differential regulation of two thiolase genes from Clostridium acetobutylicum DSM 792." J Mol Microbiol Biotechnol **2**(4): 531-41.

Wood, B. E. and L. O. Ingram (1992). "Ethanol production from cellobiose, amorphous cellulose, and crystalline cellulose by recombinant Klebsiella oxytoca containing chromosomally integrated Zymomonas mobilis genes for ethanol production and plasmids expressing thermostable cellulase genes from Clostridium thermocellum." Appl Environ Microbiol **58**(7): 2103-10.

Woods, D. R. (1995). "The genetic engineering of microbial solvent production." Trends in Biotechnology **13**: 259-264.

Yang, H., H. Haddad, et al. (2003). "A segmental nearest neighbor normalization and gene identification method gives superior results for DNA-array analysis." Proc Nat Acad Sci (USA) **100**(3): 1122-1127.

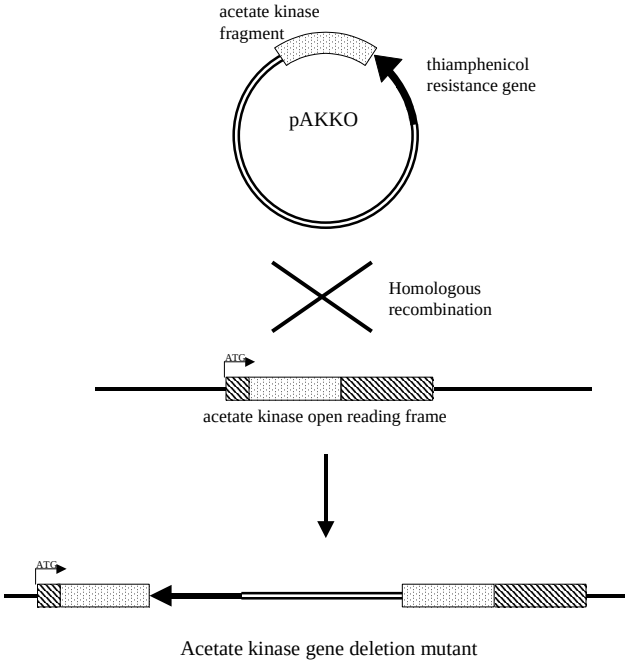
Zhao, Y. S., C. A. Tomas, et al. (2005). "Intracellular butyryl phosphate and acetyl phosphate concentrations in *Clostridium acetobutylicum* and their implications for solvent formation." Appl. Environ. Microb. **71**(1): 530-537.

Zhu, Y., X. Liu, et al. (2005). "Construction and characterization of pta gene-deleted mutant of Clostridium tyrobutyricum for enhanced butyric acid fermentation." Biotechnol Bioeng **90**(2): 154-66.

APPENDIX A: SEQUENCE DATA AND ANNOTATION OF ACETATE KINASE AND BUTYRATE KINASE M5 MUTANTS

Annotation of the plasmid pAKKO and acetate kinase disruption locus

A. Homologous recombination of non-replicating pAKKO containing a partial gene fragment with the chromosomal DNA, resulting in a disruption of the acetate kinase gene.



B. Nucleotide sequence of the acetate kinase disruption mutant. The chromosomal DNA was PCR amplified in four fragments spanning the region upstream of the acetate kinase gene to downstream of the gene and sequenced. The arrow shows the start codon of the acetate kinase gene. The sequence with the solid underlining is the genomic copy of the acetate kinase gene disrupted by the non-replicative plasmid, pAKKO. The dashed-underlined sequence is from the vector integrated into the genome.

GTATTCCCAGATCTTCAAACAGGAAACATTGGGTACAAGCTTGTTCAAAGATTTGCAAAAGCAAAAGCAATAGGACCTATATGTCAAGGATTTGCAAAACCTATTAA

TGATTTATCAAGAGGCTGTAGCTCAGAGGATATAGTAAATGTTGTTGCTATAACTGTTGTTCAGGCTCAAAGAGGTATATAAGGAGGATTTTTATGAAAAACTTAGTTA

TTAACTGCGGTAGTTCATCAATCAAATACCAGTTTATAGATATGAAGGATGAAACTGTACTCGCTAAAGGATTAGTTGAAAGAATTGGAATAAAAGGATCTGTAATAAC

CCATAAAGTAAATGGAGAAAAATATGTTACAGAAACTCCTATGGAAGATCATAAAAAGGCTATAAAGCTTGTATTAGATGCTTTATTAAATGATGAATATGGTGTTATA

AAAAATATTGATGAGATATCAGCAGTAGGACACAGAATCGTTCATGGTGGAGAAAAATATGCAAACTCAGTTTTAATAGATGAAGATGTTATGAAGTCTATAGAAGATT

GTGTGAGTCTTGCACCGCTTCATAATCCACCACACATAATAGGAATTAATGCTTGCAAGGAATTAATGCCAAACGTTCCTATGGTTGCAGTATTTGATACAGCATTCCA

TCAAACTATACCTGATTATGCATATATGTATGCTATTCCATATGAATATTATGATAAATACAAAATAAGAAAATATGGTTTTCATGGAACATCACACAAATATGTATCA

AGAACAGCTGCAGAATTTATAGGTAAAAAAGTAGAAGATTTTCACACAGGAAACAGCTATGACCATGATTACGAATTCTATGAGTCGACATTAAAAAAATAAGAGTTAC

CATTTAAGGTAACTCTTATTTTTATTACTTAAGATAATCATATATAACTTCAGCTCTAGGCAATATTATATCTGCAAGAATGTGAGAGCTAGAAACAATCTCTTTTACT

GGCAAATCATTAAGTGGCGCCAGTTACAGACAAACCTGAAGTTAACTATTTATCAATTCCTGCAATTCGTTTACAAAACGGCAAATGTGAAATCCGTCACATACTGCGT

GATGAACTTGAATTGCCAAAGGAAGTATAATTTTGTTATCTTCTTTATAATATTTCCCCATAGTAAAAATAGGAATCAAATAATCATATCCTTTCTGCAAATTCAGATT

AAAGCCATCGAAGGTTGACCACGGTATCATAGATACATTAAAAATGTTTTCCGGAGCATTTGGCTTTCCTTCCATTCTATGATTGTTTCCATACCGTTGCGTATCACTT

TCATAATCTGCTAAAAATGATTTAAAGTCAGACTTACACTCAGTCCAAAGGCTGGAAAATGTTTCAGTATCATTGTGAAATATTGTATAGCTTGGTATCATCTCATCAT

ATATCCCCAATTCACCATCTTGATTGATTGCCGTCCTAAACTCTGAATGGCGGTTTACAATCATTGCAATATAATAAAGCATTGCAGGATATAGTTTCATTCCCTTTTC

CTTTATTTGTGTGATATCCACTTTAACGGTCATGCTGTATGTACAAGGTACACTTGCAAAGTAGTGGTCAAAATACTCTTTTCTGTTCCAACTATTTTTATCAATTTTT

TCAAATACCATCTAAGTTCCCTCCCTTTTAAATTCAAGTGGATCCAATTATCTATGTTAAATGATTAACGTGCTCATTTATATTTTAACAAACTTTTCACATGAAGTTA

AGTTTTTTAGAAAATTATTTATATTTTATATTTTCTCATTATACTTTCTGCTGACTTAACTCCATCCACAGTCGACCTGCAGGAGCTTTGGCTAACACACACGCCATTC

CGGCGCGAAGGGCGAATTCGGAGCCTGCTTTTTTGTACAAAGTTGGCATTATAAAAAAGCATTGCTCATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAA

AATCATTATTTGGGGCCCGAGCTTAAGACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACATCCATGCTAGCGTTAACGCGAGAGTAGGGAACTGCCAGGCATCA

AATAAAACGAAAGGCTCAGTCGGAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTTGAACGTT

GTGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAGGCATCAAACTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAA

CTCTTCCTGGCTAGCGGTACGCGTATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGG

AGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATAC

GGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCT

CCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGC

TCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGG

TCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC

GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTA

TTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGC

AGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGACGCGTAACTCACGTTAAGGGATTTTGGTCA

TGGGTGGCTCGAGGGTTATTTGCCGACTACCTTGGTGATCTCGCCTTTCACGTAGTGGACAAATTCTTCCAACTGATCTGCGCGCGAGGCCAAGCGATCTTCTTCTTGT

CCAAGATAAGCCTGTCTAGCTTCAAGTATGACGGGCTGATACTGGGCCGGCAGGCGCTCCATTGCCCAGTCGGCAGCGACATCCTTCGGCGCGATTTTGCCGGTTACTG

CGCTGTACCAAATGCGGGACAACGTAAGCACTACATTTCGCTCATCGCCAGCCCAGTCGGGCGGCGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAGATC

CTGTTCAGGAACCGGATCAAAGAGTTCCTCCGCCGCTGGACCTACCAAGGCAACGCTATGTTCTCTTGCTTTTGTCAGCAAGATAGCCAGATCAATGTCGATCGTGGCT

GGCTCGAAGATACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGCAGTTCGCGCTTAGCTGGATAACGCCACGGAATGATGTCGTCGTGCACAACAATGGTGA

CTTCTACAGCGCGGAGAATCTCGCTCTCTCCAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAAAGCTCGCCGCGTTGTTTCATCAAGCCTTACGGTCACCGTAAC

CAGCAAATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACTGCGGAGCCGTACAAATGTACGGCCAGCAACGTCGGTTCGAGATGGCGCTCGATGACGCCAACTACC

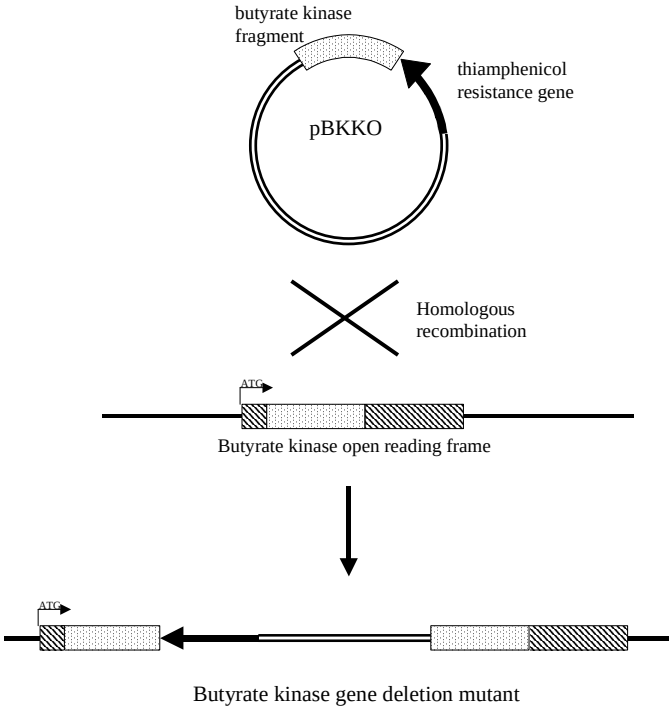
TCTGATAGTTGAGTCGATACTTCGGCGATCACCGCTTCCCTCATGATGTTTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCGTTGCTGCTCCATAACATCA

AACATCGACCCACGGCGTAACGCGCTTGCTGCTTGGATGCCCGAGGCATAGACTGTACCCCAAAAAAACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGTTACC

ACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGAGCGCATACGCTACTTGCATTACAGCTTACGAACCGAACAGGCTTATGTCCACTGGGTTCGTGCCTTCA
TCCGTTTCCACGGTGTGCGTCACCCGGCAACCTTGGGCAGCAGCGAAGTCGAGGCATTTCTGTCCTGGCTGGTCTAGAGGAGGCATGATGATATATTTTATCTTGTGC
AATGTAACATCAGAGATTTTGAGACACGGGCCAGAGCTGCCAGGAAACAGCTATGACCATGTAATACGACTCACTATAGGGGATATCAGCTGGATGGCAAATAATGATT
TTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGAGCAATTATTTTTTATAATGCCAACTTTGTACAAGAAAGCTGGGTCGAATTCGCCCTTCGCGCCAG
GACACAGAATCGTTCATGGTGGAGAAAAATATGCAAACTCAGTTTTAATAGATGAAGATGTTATGAAGTCTATAGAAGATTGTGTGAGTCTTGCACCGCTTCATAATCC
ACCACACATAATAGGAATTAATGCTTGCAAGGAATTAATGCCAAACGTTCCTATGGTTGCAGTATTTGATACAGCATTCCATCAAACTATACCTGATTATGCATATATG
TATGCTATTCCATATGAATATTATGATAAATACAAAATAAGAAAATATGGTTTTCATGGAACATCACACAAATATGTATCAAGAACAGCTGCAGAATTTATAGGTAAAA
AAGTAGAAGATTTAAAAATGGTAGTATGCCATATGGGAAATGGTGCTAGTATTACAGCTGTAGAAAATGGTAAATCAGTAGATACAAGTATGGGATTTACTCCTCTTGG
CGGTCTTGCTATGGGAACTAGAAGTGGAGATATGGACCCAGCAGTAGTAACTTTTTTAATGGATAAATTAAATATAAATGCTTCTGAAGTAAATAATCTATTAAATAAA
AAGTCAGGTATTGAAGGCTTAAGTGGAATAAGCAGCGATATGCGTGATATTAAAAAAGGAAACTATGTAGATAAAGACCCTAAAGCTATGCTAGCTTACAGTGTATTTA
ACTATAAAATAAAGCAATTTATAGGTTCATATACTGCAGTTATGAATGGATTAGACTGTTTAGTATTCACTGGTGGAATAGGTGAAAATTCATTTGAAAATAGAAGAGA
AATATGCAAAAACATGGATTATCTAGGAATAAAAATTGACGATAAGAAAAATGATGAAACTATGGGAATACCAATGGATATAAGTGCAGAAGGTTCTAAAGTTAGGGTA
CTTGTAATTCCAACTAATGAGGAGTTAATGATTGCAAGGGATACCAAAGATATAGTAGGCAAGTTAAAATAAAACTTGACATTTATTTTGCATGCTTATATAATAAATT
ATGGCTGCGTTAATGATTTAACATTAAAACAGTTTATTTTTGTGAATTTGAATAAATATAACTTATTGACTTAGGAGCTTATAATAACATGTTAGATGTTTCTGATTTA
ATAACCAAAAAAGTCGACAAAAAAGAT

Annotation of the plasmid pBKKO and butyrate kinase disruption locus

A. Homologous recombination of non-replicating pBKKO containing a partial gene fragment with the chromosomal DNA, resulting in a disruption of the butyrate kinase gene.



B. Nucleotide sequence of the butyrate kinase disruption mutant. The chromosomal DNA was PCR amplified in four fragments spanning the region upstream of the butyrate kinase gene to downstream of the gene and sequenced. The arrow shows the start codon of the butyrate kinase gene. The sequence with the solid underlining is the genomic copy of the butyrate kinase gene disrupted by the non-replicative plasmid, pBKKO. The dashed-underlined sequence is from the vector integrated into the genome.

ACATAGAAACAGGAAATGTAATGTATAAGACTTTAACATATACAACTGATTCAAAAAATGGAGGAATCTTAGTTGGAACTTCTGCACCAGTTGTTTTAACTTCAAGAGC

TGACAGCCATGAAACAAAAATGAACTCTATAGCACTTGCAGCTTTAGTTGCAGGCAATAAATAAATTAAAGTTAAGTGGAGGAATGTTAACATGTATAGATTACTAATA

ATCAATCCTGGCTCGACCTCAACTAAAATTGGTATTTATGACGATGAAAAAGAGATATTTGAGAAGACTTTAAGACATTCAGCTGAAGAGATAGAAAAATATAACACTA

TATTTGATCAATTTCAATTCAGAAAGAATGTAATTTTAGATGCGTTAAAAGAAGCAAACATAGAAGTAAGTTCTTTAAATGCTGTAGTTGGAAGAGGCGGACTCTTAAA

GCCAATAGTAAGTGGAACTTATGCAGTAAATCAAAAAATGCTTGAAGACCTTAAAGTAGGAGTTCAAGGTCAGCATGCGTCAAATCTTGGTGGAATTATTGCAAATGAA

ATAGCAAAAGAAATAAATGTTCCAGCATACATAGTTGATCCAGTTGTTGTGGATGAGCTTGATGAAGTTTCAAGAATATCAGGAATGGCTGACATTCCAAGAAAAAGTA

TATTCCATGCATCACACAGGAAACAGCTATGACCATGATTACGAATTCTATGAGTCGACATTAAAAAAATAAGAGTTACCATTTAAGGTAACTCTTATTTTTATTACTT

AAGATAATCATATATAACTTCAGCTCTAGGCAATATTATATCTGCAAGAATGTGAGAGCTAGAAACAATCTCTTTTACTGGCAAATCATTAAGTGGCGCCAGTTACAGA

CAAACCTGAAGTTAACTATTTATCAATTCCTGCAATTCGTTTACAAAACGGCAAATGTGAAATCCGTCACATACTGCGTGATGAACTTGAATTGCCAAAGGAAGTATAA

TTTTGTTATCTTCTTTATAATATTTCCCCATAGTAAAAATAGGAATCAAATAATCATATCCTTTCTGCAAATTCAGATTAAAGCCATCGAAGGTTGACCACGGTATCAT

AGATACATTAAAAATGTTTTCCGGAGCATTTGGCTTTCCTTCCATTCTATGATTGTTTCCATACCGTTGCGTATCACTTTCATAATCTGCTAAAAATGATTTAAAGTCA

GACTTACACTCAGTCCAAAGGCTGGAAAATGTTTCAGTATCATTGTGAAATATTGTATAGCTTGGTATCATCTCATCATATATCCCCAATTCACCATCTTGATTGATTG

CCGTCCTAAACTCTGAATGGCGGTTTACAATCATTGCAATATAATAAAGCATTGCAGGATATAGTTTCATTCCCTTTTCCTTTATTTGTGTGATATCCACTTTAACGGT

CATGCTGTATGTACAAGGTACACTTGCAAAGTAGTGGTCAAAATACTCTTTTCTGTTCCAACTATTTTTATCAATTTTTTCAAATACCATCTAAGTTCCCTCCCTTTTA

AATTCAAGTGGATCCAATTATCTATGTTAAATGATTAACGTGCTCATTTATATTTTAACAAACTTTTCACATGAAGTTAAGTTTTTTAGAAAATTATTTATATTTTATA

TTTTCTCATTATACTTTCTGCTGACTTAACTCCATCCACAGTCGACCTGCAGGAGCTTTGGCTAACACACACGCCATTCCAAGGGCGAATTCGACCCAGCTTTCTTGTA

CAAAGTTGGCATTATAAAAAAATAATTGCTCATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATATCCCCTATAGT

GAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCATCATGCCTCCTCTAGA

CCAGCCAGGACAGAAATGCCTCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACACCGTGGAAACGGATGAAGGCACGAACCCAGTGGACATAAGCCTGTTCGGT

TCGTAAGCTGTAATGCAAGTAGCGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTGGCGGTTTTCATGGCTTGTTATGAC

TGTTTTTTTGGGGTACAGTCTATGCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCA

GTCGCCCTAAAACAAAGTTAAACATCATGAGGGAAGCGGTGATCGCCGAAGTATCGACTCAACTATCAGAGGTAGTTGGCGTCATCGAGCGCCATCTCGAACCGACGTT

GCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGGCCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTAAGGCTTGATGAAACAACGCGGCGA

GCTTTGATCAACGACCTTTTGGAAACTTCGGCTTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTCACCATTGTTGTGCACGACGACATCATTCCGTGGCGTT

ATCCAGCTAAGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGCAGGTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTATCTTGCTGACAAA

AGCAAGAGAACATAGCGTTGCCTTGGTAGGTCCAGCGGCGGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGCTAAATGAAACCTTAACGCTATGG

AACTCGCCGCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTTACGTTGTCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGC

CGACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAGACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCAG

TTGGAAGAATTTGTCCACTACGTGAAAGGCGAGATCACCAAGGTAGTCGGCAAATAACCCTCGAGCCACCCATGACCAAAATCCCTTAACGTGAGTTACGCGTCGTTCC

ACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGT

TTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCAC

TTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGAT

AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCATTG

AGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTAT

CTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTAC

GGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCG

CAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCAC

GACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATACGCGTACCGCTAGCCAGGAAGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGGATGGC

CTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTGCCCGCCACCCTCCGGGCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTC

AGGAGAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCAGTTCCCTACTCTCGCGTTAACGCTA

GCATGGATGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTCTTAAGCTCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAA

TTGATGAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCGAATTCGCCCTTCGCGCCATCAATCCTGGCTCGACCTCAACTAAAATTGGTATTTA

TGACGATGAAAAAGAGATATTTGAGAAGACTTTAAGACATTCAGCTGAAGAGATAGAAAAATATAACACTATATTTGATCAATTTCAATTCAGAAAGAATGTAATTTTA

GATGCGTTAAAAGAAGCAAACATAGAAGTAAGTTCTTTAAATGCTGTAGTTGGAAGAGGCGGACTCTTAAAGCCAATAGTAAGTGGAACTTATGCAGTAAATCAAAAAA

TGCTTGAAGACCTTAAAGTAGGAGTTCAAGGTCAGCATGCGTCAAATCTTGGTGGAATTATTGCAAATGAAATAGCAAAAGAAATAAATGTTCCAGCATACATAGTTGA

TCCAGTTGTTGTGGATGAGCTTGATGAAGTTTCAAGAATATCAGGAATGGCTGACATTCCAAGAAAAAGTATATTCCATGCATTAAATCAAAAAGCAGTTGCTAGAAGA

TATGCAAAAGAAGTTGGAAAAAAATACGAAGATCTTAATTTAATCGTAGTCCACATGGGTGGAGGTACTTCAGTAGGTACTCATAAAGATGGTAGAGTAATAGAAGTTA

ATAATACACTTGATGGAGAAGGTCCATTCTCACCAGAAAGAAGTGGTGGAGTTCCAATAGGAGATCTTGTAAGATTGTGCTTCAGCAACAAATATACTTATGAAGAAGT

AATGAAAAAGATAAACGGCAAAGGCGGAGTTGTTAGTTACTTAAATACTATCGATTTTAAGGCTGTAGTTGATAAAGCTCTTGAAGGAGATAAGAAATGTGCACTTATA

TATGAAGCTTTCACATTCCAGGTAGCAAAAGAGATAGGAAAATGTTCAACCGTTTTAAAAGGAAATGTAGATGCAATAATCTTAACAGGCGGAATTGCGTACAACGAGC

ATGTATGTAATGCCATAGAGGATAGAGTAAAATTCATAGCACCTGTAGTTAGATATGGTGGAGAAGATGAACTTCTTGCACTTGCAGAAGGTGGACTTAGAGTTTTAAG

AGGAGAAGAAAAAGCTAAGGAATACAAATAATAAAGTCATAAATAATATAATATAACCAGTACCCATGTTTATAAAACTTTTGCCCTATAAACATGGGTATTGTTTTTT

TTTTATTTTTTTCTGATAAATTTCTTTAATATAACATAATAAACAAAATAACTAATCCACATGCCAATATTTGCACCAAGAAAATCCACTAAAACATCTGAAACAAGCG

ATGTTCTTCCTGGAACATATAATTG

APPENDIX B : LIST OF BEST-BEST MATCHES BETWEEN *C. ACETOBUTYLICUM* AND *B. SUBTILIS*

The list that follows shows the best-best matches (Chapter 5) between *C. acetobutylicum* (C. aceto) and *B. subtilis*. Gene annotation numbers are listed and the name listed is from the *B. subtilis* annotation.

C. aceto	*B. subtilis*	name	*C. aceto*	*B. subtilis*	name	*C. aceto*	*B. subtilis*	name
CAC0001	BSU00010	dnaA	CAC1425	BSU20020	yosS	CAC2711	BSU37170	acdA
CAC0002	BSU00020	dnaN	CAC1427	BSU03900	gabT	CAC2712	BSU18220	yngF
CAC0003	BSU00030	yaaA	CAC1439	BSU04850	ydcP	CAC2713	BSU05970	ydiH
CAC0004	BSU00040	recF	CAC1465	BSU21700	ypoP	CAC2714	BSU05950	ydiF
CAC0006	BSU00060	gyrB	CAC1467	BSU04250	lrpC	CAC2716	BSU19420	yojK
CAC0007	BSU00070	gyrA	CAC1472	BSU05620	ydgF	CAC2723	BSU29980	ytjP
CAC0009	BSU27890	yrxA	CAC1473	BSU33820	opuCB	CAC2725	BSU23640	yqkD
CAC0021	BSU00130	serS	CAC1479	BSU02390	ybgE	CAC2726	BSU17880	ynzC
CAC0023	BSU36160	ywqM	CAC1483	BSU05290	ydeP	CAC2737	BSU10640	sbcD
CAC0031	BSU02290	psd	CAC1486	BSU14200	ykuS	CAC2747	BSU00160	yaaH
CAC0073	BSU40650	yybG	CAC1487	BSU18030	tlp	CAC2748	BSU39110	deaD
CAC0075	BSU10600	yhjQ	CAC1495	BSU37220	ywjB	CAC2755	BSU09710	yheI
CAC0086	BSU13010	ykgB	CAC1497	BSU03190	ycgL	CAC2756	BSU09720	yheH
CAC0087	BSU04540	ydbO	CAC1512	BSU38060	ywcJ	CAC2758	BSU23370	ypuA
CAC0088	BSU21690	msrA	CAC1516	BSU34720	yvcP	CAC2769	BSU23360	ppiB
CAC0089	BSU34680	yvcT	CAC1517	BSU34710	yvcQ	CAC2772	BSU29990	ytiP
CAC0091	BSU28290	ilvC	CAC1518	BSU34700	yvcR	CAC2786	BSU24810	yqgV
CAC0092	BSU29150	ytwI	CAC1519	BSU34690	yvcS	CAC2788	BSU13860	ykvY
CAC0093	BSU37650	ywfK	CAC1521	BSU05540	ydfS	CAC2804	BSU07620	yflN
CAC0097	BSU28150	hemC	CAC1523	BSU05860	ydhR	CAC2807	BSU39070	bglS
CAC0098	BSU03280	nasF	CAC1538	BSU03550	ycxC	CAC2825	BSU25230	yqxD
CAC0099	BSU28120	hemL	CAC1544	BSU25300	cdd	CAC2830	BSU07640	yflL
CAC0100	BSU28130	hemB	CAC1545	BSU39420	dra	CAC2832	BSU09570	yhdR
CAC0103	BSU15600	cysC	CAC1546	BSU39400	pdp	CAC2834	BSU40040	yxaA
CAC0107	BSU08830	ssuB	CAC1550	BSU21680	yppQ	CAC2838	BSU05910	ydiB
CAC0112	BSU27430	glnQ	CAC1556	BSU34020	yvbX	CAC2839	BSU05920	ydiC
CAC0113	BSU36300	glcR	CAC1569	BSU13150	ykmA	CAC2840	BSU05930	ydiD
CAC0123	BSU00180	yaaJ	CAC1570	BSU21900	bsaA	CAC2841	BSU23050	ypaA
CAC0125	BSU00190	dnaX	CAC1572	BSU40190	fbp	CAC2842	BSU04780	ydcI
CAC0126	BSU00200	yaaK	CAC1573	BSU12030	yjdF	CAC2845	BSU35290	prfB
CAC0127	BSU00210	recR	CAC1576	BSU04190	ydaD	CAC2846	BSU35300	secA
CAC0129	BSU00230	bofA	CAC1584	BSU23840	yqjK	CAC2847	BSU35310	yvyD
CAC0130	BSU29070	ytaF	CAC1585	BSU39100	yxiO	CAC2850	BSU33830	opuCA
CAC0147	BSU09080	yhcH	CAC1586	BSU09610	yhdV	CAC2854	BSU27480	yrrC
CAC0154	BSU03980	mtlA	CAC1587	BSU09600	yhdU	CAC2856	BSU30550	metK
CAC0157	BSU03990	mtlD	CAC1590	BSU07570	yflS	CAC2859	BSU36420	spoIIID
CAC0158	BSU01780	glmS	CAC1592	BSU03250	ycgR	CAC2861	BSU36750	spoIID
CAC0161	BSU40480	yycB	CAC1593	BSU02850	ycdH	CAC2862	BSU36760	murAA

CAC0166	BSU14110	ykuK	CAC1596	BSU29880	malS	CAC2864	BSU36800	atpC
CAC0169	BSU08310	yfiL	CAC1604	BSU22480	mgsA	CAC2865	BSU36810	atpD
CAC0174	BSU35490	degU	CAC1610	BSU26690	brnQ	CAC2866	BSU36820	atpG
CAC0175	BSU01700	ybbI	CAC1613	BSU37230	ywjA	CAC2867	BSU36830	atpA
CAC0178	BSU11400	appC	CAC1635	BSU11030	yitL	CAC2869	BSU36850	atpF
CAC0179	BSU11360	appD	CAC1657	BSU10890	yisX	CAC2873	BSU24170	mmgA
CAC0181	BSU01670	ybbE	CAC1664	BSU30940	glgP	CAC2874	BSU35660	yvyH
CAC0182	BSU01660	ybbD	CAC1669	BSU28710	cstA	CAC2875	BSU35530	tagO
CAC0185	BSU01650	ybbC	CAC1670	BSU28920	lytT	CAC2877	BSU02860	ycdI
CAC0187	BSU35020	nagB	CAC1671	BSU27530	yrvN	CAC2878	BSU02870	yceA
CAC0188	BSU35010	nagA	CAC1673	BSU18450	gltA	CAC2879	BSU36890	upp
CAC0189	BSU35030	yvoA	CAC1674	BSU18440	gltB	CAC2880	BSU36920	ywlF
CAC0191	BSU01690	ybbH	CAC1675	BSU27520	yrzC	CAC2881	BSU36930	ywlE
CAC0192	BSU12980	ykfB	CAC1678	BSU27410	alaS	CAC2882	BSU36950	ywlC
CAC0193	BSU40490	yycA	CAC1679	BSU27400	yrzL	CAC2884	BSU37010	prfA
CAC0197	BSU04770	ydcH	CAC1680	BSU27390	yrrK	CAC2885	BSU37000	ywkE
CAC0210	BSU30200	bioB	CAC1682	BSU23520	fur	CAC2886	BSU24490	yqhQ
CAC0211	BSU32030	yuiG	CAC1683	BSU14530	ykqC	CAC2887	BSU37060	tdk
CAC0212	BSU22440	birA	CAC1684	BSU14770	ylaG	CAC2888	BSU37070	rpmE
CAC0215	BSU28820	ysdC	CAC1685	BSU27370	yrrL	CAC2889	BSU37080	rho
CAC0217	BSU27900	pheA	CAC1686	BSU27360	yrrM	CAC2892	BSU37150	pyrG
CAC0222	BSU40880	exoA	CAC1687	BSU27340	yrrO	CAC2895	BSU04560	ddl
CAC0223	BSU22170	ypsC	CAC1692	BSU15280	ftsA	CAC2898	BSU36970	spoIIR
CAC0225	BSU30390	ytsB	CAC1693	BSU15290	ftsZ	CAC2902	BSU00460	ispE
CAC0231	BSU31210	yulB	CAC1695	BSU15320	sigE	CAC2904	BSU00440	veg
CAC0232	BSU14390	fruK	CAC1696	BSU15330	sigG	CAC2905	BSU00430	yabG
CAC0234	BSU12010	manP	CAC1697	BSU15360	ylmC	CAC2914	BSU22430	panB
CAC0240	BSU08300	yfiK	CAC1698	BSU29000	ytcG	CAC2915	BSU22420	panC
CAC0249	BSU37020	ywkD	CAC1700	BSU29110	phoP	CAC2916	BSU22410	panD
CAC0252	BSU33370	yvgK	CAC1701	BSU23110	resE	CAC2918	BSU05870	ydhS
CAC0265	BSU09060	yhcF	CAC1705	BSU24990	pstS	CAC2922	BSU11690	thiG
CAC0266	BSU09070	yhcG	CAC1706	BSU24980	pstC	CAC2924	BSU11680	thiS
CAC0267	BSU03050	ldh	CAC1707	BSU24970	pstA	CAC2926	BSU00770	sul
CAC0271	BSU18040	yneP	CAC1708	BSU24950	pstBB	CAC2928	BSU30990	yuaJ
CAC0274	BSU23570	ansB	CAC1711	BSU22840	yphC	CAC2934	BSU03470	hxlR
CAC0278	BSU03790	yclM	CAC1712	BSU22830	gpsA	CAC2942	BSU30670	luxS
CAC0281	BSU33380	yvgL	CAC1713	BSU22800	spoIVA	CAC2947	BSU04260	topB
CAC0284	BSU09650	yhdZ	CAC1714	BSU02690	yccC	CAC2948	BSU14430	ykpA
CAC0285	BSU23870	yqjH	CAC1715	BSU13580	ykrV	CAC2959	BSU38200	galK
CAC0289	BSU33220	yvrH	CAC1716	BSU15660	yloC	CAC2960	BSU38860	galE
CAC0293	BSU12970	ykfA	CAC1717	BSU15670	ylzA	CAC2961	BSU38190	galT
CAC0296	BSU00240	csfB	CAC1718	BSU15680	gmk	CAC2965	BSU38570	licA
CAC0299	BSU00290	yaaQ	CAC1720	BSU15700	yloI	CAC2967	BSU36000	alsD
CAC0301	BSU00320	yaaT	CAC1721	BSU15710	priA	CAC2969	BSU38620	yxlJ
CAC0305	BSU29720	ytxE	CAC1722	BSU15720	def	CAC2970	BSU31440	patB
CAC0306	BSU00340	yabB	CAC1723	BSU15730	fmt	CAC2971	BSU29580	ytbJ
CAC0307	BSU00360	yabC	CAC1724	BSU31310	yugP	CAC2972	BSU29590	nifZ
CAC0310	BSU00370	abrB	CAC1725	BSU15740	yloM	CAC2973	BSU22100	kdgA
CAC0316	BSU11250	argF	CAC1726	BSU15750	yloN	CAC2986	BSU00420	ksgA

CAC0319	BSU33270	yvrO	CAC1727	BSU15760	prpC	CAC2987	BSU00410	rnmV
CAC0320	BSU14370	yknZ	CAC1728	BSU15770	prkC	CAC2988	BSU00400	yabE
CAC0327	BSU08720	ygaF	CAC1729	BSU15780	yloQ	CAC2989	BSU00390	yabD
CAC0330	BSU34990	lgt	CAC1730	BSU15790	rpe	CAC2991	BSU00380	metS
CAC0332	BSU05880	ydhT	CAC1731	BSU15800	yloS	CAC2994	BSU02940	yceH
CAC0359	BSU30120	yteR	CAC1733	BSU15820	rpmB	CAC3003	BSU17680	thyA
CAC0360	BSU10840	degA	CAC1734	BSU15830	yloU	CAC3004	BSU21810	dfrA
CAC0361	BSU22140	kduD	CAC1735	BSU15840	yloV	CAC3009	BSU07980	yfjS
CAC0363	BSU36170	ywqL	CAC1736	BSU15870	recG	CAC3010	BSU04580	ydbR
CAC0367	BSU37770	rocB	CAC1737	BSU15010	ylbH	CAC3012	BSU07370	yfmR
CAC0370	BSU06710	yerP	CAC1738	BSU15020	ylbI	CAC3014	BSU08790	thiC
CAC0375	BSU14000	patA	CAC1741	BSU15060	ylbM	CAC3021	BSU10340	yhfR
CAC0379	BSU05700	ydhC	CAC1742	BSU37660	pta	CAC3035	BSU36290	ywpJ
CAC0382	BSU27080	levR	CAC1743	BSU29470	ackA	CAC3037	BSU29740	ccpA
CAC0383	BSU05820	ydhN	CAC1745	BSU15080	rpmF	CAC3040	BSU36250	ywqD
CAC0384	BSU38590	licB	CAC1746	BSU15890	plsX	CAC3046	BSU35840	ywtF
CAC0386	BSU38580	licC	CAC1747	BSU15920	acpA	CAC3062	BSU36240	ywqE
CAC0391	BSU11880	yjcJ	CAC1748	BSU15930	rnc	CAC3073	BSU34250	yvfC
CAC0393	BSU22120	kdgR	CAC1751	BSU15940	smc	CAC3075	BSU24070	buk
CAC0395	BSU22110	kdgK	CAC1752	BSU15950	ftsY	CAC3076	BSU24090	ptb
CAC0396	BSU03460	hxlA	CAC1754	BSU15980	ffh	CAC3079	BSU01750	ybbP
CAC0397	BSU03450	hxlB	CAC1755	BSU15990	rpsP	CAC3083	BSU28500	trxA
CAC0422	BSU39080	licT	CAC1756	BSU16000	ylqC	CAC3087	BSU13910	ptsI
CAC0423	BSU38050	sacP	CAC1757	BSU16020	rimM	CAC3088	BSU24100	bkdR
CAC0424	BSU06170	ydjE	CAC1758	BSU16030	trmD	CAC3092	BSU01530	cwlD
CAC0425	BSU38040	sacA	CAC1759	BSU16040	rplS	CAC3093	BSU12840	pit
CAC0426	BSU05170	ydeE	CAC1761	BSU16050	ylqF	CAC3094	BSU12850	ykaA
CAC0428	BSU28730	araQ	CAC1762	BSU16060	rnhB	CAC3095	BSU11710	yjbV
CAC0430	BSU24180	yqiK	CAC1770	BSU13450	sigI	CAC3096	BSU38300	thiM
CAC0434	BSU00910	yacN	CAC1772	BSU34450	sacB	CAC3097	BSU01500	rpsI
CAC0436	BSU38800	yxkH	CAC1773	BSU34460	yveB	CAC3098	BSU01490	rplM
CAC0459	BSU40350	rocR	CAC1783	BSU10220	gltT	CAC3099	BSU01480	truA
CAC0463	BSU12900	htrA	CAC1784	BSU16110	smf	CAC3100	BSU01470	ybaF
CAC0467	BSU04230	ydaH	CAC1785	BSU16120	topA	CAC3101	BSU01460	ybaE
CAC0469	BSU23180	spmA	CAC1786	BSU16170	codY	CAC3102	BSU01450	ybxA
CAC0470	BSU23170	spmB	CAC1787	BSU16490	rpsB	CAC3103	BSU01440	rplQ
CAC0476	BSU38920	pepT	CAC1788	BSU16500	tsf	CAC3104	BSU01430	rpoA
CAC0484	BSU01770	ybbT	CAC1789	BSU16510	pyrH	CAC3105	BSU29660	rpsD
CAC0485	BSU07900	yfkH	CAC1790	BSU16520	frr	CAC3106	BSU01420	rpsK
CAC0489	BSU04620	acpS	CAC1791	BSU16530	uppS	CAC3107	BSU01410	rpsM
CAC0492	BSU17640	yncD	CAC1795	BSU16550	dxr	CAC3108	BSU01400	rpmJ
CAC0494	BSU04660	ydcE	CAC1796	BSU16560	yluC	CAC3109	BSU01390	infA
CAC0495	BSU38290	thiE	CAC1797	BSU25070	yqfY	CAC3112	BSU01370	adk
CAC0497	BSU35260	ftsE	CAC1798	BSU16590	ylxS	CAC3113	BSU01360	secY
CAC0498	BSU35250	ftsX	CAC1799	BSU16600	nusA	CAC3114	BSU01350	rplO
CAC0499	BSU19590	ctpA	CAC1800	BSU16610	ylxR	CAC3115	BSU01340	rpmD
CAC0502	BSU35170	uvrB	CAC1802	BSU16630	infB	CAC3116	BSU01330	rpsE
CAC0503	BSU35160	uvrA	CAC1803	BSU16650	rbfA	CAC3117	BSU01320	rplR
CAC0508	BSU28490	uvrC	CAC1805	BSU16660	truB	CAC3118	BSU01310	rplF

CAC0510	BSU15230	murB	CAC1806	BSU16670	ribC	CAC3119	BSU01300	rpsH
CAC0511	BSU34770	yvcJ	CAC1807	BSU16680	rpsO	CAC3120	BSU01290	rpsN
CAC0512	BSU34760	yvcK	CAC1808	BSU16690	pnpA	CAC3121	BSU01280	rplE
CAC0513	BSU34750	yvcL	CAC1810	BSU16760	dapG	CAC3122	BSU01270	rplX
CAC0514	BSU29270	ytoI	CAC1811	BSU16790	tepA	CAC3123	BSU01260	rplN
CAC0516	BSU29230	dnaE	CAC1812	BSU16800	spoIIIE	CAC3124	BSU01250	rpsQ
CAC0517	BSU29190	pfkA	CAC1814	BSU16920	pgsA	CAC3125	BSU01240	rpmC
CAC0519	BSU15500	pyrC	CAC1815	BSU16940	recA	CAC3126	BSU01230	rplP
CAC0523	BSU06730	yefA	CAC1816	BSU16960	ymdA	CAC3127	BSU01220	rpsC
CAC0531	BSU08190	yfiA	CAC1817	BSU16980	spoVS	CAC3128	BSU01210	rplV
CAC0532	BSU08200	malP	CAC1819	BSU22370	aspB	CAC3129	BSU01200	rpsS
CAC0533	BSU08180	malA	CAC1820	BSU13900	ptsH	CAC3130	BSU01190	rplB
CAC0534	BSU18830	pps	CAC1825	BSU21910	metA	CAC3131	BSU01180	rplW
CAC0536	BSU38540	dltE	CAC1832	BSU17850	lexA	CAC3132	BSU01170	rplD
CAC0557	BSU10290	yhfN	CAC1833	BSU17440	ynbB	CAC3133	BSU01160	rplC
CAC0560	BSU05310	ydeR	CAC1834	BSU17340	ymaH	CAC3134	BSU01150	rpsJ
CAC0564	BSU02550	ycbL	CAC1835	BSU17330	miaA	CAC3136	BSU01130	tufA
CAC0566	BSU12320	yjmC	CAC1836	BSU17050	mutL	CAC3138	BSU01120	fusA
CAC0570	BSU02350	gamP	CAC1837	BSU17040	mutS	CAC3139	BSU01110	rpsG
CAC0577	BSU39330	yxiA	CAC1838	BSU17010	ymcB	CAC3140	BSU01100	rpsL
CAC0579	BSU08970	prkA	CAC1843	BSU04820	ydcN	CAC3141	BSU01090	ybxF
CAC0580	BSU08980	yhbH	CAC1845	BSU13680	motB	CAC3142	BSU01080	rpoC
CAC0581	BSU09400	spoVR	CAC1846	BSU13690	motA	CAC3143	BSU01070	rpoB
CAC0590	BSU23280	ribD	CAC1848	BSU22890	cmk	CAC3145	BSU01050	rplL
CAC0591	BSU23270	ribE	CAC1849	BSU30060	ytfP	CAC3146	BSU01040	rplJ
CAC0592	BSU23260	ribA	CAC1851	BSU23160	rluB	CAC3147	BSU01030	rplA
CAC0593	BSU23250	ribH	CAC1852	BSU08000	yfjQ	CAC3148	BSU01020	rplK
CAC0594	BSU00110	yaaD	CAC1871	BSU04860	ydcQ	CAC3149	BSU01010	nusG
CAC0595	BSU00120	yaaE	CAC1947	BSU21020	yonR	CAC3150	BSU01000	secE
CAC0603	BSU33450	yvgS	CAC1958	BSU33400	yvgN	CAC3151	BSU24900	rpmGA
CAC0608	BSU23380	lysA	CAC1962	BSU01600	ybbA	CAC3152	BSU00980	sigH
CAC0611	BSU28900	ysbB	CAC1990	BSU33170	yvrB	CAC3153	BSU00970	yacP
CAC0613	BSU26990	yraD	CAC1992	BSU05960	ydiG	CAC3154	BSU00960	yacO
CAC0614	BSU26980	yraE	CAC1993	BSU36700	moaA	CAC3156	BSU00950	yazC
CAC0618	BSU30620	ytlD	CAC2059	BSU29500	ytfJ	CAC3157	BSU22630	trpA
CAC0619	BSU30610	ytlC	CAC2060	BSU23210	ypuH	CAC3158	BSU22640	trpB
CAC0626	BSU11420	trpS	CAC2061	BSU23220	ypuG	CAC3159	BSU22650	trpF
CAC0629	BSU39550	yxeH	CAC2063	BSU23480	dacF	CAC3160	BSU22660	trpC
CAC0632	BSU09240	yhcW	CAC2064	BSU23490	punA	CAC3161	BSU22670	trpD
CAC0637	BSU29670	tyrS	CAC2065	BSU23500	drm	CAC3162	BSU00750	pabA
CAC0639	BSU16080	ylqH	CAC2066	BSU23510	ripX	CAC3163	BSU22680	trpE
CAC0642	BSU11600	yjbM	CAC2068	BSU23530	spoIIM	CAC3164	BSU40330	rocE
CAC0646	BSU30320	leuS	CAC2071	BSU24220	spo0A	CAC3169	BSU28310	ilvB
CAC0665	BSU06980	yesP	CAC2072	BSU24230	spoIVB	CAC3170	BSU21870	ilvD
CAC0671	BSU27680	yrbG	CAC2073	BSU24240	recN	CAC3171	BSU28270	leuB
CAC0673	BSU15850	sdaAB	CAC2074	BSU24250	ahrC	CAC3173	BSU28260	leuC
CAC0674	BSU15860	sdaAA	CAC2076	BSU24260	yqxC	CAC3176	BSU28300	ilvH
CAC0676	BSU02270	pssA	CAC2077	BSU24270	dxs	CAC3177	BSU00940	cysS
CAC0678	BSU09230	yhcV	CAC2080	BSU24280	yqiD	CAC3178	BSU16570	proS

CAC0679	BSU06720	yerQ	CAC2082	BSU24300	yqiB	CAC3184	BSU00900	yacM
CAC0681	BSU36520	nrgB	CAC2083	BSU24310	folD	CAC3185	BSU00890	yacL
CAC0682	BSU36510	nrgA	CAC2084	BSU24320	nusB	CAC3187	BSU00880	yacK
CAC0687	BSU00930	cysE	CAC2085	BSU24330	yqhY	CAC3188	BSU00870	radA
CAC0689	BSU22340	nth	CAC2089	BSU24390	spoIIIAE	CAC3189	BSU00860	clpC
CAC0690	BSU10200	yhfE	CAC2090	BSU24400	spoIIIAD	CAC3190	BSU00850	mcsB
CAC0692	BSU12300	uxaC	CAC2091	BSU24410	spoIIIAC	CAC3191	BSU00840	mcsA
CAC0693	BSU12370	exuR	CAC2093	BSU24430	spoIIIAA	CAC3192	BSU00830	ctsR
CAC0694	BSU12310	yjmB	CAC2094	BSU24450	efp	CAC3194	BSU15200	murD
CAC0695	BSU12380	uxaB	CAC2095	BSU24460	yqhT	CAC3197	BSU00820	lysS
CAC0696	BSU12390	uxaA	CAC2105	BSU24730	comGA	CAC3198	BSU27320	greA
CAC0697	BSU18360	yoxA	CAC2111	BSU15640	yloA	CAC3199	BSU00810	yacF
CAC0699	BSU13930	splB	CAC2112	BSU15480	pyrP	CAC3200	BSU00700	yacB
CAC0700	BSU08930	cspR	CAC2113	BSU15470	pyrR	CAC3202	BSU00690	ftsH
CAC0702	BSU31540	yufN	CAC2114	BSU15460	ylyB	CAC3203	BSU00680	hprT
CAC0703	BSU31550	yufO	CAC2115	BSU15450	lspA	CAC3206	BSU00630	yabR
CAC0704	BSU31560	yufP	CAC2116	BSU30840	yteA	CAC3209	BSU00600	yabP
CAC0705	BSU31570	yufQ	CAC2117	BSU27270	mtn	CAC3210	BSU00590	yabO
CAC0707	BSU34200	sigL	CAC2118	BSU15420	divIVA	CAC3211	BSU22790	hbs
CAC0708	BSU33950	cggR	CAC2119	BSU15400	ylmG	CAC3212	BSU00580	yabN
CAC0709	BSU33940	gapA	CAC2121	BSU15380	ylmE	CAC3214	BSU00560	spoVT
CAC0710	BSU33930	pgk	CAC2123	BSU15270	sbp	CAC3215	BSU09950	prsA
CAC0711	BSU33920	tpiA	CAC2126	BSU15210	spoVE	CAC3216	BSU00550	mfd
CAC0712	BSU33910	pgm	CAC2127	BSU15190	mraY	CAC3217	BSU00530	spoVC
CAC0713	BSU33900	eno	CAC2128	BSU04570	murF	CAC3220	BSU23120	resD
CAC0715	BSU33610	rnr	CAC2129	BSU15180	murE	CAC3221	BSU00510	prs
CAC0716	BSU33600	smpB	CAC2130	BSU15170	spoVD	CAC3222	BSU00500	gcaD
CAC0718	BSU03860	ycnD	CAC2132	BSU15140	ylxA	CAC3223	BSU00490	spoVG
CAC0727	BSU02060	ybxG	CAC2133	BSU15130	yllB	CAC3224	BSU00470	purR
CAC0737	BSU37790	rocG	CAC2134	BSU40920	yyaF	CAC3225	BSU29790	murC
CAC0740	BSU16640	ylxP	CAC2137	BSU15650	yloB	CAC3229	BSU23820	yqjM
CAC0741	BSU07360	yfmS	CAC2138	BSU40550	ppaC	CAC3233	BSU10940	yitC
CAC0743	BSU39260	bglH	CAC2139	BSU16290	flgE	CAC3237	BSU38810	msmX
CAC0751	BSU03670	yclF	CAC2143	BSU16470	sigD	CAC3238	BSU30480	ytqA
CAC0763	BSU35080	yvmB	CAC2145	BSU16410	ylxH	CAC3250	BSU26810	yrpC
CAC0766	BSU27000	yraB	CAC2146	BSU16400	flhF	CAC3252	BSU18480	proH
CAC0780	BSU38460	tyrZ	CAC2147	BSU16390	flhA	CAC3253	BSU18470	proJ
CAC0788	BSU33300	fhuG	CAC2149	BSU16360	fliQ	CAC3254	BSU13130	proA
CAC0789	BSU33310	fhuB	CAC2150	BSU16350	fliP	CAC3260	BSU22360	asnS
CAC0790	BSU33320	fhuD	CAC2159	BSU16240	fliI	CAC3261	BSU13540	ogt
CAC0791	BSU33290	fhuC	CAC2161	BSU16220	fliG	CAC3267	BSU01730	sigW
CAC0792	BSU09670	dat	CAC2163	BSU16200	fliE	CAC3281	BSU08220	yfiC
CAC0794	BSU30870	ytcB	CAC2164	BSU16190	flgC	CAC3282	BSU08210	yfiB
CAC0805	BSU31260	mcpB	CAC2165	BSU16180	flgB	CAC3284	BSU35480	yviA
CAC0807	BSU05120	cspC	CAC2173	BSU34330	yveO	CAC3285	BSU09460	yhdG
CAC0820	BSU36220	ywqG	CAC2187	BSU37870	spsE	CAC3286	BSU02450	ycbB
CAC0827	BSU37120	fbaA	CAC2190	BSU37860	spsF	CAC3288	BSU32710	yurY
CAC0844	BSU26730	yrdF	CAC2203	BSU35360	hag	CAC3289	BSU32670	yurU
CAC0848	BSU35280	yvjA	CAC2206	BSU35330	fliS	CAC3291	BSU32690	csd

CAC0849	BSU19540	yodB	CAC2209	BSU35370	csrA	CAC3292	BSU32680	yurV
CAC0850	BSU05480	ydfN	CAC2210	BSU35380	yviF	CAC3297	BSU19620	yodJ
CAC0864	BSU08290	yfiJ	CAC2215	BSU16320	fliY	CAC3299	BSU31360	yugK
CAC0869	BSU34790	trxB	CAC2216	BSU16310	fliM	CAC3301	BSU03720	gerKB
CAC0872	BSU22060	pbuX	CAC2218	BSU16330	cheY	CAC3306	BSU29490	tpx
CAC0873	BSU22070	xpt	CAC2219	BSU16450	cheC	CAC3315	BSU39820	htpG
CAC0875	BSU05160	ydeD	CAC2220	BSU16430	cheA	CAC3316	BSU37240	ywiE
CAC0878	BSU39490	yxeN	CAC2221	BSU22720	cheR	CAC3325	BSU03610	yckK
CAC0879	BSU39480	yxeO	CAC2222	BSU16420	cheB	CAC3326	BSU03600	yckJ
CAC0880	BSU39500	yxeM	CAC2223	BSU16460	cheD	CAC3330	BSU02100	cypC
CAC0887	BSU14520	adeC	CAC2224	BSU16440	cheW	CAC3335	BSU32980	yusZ
CAC0892	BSU29750	aroA	CAC2232	BSU27620	yrvE	CAC3338	BSU05270	ydzF
CAC0894	BSU22700	aroB	CAC2233	BSU27500	trmU	CAC3339	BSU07420	yfmM
CAC0896	BSU22710	aroF	CAC2234	BSU27510	yrvO	CAC3340	BSU38480	ywaC
CAC0897	BSU25660	aroD	CAC2235	BSU00730	cysK	CAC3342	BSU13960	ykwC
CAC0898	BSU03150	aroK	CAC2237	BSU30970	glgC	CAC3345	BSU39990	yxaF
CAC0899	BSU24470	yqhS	CAC2238	BSU30960	glgD	CAC3346	BSU02670	lmrB
CAC0901	BSU05940	gcp	CAC2239	BSU30950	glgA	CAC3347	BSU02400	ybgF
CAC0907	BSU31280	yugU	CAC2241	BSU33490	yvgW	CAC3348	BSU02410	ybgG
CAC0908	BSU27540	yrvM	CAC2242	BSU19120	yozA	CAC3351	BSU02590	ycbP
CAC0909	BSU31230	tlpB	CAC2243	BSU10790	asnO	CAC3354	BSU18370	yoeA
CAC0921	BSU39470	yxeP	CAC2251	BSU12880	ykcB	CAC3360	BSU40670	yybE
CAC0930	BSU27250	yrhB	CAC2260	BSU37890	spsC	CAC3361	BSU36020	alsR
CAC0931	BSU27260	yrhA	CAC2261	BSU09930	yhaM	CAC3368	BSU08400	yfiU
CAC0935	BSU34930	hisZ	CAC2262	BSU10630	addA	CAC3370	BSU08340	padR
CAC0936	BSU34920	hisG	CAC2263	BSU10620	addB	CAC3372	BSU05530	ydfR
CAC0937	BSU34910	hisD	CAC2264	BSU36900	glyA	CAC3375	BSU26970	adhB
CAC0938	BSU34900	hisB	CAC2268	BSU33520	yvgZ	CAC3378	BSU04150	ycsN
CAC0939	BSU34890	hisH	CAC2269	BSU27550	aspS	CAC3379	BSU25090	yqfW
CAC0940	BSU34880	hisA	CAC2271	BSU09840	hemZ	CAC3398	BSU38870	yxkA
CAC0941	BSU34870	hisF	CAC2272	BSU24790	yqgX	CAC3409	BSU18460	gltC
CAC0944	BSU17890	tkt	CAC2273	BSU27590	yrvI	CAC3418	BSU37550	ywhA
CAC0947	BSU21790	yplQ	CAC2274	BSU27600	relA	CAC3421	BSU33540	yvaB
CAC0963	BSU31150	yubB	CAC2275	BSU27610	apt	CAC3429	BSU05420	ydfI
CAC0965	BSU09540	yhdO	CAC2282	BSU27710	tgt	CAC3430	BSU05410	ydfH
CAC0972	BSU29130	icd	CAC2283	BSU27720	queA	CAC3431	BSU05430	ydfJ
CAC0973	BSU29450	argG	CAC2284	BSU27730	ruvB	CAC3442	BSU16580	polC
CAC0974	BSU29440	argH	CAC2285	BSU27740	ruvA	CAC3445	BSU33890	yvbK
CAC0976	BSU31380	yuzA	CAC2295	BSU27820	yrbC	CAC3446	BSU08470	yfhB
CAC0977	BSU05060	lrpB	CAC2296	BSU35510	yvyE	CAC3447	BSU38250	ywbO
CAC0984	BSU32750	yusC	CAC2301	BSU22320	ponA	CAC3451	BSU17570	xynP
CAC0985	BSU32740	yusB	CAC2304	BSU23410	spoVAD	CAC3452	BSU17580	xynB
CAC0986	BSU32730	yusA	CAC2305	BSU23420	spoVAC	CAC3471	BSU32130	guaC
CAC0990	BSU00920	gltX	CAC2307	BSU23460	spoIIAB	CAC3474	BSU09010	yhcA
CAC0991	BSU27020	yraA	CAC2308	BSU23470	spoIIAA	CAC3475	BSU36600	mta
CAC0998	BSU32260	hom	CAC2315	BSU37820	spsK	CAC3481	BSU05320	ydeS
CAC1002	BSU31750	yueK	CAC2331	BSU37810	spsL	CAC3485	BSU40010	yxaD
CAC1006	BSU30630	ytkD	CAC2332	BSU37830	spsJ	CAC3490	BSU20010	yosT
CAC1012	BSU18170	yngA	CAC2333	BSU37840	spsI	CAC3494	BSU19100	yobV

CAC1022	BSU03520	srfAD	CAC2335	BSU35670	gtaB	CAC3498	BSU32570	yurL
CAC1023	BSU27860	nadC	CAC2337	BSU09310	yhxB	CAC3501	BSU32610	yurP
CAC1024	BSU27870	nadB	CAC2338	BSU14630	speA	CAC3502	BSU32560	yurK
CAC1025	BSU27850	nadA	CAC2339	BSU01720	ybbK	CAC3536	BSU40230	yydA
CAC1036	BSU29180	pyk	CAC2340	BSU28580	mutSB	CAC3538	BSU40370	yycJ
CAC1037	BSU09660	yheN	CAC2356	BSU28630	pheT	CAC3539	BSU37100	murAB
CAC1039	BSU11510	yjbE	CAC2357	BSU28640	pheS	CAC3544	BSU09120	yhcK
CAC1048	BSU30520	ytoA	CAC2358	BSU28650	ysgA	CAC3546	BSU40820	yyaL
CAC1054	BSU40320	rocF	CAC2359	BSU28850	rplT	CAC3551	BSU02750	natA
CAC1084	BSU03410	yckE	CAC2360	BSU28860	rpmI	CAC3568	BSU29200	accA
CAC1088	BSU37090	ywjI	CAC2361	BSU28870	infC	CAC3569	BSU29210	accD
CAC1089	BSU35000	hprK	CAC2362	BSU37560	thrZ	CAC3570	BSU24340	accC
CAC1090	BSU24890	yqgN	CAC2365	BSU13470	sspD	CAC3571	BSU36370	ywpB
CAC1098	BSU29090	polA	CAC2370	BSU00710	yacC	CAC3572	BSU24350	accB
CAC1099	BSU29060	ytaG	CAC2371	BSU25610	yqeM	CAC3573	BSU11340	fabF
CAC1218	BSU20270	yorS	CAC2372	BSU00450	sspF	CAC3574	BSU15910	fabG
CAC1231	BSU31170	yulF	CAC2375	BSU11390	appB	CAC3575	BSU15900	fabD
CAC1233	BSU14010	cheV	CAC2378	BSU16770	dapA	CAC3578	BSU11330	fabHA
CAC1234	BSU27910	pheB	CAC2381	BSU14180	ykuQ	CAC3585	BSU14360	yknY
CAC1235	BSU32240	thrB	CAC2388	BSU11220	argD	CAC3586	BSU16930	cinA
CAC1240	BSU28050	maf	CAC2389	BSU11210	argB	CAC3593	BSU40420	purA
CAC1241	BSU28040	radC	CAC2390	BSU11190	argC	CAC3601	BSU23610	nudF
CAC1242	BSU28030	mreB	CAC2391	BSU11200	argJ	CAC3605	BSU40070	gntP
CAC1248	BSU28000	minC	CAC2394	BSU03620	yclA	CAC3617	BSU10010	yhaG
CAC1249	BSU27990	minD	CAC2398	BSU28080	folC	CAC3619	BSU23970	yqiY
CAC1257	BSU27960	rplU	CAC2399	BSU28090	valS	CAC3621	BSU23040	fer
CAC1259	BSU27940	rpmA	CAC2410	BSU08520	yfhG	CAC3626	BSU22780	mtrA
CAC1260	BSU27920	obg	CAC2420	BSU18610	yoaH	CAC3627	BSU13720	ykvJ
CAC1261	BSU25650	yqeI	CAC2423	BSU14550	ykrA	CAC3628	BSU13000	ykfD
CAC1262	BSU25640	yqeJ	CAC2431	BSU07710	yflE	CAC3634	BSU11430	oppA
CAC1263	BSU25630	yqeK	CAC2432	BSU04480	ydbI	CAC3635	BSU11370	appF
CAC1265	BSU00480	yabJ	CAC2433	BSU33000	yvtA	CAC3642	BSU11460	oppD
CAC1266	BSU09210	yhcT	CAC2484	BSU23950	yqjA	CAC3643	BSU11450	oppC
CAC1267	BSU23190	dacB	CAC2500	BSU18070	yneS	CAC3644	BSU11440	oppB
CAC1274	BSU25550	rpsT	CAC2529	BSU39870	yxbD	CAC3652	BSU36010	alsS
CAC1275	BSU25540	gpr	CAC2537	BSU36530	ywoA	CAC3654	BSU33510	yvgY
CAC1278	BSU25510	lepA	CAC2539	BSU28330	ysnE	CAC3655	BSU33500	yvgX
CAC1279	BSU25500	hemN	CAC2540	BSU26650	czcD	CAC3657	BSU02470	ycbD
CAC1280	BSU25490	hrcA	CAC2542	BSU28680	ysfC	CAC3661	BSU12890	ykcC
CAC1281	BSU25480	grpE	CAC2552	BSU18740	yozG	CAC3663	BSU33010	cssR
CAC1282	BSU25470	dnaK	CAC2553	BSU18730	yoaS	CAC3666	BSU33570	yvaE
CAC1283	BSU25460	dnaJ	CAC2566	BSU30490	ytqB	CAC3670	BSU30290	amyC
CAC1284	BSU25450	yqeT	CAC2570	BSU34120	yvfO	CAC3671	BSU30280	amyD
CAC1285	BSU25440	yqeU	CAC2596	BSU10390	yhfW	CAC3673	BSU17590	xylR
CAC1286	BSU25430	yqeV	CAC2598	BSU35980	ywsA	CAC3683	BSU04140	pbpC
CAC1287	BSU10030	hit	CAC2602	BSU37500	speE	CAC3709	BSU31870	yukA
CAC1288	BSU25410	rpsU	CAC2612	BSU17610	xylB	CAC3715	BSU40440	dnaC
CAC1289	BSU25400	yqeY	CAC2613	BSU24850	glcK	CAC3717	BSU40500	rplI
CAC1290	BSU25360	yqfC	CAC2614	BSU34550	pgcM	CAC3718	BSU40510	yybT

CAC1292	BSU25330	yqfF	CAC2616	BSU24520	mntR	CAC3722	BSU40890	rpsR
CAC1293	BSU25320	yqfG	CAC2624	BSU32170	dapF	CAC3723	BSU40900	ssb
CAC1295	BSU25290	era	CAC2627	BSU29900	ytmQ	CAC3724	BSU40910	rpsF
CAC1299	BSU25210	dnaG	CAC2634	BSU08730	perR	CAC3726	BSU14210	ykuT
CAC1300	BSU25200	sigA	CAC2636	BSU28190	ysxC	CAC3727	BSU40950	yyaC
CAC1302	BSU25180	yqfN	CAC2637	BSU28200	lonA	CAC3729	BSU40960	spo0J
CAC1319	BSU09280	glpF	CAC2638	BSU28210	lonB	CAC3730	BSU40970	soj
CAC1320	BSU09270	glpP	CAC2639	BSU28220	clpX	CAC3731	BSU40990	yyaA
CAC1321	BSU09290	glpK	CAC2640	BSU34540	clpP	CAC3732	BSU41000	gidB
CAC1325	BSU25110	yqfU	CAC2641	BSU28230	tig	CAC3733	BSU41010	gidA
CAC1329	BSU03570	sfp	CAC2644	BSU15520	pyrAB	CAC3734	BSU41020	thdF
CAC1331	BSU12350	yjmF	CAC2645	BSU15510	pyrAA	CAC3735	BSU41030	jag
CAC1332	BSU12340	uxuA	CAC2646	BSU23310	sipS	CAC3736	BSU41040	spoIIIJ
CAC1337	BSU06900	cotJB	CAC2650	BSU15540	pyrD	CAC3737	BSU30680	ytjA
CAC1338	BSU06910	cotJC	CAC2651	BSU15530	pyrK	CAC3738	BSU41050	rnpA
CAC1339	BSU33960	araE	CAC2654	BSU15490	pyrB	CAP0006	BSU07820	treR
CAC1340	BSU33970	araR	CAC2655	BSU18190	yngC	CAP0020	BSU03700	gerKA
CAC1341	BSU28780	araD	CAC2659	BSU35110	yvlC	CAP0027	BSU04320	ydaO
CAC1345	BSU35830	ywtG	CAC2660	BSU14860	pycA	CAP0029	BSU31480	yuxJ
CAC1346	BSU28800	araA	CAC2664	BSU28350	ysnB	CAP0030	BSU36490	ywoC
CAC1347	BSU37110	ywjH	CAC2665	BSU28360	ysnA	CAP0036	BSU08950	yhbE
CAC1353	BSU07700	nagP	CAC2669	BSU06690	gatB	CAP0037	BSU08940	yhbD
CAC1354	BSU22230	ypqE	CAC2670	BSU06680	gatA	CAP0047	BSU10240	yhfI
CAC1355	BSU13880	glcT	CAC2671	BSU06670	gatC	CAP0050	BSU34390	pnbA
CAC1361	BSU30210	bioD	CAC2673	BSU06620	ligA	CAP0052	BSU32870	yusO
CAC1362	BSU30230	bioA	CAC2674	BSU06610	pcrA	CAP0059	BSU31050	gbsB
CAC1390	BSU06420	purE	CAC2675	BSU06580	yerC	CAP0061	BSU26900	yraL
CAC1391	BSU06450	purC	CAC2679	BSU29930	amyX	CAP0063	BSU04360	mntH
CAC1392	BSU06490	purF	CAC2680	BSU31350	pgi	CAP0067	BSU27050	levF
CAC1393	BSU06500	purM	CAC2685	BSU34570	yvdK	CAP0068	BSU27040	levG
CAC1394	BSU06510	purN	CAC2687	BSU19220	yocI	CAP0073	BSU21470	sunT
CAC1395	BSU06520	purH	CAC2688	BSU10900	yisY	CAP0090	BSU30450	ytrB
CAC1396	BSU06530	purD	CAC2692	BSU40850	maa	CAP0103	BSU02970	yceK
CAC1404	BSU14380	fruR	CAC2700	BSU06360	guaA	CAP0104	BSU25790	arsB
CAC1407	BSU39270	bglP	CAC2701	BSU00090	guaB	CAP0105	BSU25780	arsC
CAC1412	BSU02910	yceE	CAC2703	BSU06030	groEL	CAP0118	BSU18150	ynfF
CAC1414	BSU02890	yceC	CAC2704	BSU06020	groES	CAP0120	BSU18160	xynD
CAC1415	BSU02920	yceF	CAC2708	BSU24160	mmgB	CAP0131	BSU32880	yusP
CAC1422	BSU31300	yugS	CAC2709	BSU28520	etfA	CAP0163	BSU19730	yodS
CAC1423	BSU09450	yhdF	CAC2710	BSU28530	etfB	CAP0164	BSU38980	scoB
						CAP0168	BSU03040	amyE

APPENDIX C : COMPILATION OF INFERRED REGULONS FOR SPO0A AND SPORULATION RELATED SIGMA FACTORS

The following lists shows the gene name for *B. subtilis* (Bsu) and the gene accession number for *C. acetobutylicum* (Cac).

SigH regulon		Spo0A regulon		SigF regulon		SigG regulon	
Bsu gene	Cac number	Bsu gene	Cac number	Bsu gene	Cac number	Bsu gene	Cac number
yoxA	CAC0697	ykuK	CAC0166	csfB	CAC0296	exoA	CAC0222
sigA	CAC1226	ywkD	CAC0249	pgk	CAC0710	gerKA	CAC0596
dnaG	CAC1299	tkt	CAC0944	tpiA	CAC0711	gerKB	CAC0597
glgP	CAC1664	ycgL	CAC1497	pgm	CAC0712	gerKC	CAC0598
ftsA	CAC1692	yocI	CAC2687	dacF	CAC0993	yraD	CAC0613
ftsZ	CAC1693	accA	CAC3568	gpr	CAC1275	yraD	CAC0613
spoVS	CAC1817	accD	CAC3569	spoIIP	CAC1276	yraE	CAC0614
spo0A	CAC2071	yueK	CAC1002	tlp	CAC1487	yhcV	CAC0678
glgC	CAC2237	dnaA	CAC0001	sigG	CAC1696	yuzA	CAC0976
glgD	CAC2238	dnaN	CAC0002	ytfJ	CAC2059	dacF	CAC0993
glgA	CAC2239	yaaT	CAC0301	spoIVB	CAC2072	gpr	CAC1275
sigF	CAC2306	yabC	CAC0307	yhfW	CAC2596	spoIIP	CAC1276
spoIIAB	CAC2307	abrB	CAC0310	lonB	CAC2638	tlp	CAC1487
spoIIAA	CAC2308	yknZ	CAC0320	yqhQ	CAC2886	ydfS	CAC1521
yvyD	CAC2847	rocR	CAC0459	spoIIR	CAC2898	sigG	CAC1696
spoVG	CAC3223	ftsE	CAC0497	ksgA	CAC2986	spoIVB	CAC2072
bsaA	CAC1549	ftsX	CAC0498	mcsA	CAC3191	yteA	CAC2116
ymaH	CAC1834	yaaD	CAC0594	yyaC	CAC3727	spoVAE	CAC2302
yveO	CAC2173	yaaE	CAC0595	arsB	CAP0104	spoVAD	CAC2304
yvfC	CAC3073	tkt	CAC0944	arsC	CAP0105	spoVAC	CAC2305
yoeA	CAC3354	rocF	CAC1054			sspA	CAC2365
ycsN	CAC3378	dnaG	CAC1299			sspD	CAC2365
		fruR	CAC1404			sspF	CAC2372
		yrrL	CAC1685			splB	CAC2900
		ylmC	CAC1697			cwlD	CAC3092
		codY	CAC1786			ydfR	CAC3372
		divIVA	CAC2118			yvaB	CAC3421
		ylmG	CAC2119			spoVT	CAC3649
		ylmE	CAC2121				
		flgC	CAC2164				
		flgB	CAC2165				
		yrvI	CAC2273				
		relA	CAC2274				
		apt	CAC2275				
		yvyE	CAC2296				
		yqxD	CAC2825				
		metS	CAC2991				
		ykaA	CAC3094				
		rocE	CAC3164				
		spo0J	CAC3729				
		soj	CAC3730				
		rpmH	CAC3739				

sigK regulon		SigE regulon		SigE regulon	
Bsu gene	Cac number	Bsu gene	Cac number	Bsu gene	Cac number
asd	CAC0022	bofA	CAC0129	glnQ	CAC0112
yhJQ	CAC0075	spoVD	CAC0329	ybbE	CAC0181
ytwI	CAC0092	spmA	CAC0469	ybbD	CAC0182
ribH	CAC0593	spmB	CAC0470	ybbC	CAC0185
ytlD	CAC0618	murB	CAC0510	patA	CAC0375
ytlC	CAC0619	yjmC	CAC0566	yncD	CAC0492
ytcB	CAC0794	spoVR	CAC0581	yhfN	CAC0557
yrhB	CAC0930	exuR	CAC0693	prkA	CAC0579
yrhA	CAC0931	dacB	CAC1267	yhbH	CAC0580
ydfS	CAC1521	spoIIP	CAC1276	birA	CAC0589
pdp	CAC1546	yjmF	CAC1331	yclF	CAC0751
sigK	CAC1689	uxuA	CAC1332	nadA	CAC1025
dapG	CAC1810	cotJB	CAC1337	yjbE	CAC1039
spsE	CAC2187	cotJC	CAC1338	yabJ	CAC1265
spsF	CAC2190	glgP	CAC1664	yqfC	CAC1290
asnO	CAC2243	spoIVA	CAC1713	ycgL	CAC1497
spsC	CAC2260	mmgB	CAC2009	phoP	CAC1506
spsK	CAC2315	spoIIM	CAC2068	phoR	CAC1507
spsL	CAC2331	spoIIIAE	CAC2089	yxiO	CAC1585
spsJ	CAC2332	spoIIIAD	CAC2090	mgsA	CAC1604
spsI	CAC2333	spoIIIAC	CAC2091	yloB	CAC2137
cotF	CAC2683	spoIIIAA	CAC2093	spsC	CAC2260
yojK	CAC2716	spoVE	CAC2126	hemZ	CAC2271
ppiB	CAC2769	murG	CAC2231	dapB	CAC2379
gntP	CAC2835	glgC	CAC2237	ydbI	CAC2432
yabG	CAC2905	glgD	CAC2238	ysnE	CAC2539
thyA	CAC3003	glgA	CAC2239	sipS	CAC2646
yitC	CAC3233	asnO	CAC2243	ysnB	CAC2664
yrpC	CAC3250	yaaH	CAC2747	ysnA	CAC2665
proJ	CAC3253	SpoIIID	CAC2859	yisY	CAC2688
ydfR	CAC3372	spoIIID	CAC2859	acdA	CAC2711
guaC	CAC3471	spoIID	CAC2861	yngF	CAC2712
		mmgA	CAC2873	deaD	CAC2748
		cwlD	CAC3092	ywlC	CAC2882
		yabP	CAC3209	purR	CAC3224
				proH	CAC3252
				proJ	CAC3253
				yodS	CAP0163
				scoB	CAP0164

APPENDIX D : CONSTRUCTION AND CHARACTERIZATION OF A SECOND BUTYRATE KINASE MUTANT IN *C. ACETOBUTYLICUM*

D.1 Introduction

A major limitation in the genetic repertoire of *C. acetobutylicum* is the lack of a well-developed and reproducible gene knockout method. Gene knockouts have been performed in the past, but simple non-replicating plasmid approaches have been difficult to reproduce (Green and Bennett 1996; Green, Boynton et al. 1996), and replicating vectors are time-consuming requiring extensive screening (Harris, Welker et al. 2002). Recent efforts applying the group II intron to clostridia have shown good initial results, but this technique is not readily modified for high-throughput, genome wide studies (Heap, Pennington et al. 2007; Shao, Hu et al. 2007). Additionally, the group II intron system is commercially licensed and limits the potential for use in developing industrially relevant strains.

We aimed to improve upon the gene knockout technology. A newly designed antibiotic resistance gene was used to increase expression using a highly expressed endogenous promoter. As previous efforts utilizing antibiotic resistance markers had limited success in isolating genomic integrants, it was thought that by increasing the expression of the antibiotic resistance genes might result in greater efficiency of mutant isolation. Although antibiotic markers for erythromycin and thiamphenicol resistance have been used successfully with replicating cloning vectors, these plasmid-borne markers have multiple gene copies in each cell, while chromosomal integrants would have only one gene copy. To increase expression and translation of the thiamphenicol resistance gene, an optimized Shine-Delgarno sequence (AGGAGG) replaced the

endogenous sequence and the gene was put under the control of the high expression *ptb* promoter.

For the choice of initial targets, we wanted to have a gene target that had been disrupted before (to test and compare phenotypic data) and, if possible, to aid the design and development of future strains. One target gene chose was the butyrate kinase gene, *buk*. *buk* is part of a bi-cistronic operon also containing the *ptb* gene. Together these two genes encode for the proteins that convert butyryl-CoA to butyrate. The initial study that disrupted the *buk* gene used a non-replicating plasmid and was integrated into the chromosome through a single crossover event and screened using the erythromycin marker. This *buk* disruption strain, PJC4BK, was characterized using southern hybridizations, enzyme assays, and grown in bioreactor experiments to test for impact on butyrate formation and other fermentation products (Green, Boynton et al. 1996). This strain produces high levels of butanol and the generation of a *buk* mutant with another antibiotic marker (thiamphenicol) would aid the construction of improved strains that did not rely on erythromycin, the most commonly used antibiotic in cloning vectors. Here, we describe our own efforts to produce a *buk* gene disruption strain, BKKO, and the unexpected results from that strain.

D.2 Materials and Methods

D.2.1 Plasmid and strain construction

A *buk* fragment was amplified using the primers bk_fwd (ATATATGGCGCGCCATCAATCCTGGCTCGACCTCAACT) and bk_rev (ATATATGGCGCGCCCGTTGTACGCAATTCCGCCTGTTAAG). The PCR fragment was then ligated into the pGW8-TOPO vector and digested with (*Nsi*I). The newly designed

antibiotic marker was PCR amplified and ligated into the digested vector as previously described (Chapter 4) producing pBKKO2F. The plasmid was methylated and transformed as previously described. Strains and plasmids used in this study are shown in Table D.1

D.2.2 Enzyme assays

Cells were collected by centrifugation (5,000g, 4^{o}C) for 10 minutes and frozen. Cells were thawed on ice and resuspended in lysis buffer (10 mM potassium phosphate pH 7.2 and 1 mM dithiothreitol) to an OD equivalent of 10 (e.g. 10 mL of OD 1.0 cells resuspended in 1 mL lysis buffer). 1 mL of cell suspension was centrifuged and resuspended in lysis buffer. Cells were lysed by sonication (power setting 4 for 10 minutes). The lysate was centrifuged twice (15,000g for 20 minutes at 4^{o}C) and stored at 4^{o}C. Protein was used the same day for quantification and enzyme assays. Protein quantification used the Bradford method from a Bio-Rad protein assay kit.

D.2.3 PTA assay

PTA activity was determined using a coupled reaction (Brown, Jones-Mortimer et al. 1977). Acetyl-CoA formed by PTA is trapped by condensation with oxaloacetate in the presence of citrate synthase. This reaction is coupled to the reaction producing oxaloacetae and NADH from L-malate. The formation of oxaloacetate was measured at 340 nm as the formation of NADH. The extinction coefficient for this reaction is 6.22 mM^{-1} cm^{-1}. The reaction contained 100 mM Tris (pH 8), 5 mM $MgSO_4$, 0.5 mM NAD, 0.5 mM CoA, 5 mM L-malate, 10 mM acetyl-phosphate, 50 U L-malic dehydrogenase, 5 U citrate synthase, and 1 μL crude extract in 1 mL. Acetyl-phosphate was added to the other components to initiate the reaction. One unit of PTA activity is defined as the amount of enzyme catalyzing the formation of 1 μmol of NADH

Strain or Plasmid	Relevant Characteristics[a]	Source or Reference[b]
Bacterial Strains		
C. acetobutylicum		ATCC
PJC4BK	*buk*-	(Green, Boynton et al. 1996)
BKKO2F	*buk*-	This study
E. coli		
Top10		Invitrogen
ER2275		New England Biolabs
Plasmids		
pAN1	Cm^r, Φ3T I gene, p15A origin	(Mermelstein and Papoutsakis 1993)
pSOS94[c]	acetone operon (*ptb* promoter)	Soucaille and Papoutsakis, unpublished
pLHKO [c]	Cm^r	(Harris 2001)
pGW8-TOPO		Invitrogen
pSOS94-CMptb [c]	Cm^r (*ptb* promoter)	This study
pBKKO2F	*buk* gene fragment, Cm^r (*ptb* promoter)	This study

[a]Cm^r, chloramphenicol/thiamphenicol resistance gene; *ptb*, phosphotransbutyrylase gene; *aad*, alcohol/aldehyde dehydrogenase gene; *thl*, thiolase gene; *adc*, acetoacetate decarboxylase gene; *ack*, acetate kinase gene; *buk*, butyrate kinase gene.

[b]ATCC, American Tissue Culture Collection, Rockville, MD

[c]contains the following: ampicillin resistance gene; macrolide, lincosimide, and streptogramin B resistance gene; *repL*, pIM13 Gram-positive origin of replication; ColE1 origin of replication

Table D.6.1 Bacterial strains and plasmids used in this study

per minute at room temperature. PTA activity is reported as units of activity per mg of protein in the crude extract.

D.2.4 PTB assay

PTB activity was determined by monitoring the formation of a complex between CoA and 5'5'-dithiobis (2-nitrobenzoic acid) (DTNB) at 412 nm (Wiesenborn, Rudolph et al. 1989). The extinction coefficient is 13.6 mM^{-1} cm^{-1}. The reaction contains 100 mM potassium phosphate (pH 7.4), 0.2 mM butyryl-CoA, 0.08 mM DTNB, and 1 μL crude extract. Butyryl-CoA was used to initiate the reaction. One unit of PTB activity is defined as the amount of enzyme catalyzing the formation of 1 μmol of the DTNB-CoA complex per minute at room temperature. Specific PTB activity is reported as units of activity per mg of protein in the crude extract.

D.2.5 AK and BK assay

AK activity was determined using the ability of acyl-phosphates to rapidly form hydroxamic acids at neutrality (Rose 1955). Hydroxamic acids are measured by the formation of a colored ferric-hydroxamate complex in acid solution. 300 μL of fresh substrate solution (5 mL of 3.2 M potassium acetate (pH 7.4), 1 mL of 1 M Tris (pH 7.4), and 200 μL of 1 M $MgSO_4$) and 350 μL neutral hydroxylamine solution (5 mL of 4 M hydroxylamine-HCl and 5 mL of 4 M potassium hydroxide) is combined with 250 μL of 10 fold diluted crude extract and incubated at 29^{o}C. 100 μL of 0.1 M ATP is used to initiate the reaction. At 5 to 10 minute intervals, 200 μL of the reaction is transferred to 200 μL of 10% TCA to stop the reaction. The mixture is combined with 400 μL of 1.25% $FeCl_3$ in 1 N HCl. The ferric-hydroxamate complex is quantitated at 540 nm using an extinction coefficient of 0.69 mM^{-1} cm^{-1}. One unit of AK activity

is defined as the amount of enzyme catalyzing the formation of 1 μmol of the ferric-hydroxamate complex per minute at 29^{o}C. Specific activity is reported as units of activity per mg of crude extract. BK activity is performed in an identical fashion by substituting potassium butyrate for potassium acetate.

D.3 Results

D.3.1 Transformation

In the past, efforts to transform *C. acetobutylicum* with non-replicating plasmids for the purpose of genomic integrations, large amounts of DNA were used (15 μg). We used similar amounts of DNA and despite our attempts to increase the expression of the antibiotic resistance gene, very few recombinants were isolated (less than 1 transformant per μg). However, unlike previous attempts to duplicate past gene deletions, transformants were isolated, despite the low efficiency.

Similar studies were performed in the pSOL1 deficient strain M5 and higher transformation efficiencies were achieved (as high as 10 transformants/μg). The M5 strain also had higher transformation efficiencies with replicating plasmids, suggesting that chromosomal integrations could be performed at higher efficiencies in the WT, if the transformation protocol could be improved. The transformation efficiency of the replicating plasmids was found to be about 4 orders of magnitude higher than for the chromosomal integrations. This suggests that the DNA is successfully being introduced into the cell, but the cellular machinery does not readily integrate this DNA into the chromosome.

D.3.2 PCR confirmation

Transformants from the electroporation were initially screened using colony PCR. This confirmed the integration of the antibiotic selection marker into the genomic locus desired. Once the initial screening identified positive clones for the integration, genomic DNA was isolated and sequenced to confirm the exact nature of the insertion. A set of four primer pairs was used to span the *buk* gene locus, beginning upstream of the homologous regions and ending downstream of the second homologous region. The integration of the plasmid into the chromosome was determined to be a single crossover (Fig D.1). A single crossover event of the plasmid could occur using either of the two homologous regions contained on the plasmid. The BKKO strain used the second region of homology for integration and did not undergo a second crossover event.

A double crossover mutant would be desirable. A single crossover mutant still contains both the genomic and plasmid borne *buk* fragments and a crossover event could occur to excise the plasmid DNA resulting in a reversion to the wildtype. To encourage this second crossover to occur, the strain was grown in liquid media and vegetatively transferred for 100 generations. Following the 100 generations, individual colonies were grown and analyzed using colony PCR. Cells were grown in liquid media without antibiotic. 96 mutants were analyzed and none were observed to have undergone any genomic rearrangement from the original single crossover. This shows that the single integration is very stable and also that recombination machinery within the cell is not very active.

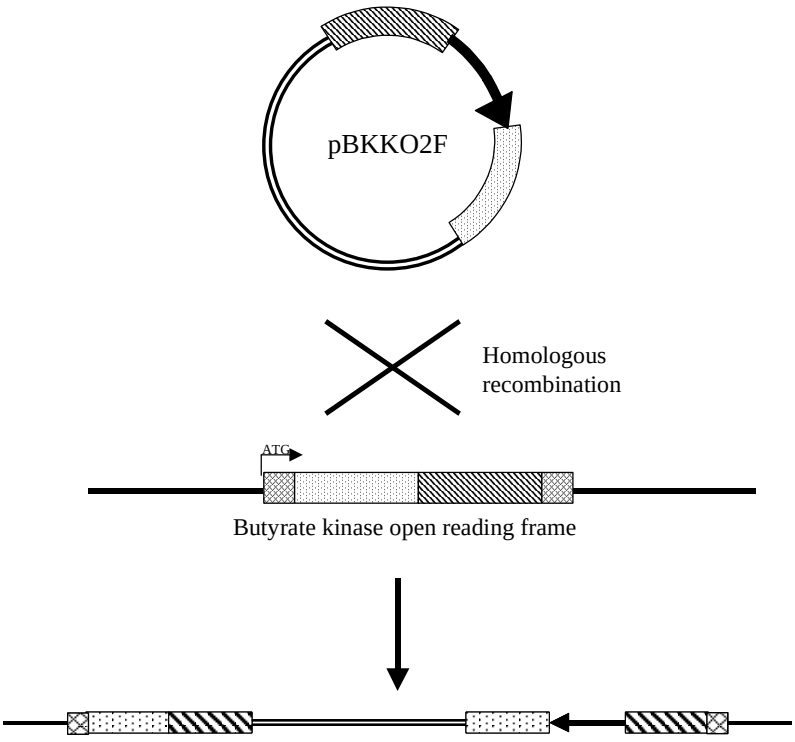


Figure D.1 Genetic sequence and homologous recombination of BKKO2F

D.3.3 Product formation

The BKKO2F and PJC4BK strains were grown in static flask and sampled for product formation. The average results for two replicates and wildtype data are shown in Table D.2. The PCJ4BK strain shows enhanced production of both butanol and ethanol as has been shown previously. Additionally, acetone formation is decreased dramatically compared to the WT. Both acetate and butyrate are produced at very low levels and nearly completely re-assimilated by PJC4BK. The BKKO2F strain shows a dramatic shift in product formation that has never been seen before. Butanol and acetone are barely produced reaching levels of just five and three mM. No butyrate was detected in any timepoint, while acetate levels were comparably high in this strain. Ethanol formation was the dominant product produced reaching 95 mM, but was responsible for about two-thirds of the products produced, an ethanol selectivity never before observed in this strain.

D.3.4 Enzyme assays

Based on the product formation a major difference is the lack of butyrate formation in the BKKO strain. The PJC4BK strain has previously been shown to have enhanced PTB activity and it was thought that the BKKO integration may also lead to decreased PTB activity. This decreased PTB activity might help to explain the lack of butyrate and possible further implicate butyryl-phosphate as a signaling molecule and necessary for strong butanol production.

Strains	Product Concentrations (mM)				
	Butanol	Ethanol	Acetone	Acetate	Butyrate
824	162	23	94	17	19
PJC4BK	180	96	50	4	13
BKKO2F	5	95	3	44	0

Table D.6.2 Product formation of *buk*- strains and the wildtype strain

All enzymes in both the acetate and butyrate formation pathways were assayed for their activity in both mutant strains and also the wildtype strain (Fig D.2). The butyrate kinase assay for all strains shows exactly what is expected, namely nearly a lack of activity in both mutant strains, while high activity is present in the WT strain. Comparing the PTB activity there does not seem to be a significant difference between the strains. PTB activity is highest in the 824 strain, but there is not a dramatic difference between the strains, and certainly there are significant levels of PTB activity in the BKKO strain.

In the acetate formation pathway, acetate kinase activity is highest in the BKKO strain, but both the WT and the PJC4BK strain show a bimodal activity profile. Examining the PTA activity, PJC4BK has the highest activity in 3 out of 4 timepoints analyzed, but PTA activity is high in all strains tested. Overall, the dominant result is the lack of BK activity in the two mutants, but no other insight is readily available from these assays.

D.4 Discussion

Here we have constructed a new butyrate kinase mutant and compared it against the WT and another butyrate kinase mutant. Despite close similarities in the genomic disruption both occurring from single crossovers (although both used different vector backbones), and similar enzyme activity profiles, dramatic differences in product formation are evident. It was initially suggested that the BKKO strain may also suffer from a lack of PTB activity, but this has not been supported by the enzyme assays.

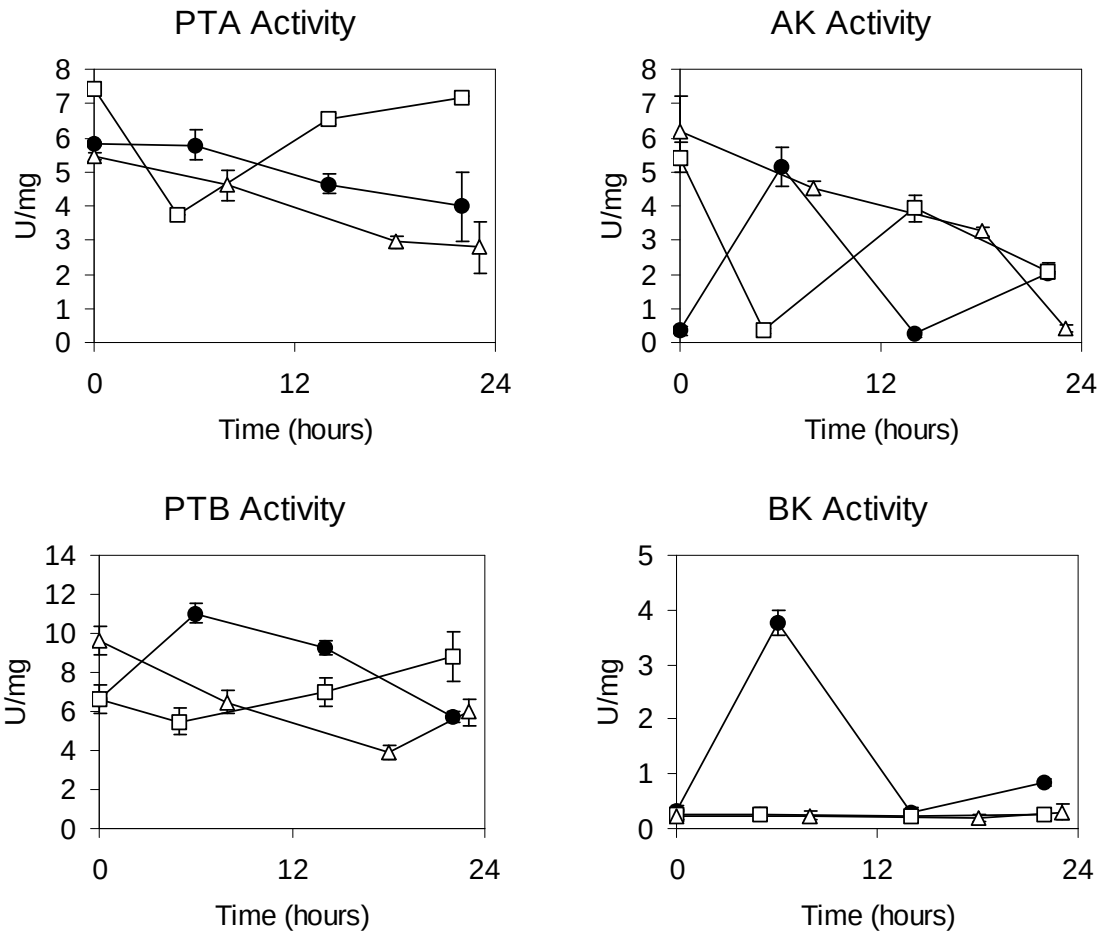


Figure D.2 Enzyme assays of wildtype and *buk*- strains.

Enzyme activity of WT (closed circles), PJC4BK (open squares), and BKKO2F (open triangles) are shown. Growth of all strains were normalized to an OD of 1.0 at hour zero.

Despite the questions regarding the mechanism of the lack of butyrate formation, the BKKO strain may be a valuable strain for probing ethanol formation in this strain and serve as a basis for ethanol production. Large ethanol titers have never been observed without significant butanol formation. In those previous cases, it has been suggested that increased *aad* expression or *aad* overexpression was responsible for the large amounts of ethanol. *C. acetobutylicum* has a pyruvate decarboxylase gene and it may be responsible for the large ethanol titers observed here, suggesting a new model for ethanol formation in this strain. It is also curious that no butyrate is formed in BKKO2F. Typically, when one acid pathway is downregulated or disrupted, the other pathway exhibits cross-reactivity, allowing both acids to be formed. Additionally, almost no acetone is formed in this strain. It may be that the genes for acetone formation are not induced in this strain. The in vitro data suggest that acetate is a preferred substrate for re-assimilation and acetone formation, but despite available acetate, very little acetone was formed in the BKKO2F strain that produces no butyrate.