AN INVESTIGATION OF THE ROLES OF SMALL RNA ON SOLVENT TOLERANCE AND PRODUCTION IN CLOSTRIDIUM ACETOBUTYLICUM

by

Alexander Joseph Jones

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Science in Biological Sciences

Winter 2015

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

AN INVESTIGATION OF THE ROLES OF SMALL RNA ON SOLVENT TOLERANCE AND PRODUCTION IN CLOSTRIDIUM ACETOBUTYLICUM

by

Alexander Joseph Jones

Approved: _____
Eleftherios Terry Papoutsakis, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____
Robin W. Morgan, Ph.D.
Chair of the Department of Biological Sciences

Approved: _____
George H. Watson, Ph.D.
Dean of the College of Arts and Sciences

Approved: _____
James G. Richards, Ph.D.
Vice Provost for Graduate and Professional Education

ACKNOWLEDGMENTS

First, I offer my thanks to the One who answers my constant prayers and provides me the strength I need, and for the capacity to know and celebrate that Science and Faith are not mutually exclusive;

My love and thanks to Mom and Dad, for all the love, support, and encouragement you have shown me all my life, to Hannah and Zach, for the help you may not even know that you give me, and to Alysia, whose constant love and support mean so much to me – you have all given me such support and motivation, and I am blessed to have all of you in my life;

I owe much gratitude to so many teachers throughout the course of this academic and scientific journey, but especially to Greg McDivitt, for your enthusiasm for science learning that hooked me for life and drove me to want a better and deeper understanding of the world around me, and Brother Tom McPhillips, PhD, for your unending passion, guidance, and friendship – I am forever a Biology Fan;

Thank you to all the lab members of the Papoutsakis group – Alan, Brian, Chen-Yuan, Elli, Jin Lin, Keerthi, Kyle, Lisa, Matt, Mohab, Nich, Sergios, Stefan, Stephanie, and Vanessa – for camaraderie and fellowship in and out of the lab, for the stream of support, advice, and assistance from the smallest lab task to the largest conceptual challenge; thanks especially to Mohab and Sergios, for helping me transition to the lab and to graduate school, and making me feel at home, to Alan, for your ceaseless help and friendship, especially during the setbacks, and most of all to Keerthi, whose mentorship and instruction for this project has meant so much;

Thank you also to my committee members, Dr. Boyd and Dr. Herson, for your support and feedback, as well as Dr. Duncan for your assistance and direction throughout my time at UD;

And most of all, my gratitude, appreciation, and respect to my advisor, E. Terry Papoutsakis, PhD, for your patience, guidance, and support; with your direction and encouragement, I have become a truly better scientist, scholar, and researcher.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Clostridium acetobutylicum is a Gram-positive, anaerobic, spore-forming bacterium with considerable engineering potential regarding its fermentative metabolism and is a model organism for the ABE fermentation and the anaerobic Firmicutes. Much recent work has been devoted to elucidating and optimizing the regulatory and stress response mechanisms in *C. acetobutylicum*, with increasing focus on transcriptomics.

6S RNA and tmRNA are small, noncoding RNA molecules with observed influences on stress response in many prokaryotic organisms. 6S RNA regulates gene expression during the transition from exponential to stationary phase growth, and tmRNA contributes to quality control activities in cells under stress as a component of a ribosomal rescue system. In this study, we investigated the influence of 6S RNA (sCA_C1377) and tmRNA (sCA_C834) overexpression on tolerance to butanol in *C. acetobutylicum* ATCC 824. Strains overexpressing each sRNA sequence exhibited sustained higher percent survival under toxic concentrations of butanol stress through late stationary phase growth; this phenotype was particularly evident for 6S RNA overexpression. In addition to higher butanol tolerance, overexpression of 6S RNA also resulted in an increase in butanol production. Overexpression of a homologous 6S RNA sequence showed no discernible effects on tolerance or solvent production.

In recent years, a sRNA sequence called *solB* (sCA_P176) has been under discussion as a negative regulator of solventogenesis genes in *C. acetobutylicum*. To investigate this putative regulatory function, we generated a *solB* overexpression strain

and a *solB* deletion strain, the latter generated using a novel allelic exchange system in clostridia (Al-Hinai, et al, 2012). Overexpression of *solB* generated the anticipated solvent-negative phenotype, and quantitative PCR and Northern analysis revealed strong expression reduction of the solventogenic genes (*adhE1*, *ctfA*, and *adc*). Surprisingly, deletion of *solB* similarly produced no solvents, although without loss in solventogenic gene transcripts. Antisense knockdown of *solB* generated the same solvent-negative metabolic phenotype. Putative secondary structures and base-pairing predictions suggested respectable interaction between *solB* and the UTRs of each target transcript. Also, QPCR data revealed a possible early induction of solventogenic gene transcription in the deletion strain. Overall, this study indicates a clear role of *solB* on solventogenesis, but one that is more complex than initially thought; it is possible, given the results presented here, that a bimodal function for *solB* exists.

Chapter 1

INTRODUCTION

1.1 History, General Characteristics, and Relevance of Clostridia

The genus Clostridium consists of Gram-positive, rod-shaped, anaerobic, endospore-forming bacteria in the phylum Firmicutes. (Tracy, et al, 2012) Clostridia first appeared approximately 2.7 billion years ago, before the Great Oxidation Event, and are the evolutionary predecessors of the aerobic Firmicutes (i.e. the Bacilli). (Paredes, et al, 2005; Tracy, et al, 2012) Vegetative clostridial cells are strictly anaerobic, but upon exposure to a harmful or stressful environment, a clostridial cell can generate an inert, protective endospore and survive long-term exposure to oxygen, pH changes, or other harmful, toxic, or stressful environments. Clostridia are largely saprotrophic organisms, and they can be found in most anaerobic environments that contain organic material, including soils, aquatic sediments, and anaerobic human and animal tissue. (Tracy, et al, 2012)

Clostridia are relevant to human health and physiology as well as for their range in metabolic processes. Some are known human pathogens, like *C. botulinum*, *C. tetani*, *C. perfringens*, and *C. difficile* (Paredes, et al, 2005). The majority of clostridia, however, are harmless to humans; some species, like *C. sporogenes*, have been investigated as a drug delivery method in the treatment of anoxic cancer tissues. Most clostridia, though, garner attention and importance for the considerable industrial potential of their wide variety of metabolic processes. These include the cellulolytic clostridia (*C. phytofermentans* and *C. thermocellum*) (Paredes, et al, 2005), acetogens

like *C. carboxidivorans*, and the solvent producers, or solventogenic clostridia, which include *C. acetobutylicum* and *C. beijerinkii*.

This impressive array of metabolic processes means clostridia can collectively utilize a wide variety of carbon substrates and yield many useful metabolites, providing rich potential for industrial development of solvent production, particularly biofuels. Clostridial metabolite products include an array of commodity chemicals, including some that are usable directly or indirectly as biofuels; in fact, clostridial fermentation was the primary source of butanol and acetone for 40 years during the early 20th century. (Papoutsakis, 2008) For an excellent and extensive review of the wide array of metabolic substrates, products, and programs that exist in the clostridia genus, see Tracy, et al, 2012. Engineering industrial strains of Clostridium, especially ones capable of utilizing multiple varieties of feedstock, represents a major endeavor in commercial bio-based chemical production. (Gheshlaghi, et al, 2009; Green, 2011; Ezeji, et al, 2007) Greater substrate utilization gives clostridia a clear advantage over the organisms more commonly used in biorefinery applications (i.e. *Escherichia coli*, *Saccharomyces cerevisiae*, and other yeasts). Although deficiencies in genetic tools for the manipulation and study of clostridia have limited research in this field, many significant advances have been made in the last 10 years, with new molecular techniques only adding to this promising and upward trend. (Tracy, et al, 2012)

1.2 *Clostridium acetobutylicum* and the Utilization of ABE Fermentation

The saccharolytic, solventogenic *Clostridium acetobutylicum* was first isolated between 1912 and 1914, and by the 1920s its fermentation of acetone, butanol, and ethanol (or ABE fermentation) was used extensively for the production of acetone, especially for gunpowder in World War I. (Nolling, et al, 2001) Continued

development of *C. acetobutylicum* and its ABE fermentation led to a rapid expansion of industrial butanol fermentation facilities worldwide. This process thrived as the largest source of butanol until the 1950s, when it fell out of favor as the growing dominance of petrochemical refining methods offered an easier and more economical source of commodity chemicals. (Nolling, et al, 2001) Interest in clostridial fermentation was renewed beginning in the 1970s, following the oil crisis of 1973, rising oil prices, and greater concerns about conservation of fossil fuels and environmental impacts. Studies aiming to better understand these organisms and utilize their metabolism have grown since the 1980s, especially regarding the potential of biobutanol. Butanol is a widely used chemical, and it can be used as a direct replacement for gasoline, with no modification to existing infrastructure. (Lee, et al, 2008) Compared to ethanol (another fuel additive), butanol boasts a higher energy profile and is less volatile and corrosive. (Durre, 2007) The biggest requirement to make clostridial fermentations economically and industrially competitive again is to increase solvent production titers to a more viable level (Papoutsakis, 2008), and research has been conducted in several areas to achieve this goal, (Lutke-Eversloh and Bahl, 2011) from efforts to abolish sporulation to increasing solvent tolerance or aerotolerance phenotypes. (Hillmann, et al, 2008; Tracy, et al, 2012)

In addition to its promising potential for industrial-scale biofuel and solvent production, *C. acetobutylicum* is relevant as a chosen model organism for clostridia, particularly the solventogenic clostridia, for understanding stationary-phase processes (like sporulation and solventogenesis) and the development of genetic tools for studying and manipulating these organisms.

1.3 Genetic Organization and Molecular Engineering of *C. acetobutylicum*

The genome of *C. acetobutylicum* consists a single 3.94-Mb circular chromosome, containing 3740 open reading frames (ORFs) and 107 RNA genes, and a 192-kb megaplasmid with an additional 178 protein-coding genes. (Nolling, et al, 2001) Like all prokaryotes, the genes of the *C. acetobutylicum* genome are grouped in polycistronic operons, also referred to as transcriptional units (TU); a map of the TUs for *C. acetobutylicum* was determined by genome-wide computational analysis, (Paredes, et al, 2004) revealing a total of 2268 TUs, with 128 of them present on the megaplasmid. This megaplasmid, termed pSOL1, contains the primary genes for solvent production. It has been determined that continuous cultures or repeated vegetative transfers of *C. acetobutylicum* can lead to the loss of pSOL1 and subsequent degeneration into a non-solvent-producing, asporogenous strain (M5). (Cornillot, et al, 1997; Papoutsakis, 2008)

As mentioned above, many tools have been developed over the past 20 years for the genetic manipulation and study of clostridia. Accumulating advances in molecular techniques have advanced the study of *C. acetobutylicum* and the development of genetic tools for its study and manipulation. These advances are reviewed by Tracy, et al (2012), and Papoutsakis (2008) and have contributed significantly to the understanding of cellular processes like sporulation, acid and solvent production, gene regulation, and stress response. Despite these advances, much more remains to be elucidated, particularly in the area of regulatory mechanisms, in order to more fully understand and manipulate *C. acetobutylicum*. More recent studies have involved the use of RNA Deep Sequencing; as described in subsequent sections, this technique has helped identify RNA sequences, referred to as small or non-coding RNA (and hereafter called sRNA in this study), that have been

determined to play increasingly essential and prevalent roles in the expression and regulation of genes and cellular processes.

1.4 Overview of Clostridial Growth and Sporulation

The process of sporulation in *C. acetobutylicum* is a mid- to late-stationary phase event and one that is closely related to the production of solvents. (Lee, et al, 2008) In *Bacillus subtilis*, the shift to stationary phase events like sporulation are triggered by nutrient limitation (Sonenshein, 1989), but in clostridia, low intracellular pH and rising butyrate levels are the conditions that induce the shift to stationary-phase processes, namely, solvent formation, sporulation initiation, and granulose accumulation. (Paredes, et al, 2005) A proposed model (Grupe and Gottschalk, 1992) indicates that these conditions result in changes to ATP and NAD(P)H concentrations, which (as indicators of energy sufficiency) are the two signals for phase transition in *C. acetobutylicum*.

Understanding of the clostridial sporulation pathway remains incomplete, but much work has been done in the last 10 years to elaborate on its mechanisms and control. Spore formation in clostridia is a pathway of cell differentiation that begins in a similar way to mitosis, except that the cell divides asymmetrically into two morphologically distinct forms called the mother cell and the prespore. The prespore becomes engulfed by the mother cell as differentiation and spore maturation progresses, becoming a forespore and then an endospore before the terminally differentiated free spore is finally released. A thorough investigation of the transcriptional profile of sporulation in *C. acetobutylicum* was conducted by Jones, et al, (2008) and is summarized here: exponential phase in clostridia occurs from about six through 10 hours of growth, with its characteristic acid production peaking at

around 16 hours. The transition from exponential to stationary phase growth occurs between 10 and 18 hours and includes the beginning of solvent formation gene expression. Early stationary phase (18-24 hours) exhibits high expression of solvent production and stress response genes, and middle stationary phase (24-36 hours) is when sporulation initiation genes begin to be expressed. Early and middle stationary phase growth (18 to 36 hours) is characterized by a cell morphology called the clostridial form, a fattened, cigar-shaped morphology that is believed to be the active solvent-producing form. By late stationary phase (36+ hours) spore maturation genes and late-phase vegetative cells are highly expressed as the sporulation process nears completion, with mature free spores observed as early as 44 hours.

As in the *B. subtilis* sporulation model, the master transcriptional regulator Spo0A regulates both sporulation and solventogenesis genes in *C. acetobutylicum*, and it appears to orchestrate a careful temporal balance between these two stationary phase events. (Harris, et al, 2002) Spo0A is part of the two-component signal-transduction protein family, with homologues only in sporulating bacteria. Disruption of *spo0A* in *C. acetobutylicum* blocks sporulation and inhibits solventogenesis, indicating its essential role in the activation of both stationary-phase processes. (Brown, et al, 1994; Harris, et al, 2002) Spo0A is activated via phosphorylation by histidine kinases, and although most of the downstream genetic components of the sporulation cascade are also generally conserved from the *B. subtilis* model, clostridial sporulation follows a different model. (Paredes, et al, 2005; Lee, et al, 2008) Spo0A activates expression of the solvent formation genes (described in the next section) as well as the major sporulation initiation genes, which include the sporulation-specific sigma factors (σF, σE, σG, σH, and σK). (Papoutsakis, 2008) Peak expression of *spo0A* occurs during the

transition phase at about 10 hours, just ahead of its target genes, the solvent formation genes and *spoIIAA*, the operon containing the first sporulation sigma factor, σF, which is active only in the prespore morphology. (Alsaker, et al 2005; Jones, et al, 2008) Each subsequent sigma factor relies on the previous one for activation, and all are active at specific times and are specific to either the mother cell or the prespore/endospore. (Papoutsakis, 2008; Bi, et al, 2011; Al-Hinai, et al, 2014)

Solvent production by the clostridial form occurs for a comparatively short period before sporulation begins; also, sporulating strains are not suitable for continuous or fed-batch fermentations because of their tendency to degenerate. (Papoutsakis, 2008) Therefore, the ideal industrial solventogenic clostridial strain should exhibit abolished sporulation with preserved solvent production, although the pursuit of this more efficient strain is difficult due to the common *spo0A* regulation of both stationary phase processes. Molecular engineering efforts therefore have focused downstream of the shared transcriptional regulator, and inactivation studies of the sporulation-specific sigma factors have resulted in blocked sporulation and the preservation of solvent production. (Jones, et al, 2011; Tracy, et al, 2011; Bi, et al, 2011; Al-Hinai, et al, 2014)

1.5 Review of Metabolism in *C. acetobutylicum*

1.5.1 Overview of Clostridial ABE Fermentation

The metabolism of *C. acetobutylicum* is biphasic in nature (Figure 1.1). After the generation of acetyl-CoA via glycolysis, clostridial fermentation begins with acidogenesis, the acid production phase in which the cell converts acetyl-CoA into acid products, specifically acetate and butyrate, utilizing substrate-level

phosphorylation for ATP generation (Tracy, et al, 2012). Bennett and Rudolph (1995) review the central carbon metabolism pathway, highlighting the conversion of acetyl-CoA to butyryl-CoA (an intermediate process leading to the production of butyrate). Acidogenesis is an exponential-phase process, occurring from around six to 10 hours of culture growth. The enzymes and intermediate products of acidogenesis are outside the scope of this thesis, but they are presented in more detail by Tracy, et al, (2012) and other reviews of clostridial metabolism.

The secondary phase of ABE fermentation is solvent formation, or solventogenesis; like sporulation initiation, it is a stationary-phase event whose genes are activated in late exponential phase by the transcriptional master regulator Spo0A (Ravagnani, et al, 2000) and triggered by the drop in pH and accumulation of butyrate generated in the acidogenesis phase. In this phase the cell produces butanol, acetone, and ethanol in combination with the reuptake and conversion of the acid products (butyrate and acetate). (Papoutsakis, 2008; Gheshlaghi, et al, 2009) At the onset of solventogenesis, concentrations of butyryl-CoA increase while free CoA and acetyl-CoA decrease, in conjunction with a drop in acidogenic enzyme activity and a sharp rise in solventogenic enzyme activity. (Boynton, et al, 1994) Significant gene expression changes coincide with this metabolic shift, including the expression of the solvent formation genes. (Durre, et al, 1987; Durre, et al, 1995) These genes, and product changes resulting from simple alterations to their expression, affect the transcriptional profiles of several other cellular processes, including stress response, sporulation, motility, and fatty-acid biosynthesis gene expression (Tummala, et al, 2003a, c; Wang, et al 2013), pointing to the complex regulation of and intricate

connection between the solventogenic genes and other major cellular processes in *C. acetobutylicum*.

1.5.2 The Solventogenic Enzymes

Transcriptomic studies of the metabolic shift to solventogenesis show that expression of the major solventogenic genes begins at around 10 hours and steadily increases throughout stationary phase, paired with the up-regulation of sporulation initiation and lipid synthesis genes and the down-regulation of translation-related genes in stationary phase. (Alsaker and Papoutsakis, 2005) Solventogenesis consists of four genes (Figure 1.2) that code for the three major enzymes responsible for primary solvent production and are all encoded locally on the pSOL1 megaplasmid. (Cornillot, et al, 1997) The first gene, *adhE1*, coding for a bi-functional acetaldehyde-CoA/alcohol dehydrogenase, functions in butanol and ethanol production. The adjacent genes *ctfA* and *ctfB* code for the two subunits of the CoA-transferase (CoAT) enzyme that catalyzes the transfer of coenzyme A (CoA) between acetoacetyl-CoA and either butyrate or acetate. The *ctfA* gene lies 63 base pairs downstream of *adhE1*, and these two genes, along with *ctfB*, comprise the *sol* operon, a 4.1-kbp polycistronic transcript on the pSOL1 coding strand. Lastly, *adc* codes for the enzyme acetoacetate decarboxylase, the enzyme responsible for the final step in acetone production; *adc* exists as a monocistronic operon 65 bp downstream of *ctfB* but encoded on the complementary strand. All four genes are collectively designated as the *sol* locus (Figure 1.2) and will be referred to as such in this thesis. All *sol* locus genes are expressed in solventogenic cells but show no detectable expression in acidogenic cells (Gerischer, et al, 1992; Fischer, et al, 1993); in fact, transcription is observed three to four hours before the detection of solvent products. (Gerischer and Durre, 1992)

Butyraldehyde dehydrogenases (BYDH) catalyze the conversion of butyryl-CoA to butyraldehyde (releasing CoA and oxidizing NAD(P)H), and butanol dehydrogenases the conversion of butyraldehyde to butanol. *C. acetobutylicum* exhibits both activities during solventogenesis, and **AdhE1** is the monomeric bi-functional acetaldehyde-CoA/alcohol dehydrogenase that performs both functions. The gene for this enzyme, ***adhE1*** **(CA_P0162)**, is a 2,589 bp ORF, located 63 bp upstream of *ctfA* on the coding strand of pSOL1 and coding for an 873-amino acid protein. (Fischer, et al, 1993; Nair, et al, 1994) Homology searches (Nair, et al, 1994) determined that its N-terminal region (739 bp) shows considerable homology to an aldehyde dehydrogenase domain, and the C-terminal region (663 bp) shows strong conservation with bacterial and yeast alcohol dehydrogenases. Primer extension analyses (Fischer, et al, 1993; Nair, et al, 1994) identified two transcriptional start sites (TSS) 83 and 243 bp upstream of the start codon, and their corresponding distal and proximal promoters exhibit different activity levels. The distal promoter, which shows strong homology to clostridial and Gram-positive consensus promoter sequences (Young, et al, 1989), exhibits very weak transcriptional activity; alternatively, the proximal promoter sequence is less conserved but exhibits most of the observed transcriptional activity for the *sol* operon. (Fischer, et al, 1993; Nair, et al, 1994; Scotcher, et al, 2003) Restoration of butanol production, but not acetone production, to near wild type levels in the M5 degenerate strain (which, due to the loss of the pSOL1 megaplasmid, lacks all solventogenic genes) by plasmid expression of *adhE1* (Nair and Papoutsakis, 1994) suggests that butanol production in solventogenic cultures originates largely from AdhE1, and this enzyme heavily prioritizes butanol production over ethanol production. The gene for the AdhE1 enzyme (*adhE1*, CA_P0162) should

not be confused with *adhE* (or *adhE2*, CA_P0035), the gene for a second acetaldehyde/alcohol dehydrogenase that is also encoded on pSOL1; this enzyme functions primarily in alcohologenic growth and is described below.

The neighboring genes ***ctfA*** and ***ctfB*** **(CA_P0163** and **CA_P0164)**, separated by one base pair, code for the two subunits of acetoacetyl-CoA:acetate/butyrate:CoA transferase, or CoA transferase (**CoAT**). The combined ORF sequences of *ctfA*, *ctfB*, and *adhE1* comprise the 4.1-kbp transcript of the *sol* operon. (Gerischer and Durre, 1992) The two subunits of CoAT were first purified and characterized by Wiesenborn, et al, (1989) and cloned by Cary, et al (1990). Knockdown of *ctfB* results in partial degradation of the entire *sol* operon transcript, affecting AdhE1 activity and causing a subsequent reduction in solvent production. (Sillers, et al, 2008) By catalyzing the transfer of CoA from acetoacetyl-CoA to either butyrate or acetate, CoA transferase facilitates the uptake of these acids, detoxifying their inhibitory effects and leading to the non-growth-associated solvent production of stationary phase. CoAT functions together with AADC in the production of acetone, and the uptake of butyrate and acetate by CoAT also facilitates the production of butanol and ethanol by AdhE1.

The gene ***adc*** **(CA_P0165)** codes for the enzyme acetoacetate decarboxylase (**AADC**), which catalyzes the final step of acetone production, namely, the decarboxylation of acetoacetate into acetone, directly with no cofactor necessary. (Davies, et al, 1943) The *adc* gene is a 865-bp monocistronic operon that contains the 735 bp AADC ORF and is transcribed at a single transcriptional start site from a well-conserved promoter sequence. (Gerischer and Durre, 1992) Spo0A activity on this *adc* promoter peaks at transitional phase (10 to 11 hours) (Ravagnani, et al, 2000), in agreement with the previously described transcriptional studies of *sol* locus gene

expression. A rho-independent, bidirectional terminator sequence exists between the *ctfB* and *adc* coding sequences. (Petersen, et al, 1993) AADC serves as the key enzyme in the acetone formation pathway, with its activity driving bidirectional CoAT activity toward the formation of acetoacetate. (Hartmanis, et al, 1984) However, AADC does not appear to be the rate-limiting enzyme in acetone production; antisense RNA knockdown of AADC does not affect acetone production, but similar knockdown of CoAT severely diminishes acetone titers, indicating that CoAT is the rate-limiting enzyme of the acetone formation pathway. (Tummala, et al, 2003b)

1.5.3 Metabolic Engineering of the Solventogenic Genes

There have been several targeted inactivation and overexpression studies of different combinations of solventogenic and acidogenic enzymes in an effort to optimize solventogenesis, with a focus on butanol production. Inactivation and antisense knockdown of butyrate kinase (*buk*, expressed in acidogenesis) resulted in earlier onset of solventogenesis, lower butyrate production, and an increase in butanol titers, (Green, et al, 1996; Desai and Papoutsakis, 1999) but antisense knockdown of the acidogenic phosphotransbutyrylase (*ptb*) significantly reduced solvent production. (Desai and Papoutsakis, 1999) Additional work utilizing metabolic flux analysis (Desai, et al, 1999) has further elaborated on the influence of acid formation genes on solvent production and the potential of targeting these genes for the manipulation of carbon flow. Individual, targeted mutagenesis of the solventogenic genes unsurprisingly has yielded reductions in the corresponding solvent products. (Cooksley, et al, 2012) However, knockdown of CoAT, combined with overexpression of AdhE1, yields a higher butanol/acetone ratio, decreased acetone yields, and a significant increase in ethanol production. (Sillers, et al, 2008)

Inactivation of *adc* and overexpression of the *sol* operon also generates higher butanol and ethanol titers. (Hou, et al, 2013) These studies exemplify the potential that exists in manipulating clostridial metabolism to redirect carbon flow toward and optimize solvent production and the complex balance among the genes and enzymes involved. While this work is promising and sheds light on the genes, enzymes, and processes of solvent formation, the desired levels of sustainable solvent production for industrial use has not yet been achieved.

1.5.4 Additional Butanol Dehydrogenases and Alcohologenic Growth

Two additional enzymes functioning as butanol dehydrogenases (BDH) have been identified and characterized in the *C. acetobutylicum* genome. (Petersen, et al, 1991; Walter, et al, 1992) The adjacent genes *bdha* and *bdhb*, coding for BDH I and II, are encoded on the chromosome but as two neighboring operons under separate regulatory systems. BDH I and BDH II are NADH- and NADPH-dependent, respectively, and their expression is induced at the onset of solventogenesis just prior to the accumulation of butanol. (Walter, et al, 1992) While BDH I is thought only to fill a minor role as a sink for reducing equivalents, BDH II is believed to have a role in butanol formation, active later than the *sol* locus enzymes. (Durre, et al, 2002) However, studies of inactivation mutants targeting *adhE1* and *bdha/bdhb* (Cooksley, et al, 2012) revealed that loss of AdhE1 results in severe loss of solvent production, but the inactivation mutants of BDH I and BDH II exhibited no significant change in solvent production, suggesting that *adhE1* is the primary solventogenic enzyme responsible for solvent production.

In clostridia, a secondary solvent formation pathway exists, termed alcohologenesis, which occurs under high NADH/NAD+ ratios or upon the addition of

redox dyes like methyl vialogen. (Fontaine, et al, 2002) Alcohologenic growth involves a different set of genes/enzymes than solventogenic growth, and although the pathway has not been fully characterized, the genes are believed to include the secondary *adhE2* (CAP0035), pyruvate decarboxylase (*pdc*), and ethanol dehydrogenase (*edh*). (Lee, et al, 2008) It has been shown that the redox-sensing transcriptional regulator Rex (CA_C2713) in *C. acetobutylicum* maintains NADH homeostasis in the cell as well as regulates expression of select secondary metabolic genes, including *adhE2*, on the basis of cellular redox ratio (NADH/NAD+). (Wietkze and Bahl, 2012; Zhang, et al, 2014) Sequencing and molecular characterization (Fontaine, et al, 2002) have identified *adhE2* as a 2577 bp ORF, present on the megaplasmid but encoded on the complementary strand. Like *adhE1*, it has two TSS, with a distal promoter that is weaker relative to the proximal promoter, and exhibits similarly high identity with other *adhE* domains, including a 66.1% identity with *adhE1* in *C. acetobutylicum*. Although plasmid-based expression of *adhE2* in a pSOL1-deficient strain did partially restore butanol production, *adhE2* was also found to be natively expressed only in alcohologenic cultures; conversely, *adhE1* expression was found only in solventogenic cultures. (Fontaine, et al, 2002) Furthermore, inactivation of *adhE1* in *C. acetobutylicum* (Cooksley, et al, 2012) yielded significant loss in solvent production while similar inactivation of *adhE2* revealed no significant change, in support that *adhE1* functions as the primary solventogenic gene.

1.6 Solvent Toxicity and Mechanisms of Tolerance in *C. acetobutylicum*

Solvent toxicity and tolerance is a major concern in biofuel development, bioremediation, and other applications of bacterial metabolism. The toxicity of metabolites, particularly alcohols and carboxylic acids, commonly limits product titers

and restricts fermentation performance, affecting the economy of these processes. (Jones and Woods, 1986; Papoutsakis, 2008; Nicoloau, et al, 2010)

Rather than a solvent's chemical structure, it is the concentration that accumulates in a cell's membrane that determines its level of toxicity; of the three organic solvents produced by *C. acetobutylicum*, butanol is the most toxic. (Jones and Woods, 1986; Sardessai and Bhosle, 2002) Studies of toxicity and tolerance in *E. coli* and other Gram negative species indicate that butanol and other alcohols intercalate within a cell's phospholipid bilayer, with deleterious effects on the composition and crucial functioning of the membrane. Intercalation of butanol molecules increases membrane fluidity, displaces integral proteins, and disrupts membrane homeostasis and energy production, possibly resulting in cell death. (Papoutsakis, 2008; Nicoloau, et al, 2010) In Gram-positive organisms, butanol stress results in a higher ratio of saturated to unsaturated fatty acids in the membrane, to counteract the higher membrane fluidity and increase membrane order. In *C. acetobutylicum*, butanol also inhibits nutrient uptake (specifically glucose transporters), impairs membrane ATPases, and inhibits ATP synthesis; in so doing, it destroys the cell's ability to maintain a constant internal pH of 6.2, weakening the proton gradient, disrupting electrochemical potential, and reducing energy production. (Bowles and Ellefson, 1985; Borden and Papoutsakis, 2007; Papoutsakis, 2008; Nicoloau, et al, 2010)

Solvent tolerance, therefore, is a multi-genic trait more complex than any single gene and a complex mechanism to investigate and manipulate. Cellular responses to solvent stress include contributions from a range of cellular processes and elements, including induction of stress response mechanisms (particularly heat shock proteins), molecular efflux pumps, metabolic detoxification or neutralization of the

stressor, alteration of cell membrane properties and composition, changes in energy metabolism, and adjustments to transcriptional or translational processes. (Papoutsakis, 2008; Nicoloau, et al, 2010) Molecular efflux pumps are prevalent in Gram-negative stress responses, (Ramos, et al, 2002) but very little is known of their role (if any) in Gram-positive systems. (Nicoloau, et al, 2010) Gram-positive organisms possess only a single cell membrane, but the thicker layer of peptidoglycan present in this membrane may possibly facilitate tolerance. (Sardessai and Bhosle, 2002; Nicoloau, et al, 2010) In response to solvent stress, it has been shown that clostridia up-regulate energy-generating processes, maintaining glucose uptake and repressing protein biosynthesis genes (presumably to redirect ATP toward stress response). (Nicoloau, et al, 2010) Another method of combating stress, metabolic detoxification, is illustrated by the removal of inhibitory acid accumulation by the solventogenic phase of clostridial fermentation; however, manipulation of this method to combat butanol stress in *C. acetobutylicum* would counteract the primary goal of its metabolic engineering. (Nicoloau, et al, 2010)

One of the most prevalent and understood processes relating to solvent tolerance in prokaryotes is the activity of stress response systems. Studies have shown strong links between general stress response proteins, or heat shock proteins (HSP) (also called molecular chaperones) and solvent tolerance, and that changes in tolerance phenotype can be effected by targeting these stress response genes, not membrane function alterations. (Sardessai and Bhosle, 2002; Paredes, et al, 2005; Nicoloau, et al, 2010; Wang, et al, 2013) The benefit imparted by HSP under stress conditions is likely the stabilization of proteins and enzymes that are necessary for cell activity but are also sensitive to solvent stress. (Nicoloau, et al, 2010)

Recent studies have sought to characterize and manipulate general and specific stress response systems, revealing much about stress response and solvent tolerance networks as well as achieving some success in engineering phenotypes of higher tolerance. Microarray-based transcriptional analyses have revealed the general stress response in *C. acetobutylicum* to include increased expression of solvent genes and HSP genes (i.e. *groES*, *dnaKJ*, *hsp18*, and *hsp90*), decreased expression of translation-machinery and pyrimidine synthesis genes, and dose-dependent down-regulation of select lipid biosynthesis genes. (Tomas, et al, 2004; Alsaker, et al, 2010) Overexpression of HSP genes has been shown to increase solvent tolerance as well as solvent titers in *C. acetobutylicum* and several other organisms. The most studied example is the GroESL general stress response proteins. Overexpression of GroESL, as well as their co-overexpression with other HSP proteins (ClpB and GrpE), in *E. coli* imparts significant increases in tolerance toward a range of organic solvents, including ethanol and butanol. (Zingaro and Papoutsakis, 2012, 2013) Overexpression of the *groESL* operon genes in *C. acetobutylicum* reduces growth inhibition by 85% relative to wild type, more than doubles the time of active glucose metabolism, and generates 40% and 33% higher solvent titers than wild type and plasmid control, respectively. (Tomas, et al, 2003, 2004) Also, HSP genes are among those that are up-regulated upon *spo0A* overexpression, a condition that exhibits higher butanol tolerance. (Alsaker, et al, 2004) Finally, expression of the HSP genes under stress has appeared to be dose-dependent, suggesting that *C. acetobutylicum* possesses a system for sensing butanol levels and regulating accordingly. (Tomas, et al, 2004)

Most chemical stresses elicit such general stress response systems, but many are believed to trigger more specialized responses as well, (Nicoloau, et al, 2010) and

several studies, frequently stemming from efforts to characterize general stress response and solvent tolerance, have identified additional genes (Alsaker, et al, 2004; Borden and Papoutsakis, 2007; Jia, et al, 2012; Xu, et al, 2014) as well as transcription factors (Borden and Papoutsakis, 2007; Wang, et al, 2013). These studies are not only adding to our understanding of stress response networks and solvent tolerance but also expanding it to include stressor-specific response systems. Microarray-based transcriptional analysis of butanol stress response indicates the butanol-stress-specific up-regulation of glycerol metabolism genes and (surprisingly) the down-regulation of saturated fatty-acid biosynthesis genes, as well as severely impaired metabolism (i.e. solvent production, glucose uptake, and acid-reassimilation) despite growth to 80% cell density. (Alsaker, et al, 2010) Transcriptional analysis of stress cultures, combined with target predictions for known stress transcription factors (TF), characterized additional genes and TFs in stress response systems, identifying 164 transcriptional regulators differentially expressed under butanol or butyrate stress. Thus, more specific stress response networks, including predicted regulons for these stress-specific TFs, were elucidated. These newly identified stressor-specific elements included transcriptional regulators for heat shock (HrcA and CtsR), SOS response (LexA), redox sensing (Rex), peroxide sensing (PerR), and amino acid and purine metabolism (ArgR, HisR, CymR, and PurR). (Wang, et al, 2013) Collectively, this study shows the beginning of a comprehensive model of the transcriptional network of stress responses in *C. acetobutylicum*. Differential regulation of a variety of other genes, including dose- and time-dependent expression, has been observed under both solvent and acid stresses, thus further illustrating the complexity of stress response. (Alsaker, et al, 2010; Wang, et al, 2013)

1.6.1 Small RNA in Stress Response and Solvent Tolerance

There are other levels of gene regulation aside from transcriptional, with epigenetics, translation, and mRNA stability all offering possible locations for gene regulation. (Wang, et al, 2013) However, a more novel field of study regarding gene regulation lies in non-coding small RNA research (for a detailed review of sRNA regulators and mechanisms, refer to Section 1.7).

Small RNAs involve lighter metabolic costs on the cell compared to protein-based expression, making sRNA regulators a promising and effective engineering strategy, and several studies have demonstrated the ability of sRNA expression to alter solvent tolerance and other cellular processes, including solvent formation. Recent work (Gaida, et al, 2013; Bak, et al, 2014) has demonstrated that co-overexpression of DsrA, RprA, and ArcZ (sRNAs that activate translation of the RpoS stress response sigma factor) in *E. coli* resulted in a supra-additive increase in acid tolerance, plus protection against carboxylic acid and oxidative stress. This observed sRNA-based tolerance phenotype was neither repressible by glucose nor dependent on amino-acid activation or acid induction, as other characterized acid resistance systems typically are. In addition to acid stress, RpoS is also known for its role in responses to other stresses, including starvation, temperature, and hyperosmolarity. (Gaida, et al, 2013) RpoS in *Salmonella* has been shown (Levi-Meyrueis, et al, 2014) to participate in a regulatory stress response network that includes several sRNA and affects many metabolic genes and processes as a competitive advantage in stationary phase.

Several regulatory sRNA have been shown in other fermentative bacteria to be associated with metabolism and solvent stress response, such as ethanol production in the facultative anaerobe *Zymomonas mobilis*. (Cho, et al, 2014) In *C. acetobutylicum*, comparative genomic analysis using predictions of promoters and rho-independent

terminators identified 113 sRNA (101 on the chromosome and 12 on pSOL1), including 32 that match previously predicted RNA sequences. An additional 46 novel sRNAs have since been identified in *C. acetobutylicum* (Venkataramanan, et al, 2013), and of the 159 total sRNA sequences, over 60% were overexpressed under butanol and butyrate stress; this pattern suggests strong involvement of clostridial sRNA in metabolite stress response. Specifically, many of sRNA exhibited stress and dose dependent differential expression, and several others showed consistent expression patterns regardless of stress, pointing to roles in both specific and general stress responses.

Some of the identified sRNA sequences, by virtue of previous characterization in other organisms, can be related to stress response roles in *C. acetobutylicum*, including 6S RNA (sCAC1377) and tmRNA (sCAC834). 6S RNA is known to negatively regulate genes under the control of exponential phase sigma factors (i.e. σ70 or σA) at the onset of solventogenesis as well as under stress conditions, and transcriptional analysis confirms its stress-related expression. Functioning in complex with three proteins to rescue stalled ribosomes in the trans-translation quality control system, tmRNA is also up-regulated under stress and is among the most highly up-regulated sRNA sequences reported, suggesting a role in stress response. Notably, its partner proteins were not up-regulated, suggesting the function of regulating trans-translation under stress conditions falls on the tmRNA sequence. This was also the first experimental evidence of tmRNA in *C. acetobutylicum*. Both 6S RNA and tmRNA, and their effects on solvent tolerance in *C. acetobutylicum*, are investigated more closely in this thesis and are described in more detail in Chapter 3.

1.7 Small RNA Regulation in Prokaryotes

Small RNA (sRNA) are a class of single-stranded, non-coding RNA who have garnered attention and study in the last 30 years for their extensive and essential regulatory roles in numerous pathways. Prokaryotic sRNA regulators range from 50 to 500 base pairs in size and can exhibit secondary structures, including stem loops and central bubbles, which can be critical to the functional interaction between the sRNA and its target. Many sRNA regulators also require a chaperone for effective function. Regulatory sRNA are triggered by stimuli like anaerobiosis, oxidative and pH stress, nutrient deprivation, iron availability, and osmotic imbalance. (Beisel and Storz, 2010) Small RNA are prevalent throughout the microbial world in Gram-negative and Gram-positive organisms alike, (Georg and Hess, 2011) and regulatory sRNA influence gene expression positively and negatively in a wide range of cellular processes, from pathogenicity to biofilm formation. (Chambers and Sauer, 2013; Ghaz-Jahanian, et al, 2013; Majdalani, et al, 2005; White, et al, 2010) Within these processes, sRNA do not always play the central role, flipping the proverbial 'master switch;' rather, many sRNA regulators serve as steps in the process, linking a stimulus or response to a particular pathway and enabling more finely tuned regulation. Regulatory networks are significantly interconnected, and although regulatory sRNA represent a group of regulators whose complete role is not yet fully understood, their participation as recipients, transducers, and originators of regulatory action has become apparent. (Mandin and Guillier, 2013; Beisel and Storz, 2010; Gopel and Gorke, 2012) Their particular mechanisms of regulation, while present across species and shared among all cellular processes, are complex and varied. Regulators repress transcript expression by blocking ribosomal access or by initiating total degradation of the transcript, and alternatively they may stabilize transcripts or activate translation, typically by utilizing

the same regions used for repression in other systems. RNA regulators may be trans-acting or antisense, part of any number of regulatory network motifs, or even contain coding for a protein regulator.

Small RNA regulation is similar to protein regulation, with both types of regulators requiring specific binding sequences to function, but in weighing direct regulation by both methods, sRNA are usually metabolically cheaper than protein regulators, as they involve shorter transcription and no translation. (Beisel and Storz, 2010; Gaida, et al, 2013) Regulation by sRNA also offers greater speed and flexibility, with the ability to rapidly alter metabolic pathways in response to changing environmental conditions and stimuli. Also, sRNA regulation can be reversible and operates with less noise and cell-to-cell variability than protein-based transcriptional repression. (Beisel and Storz, 2010) Advances in understanding sRNA regulation, particularly in metabolism, is critical to understanding and manipulating prokaryotes in healthcare, bioengineering of alternative energy and bioremediation strategies, and other fields of study.

1.7.1 Types of sRNA Regulators

Regulatory sRNA can be divided into two main classes: those that modulate protein activity, and those that interact with mRNA transcripts, or base-pairing sRNA. The smaller class of sRNA – those that interact with protein targets – usually regulates enzymatic activity by competing with the proteins' mRNA targets by binding to and sequestering the proteins to allow mRNA translation. (Chambers and Sauer, 2013; Majdalani, et al, 2005; Pichon and Feldon, 2007; Storz, et al, 2011) Modulation of enzymatic activity by sRNA, along with transcription-level regulation, is perhaps best

exemplified by 6S RNA, a sRNA that interacts with RNA polymerase through promoter competition and is explained in detail in Chapter 3.

The most prevalent variety of sRNA, representing at least one-third of all characterized sRNA, are the base-pairing regulators. (Majdalani, et al, 2005) They include trans-acting sRNA, encoded on genomic regions distant to their target, and cis-acting or antisense sRNA (asRNA), which are encoded on the complementary strand directly opposite their targets. Both trans-acting and anti-sense sRNA regulators base-pair stoichiometrically with their targets by hydrogen bonding. (Georg and Hess, 2011) Trans-acting sRNA are characterized by short, imperfectly complementary regions of seven to 10 base pairs each, while asRNA exhibit more complete and extended pairing. Anti-sense sRNA can exist at barely detectable levels or levels comparable to that of mRNA transcripts, but are as prevalent in bacteria as they are in Eukarya and Archaea. (Georg and Hess, 2011) Most base-pairing sRNA are non-coding, but a few base-pairing and protein-coding sRNA regulators have been identified in recent years. (Vanderpool, et al, 2011)

All base-pairing sRNA regulators either activate or repress a gene by targeting its mRNA transcript through a multi-step, often structure-dependent mechanism. (Storz, et al, 2011) Typically, sRNA possess a short, single-stranded seed-pairing region of a few nucleotides (6-7 bp) and found on its 5' end that stimulates the initial pairing between the sRNA and its mRNA target. This initial interaction then leads to further base-pairing to complete the regulatory sRNA-mRNA duplex.

This base-pairing continuation is mediated for many trans-acting sRNA by the RNA chaperone protein host factor Qβ (Hfq), which plays a vital role in the activity of many sRNA. (Desnoyers, et al, 2013) Hfq binds to both sRNA and mRNA at mutual

AU-rich regions to stabilize their interaction, especially in cases of noncontiguous binding. (Beisel and Storz, 2010; Majdalani, et al, 2005; Pichon and Feldon, 2007) A poly-U sequence, protecting the sRNA from exonucleolytic degradation, is found on its 3' end and has also been shown in some instances to bind Hfq. Concentrations of sRNA, mRNA, and Hfq, as well as their association and dissociation rates, are thought to contribute to regulatory activity, but in ways not yet fully understood. Most trans-acting sRNA in Gram-negative bacteria require Hfq for effective binding. (Caron, et al, 2010) Hfq homologs have been observed in several Gram-positive organisms as well, including Firmicutes like *Listeria monocytogenes* and *C. acetobutylicum*; nevertheless, Hfq seems to play a less important role in Gram-positive species than in Gram-negative ones. (Storz, et al, 2011) This may result from its need for distinct AU-rich sequences and the characteristic AT-rich genomes of Gram-positive organisms. (Desnoyers, et al, 2013)

1.7.2 Mechanisms of RNA Regulation by Base-Pairing sRNA

1.7.2.1 Translation Repression and Transcript Degradation

Translation of an mRNA can be inhibited by the interaction of another RNA molecule with the ribosomal binding site (RBS) of the transcript, and many base-pairing sRNA regulate in this manner, binding the RBS or other nearby regions, like translational enhancer elements present in the 5' UTR, to block translation initiation. (Beisel and Storz, 2010; Desnoyers, et al, 2013; Storz, et al, 2011) Although simple translational repression of a target gene occurs for many base-pairing sRNA, like Spot-42, OmpAB, and GcvB, most such gene down-regulation is paired with the degradation of the transcript by a ribonuclease like RNase E. (Desnoyers, et al, 2013)

RNA degradation enables rapid and more permanent repression of gene expression. This coupling is sufficiently common that mRNA degradation is utilized in the identification of sRNA regulatory targets. The specifics of sRNA-bound mRNA degradation remain largely unknown, but two models have been developed: passive degradation, due to the lack of steric protection against nucleases provided by a bound ribosome, and active degradation, resulting from the direct activity of interacting sRNA, Hfq, and RNase E. (Lalaouna, et al, 2013) Regulatory sRNA may bind to the transcript's RBS or to other distal regions, including its coding sequence, to stimulate degradation. Studies identifying rate-limiting cleavage sites at the RBS and at these distal sites suggest that the sRNA-Hfq-RNase E complex actively stimulates degradation, regardless of cleavage site, as part of translational repression, and in other cases, transcript degradation, rather than a necessary mechanism of repression, is simply a removal of inhibited mRNA. (Beisel and Storz, 2010; Desnoyers, et al, 2013) This variety in regulatory action, seen in some individual cases like RyhB or SgrS regulation in *E. coli*, (Caron, et al, 2010; Rice, et al, 2012) suggests that mRNA transcripts contain the necessary 'instructions' for their specific and complex regulation.

1.7.2.2 Translational Activation and Transcript Stability

Some regulatory sRNA increase a transcript's stability by blocking cleavage sites or by facilitating cleavage of polycistronic transcripts to yield new end structures that increase stability or improve translation. (Storz, et al, 2011) As with their mRNA targets, the stability of sRNA regulators themselves sometimes relies on their interaction with their targets. Some sRNA are degraded when in a duplex, while others are stable in a duplex and degraded when unbound. This dynamic highlights the

specificity of ribonucleases for particular RNA molecules and/or duplexes and its importance to regulation. (Storz, et al, 2011) Some sRNA activate gene expression, and like sRNA repressors, these positive regulators commonly target a transcript's RBS, binding to hairpin loops that block the RBS to release the loop and expose the RBS for translation. (Beisel and Storz, 2010) Primary examples of these translational activators include the sRNA that stimulate RpoS, the stationary-phase stress response sigma factor; RprA, ArcZ, and DsrA are sRNA regulators of RpoS that allow for different environmental conditions and stresses to efficiently affect a single target – an important aspect of stress response and adjustment of metabolism and a key benefit of sRNA-based regulation. (Chambers and Sauer, 2013; Majdalani, et al, 2005) Although it is very uncommon, some sRNA regulators, like RyhB, RprA, and DsrA, can repress and activate different targets. (Beisel and Storz, 2010; Caron, et al, 2010)

1.7.3 Small RNA Regulators in Gram-Positive Organisms

Examples of nearly every variety of sRNA regulator have been identified in Gram-negative and Gram-positive organisms alike, from base-pairing to protein-modulating and from repression and degradation to activation, but there are comparatively fewer characterized examples in Gram-positive organisms. (Rice, et al, 2012; Vanderpool, et al, 2011) This fact might indicate a preference, essentiality, or greater complexity of sRNA regulation in Gram-negative organisms that is absent in Gram-positive species, or it may be possible that equivalent instances of sRNA regulation simply have not yet been documented in Gram-positive organisms. More extensive study of sRNA in Gram-positive organisms must be conducted before more concrete comparisons and conclusions can be made. Understanding sRNA can help fill in the gaps, answer the questions, and provide the finer details of complicated

processes in prokaryotic molecular biology. This more accurate perspective will facilitate continuing advancements in medical treatment, pharmaceutical development, bioengineering, and other biotechnological pursuits. Continued progress in manipulating microbial metabolism in these fields will become increasingly dependent on studying the variety and complexity of their regulation by sRNA.

1.8 Previous Work on Metabolic Regulation in *C. acetobutylicum*

Twenty years ago, a 957-bp coding sequence, subsequently named *solR*, was identified on the pSOL1 megaplasmid, located 643 bp upstream of the *adhE1* gene of the *sol* operon, with two inverted repeat stem loop sequences downstream. (Nair, et al, 1994) A single TSS for *solR* was identified 35 bp upstream of the start codon, with a putative clostridial promoter (TCGATA-17bp-TATTAT) and ribosomal binding site found seven bp and 11 bp, respectively, upstream. (Fischer, et al, 1993; Nair, et al, 1999) Northern analysis (Fischer, et al, 1993; Nair, et al, 1999) also observed two transcript products for *solR* (1.1 kb and 1.3kb), suggesting that transcription occurs through both of the identified stem loop terminator sequences. The predicted 36.9 kDa SolR protein exhibits a predicted helix-turn-helix (HTH) DNA-binding motif, with homology searches showing some similarity with the HTH-motifs of several common bacterial DNA-binding proteins. (Nair, et al, 1999) Northern analysis of *solR* (Fischer, et al, 1993; Nair, et al, 1999) showed that *solR* expression is low in wild type but detected in exponential and stationary phase growth; also, consistently minimal *solR* expression under *spo0A* inactivation and overexpression alike, combined with no identified 0A box for *solR*, suggests no interaction between Spo0A and *solR*. (Harris, et al, 2002)

Overexpression of *solR* in *C. acetobutylicum* ATCC 824 resulted in severely attenuated solvent production, and no observable expression of the *adhE1* or *adc* genes indicated the likely mechanism of this phenotype. (Nair, et al, 1999) Additionally, disruption of *solR* by genomic integration of the non-replicative suicide-plasmid pO1x displayed three- to four-fold increases in butanol, acetone, and ethanol titers, earlier induction of AdhE1 activity, a longer solventogenesis phase, and more biomass than control. (Nair, et al, 1999; Harris, et al, 2001) Therefore, SolR was characterized as a putative regulator of solventogenesis (i.e. the *sol* locus genes). Nevertheless, despite earlier predictions, no DNA targets were successfully predicted for interaction (Nair, et al, 1994; Nair, et al, 1999); furthermore, additional studies (Thormann and Durre, 2001) were unable to observe any DNA-binding activity of purified SolR to the *adc* promoter or the distal promoter region of the *sol* operon, and subsequent homology searches revealed similarity of SolR to O-linked GlcNAc transferases from a variety of eukarya and archaea. (Thormann and Durre, 2001) In support of these findings, overexpression of *solR* revealed a sharp reduction in glycosylated exoprotein levels, and further experimentation demonstrated the localization of SolR to the cellular membrane, likely as an extracellular protein (Thormann and Durre, 2001) Overall, these results suggested that SolR is an extracellular membrane protein involved in protein glycosylation patterns rather than as a repressor of solventogenesis. However, the effects of the *solR* recombinant strains needed to be explained in order to eliminate the possibility of SolR as a solventogenic regulator. It was discovered that the sequence used in the *solR* overexpression strain (Nair, et al, 1999) also expressed a portion of the intergenic region downstream of *solR*, and subsequent overexpression of 430 bp of this region between *solR* and *adhE1*

resulted in the same solvent-negative phenotype to suggest that a regulatory sequence in this intergenic region was responsible for the phenotype and therefore the regulation of solventogenesis. (Thormann, et al, 2002) Also, complementation of the *solR* inactivation mutants with complete as well as truncated *solR* sequences (Scotcher, et al, 2003) reduced solvent production to wild type levels, further suggesting that the complete *solR* sequence has no impact on solvent production. However, the higher solvent titers of these *solR* inactivation strains remained unknown.

Other studies have focused on transcriptional activation by Spo0A, and possibly another activator, as the regulatory mechanism controlling solventogenesis. Spo0A had been previously identified as a transcription factor affecting the *adc* and *sol* operons (Ravagnani, et al, 2000; Harris, et al, 2002), and additional studies found that an 0A box and two nearby imperfect repeats (R1 and R3), found in the intergenic region upstream of *adhE1*, are essential for successful transcription of the *sol* operon. (Thormann, et al, 2002; Scotcher, et al, 2003) However, inactivation of *spo0A* does not totally eliminate expression of the *sol* locus genes, and its overexpression shows earlier induction of the *sol* operon with unequal patterns of subsequent down-regulation (Harris, et al, 2002); therefore, the possibility exists that Spo0A is one of many regulators of solvent formation in *C. acetobutylicum* and that the *sol* locus genes may be differentially regulated. While the regulatory role of SolR may be in doubt, its influence remains incompletely understood (especially considering the solvent titers of the inactivation mutants), and although transcriptional regulation involving Spo0A may be likely, much remains unknown about the regulation of solventogenesis.

Over the last five years, a sRNA sequence has been discussed as a potential post-transcriptional regulator of the solventogenic gene transcripts. A recent doctoral

dissertation (Zimmermann, 2013) presented a characterization of *solB*, a 196-bp sRNA located between *solR* and *adhE1*, showing by primer extension and Northern analysis a transcriptional start site and a predicted promoter matching the clostridial consensus (Figure 1.3). A putative secondary structure for *solB* has been identified, and a binding mechanism utilizing a hairpin loop sequence has been predicted, and although this sequence exhibits complementation with the UTRs of the *spo0A*, *adhE2*, and *bdhB* transcripts, (Schiel, et al, 2010; Durre, 2012) confirmation of these binding interactions was unsuccessful. (Zimmermann, 2013, [Dissertation]) Down-regulation of *solB* under butyrate stress and up-regulation under butanol stress has been shown and is opposite to the pattern observed for *sol* operon expression under the same conditions (Venkataramanan, et al, 2013); this supports the hypothesis that *solB* is a repressor of solventogenesis and suggests that it could be part of a specific or specialized stress response in *C. acetobutylicum*.

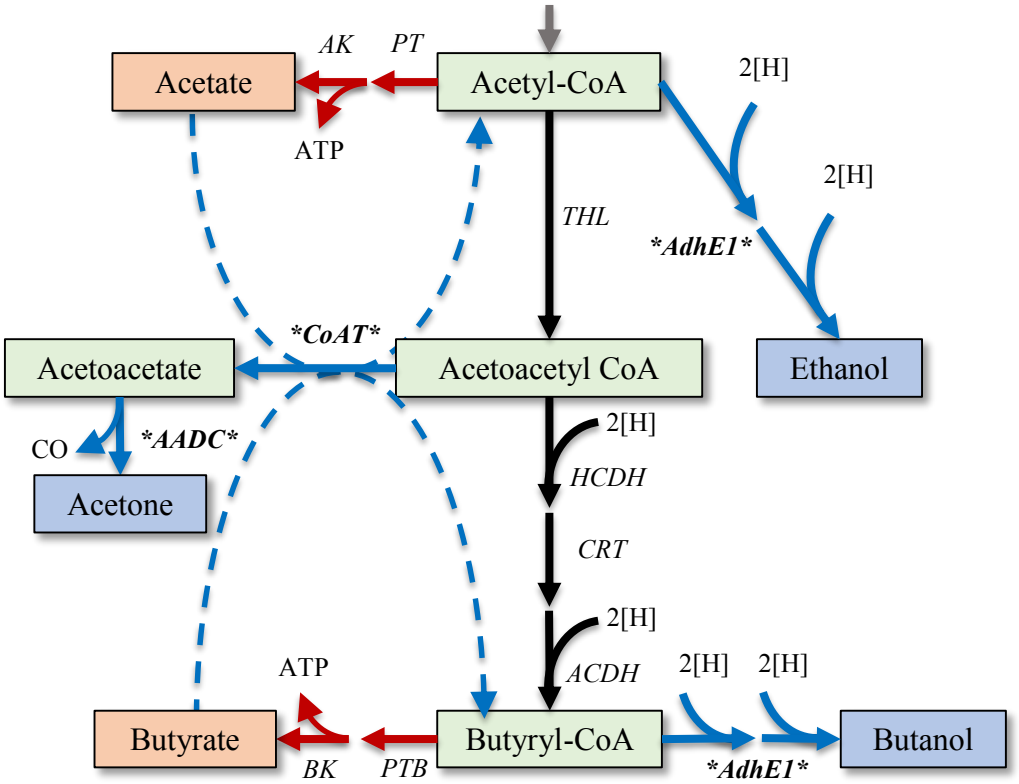


Figure 1.1 – General flowchart of major ABE fermentation intermediates, products, and enzymes, with the solventogenic enzymes in bold and marked with asterisks.

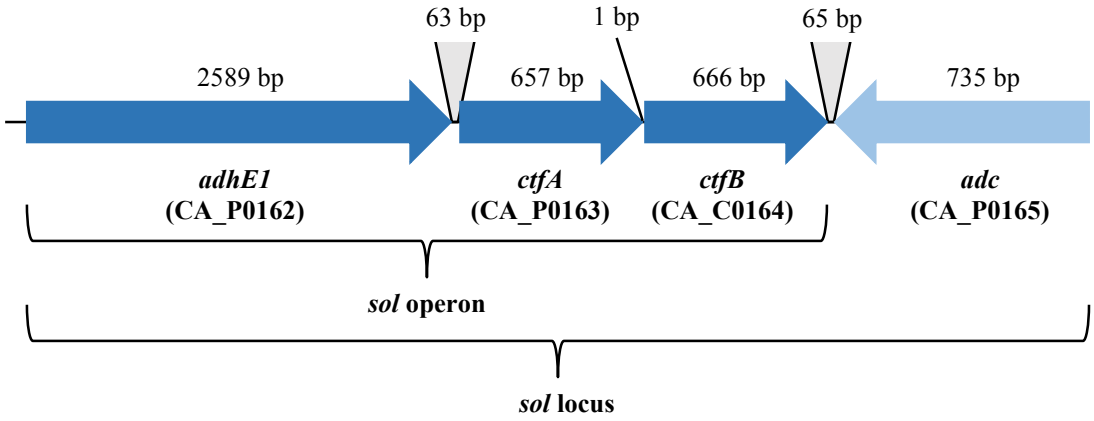


Figure 1.2 – Genomic location and organization of the *sol* locus genes in *C. acetobutylicum*. Shown are coordinates 175,810 to 180,582 of the pSOL1 megaplasmid, with the view aligned to the coding strand.

ATTAAGATATAGCTTCTTTTATGTAGTATTATTTCAGAAGTCTACAAATTAAGTTTATA

TTTAGACCCTGGGGTGTAACTATAGTATTTAATATTGGTACTATTAATTAGGGTTATAT

ATACTAGAACTTATCATGGTAAACATAAATATAAACTCAATTCTATTTATGCTCCTATA

AAATTTTATAATATAGGA**AAACTGCTAAATGTAAA**TTATACG**TTTACATTTAGCAGTTT**

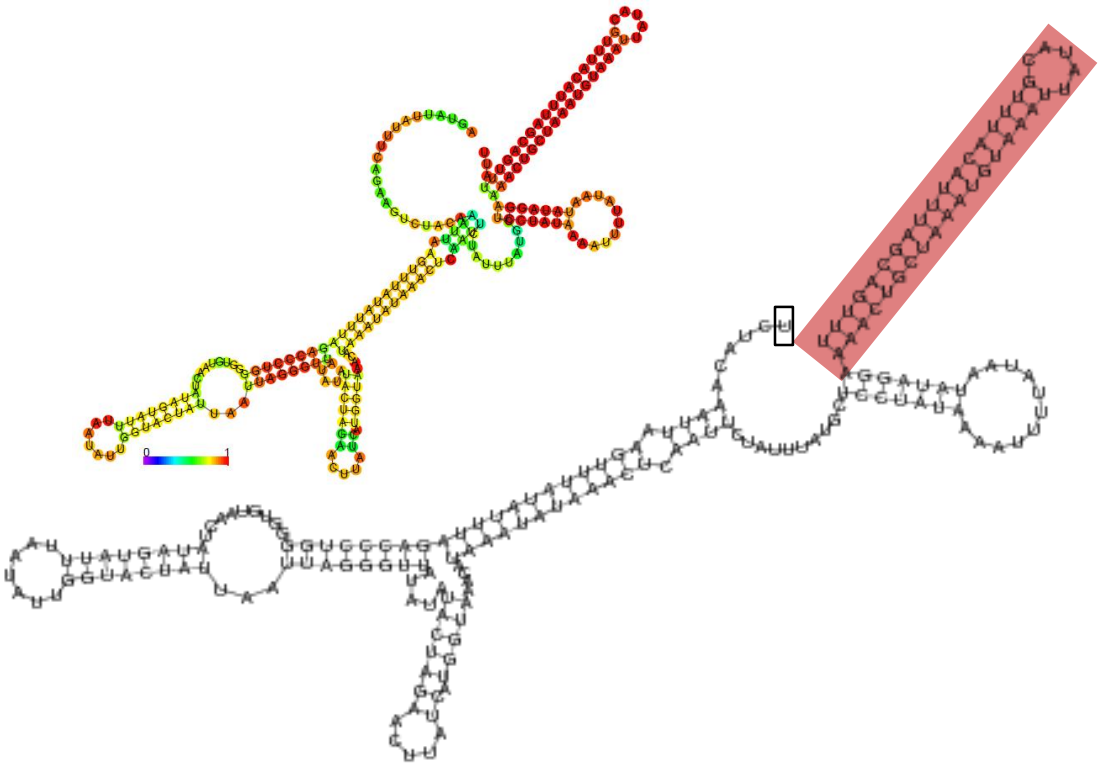


Figure 1.3 – Sequencing and secondary structure of *solB* (sCA_P176). Predicted promoter sequence is underlined, with -35 and -10 regions shown with double lines, transcriptional start site is boxed, and the terminal stem loop highlighted in red; inset image indicates base-pairing probabilities (blue<yellow<red); predicted ΔG = -51.5 kcal/mol.

Chapter 2

EXPERIMENTAL MATERIALS AND METHODS

2.1 Bacterial Strains and Plasmids

Relevant characteristics and sources of all bacterial strains and plasmids used in these studies are listed in Table 2.1.

2.2 Culture Conditions

E. coli strains were grown aerobically at 37°C and 220 rpm in liquid or solid LB media, supplemented with ampicillin (Amp) (50 μg/mL), kanamycin (Kan) (25 μg/L), or chloramphenicol (Cm) (35 μg/mL) when necessary.

Colonies of *C. acetobutylicum* were grown anaerobically at 37°C on solid 2xYTG media (pH 5.8) for at least five days before being transferred to 10 mL of liquid CGM. Recombinant strains were grown similarly but in media supplemented with erythromycin (Em) (40 μg/mL in solid media, 50 μg/mL in liquid media) or thiamphenicol (Th) (5 μg/mL) when necessary. To induce expression of the lactose-inducible promoter of the pKO_mazF vector, 40 mM of β-lactose (Sigma-Aldrich) was added to the culture media.

2.3 Construction of Overexpression and Anti-sense Vectors

All primers and their sequences utilized in these studies are presented in Table 2.2.

The overexpression plasmids for each small RNA, as well as the plasmids for anti-sense knockdown of *solB*, were generated by first amplifying the small RNA coding sequences by standard PCR and subsequently appending the 5'-end of each

with a BamHI restriction site and the 3'-end with a KasI restriction site using overhangs on PCR primers consisting of the restriction enzyme recognition sequence followed by six additional purine nucleotides to ensure efficient cleavage by the enzyme. To generate the anti-sense *solB* plasmid, the restriction sites were appended to the opposite ends of the PCR product (i.e. BamHI on the 3' end and KasI on the 5' end), allowing it to ligate in reverse orientation into the expression vector. All PCR amplification products were confirmed by agarose gel electrophoresis and then double digested with BamHI and KasI (NEB). The p94MCS expression vector was similarly double-digested with BamHI and KasI and dephosphorylated with Antarctic phosphatase (NEB), and each digested PCR product was individually ligated into the p94MCS vector using the Instant Sticky-End Master Mix (NEB). Each plasmid product was transformed into chemically competent Turbo or 10-beta *E. coli* cells (both NEB) and harvested (Spin Miniprep Kit, Qiagen). Successful construction of the resulting overexpression plasmids was confirmed by one-hour restriction digest with either SalI (NEB) (p94_solB, p94_6S, p94_6Sb) or BsgI (p94_tm) and agarose gel electrophoresis as well as by Sanger sequencing analysis. Each plasmid was then transformed by electroporation into BL21(pAN3) competent cells for methylation (Mermelstein and Papoutsakis, 1993), harvested, and confirmed by additional restriction digest analysis as above. The methylated plasmids were transformed by electroporation into *Clostridium acetobutylicum* ATCC 824 electrocompetent cells according to previously described protocol (Mermelstein, et al, 1992) and plated on 2xYTG solid media supplemented with erythromycin (40 μg/mL). Colony transformants were selected for outgrowth in supplemented liquid CGM media and used to generate duplicate frozen stocks for each strain (in 15-20% glycerol in CGM at

-80°C) as well as to extract plasmid samples for validation by PCR amplification of the multiple cloning site (MCS) and appropriate insert (Figure 2.1).

2.4 Construction of the *solB* Knockout Plasmid (pKO_mazF::solB)

The pKO_mazF::solB plasmid was constructed utilizing the procedure developed and described previously (Al-Hinai, et al, 2012). Two regions of homology (RH) were amplified by standard PCR from genomic DNA of *C. acetobutylicum* ATCC 824. The 936-bp upstream RH corresponds to nucleotides -986 to -50 relative to the transcriptional start site of *solB*, and the 901-bp downstream RH corresponds to nucleotides +21 to +901 relative to the end of the *solB* sequence. These PCR products were confirmed by agarose gel electrophoresis, and restriction enzyme recognition sites were added to both ends of both RHs as described above; the upstream RH was appended with an XmaI site at its 5' end and a NotI site on its 3' end, and the downstream RH was appended with recognition sites for AvaII and SphI at its 5' and 3' ends, respectively. These appended products were again confirmed by agarose gel electrophoresis and were then double-digested with the appropriate restriction enzymes (NEB), and the empty pKO_mazF vector was double-digested with AvaII and SphI and dephosphorylated. The downstream RH was ligated into the digested pKO_mazF vector, transformed into chemically competent Turbo *E. coli* cells (NEB), and harvested as described above. Resulting vector samples were screened by digestion with EcoRI (NEB) and agarose gel electrophoresis to confirm successful insertion of the downstream RH. Successful constructions were then doubled-digested with XmaI and NotI enzymes and dephosphorylated, and the above ligation and transformation was repeated for the upstream RH. Transformants were screened as above for successful ligation, and confirmed plasmid constructs were transformed into

BL21(pAN3) cells for methylation as described above. The harvested methylated vector was confirmed by Sanger sequencing analysis of the ligation sites, and the validated pKO_mazF::solB vector was then transformed into *C. acetobutylicum* ATCC 824 electrocompetent cells as above. After three days of growth on 2xYTG solid media supplemented with thiamphenicol (5 μg/mL), triplicate colonies were selected for growth in liquid 2xYTG media and subsequent vegetative transfers with fresh media supplemented with thiamphenicol (5 μg/mL), as described previously (Al-Hinai, 2012), to facilitate the integration of the Th^R cassette. The final culture transfers were grown with an additional supplement of 40 mM β-lactose to cure the cells of the pKO_mazF vector and then were serially diluted and grown on 2xYTG solid media supplemented with thiamphenicol (5 μg/mL) and lactose (40 mM) to complete the curing of remaining vector backbone. Colonies were selected for generation of glycerol frozen stocks and genomic DNA samples, which were screened by PCR amplification for both crossover integration sites and the solB sequence (Figure 2.3). PCR products of the integration sites were further validated by adenylation (Taq) and ligation into the pCR4-TOPO-TA sequencing vector (Invitrogen), and submission for Sanger sequencing analysis (Figure 3.2). Frozen stocks of validated cultures were grown on 2xYTG solid media supplemented with thiamphenicol (5 μg/mL) and lactose (40 mM).

2.5 Butanol Tolerance Assay

Colonies from at least 5 days of growth were selected to inoculate starter cultures in static conical tubes of 10 mL CGM, supplemented with erythromycin (50 μg/mL) when appropriate, and grown anaerobically at 37°C until turbid (an $A_{600} \geq 0.5$). A 50-mL preculture tube of CGM, supplemented with erythromycin (50 μg/mL)

when appropriate, was inoculated with the full 10 mL starter culture, creating a 1:6 dilution; this preculture was grown anaerobically at 37°C to an A_{600} of 1.0 ± 0.1 (mid-exponential phase), at which point 9.8 mL of the preculture was aliquoted into each of four individual 15-mL conical tubes, each containing appropriate volumes of butanol and CGM to generate final concentrations of 0%, 1% (109.3 mM), 1.5% (163.9 mM), and 2% (218.6 mM) (v/v) butanol. Using the remaining preculture, duplicate serial dilutions were prepared with non-supplemented CGM media, and the lowest dilution (10^{-5}) was plated on 2xYTG solid media, supplemented with erythromycin (5 µg/mL) when appropriate. Also, one mL samples from the remaining preculture were collected for HPLC analysis of metabolite concentrations. At 24 and 48 hours of static growth, triplicate serial dilutions (10^{-1} through 10^{-5}) were prepared from each butanol challenge tube using non-supplemented CGM media, and three dilutions of each series replicate were plated on 2xYTG solid media, supplemented with erythromycin (50 µg/mL) when appropriate, and incubated anaerobically at 37°C for 24-36 hours before recording counts of colony forming units (CFU). The A_{600} of each tube was also measured by spectrophotometer at 24 and 48 hours, and one-mL samples were collected at both time points for HPLC analysis. Note: all challenge tubes were vigorously mixed by thorough vortexing prior to sampling for serial dilutions, HPLC samples, and absorbance measurements. The percent survival of each challenge level was determined by calculating the concentration of the preculture and each challenge tube using the CFU counts and dilution factors of each plate and dividing the concentration of each challenge tube by the concentration of the preculture at the time of aliquoting. All tubes and pipettes were allowed to deoxygenate in the anaerobic

chamber for at least 1 hour, all plates deoxygenated for 3.5 to 4 hours, and liquid media deoxygenated for at least 36 hours.

2.6 Culture Growth Conditions for RNA Extractions

Starter cultures were inoculated in 10 mL CGM conical tubes, supplemented with erythromycin (50 μg/mL) or thiamphenicol (5 μg/mL) when appropriate, from colonies at least 5 days old and grown anaerobically at 37°C until turbid (an $A_{600} \geq 0.5$). A 150-mL static culture bottle of CGM supplemented with erythromycin (50 μg/mL) or thiamphenicol (5 μg/mL), when appropriate, was inoculated with the full 10-mL starter culture, creating a 1:16 dilution; the A_{600} of this culture was immediately recorded and one-mL samples collected for HPLC analysis (this represents the t_0 time point). With the A_{600} monitored and recorded regularly, additional one-mL samples for HPLC analysis as well as 15-mL culture samples were collected at 6, 12, and 24 hours (corresponding to the mid-exponential, transition, and stationary phases of growth). The 15-mL samples were immediately centrifuged for 10 minutes at 8,000xg, the supernatant was decanted, and the resultant cell pellets were stored at -80°C for up to 10 days.

2.7 RNA Extraction

Total RNA, including small RNA, was extracted and purified using the Qiagen miRNeasy Mini Kit, as described (Jones, et al, 2008), but with the variations that samples were not split and diluted after Trizol addition and that Buffer RWT (optimized for miRNA and sRNA retention) was used in place of Buffer RW1. After RNA isolation, genomic DNA contamination was removed from all samples by treatment with the TURBO DNase Kit (Ambion), which was removed either by

manufacturer-supplied inactivation reagent or by an additional phenol-chloroform extraction and miRNeasy isolation. Both methods of DNase removal produced RNA samples of equal purity and comparable concentrations. Finally, RNA samples were further purified by ethanol precipitation at -20°C before storage at -80°C.

2.8 Two-Step Quantitative Real-Time Polymerase Chain Reaction (QPCR)

Generation of cDNA was performed with 2 μg of RNA template and the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) and stored at -20°C. All primer sets used for quantitative Real-Time PCR (QPCR) are listed in Table 2.2, with each primer set first validated for efficiency on a test plate. Each cDNA sample was tested in triplicate using iTaq Universal SYBR Green Supermix (BioRad). All cDNA samples intended for direct comparison were run on the same plate using the same master mix preparation to minimize variation in expression results. Fold changes in target expression were calculated relative to plasmid control for all overexpression strains and relative to wild type ATCC 824 for the remaining strains. Expression of target transcripts was normalized to expression of the clostridial housekeeping gene CA_C3571, and fold changes were calculated following the delta-delta-Ct (ΔΔCt) quantification method (Schmittgen, et al, 2008). Primer efficiencies were assumed be to 100% for these calculations because their calculated efficiencies fell within the accepted range for assumption of 100% efficiency (Schmittgen, et al, 2008).

2.9 Northern Blot Analysis

Northern blotting was performed as described previously (Venkataramanan, et al, 2013). Five or 15 μg of total RNA samples, thawed on ice from storage at -80°C, were mixed with 2x TBE urea sample buffer, denatured at 65°C for 15 minutes, and

immediately loaded into a pre-run 5% precast polyacrylamide Ready Gel TBE-urea (BioRad, Hercules, CA) and resolved in a BioRad Mini PROTEAN Tetra Cell apparatus at 170 V for 45 or 65 minutes. Gels were then stained with ethidium bromide and photographed with an AlphaImager HP UV-Vis camera (AlphaInnotech). Blotting of RNA samples onto positively-charged nylon membrane (Ambion) was performed at 400 mA for 90 minutes at 4°C in the transfer apparatus. Upon completion of the RNA transfer, the membrane was dried for two hours at 80°C before storage at room temperature.

Sequences and relevant characteristics of the single-stranded oligo DNA probes used for hybridization can be found in Table 2.3. Radioactive probes were generated using a phosphatase-minus T4 polynucleotide kinase (Affymetrix) to label the oligonucleotide with [γ-^{32}P]ATP. Excess ATP was removed with NucAway spin columns (Ambion), and the radioactive oligonucleotide probe was mixed with prewarmed ultrasensitive hybridization buffer (Ambion) and incubated with the membrane for 16 to 22 hours at 42°C under gentle agitation in a hybridization incubator (Fisher Scientific). The hybridized membrane was then washed twice with 2x SSC, 0.1% SDS and once with 0.1x SSC, 0.1% SDS, all for 15 minutes at 42°C under gentle agitation. The membrane was then exposed to a storage phosphor screen for three to 48 hours in an exposure cassette (Molecular Dynamics), and the final image was collected using a phosphor imager (Typhoon 9400 Variable Mode Imager, GE Healthcare).

2.10 Polymerase Chain Reaction (PCR)

Standard PCR reactions utilized the Q5 Hot Start High-Fidelity DNA Polymerase (NEB), 200 μM dNTPs, 0.5 μM of each primer, and 50-300 ng of DNA

template. Reactions were run in a Bio Rad C1000 Touch Thermal Cycler; QPCR included the addition of the CFX96 Real-Time System (BioRad). Standard PCR reaction conditions consisted of initial denaturation (98°C for 30 seconds) followed by 30 cycles consisting of 10 seconds of denaturation at 98°C, 45 seconds for annealing at a temperature specific to the primer set (typically three degrees greater than the lower of the two primers' annealing temperatures), and extension at 72°C for a variable time dependent on the size of the product, with a final extension step of five minutes at 72°C.

2.11 Characterization of Recombinant Strain Growth

Starter cultures were inoculated in 10 mL CGM conical tubes, supplemented with erythromycin (50 μg/mL) or thiamphenicol (5 μg/mL) when appropriate, from colonies at least five days old from solid media, and grown anaerobically at 37°C until turbid (an $A_{600} \geq 0.6$). A 30-mL static culture bottle of CGM supplemented with erythromycin (50 μg/mL) or thiamphenicol (5 μg/mL), when appropriate, was inoculated with the full 10-mL starter culture, creating a 1:4 dilution; the A_{600} of this culture was immediately recorded and one-mL sample collected for HPLC analysis (this represents the t_0 time point). With the A_{600} monitored and recorded regularly, additional one-mL samples for HPLC analysis were collected at every two to three hours for the first 24 hours of growth, followed by sampling every six or 12 hours through 72 hours of growth. Cell density and metabolite concentrations were used to generate growth curves for each recombinant strain as well as ATCC 824 and vector control.

2.12 Confirmation of Recombinant Strains by Sequencing

A 10 mL static culture was grown of the solRB or 824ΔsolB strain under erythromycin (50 µg/mL) or thiamphenicol (5 µg/mL) stress, respectively, until turbid, then genomic DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen). A region of the genomic DNA was amplified by standard PCR, and amplification was confirmed by agarose gel electrophoresis. The confirmed product was purified using a PCR Purification Kit (Qiagen) and adenylated using GoTaq Hot Start Polymerase (Promega) at 72°C for 20 minutes. The PCR product was then ligated into the pCR4-TOPO-TA sequencing vector (Life Technologies) via incubation at room temperature for 30 minutes. The resulting sample was placed on ice and subsequently transformed into chemically competent 10-beta cells (NEB) as in Section 2.3. Replicate colonies from the plated transformation were selected for liquid cultures, from which the ligated vector was extracted and confirmed via one hour double digest with SpeI and NotI (NEB). The successfully ligated vector was then submitted to the UD Genotyping and Sequencing Center for Sanger sequencing.

2.13 General Analytical Methods

Cell density was measured at A_{600} using a Beckman Coulter DU 730 spectrophotometer (A23616). Culture supernatants were measured for glucose, acetate, ethanol, butyrate, acetone, and butanol concentrations in an Agilent Technologies 1200 Series high-performance liquid chromatograph (HPLC), using an H_2SO_4 mobile phase and a 45 minute separation method (Tomas, et al, 2003).

Table 2.1 – Bacterial Strains and Plasmids Used in this Study

Strains or Plasmids	**Relevant Characteristics**	**Sources**
Strains		
E. coli		
Turbo	*lacIq endA1 Δ(hsdS-mcrB)5*	NEB
10-beta Competent	*recA1 endA1 Δ(mrr-hsdRMS-mcrBC)*	NEB
BL21(pAN3)	pAN3, Km^R, ϕ3T I	Al-Hinai, 2012
C. acetobutylicum		
ATCC 824	Wild-type strain	ATCC
824(p94MCS)	p94MCS, MLS^R	Al-Hinai, 2012
824(p94_solB)	p94_solB, MLS^R	This study
824(p94_tm)	p94_tm, MLS^R	This study
824(p94_6S)	p94_6S, MLS^R	This study
824(p94_6Sb)	p94_6Sb, MLS^R	This study
824(p94_as(solB)_F)	215-bp antisense sequence against *solB*	This study
824(p94_as(solB)_S)	100-bp antisense sequence against *solB*	This study
SolRB	Disrupted CA_P0161, MLS^R, Tc^R	Nair, 1994a
824(pKO_mazF::solB)	pKO_mazF::solB, Th^R	This study
824ΔsolB	*solB* deletion, Th^R	This study
Plasmids		
p94MCS	Amp^R; MLS^R; *ptb* promoter; MCS	Al-Hinai 2012
p94_solB	Amp^R; MLS^R; *ptb* promoter; *solB*	This study
p94_tm	Amp^R; MLS^R; *ptb* promoter; *tmRNA*	This study

Table 2.1 continued

Strains or Plasmids	**Relevant Characteristics**	**Sources**
p94_6S	Amp^R; MLS^R; *ptb* promoter; *6S*	This study
p94_6Sb	Amp^R; MLS^R; *ptb* promoter; *6Sb*	This study
p94_as(solB)_F	Amp^R; MLS^R; *ptb* promoter; 215-bp antisense *solB*	This study
p94_as(solB)_S	Amp^R; MLS^R; *ptb* promoter; 100-bp antisense *solB*	This study
pKO_mazF	Th^R-FRT; repL; ori; *bgaR* and P_{bgaL} upstream of *mazF*	Al-Hinai 2012
pKO_mazF::solB	Th^R-FRT; repL; ori; bgaR and PbgaL upstream of mazF; Th^R flanked by *solB* ~900-bp upstream and downstream regions of homology	This study
pAN3	Km^R; ϕ3T I	Al-Hinai 2012
pCR4-TOPO-TA	Km^R; Amp^R; covalently bound topoisomerase; pUC ori	Life Technologies

lacIq, lac repressor; endA1, nonspecific endonuclease deleted; Δ*(hsdS-mcrB)5*, host-specific restriction and methylcytosine specific restriction deleted; *recA1*, homologous recombination deleted; *Δ(mrr-hsdRMS-mcrBC)*, host-specific restriction and methylcytosine specific restriction deleted; Km^R, kanamycin resistance; ϕ3T I, *B. subtilis* phage ϕ3T I methyltransferase gene; MLS^R, macrolide-lincosamide-streptogramin B resistance; Tc^R, tetracycline resistance; Th^R, thiamphenicol resistance; Amp^R, ampicillin resistance; *ptb*, phosphotransbutyrylase; MCS, multiple cloning site; Th^R-FRT, FRT-flanked thiamphenicol resistance; repL, pIM13 Gram-positive origin of replication; ori, ColE1 origin of replication; bgaR, beta-galactosidase regulator; P_{bgal}, promoter of beta-galactosidase; *mazF*, codon-optimized E. coli MazF toxin gene; pUC ori, pUC Gram-negative origin of replication.

NEB, New England Biolabs, Ipswich, MA; Life Technologies, Grand Island, NY; ATCC, American Type Culture Collection, Manassas, VA

Table 2.2 – Primer Sequences Used in this Study

Primer Name	Sequence (5'->3')	Tm	Description
tm_F	GGAACCTGTGGACCCGCGTTA	61.8	amplify tmRNA sequence
tm_R	AATAGGGAAAAAACCCTCGACATCCG	58.9	amplify tmRNA sequence
tm_F_BamHI	atatatggatccGGAACCTGTGGACCCGCGTTA	65.7	append 5' BamHI site to tmRNA amplicon
tm_R_KasI	tataatggcgccAATAGGGAAAAAACCCTCGACATCCG	65.6	append 3' KasI site to tmRNA amplicon
6S_F	GCCATTTCAAAAAGTATTGATTTTATT	50.1	amplify 6S sequence
6S_R	AAAAAACACAAAGAAGAGAGGCAT	53.6	amplify 6S sequence
6S_F_BamHI	tatataggatccGCCATTTCAAAAAGTATTGATTTTATT	57.4	append 5' BamHI site to 6S amplicon
6S_R_KasI	tataatggcgccAAAAAACACAAAGAAGAGAGGCAT	62.9	append 3' KasI site to 6S amplicon
6SB F	GCACCAAAATAAATGATTTTCTGAAGTAT	53.3	amplify 6Sb sequence
6SB R	TCTCTCTACATTCAAATGTAAATCCAAAATA	53.5	amplify 6Sb sequence
6SB_BamHI F	atatatggatccGCACCAAAATAAATGATTTTCTGAAGTAT	59.7	append 5' BamHI site to 6Sb amplicon
6SB_KasI R	tataatggcgccTCTCTCTACATTCAAATGTAAATCCAAAATA	61.3	append 3' KasI site to 6Sb amplicon
solB F	GTAGTATTATTTCAGAAGTCTACAAATT	49.8	amplify solB sequence
solB R	AAAATATGAAGGTTTAAAATAAACTGC	49.7	amplify solB sequence
as_solB_short_F	GTAGTATTATTTCAGAAGTCTACAAAT	49.3	amplify 100-bp solB sequence
as_solB_short_R	GTATATATAACCCTAATTAATAGTACCAAT	49.2	amplify 100-bp solB sequence
solB-BamHI F[a]	atatatggatccGTAGTATTATTTCAGAAGTCTACAAATT	57.4	append 5' BamHI site to solB amplicon
solB-KasI R	tataatggcgccAAAATATGAAGGTTTAAAATAAACTGC	59.6	append 3' KasI site to solB amplicon
asSOLB-KasI-F	tataatggcgccGTAGTATTATTTCAGAAGTCTACAAATT	60.2	append 5' KasI site to solB amplicon
asSOLB-BamHI-R	atatatggatccAAAATATGAAGGTTTAAAATAAACTGC	56.8	append 3' BamHI site to 215-bp solB amplicon
as_solB_short_BamHI-R	atatatggatccGTATATATAACCCTAATTAATAGTACCAAT	56.8	append 3' BamHI site to 100-bp solB amplicon
solB RH1 F	CCTATTGCGATATGTATAATACTTCC	51.2	amplify solB upstream region of homology
solB RH1 Rv2	GATATAGCTTCTTTTATACTAAAAATTTTCCG	52.0	amplify solB upstream region of homology
solB RH1 F2	aatatacccgggCCTATTGCGATATGTATAATACTTCC	61.1	append 5' XmaI site to solB upstream region of homology

Table 2.2 Continued

solB RH1 Rv2_2	aattatgcggccgcGATATAGCTTCTTTTATACTAAAAATTTTCCG	63.0	append 3' NotI site to solB upstream region of homology
solB RH2 F1	TCTAAATATACTGATAATTCCTAAA	45.3	amplify solB downstream region of homology
solB RH2 R1	GTTCATCTATCCAACCTATTA	46.6	amplify solB downstream region of homology
solB RH2 F2	ggaccTCTAAATATACTGATAATTCCTAAA	52.6	append 5' AvaII site to solB downstream region of homology
solB RH2 R2	aataatgcatgcGTTCATCTATCCAACCTATTA	58.2	append 3' SphI site to solB downstream region of homology
pKOmazF_seq_RH1	CGGTCATGCTGTATGTACAAGG	55.6	confirmation primer for knockout vector
pKOmazF_seq_RH2	GGCGCCACTTAATGATTTGCCCAG	61.0	confirmation primer for knockout vector
solB UP conf F	CTCGATAATTTTCTTTGTATGTGGTAT	51.8	solB mutant forward confirmation and sequencing primer
solB DN conf R	TCATCAATTATTACTGGGGTGTTACC	54.8	solB mutant reverse confirmation and sequencing primer
Th_seq out F	GTAAACGAATTGCAGGAATTGA	51.7	forward confirmation primer that pairs with integrated Th^R marker
Th_seq out R	TGGTCAAAATACTCTTTTCTGTT	50.5	reverse confirmation primer that pairs with integrated Th^R marker
solR453 Seq F	GAGTTGAATTTAGCATGAATTTATTA	48.6	forward primer to amplify 5' crossover integration in solRB
Tc239 Seq R	CATAGAAATTGCATCAACGCATA	51.9	reverse primer to amplify 5' crossover integration solRB
Em373 Seq F	CAATTGTTTTATTCTTTGGTTGAGTAC	51.7	forward primer to amplify 3' crossover integration in solRB
solR1361 Seq R	AATTTTCCGTTAAGTATTTTTTTATCAT	49.4	reverse primer to amplify 3' crossover integration in solRB
Tc300 Seq F	GATATCGTCCATTCCGACAGCAT	56.8	amplify internal fragment of *solR* disruption in solRB
Em450 Seq R	CGATCGTTGTCAGAAGTAAGTTGG	55.7	amplify internal fragment of *solR* disruption in solRB
Em700 Seq F	CTGGATGGAGGCGGATAAAGTT	56.9	amplify internal fragment of *solR* disruption in solRB
Tc3_F	GAAATCTAACAATGCGCTCAT	51.3	amplify integrated pO1x vector in solRB
Em_UP_Rev	TAACGAGTGAAAAAGTACTCAAC	50.8	amplify integrated pO1x vector in solRB
RH3_cen_rev	CTTGTCACAGCTTTGTTAAGAGATTTG	54.8	sequence internal fragment of solRB 3' amplicon
RH3_cen_for	GAACAATGAATTTTTTGATATTCGTAA	49.7	sequence internal fragment of solRB 3' amplicon
p94_seq_F	TGCAGGTCGACTGTGGATGGAG	61.1	p94MCS sequencing primer for confirmation of ligation at MCS
T3	ATTAACCCTCACTAAAGGGA	50.3	5' to 3' sequencing primer of pCR4-TOPO-TA cloning site

Table 2.2 Continued

T7	TAATACGACTCACTATAGGG	47.5	3' to 5' sequencing primer of pCR4-TOPO-TA cloning site
qPCR 6S F	CGAGCACCCACCTATCATTAG	54.9	6S RNA QPCR forward primer
qPCR 6S R	CACAAAGAAGAGAGGCATCCA	54.8	6S RNA QPCR reverse primer
qPCR tm F	GCTTTGAGTAAGGAACGGAATTTAT	53.6	tmRNA QPCR forward primer
qPCR tm R	TCTCCGAGTGCAGTTTGTATTT	54.3	tmRNA QPCR reverse primer
solB qRTPCR F	ATACTAGAACTTATCATGGTAAACA	49.5	solB QPCR forward primer
solB qRTPCR R	ACGTATAATTTACATTTAGCAGTTT	49.2	solB QPCR reverse primer
adhE1 qRTPCR F	CAAGGTGCAAGTGGAGATCTAT	54.5	adhE1 QPCR forward primer
adhE1 qRTPCR R	TGGACCAACATTCTCGGAAAC	55.2	adhE1 QPCR reverse primer
adc qRTPCR F	AGTTGTGCCAGAGCCTTTAG	55.0	adc QPCR forward primer
adc qRTPCR R	GCCTGTCCGCTTTCTGTATAA	54.8	adc QPCR reverse primer
ctfA qRTPCR F	CTCTTACAGCCGATGTAGCATTA	54.3	ctfA QPCR forward primer
ctfA qRTPCR R	CTGCCATTGCCATATAGGGATTA	54.6	ctfA QPCR reverse primer
solR qPCR F	TCCATGGGCAGTACAAACTC	54.8	*solR* QPCR forward primer
solR qPCR R	TCAGGTTCTGCGTCTAAAGC	54.9	*solR* QPCR reverse primer
Q-PCR CAC2071 for	GAGAAAGAGGCAAGCTTTGCAGGT	59.9	spo0A QPCR forward primer
Q-PCR CAC2071 rev	TGTGCTGGAACACCTATTTGATG	55.4	spo0A QPCR reverse primer

Table 2.3 – Oligonucleotide Probes for Northern Blot Analysis.

Probe Name	**Sequence (5'->3')**	**Tm**	**Target**
sCAC1377_sRNA_NB_alt	TACCGCTGCTCTTTCATTGACACC	59.4	6S RNA
sCAC834_sRNA_NB_alt	TTCCCTCCATGAAACGCCTACAAACA	60.5	tmRNA
solB N1	ATAGTTACACCCCAGGGTCT	54.8	*solB*
ctfa N1	CAATTAATTTGGTTGGAGTGCC	52.7	*ctfA*
adc N1	GAAGTGCATCCATATCTGTACG	53.1	*adc*
adhE1 N2	GTTTTAGCTGCTAGTATTGTGGA	52.8	*adhE1*
solR N probe 1	ATATTCTTCGGATTTTTTATAATCTCC	49.9	*solR*
CAP0162_NB_dist	AATTATTAATTACTAATTACTATGCTT	45.7	*adhE1* distal UTR
Ca_P0162 prox_UTR v2	CACTTCTTTCTAAAATATTTATTATAT	44.6	*adhE1* proximal UTR

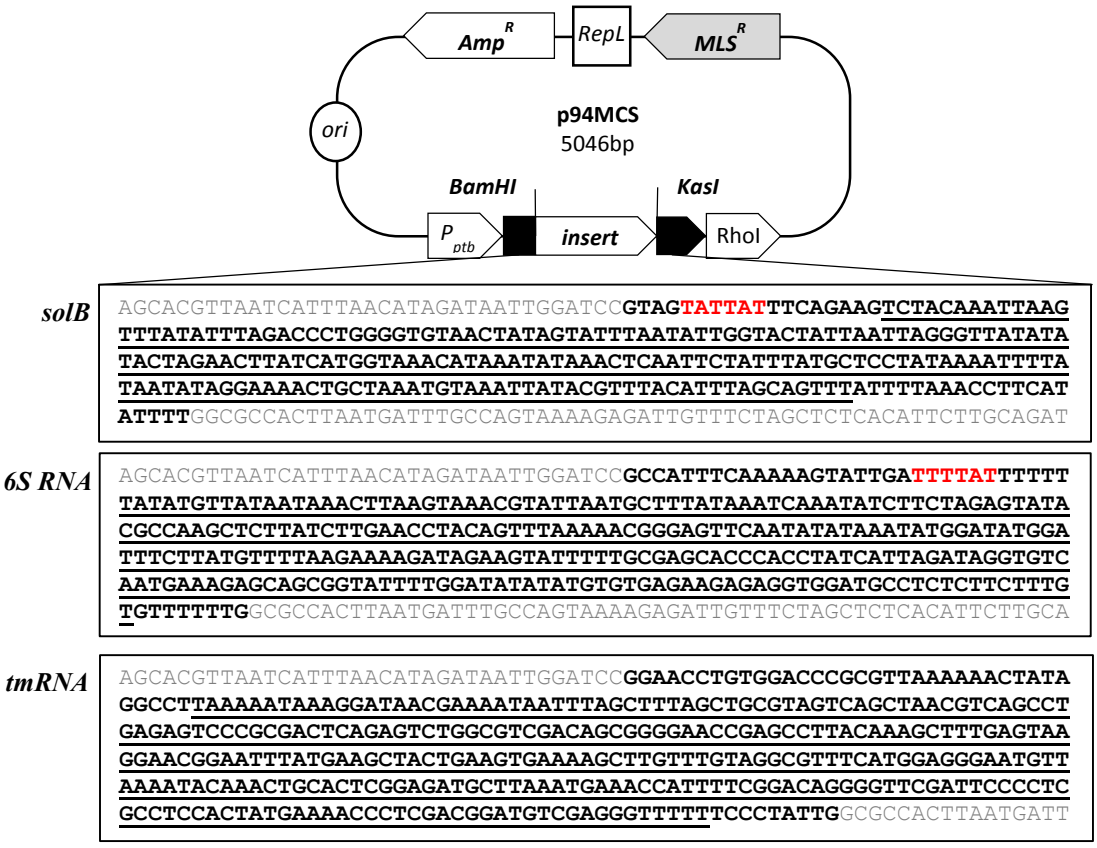


Figure 2.1 – Vector construction and sequencing of overexpression plasmids. Each vector was confirmed by restriction enzyme digest as well as submitted for Sanger sequencing. The MCS site of each vector, containing the appropriate insert, was sequenced and appear below the vector diagram. Total ligated sequence is in bold and in black, vector backbone and ligation sites are in gray, coding sequence for each sRNA is underlined, and predicted -10 sites for promoters are in red. SolB = 234-bp insert, 196-bp coding sequence; 6S RNA = 304-bp insert, 265-bp coding sequence; tmRNA = 345-bp insert, 300-bp coding sequence.

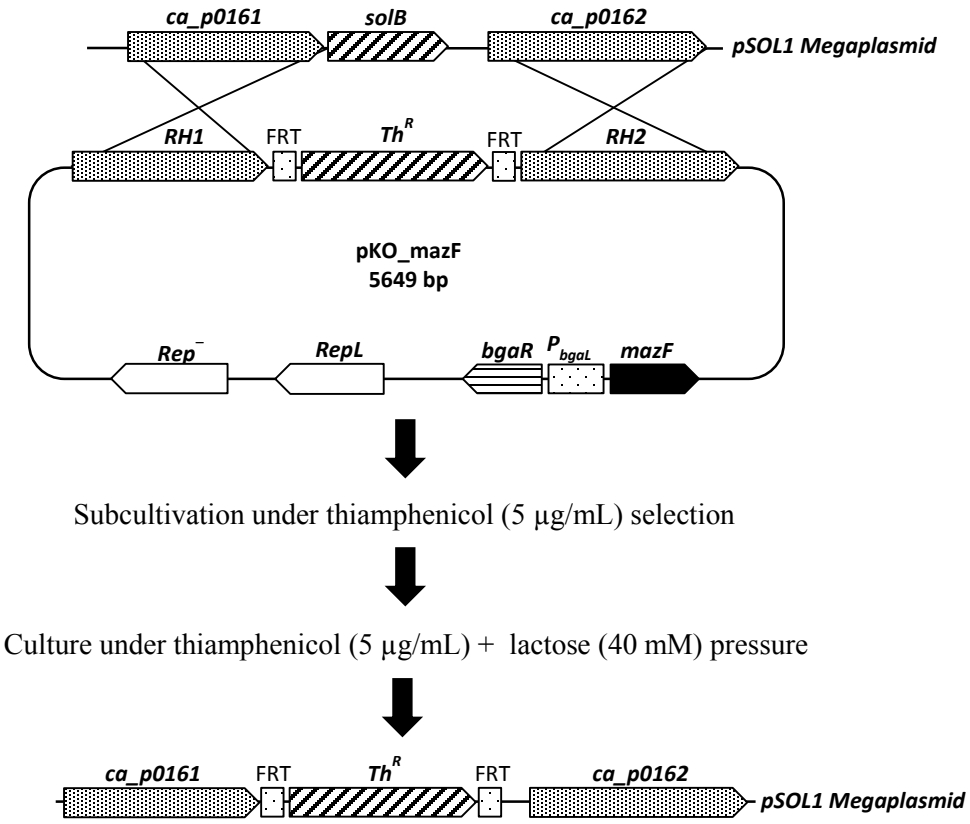


Figure 2.2 – Construction of pKO_mazF::solB allelic exchange knockout vector. Figure adapted from Al-Hinai, et al, 2012.

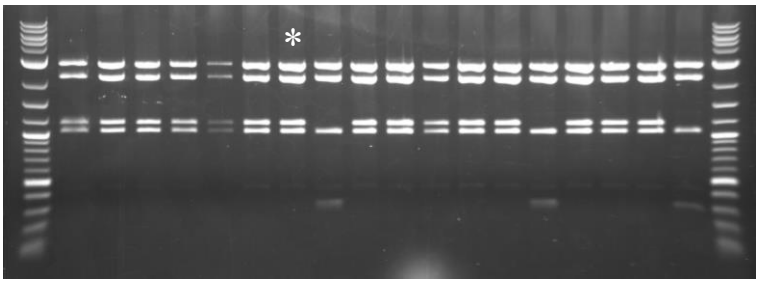


Figure 2.3 – Agarose gel of EcoRI confirmation digests of the pKO_mazF::solB vector (the prep in lane 8 was chosen for subsequent methylation and transformation into ATCC 824).

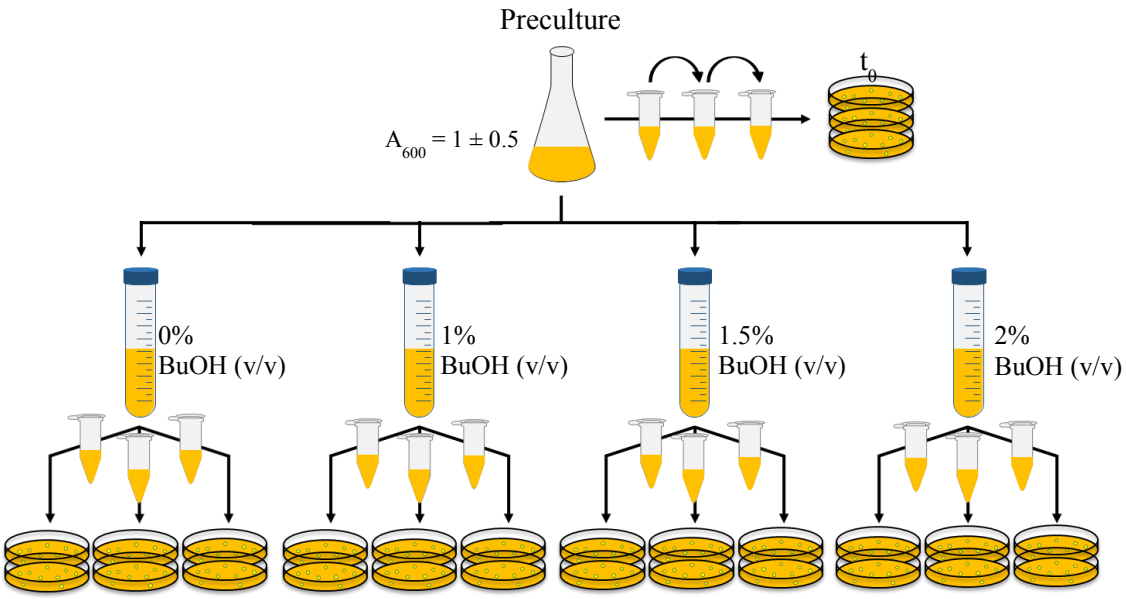


Figure 2.4 – Butanol tolerance assay design. Serial dilutions and plating were performed at 24 and 48 hours; triplicate dilution series were generated in 1.5-mL Eppendorf tubes using 900 µL of CGM and 100 µL of culture or previous dilution.

Chapter 3

CHARACTERIZING THE ROLES OF SELECT SMALL RNA IN BUTANOL TOLERANCE AND PRODUCTION IN CLOSTRIDIUM ACETOBUTYLICUM

6S RNA is a small, noncoding RNA with homologs in hundreds of bacterial species. (Barrick, et al, 2005) 6S RNA facilitates changes in gene expression patterns in the transition from exponential to stationary phase growth; it has been shown in *E. coli* and *B. subtilis* (Gildehaus, et al, 2007; Neusser, et al, 2010) that 6S RNA binds to and sequesters the exponential-phase RNA polymerase holoenzyme; this sequestration allows for a corresponding increase in concentration and activity of the stationary phase RNA polymerase holoenzyme, thereby effecting phase transition. 6S RNA possesses only a moderately conserved primary sequence but a very distinctive secondary structure, and its crucial functional motif is the central loop (Figure 3.1). To interact with RNA polymerase, 6S RNA uses this central loop to mimic the open promoter complex of a DNA molecule entering transcription. The generally well-conserved sequence of the 3' side of the loop has been shown to interact directly with RNA polymerase and is thus believed to be read as a DNA promoter. (Barrick, et al, 2005; Gildehaus, et al, 2007) The 5' side of the loop is very poorly conserved and is believed to serve in stabilizing the loop's molecular structure. (Barrick, et al, 2005)

Studies of 6S RNA indicate a strong influence on stress response and cell survival. Deletion of 6S RNA in *E. coli* (Neusser, et al, 2010) and *Legionella pneumophila* (Faucher, et al, 2010) have shown reduced viability, inhibited outgrowth from stationary phase, and impaired growth in host cells, indicating that 6S RNA is necessary to sustain cell viability in competitive or stressful environments. Transcriptional analysis of genes differentially expressed in 6S RNA deletion mutants

(i.e. nutrient acquisition, transport, amino acid metabolism, detoxification, stress response, and DNA repair) also indicate a functional influence on stress response. (Neusser, et al, 2010; Faucher, et al, 2010) Also, stress-associated down-regulation of several ribosomal biosynthesis genes suggests that loss of 6S RNA results in increased stress to the cell, further indicating its connection to general stress response.

It has been proposed that 6S RNA may sequester RNA polymerase to prevent its degradation until favorable growth conditions return. (Gildehaus, et al, 2007) Thus, 6S RNA binding with RNA polymerase is reversible, enabling the cell to emerge from stationary phase when environmental conditions improve. 6S RNA molecules serve as a template for the transcription of product RNA (pRNA), which are essential for cell viability and successful outgrowth from stationary phase in *E. coli* and *B. subtilis*. (Gildehaus, et al, 2007; Cavanagh, et al, 2012) They regulate 6S RNA activity as the mechanism for its release from RNA polymerase; interaction with pRNA alters 6S RNA structure, preventing its binding to RNA polymerase. (Cavanagh, et al, 2012)

Unlike most genera, a select few organisms, primarily *B. subtilis* and *C. acetobutylicum*, contain multiple copies of the 6S RNA coding sequence. (Barrick, et al, 2005) *B. subtilis* contains two homologs, 6S-1 and 6S-2, that show divergent temporal expression patterns and distinct, albeit complex and incompletely understood, roles in the cell. 6S-1 alone serves as a pRNA template and also influences initiation of sporulation, while the function of 6S-2 is poorly understood. (Cavanagh, et al, 2012, 2013) Better understanding of the activity, regulation, and interaction of 6S RNA will undoubtedly help elucidate not only the details of stress response, sporulation, and other cellular processes but also the mechanisms of protein modulation by sRNA.

Transfer-messenger RNA (tmRNA), also called 10Sa RNA or SsrA RNA, is a noncoding small RNA that functions in a quality control system that rescues ribosomes stalled in incomplete translation. In *B. subtilis*, the gene for tmRNA, *ssrA*, is part of a larger operon that also contains *smpB*, the gene coding for a ribonucleoprotein that acts in complex with tmRNA. This polycistronic operon exhibits complex expression and contains several promoters, including one positioned for *ssrA* and *smpB* alone. (Shin and Price, 2007) The sequence and structure of tmRNA illustrate the remarkable mechanism by which it rescues stalled ribosomes. The 3' and 5' ends of the transcribed tmRNA molecule form a partial tRNA structure, which is aminoacylated with alanine. (Komine, et al, 1994) This alanyl-tmRNA molecule binds EF-Tu and GTP to interact with a stalled ribosome. In addition to the tRNA-like region, tmRNA also contains an mRNA-like sequence coding for a short peptide called tag-peptide. The stalled ribosome translates this mRNA-like region, adding the tag-peptide sequence to the incompletely translated nascent protein. (Fujihara, et al, 2002) Completion of this translation releases the ribosome for recycling and effectively translates a single (misfolded) protein from two mRNA; therefore, this ribosomal rescue system is referred to as *trans*-translation. Furthermore, tmRNA targets the misfolded protein for degradation by virtue of the tag-peptide. (Jannsen and Hayes, 2012; Himeno, et al, 2014)

The importance of tmRNA activity in cellular stress response and survival has been well studied. In *B. subtilis*, tmRNA is more highly expressed at higher temperatures (50°C), with growth of a tmRNA deletion mutant considerably inhibited at high (52°C) and low (16°C) temperature as well as under ethanol stress. (Fujihara, et al, 2002; Shin and Price, 2007; Muto, et al, 2000) Also, proteins related to stress

response, including PerR and GsiB, have been shown to be tagged by tmRNA in *B. subtilis*. (Fujihara, et al, 2002) Deletion of tmRNA in *E. coli* reduces motility as well as growth rate at high temperature. (Komine, et al, 1994) Similarly, tmRNA disruption in *Francisella novicida*, a Gram-negative pathogen, showed decreased growth rates and greater sensitivity to high and low temperature conditions as well as to antibiotics that target translation. (Svetlanov, et al, 2012) Also, *Mycobacterium smegmatis* has been observed to increase expression levels of tmRNA as a direct result of ribosome-targeting antimicrobials. (Andini et al, 2011) These phenotypes are strong indications that tmRNA, while not necessarily essential for cell viability in these species, plays an influential role in cell survival under stress conditions and suboptimal environments.

The tmRNA-based ribosomal rescue system is unique in its use of small RNA and its ability to tag misfolded proteins to prevent their accumulation; furthermore, it is the only such ribosome rescue system found in all classes of bacteria. (Himeno, et al, 2014) Although tmRNA and its significance to cell survival under stress in many organisms has been well documented and characterized (Jannsen and Hayes, 2012), a tmRNA-like sRNA sequence was only recently identified in *C. acetobutylicum* (Chen, et al, 2011; Venkataramanan, et al, 2013). Therefore, further study of this sRNA and its significance in this Gram-positive organism is needed, especially given its characterized influence on cell survival and stress response in other organisms.

3.1 Two Additional 6S RNA Homologues Exist in *Clostridium acetobutylicum*

Despite most organisms possessing a single copy of 6S RNA, three homologues of 6S RNA have been identified in *Clostridium acetobutylicum* by computational analysis (Barrick, et al, 2005); two homologues share an identical primary sequence (212 bp), and the third unique homologue (265 bp) matches the 6S

RNA that has been studied previously (Chen, et al, 2011; Venkataramanan, et al, 2013). All three homologs are located on the chromosome and are dispersed throughout the genome. In support of the initial study that identified them, all three homologues were validated against a covariance model generated from the 6S RNA consensus sequence on the Rfam database using the Infernal RNA modeling program (Nawrocki and Eddy, 2013). The previously studied 6S RNA homologue is identified as 6S-C, and the two additional homologues are named 6S-A and 6S-B; in this study, '6S' is used in reference to 6S-C and '6Sb' is used in reference to the sequence corresponding to homologues 6S-A and 6S-B.

Pairwise local alignment (using the Smith-Waterman algorithm) of the two clostridial 6S RNA sequences showed considerable conservation between them (71.2% identity and similarity, 10.1% gaps), revealing stronger sequence alignment than the two 6S RNA homologues of *B. subtilis* (53.4% identity and similarity, 38.9% gaps). It is possible that the homologues of *B. subtilis* have diverged more extensively over time than the clostridial homologues. Secondary structure predictions for all three clostridial 6S RNA homologues (Figure 3.2) confirmed, with minor variation, the conservation of the central loop motif that is essential for 6S RNA function, providing further validation of the homologues. Furthermore, the putative sequence for the central loop in *C. acetobutylicum* was identified, and although its exact position varies slightly among the homologues, it is well conserved (Figure 3.2). The binding motif for sigma factor A, once it is identified, can be used to validate this putative central loop sequence; alternatively, this sequence, as the region responsible for interaction with the RNA polymerase and sigma factor, may be helpful in determining the sigA binding motif.

Moreover, the two additional 6S RNA homologues showed an increase in expression under butanol stress as part of the Wasabi database of previously performed RNA Deep Sequencing (Venkataramanan, et al, 2013). According to this data, both homologue sequences also overlap with putative genes CA_C1101 (6S-A) and CA_C1229 (6S-B), which both code for hypothetical proteins; however, expression of these ORFs in the RNA-Seq data is only seen in the regions shared with the 6S RNA homologues. It may be that the ORF prediction tools, which in part use promoter sequences to identify hypothetical proteins, erroneously reported the central loop motifs of the 6S RNA sequences as promoters of actual protein coding regions.

To date, no definitive studies have been performed to investigate or characterize the functions of the three 6S RNA homologues in *C. acetobutylicum*. However, several studies have examined the two 6S RNA homologues present in *B. subtilis* and their divergent functions (Cavanagh, et al, 2012, 2013). Although both homologues bind to the RNA polymerase holoenzyme, it has been shown that deletion of 6S-1, but not 6S-2, results in an earlier induction of the sporulation pathway; it has been speculated that these effects on sporulation are secondary effects triggered by a change in nutrient uptake efficiency. Also, synthesis of pRNA was shown to be more efficient from the 6S-1 template. These studies suggest that the two 6S RNA homologues target different promoters and regulate different genes. While further studies must be conducted to fully elucidate the roles of the 6S RNA homologues in *B. subtilis*, it can be inferred from the current findings that the homologues in *C. acetobutylicum*, a closely related species, may similarly exhibit a complex pattern of divergent function and possibly regulate key cellular processes relating to metabolism

and/or sporulation initiation. It was therefore worthwhile to consider any significance of all 6S RNA homologues in this study as well as in consideration of future work.

3.2 Overexpression of tmRNA (sCA_C834) and 6S RNA (sCA_C1377) in *Clostridium acetobutylicum* Results in Increased Cell Survival under Butanol Stress as well as an Increase in Butanol Production

Overexpression plasmids for the tmRNA, 6S RNA, and 6Sb sequences were designed, constructed, and confirmed as described in Chapter 2 (Figure 2.1), and these plasmids were transformed into ATCC 824 as described above. The increased expression levels of tmRNA and 6S RNA were validated in their respective overexpression strains using both two-step QPCR and Northern analysis, as described in Chapter 2; QPCR primer sets and Northern oligonucleotide probe sequences are listed in Tables 2.2 and 2.3, respectively. Fold changes in sRNA transcript expression were calculated for overexpression strains relative to the expression levels in a plasmid control strain (824(p94MCS)) carrying an empty p94MCS expression vector. All transcript expression was normalized against expression of the constitutively expressed clostridial housekeeping gene CA_C3571.

Significant overexpression was seen in both recombinant strains, and this increased expression was maintained throughout growth. Fold change calculations derived from QPCR Ct values reveal 1040-fold, 2260-fold, and 2190-fold increases in 6S RNA transcript expression in 824(p94_6S) at six, 12, and 24 hours, respectively. Weaker yet still significant overexpression of the tmRNA transcript is shown in 824(p94_tm), with 17-fold increases at six and 12 hours and a 5-fold increase at 24 hours (Table 3.1).

Northern analysis also demonstrates a clear increase in expression of the sRNA transcripts in their respective overexpression strains relative to 824(p94MCS) and wild

type ATCC 824 at both exponential phase (6 hours) and stationary phase (24 hours) (Figure 3.3). In addition to the intense bands corresponding to the tmRNA (300 bp) or 6S RNA (265 bp) transcripts, several additional, fainter bands appear above and below the target bands in both overexpression strains. The lower bands, corresponding to transcripts of smaller size, are likely the result of post-transcriptional processing or trimming of the sRNA product; this trimming has been suggested for both tmRNA and 6S RNA in previously published studies (Venkataramanan, et al, 2013; Komine, et al, 1994). In addition, the oligonucleotide probe used for 6S RNA aligns with the 6S sequence (75.0% identity and similarity) as well as the 6Sb homologue sequence (76.5% identity and similarity), offering a partial explanation for the additional bands seen previously (Venkataramanan, et al, 2013) as well as in this study. The larger bands present in both tmRNA and 6S RNA Northern blots are believed to be transcriptional products from additional promoters on the p94MCS vector backbone located upstream of its cloning site; using the Softberry promoter prediction tool (Solovyev and Salamov, 2011), several putative promoter sequences were predicted in the 800-bp region upstream of the cloning site. Not only do the transcripts generated from several of these predicted promoters correspond in approximate size to the additional bands on the Northern blots, but one of these bands (a product of approximately 800 bp) also appears in the 6-hour sample of 824(p94MCS), further validating this explanation of their origins on the vector backbone. The activity from these additional promoter sequences may be silenced fairly easily by the insertion of a Rho-independent terminator directly upstream of the *ptb* promoter of the cloning site on p94MCS; however, these additional transcripts do not present a major cause for

concern or complication regarding the intended use of these plasmids (i.e. the overexpression of the sRNA sequence) and therefore were not corrected for this study.

Biological replicate cultures of 824(p94_tm), 824(p94_6S), and 824(p94MCS) were subjected to the butanol tolerance assay, with a minimum of n = 3 and maximum of n = 7 used for all strains. Percent survival was calculated from CFU counts at 24 and 48 hours relative to CFU counts taken from the initial culture (t_0); results are presented graphically in Figure 3.4. Increased cell survival is observed in 824(p94_6S) relative to 824(p94MCS) after 24 hours under 2% (v/v) butanol stress, and after 48 hours under 1% and 1.5% (v/v) butanol stress; cell survival at 48 hours under 2% (v/v) butanol stress, although also heightened, is not statistically significant due to standard error. Increased cell survival is also seen in 824(p94_tm) relative to plasmid control under 1% and 2% (v/v) butanol stress, although this increase in tolerance is not manifested until 48 hours. In fact, much of the overall statistically significant increase in cell survival under butanol stress for both strains occurred at 48 hours.

Two replicate cultures of 824(p94_6Sb) were also tested in the butanol stress assay, but no significant change in cell survival relative to 824(p94MCS) was observed (data not shown). Not only does overexpression of 6S RNA and tmRNA in ATCC 824 afford increased tolerance to butanol, but also this tolerance phenotype is particularly significant in that it persists under very high butanol stress and into mid- to late stationary phase. It can be concluded that 6S RNA and tmRNA contribute to butanol tolerance in *C. acetobutylicum*. This function in stress response, given the level and duration of tolerance seen upon overexpression, may be relevant or hold potential in the pursuit of engineering an industrial strain capable of superior solvent production.

3.3 Overexpression of 6S RNA, but not tmRNA or 6Sb, Yields Higher Butanol Production

In order to further characterize the effects of overexpression of 6S RNA and tmRNA in ATCC 824, the metabolite profile of each strain was analyzed by HPLC. Static flask cultures were grown for 72 hours and measured regularly for cell density and the collection of supernatant samples for HPLC analysis. To determine any significant change in metabolite concentrations, the plasmid control strain 824(p94MCS) was used for comparison rather than wild type since higher solvent titers have been observed and documented routinely for many years in plasmid-containing clostridial strains, including plasmid control strains; it is believed that the presence of any plasmid results in sufficient metabolic pressure on the transformed cell to produce slightly higher solvent titers. (Nair, et al, 1994) Overexpression of 6S RNA yielded higher butanol production relative to plasmid control; in fact, butanol production in 824(p94_6S) was comparable to that observed in the solRB high-solvent producing strain (Figure 3.5, Table 3.2). Overexpression of tmRNA, however, appears to reduce solvent titers to wild type levels, and overexpression of the 6Sb homologue sequence showed no change in solvent titers relative to plasmid control. Growth curves were also generated for each strain (Figure 3.6), indicating that 6S RNA overexpression facilitates growth to a higher cell density than the other overexpression strains or control strains.

3.4 Conclusions on the Activity of 6S RNA and tmRNA in Solventogenesis and Solvent Tolerance

Overexpression of 6S RNA, but decidedly not the additional 6Sb RNA homologue sequence, results in a phenotype of higher tolerance to butanol and increased butanol production in ATCC 824. Overexpression of tmRNA similarly

generates a phenotype of higher butanol tolerance, but paired with a reduction of solvent production to wild type levels. These results suggest a function for 6S RNA as well as tmRNA in butanol stress response in *C. acetobutylicum*, and further that 6S RNA either plays a direct role in butanol production or has a related function that positively affects production.

The results presented here suggest a role for tmRNA in butanol tolerance, but it is unclear why its overexpression also appears to negatively affect solvent production. A possible explanation for this reduction in solvent titers is an inhibition of solventogenic gene translation due to the overabundance of tmRNA transcripts competing for the active sites of functioning as well as stalled ribosomes in the cell. The increased butanol tolerance observed in the tmRNA overexpression strain may result from the cell's enhanced ability to correct for cell damage resulting from butanol accumulation – i.e., the overabundance of tmRNA can rescue more stalled ribosomes and facilitate the degradation of more misfolded and partially translated proteins than the wild type or plasmid control strain. The prominence of tmRNA activity in stress response in other organisms has been observed (Janssen and Hayes, 2012), and this explanation for butanol tolerance in ATCC 824 agrees with these existing observations.

It is likely that 6S RNA imparts the observed increase in butanol tolerance as well as butanol production by its suppression of sigA-based exponential phase transcription and subsequent facilitation of stationary phase gene expression, as described above (Figure 3.1). These genes include the *sol* locus genes, responsible for solventogenesis, as well as stress response genes that impart protection against the accumulation of solvents during late growth as the cell prepares for sporulation.

Overexpression of 6S RNA likely intensifies the expression of these stationary-phase genes to yield not only a more robust stress response but also increased solvent production. It is also possible that this stationary phase gene expression facilitated by 6S RNA may be induced earlier or persist longer in the overexpression strain to impart the observed phenotype. It is interesting, however, that 6S RNA overexpression increases only butanol production, with acetone and ethanol concentrations remaining largely unaltered; the reason for this dichotomy is unknown. Additional work must be done to investigate the interaction of 6S RNA with solvent formation genes, particularly the *sol* locus. Most illuminating would be transcriptomic studies utilizing QPCR and RNA Deep Sequencing to examine possible targets affected by 6S RNA overexpression. In this way, the interactions and mechanistic details of 6S RNA in clostridia, especially within the context of solventogenesis, can be further elucidated.

The changes in butanol production and tolerance result only from overexpression of the first 6S RNA homologue; the additional homologues do not seem to affect these processes. It is possible that the additional homologues serve divergent or secondary functions, as has been observed in studies of *B. subtilis* (Cavanagh, et al, 2012, 2013). It has been described that 6S RNA, although universally observed as a transcriptional regulator, can function in varying processes depending on the species (Cavanagh and Wassarman, 2014). It is possible, therefore, that the first 6S RNA homologue in *C. acetobutylicum* regulates genes involved in solventogenesis and solvent tolerance mechanisms, and the additional homologues function in other cellular processes. This dynamic would explain why overexpression of one sequence does not impart the same phenotype as overexpression of the other.

These speculations and observations of the effects of 6S RNA and tmRNA on butanol tolerance and production present a fascinating area for further study to elaborate not only on the functions of sRNA in the Gram-positive clostridia but also on their regulation of solvent production. Much of the work and results presented here, while promising, is preliminary. The metabolite data were taken from small-scale cultures without persistent pH control; larger scale batch fermentation studies will more clearly define the observed effects in both strains regarding solvent production. Also, co-overexpression of 6S RNA and tmRNA may augment the phenotypes of increased tolerance and increased butanol production. Co-overexpression of both 6S RNA homologue sequences may also be revealing not only in examining solvent tolerance and production but also in more fully understanding 6S RNA function in clostridia. These and other future directions are elaborated further in Chapter 5.

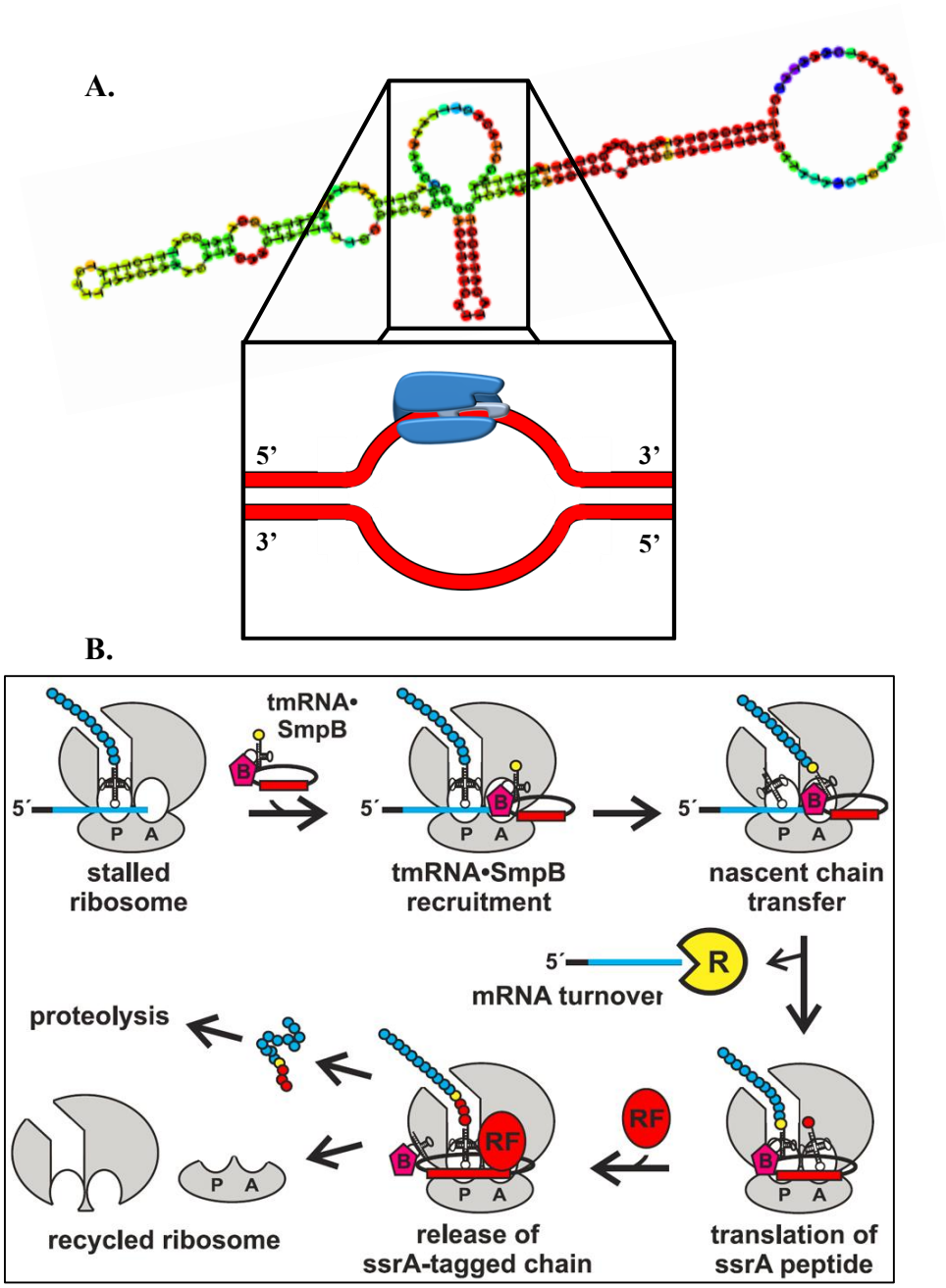


Figure 3.1 – Mechanisms of action of (A) 6S RNA and (B) tmRNA. (A) Shown is the central loop motif on the structure of 6S-C in *C. acetobutylicum* (sCA_C1377), bound to the exponential-phase RNA polymerase; (B) tmRNA ribosomal rescue process (from Janssen and Hayes, 2012), showing the tmRNA-SmpB complex entering the ribosomal A site for translation of the tag-peptide and then release of the ribosome, mRNA, and nascent protein, facilitated by release factors (RF).

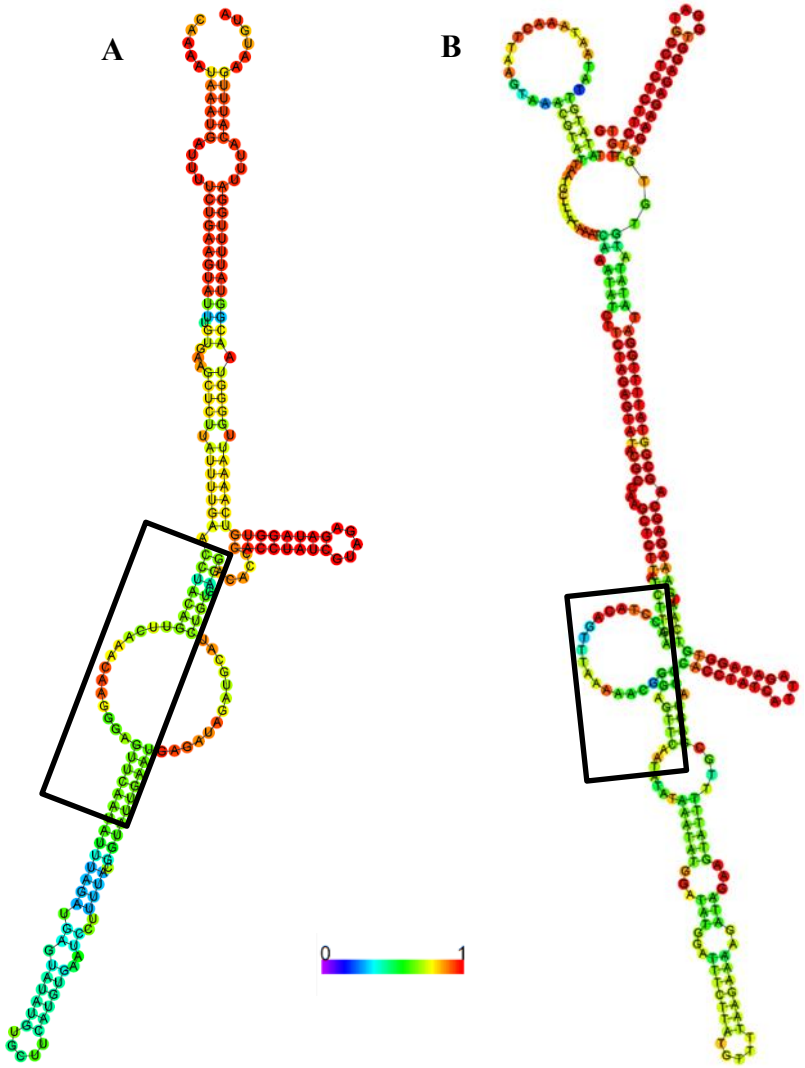


Figure 3.2 – Secondary structure predictions for 6S RNA homologues in *C. acetobutylicum*. The sequences for (A) 6S-A, 6S-B, and (B) 6S-C exhibit the conserved central bubble motif characteristic of 6S RNA (boxed); the primary sequence of the central bubble for both sequences is very highly conserved, supporting both sequences as 6S RNA homologues and also identifying a possible consensus sequence for clostridial 6S RNA; minimum free energy for the 6S-A/6S-B structure is -52.60 kcal/mol, and for the 6S-C structure is -67.30 kcal/mol; secondary structures generated using Vienna RNA Websuite (Gruber, et al, 2008).

Table 3.1 – Fold changes in expression of tmRNA and 6S RNA in overexpression strains relative to plasmid control (824(p94MCS)). Transcript expression is normalized to the constitutively expressed housekeeping gene CA_C3571, and derived from average Ct values from QPCR (n = 2).

Time	*tmRNA* in 824(p94_tm)	*6S RNA* in 824(p94_6S)
Exponential t_6	+ 17	+ 1040
Transitional t_{12}	+ 17	+ 2260
Stationary t_{24}	+ 5	+ 2190

Table 3.2 – Final metabolite levels in recombinant strains and ATCC 824 after 72 hours of static culture growth (40 mL). Numbers in parentheses represent standard deviation (n=3).

	Final Product Titers (mM)					
Strain	**Butanol**	**Acetone**	**Ethanol**	**Acetate**	**Butyrate**	**Glucose**
ATCC 824	135.0 (±27.0)	70.8 (±4.8)	19.2 (±5.5)	1.2 (±6.6)	6.2 (±5.6)	276.9 (±52.1)
p94MCS	150.4 (±16.1)	73.9 (±4.9)	29.6 (±4.9)	-1.3 (±1.9)	3.6 (±4.3)	311.8 (±24.2)
824(p94_tm)	104.2 (±35.6)	61.7 (±13.8)	16.2 (±5.1)	11.8 (±2.1)	14.2 (±5.8)	231.5 (±58.7)
solRB	179.1 (±17.7)	88.2 (±12.3)	30.6 (±0.5)	-4.4 (±1.2)	2.4 (±1.8)	372.1 (±23.7)
824(p94_6S)	178.9 (±6.4)	78.8 (±1.1)	26.8 (±0.7)	-2.3 (±1.4)	2.7 (±1.0)	358.8 (±13.6)
824(p94_6Sb)	148.9 (±19.5)	78.1 (±5.5)	23.9 (±1.9)	-2.2 (±2.4)	4.8 (±1.9)	326.5 (±14.7)

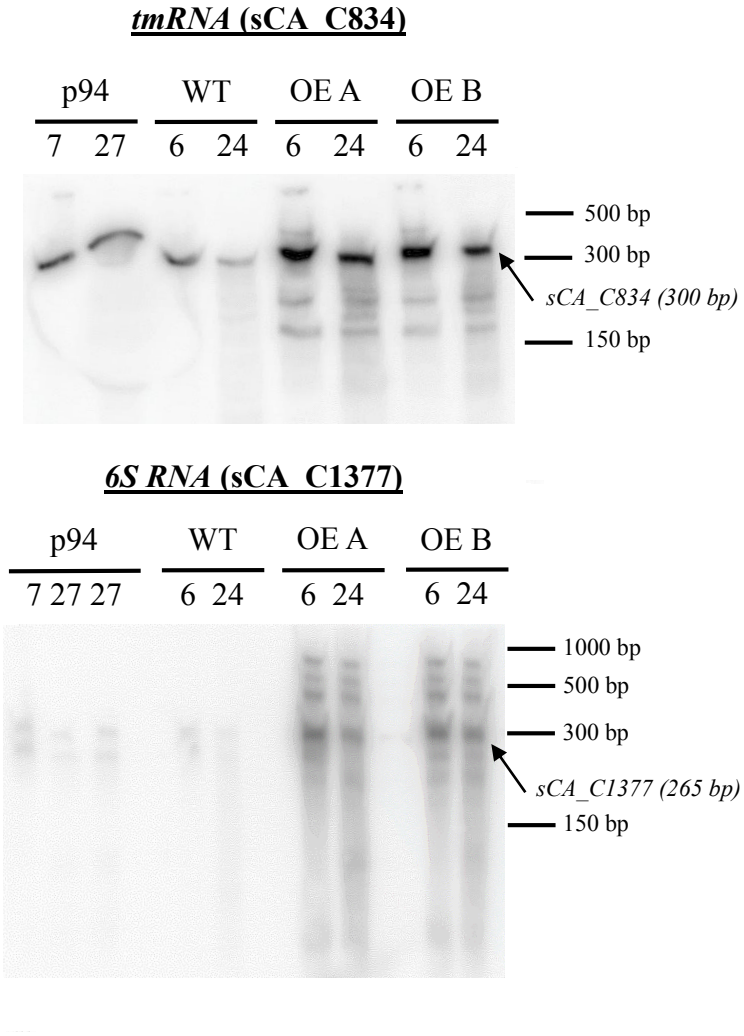


Figure 3.3 – Northern blot analysis of 824(p94_6S) and 824(p94_tm). OE, overexpression strain replicates; five µg samples of total RNA; sampling times for plasmid control was adjusted based on cell density of the culture to correspond with the other strains; additional bands are due to post-transcriptional processing and additional promoter activity on the vector backbone.

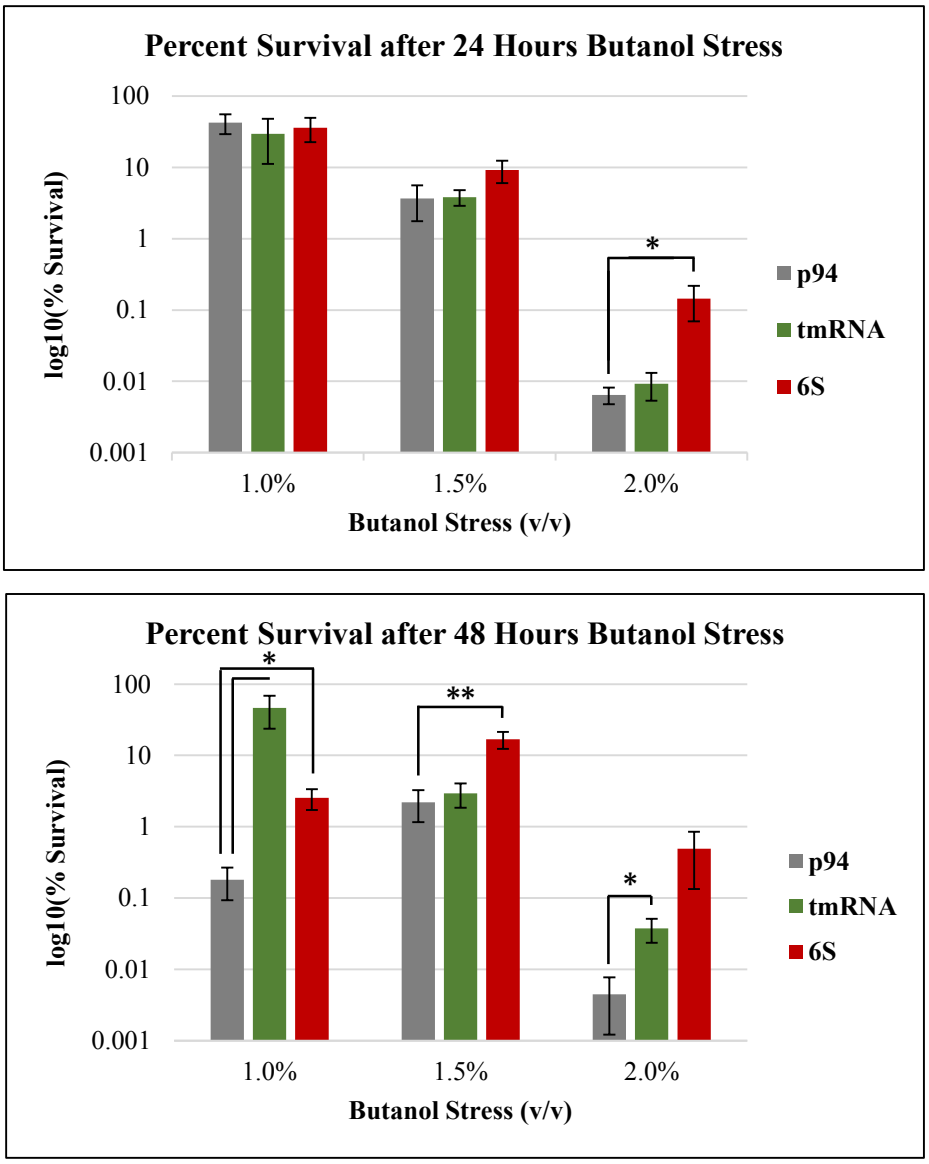


Figure 3.4 – Percent survival of 824(p94_6S), 824(p94_tm), and 824(p94MCS) ($3 \leq n \leq 7$). Error bars show standard error of the mean; statistical significance * = $p \leq 0.05$, ** = $p \leq 0.01$, from two-tailed unpaired t-test.

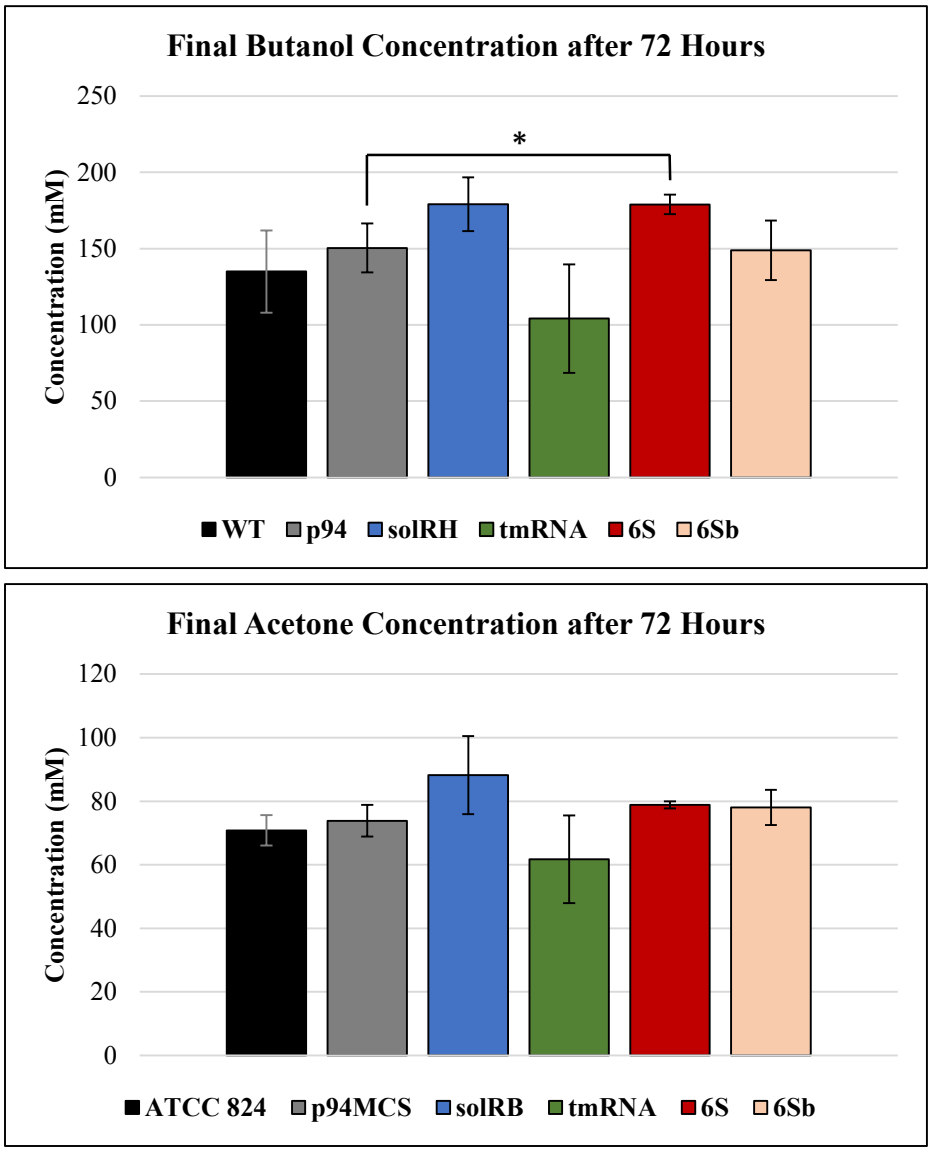


Figure 3.5 – Final solvent concentrations of recombinant strains and ATCC 824 (WT). Error bars represent standard deviation (n=3); * = $p \leq 0.05$, from two-tailed unpaired t-test.

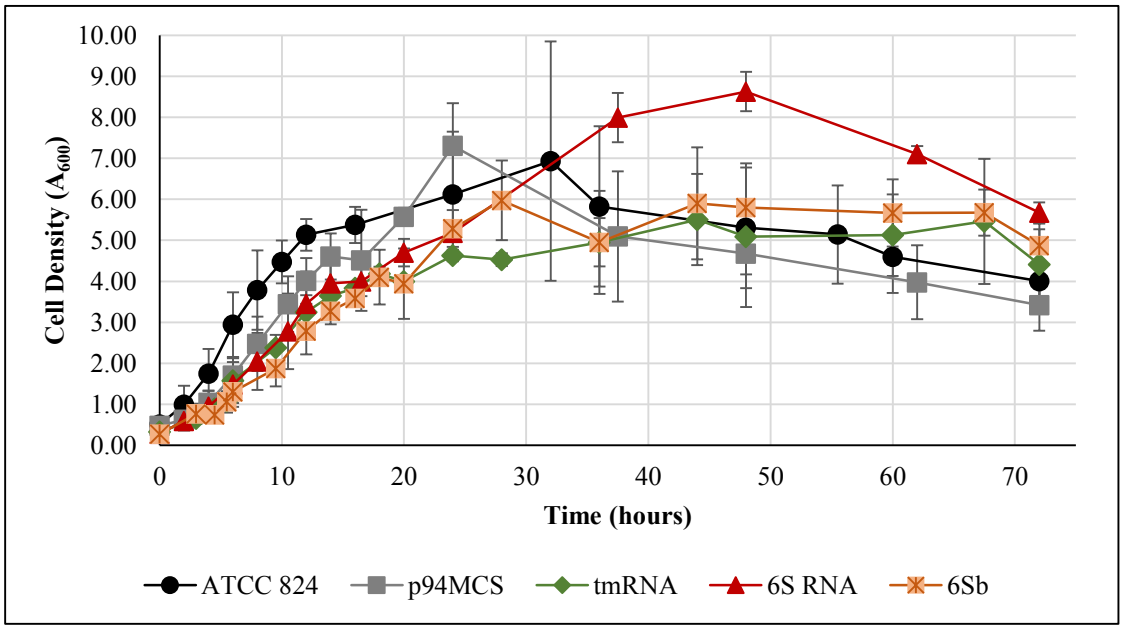


Figure 3.6 – Growth curves for 72-hour cultures of ATCC 824, 824(p94MCS), 824(p94_tm), 824(p94_6S), and 824(p94_6Sb) (40 mL). Error bars represent standard deviation (n = 3)

Chapter 4

INVESTIGATING THE INTERACTION BETWEEN SOLB AND SOLVENTOGENESIS IN CLOSTRIDIUM ACETOBUTYLICUM

As described in Chapter 1, the *solR* gene was thought to code for a repressor of the *sol* locus genes and therefore of solventogenesis. Studies investigating this hypothesis included the construction and characterization of a *solR* overexpression strain as well as the solRB strain, in which the *solR* ORF was disrupted by the insertion of the suicide vector pO1x (Nair, et al, 1999; Harris, et al, 2001). Overexpression of *solR* resulted in attenuated solvent production, and its disruption yielded increased solvent titers in solRB, seemingly confirming *solR* as a repressor of solventogenesis. However, subsequent studies were unable to show interaction between SolR and solvent gene DNA sequences (Thormann, et al, 2001) and suggested instead that SolR is a membrane-bound, extracellular protein involved in glycosylation, largely disproving a regulatory role for *solR*. Recent thought (Durre, et al, 2012; Zimmermann, 2013) has pointed instead to *solB*, a small RNA sequence downstream of *solR*, as a presumed regulator of solventogenesis in *C. acetobutylicum* (Figure 1.3). The *solR* overexpression strain was found (Thormann, et al, 2002) to have included with the *solR* ORF a downstream regulatory region that includes the *solB* sequence, and it was shown (Thormann, et al, 2002) that the observed loss of solvent production was the result of overexpression of this region rather than of *solR*. A viable explanation for the high-solvent phenotype observed in solRB, however, as well as whether the *solB* sequence was affected in this strain, has not been offered and remains unknown.

4.1 The Disruption of the *solR* ORF in the High-Solvent Producing solRB Strain did not Affect the Downstream *solB* Sequence

In order to resolve this longstanding problem regarding solRB and the connections among *solR*, *solB*, and solventogenesis, the disrupted region of the solRB genome was amplified and sequenced. Each end of the pO1x suicide-vector integration was successfully amplified by PCR (Figure 4.2). Sanger sequencing of these integration sites revealed that the disruption of *solR* by the pO1x suicide vector occurred entirely within the *solR* ORF (Figure 4.3). Also, successful PCR amplification using the upstream site's reverse primer and downstream site's forward primer (Figure 4.2, final two lanes) indicates the presence of multiple copies of pO1x in the solRB genome. These results collectively confirm that only the *solR* ORF was disrupted in solRB and that the downstream region was not disrupted or otherwise altered; however, they do not suggest a clear explanation of the seemingly anomalous solRB phenotype, namely, its high level of solvent production, or evidence of a regulatory role for *solR* or the region immediately downstream, including *solB*.

With the genome of the solRB strain more clearly defined, and since a connection between *solR* and solventogenesis remains undetermined, *solB* provides a more promising explanation as a putative regulator of the solventogenic genes (i.e. the *sol* locus genes, Figure 1.2) and a potentially more illuminating investigation of the regulation of solventogenesis than an explanation based on the role of *solR*. Although a clear explanation of the high-solvent phenotype or the role of *solR* was not identified from the sequencing of the solRB genome, the expression of *solR* was investigated in the experiments of this study, and those results are presented in Section 4.8 below.

4.2 Overexpression and Deletion of *solB* in *C. acetobutylicum* ATCC 824

The first phase of the study of *solB* as a regulator of solventogenesis involved testing whether its overexpression and deletion generate the same changes in solvent production observed in the similar, previous study of *solR* (Nair, et al, 1999). Overexpression of *solB* in *C. acetobutylicum* was achieved through cloning a 234-bp sequence consisting of the complete 196-bp *solB* coding sequence (Zimmermann, 2013 [Dissertation]) with the neighboring 18-bp upstream region and 20-bp downstream region, into the p94MCS expression vector as described in Chapter 2 (Figure 2.1); transformation of this plasmid into ATCC 824 generated the overexpression strain 824(p94_solB). A *solB* deletion strain, hereby named 824ΔsolB, was generated using the pKO_mazF allelic exchange system previously described (Al-Hinai, et al, 2012) and the amplified regions of homology described in Chapter 2. The 824(p94_solB) strain construction was confirmed by Sanger sequencing (Figure 2.1), and the loss of *solB* in 824ΔsolB was confirmed by PCR analysis (Figure 4.4) and subsequent Sanger sequencing (Figure 4.5). PCR analysis confirmed both ends of the allelic exchange integration and demonstrated the absence of *solB* from 824ΔsolB genomic DNA and cDNA samples. Relative *solB* expression in both recombinant strains was also validated by Northern blot and QPCR analysis. Northern blot analysis (Figure 4.6) revealed significantly increased *solB* expression in 824(p94_solB); also included are the previously explained artifacts (see Section 3.2) generated from additional promoter activity on the vector backbone. There were no detectable bands for *solB* in 824ΔsolB or the wild type. The lack of detectable *solB* expression in the wild type agrees with minimal native expression of *solB* observed in multiple RNA Deep-Sequencing (RNA-Seq) studies (see Appendix A); this indicates that the low

levels of *solB* mRNA in ATCC 824, although sufficient for minimal detection by RNA-Seq, is below the threshold detection for Northern analysis.

In agreement with the Northern blot analysis, Ct values and fold changes generated from QPCR experiments (Figure 4.7) substantiate the significant increase in *solB* transcript in 824(p94_solB) and the absence of its expression in 824ΔsolB. Expression of *solB* is investigated and discussed further in subsequent experiments (Section 4.6 below).

4.3 Overexpression as well as Deletion of *solB* in ATCC 824 both Result in Severe Attenuation of Solvent Production

Based on the hypothesis that *solB* serves as a repressor of solventogenesis, it should follow that its overexpression will yield a solvent-negative phenotype and its deletion will produce higher solvent titers (or, at minimum, no change in solvent production). To test this hypothesis, the effects of *solB* overexpression and deletion on metabolite concentrations was investigated by HPLC.

Three biological replicates of static cultures of ATCC 824, 824(p94MCS), 824(p94_solB), 824ΔsolB, and solRB were grown for 72 hours and assayed regularly for both cell density and for metabolite concentrations in the supernatant samples by HPLC. As anticipated, the overexpression strain 824(p94_solB) exhibited severely attenuated production of butanol (19.8 mM), acetone (6.6 mM), and ethanol (3.7 mM) relative to both wild type and plasmid control (Figure 4.8, Table 4.1). Acid production was sustained throughout the stationary phase, with butyrate and acetate concentrations peaking at 51 mM and 15 mM, respectively, and remaining constant after approximately 24 hours (Figure 4.9, Table 4.1). That is, there was no acid re-

uptake, which suggests that the expression or activity of the enzyme CoA-transferase (CoAT) was very low.

The deletion strain 824ΔsolB, in sharp contrast to the expected outcome, exhibited an even more severe solvent-deficient phenotype (Figure 4.8, Table 4.1), showing concentrations of butanol (5.5 mM), acetone (0.0 mM), and ethanol (2.7 mM) even lower than those of 824(p94_solB). This loss of solvent production was again associated with sustained acid production, with butyrate and acetate concentrations following the same pattern of accumulation and peaking at 65.7 mM and 12.5 mM, respectively (Figure 4.9, Table 4.1). Glucose consumption in both strains was reduced about three-fold as well, indicating that overall metabolism was constrained or inhibited.

The transition phase of clostridial growth, occurring between 10 and 12 hours and representing the metabolic shift from acid fermentation to solvent production, marks the point at which both 824(p94_solB) and 824ΔsolB diverge from wild type growth. The continued increase in acid production after 12 hours indicates that transition did not occur in either strain and that this loss or severe inhibition of the solventogenic phase of metabolism is a result of both *solB* overexpression and deletion. Growth curves of all strains, generated from cell density measurements (Figure 4.10), reveal that cell growth peaks at lower cell densities for 824(p94_solB) and 824ΔsolB than for the other strains; this stunted growth is not surprising, given the abnormally high levels of acid metabolites, which are known to inhibit cell growth (Papoutsakis, et al, 1987; Borden, et al, 2010).

It is most unusual that overexpression and deletion alike of *solB* result in the same solvent-negative phenotype. Since this outcome was not the anticipated

phenotype for the deletion strain, cultures of 824ΔsolB were grown in media lacking thiamphenicol pressure in an effort to examine antibiotic selection as a factor in the unexpected results; however, all three biological replicates that were tested retained the solvent-negative phenotype (data not shown). It is evident that *solB* affects solventogenesis, but apparently through a much more complex interaction than was initially hypothesized.

4.4 Overexpression, not Deletion, of *solB* Results in Decreased Transcript Levels of *sol* Locus Genes

To investigate further the effects of *solB* overexpression and deletion on solventogenesis and to possibly determine the molecular mechanism responsible for the surprising and perplexing metabolite data, transcript levels of three *sol* locus genes – *adhE1*, *ctfA*, and *adc* – were measured in all five experimental strains. The transcript for *ctfB*, although also involved in solventogenesis, was not tested because it is known to be coupled with *ctfA* both as a polycistronic transcript and as the second subunit of the CoA transferase enzyme; therefore, expression of *ctfA* was deemed to be representative of both *ctfA* and *ctfB* expression. Relative transcript levels of the *sol* locus genes were measured by Northern analysis and QPCR in ATCC 824, 824(p94MCS), 824(p94_solB), 824ΔsolB, and solRB. Preliminary Northern blots revealed strong expression of all genes in both ATCC 824 and 824(p94MCS) samples (see Appendix B); to conserve gel lanes, only ATCC 824 samples were included for comparison in all subsequent Northern blots.

Transcription of *adc* peaks at approximately 12 hours, as the cell enters stationary phase and begins solventogenesis, and then steadily decreases throughout stationary phase (Gerischer and Durre, 1992). This expression pattern was indicated in

Northern blot analysis by the *adc* bands in ATCC 824 samples (Figure 4.11a). Bands of comparable size and intensity are seen in solRB as well as 824ΔsolB samples, indicating that *adc* transcription occurs in these strains at levels relatively unchanged from wild type. However, a considerably lighter band appears in 824(p94_solB) samples, corresponding to considerably lower transcript expression. Expression data for *adc* from QPCR (Figure 4.11b) follows this same trend for 824(p94_solB), namely, reduction in *adc* expression of about 5-fold at six and 24 hours and 2-fold at 12 hours relative to the plasmid control strain. Ct values and fold change calculations for both 824ΔsolB and solRB, however, show a trend of slightly higher *adc* expression in these strains relative to wild type. Additionally, these fold changes, particularly the six-fold changes at six hours, suggest that transcription of *adc* may be induced earlier in 824ΔsolB and solRB. Overall, *adc* is expressed at higher levels than wild type both in the solRB strain and upon *solB* deletion, but it is down-regulated upon *solB* overexpression.

A similar but not identical pattern was observed in the analysis of the *adhE1* transcript (Figure 4.12). The *solB* overexpression strain showed undetectable levels of *adhE1* transcript on the Northern blot (Figure 4.12a) as well as higher Ct values and reduced expression by 20-fold and 6.5-fold at six and 12 hours, respectively, according to QPCR analysis (Figure 4.12b-c). Like the wild type, bands are present on the Northern blot in 824ΔsolB and solRB samples, indicating *adhE1* expression in these strains. Although the *adhE1* bands in 824ΔsolB appear less strong than those in solRB or the wild type, the transcript is nevertheless present and in clear contrast with the 824(p94_solB) samples. Ct values and fold change calculations revealed further insight into *adhE1* transcript levels in both solRB and 824ΔsolB (Figure 4.12b-c). In a

partial divergence from the trends in *adc* expression, levels of *adhE1* in 824ΔsolB seem to fluctuate, from a 3-fold increase from wild type levels at six hours to a 2-fold reduction at 12 hours to no change from wild type levels at 24 hours. This is juxtaposed with the positive 16-fold and 2-fold changes in solRB at six and 12 hours, respectively, before an 8-fold reduction in *adhE1* expression at 24 hours. Again, the increased expression seen in both strains at six hours supports the possibility of earlier expression. It is unclear why *adhE1* expression seems to fluctuate in 824ΔsolB; additional replicates and testing will be necessary to more finely characterize this expression pattern. Nevertheless, when viewed within the overall context of transcript expression at all time points and in Northern blot analysis, *adhE1* is clearly expressed, albeit erratically, in the *solB* deletion strain, in sharp contrast with its significant reduction in 824(p94_solB), despite yielding nearly identical metabolic phenotypes.

Of particular note is that in all three strains exhibiting *adhE1* expression, two bands appear on the Northern blot, with both appearing larger than 1000 base pairs (Note: in these blot images, bands at or above the 1000-bp mark are compressed and therefore are greater than or equal to 1000 base pairs in size). One represents the expected polycistronic *sol* operon transcript (3973 bp) and the other, smaller band likely corresponds to a smaller transcript containing only the *adhE1* gene (2589 bp). In recently conducted RNA-Seq studies, the 63-bp intergenic region between *adhE1* and *ctfA* revealed sustained low coverage (Ralston, personal communication), and upon further examination, a previously undiscovered, putative terminator sequence was identified within the first 40 base pairs of the *ctfA* coding sequence (Ralston, personal communication). This terminator had not been previously identified due to its location within a coding region, but the presence and relative sizes of the two bands seen in

Northern blot analysis suggest that it is both present and active, generating the originally identified polycistronic transcript as well as a truncated transcript containing only *adhE1*. The ability to express *adhE1* without *ctfAB* or at a higher ratio adds greater complexity to the expression of the *sol* locus and might enable the cell to more meticulously regulate enzyme production and metabolism under certain conditions or in response to changes in the environment or substrate or metabolite concentrations.

The analysis of *ctfA* transcript levels exhibited similar expression to the previous two gene transcripts. The Northern blot analysis of *ctfA* generated weaker results than the previous blots, and although the images in Figure 4.13a are difficult to interpret with total confidence, they do reveal the presence of *ctfA* transcript, albeit faint, in wild type, solRB, and 824ΔsolB samples, but no clearly visible transcript for 824(p94_solB). In full support of these assessments (and in agreement with the observations for *adhE1* expression), Ct values and fold changes for *ctfA* from QPCR analysis (Figure 4.13b-c) indicate substantial reduction in transcript levels in 824(p94_solB) at six and 12 hours, no significant changes from wild type for 824ΔsolB, and positive fold changes for solRB at six and 12 hours followed by a 2-fold reduction at 24 hours.

As seen in the Northern blot analysis of *adhE1* expression, two faint bands are present in each strain showing *ctfA* transcription; this is particularly evident in the 12 hour solRB sample of the lower image in Figure 4.13a. In addition to the terminator sequence previously described, a putative promoter (5'-CTTCAT[15bp]TATAAT-3') nearly identical to the consensus sequence has been discovered within the short intergenic region between the *adhE1* and *ctfA* coding regions (Ralston, personal communication). The two bands shown by Northern analysis, given their relative

sizes, likely correspond to the polycistronic *sol* operon transcript (3973 bp) and a smaller transcript originating at this additional promoter and consisting only of the *ctfA* and *ctfB* genes (1324 bp total). Although additional Northern blot analyses will be needed to verify this conclusion, these results suggest that this sequence is an active promoter and that these genes are transcribed both as part of a full *sol* operon transcript and as a separate transcript.

It is important to note that in the *solB* overexpression strain, the *sol* operon genes are being down-regulated, but likely are not fully absent; low expression of the genes is probably still occurring. This is evident in the raw Ct values, which are significantly lower in 824(p94_solB) yet visibly larger than the nonspecific amplification values of the non-template controls (Figures 4.11b, 4.12b, 4.13b).

The results presented here suggest that *solB* overexpression produces the observed solvent-negative phenotype by decreasing transcript levels of the *sol* locus genes, either through blocking their transcription or by targeting the mRNA transcript for degradation. However, it is intriguing that, although solventogenesis is absent with *solB* deletion as well, significant, sustained reduction in transcript levels are observed only upon *solB* overexpression; the transcript levels differ widely between the 824(p94_solB) and 824ΔsolB strains, but the metabolite profile remains the same. It appears that while overexpression of *solB* represses solventogenesis through decreasing transcript levels, deletion of *solB* generates the same metabolic phenotype but with a different effect on transcript levels. Furthermore, it is fascinating that *sol*-operon transcript levels are comparable in both 824ΔsolB and solRB (Figure 4.14), considering that solvent production is negligible in the former but amplified in the

latter. This further supports a much more complex regulation regarding the role of *solB* in solventogenesis.

4.5 Expression Patterns of the *sol* Locus Genes in 824ΔsolB during Early Growth may Suggest a Change in Induction Time

The fold change calculations for *sol* locus expression at six-hours in 824ΔsolB suggest an earlier induction of transcription relative to wild type. Therefore, RNA samples were collected during early growth – between two and 12 hours – and were examined by QPCR to explore a potential early induction of *sol*-locus genes in this strain and to investigate any possible influence of *solB* overexpression and deletion on transcription during early growth.

As was observed in the previous QPCR analyses, expression of *adc*, *adhE1*, and *ctfA* at early time points was lower in 824(p94_solB) (Figure 4.15); Ct values and fold change calculations show a reduction in transcript expression that increases over time as expression of the *sol* locus genes begins and increases in the wild type but remains inhibited in 824(p94_solB) (Figure 4.15).

In the 824ΔsolB samples, transcript levels for all three genes during early growth are considerably higher at two and three hours relative to wild type, followed by insignificant fold changes in expression starting at four hours (Figure 4.15). The relative Ct values and fold change calculations for six and 12 hours reveal comparable levels of transcript expression suggested by the previous QPCR experiments and Northern blot analyses, but the larger fold changes at two hours for *adc*, *adhE1*, and *ctfA* (37-, 156-, and 125-fold, respectively), and to a lesser extent three and four hours, do suggest an early induction of transcription. While these results may indicate early induction, the values also may simply be error propagated from differences in early

growth rates between wild type and 824ΔsolB or due to variations in very early growth between replicates. These results, and the apparent possible statistical and temporal variance between samples (i.e. the large standards of error for the wild type Ct values) indicate the possibility of early induction of *sol* locus expression in the 824ΔsolB strain but also that additional analysis will be needed to assist in validating these observations.

4.6 Expression of *solB* is Minimal yet Detectable in ATCC 824

In addition to the genes of the *sol* locus, expression of *solB* during early growth was examined by QPCR. Overexpression and deletion of *solB* in 824(p94_solB) and 824ΔsolB, respectively, were validated strongly at these early time points (Figure 4.16). From WT samples, Ct values indicate that *solB* expression is very low but detectable at early time points (Figure 4.16), as was seen in the time points of the previous analysis (Figure 4.7). A slight increase in *solB* expression at six and 12 hours relative to the earlier time points suggests a possible pattern of increasing expression as growth progresses, but even considering this small increase, expression of *solB* remains comparatively low according to these QPCR data. This minimally detectable *solB* expression matches both the previous Northern blot analysis and QPCR (Figures 4.6 and 4.7, respectively) and what has been consistently observed in standard and strand-specific RNA-Seq experiments (Appendix A). It is striking that an RNA molecule that is expressed at such low levels appears to influence solvent production and the expression of the solventogenic genes.

4.7 Transcriptional Expression of the Master Regulator *spo0A* was not Affected during Early Growth in *solB* Overexpression and Deletion Strains

As previously mentioned, two possible explanations for the decrease in *sol*-operon expression seen in 824(p94_solB) are, first, that *solB* targets the transcripts for degradation or, second, that *solB* blocks their transcription. Translational repression by target degradation is a common regulatory mechanism for numerous prokaryotic sRNA (Desnoyers, et al, 2013), but transcriptional repression may result indirectly from a trans-acting sRNA repressing the translation of a transcriptional activator to block the transcription of its target genes.

The transcription factor *spo0A* is a master regulator in sporulating clostridia and all endospore formers; it regulates the initiation of major stationary-phase pathways and cellular processes like granulose production, the sporulation cascade, and solventogenesis (Tracy, et al, 2011). *spo0A* has been demonstrated in previous studies to be an essential transcriptional activator of the *sol*-locus genes (Harris, et al, 2002; Durre, et al, 2002) and has been examined as a possible target for *solB* regulatory activity (Zimmermann, 2013, dissertation). Therefore, of particular interest was to investigate the possibility that *solB* represses transcription of the *sol*-locus genes indirectly by interacting with the *spo0A* mRNA and reducing its concentration.

To examine a possible effect of *solB* on *spo0A* transcript levels, the *spo0A* transcript was assayed by QPCR at several early time points of exponential growth. In the WT strain, *spo0A* is expressed briefly in early exponential phase and then again beginning in middle to late exponential phase when it activates transcription of its stationary-phase targets, thereby initiating the transition phase. (Jones, et al, 2008) Results from the QPCR analysis of *spo0A* indicate that in both 824(p94_solB) and 824ΔsolB, *spo0A* transcript levels remain nearly identical to expression levels in the

WT strain, as shown by comparable Ct values as well as insignificant fold changes at each time point (Figure 4.17). These results indicate that *solB* does not affect the levels of *spo0A* mRNA in ATCC 824.

It is still possible, however, that *solB* may affect Spo0A protein levels, i.e. by regulation at the translational level. Although translational down-regulation by regulatory sRNA usually occurs via targeted transcript degradation, there are examples of sRNA-mediated translational down-regulation without degradation of the transcript, involving instead the occlusion of the ribosomal binding site, as with GcvB or Spot-42 in *E. coli* (Desnoyers, et al, 2013). Western blot analysis of *spo0A* protein level expression in 824ΔsolB, 824(p94_solB), and wild type would effectively test this hypothesis and is currently underway.

4.8 Expression of *solR* is Significantly Reduced in all Recombinant Strains

The relative expression of the *solR* gene was examined in all recombinant strains, due to its role in previous studies of the regulation of solventogenesis. Northern blot analysis showed no significant or detectable levels of *solR* transcript in wild type, 824(p94_solB), 824ΔsolB, or solRB (Figure 4.18a). Ct values and fold change calculations from QPCR analysis revealed average transcript levels in wild type but significant reduction in *solR* expression in all three recombinant strains (Figure 4.18b), but especially in 824(p94_solB). Expression of *solR* at early time points (Figure 4.18c) followed the same pattern of *solR* expression reduction in both 824(p94_solB) and 824ΔsolB, although fold changes in the latter were less prominent than at later time points. Although previous studies have intimated that *solR* is not involved in the regulation of solventogenesis, it appears that its transcription may somehow be affected by the overexpression and deletion of *solB*; interestingly, these

solB recombinant strains appear to show lower *solR* expression than the solRB strain. It may be significant that a stem loop sequence between the *solR* and *solB* coding regions was truncated in 824ΔsolB, possibly inhibiting the completion of successful transcription of *solR*. However, the effects of *solB* overexpression are unclear. Additional study will be necessary to elucidate any connection of *solR* to the *solB* regulatory system.

4.9 Knockdown of *solB* Expression by Antisense RNA also Results in the Loss of Solvent Production

The unexpected solvent-negative phenotype of the *solB* deletion strain may potentially be only the result of frameshift disruption of the ORFs of the downstream *sol* locus, the deletion or disruption of regulatory sequences surrounding the deleted *solB* region (Thormann, et al, 2002; Durre, et al, 2012), or any side effects from altering the cell's genomic DNA. Since the replacement of the *solB* region with the thiamphenicol selection cassette was substantially upstream of both the *sol* locus and its promoter elements, no frameshift mutation of those ORFs is possible. Also, the Sanger sequencing data from 824ΔsolB indicate that the recombination events left all downstream transcriptional elements intact. Furthermore, as previously mentioned, 824ΔsolB cultures were grown without thiamphenicol selection and exhibited no change in solvent phenotype, suggesting that the antibiotic pressure is not a factor in the attenuated solvent production.

A variety of regulatory sequences and transcription factor binding sites surround the *solB* sequence (Thormann, et al, 2002; Durre, et al, 2012). It could be possible that one or more of these sequences may have been truncated or deleted along with the *solB* sequence in 824ΔsolB. Alternatively, the allelic exchange of the

thiamphenicol resistance marker for *solB* may have somehow allosterically interfered with activity of one or more of these nearby regulatory regions in such a way that *sol* locus transcription was altered without being inhibited; for instance, generation of damaged or altered transcripts incapable of successful translation might impede solvent production regardless of successful transcript expression.

To explore the effects of *solB* removal through a different and less invasive approach, as well as to examine the possibility of interference to neighboring sequences in 824ΔsolB, two *solB* knockdown strains were developed utilizing an antisense RNA strategy. The first strain, 824(p94_as(solB)F), expresses an antisense RNA complementary to the full *solB* sequence, and the second strain, 824(p94_as(solB)S), expresses a shorter antisense RNA complementary to the first 100 base pairs of *solB*. Metabolite analysis by HPLC of triplicate 72-hour cultures of both knockdown strains yielded total loss of solvent production, even more severe than what was observed with the *solB* overexpression and deletion strains (Table 4.2). No butanol or acetone were detected in either knockdown strain, compared to only 5.5 mM butanol and 0 mM acetone in 824ΔsolB. Acid concentrations are also within range of the values for both 824(p94_solB) and 824ΔsolB, and glucose consumption in both knockdown strains was similarly low. These results indicate that a severe solvent-negative phenotype is observed whether *solB* is removed by antisense RNA knockdown or total deletion of the coding sequence from the genome. Additionally, this experiment supports the conclusion that the phenotypic effects on solventogenesis are the result of the loss of *solB* expression and not the loss or interference of a peripheral regulatory sequence.

4.10 Transcript Secondary Structures and *solB* Binding Predictions Indicate very Favorable Regulatory Interactions

As seen through metabolite and transcriptional analyses, *solB* impacts solventogenesis and the *sol*-locus genes in some complex regulatory manner. To elaborate on these findings and consider potential mechanisms of interaction, a variety of investigations of the transcripts and their 5' untranslated regions (UTRs) were conducted. As mentioned in Chapter 1, primer extension (Nair, et al, 1994) and RNA-Seq experiments (Ralston, unpublished data) have identified a distal and a proximal promoter for the *adhE1* gene, with transcriptional start sites 243 and 83 base pairs upstream of the first codon, respectively. Several studies (Nair, et al, 1994; Scotcher, et al, 2003) have investigated the roles and relationship of these two promoters, offering the clearest conclusion that most transcription initiates at the proximal promoter. Northern analysis was conducted using oligo DNA probes complementary to the 5' UTRs corresponding to each promoter to investigate whether any dynamic between the two promoters or their respective UTRs contributes to the phenotype observed for *solB* overexpression and deletion (see Appendix C). Although the results were largely inconclusive, they do suggest that activity involving the 5' UTR of the *adhE1* transcript may be influenced by *solB* deletion and overexpression.

To continue these investigations, putative secondary structure folding was predicted for the 5' UTR of each transcript of the *sol* locus using the Vienna RNA Websuite (Gruber, et al, 2008). Also, predicted binding interactions between *solB* and the UTR of each transcript were generated using the IntaRNA online prediction program (Busch, et al, 2008; Wright, et al, 2014). These predictions were performed to identify possible molecular interactions between *solB* and each of the genes

responsible for solventogenesis and to explore possible mechanistic details of these regulatory interactions.

The secondary structure folding predictions, which consisted of the 5' UTR and first 100 base pairs of the gene's coding region, revealed significant stem loop motifs in both *adhE1* and *adc* transcripts (Figures 4.19a and 4.20a). Large, negative free energy values associated with both structure predictions (-23.7 kcal/mol for *adhE1*, -26.8 kcal/mol for *adc*) indicate a strong probability that these transcript UTRs exhibit these folding motifs naturally. Additionally, the predicted folding of both transcripts occludes the ribosomal binding site, suggesting a regulatory role for the secondary structure of these transcripts, particularly a potential need for translational activation.

The predicted interactions of *solB* with the UTRs of *adhE1* and *adc* as well as the intergenic region between *adhE1* and *ctfA* revealed strong putative binding between *solB* and each transcript, as indicated by large negative free energy calculations. For each binding prediction, the hybridization energy reflects the free energy associated with the interaction of the paired sequences only, while the overall free energy accounts for the effects on the free energy of the entire transcripts and their secondary structures. For these predictions, the full *solB* coding sequence was queried against the full 5' UTR sequence (or intergenic region) plus the first 50 base pairs of the gene. The predicted binding of *solB* with the proximal UTR of the *adhE1* transcript exhibits a predicted hybridization energy of -14.8 kcal/mol and overall free energy of -4.09 kcal/mol and consists of highly analogous base-pairing, as shown in Figure 4.19c. The *solB*-binding region of the *adhE1* UTR includes the ribosomal binding site and pairs with a strongly predicted central region of *solB* (Figure 4.21a).

On the *adc* transcript UTR, the same central region of *solB* interacts with near perfect identity with a sequence 39 base pairs upstream of the *adc* ribosomal binding site (Figure 4.20), with a hybridization energy of -11.2 kcal/mol and overall free energy of -4.40 kcal/mol. The intergenic region between *adhE1* and *ctfA* was also tested for interaction with *solB*, considering the terminator, promoter, and ribosomal binding site identified in this 63-base pair region (Figure 4.21). Again, strong interaction with *solB* was predicted for a sequence within the *ctfA* coding region 93 base pairs downstream of the ribosomal binding site, showing a hybridization energy of -11.5 kcal/mol and overall free energy of -6.69 kcal/mol; however, this binding prediction involves a different yet still strongly predicted region of *solB* secondary structure (Figure 4.21a).

These very strong predicted interactions of each transcript with *solB*, combined with the secondary structure predictions for each transcript's UTR, supports the existence of a regulatory function for *solB*. Specifically, the binding predictions indicate that *solB* pairing with each transcript is not only very likely but also feasibly involves the ribosomal binding site of each transcript; combined with the transcripts' secondary structure predictions, which indicate a clear possibility of stem loop motifs in the 5' UTRs that block the ribosomal binding site and therefore necessitate translational activation, these predictions reinforce the existence of a molecular interaction between *solB* and the *sol* locus, one that potentially involves translation. Western blot analysis targeting the *sol* locus enzymes (i.e. AADC, AdhE1, CoAT) in ATCC 824 and in 824ΔsolB and 824(p94_solB) will determine the possibility and existence of this interaction.

4.11 Conclusions on the Possible Regulatory Influence of *solB* in ATCC 824

It is clear from the metabolite data and transcript analyses that *solB*, although minimally expressed in wild type cultures, plays a complex regulatory role in solventogenesis in ATCC 824. It has been shown that overexpression of *solB* results in attenuation of solvent production paired with a decrease in *sol* locus transcript levels. Deletion of *solB*, by allelic exchange and antisense RNA knockdown alike, generates the same solvent-negative phenotype without a significant change in *sol* locus transcription. The specific mechanism by which *solB* affects solventogenesis ultimately remains unknown, but this study has presented considerable results to further this ongoing discussion. With these findings, several speculations can be made on the molecular function of *solB* and will serve to encourage and guide further study.

The solvent-negative phenotype and decrease in *sol* locus transcripts seen in the *solB* overexpression strain 824(p94_solB) provide strong evidence for *solB* as a repressor of solventogenesis; the loss of solvent production is likely due to loss of *sol* locus transcripts, and *solB* may stimulate this outcome by either blocking *sol* locus transcription (via blocking translation of *spo0A* transcripts) or by targeting the *sol* locus transcripts for degradation. Strong interaction between *solB* and its target transcripts, particularly at their ribosomal binding sites, supports this mechanism, since repression at such sites has been documented for many sRNA repressors (Desnoyers, et al, 2013). However, this function as a simple repressor of the *sol* locus genes is contradicted by the solvent-negative phenotype in the *solB* deletion strain. It is possible that another regulatory molecule works in conjunction with *solB*, and in the absence of *solB* this partner overcompensates, generating the observed phenotype. Alternatively, it must be considered that *solB* may act instead as an activator, rather than a repressor, of the *sol* locus. Given the stem loop structures predicted in the UTRs

of each *sol* locus transcript, *solB* may interact with these UTRs, as predicted, but to free the ribosomal binding sites and enable translation initiation, rather than to block translation and facilitate target degradation. This mechanism of translational activation by regulatory sRNA has also been well-documented in prokaryotes (Desnoyers, et al, 2013), but again, while this speculated role for *solB* would explain the lack of solvent production and preserved transcript levels observed for *solB* deletion, it does not account for the same solvent-negative phenotype in *solB* overexpression.

A second possible explanation may lie in a bimodal function for *solB* dependent on its concentration in the cell. At low levels, *solB* may act as an activator of translation for the *sol* locus genes, as described above, but at high expression levels, *solB* might facilitate the degradation of its target transcripts. Low levels of *solB* may be necessary in the cell to activate the translation of the *sol* locus transcripts, but as its expression and binding increases, it begins to titrate out its target transcripts and increase their degradation. This hypothesis explains the solvent-negative phenotype observed for both *solB* overexpression and deletion; furthermore, it agrees with the slight increase in *solB* expression seen in wild type QPCR data – the minimal *solB* expression levels seen through 12 hours of growth would facilitate rapid translational activation of the sol locus genes as their transcription occurs, and the increase in *solB* expression at later time points (i.e. 24 hours) might facilitate the degradation of these transcripts once they have been translated and are thus no longer necessary. It may be interesting to perform QPCR targeting *solB* expression into very late stationary phase growth to more clearly observe this potential expression pattern; however, RNA degradation in late growth cultures of ATCC 824 may limit the success and effectiveness of such an experiment.

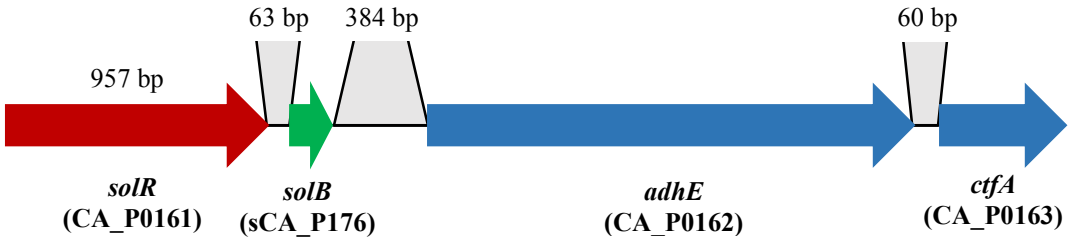


Figure 4.1 – Genomic location and organization of *solR*, *solB*, and the first *sol* locus genes in *C. acetobutylicum*. Shown are coordinates 174,190 to 179,115 of the pSOL1 megaplasmid coding strand.

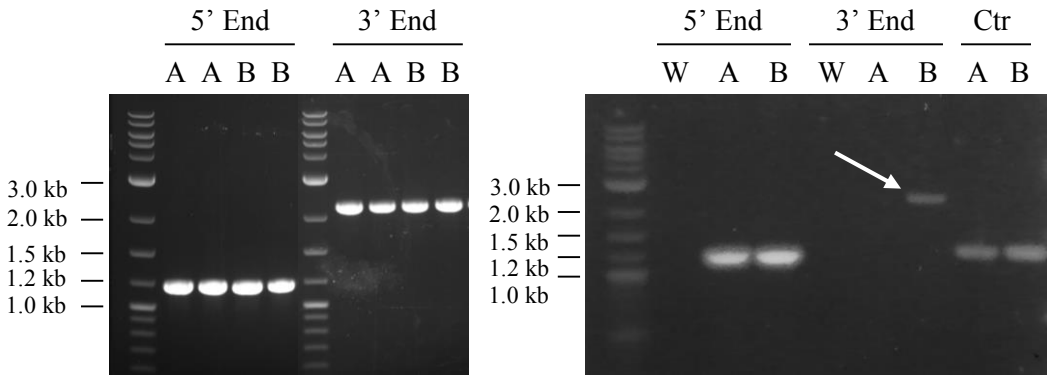


Figure 4.2 – PCR confirmation of *solR* disruption regions in two replicate solRB cultures (A and B). The 5' end of the disruption (~1.2 kb) was amplified with primers *solR453 Seq F* and *Tc239 Seq R*, and the 3' end (~2.2 kb) amplified with *Em373 Seq F* and *solR1361 Seq R*; a central region (Ctr), indicating the presence of two copies of the pO1x suicide vector, was amplified with *Em373 Seq F* and *Tc239 Seq R* (see Figure 4.2); the templates for all reactions are genomic DNA preps from ATCC 824 (W) and two solRB replicate cultures (A and B); the absence of product for a 3' end reaction indicates a solitary failed reaction, given the successful amplification of the same region in the additional replicate reactions.

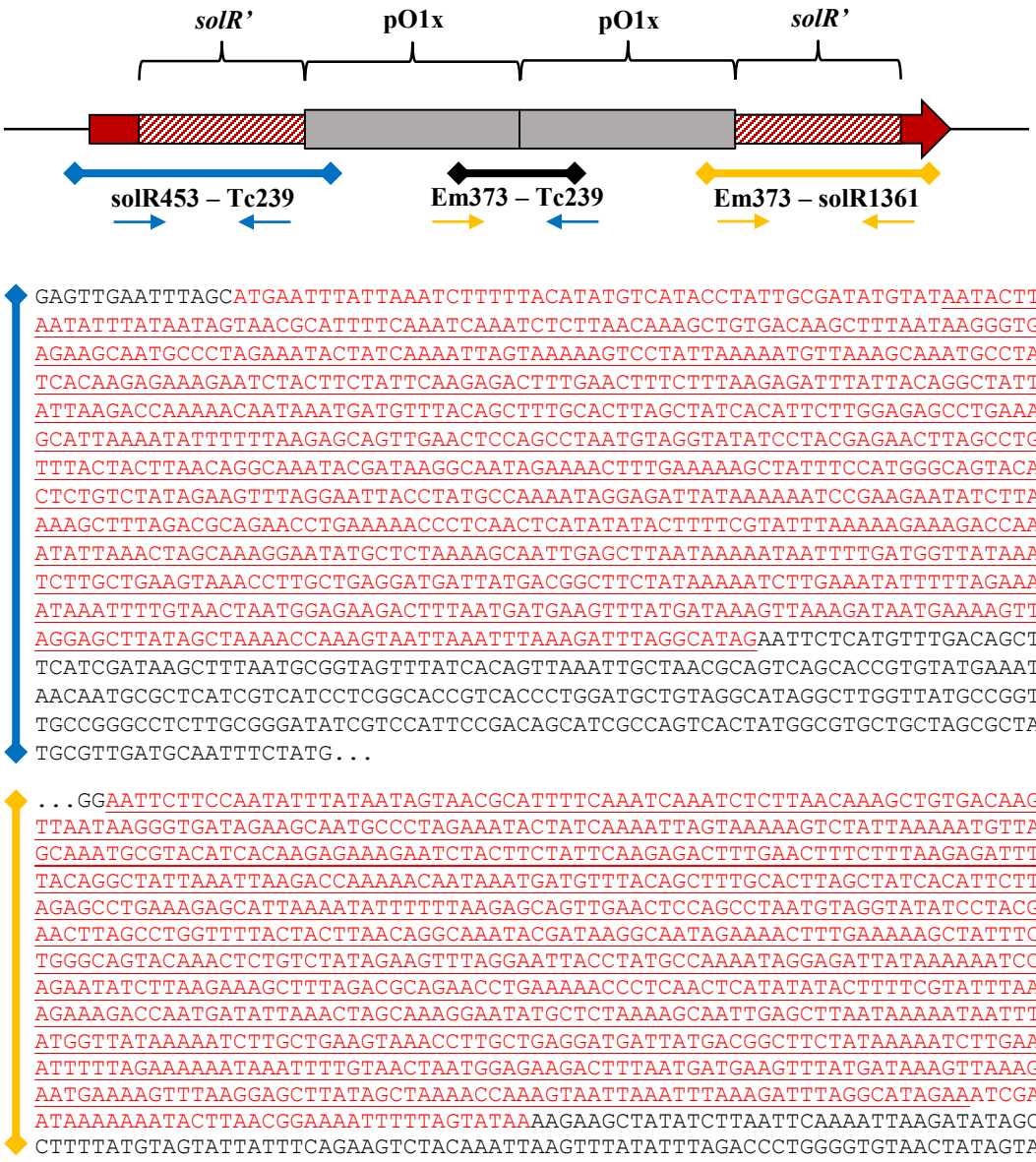


Figure 4.3 – Sequencing schematic of *solR* disruption in the solRB strain. The *solR* coding region is shown in red, with the region of homology (*solR'*) used in the pO1x suicide vector shown striped or underlined; the 5' and 3' crossover events (blue and orange bars, respectively) as well as the central region spanning both vector copies were amplified and sequenced (central region sequencing not shown).

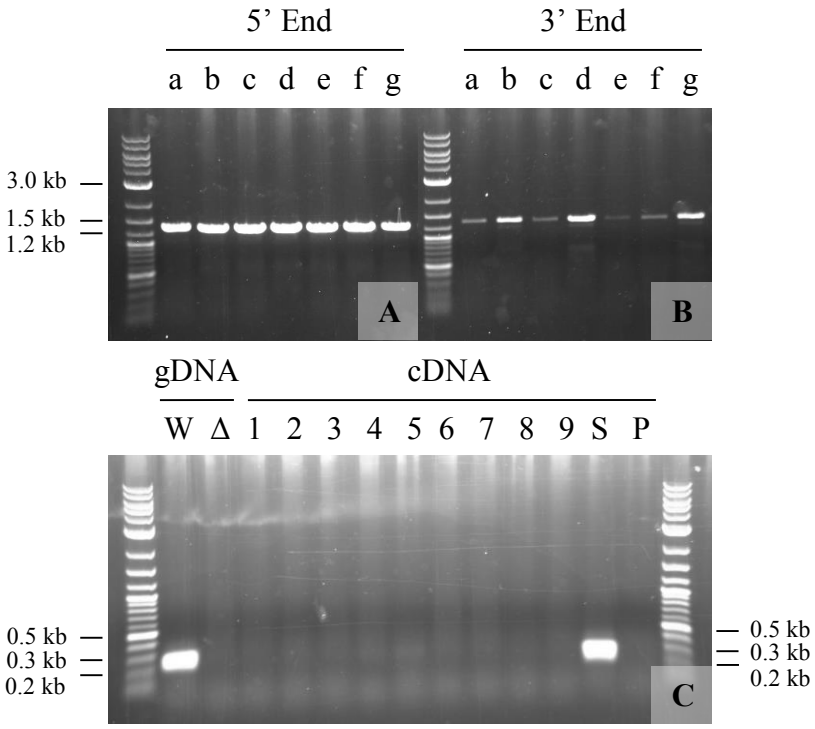


Figure 4.4 – Agarose gels showing PCR confirmation of 824ΔsolB. Amplification products of the (A) 5' and (B) 3' integration sites of the Th^R cassette in genomic DNA preps from 824ΔsolB replicate strains a-g (strain b was chosen for further study); PCR amplification of *solB* sequence (C) in genomic DNA preps of ATCC 824 (W) and 824ΔsolB strain B (Δ), as well as cDNA generated from 3 replicate cultures of 824ΔsolB (1,4,7 = 6 hours; 2,5,8 = 12 hours; 3,6,9 = 24 hours) and cDNA from exponential phase cultures (6 hours) of 824(p94_solB) (S) and 824(p94_MCS) (P); p94MCS is the empty overexpression vector, transformed and used as plasmid control strain.

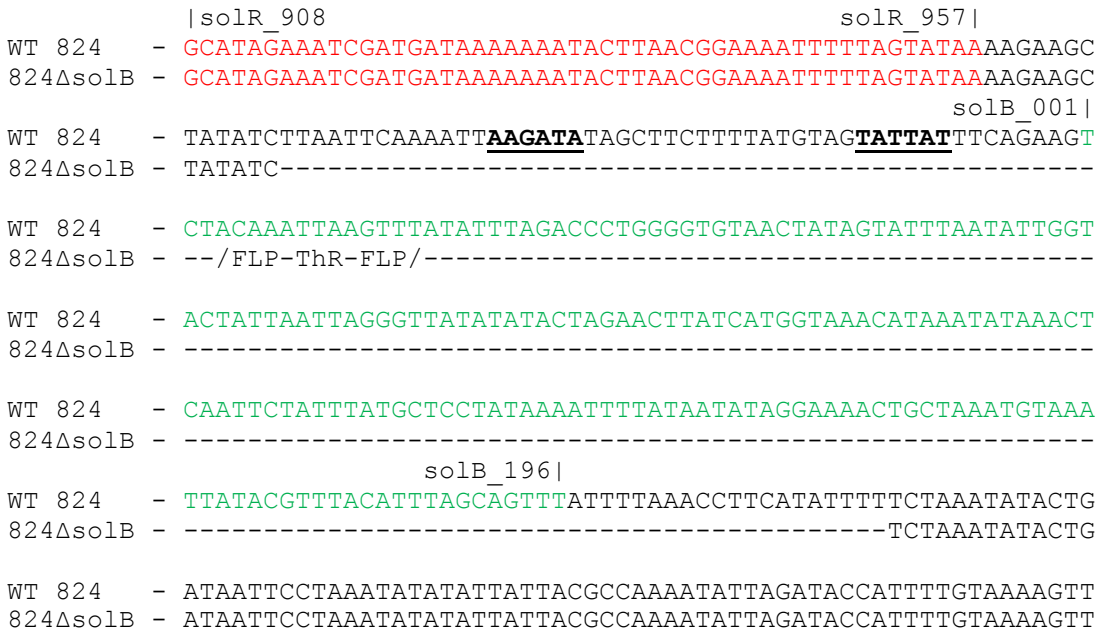


Figure 4.5 – Sequencing of the *solB* region in 824ΔsolB, aligned with the ATCC 824 genome. Deletion of *solB* also included the removal of 50 bp upstream, including its sigA promoter consensus sequence (bold and underlined) (Zimmermann, 2013, dissertation), and 20 bp downstream, but did not interfere with any previously identified regulatory sequences, the distal promoter of the sol operon, or the *solR* coding region (in red); *solB* sequence in ATCC 824 genome is shown in green; FLP-ThR-FLP indicates the 1169-bp thiamphenicol resistance cassette, containing flippase recognition sequences, inserted in the 824ΔsolB genome.

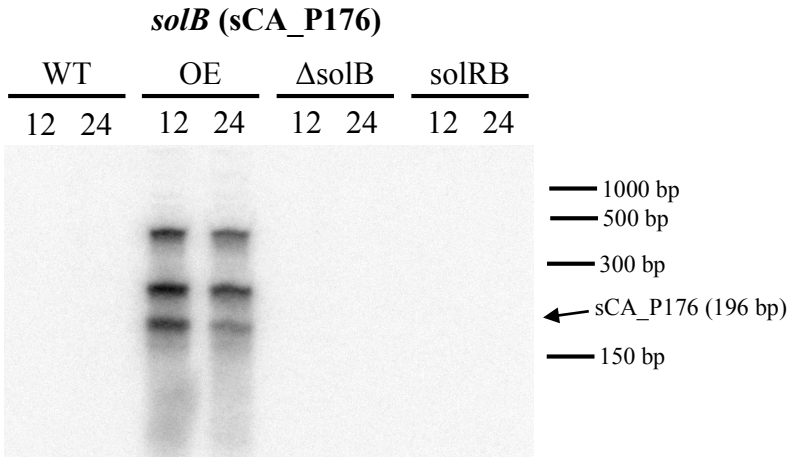


Figure 4.6 – Northern blot analysis of *solB* expression in ATCC 824, 824(p94_solB) (OE), 824ΔsolB, and solRB strains. Each sample contained five µg of total RNA; larger bands are due to additional promoter activity on the vector backbone.

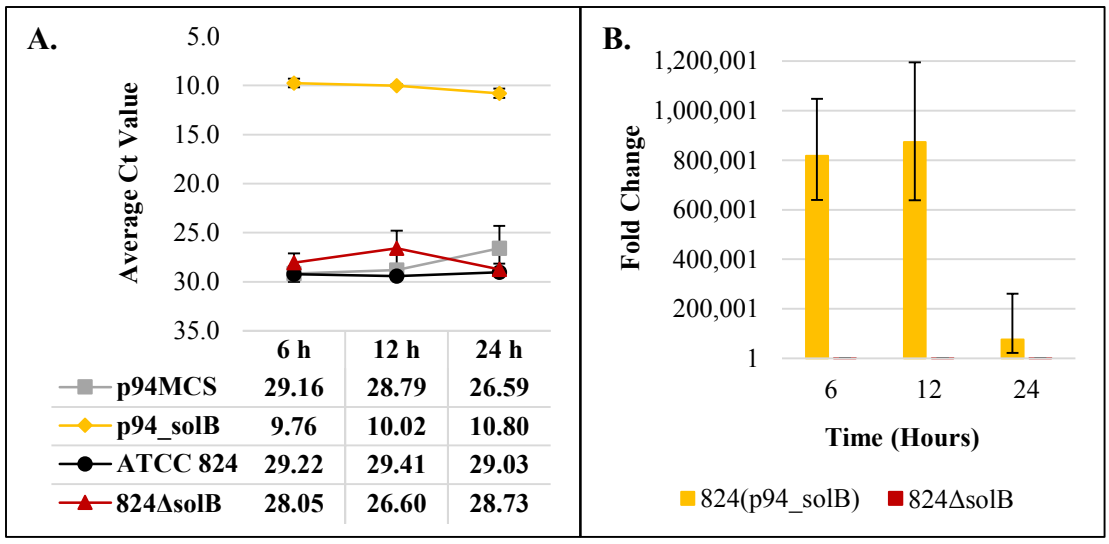


Figure 4.7 – Ct values and fold changes in *solB* expression in 824(p94_solB) and 824ΔsolB. Fold changes in the overexpression strain and knockout strain are relative to plasmid control and wild type, respectively, with *solB* expression normalized to the constitutively expressed housekeeping gene CA_C3571; average Ct value for *solB* in non-template controls is 31.96 ± 0.41; error bars in both graphs represent standard deviation (n = 3).

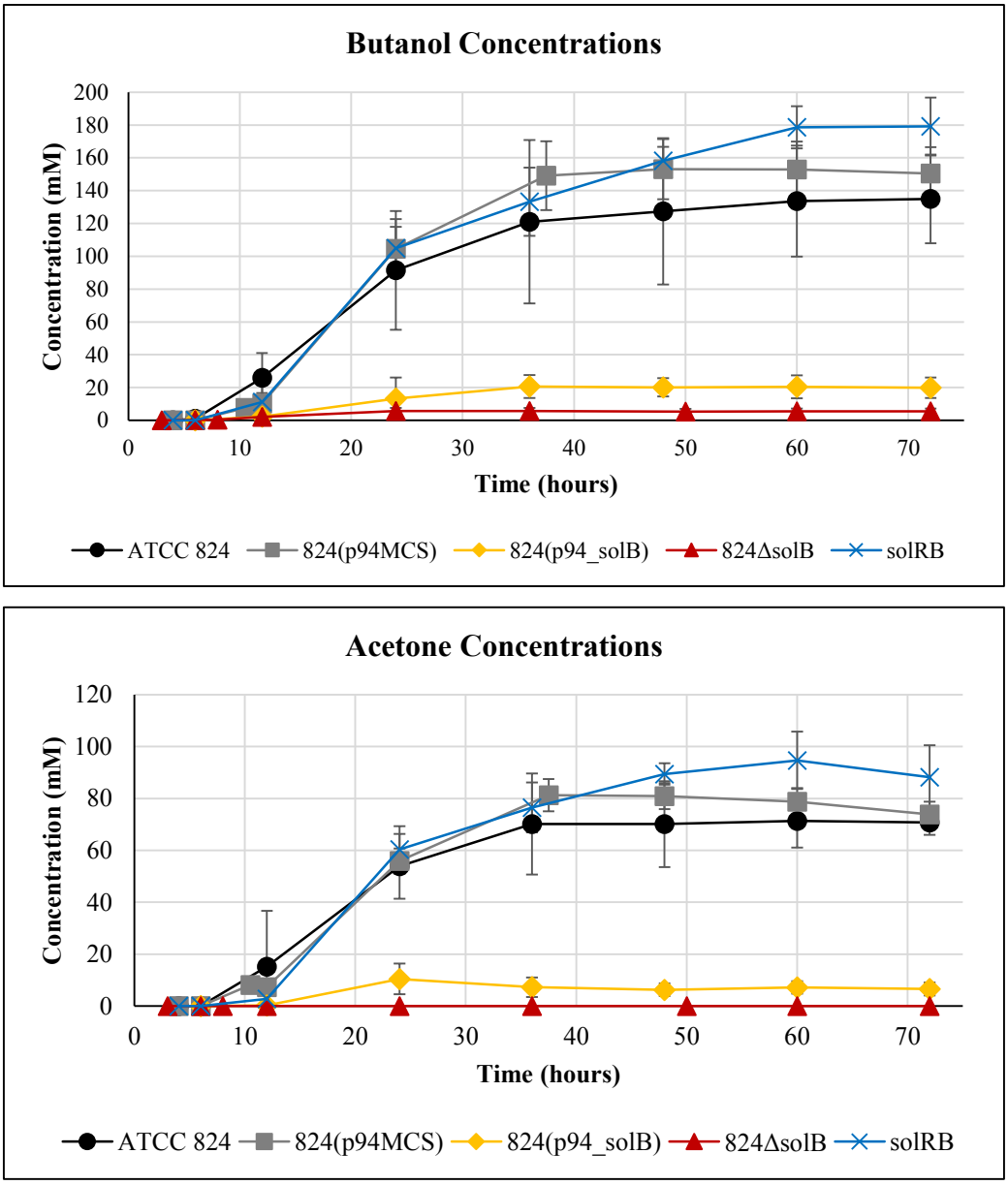


Figure 4.8 – Solvent production in recombinant strains during 72 hours of static flask growth (40 mL). Concentration data are listed in Table 4.1; error bars represent standard deviation (n = 3).

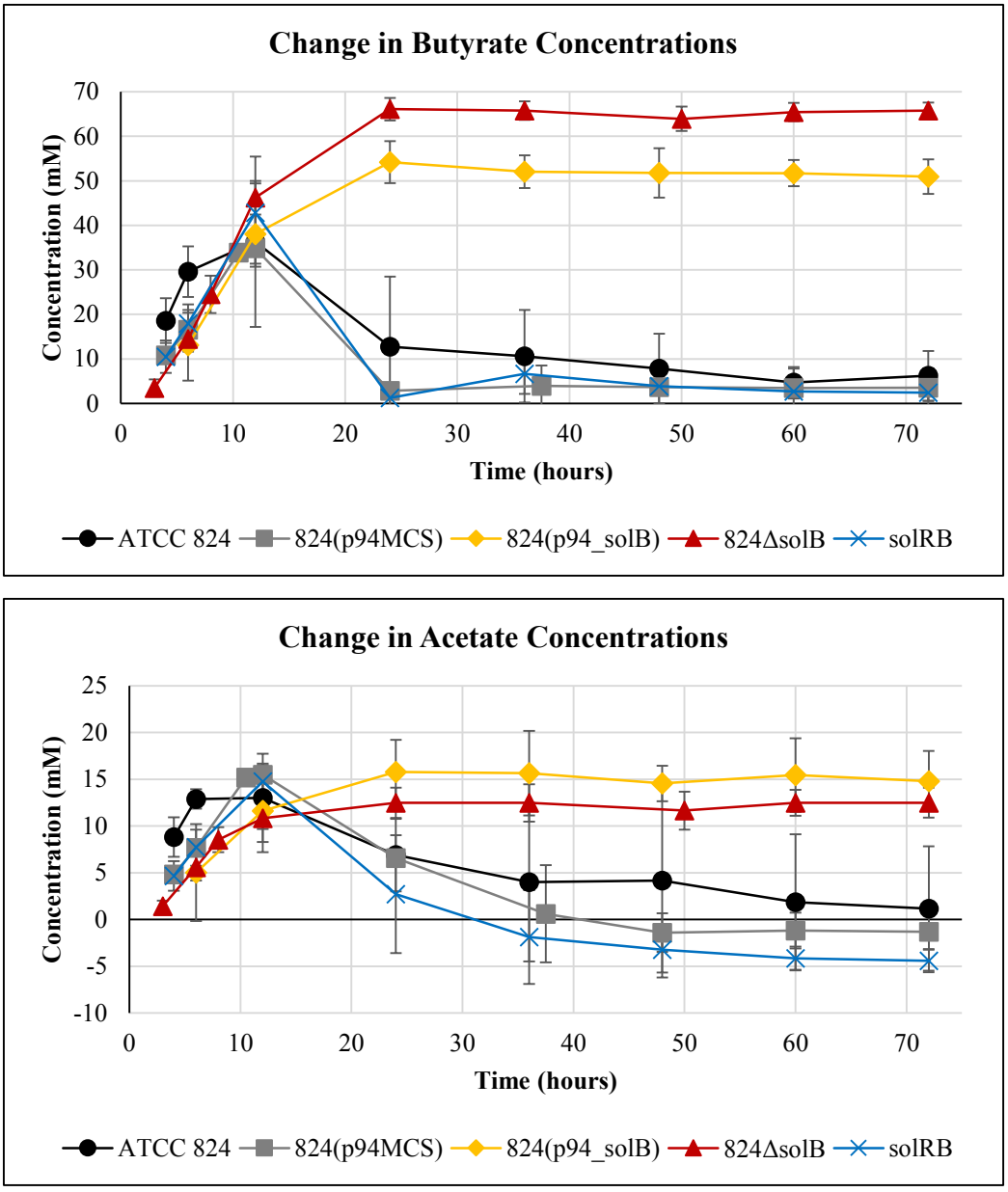


Figure 4.9 – Changes in acid concentration in recombinant strains during 72 hours of static flask growth (40 mL). Concentration data are listed in Table 4.1 and were calculated relative to the first time point (t_0); error bars represent standard deviation (n = 3).

Table 4.1 – Final metabolite production and glucose utilization in recombinant strains and ATCC 824 after 72 hours of static flask growth (40 mL). Values represent change in concentration relative to the first time point (t_0), except for glucose values, which represent total amount consumed after 72 hours; numbers in parentheses represent standard deviation (n=3).

	Final Change in Metabolite Concentration (mM)					
Strain	**Butanol**	**Acetone**	**Ethanol**	**Acetate**	**Butyrate**	**Glucose**
ATCC 824	135.0 (±27.0)	70.8 (±4.8)	19.2 (±5.5)	1.2 (±6.6)	6.2 (±5.6)	276.9 (±52.1)
824(p94MCS)	150.4 (±16.1)	73.9 (±4.9)	29.6 (±4.9)	-1.3 (±1.9)	3.6 (±4.3)	311.8 (±24.2)
solRB	179.1 (±17.7)	88.2 (±12.3)	30.6 (±0.5)	-4.4 (±1.2)	2.4 (±1.8)	372.1 (±23.7)
824(p94_solB)	19.8 (±6.2)	6.6 (±2.5)	3.7 (±1.2)	14.8 (±3.2)	50.9 (±3.9)	105.2 (±5.3)
824ΔsolB	5.5 (±1.8)	0 (±0)	2.7 (±0.6)	12.5 (±1.6)	65.7 (±1.8)	85.8 (±10.4)

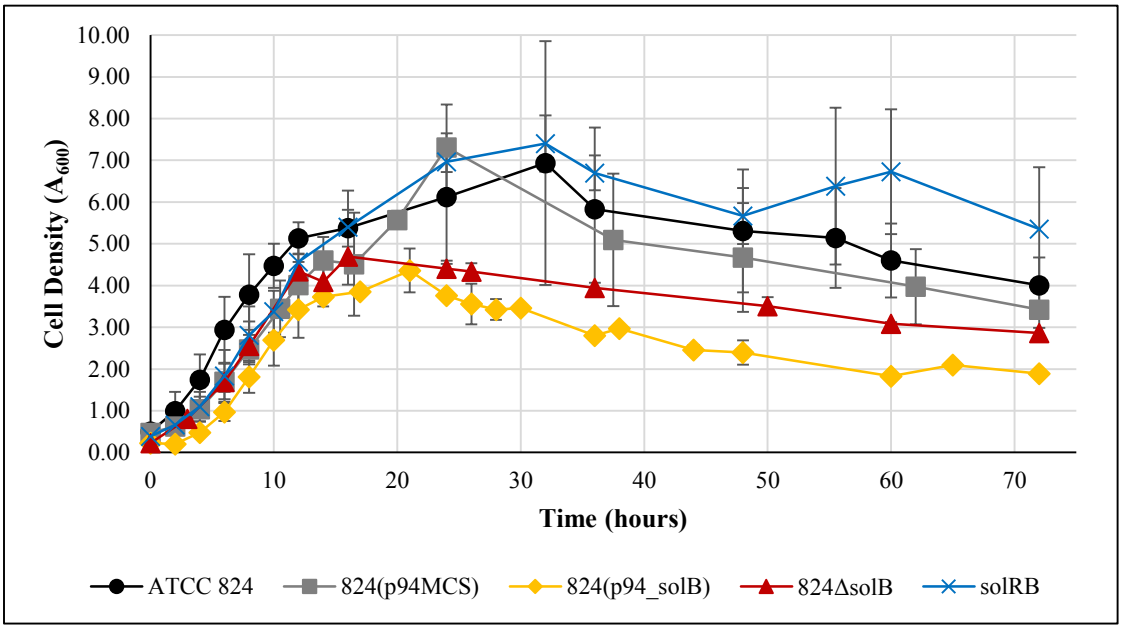


Figure 4.10 – Growth curves for 72-hour cultures of ATCC 824, 824(p94MCS), 824(p94_solB), 824ΔsolB, and solRB (40 mL). Error bars represent standard deviation (n = 3)

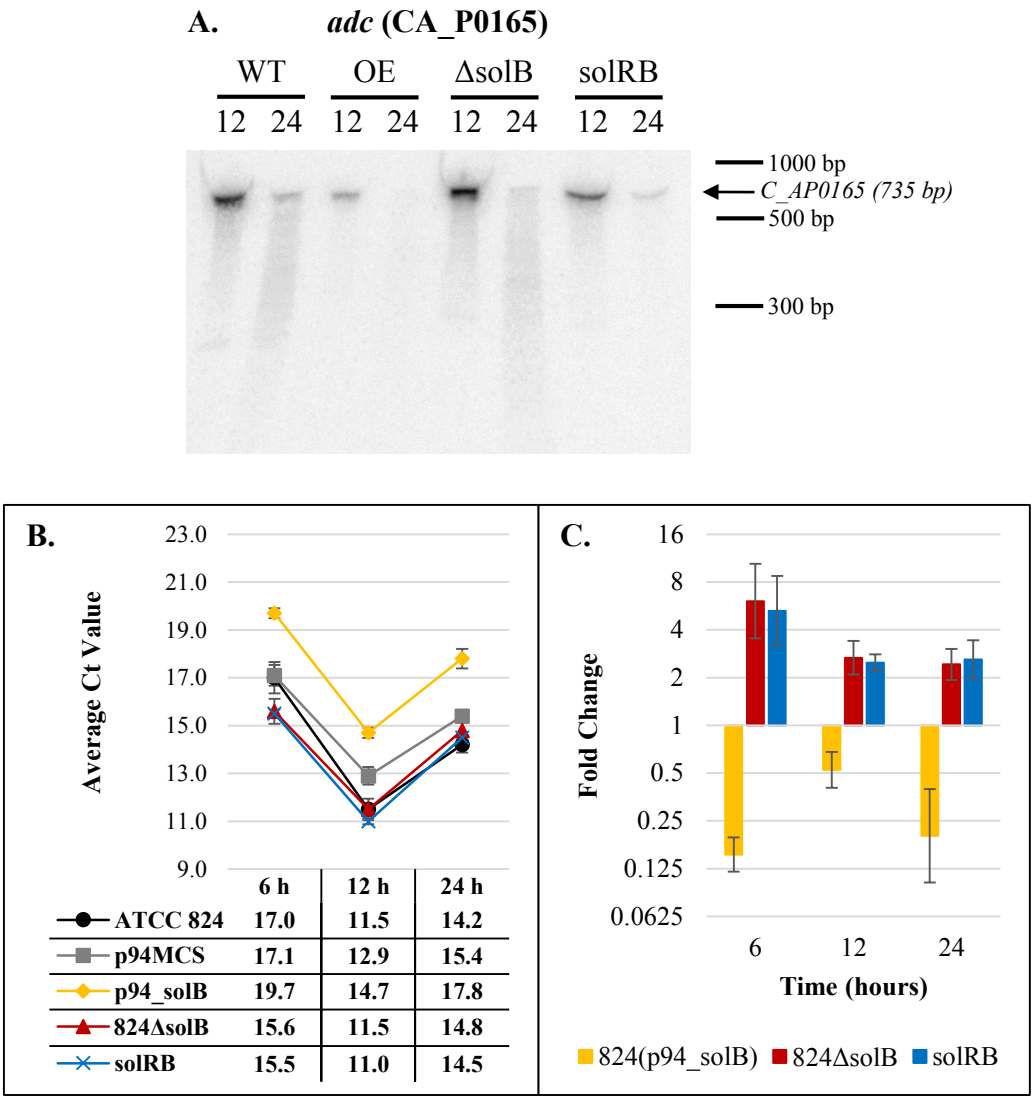


Figure 4.11 – Transcriptional analysis of *adc* in recombinant strains and WT *C. acetobutylicum*. (A) Northern analysis of *adc* transcript levels at 12 and 24 hours in ATCC 824 (WT), 824(p94_solB) (OE), 824ΔsolB, and solRB; five µg of total RNA samples; (B) Ct values and (C) fold changes in *adc* expression in ATCC 824, 824(p94MS), 824(p94_solB), 824ΔsolB, and solRB; the average Ct value for *adc* in non-template controls was 30.53 ± 0.17; fold changes in 824(p94_solB) are relative to plasmid control (p94MCS), and 824ΔsolB and solRB are relative to WT; *adc* expression normalized to the constitutively expressed CA_C3571 housekeeping gene; error bars represent standard deviation (n = 3).

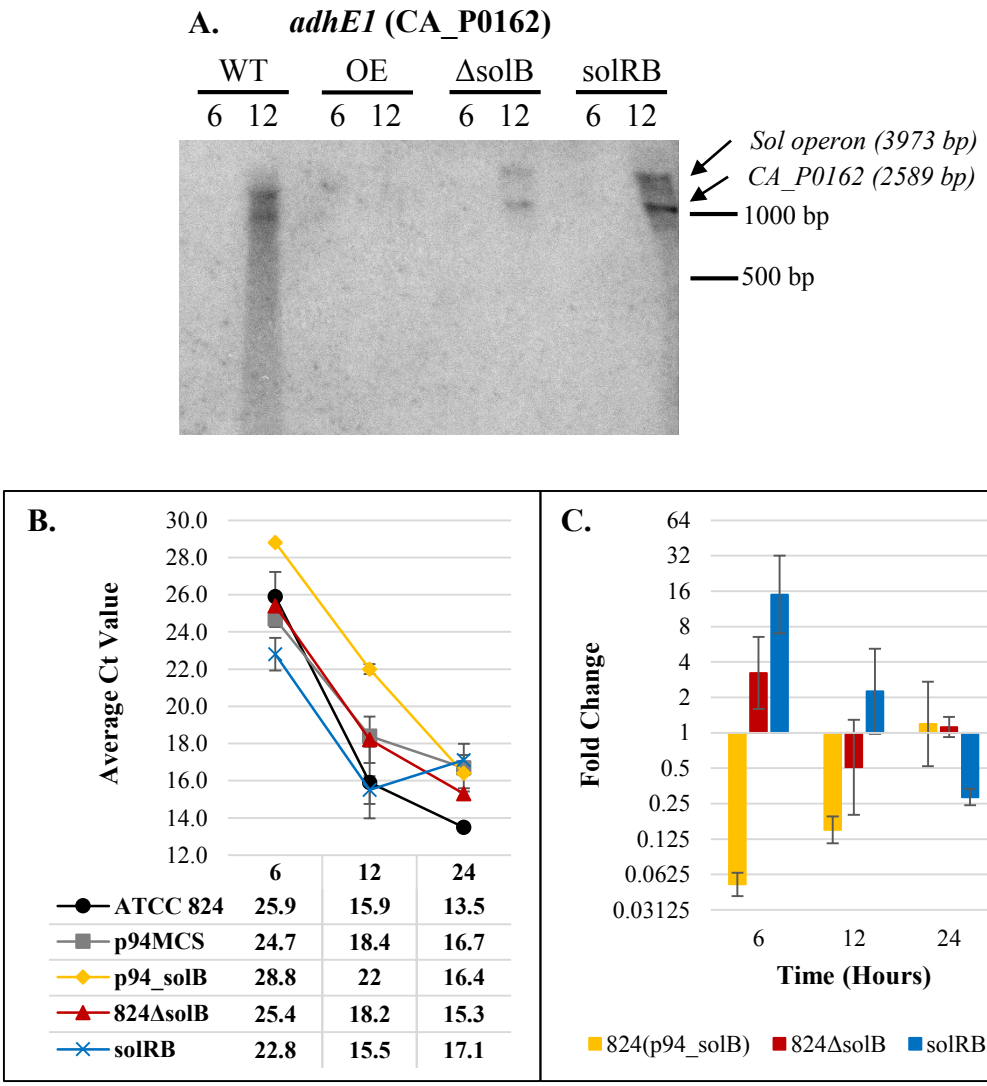
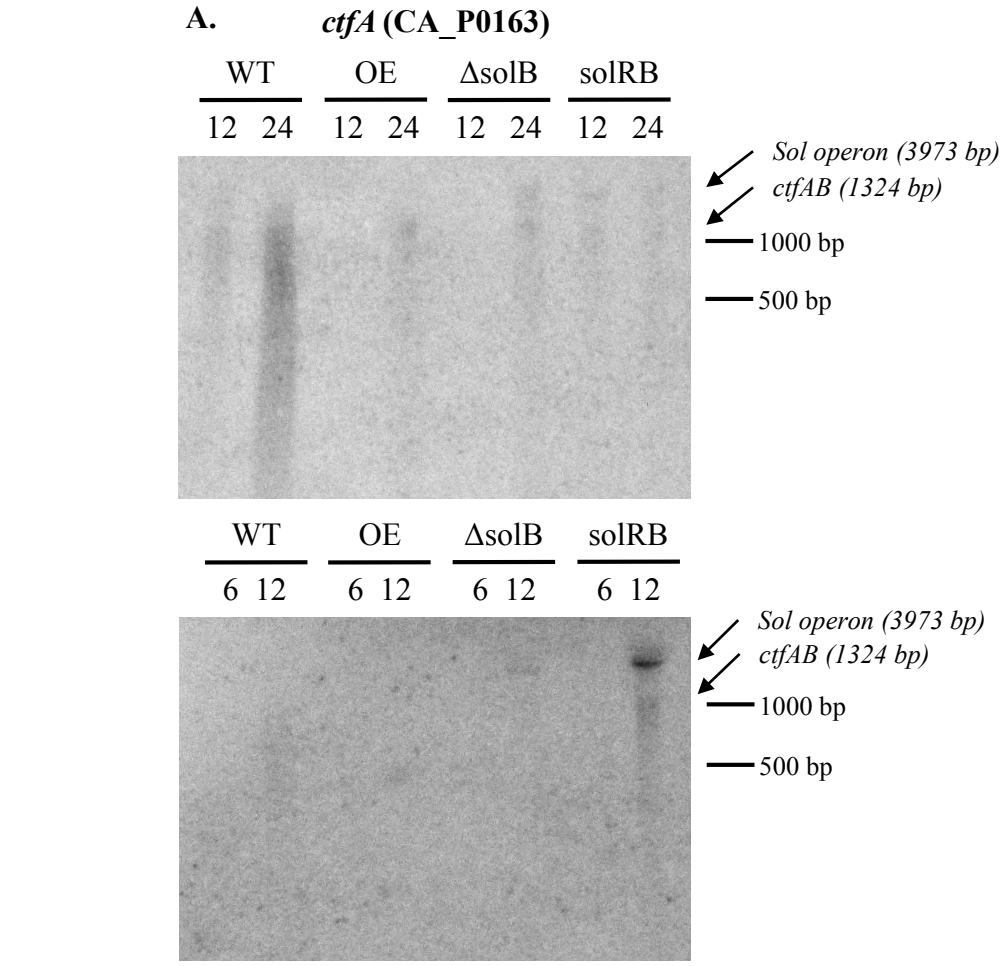


Figure 4.12 – Transcriptional analysis of *adhE1* in recombinant strains and WT *C. acetobutylicum*. (A) Northern analysis of *adhE1* transcript levels at 6 and 12 hours in ATCC 824 (WT), 824(p94_solB) (OE), 824ΔsolB, and solRB; fifteen µg of total RNA samples; note that larger bands (1000+ bp) are compressed in these blots; (B) Ct values and (C) fold changes in *adhE1* expression in ATCC 824, 824(p94MCS), 824(p94_solB), 824ΔsolB, and solRB; the average Ct value for *adhE1* in non-template controls was 29.58 ± 0.12; fold changes in 824(p94_solB) are relative to plasmid control (p94MCS), and 824ΔsolB and solRB are relative to WT; *adhE1* expression normalized to the constitutively expressed CA_C3571 housekeeping gene; error bars represent standard deviation (n = 3).


A.
ctfA (CA_P0163)
WT
OE
ΔsolB
solRB
12 24 12 24 12 24 12 24
Sol operon (3973 bp)
ctfAB (1324 bp)
1000 bp
500 bp
WT
OE
ΔsolB
solRB
6 12 6 12 6 12 6 12
Sol operon (3973 bp)
ctfAB (1324 bp)
1000 bp
500 bp

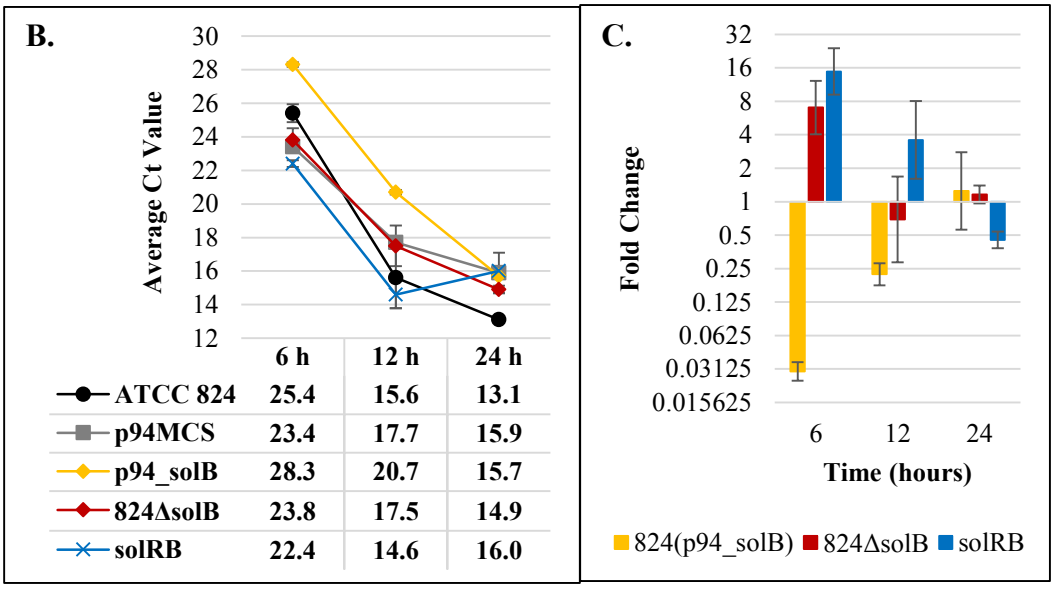

B.
Average Ct Value
30
28
26
24
22
20
18
16
14
12
6 h 12 h 24 h
ATCC 824 25.4 15.6 13.1
p94MCS 23.4 17.7 15.9
p94_solB 28.3 20.7 15.7
824ΔsolB 23.8 17.5 14.9
solRB 22.4 14.6 16.0
C.
Fold Change
32
16
8
4
2
1
0.5
0.25
0.125
0.0625
0.03125
0.015625
6 12 24
Time (hours)
824(p94_solB) 824ΔsolB solRB

Figure 4.13 – Transcriptional analysis of *ctfA* in recombinant strains and WT *C. acetobutylicum*. (A) Northern analysis of *ctfA* transcript levels at 12 and 24 hours and 6 and 12 hours in ATCC 824 (WT), 824(p94_solB) (OE), 824ΔsolB, and solRB; fifteen µg of total RNA samples; (B) Ct values and (C) fold changes in *ctfA* expression in ATCC 824, 824(p94MCS), 824(p94_solB), 824ΔsolB, and solRB; the average Ct value for *ctfA* in non-template controls was 30.85 ± 0.16; fold changes in 824(p94_solB) are relative to plasmid control (p94MCS), and 824ΔsolB and solRB are relative to wild type; *ctfA* expression normalized to the constitutively expressed CA_C3571 housekeeping gene; error bars represent standard deviation (n = 3).

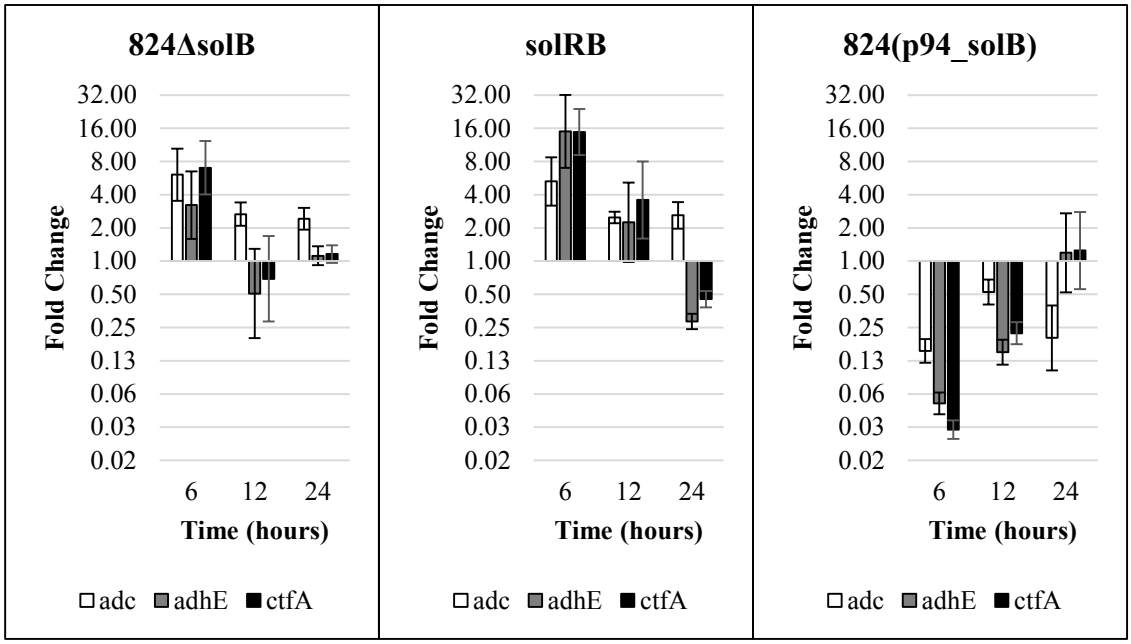
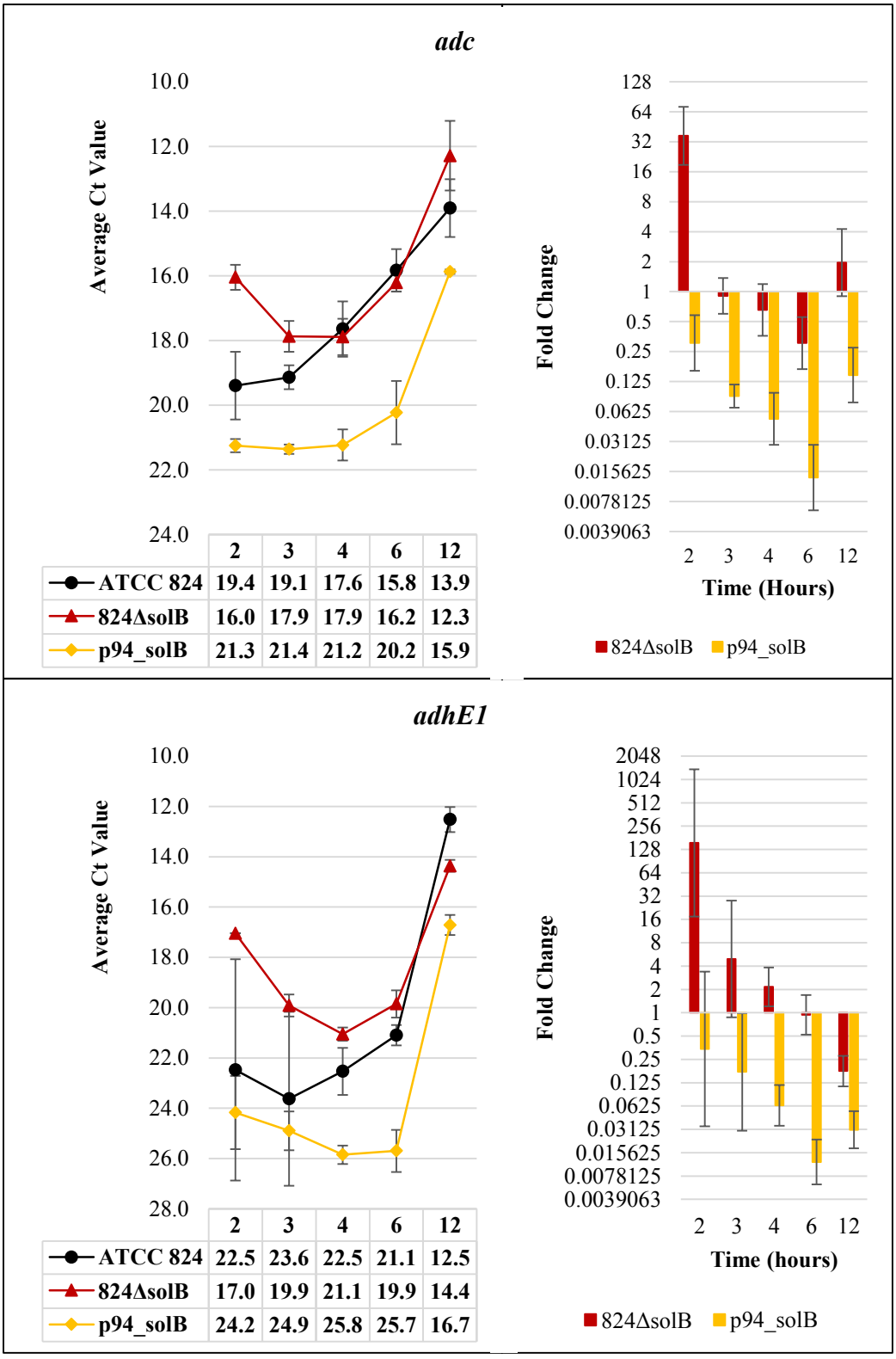


Figure 4.14 – Comparison of fold changes in sol locus gene expression in 824ΔsolB, solRB, and 824(p94_solB). Fold change values are taken from Figures 4.10c, 4.11c, and 4.12c; error bars represent standard deviation (n = 3).


adc
Average Ct Value
10.0
12.0
14.0
16.0
18.0
20.0
22.0
24.0
2 3 4 6 12
ATCC 824 19.4 19.1 17.6 15.8 13.9
824ΔsolB 16.0 17.9 17.9 16.2 12.3
p94_solB 21.3 21.4 21.2 20.2 15.9
Fold Change
128
64
32
16
8
4
2
1
0.5
0.25
0.125
0.0625
0.03125
0.015625
0.0078125
0.0039063
2 3 4 6 12
Time (Hours)
824ΔsolB
p94_solB
adhE1
Average Ct Value
10.0
12.0
14.0
16.0
18.0
20.0
22.0
24.0
26.0
28.0
2 3 4 6 12
ATCC 824 22.5 23.6 22.5 21.1 12.5
824ΔsolB 17.0 19.9 21.1 19.9 14.4
p94_solB 24.2 24.9 25.8 25.7 16.7
Fold Change
2048
1024
512
256
128
64
32
16
8
4
2
1
0.5
0.25
0.125
0.0625
0.03125
0.015625
0.0078125
0.0039063
2 3 4 6 12
Time (hours)
824ΔsolB
p94_solB

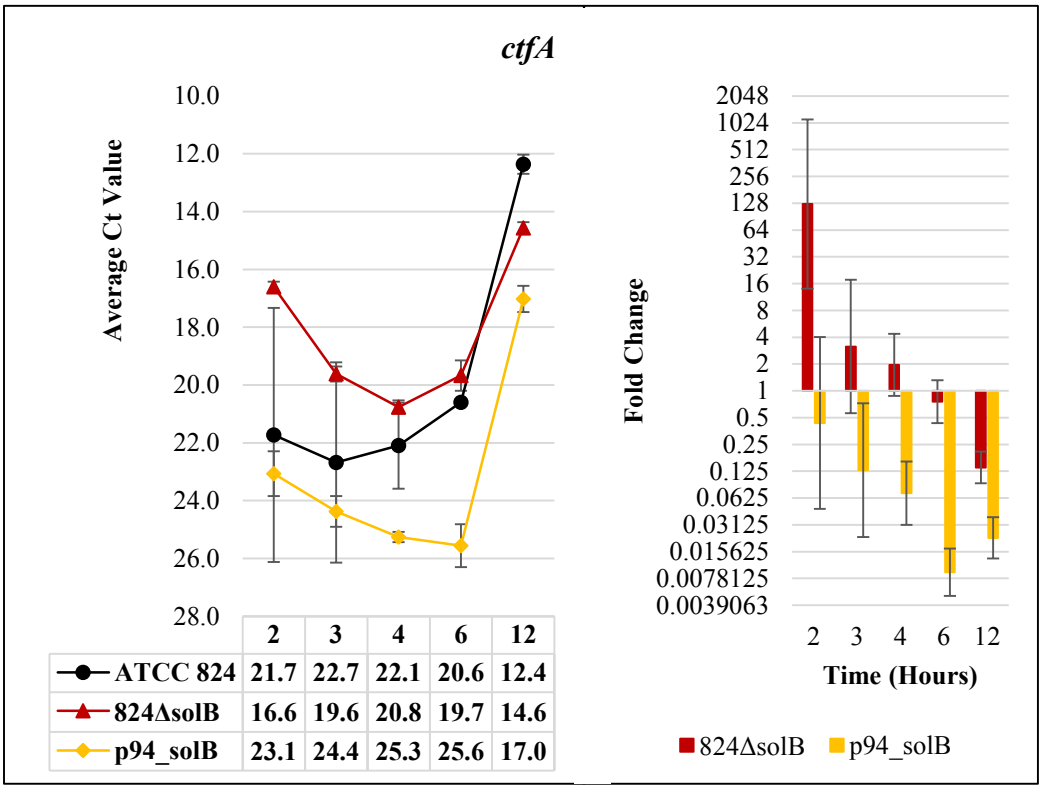


Figure 4.15 – Transcriptional analysis of *sol* operon genes during early growth. Ct values (left) and fold changes (right) of sol operon genes; no detectable expression of any gene in non-template controls; fold changes in both 824(p94_solB) and 824ΔsolB are relative to wild type and normalized to CA_C3571; error bars in all graphs represent standard deviation (n = 2).

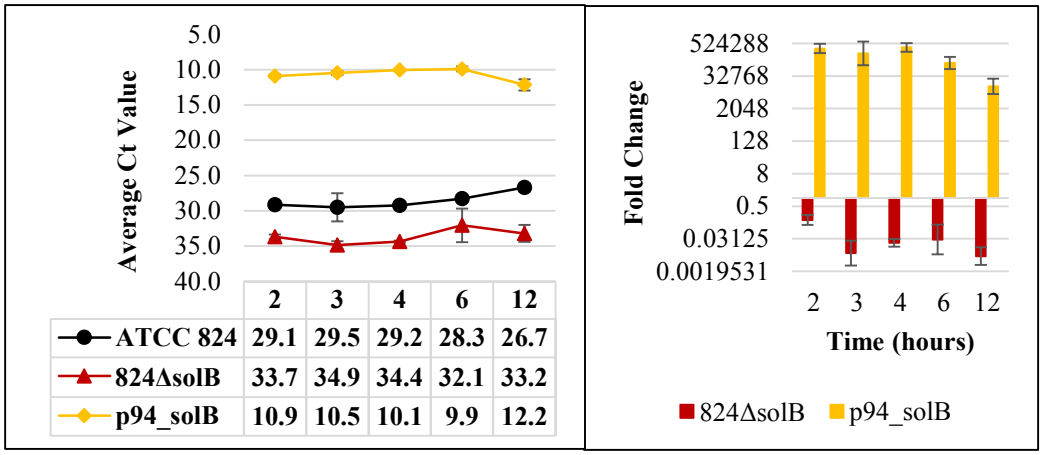


Figure 4.16 – Transcriptional analysis of *solB* expression during early growth. (A) Average Ct values and (B) fold changes of *solB* in 824(p94_solB) and 824ΔsolB; the average Ct value of *solB* in non-template controls was 33.05 ±1.30; fold changes in both strains are relative to wild type and normalized to CA_C3571; error bars in all graphs represent standard deviation (n = 2).

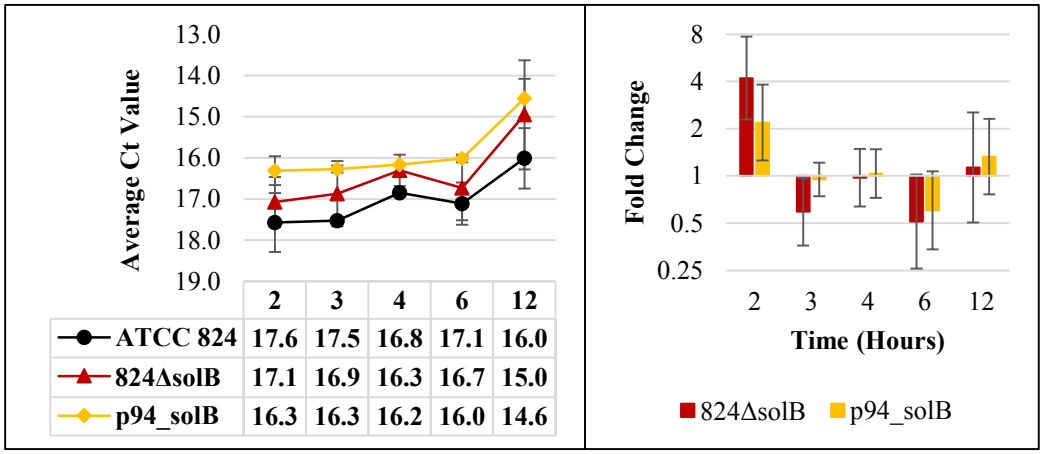


Figure 4.17 – Transcriptional analysis of *spo0A* expression during early growth. (A) Average Ct values and (B) fold changes of *solB* in 824(p94_solB) and 824ΔsolB; the average Ct value of *spo0A* in non-template controls was 37.44 ±0.55; fold changes in both strains are relative to wild type and normalized to CA_C3571; error bars in all graphs represent standard deviation (n = 2).

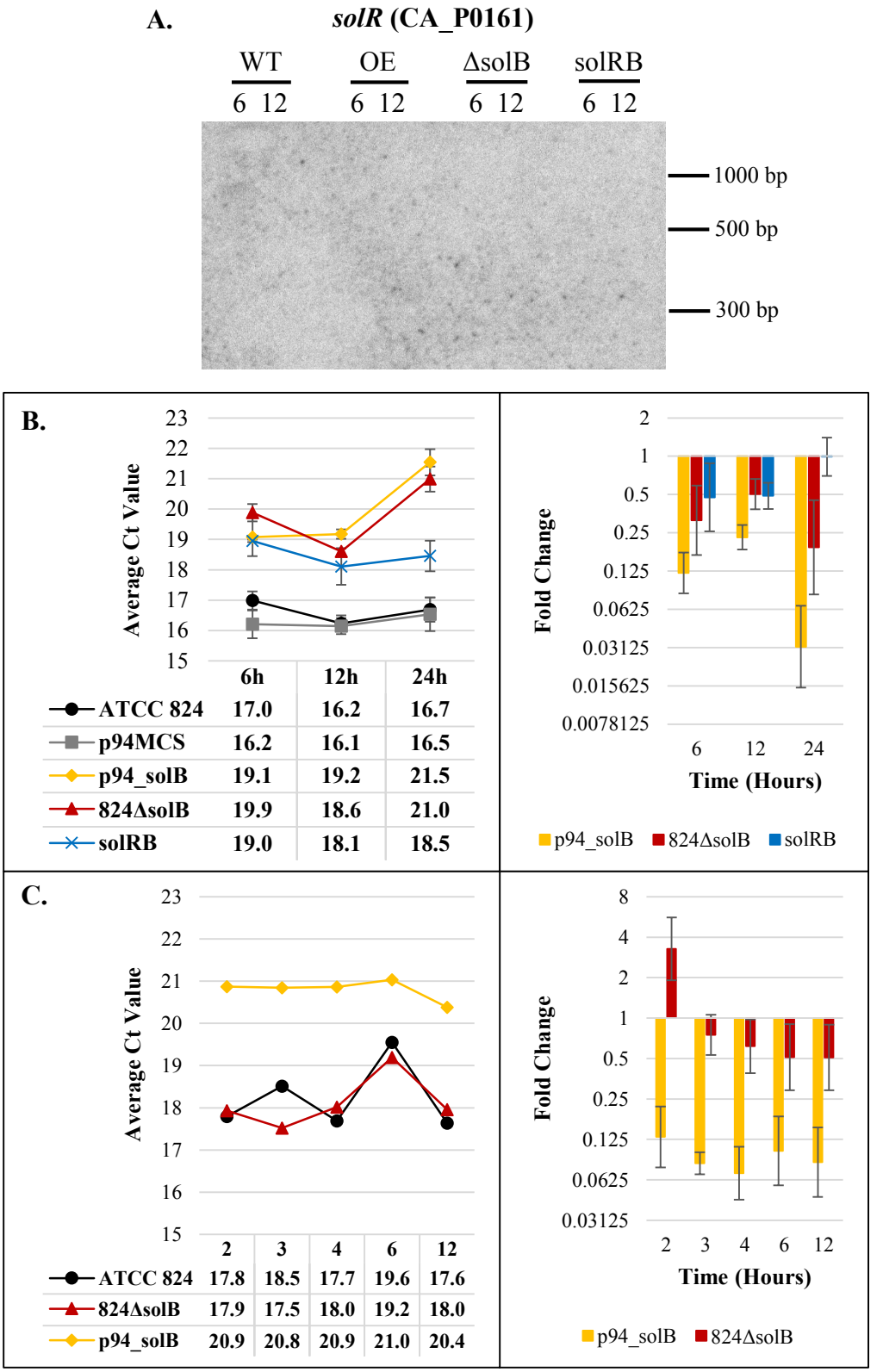
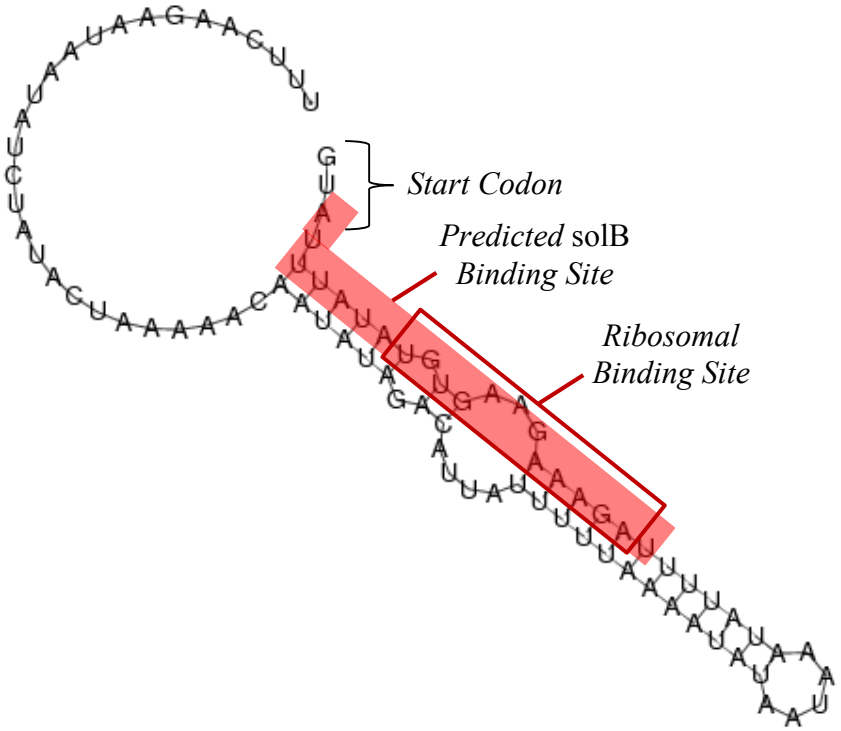

A.
solR (CA_P0161)
WT
OE
ΔsolB
solRB
6 12
6 12
6 12
6 12
1000 bp
500 bp
300 bp
B.
Average Ct Value
23
22
21
20
19
18
17
16
15
6h 12h 24h
ATCC 824 17.0 16.2 16.7
p94MCS 16.2 16.1 16.5
p94_solB 19.1 19.2 21.5
824ΔsolB 19.9 18.6 21.0
solRB 19.0 18.1 18.5
Fold Change
2
1
0.5
0.25
0.125
0.0625
0.03125
0.015625
0.0078125
6 12 24
Time (Hours)
p94_solB 824ΔsolB solRB
C.
Average Ct Value
23
22
21
20
19
18
17
16
15
2 3 4 6 12
ATCC 824 17.8 18.5 17.7 19.6 17.6
824ΔsolB 17.9 17.5 18.0 19.2 18.0
p94_solB 20.9 20.8 20.9 21.0 20.4
Fold Change
8
4
2
1
0.5
0.25
0.125
0.0625
0.03125
2 3 4 6 12
Time (Hours)
p94_solB 824ΔsolB

Figure 4.18 – Transcriptional analysis of *solR* in recombinant strains and WT *C. acetobutylicum*. (A) Northern analysis of *solR* transcript levels at 6 and 12 hours in ATCC 824 (WT), 824(p94_solB) (OE), 824ΔsolB, and solRB; fifteen µg of total RNA samples; (B) Ct values (left) and fold changes (right) in *solR* expression at 6, 12, and 24 hours in ATCC 824, 824(p94MS), 824(p94_solB), 824ΔsolB, and solRB; (C) Ct values (left) and fold changes (right) of *solR* during early growth; the average Ct value for *solR* in non-template controls was 31.36 ± 2.9 in (B) and 33.82 ± 0.42 in (C); fold changes in 824(p94_solB) are relative to plasmid control (B) or wild type (C), and all 824ΔsolB and solRB values are relative to wild type; *solR* expression normalized to the constitutively expressed CA_C3571 housekeeping gene; error bars represent standard deviation (B, n = 3) (C, n = 2).

A.



B.

UUGCAGUUUACAAUAUCUAUCUCCAAAUCUGC***U***UUCAAGAAUAAUAUCUAUACUA
AAAACAAUAUAGACAUUAUUUUUAAAAUAUAAUAAAUAUUUU**AGAAAGAAGUGU**A
UAUUU*AUG*

C.

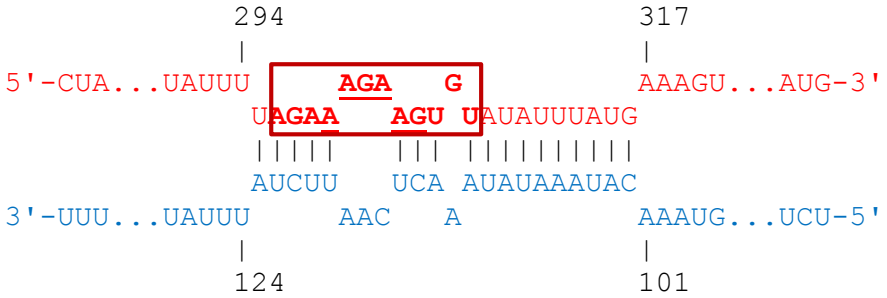
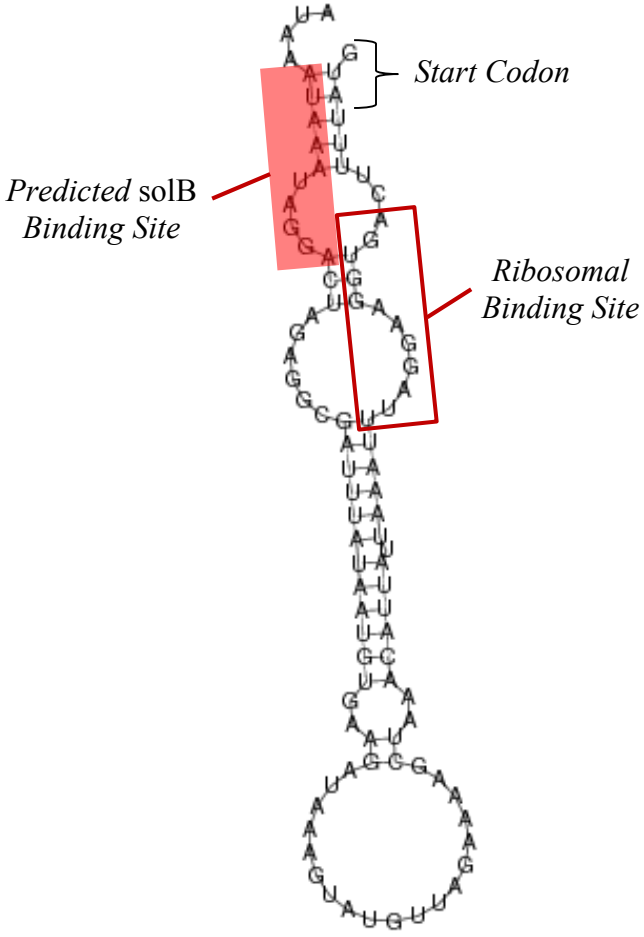


Figure 4.19 – Predictions of secondary structures and binding interactions of the *adhE1* transcript with *solB*. (A) Predicted secondary structure for the *adhE1* UTR+100 bp has MFE of -23.7 kcal/mol, and the region predicted for solB binding interaction is highlighted; (B) the sequence of the *adhE1* proximal UTR, showing the proximal promoter (bold, with -35 and -10 underlined), TSS (bold, italics), RBS (bold, boxed), and start codon (italics); (C) predicted interaction of *solB* (blue) with *adhE1* UTR (red), with RBS boxed; the hybridization energy for the predicted binding is -14.8 kcal/mol, and the overall free energy is -4.09 kcal/mol.

A.



B.

UUUACUUAAAAAAACAAUAUGUGUUAUAAUGUAAAU**A**UAAAUAAAUAGGACUAG
AGGCGAUUUAUAAUGUGAAGAUAAAGUAUGUUAGAAAAGCUAAACAUUAUUAAA
U**UUAGGAAGGUGA**CUUUU*AUG*UUAAAGG

C.

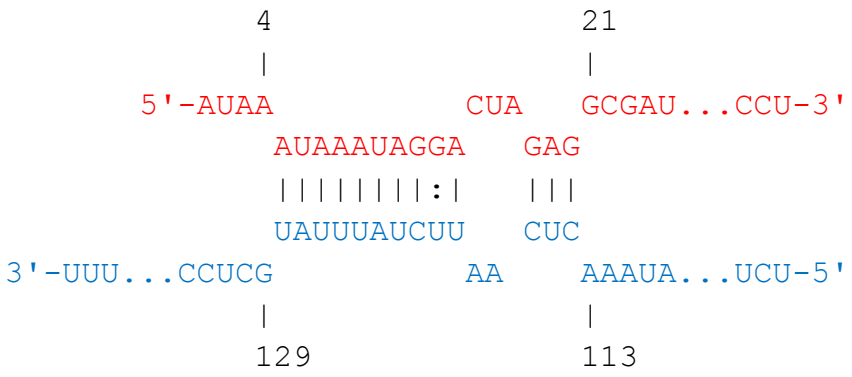
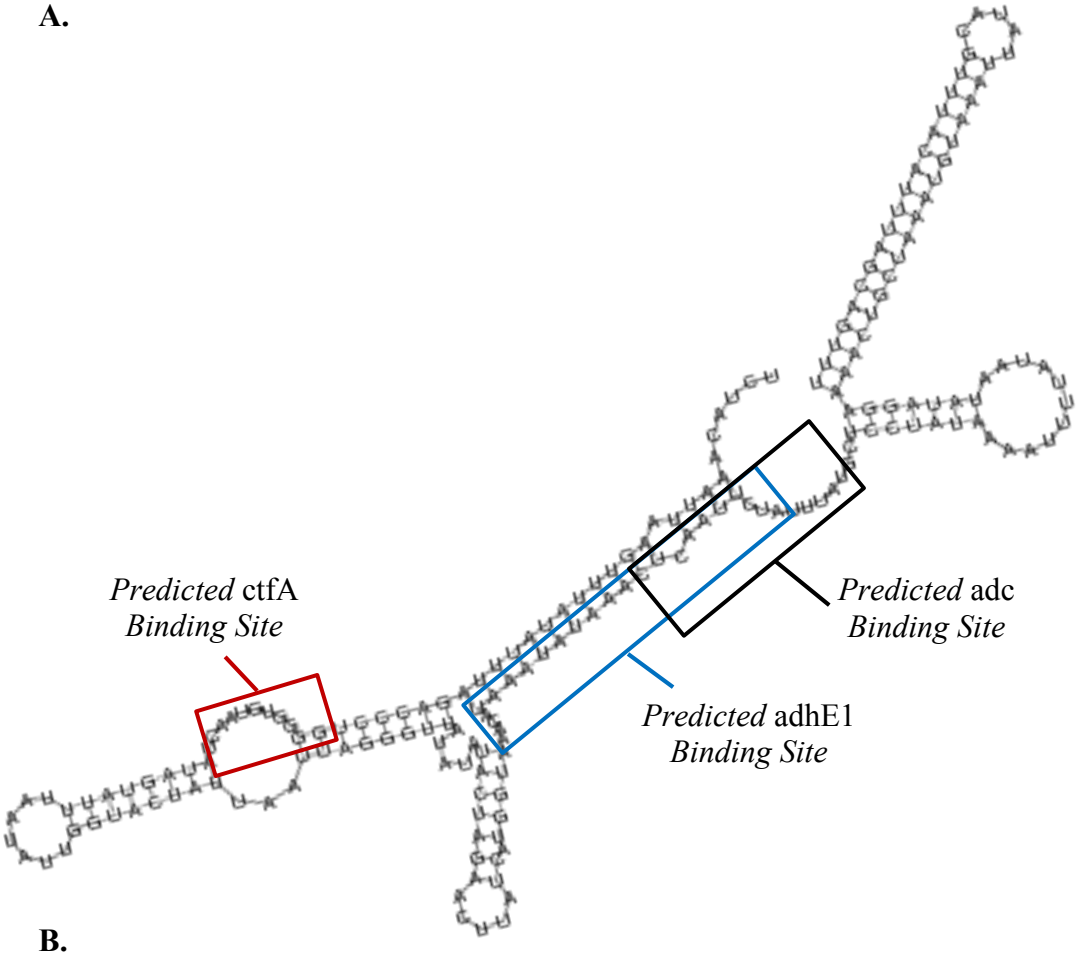


Figure 4.20 – Predictions of secondary structure and binding interactions of the *adc* transcript with *solB*. (A) Predicted secondary structure for the *adc* UTR+100 bp has an MFE of -26.8 kcal/mol, and the region predicted for *solB* binding interaction is highlighted; (B) the sequence of the *adc* UTR, showing the promoter (bold, with -35 and -10 underlined), TSS (bold, italics), RBS (bold, boxed), and start codon (italics); (C) predicted interaction of *solB* (blue) with *adc* UTR (red); the hybridization energy for the predicted binding is -11.2 kcal/mol, and the overall free energy is -4.40 kcal/mol.

A.



B.

AUUUCCUCUUACAAGUGAGAUAAAAGAAAUGUAUAUAAAUUGUUUUAAAAAACAACCU
*UAA*ACC**UUCAUAUUUCAACUACUUUUUAUAAU**UUUAAUAAAGAAUUUAAA**AGGAGG**GA
UUAAA*AUG*AACUCUAAAAUAAUUAGAUUU

C.

```
                  170       180
                   |         |
5'-AAA...AAACU  G         -3'
                GU GCACUCC
                || :|||:||
                CA UGUGGGG
3'-UUU...GAUAU  A        UCCCA...UCU-5'
                   |         |
                   40        29
```

Figure 4.21 – Predictions of binding interactions of the *ctfA* UTR, *adc* UTR, and *adhE1* proximal UTR with *solB*. (A) Predicted secondary structure for *solB* has an MFE of -51.0 kcal/mol, and the regions predicted for *ctfA* (red box), *adhE1* (blue box), and *adc* (black box) binding interaction are identified; (B) the sequence of the *ctfA* UTR, showing the terminal end of *adhE1* (italics), the predicted promoter (bold, with -35 and -10 underlined), RBS (bold, boxed), and *ctfA* start codon (italics); (C) the predicted interaction of *solB* (blue) with *ctfA* UTR (red); this predicted binding region of *ctfA* lies 93 bp downstream of the RBS; the hybridization energy for this binding prediction is -11.3 kcal/mol and overall free energy is -6.69 kcal/mol.

Table 4.2 – Final metabolite production and glucose utilization for recombinant strains and ATCC 824 after 72 hours of static flask growth (40 mL). Values represent change in concentration relative to the first time point (t_0), except for glucose values, which represent total consumption after 72 hours; numbers in parentheses represent standard deviation (n=3).

	Final Metabolite Titers (mM)					
Strain	**Butanol**	**Acetone**	**Ethanol**	**Acetate**	**Butyrate**	**Glucose**
ATCC 824	135.0 (±27.0)	70.8 (±4.8)	19.2 (±5.5)	1.2 (±6.6)	6.2 (±5.6)	276.9 (±52.1)
824(p94MCS)	150.4 (±16.1)	73.9 (±4.9)	29.6 (±4.9)	-1.3 (±1.9)	3.6 (±4.3)	311.8 (±24.2)
824(p94_solB)	19.8 (±6.2)	6.6 (±2.5)	3.7 (±1.2)	14.8 (±3.2)	50.9 (±3.9)	105.2 (±5.3)
824ΔsolB	5.5 (±1.8)	0.0 (±0.0)	2.7 (±0.6)	12.5 (±1.6)	65.7 (±1.8)	85.8 (±10.4)
p94_as(solB)F	0.0 (±0.0)	0.0 (±0.0)	5.7 (±3.5)	12.3 (±1.6)	46.4 (±1.8)	72.7 (±15.8)
p94_as(solB)S	-0.1 (±0.0)	0.0 (±0.0)	3.8 (±1.9)	19.2 (±0.8)	58.4 (±0.9)	76.3 (±13.9)

Chapter 5

CONCLUSIONS AND FUTURE DIRECTIONS

The overall purpose of this work was to examine the influence of 6S RNA and tmRNA on solvent tolerance in *C. acetobutylicum* and to investigate the putative regulatory role of *solB* on solventogenesis in the same organism. Small RNA are playing a rapidly increasing role in our understanding of prokaryotic life, especially in regulatory functions. *C. acetobutylicum* has been studied as a model for Firmicutes and for solventogenic fermentation, especially as a promising subject for the bioengineering of industrial-scale solvent production and biofuel development. Given its potential and the recent advances in sRNA studies, it is a clear choice to investigate the roles of sRNA in the processes related to efficient clostridial solvent production. It has been shown here that all three sRNA sequences have prominent roles in solvent formation and/or tolerance, but many more areas for further study of these regulators exist and are presented below. In particular, although an obvious role for *solB* in solventogenesis has been demonstrated, its specific regulatory function as well as its mechanism of action remain unclear. Several speculations as well as recommendations for future work in elucidating this function will be addressed.

5.1 General Conclusions

This study aimed to investigate the functions of three small RNAs in butanol tolerance and solventogenesis in *C. acetobutylicum*; ultimately, it was shown that both 6S RNA and tmRNA contribute positively to solvent tolerance, with overexpression of

the former also enhancing butanol production. Next, it was determined that *solB* plays a clear and influential role in solventogenesis in *C. acetobutylicum* that is more complex than the hypothesis suggests.

5.1.1 6S RNA and tmRNA Function in Stress Response and Solvent Tolerance

The small RNA sequences 6S RNA (sCA_C1377) and tmRNA (sCA_C834) were previously shown by comparative analysis of RNA-Seq results to be up-regulated under butanol and butyrate metabolite stress, suggesting a role in chemical stress response for both sRNA (Venkataramanan, et al, 2013). Building on these observations, the present study, described in detail in Chapter 3, demonstrated that overexpression of both 6S RNA and tmRNA sequences positively affects butanol tolerance, thus supporting the hypothesis that both sRNA function in chemical stress response. In addition, 6S RNA overexpression was found to positively affect butanol production. Studies of 6S RNA in *B. subtilis* and *E. coli* have suggested that 6S RNA regulates nutrient uptake efficiency (Faucher, et al, 2010; Cavanagh, et al, 2012, 2013); it is possible, therefore, that overexpression of 6S RNA in *C. acetobutylicum*, in addition to stimulating stationary-phase processes, results in more efficient nutrient absorption, which would contribute to a more rapid and efficient stimulation of stationary phase metabolism to yield the observed higher butanol production. Also clearly demonstrated by the present study is that only one 6S RNA homologue influences stress response and butanol production in *C. acetobutylicum*; the additional 6S RNA sequence may affect other cellular processes, fill some peripheral or supportive role, or possibly simply remain as a redundant, vestigial sequence that no longer serves a functional purpose.

5.1.2 Disruption of *solR* in solRB did not Impact Other Regions

The present study successfully revealed that the disruption of *solR* (CA_P0161) in the solRB strain did not impact any neighboring regulatory sequences, including the presumed regulatory sRNA sequence *solB*, and therefore indicating that the augmented solvent production of solRB was not a result of any effect on *solB*. Although a regulatory role for *solR* in solventogenesis has generally been discredited, as described in Chapter 1, the anomalous phenotype of solRB remains unknown. However, a speculated justification for this phenotype may be made: since all plasmid-containing strains routinely show an increase in solvent production relative to wild type levels, the higher solvent titers of solRB could be simply the result of the presence of two full copies of the pO1x suicide vector, each containing two antibiotic resistance markers; the integration of multiple, complete copies of the vector into the genome may provide sufficient metabolic pressure to generate the observed increase in solvent titers, especially given their presence in the genome rather than as independent plasmids.

5.1.3 The sRNA *solB* has a Complex Regulatory Role in Solventogenesis

Finally, this study has shown that *solB*, a sRNA sequence encoded upstream of the *sol* locus genes, plays a clear regulatory role in solventogenesis, and that this role is more complex than the initial hypothesis. Overexpression and deletion of *solB* both attenuate solvent production, but only increased *solB* levels results in a decrease in *sol* locus transcripts. These results for *solB* overexpression are reminiscent of those of targeted inactivation of both *adhE1* (Cooksley, et al, 2012) and *spo0A* (Harris, et al, 2002), which both showed minimal butanol, acetone, and ethanol production in conjunction with increased acid levels and a reduction in *sol* locus gene transcripts.

There are strongly predicted stem loops in the *sol* locus transcript UTRs, and strongly predicted interactions between *solB* and each *sol* locus transcript. All of these results indicate a strong, albeit still unknown, regulatory function. Interest in the role of the transcripts' UTRs is a relevant inquiry, since a recent study (Ito, et al, 2013) has elaborated on the crucial regulatory role of a stem loop sequence in the 5' UTR of *adhE* in *E. coli*. Previous studies (Thormann, et al, 2002; Scotcher, et al, 2003) have discussed the possibility of mRNA processing at the *adhE1* 5' UTR, suggesting too that processing may be dependent on its secondary structure. Although more recent 5' phosphorylation studies (Ralston, personal communication) suggest that the proximal promoter of *adhE1* is not an mRNA processing site, the UTR may still be necessary in some way for proper processing of a functioning transcript.

As described in Chapter 4, most common sRNA regulatory functions become contradicted when the phenotypes of both *solB* overexpression and deletion strains are considered. The best universally applied hypothesis to explain the *solB* regulatory action, based on the results of this study, is one of bimodal regulation dependent on *solB* concentration. At low concentrations, *solB* activates translation of the *sol* locus by freeing the ribosomal binding sites of each transcript, but at higher levels *solB* begins to facilitate the transcripts' degradation. Analysis of early growth has refuted the hypothesis that *solB* affects *spo0A* transcript levels; however, the possibility remains that *solB* could be a repressor of *spo0A* translation. Coupled with a role as activator of the *sol* locus genes, this model represents another possible bimodal regulatory function of *solB*. Translation of *spo0A* would be repressed by *solB* until the transitional phase, when other cell signals overpower *solB* repression and *spo0A* is translated effectively. From this point onward, *solB* would activate the translation of

sol locus transcripts as they are generated by *spo0A*. According to this model, overexpression of *solB* would result in the absence of *sol* locus transcripts due to extensive and sustained repression of Spo0A protein levels, and deletion of *solB* would yield successful *sol* locus transcription but not translation, due to the absence of their sRNA activator. Both of these proposed mechanisms of action for *solB* hold up against all the results observed in this study, but they are only speculation. These models can serve as hypotheses for further studies into the regulatory function of *solB*, and several recommendations for such future work are presented below.

A recent doctoral dissertation (Zimmermann, 2013) reported a solvent-negative *solB* overexpression strain (similar to that which is presented in this thesis) and a *solB* deletion strain exhibiting no change in solvent production or growth conditions. The differing phenotypes of the deletion strains of this dissertation and the present study are likely due to differences in the particular sequences deleted in each strain; notably the construction of this deletion strain, as presented, includes additional sequences, including the *repL* origin of replication, inserted in the genome immediately upstream of the *sol* operon. It is possible that this added sequencing positively affected the expression of the *sol* locus to maintain solvent production; also, the author speculates whether one or more other sRNA or regulatory elements could also be in play.

5.2 Recommendations and Future Directions

5.2.1 Optimizing the sRNA Overexpression Strains

With regard to the investigation of the overexpression of 6S RNA and tmRNA and their effects on butanol tolerance and solvent production, several areas of further work are warranted and recommended. As mentioned in Chapter 3, the additional

promoter activity on the p94MCS vector backbone can be eliminated by the addition of a Rho-independent terminator immediately upstream of the cloning site. In reference to the disparity between sRNA overexpression levels in 824(p94_6S) and 824(p94_tm) (Table 3.1), both sequences should be overexpressed to more equal levels so that results from solvent tolerance studies and metabolite analyses can be directly compared and evaluated more meaningfully. Alternative promoters can be tested in place of the *ptb* promoter on the p94MCS vector in an effort to bring expression of both sRNA to equal levels. Also, multiple copies of the tmRNA sequence can be cloned onto the same expression plasmid backbone in an effort to increase its overexpression to better match that of 6S RNA.

5.2.2 Expanding the Characterization of sRNA Effects on Solvent Tolerance and Production

As mentioned in Chapter 3, larger scale, pH-controlled batch fermentations of both 6S RNA and tmRNA overexpression strains will provide more controlled growth conditions that will provide a more precise characterization of the metabolite profiles and cell densities of both strains. To further develop the tolerance phenotype observed in this study, co-overexpression of 6S RNA and tmRNA would effectively investigate an intensified tolerance phenotype in ATCC 824. Relating to the higher butanol production observed in 824(p94_6S), co-overexpression of 6S RNA with one or more *sol* locus genes, like *adhE1*, may enhance the increased solvent production phenotype. Both co-overexpression studies, if performed by sequence co-expression from the same plasmid, would impose minimal additional pressure to the cell but possibly impart a significant increase to industrially relevant phenotypes.

In addition, investigations into the interaction(s) between 6S RNA and the *sol* locus genes will offer insight into the influence of 6S RNA on solvent formation and also possibly its function in solvent tolerance. Comparative analysis of the transcriptomic profile of 824(p94_6S) with wild type ATCC 824 using techniques including QPCR, microarray analysis, and RNA Deep Sequencing may potentially elucidate more detail on 6S RNA function, including the identification of one or more of the solventogenic genes as a target of 6S RNA action. Such research may also lead to discoveries of 6S RNA influences on other relevant genes and cellular processes.

Finally, a set of follow-up experiments in studying tolerance phenotypes in 6S RNA and tmRNA overexpression strains will be to test both strains in tolerance assays using the other solvents and metabolites generated by *C. acetobutylicum*, namely, acetone, ethanol, and the acid intermediates acetate and butyrate. Given the observable and statistically significant phenotypes arising under butanol stress, testing these strains with other metabolite stresses may reveal more extensive influences of tmRNA and 6S RNA on clostridial fermentation.

5.2.3 Investigating *solB* Effects on Translation by Western Blot Analysis

With the details of the regulatory role of *solB* uncertain, it is possible that its regulation may occur, partially or wholly, at the protein level. To investigate the possible effects of *solB* on translation as well as advance the understanding of *solB* regulatory action in general, Western blot analysis will reveal any translational regulation of the *sol* locus and/or *spo0A* in 824ΔsolB and 824(p94_solB). This investigation will elucidate any element of *solB* regulation at the translational level, particularly revealing whether *solB* regulation involves translational repression, as speculated above. Also, examination of Spo0A protein levels in both recombinant

strains will support or refute the hypothesis that *solB* influences *sol* locus transcription indirectly by repressing its activator.

5.2.4 Examining Sporulation Ability in Recombinant Strains

Relevant to continued understanding of the roles of 6S RNA and tmRNA as well as the overexpression and deletion of *solB* is the determination of the effect, if any, of the overexpression of each sRNA on the cell's ability to sporulate. This quality of each strain can be easily observed through microscopy techniques like phase contrast microscopy with sufficiently grown cultures (i.e., 72+ hours of growth).

5.2.5 Additional Recommendations and Thoughts

There is potential for advancing the understanding of *solB* function by analyzing the transcriptomic profile of both 824ΔsolB and 824(p94_solB) with RNA-Seq; this approach would yield a more detailed and global characterization of the effects of *solB* in the cell.

The allelic exchange system used in this study for deleting *solB* includes recognition sequences for the flippase enzyme in order to remove the thiamphenicol cassette, generating a markerless deletion; since no frameshift mutation has occurred in 824ΔsolB, and an insertion of a large sequence in the genome is also present in solRB, utilizing this method to remove the thiamphenicol marker is likely not necessary in 824ΔsolB, but it may nevertheless be worth considering, if only to prove definitively that the presence of the cassette has no effect on the observed phenotype.

To test the idea that *solB* overexpression results in transcript degradation, an experiment can be designed in which *solB* is overexpressed and one or more ribonuclease is completely or conditionally knocked out. In this recombinant strain,

the high *solB* levels would target *sol* locus transcripts for degradation, but the lack of ribonuclease activity would inhibit the cell's ability to degrade them. This complementation of *solB* overexpression with the inhibition of transcript degradation processes would aim to restore solvent production in the *solB* overexpression strain.

REFERENCES

1. Al-Hinai MA, Fast AG, Papoutsakis ET. Novel System for Efficient Isolation of Clostridium Double-Crossover Allelic Exchange Mutants Enabling Markerless Chromosomal Gene Deletions and DNA Integration. Applied and Environmental Microbiology 2012;78(22):8112-8121.

2. Al-Hinai MA, Jones SW, Papoutsakis ET. sigma(K) of Clostridium acetobutylicum Is the First Known Sporulation-Specific Sigma Factor with Two Developmentally Separated Roles, One Early and One Late in Sporulation. Journal of Bacteriology 2014;196(2):287-299.

3. Alsaker KV. Design, Optimization and Validation of Genomic DNA Microarrays for Examining the *Clostridium acetobutylicum* Transcriptome. 2005(10):432-443.

4. Alsaker KV, Papoutsakis ET. Transcriptional program of early sporulation and stationary-phase events in Clostridium acetobutylicum. J Bacteriol 2005;187(20):7103-18.

5. Alsaker KV, Paredes C, Papoutsakis ET. Metabolite stress and tolerance in the production of biofuels and chemicals: gene-expression-based systems analysis of butanol, butyrate, and acetate stresses in the anaerobe Clostridium acetobutylicum. Biotechnol Bioeng 2010;105(6):1131-47.

6. Alsaker KV, Spitzer TR, Papoutsakis ET. Transcriptional analysis of spo0A overexpression in Clostridium acetobutylicum and its effect on the cell's response to butanol stress. Journal of Bacteriology 2004;186(7):1959-1971.

7. Andini N, Nash KA. Expression of tmRNA in mycobacteria is increased by antimicrobial agents that target the ribosome. Fems Microbiology Letters 2011;322(2):172-179.

8. Bak G, Han K, Kim D, Lee Y. Roles of *rpoS*-activating small RNAs in pathways leading to acid resistance of *Escherichia coli*. *MicrobiologyOpen* 2014;3(1):15-28.

9. Barrick JE, Sudarsan N, Weinberg Z, Ruzzo WL, Breaker RR. 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. RNA 2005;11(5):774-84.

10. Behrens S, Widder S, Mannala GK, Qing XX, Madhugiri R, Kefer N, Abu Mraheil M, Rattei T, Hain T. Ultra Deep Sequencing of Listeria monocytogenes sRNA Transcriptome Revealed New Antisense RNAs. Plos One 2014;9(2):10.

11. Beisel CL, Storz G. Base pairing small RNAs and their roles in global regulatory networks. Fems Microbiology Reviews 2010;34(5):866-882.

12. Bennett GN, Rudolph FB. The central metabolic pathway from acetyl-CoA to butyryl-CoA in Clostridium acetobutylicum. FEMS Microbiology Reviews 1995;17(3):241-249.

13. Bi C, Jones SW, Hess DR, Tracy BP, Papoutsakis ET. SpoIIE is necessary for asymmetric division, sporulation, and expression of sigmaF, sigmaE, and sigmaG but does not control solvent production in Clostridium acetobutylicum ATCC 824. J Bacteriol 2011;193(19):5130-7.

14. Borden JR, Jones SW, Indurthi D, Chen YL, Papoutsakis ET. A genomic-library based discovery of a novel, possibly synthetic, acid-tolerance mechanism in Clostridium acetobutylicum involving non-coding RNAs and ribosomal RNA processing. Metabolic Engineering 2010;12(3):268-281.

15. Borden JR, Papoutsakis ET. Dynamics of genomic-library enrichment and identification of solvent tolerance genes for Clostridium acetobutylicum. Appl Environ Microbiol 2007;73(9):3061-8.

16. Bowles LK, Ellefson WL. Effects of Butanol on Clostridium-acetobutylicum. Applied and Environmental Microbiology 1985;50(5):1165-1170.

17. Boynton ZL, Bennett GN, Rudolph FB. Intracellular Concentrations of Coenzyme A and Its Derivatives from Clostridium acetobutylicum ATCC 824 and Their Roles in Enzyme Regulation. Appl Environ Microbiol 1994;60(1):39-44.

18. Brown DP, Ganova-Raeva L, Green BD, Wilkinson SR, Young M, Youngman P. Characterization of spo0A homologues in diverse Bacillus and Clostridium species identifies a probable DNA-binding domain. Mol Microbiol 1994;14(3):411-26.

19. Busch A, Richter AS, Backofen R. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. Bioinformatics 2008;24(24):2849-56.

20. Caron MP, Lafontaine DA, Masse E. Small RNA-mediated regulation at the level of transcript stability. Rna Biology 2010;7(2):140-144.

21. Cary JW, Petersen DJ, Papoutsakis ET, Bennett GN. Cloning and expression of Clostridium acetobutylicum ATCC 824 acetoacetyl-coenzyme A:acetate/butyrate:coenzyme A-transferase in Escherichia coli. Appl Environ Microbiol 1990;56(6):1576-83.

22. Cavanagh AT, Sperger JM, Wassarman KM. Regulation of 6S RNA by pRNA synthesis is required for efficient recovery from stationary phase in E. coli and B. subtilis. Nucleic Acids Res 2012;40(5):2234-46.

23. Cavanagh AT, Wassarman KM. 6S-1 RNA function leads to a delay in sporulation in Bacillus subtilis. J Bacteriol 2013;195(9):2079-86.

24. Cavanagh AT, Wassarman KM. 6S RNA, a global regulator of transcription in Escherichia coli, Bacillus subtilis, and beyond. Annu Rev Microbiol 2014;68:45-60.

25. Chambers JR, Sauer K. Small RNAs and their role in biofilm formation. Trends in Microbiology 2013;21(1):39-49.

26. Chen YL, Indurthi DC, Jones SW, Papoutsakis ET. Small RNAs in the Genus Clostridium. Mbio 2011;2(1):11.

27. Cho SH, Lei R, Henninger TD, Contreras LM. Discovery of Ethanol-Responsive Small RNAs in Zymomonas mobilis. Applied and Environmental Microbiology 2014;80(14):4189-4198.

28. Cooksley CM, Zhang Y, Wang H, Redl S, Winzer K, Minton NP. Targeted mutagenesis of the Clostridium acetobutylicum acetone-butanol-ethanol fermentation pathway. Metab Eng 2012;14(6):630-41.

29. Cornillot E, Nair RV, Papoutsakis ET, Soucaille P. The genes for butanol and acetone formation in Clostridium acetobutylicum ATCC 824 reside on a large plasmid whose loss leads to degeneration of the strain. J Bacteriol 1997;179(17):5442-7.

30. Davies R. Studies on the acetone-butanol fermentation: 4. Acetoacetic acid decarboxylase of Cl. acetobutylicum (BY). Biochem J 1943;37(2):230-8.

31. Desai RP, Harris LM, Welker NE, Papoutsakis ET. Metabolic flux analysis elucidates the importance of the acid-formation pathways in regulating solvent production by Clostridium acetobutylicum. Metab Eng 1999;1(3):206-13.

32. Desai RP, Nielsen LK, Papoutsakis ET. Stoichiometric modeling of Clostridium acetobutylicum fermentations with non-linear constraints. J Biotechnol 1999;71(1-3):191-205.

33. Desai RP, Papoutsakis ET. Antisense RNA strategies for metabolic engineering of Clostridium acetobutylicum. Appl Environ Microbiol 1999;65(3):936-45.

34. Desnoyers G, Bouchard MP, Masse E. New insights into small RNA-dependent translational regulation in prokaryotes. Trends in Genetics 2013;29(2):92-98.

35. Dürre P. Biobutanol: an attractive biofuel. Biotechnol J 2007;2(12):1525-34.

36. Durre P, Bohringer M, Nakotte S, Schaffer S, Thormann K, Zickner B. Transcriptional regulation of solventogenesis in Clostridium acetobutylicum. Journal of Molecular Microbiology and Biotechnology 2002;4(3):295-300.

37. Dürre P, Fischer RJ, Kuhn A, Lorenz K, Schreiber W, Stürzenhofecker B, Ullmann S, Winzer K, Sauer U. Solventogenic enzymes of Clostridium acetobutylicum: catalytic properties, genetic organization, and transcriptional regulation. FEMS Microbiol Rev 1995;17(3):251-62.

38. Durre P, Kuhn A, Gottwald M, Gottschalk G. Enzymatic investigations on butanol dehydrogenase and butyraldehyde dehydrogenase in extracts of Clostridium acetobutylicum. Appl Microbiol Biotechnol 1987;26(3):268-272.

39. Durre, P, Linder, S, Zimmermann, T, Schiel-Bengelsdorf, B: Master and Commander: Spo0A vs. SolB, a small regulatory RNA, in controlling the shift to solventogenesis in Clostridium acetobutylicum. Programme for the International Conference on the Genetics, Physiology and Biotechnology of Solvent- and Acid-forming Clostridia (Clostridium XII), 10 to 12 September 2012. Nottingham, UK.

40. Ezeji TC, Qureshi N, Blaschek HP. Bioproduction of butanol from biomass: from genes to bioreactors. Curr Opin Biotechnol 2007;18(3):220-7.

41. Faucher SP, Friedlander G, Livny J, Margalit H, Shuman HA. Legionella pneumophila 6S RNA optimizes intracellular multiplication. Proc Natl Acad Sci U S A 2010;107(16):7533-8.

42. Fischer RJ, Helms J, Dürre P. Cloning, sequencing, and molecular analysis of the sol operon of Clostridium acetobutylicum, a chromosomal locus involved in solventogenesis. J Bacteriol 1993;175(21):6959-69.

43. Fontaine L, Meynial-Salles I, Girbal L, Yang X, Croux C, Soucaille P. Molecular characterization and transcriptional analysis of adhE2, the gene encoding the NADH-dependent aldehyde/alcohol dehydrogenase responsible for butanol production in alcohologenic cultures of Clostridium acetobutylicum ATCC 824. J Bacteriol 2002;184(3):821-30.

44. Fujihara A, Tomatsu H, Inagaki S, Tadaki T, Ushida C, Himeno H, Muto A. Detection of tmRNA-mediated trans-translation products in Bacillus subtilis. Genes to Cells 2002;7(3):343-350.

45. Gaida SM, Al-Hinai MA, Indurthi DC, Nicolaou SA, Papoutsakis ET. Synthetic tolerance: three noncoding small RNAs, DsrA, ArcZ and RprA, acting supra-additively against acid stress. Nucleic Acids Res 2013;41(18):8726-37.

46. Georg J, Hess WR. cis-Antisense RNA, Another Level of Gene Regulation in Bacteria. Microbiology and Molecular Biology Reviews 2011;75(2):286-+.

47. Gerischer U, Dürre P. mRNA analysis of the adc gene region of Clostridium acetobutylicum during the shift to solventogenesis. J Bacteriol 1992;174(2):426-33.

48. Ghaz-Jahanian MA, Khodaparastan F, Berenjian A, Jafarizadeh-Malmiri H. Influence of Small RNAs on Biofilm Formation Process in Bacteria. Molecular Biotechnology 2013;55(3):288-297.

49. Gheshlaghi R, Scharer JM, Moo-Young M, Chou CP. Metabolic pathways of clostridia for producing butanol. Biotechnol Adv 2009;27(6):764-81.

50. Gildehaus N, Neusser T, Wurm R, Wagner R. Studies on the function of the riboregulator 6S RNA from E. coli: RNA polymerase binding, inhibition of in vitro transcription and synthesis of RNA-directed de novo transcripts. Nucleic Acids Res 2007;35(6):1885-96.

51. Girbal L, Mortier-Barrière I, Raynaud F, Rouanet C, Croux C, Soucaille P. Development of a sensitive gene expression reporter system and an inducible promoter-repressor system for Clostridium acetobutylicum. Appl Environ Microbiol 2003;69(8):4985-8.

52. Gopel Y, Gorke B. Rewiring two-component signal transduction with small RNAs. Current Opinion in Microbiology 2012;15(2):132-139.

53. Green EM. Fermentative production of butanol--the industrial perspective. Curr Opin Biotechnol 2011;22(3):337-43.

54. Green EM, Boynton ZL, Harris LM, Rudolph FB, Papoutsakis ET, Bennett GN. Genetic manipulation of acid formation pathways by gene inactivation in Clostridium acetobutylicum ATCC 824. Microbiology-Uk 1996;142:2079-2086.

55. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The Vienna RNA websuite. Nucleic Acids Res 2008;36(Web Server issue):W70-4.

56. Grupe H, Gottschalk G. Physiological Events in Clostridium acetobutylicum during the Shift from Acidogenesis to Solventogenesis in Continuous Culture and Presentation of a Model for Shift Induction. Appl Environ Microbiol 1992;58(12):3896-902.

57. Harris LM, Blank L, Desai RP, Welker NE, Papoutsakis ET. Fermentation characterization and flux analysis of recombinant strains of Clostridium acetobutylicum with an inactivated solR gene. Journal of Industrial Microbiology & Biotechnology 2001;27(5):322-328.

58. Harris LM, Welker NE, Papoutsakis ET. Northern, morphological, and fermentation analysis of spo0A inactivation and overexpression in Clostridium acetobutylicum ATCC 824. Journal of Bacteriology 2002;184(13):3586-3597.

59. Hartman AH, Liu H, Melville SB. Construction and characterization of a lactose-inducible promoter system for controlled gene expression in Clostridium perfringens. Appl Environ Microbiol 2011;77(2):471-8.

60. Hartmanis MG, Gatenbeck S. Intermediary Metabolism in Clostridium acetobutylicum: Levels of Enzymes Involved in the Formation of Acetate and Butyrate. Appl Environ Microbiol 1984;47(6):1277-83.

61. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP. The ClosTron: a universal gene knock-out system for the genus Clostridium. J Microbiol Methods 2007;70(3):452-64.

62. Hillmann F, Fischer RJ, Saint-Prix F, Girbal L, Bahl H. PerR acts as a switch for oxygen tolerance in the strict anaerobe Clostridium acetobutylicum. Mol Microbiol 2008;68(4):848-60.

63. Himeno H, Kurita D, Muto A. tmRNA-mediated trans-translation as the major ribosome rescue system in a bacterial cell. Frontiers in genetics 2014;5:66.

64. Hou X, Peng W, Xiong L, Huang C, Chen X, Zhang W. Engineering Clostridium acetobutylicum for alcohol production. J Biotechnol 2013;166(1-2):25-33.

65. Ito K, Hamasaki K, Kayamori A, Nguyen PA, Amagai K, Wachi M. A secondary structure in the 5' untranslated region of adhE mRNA required for RNase G-dependent regulation. Biosci Biotechnol Biochem 2013;77(12):2473-9.

66. Janssen BD, Hayes CS. THE tmRNA RIBOSOME-RESCUE SYSTEM. In: Marintchev A, editor. Advances in Protein Chemistry and Structural Biology, Vol 86: Fidelity and Quality Control in Gene Expression. Volume 86, Advances in Protein Chemistry and Structural Biology. San Diego: Elsevier Academic Press Inc; 2012. p 151-191.

67. Jia KZ, Zhang YP, Li Y. Identification and Characterization of Two Functionally Unknown Genes Involved in Butanol Tolerance of Clostridium acetobutylicum. Plos One 2012;7(6):10.

68. Jones DT, Woods DR. Acetone-butanol fermentation revisited. Microbiol Rev 1986;50(4):484-524.

69. Jones SW, Paredes CJ, Tracy B, Cheng N, Sillers R, Senger RS, Papoutsakis ET. The transcriptional program underlying the physiology of clostridial sporulation. Genome Biology 2008;9(7):21.

70. Jones SW, Tracy BP, Gaida SM, Papoutsakis ET. Inactivation of sigma(F) in Clostridium acetobutylicum ATCC 824 Blocks Sporulation Prior to Asymmetric Division and Abolishes sigma(E) and sigma(G) Protein Expression but Does Not Block Solvent Formation. Journal of Bacteriology 2011;193(10):2429-2440.

71. Komine Y, Kitabatake M, Yokogawa T, Nishikawa K, Inokuchi H. A tRNA-like structure is present in 10Sa RNA, a small stable RNA from Escherichia coli. Proc Natl Acad Sci U S A 1994;91(20):9223-7.

72. Lalaouna D, Simoneau-Roy M, Lafontaine D, Masse E. Regulatory RNAS and target mRNA decay in prokaryotes. Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms 2013;1829(6-7):742-747.

73. Lee SY, Bennett GN, Papoutsakis ET. Construction of Escherichia coli Clostridium acetobutylicum Shuttle Vectors and Transformation of Clostridium acetobutylicum Strains. Biotechnology Letters 1992;14(5):427-432.

74. Lee SY, Mermelstein LD, Bennett GN, Papoutsakis ET. Vector construction, transformation, and gene amplification in Clostridium acetobutylicum ATCC 824. Ann N Y Acad Sci 1992;665:39-51.

75. Lee SY, Park JH, Jang SH, Nielsen LK, Kim J, Jung KS. Fermentative butanol production by clostridia. Biotechnology and Bioengineering 2008;101(2):209-228.

76. Levi-Meyrueis C, Monteil V, Sismeiro O, Dillies M, Monot M, Jagla B, Coppee J, Dupuy B, Norel F. Expanding the RpoS/sigmaS-Network by RNA Sequencing and Identification of sigmaS-Controlled Small RNAs in *Salmonella*. PLoS ONE 2014;9(5):e96918.

77. Lütke-Eversloh T, Bahl H. Metabolic engineering of Clostridium acetobutylicum: recent advances to improve butanol production. Curr Opin Biotechnol 2011;22(5):634-47.

78. Majdalani N, Vanderpool CK, Gottesman S. Bacterial small RNA regulators. Critical Reviews in Biochemistry and Molecular Biology 2005;40(2):93-113.

79. Mandin P, Guillier M. Expanding control in bacteria: interplay between small RNAs and transcriptional regulators to control gene expression. Current Opinion in Microbiology 2013;16(2):125-132.

80. Mermelstein LD, Papoutsakis ET. In vivo methylation in Escherichia coli by the Bacillus subtilis phage phi 3T I methyltransferase to protect plasmids from restriction upon transformation of Clostridium acetobutylicum ATCC 824. Appl Environ Microbiol 1993;59(4):1077-81.

81. Mermelstein LD, Welker NE, Bennett GN, Papoutsakis ET. Expression of cloned homologous fermentative genes in Clostridium acetobutylicum ATCC 824. Biotechnology (N Y) 1992;10(2):190-5.

82. Muto A, Fujihara A, Ito K, Matsuno J, Ushida C, Himeno H. Requirement of transfer-messenger RNA for the growth of Bacillus subtilis under stresses. Genes to Cells 2000;5(8):627-635.

83. Nair RV, Bennett GN, Papoutsakis ET. Molecular Characterization of an Aldehyde/Alcohol Dehydrogenase Gene from Clostridium acetobutylicum ATCC-824. Journal of Bacteriology 1994;176(3):871-885.

84. Nair RV, Green EM, Watson DE, Bennett GN, Papoutsakis ET. Regulation of the sol locus genes for butanol and acetone formation in Clostridium acetobutylicum ATCC 824 by a putative transcriptional repressor. Journal of Bacteriology 1999;181(1):319-330.

85. Nair RV, Papoutsakis ET. Expression of plasmid-encoded aad in Clostridium acetobutylicum M5 restores vigorous butanol production. J Bacteriol 1994;176(18):5843-6.

86. Nawrocki EP, Eddy SR. Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics 2013;29(22):2933-5.

87. Neusser T, Polen T, Geissen R, Wagner R. Depletion of the non-coding regulatory 6S RNA in E. coli causes a surprising reduction in the expression of the translation machinery. BMC Genomics 2010;11:165.

88. Nicolaou SA, Gaida SM, Papoutsakis ET. A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: From biofuels and chemicals, to biocatalysis and bioremediation. Metab Eng 2010;12(4):307-31.

89. Nölling J, Breton G, Omelchenko MV, Makarova KS, Zeng Q, Gibson R, Lee HM, Dubois J, Qiu D, Hitti J and others. Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum. J Bacteriol 2001;183(16):4823-38.

90. Papoutsakis ET. Engineering solventogenic clostridia. Curr Opin Biotechnol 2008;19(5):420-9.

91. Papoutsakis ET, Bussineau CM, Chu IM, Diwan AR, Huesemann M. Transport of substrates and metabolites and their effect on cell metabolism (in butyric-acid and methylotrophic fermentations). Ann N Y Acad Sci 1987;506:24-50.

92. Paredes CJ, Alsaker KV, Papoutsakis ET. A comparative genomic view of clostridial sporulation and physiology. Nature Reviews Microbiology 2005;3(12):969-978.

93. Paredes CJ, Rigoutsos I, Papoutsakis ET. Transcriptional organization of the Clostridium acetobutylicum genome. Nucleic Acids Res 2004;32(6):1973-81.

94. Paredes CJ, Senger RS, Spath IS, Borden JR, Sillers R, Papoutsakis ET. A general framework for designing and validating oligomer-based DNA microarrays and its application to Clostridium acetobutylicum. Appl Environ Microbiol 2007;73(14):4631-8.

95. Petersen DJ, Cary JW, Vanderleyden J, Bennett GN. Sequence and arrangement of genes encoding enzymes of the acetone-production pathway of Clostridium acetobutylicum ATCC824. Gene 1993;123(1):93-7.

96. Pichon C, Felden B. Proteins that interact with bacterial small RNA regulators. Fems Microbiology Reviews 2007;31(5):614-625.

97. Ramos JL, Duque E, Gallegos MT, Godoy P, Ramos-Gonzalez MI, Rojas A, Teran W, Segura A. Mechanisms of solvent tolerance in gram-negative bacteria. Annu Rev Microbiol 2002;56:743-68.

98. Ravagnani A, Jennert KCB, Steiner E, Grunberg R, Jefferies JR, Wilkinson SR, Young DI, Tidswell EC, Brown DP, Youngman P and others. Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. Molecular Microbiology 2000;37(5):1172-1185.

99. Rice JB, Balasubramanian D, Vanderpool CK. Small RNA binding-site multiplicity involved in translational regulation of a polycistronic mRNA. Proceedings of the National Academy of Sciences of the United States of America 2012;109(40):E2691-E2698.

100. Sardessai Y, Bhosle S. Tolerance of bacteria to organic solvents. Res Microbiol 2002;153(5):263-8.

101. Schiel B, Nold N, Dürre P: Identification of a small noncoding RNA in Clostridium acetobutylicum. Programme for the 3rd Joint Conference of the German Society for Hygiene and Microbiology (Jahrestagung der Deutschen Gesellschaft für Hygiene und Mikorobiologie [DGHM]) and the Association for General and Applied Microbiology (Jahrestagung der Vereinigung für Allgemeine und Angewandte Mikrobiologie [VAAM]), 28 to 31 March 2010. Hannover, Germany: GRV09:135; 2010.

102. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 2008;3(6):1101-8.

103. Scotcher MC, Huang KX, Harrison ML, Rudolph FB, Bennett GN. Sequences affecting the regulation of solvent production in Clostridium acetobutylicum. Journal of Industrial Microbiology & Biotechnology 2003;30(7):414-420.

104. Shao L, Hu S, Yang Y, Gu Y, Chen J, Jiang W, Yang S. Targeted gene disruption by use of a group II intron (targetron) vector in Clostridium acetobutylicum. Cell Res 2007;17(11):963-5.

105. Shin JH, Price CW. The SsrA-SmpB ribosome rescue system is important for growth of Bacillus subtilis at low and high temperatures. Journal of Bacteriology 2007;189(10):3729-3737.

106. Sillers R, Al-Hinai MA, Papoutsakis ET. Aldehyde-Alcohol Dehydrogenase and/or Thiolase Overexpression Coupled With CoA Transferase Downregulation Lead to Higher Alcohol Titers and Selectivity in Clostridium acetobutylicum Fermentations. Biotechnology and Bioengineering 2009;102(1):38-49.

107. Solovyev V, Salamov A. In: Li, RW, editor. Automatic Annotation of Microbial Genomes and Metagenomic Sequences: Nova Science Publishers; 2011. 61-78 p.

108. Sonenshein AL. In: Smith I, Slepecky RA, Setlow P, editors. Regulation of Prokaryotic Development. Washington, DC: ASM Press; 1989. p 109-130.

109. Soutourina OA, Monot M, Boudry P, Saujet L, Pichon C, et al. (2013) Genome-Wide Identification of Regulatory RNAs in the Human Pathogen Clostridium difficile. PLoS Genet 9(5): e1003493. doi:10.1371/journal.pgen.1003493

110. Storz G, Vogel J, Wassarman KM. Regulation by Small RNAs in Bacteria: Expanding Frontiers. Molecular Cell 2011;43(6):880-891.

111. Sullivan L, Bennett GN. Proteome analysis and comparison of Clostridium acetobutylicum ATCC 824 and Spo0A strain variants. J Ind Microbiol Biotechnol 2006;33(4):298-308.

112. Svetlanov A, Puri N, Mena P, Koller A, Karzai AW. Francisella tularensis tmRNA system mutants are vulnerable to stress, avirulent in mice, and provide effective immune protection. Molecular Microbiology 2012;85(1):122-141.

113. Thormann K, Durre P. Orf5/SolR: a transcriptional repressor of the sol operon of Clostridium acetobutylicum? Journal of Industrial Microbiology & Biotechnology 2001;27(5):307-313.

114. Thormann K, Feustel L, Lorenz K, Nakotte S, Durre P. Control of butanol formation in Clostridium acetobutylicum by transcriptional activation. Journal of Bacteriology 2002;184(7):1966-1973.

115. Tomas CA, Alsaker KV, Bonarius HP, Hendriksen WT, Yang H, Beamish JA, Paredes CJ, Papoutsakis ET. DNA array-based transcriptional analysis of asporogenous, nonsolventogenic Clostridium acetobutylicum strains SKO1 and M5. J Bacteriol 2003;185(15):4539-47.

116. Tomas CA, Beamish J, Papoutsakis ET. Transcriptional analysis of butanol stress and tolerance in Clostridium acetobutylicum. Journal of Bacteriology 2004;186(7):2006-2018.

117. Tomas CA, Welker NE, Papoutsakis ET. Overexpression of groESL in Clostridium acetobutylicum results in increased solvent production and tolerance, prolonged metabolism, and changes in the cell's transcriptional program. Applied and Environmental Microbiology 2003;69(8):4951-4965.

118. Tracy BP, Jones SW, Fast AG, Indurthi DC, Papoutsakis ET. Clostridia: the importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. Current Opinion in Biotechnology 2012;23(3):364-381.

119. Tracy BP, Jones SW, Papoutsakis ET. Inactivation of sigma(E) and sigma(G) in Clostridium acetobutylicum Illuminates Their Roles in Clostridial-Cell-Form Biogenesis, Granulose Synthesis, Solventogenesis, and Spore Morphogenesis. Journal of Bacteriology 2011;193(6):1414-1426.

120. Tummala SB, Junne SG, Papoutsakis ET. Antisense RNA downregulation of coenzyme A transferase combined with alcohol-aldehyde dehydrogenase overexpression leads to predominantly alcohologenic Clostridium acetobutylicum fermentations. J Bacteriol 2003;185(12):3644-53.

121. Tummala SB, Junne SG, Paredes CJ, Papoutsakis ET. Transcriptional analysis of product-concentration driven changes in cellular programs of recombinant Clostridium acetobutylicumstrains. Biotechnol Bioeng 2003;84(7):842-54.

122. Tummala SB, Welker NE, Papoutsakis ET. Development and characterization of a gene expression reporter system for Clostridium acetobutylicum ATCC 824. Appl Environ Microbiol 1999;65(9):3793-9.

123. Tummala SB, Welker NE, Papoutsakis ET. Design of antisense RNA constructs for downregulation of the acetone formation pathway of Clostridium acetobutylicum. J Bacteriol 2003;185(6):1923-34.

124. Tyurin M, Padda R, Huang KX, Wardwell S, Caprette D, Bennett GN. Electrotransformation of Clostridium acetobutylicum ATCC 824 using high-voltage radio frequency modulated square pulses. J Appl Microbiol 2000;88(2):220-7.

125. Vanderpool CK, Balasubramanian D, Lloyd CR. Dual-function RNA regulators in bacteria. Biochimie 2011;93(11):1943-1949.

126. Venkataramanan KP, Jones SW, McCormick KP, Kunjeti SG, Ralston MT, Meyers BC, Papoutsakis ET. The Clostridium small RNome that responds to stress: the paradigm and importance of toxic metabolite stress in C. acetobutylicum. Bmc Genomics 2013;14.

127. Walter KA, Bennett GN, Papoutsakis ET. Molecular Characterization of 2 Clostridium acetobutylicum ATCC 824 Butanol Dehydrogenase Isozyme Genes. Journal of Bacteriology 1992;174(22):7149-7158.

128. Walter KA, Bennett GN, Papoutsakis ET. Molecular characterization of two Clostridium acetobutylicum ATCC 824 butanol dehydrogenase isozyme genes. J Bacteriol 1992;174(22):7149-58.

129. Wang QH, Venkataramanan KP, Huang HZ, Papoutsakis ET, Wu CH. Transcription factors and genetic circuits orchestrating the complex, multilayered response of Clostridium acetobutylicum to butanol and butyrate stress. Bmc Systems Biology 2013;7:17.

130. Wang Y, Li X, Blaschek HP. Effects of supplementary butyrate on butanol production and the metabolic switch in Clostridium beijerinckii NCIMB 8052: genome-wide transcriptional analysis with RNA-Seq. Biotechnol Biofuels 2013;6(1):138.

131. White AP, Weljie AM, Apel D, Zhang P, Shaykhutdinov R, Vogel HJ, Surette MG. A global metabolic shift is linked to Salmonella multicellular development. PLoS One 2010;5(7):e11814.

132. Wiesenborn DP, Rudolph FB, Papoutsakis ET. Coenzyme A transferase from Clostridium acetobutylicum ATCC 824 and its role in the uptake of acids. Appl Environ Microbiol 1989;55(2):323-9.

133. Wietzke M, Bahl H. The redox-sensing protein Rex, a transcriptional regulator of solventogenesis in *Clostridium acetobutylicum*. Appl Microbiol Biotechnol 2012;96:749-761.

134. Wright PR, Georg J, Mann M, Sorescu DA, Richter AS, Lott S, Kleinkauf R, Hess WR, Backofen R. CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. Nucleic Acids Res 2014;42(Web Server issue):W119-23.

135. Xu M, Zhao J, Yu L, Tang IC, Xue C, Yang ST. Engineering Clostridium acetobutylicum with a histidine kinase knockout for enhanced n-butanol tolerance and production. Appl Microbiol Biotechnol 2014.

136. Young M, Minton NP, Staudenbauer WL. Recent advances in the genetics of the clostridia. FEMS Microbiol Rev 1989;5(4):301-25.

137. Zhang L, Nie X, Ravcheev DA, Rodionov DA, Sheng J, Gu Y, Yang S, Jiang W, Yang C. Redox-Responsive Repressor Rex Modulates Alcohol Production and Oxidative Stress Tolerance in Clostridium acetobutylicum. J Bacteriol 2014;196(22):3949-63.

138. Zingaro KA, Papoutsakis ET. Toward a semisynthetic stress response system to engineer microbial solvent tolerance. MBio 2012;3(5).

139. Zingaro KA, Terry Papoutsakis E. GroESL overexpression imparts Escherichia coli tolerance to i-, n-, and 2-butanol, 1,2,4-butanetriol and ethanol with complex and unpredictable patterns. Metab Eng 2013;15:196-205.

Appendix A

EXPRESSION OF SOLB BY RNA DEEP SEQUENCING ANALYSIS

RNA Deep-Sequencing studies, both standard and strand-specific, (Venkataramanan, et al, 2013; Ralston, unpublished thesis) have investigated the transcriptome and small RNome expressed in *C. acetobutylicum* under butanol, butyrate, and no stress. Their data sets cover mid- to late-exponential phase but do not represent early growth or late stationary phase growth. Expression reads from these experiments consistently show very low, yet detectable, expression of the *solB* sequence (Figure A.1), and they reveal no significant expression in the intergenic region between *solB* and the transcriptional start sites for the *sol* operon. These data sets illustrate the impressive difference in expression level between the highly expressed *sol* operon (represented by *adhE1* in Figure A.1) and the minimally expressed *solR* and *solB*; nevertheless, expression of these latter sequences is detected.

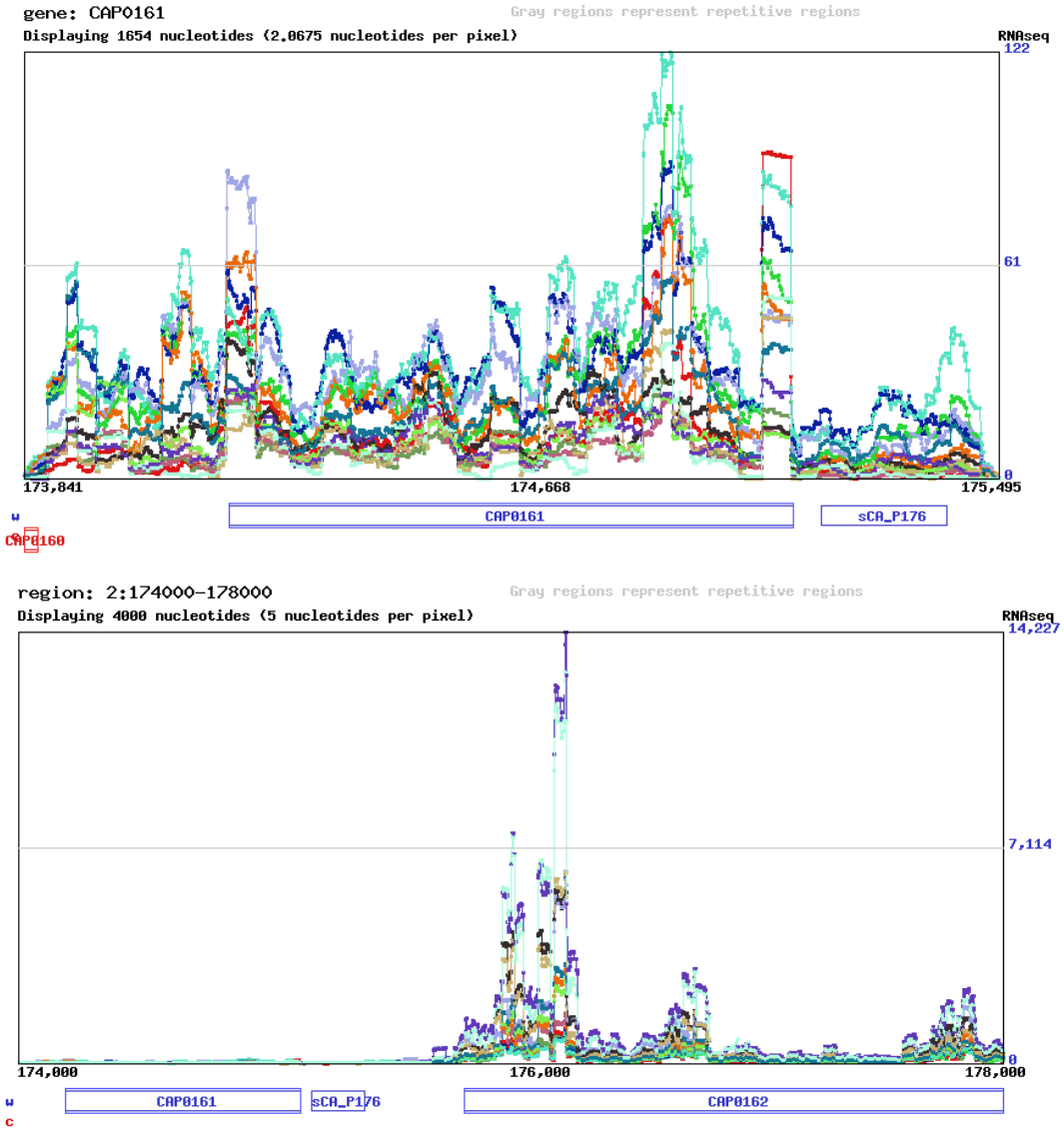


Figure A.1 – Total expression reads for the *solR-solB* region shown alone (above) and with the downstream *adhE1* sequence (below). Images taken from the custom Wasabi database software for standard RNA Deep-Sequencing data (Venkataramanan, et al, 2013); colored plotlines represent data sets from butyrate, butanol, and no stress at 15, 60, and 75 minutes post-stress; y-axis represents the number of reads at a given base-pair position indicated on the x-axis.

Appendix B

NORTHERN BLOT ANALYSIS OF SOL LOCUS GENES IN WILD TYPE AND PLASMID CONTROL STRAINS

Preliminary Northern blot experiments were run with RNA samples from both wild type and 824(p94MCS) control strain cultures (Figure B.1). Analysis for each *sol* locus gene revealed strong expression in both strains. Expression bands for each target were slightly more intense in 824(p94MCS) samples; this was expected, given the slight increase in solvent production observed for plasmid-transformed strains, including 824(p94MCS). In the interest of maximizing space on each membrane for experimental samples, wild type RNA samples were used subsequently as the sole control samples for each membrane.

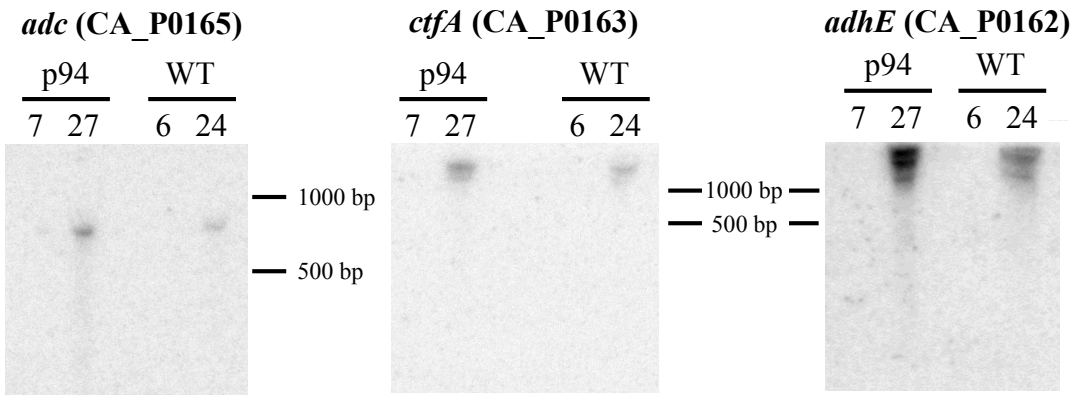


Figure B.1 – Preliminary Northern blot analysis of *sol* locus genes in ATCC 824 and 824(p94MCS) control strain. Samples of ATCC 824 (WT) and 824(p94MCS) (p94) were taken at exponential phase (A_{600} = 1) at 6 or 7 hours and at stationary phase (A_{600} = 4) at 24 or 27 hours; five µg of total RNA; note that larger sizes (1000+ bp) are compressed in these blots.

Appendix C

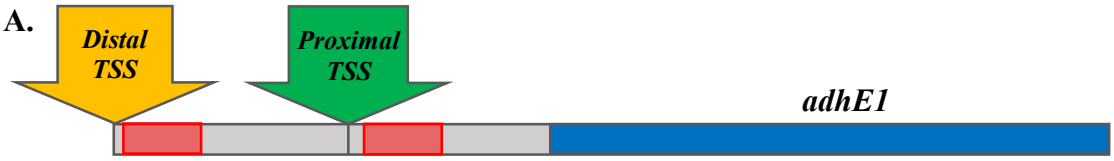
NORTHERN ANALYSIS OF THE DISTAL AND PROXIMAL 5' UNTRANSLATED REGIONS OF THE ADHE1 TRANSCRIPT

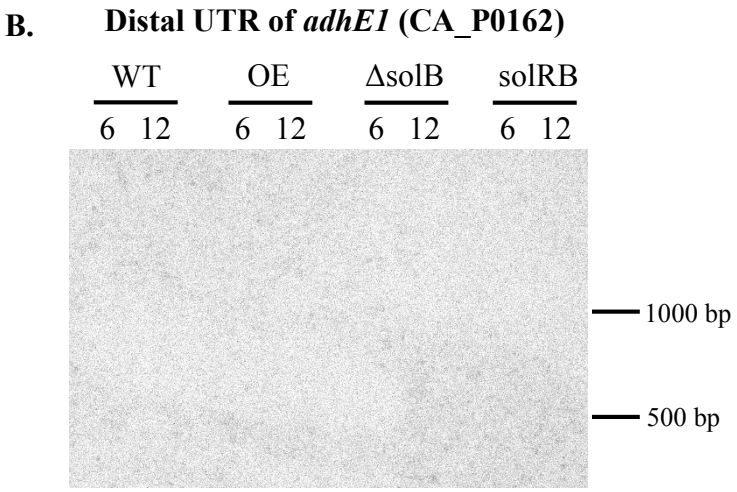
C.1 The 5' UTR of the *adhE1* Transcript only Appears in Wild Type and solRB Strains in Northern Blot Analysis

To investigate the possible role of target transcript UTRs on *solB* regulation, Northern blot analysis was performed using probes complementary to the two putative UTRs of the adhE1 transcript, generated from the distal and proximal promoter sequences (Figure C.1a).

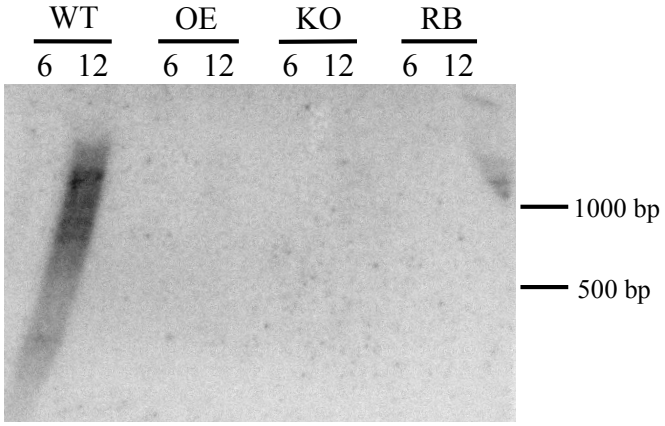
Northern analysis revealed no expression of the distal UTR in any strain tested (Figure C.1b), indicating that transcriptional activity at the distal UTR is either non-existent or below the detection limit of Northern analysis. This is in agreement with all previous observations regarding the distal promoter [sources]. However, analysis of the proximal UTR revealed expression but only in the wild type and solRB samples (Figure C.1c-d). In addition to the large band corresponding to the complete UTR+*adhE1* transcript, there appears to be several smaller fragments, including three bands approximately 700-900 bp in size and a band less than 500 bp in size. These observations are striking because the bands appear only in the solvent-producing strains, which may suggest an influence of the *adhE1* transcript UTR in the observed phenotypes and therefore its importance for successful translation and/or solvent production. The additional bands also suggest possible processing events surrounding the UTR sequence. Perhaps this UTR region contains a vital translation enhancer or ribosomal standby site. Whether the UTR possibly affects *adhE1* translation or interacts with *solB*, the details of this potential function remain to be discovered. It

could be possible that the UTR is removed or damaged in the *solB* overexpression or deletion strains, which may contribute to the phenotype of one or both of these strains.





C. Proximal UTR of *adhE1* (CA_P0162)



D. Proximal UTR of *adhE1* (CA_P0162)

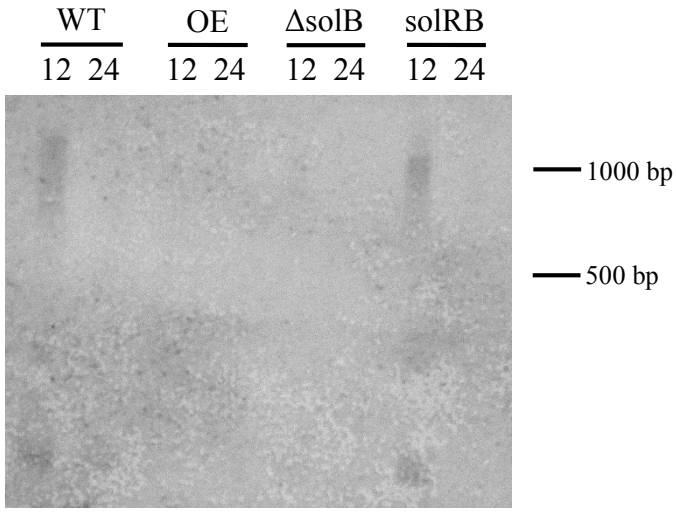


Figure C.1 – Investigation of the 5' UTR of the *adhE1* transcript in recombinant strains. (A) Schematic of *adhE1* UTR and locations of probe hybridization sites on the distal and proximal UTRs (red); Northern analysis of distal UTR expression at 6 and 12 hours (B) in ATCC 824 (WT), 824(p94_solB) (OE), 824ΔsolB, and solRB; Northern analysis of proximal UTR expression at 6 and 12 hours (C) and 12 and 24 hours (D), using the same samples as (B); fifteen μg of total RNA samples.