ENGINEERING *ACTINOBACILLUS SUCCINOGENES* FOR SUCCINATE PRODUCTION- A FOCUS ON SUCCINATE TRANSPORTERS AND SMALL RNAs

By

Rajasi Virendra Joshi

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Microbiology and Molecular Genetics- Doctor of Philosophy

2017

ABSTRACT

ENGINEERING *ACTINOBACILLUS SUCCINOGENES* FOR SUCCINATE PRODUCTION- A FOCUS ON SUCCINATE TRANSPORTERS AND SMALL RNAS

By

Rajasi Virendra Joshi

An important aspect of any industrial scale bio-based production is the choice of biocatalyst used. Many commercially relevant microorganisms and industrial strains have been engineered to optimize the production of bio-based chemicals. One such chemical is succinate, listed as one of the top 12 building block chemicals from biomass by the US Department of Energy. Succinate is considered an important platform chemical as it has a number of applications and, most importantly, is a precursor to high-volume value-added commodity chemicals. Bio-based succinate is currently being produced at industrial scale levels using engineered microorganisms such as *E. coli* and *S. cerevisiae*. *Actinobacillus succinogenes* is one of the best natural succinate producers, which can grow on a wide variety of substrates, and, with the advance in genetic tools, can possibly be engineered for increased succinate production.

Very few studies have focused on using succinate exporters as metabolic engineering targets for succinate production. Only a handful of studies have been carried out in *E. coli* and *C. glutamicum* with none whatsoever in *A. succinogenes*. With a combination of proteomics and transcriptomics we have identified candidate succinate transporters in *A. succinogenes*. Four of the top hits in our proteomics analysis were Asuc_1999, Asuc_0142, Asuc_2058 and Asuc_1990-91. To carefully tune the expression of these membrane proteins, we generated a library of promoters covering a large range of strengths below the strong, constitutive promoter (p_{pckA}) we had been using. Some of the promoters were truncated versions of p_{pckA}, and others were identified from our transcriptomics data. These promoters were tested using *lacZ* as the

reporter gene in an *A. succinogenes* Δ*lacZ* background. Promoters ranged from p_{pckA} as the highest down to p_{Asuc_0701} with a strength 209-fold lower than p_{pckA}. The four succinate transporter candidates over-expressed under $p_{pckA\text{-}92}$, a truncated version of p_{pckA}, increased the succinate yield in glucose cultures compared to the control strain carrying the empty vector.

Synthetic small RNAs (sRNAs) are another tool for metabolically engineering industrially relevant microorganisms. However, no sRNAs have been identified in *A. succinogenes* and only a few have been identified in other members of the Pasteurellaceae family. We identified sRNAs in *A. succinogenes* grown anaerobically on glucose and microaerobically on glycerol by RNA sequencing. We found 260 sRNAs in total, of which 39 were predicted by at least one of five computational programs. We validated 14 sRNAs identified from sequencing with RT-PCR. Additionally we identified probable Hfq-binding sRNAs through their Rho-independent terminators, a key feature of Hfq-binding sRNAs. Using additional characteristics of Hfq-binding sRNAs, we designed synthetic sRNAs targeting *lacZ* expression as a proof of concept. One plasmid-borne synthetic sRNA caused a 32% decrease in β-galactosidase activity in lactose-grown cultures. Using the same sRNAs scaffolds, we generated synthetic sRNAs that targeted the *ackA* and *pta* mRNAs to decrease the production of acetate, one of the major by-products of succinate production. One of the synthetic sRNAs targeting *ackA* caused a 14% decrease in the acetate yield of glucose-grown cultures.

In summary, we have identified candidate succinate transporters and seen an increase in succinate production upon their overexpression. In the process, we have developed a promoter library for tunable expression of genes in *A. succinogenes*. We have also shown that sRNAs can be used as a tool for metabolic engineering in *A. succinogenes,* although additional studies are needed to make it more tunable and robust.

This dissertation is dedicated to my entire family, especially my husband, Ashwin for his patience, constant encouragement and morale boosting pep talks. I thank my parents, sister, brother-in law and in-laws for always believing in me and supporting me throughout this journey.

ACKNOWLEDGEMENTS

This dissertation has been possible only because of the contributions of several people who have guided and helped me throughout these years. First and foremost, I would like to express my sincere gratitude to my mentor, Dr. Claire Vieille for giving me the opportunity to work in her lab during my Master's and for my PhD. I thank her for being a great mentor and for always taking out the time to discuss any idea, big or small, despite being busy with numerous other things. Her intellect, passion for science and positive attitude has always inspired and motivated me, especially during the more trying times of this project. I will always be indebted to her, for her guidance and thoughtful input throughout my PhD.

I am thankful to all the members of the Vieille lab- Maris Laiveneiks, Dr. Bryan Schindler, Dr. Justin Beauchamp, Dr. Nikolas Mcpherson and Peggy Wolf for a great work atmosphere with lots of support and help. I am truly grateful for the time taken by Dr. Bryan Schindler to train me during my initial years in the lab even when he was busy with his own experiments. I am also grateful to Maris Laivenieks, who was always willing to help me, especially during my initial struggles while building a chemostat and mixing gas tanks. I would like to thank Dr. Nikolas McPherson for being an excellent lab member and for all the helpful discussions we had while working with *Actinobacillus succinogenes* during the course of both of our PhDs. I must acknowledge Dr. Justin Beauchamp for helping me with the intial RNAseq analysis. I wish to thank Barry Tigner, who has saved my experiments numerous times, by fixing last minute instrument fails.

I must thank my committee members, Drs. Gemma Reguera, Michael Bagdasarian, Shannon Manning, Yair Shachar-Hill for their expert guidance and constructive feedback. I

would like to thank Dr. Donna Koslowsky for letting me use her lab and Laura Kirby, her grad student for helping me set up Northern blots.

I would like to thank all the undergraduates who helped me with my project, Alex Yang, Jean Kim, Maeva Bottex, Matthias Schulz, Kanupriya Tiwari, Joel Tencer, Jacob Gibson, Aaron Walters, Jack Peleman and Corey Kennelly. It was a pleasure interacting with each of them and training them has made me a better researcher.

I would like to thank Betty Miller, graduate secretary of Microbiology and Molecular Genetics for helping me navigate through the paperwork required and other assistance throughout these years.

I am grateful for the lasting friendships I have made here at MSU. I feel fortunate to have had a strong support system in them, sharing the ups and downs of graduate life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Bio-based succinate production

A huge push towards bio-based succinate production has taken place in the last ten years. Succinate, a C4-dicarboxylic acid, has consistently been among the top ten bio-based chemicals since 2004 (1-3). Succinate is currently produced by two methods. The majority of succinate is produced by catalytic hydrogenation of petroleum-based maleic anhydride, but bio-based succinate production by fermentation of glucose by natural/engineered succinate producers is also catching up, as a number of companies are either in the process of setting up or have already started large-scale fermentation facilities for succinic acid. The global production rate for succinate is between 30,000-50,000 metric tons per year. The market volume for succinate in 2011 was 40,000 metric tons, out of which only 1,150 metric tons were bio-based succinate (4). Market price for petroleum-derived succinate is between $6,000 and $9,000/metric ton (5). If all announced large-scale fermentation projects come to fruition, an estimated total of 140,000 metric tons per year of bio-based succinate will be available globally (5). Large-scale production facilities are estimated to produce a minimum of 642,450 metric tons in 2020.

As of today, four major companies have commercial-scale production facilities for bio-based succinic acid—BioAmber, Myriant, Succinity GmbH, and Reverdia. BioAmber is a partnership between DNP Green Technology and the French agricultural cooperative. In 2008, they started a succinic acid plant in Pomacle, France, with an annual capacity of 2,000 metric tons (Chemicals from Biomass) using *Escherichia coli* (licensed DuPont technology). In August 2015, BioAmber announced the opening of the world's largest succinate plant in Sarnia, Canada, as a joint venture with Mitsui & Co., Ltd. The company has hit its target for operational milestones, which were laid out at the start, and sold succinic acid worth $3.7 million in the third quarter of 2016

(Michael McCoy, ACS). BioAmber has two additional production plants planned or underway with an annual capacity of 65,000 metric tons in USA or Brazil and in Thailand (5).

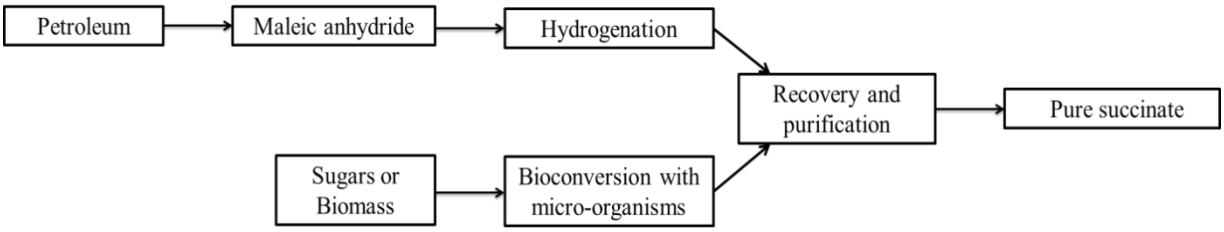

Figure 1.1 Routes to succinate production

Myriant started its flagship production facility in Lake Providence, Louisiana, in 2013 with a 13,607 metric tons succinate production capacity per year. It had plans to expand this production facility to have an additional 63,502 metric tons capacity per year (6) by 2015. In partnership with ThyssenKrupp Uhde, Myriant also started a production facility in Leuna, Germany, with a 1,500 metric tons production capacity per year (6). A jointly owned facility in Nanjing, China, between Myriant and China Nation BlueStar is being planned with a potential to scale up capacity to 100,000 metric tons per year (5).

Succinity GmbH is a joint venture between BASF and Corbian Purac. BASF has patents on a genetically engineered strain of *Basfia succiniciproducens* to produce succinate from either glycerol or mixed sugars while fixing the greenhouse gas CO_2. A production facility in Montmeló, Spain, has the capacity to produce 10,000 metric tons of succinate per year using crude glycerol as the feedstock. A larger 50,000 metric tons per year facility was planned but no specifics are known thus far (5).

Reverdia is a joint venture between Royal DSM, Netherlands, and Roquette Frères, France. In 2012, Reverdia started a production facility to produce 10,000 metric tons of succinate per year in Cassano Spinola, Italy, which uses the yeast *Saccharomyces cerevisiae* to convert commodity sugars to succinate (5).

1.2 Substrates and feedstocks for succinate production

Substrates and feedstocks used in fermentations for succinate production can significantly contribute to production costs and play a role in the economic feasibility of succinate production. Abundance, cost, and availability of substrates are some of the important factors that make a substrate more or less attractive for production. Recently, the focus has been more on using renewable sources as substrates to lower the costs of production. Succinate production has been studied from varied substrates such as glucose, sucrose, lactose, fructose, glycerol, lignocellulosic hydrolysates, whey, cassava, and sugarcane molasses (7-15).

1.3 Applications for succinate and its derivatives

Applications for succinate and its derivatives are many. Succinate is used to make a wide variety of products such as paints, coatings, adhesives, sealants, foods and flavors, cosmetics, nylons, industrial lubricants, phthalate-free plasticizers, dyes and pigments, as well as pharmaceuticals compounds (16). As bio-based succinate becomes more inexpensive, it can be used as a building block for a variety of chemicals such as adipic acid, 1,4-butanediol, γ-butyrolactone, tetrahydrofuran, n-methylpyrrolidone, poly-butylene succinate, 2-pyrrolidone, and polyamides (17, 18). About two-thirds of the bio-based succinate produced by year 2020 is expected to be used as an intermediate for producing 1,4-butanediol, tetrahydrofuran, and polyesters (19).

1.4 Succinate-producing microorganisms

Succinate is an important metabolite in cellular metabolism, as an intermediate metabolite in the tricarboxylic acid (TCA) cycle and the glyoxylate shunt. Some microbes also produce it as an end product during anaerobic fermentations. Most bacterial species that naturally produce large amounts of succinate, such as *Actinobacillus succinogenes*, *Mannheimia succiniciproducens*, and *B. succiniciproducens* were isolated from the cow rumen (20-22). These natural succinate producers have been studied and engineered to further enhance succinate production. *E. coli* has been engineered for succinate production by mimicking the metabolism of these natural producers. *E. coli*, *Corynebacterium glutamicum*, and *S. cerevisiae* have also been engineered for succinate production by disrupting the TCA cycle and forcing flux through the glyoxylate shunt. Huge advances have been made in both natural producers and engineered strains towards succinate production.

1.4.1 Native succinate producers

1.4.1.1 Anaerobiospirillum succiniciproducens

Anaerobiospirillum succiniciproducens was one of the first bacterial species studied for succinate production. When grown at pH 6.2 and high CO_2 concentrations, *A. succiniproducens* produced high amounts of succinate, as opposed to lactate (23). Continuous cultures of *A. succiniciproducens* grown on glycerol (supplemented with yeast extract), yielded 1.17 mol succinate per mol glycerol with a productivity of 2.1 g L^{-1} $h^{-}1$. Using a three-stage continuous cell recycle bioreactor, the succinate yield increased to 1.35 mol mol^{-1} glucose. Production rate and titer were 10.4 g L^{-1} h^{-1} and 83 g L^{-1}, respectively (24).

Several features make this Gram-negative species unfit for industrial succinate production, though. It is highly sensitive to oxygen, making it difficult to handle. It is known to cause rare, but potentially lethal cases of bacteremia and diarrhea in humans (25), and no genetic tools are available, making it an unsuitable host for engineering.

1.4.1.2 Actinobacillus succinogenes

A. succinogenes is a gram-negative, non-motile, osmotolerant, capnophilic, facultatively anaerobic bacterium that was isolated from a bovine rumen. It has the ability to grow on a broad range of substrates, such as glucose, lactose, xylose, arabinose, fructose, and glycerol (20). It belongs to the Pasteurellaceae family, along with *Haemophilus*, *Aggregatibacter*, *Mannheimia*, and *Pasteurella*. *A. succinogenes's* fermentation pathways are well studied. *A. succinogenes* produces phosphoenolpyruvate (PEP) from glucose via the Embden-Meyerhof Parnas pathway. PEP is then either dephosphorylated to pyruvate by pyruvate kinase, or converted into oxaloacetate by PEP carboxykinase (PEPCK, encoded by *pckA*) (26). These two branches are interconnected by two decarboxylating enzymes, malic enzyme and oxaloacetate decarboxylase (27). Oxaloacetate is further reduced to malate, malate to fumarate, and fumarate is eventually reduced to succinate (Figure 1.2). *A. succinogenes* has an incomplete TCA cycle that lacks citrate synthase and isocitrate dehydrogenase (28). It also lacks the glyoxylate shunt (27). Metabolic flux analyses and genome annotation have led to a deeper understanding of *A. succinogenes*'s pathways and metabolism (27, 29, 30). *A. succinogenes* is capable of natural transformation and a knockout method has been developed (31).

A. succinogenes does not ferment glycerol, but is can grow on glycerol by respiration. With nitrate as the terminal electron acceptor *A. succinogenes* produced acetate and CO_2 as the main

products, together with small amounts of succinate, but no ethanol. In contrast, succinate became the main secreted product of glycerol-grown cultures with dimethyl sulfoxide (DMSO) as the electron acceptor, with a yield of 59% of the maximum theoretical yield, with formate, acetate, and CO_2 as by-products. Similarly, succinate yields reached as high as 67% of the maximum theoretical yield under microaerobic conditions (1% O_2, batch), with acetate, CO_2 and formate as by-products. Deleting the pyruvate-formate lyase (Δ*pflB*) increased the succinate yield to 76% in microaerobic cultures (1% O_2, batch) (15). Carvalho et al. grew *A. succinogenes* on glycerol with DMSO as the external electron acceptor under fed-batch conditions, and obtained a succinate titer of 49.6 g L^{-1} and a succinate yield of 0.5 mol mol^{-1} glycerol (32).

The Nicol group developed a novel extended recycle biofilm reactor (continuous system) for fermentative succinate production. Using glucose as the substrate and CO_2 to maintain anaerobic conditions, fermentations were carried out with media supplemented with yeast extract and corn steep liquor. At a 0.56 h^{-1} dilution rate, the succinate yield reached 1.05 mol mol^{-1} glucose with a 6.35 g L^{-1} h^{-1} productivity (33). The succinate yield was shown to be an increasing function of glucose consumption in that system, and the succinate yield could be increased to 1.39 mol mol^{-1} glucose with a 48.5 g L^{-1} succinate titer (34). In a separate study, Guettler et al. overexpressed the native glucose-6-phosphate dehydrogenase (G6PDH) in a pyruvate–formate lyase mutant strain to direct more flux through the pentose phosphate pathway. Under anaerobic conditions, the recombinant strain produced succinate with productivity, titer, and yield of 2.19 g L^{-1} h^{-1}, 100.6 g L^{-1}, and 1.36 mol mol^{-1} glucose, respectively (35).

Vlysidis et al. first conducted batch experiments using a wide range of initial glycerol concentrations to test substrate inhibition on the process. One of the best results was found with a starting concentration of 37 g L^{-1} glycerol and yeast extract supplementation. These batch

cultures produced a succinate titer of 29.3 g L^{-1}, a productivity of 0.27 g L^{-1} h^{-1}, and a yield of 0.62 mol mol^{-1} glycerol. Next Vlysidis et al. developed an unstructured model by fitting a set of kinetic equations to the experimental data they found in the batch experiments done in small anaerobic reactors. This model was developed to predict experimental behavior and was validated by performing experiments in a scaled up bench top reactor. The model was able to predict the bench-top experiments without any additional fitting for a variety of conditions and can safely be used for future experiments (36, 37). An unstructured model was also developed to predict *A. succinogenes* fermentations on mixtures of C5 and C6 sugars found in spent sulfite liquor (38). Batch anaerobic fermentations were carried out using a mixture of xylose (72.6%), galactose, glucose, mannose, and arabinose with yeast extract. Succinate was produced with a titer of 27.4 g L^{-1}, productivity of 0.45 g L^{-1} h^{-1}, and yield of 0.7 g g^{-1} total sugars. Simulations were carried out using the kinetic parameters obtained from these experimental studies. Scaled up experiments using a 2-L lab-scale bioreactor using varying mixed sugar concentrations (15-50 g L^{-1}) were performed to validate the predictive nature of their model. The R^2 value of the model validations with the experimental results was 0.93, indicating that this model can effectively predict batch fermentations (38).

The Bechkam group has studied the behavior of *A. succcinogenes* on xylose-enriched hydrolysates, a feedstock being widely studied as a substrate for production of value-added chemicals. This group grew *A. succinogenes* on deacetylated dilute acid-pretreated corn stover hydrolysate and obtained a succinate yield of 0.74 g g^{-1} sugars. The maximum succinate productivity and titer were 1.27 g L^{-1} h^{-1} and 42.8 g L^{-1}, respectively (39). More recently, they have made Δ*ackA*, Δ*pflB* and Δ*pflB*Δ*ackA* strains of *A. succinogenes*. When grown on mock biomass hydrolysates all strains showed a decrease in succinate titer, yield and productivity

compared to the wild-type strain indicating that, removal of heterofermentative pathways may not lead to increase in flux towards succinate (40). In the same study, they also overexpressed the PEP carboxykinase, malate dehydrogenase and fumarase in the wild-type strain as well as in all three of the above strains. They did not see any significant increase in yield, productivity and titers for either of Δ*ackA*, Δ*pflB* and Δ*pflB*Δ*ackA* strains compared to the wild-type. All three of the wild-type strains expressing PEP carboxykinase, malate dehydrogenase and fumarase did show a slight increase in succinate yield and titer as compared to the wild-type strain. It is interesting to note that acetate accumulation was seen after a lag period in all *ackA* mutant backgrounds, suggesting an alternate route to acetate in *A. succinogenes*. Many other studies have tested for succinate production by *A. succinogenes* on different renewable carbon sources, such as straw hydrolysate, crop stalk waste, corn stover, cheese whey, and sugarcane bagasse (12, 41-44).

1.4.1.3 Mannheimia succiniciproducens

In 2002, Lee et al. reported the isolation and characterization of a novel succinate-producing Pasteurellaceae species, *Mannheimia succiniciproducens* MBEL55E, from cow rumen. *M. succiniciproducens* produces succinate as its major fermentation product, along with acetate, formate, and lactate as by-products (22). *M. succiniciproducens*'s genome was sequenced and its metabolism was extensively studied (45). Genome-based metabolic engineering of *M. succiniciproducens* was carried out to study its metabolism and increase succinate production (45). In-silico flux analysis identified PEPCK as the major CO_2-fixing enzyme instead of PEP carboxylase. PEPCK is the anaplerotic enzyme in the reductive TCA,

thus playing an important role in succinate production. Other enzymes responsible for formation of by-products during fermentation were also identified.

During anaerobic fermentation on glucose in batch conditions, the wild-type strain produced 10.5 g L^{-1} of succinate, 4.9 g L^{-1} of acetate, 4.1 g L^{-1} of formate, and 3.5 g L^{-1} of lactate (45). LPK7 is a metabolically engineered strain of *M. succiniciproducens*, carrying *ldhA*, *pflB*, *pta*, and *ackA* deletions (45). In fed-batch cultures this strain produced succinate with yield, productivity, and titer of 1.16 mol mol^{-1}glucose, 1.8 g L^{-1} h^{-1}, and 52.4 g L^{-1}, respectively. With glycerol as the carbon source, the PALK strain (Δ*ldhA* Δ*pta-ackA*) had a succinate yield of 0.88 mol mol^{-1} glycerol with an overall productivity of 0.13 g L^{-1} h^{-1}. The PALK strain grown in fed-batch conditions using glucose and glycerol as co-substrates along with 6.84 M magnesium hydroxide and 1.57 M ammonia for pH control gave a succinate yield of 1.15 mol mol^{-1} glucose equivalent, with an overall productivity of 3.5 g L^{-1} h^{-1} and a titer of 90.7 g L^{-1}, together with acetate (2.3 g L^{-1}) and pyruvate (4.0 g L^{-1}) as byproducts (46).

Recent simulations (47) based on omics studies and metabolism reconstructions predicted that sucrose and glycerol used as co-substrates would increase succinate production while reducing by-product formation. In this study, the PALKF (Δ*ldhA* Δ*pta-ackA* Δ*fruA*) strain was constructed to divert the majority of the flux towards succinate production, minimize by-product formation, and deregulate catabolite repression. The PALKF strain was grown in fed-batch conditions as low and medium density cultures. The succinate yields obtained were 1.56 and 1.64 mol mol^{-1} glucose equivalent with overall productivities of 2.50 and 6.02 g L^{-1} h^{-1} for low and medium density cultures, respectively. The PALKG strain (expressing the allosteric inhibition-free *E. coli glpK22*) (48, 49) was also developed, which gave a higher productivity (3.34 g L^{-1} h^{-1}) than the PALKF strain, but produced more by-products. A membrane cell recycle

bioreactor system was also developed in which the PALKF strain had a succinate productivity of 38.6 g L^{-1} h^{-1} (47). Overall, Dr. Sang Yup Lee's group was able to make a homo-succinate production possible with high productivities using sucrose and glycerol as co-substrates.

1.4.1.4 Basfia succiniciproducens

B. succiniciproducens DD1 was isolated from the bovine rumen and characterized as another succinate-producing *Pasteurellaceae* species (50). *B. succiniciproducens* DD1 and *M. succiniciproducens* MBEL55E had similar genome sizes, and their homologous proteins shared 95% similarity (21). When grown on glucose or sucrose as the carbon source, *B. succiniciproducens* DD1 produced succinate with a 5.8 g L^{-1} titer, a 1.5 g L^{-1} productivity, and a 0.6 g g^{-1} sucrose yield. With glycerol as the carbon source the succinate titer was higher, at 8.4 g L^{-1}, with a productivity of 0.9 g L^{-1} h^{-1} and a yield of 1.2 g g^{-1} glycerol (50). A continuous cultivation process was developed on crude glycerol supplemented with yeast extract and was maintained for 80 days. At a dilution rate of 0.018 h^{-1} the succinate titer was 5.21 g L^{-1}, the productivity was 0.094 g L^{-1} h^{-1}, and the yield was 1.02 g g^{-1} glycerol (50, 51). Metabolic flux analysis showed unwanted fluxes through pyruvate-formate lyase and lactate dehydrogenase. A Δ*ldhA* Δ*pflD* strain was developed that showed a succinate yield (1.08 mol mol^{-1} glucose) reaching 62% of the maximum theoretical yield (52). Succinity has a patent for the engineered strain and is using it for commercial scale production of succinate in its facilities.

1.4.2 Engineered succinate producers

1.4.2.1 Escherichia coli

E. coli is metabolically versatile and can consume glucose both anaerobically and aerobically. Under anaerobic conditions it carries out a mixed acid fermentation when grown on glucose to produce formate, lactate, ethanol, acetate, and succinate, where succinate is only a minor fermentation product (53). Under aerobic conditions it only produces succinate if the glyoxylate shunt is operational (54). The theoretical maximum yield for succinate in *E. coli* grown fermentatively on glucose is 1 mol mol^{-1} glucose (54). Wild-type *E. coli* produces only 0.11 mol succinate per mol glucose in these conditions (55). The succinate yield on glucose is generally limited by the availability of reducing equivalents (54). Although *E.coli* is not a natural succinate producer, strategies have been developed to improve succinate production in *E. coli* by engineering, evolution, or optimizing the production conditions. In one of the earlier studies by Millard et al., overexpressing PEP carboxylase (PEPC) increased succinate production 3.75-fold, but concomitantly decreased glucose uptake (56). A wild-type *E. coli* strain overexpressing a heterologous pyruvate carboxylase (from *Rhizobium etli*) produced 1.77 g L^{-1} succinate compared to 1.18 g L^{-1} by the wild-type strain (57). Strain AFP111 carrying knockout mutations in *pflB* and *ldhA* was unable to ferment glucose (53). A spontaneous mutation in the *pts*G gene restored the ability of AFP111 derivative NZN111 to ferment glucose while producing 1 mol of succinate, as well as 0.5 mol of acetate and ethanol each per mol of glucose (58). Strain AFP111(pTrc99A-*pyc*), which overexpresses the *R. etli* pyruvate carboxylase, produced 99.2 g L^{-1} of succinate with yield and productivity of 1.74 mol mol^{-1} glucose and 1.3 g L^{-1} h^{-1}, respectively, when grown in dual phase conditions (cultures grown aerobically to generate biomass, then transferred to anaerobic conditions conducive to succinate production) (59). Kim

et al., 2004 demonstrated that PEPCK can replace PEPC as the PEP-carboxylating enzyme in *E. coli*. Overexpressing the *A. succinogenes* 130Z PEPCK in *E. coli* K-12 *ppc*::kan increased succinate production 6.5-fold (60).

Dr. Lonnie Ingram's group combined metabolic evolution and gene knockouts to develop a good biocatalyst for succinate production. First the genes responsible for the formation of major by-products were knocked out, namely acetate kinase (*ackA*), lactate dehydrogenase (*ldhA*), and alcohol dehydrogenase (*adhE*). This strain, called KJ012, grew poorly on glucose minimal medium under anaerobic conditions. In this strain, however, ATP production was coupled to succinate production. Strain KJ012 was evolved to grow anaerobically in glucose minimal medium for over 2,000 generations to allow for growth-based selection. The evolved strain grew much better on glucose minimal medium under anaerobic conditions, and produced more succinate than its parent strain. Additionally, the genes *focA* and *pflB* were deleted as well, increasing the succinate production levels even higher. This new strain, KJ060, produced a succinate yield of 1.41 mol mol^{-1} glucose, with productivity and titer of 0.9 g L^{-1} h^{-1} and 86.6 g L^{-1} respectively. The succinate yield reached 1.61 mol mol^{-1} glucose when higher inoculums were used (61). KJ134 (Δ*ldhA* Δ*adhE* Δ*focA-pflB* Δ*mgsA* Δ*poxB* Δ*tdcDE* Δ*citF* Δ*aspC* Δ*sfcA* Δ*pta-ackA*) is a derivative of KJ060. This strain had additional deletions in *pta*, *tdcE*, and *tdcD* to further reduce acetate production. The carbon flux to oxaloacetate was also increased by knocking out genes coding for aspartate aminotransferase (*aspC*) and NADP-linked malic enzyme (*sfcA*). KJ134 produced 1.53 mol succinate per mol of glucose with a productivity of 0.75 g L^{-1} h^{-1} and a titer of 71.56 g L^{-1} (62). Additionally, strain KJ134 showed an 80% reduction in acetate production as compared to KJ073 (Δ*ldhA*, Δ*adhE*, Δ*ackA*, Δ*focA*, Δ*pflB*, Δ*mgsA*,

Δ*poxB*). When grown on 2% glycerol, KJ073 produced 15 g L^{-1} succinate from 16.9 g L^{-1} glycerol, a yield of 89% of the theoretical maximum (61).

During the evolution of strains KJ060 and KJ073, the authors noticed changes in metabolism that led to increased succinate production and ATP yield per glucose (63). They found a spontaneous mutation in the promoter of *pckA*, encoding gluconeogenic PEPCK, that allowed PEPCK to replace PEPC as the main carboxylating enzyme. In addition, glucose uptake through the PEP-dependent phosphotransferase system had been inactivated (spontaneous mutation) and replaced by galactose permease and glucokinase. These *E. coli* strains now had a pathway similar, for succinate production, to the natural succinate-producing organisms found in cow rumens.

The Ingram group confirmed that these changes led to higher succinate amounts by engineering an *E. coli* strain with the mutations seen above. Strain XZ647, with only two mutations (*ptsI* truncation and constitutive *pckA*), was able to produce succinate with a yield of 0.89 mol mol^{-1} of glucose. An additional *pflB* deletion yielded strain XZ721, whose succinate yield increased to 1.25 mol mol^{-1}glucose (64). When grown on glycerol, strain XZ721 had a yield of 0.8 mol mol^{-1} of glycerol (63, 65).

Dr. Ramon Gonzalez's group also engineered *E. coli* for succinate production from glycerol. Their approach included blocking the pathways to major by-products and overexpressing a heterologous pyruvate carboxylase to increase succinate production by driving the flux from pyruvate to oxaloacetate. The final strain, Δ*adhE* Δ*pta* Δ*poxB* Δ*ldhA* Δ*ppc* [pZS-*pyc*], overexpressed the *Lactobacillus lactis* pyruvate carboxylase and produced succinate with a yield of 0.54 mol mol^{-1} glycerol in microaerobic conditions (14).

1.4.2.2 Corynebacterium glutamicum

C. glutamicum is a gram-positive, non-motile, spore-forming, facultatively anaerobic bacterium that belongs to the Actinomycetes subdivision of eubacteria. It has been widely studied and used as an industrial strain for the production of amino acids and other organic acids (66). Okino et al. observed that *C. glutamicum* incubated in glucose-mineral medium under oxygen-limited conditions (without growth) produced succinate, lactate, and acetate (67). Increased bicarbonate concentrations led to an increased succinate yield and a decreased lactate yield. The ability of this species to produce organic acids in arrested growth conditions allowed for a bioreactor design in which high cell density led to a high succinate volumetric productivity of 11.7 g L^{-1} h^{-1} (67). The *C. glutamicum* strain Δ*ldhA*-pCRA717, a lactate dehydrogenase knockout that overexpresses pyruvate carboxylase, produced 146 g L^{-1} succinate in 46 h and the succinate and acetate yields were 1.4 and 0.29 mol mol^{-1} glucose, respectively (68). The succinate productivity was 3.2 g L^{-1} h^{-1}.

Litsanov et al. constructed the BOL-1 strain, in which all known pathways leading to acetate and lactate synthesis were deleted, and the BOL-2 strain where the pyruvate carboxylase gene was integrated chromosomally into the BOL-1 strain. Integrating the *Mycobacterium vaccae* NAD^+-coupled formate dehydrogenase into the BOL-2 chromosome gave strain BOL-3. Additionally, a metabolic blockage of glycolysis caused by NADH inhibition of GAPDH activity in glucose-grown BOL-3 was relieved by overexpressing the native glyceraldehyde-3-phosphate dehydrogenase gene on a plasmid. This final strain grown in the presence of glucose and formate in fed-batch conditions produced 134 g L^{-1} of succinate in 53 h. The succinate yield was 1.67 mol mol^{-1} glucose with very little accumulation of other by-products (0.1 mol mol^{-1} glucose) (69).

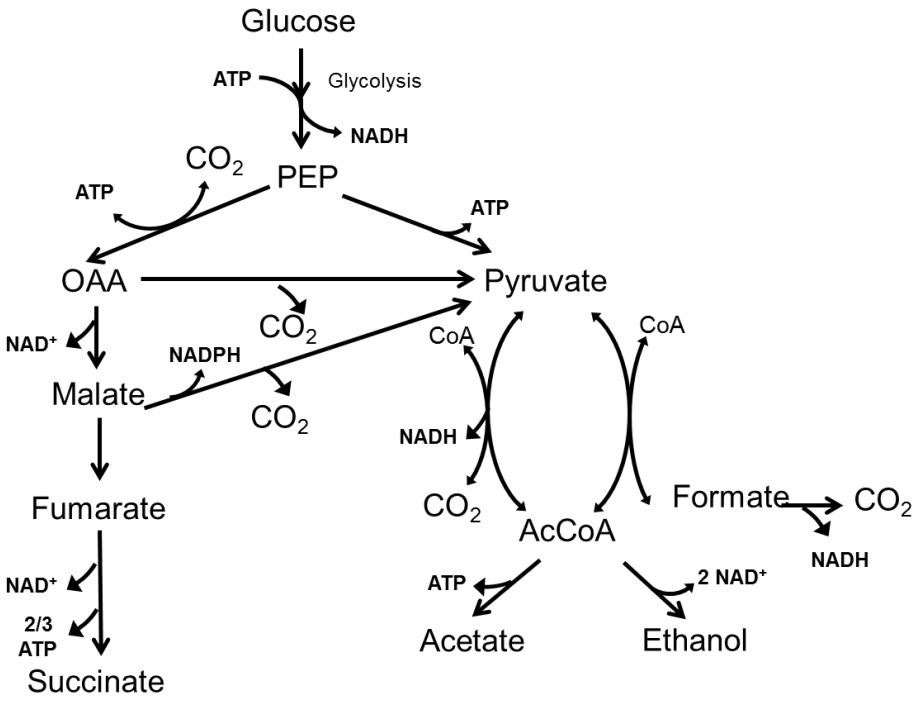

Figure 1.2 Metabolic map of *Actinobacillus succinogenes*. AcCoA, acetyl CoA; CoA, Coenzyme A; OAA, oxaloacetate; and PEP, phosphoenolpyruvate.

Another study came up with a dual synthesis approach for succinate production without using formate. The glyoxylate pathway citrate synthase genes were overexpressed to direct more carbon toward the glyoxylate pathway. The succinate exporter SucE was overexpressed as well. This strain produced 109 g L^{-1} succinate with an overall volumetric productivity of 1.11 g L^{-1} h^{-1} and a yield of 1.32 mol mol^{-1} glucose (70). All these studies were carried out under anaerobic conditions.

More recently, a few studies have focused on aerobic production of succinate in *C. glutamicum*. Strain BL-1 (pAN6-*pyc*^{P458S}*ppc*) has in-frame deletions of *pqo*, *pta-ackA*, *sdhCAB*, and *cat*, and overexpresses the pyruvate carboxylase (*pyc*) and PEPC (*ppc*) genes. This strain is the first known to produce succinate aerobically in minimal glucose medium. Grown under

conditions that decoupled succinate production from growth, this strain produced a titer of 10.6 g L^{-1} succinate with a yield of 0.45 mol mol^{-1} glucose (71). Strain BL-1(pVWEx1-*glpFKD*) overexpresses the *E. coli* glycerol facilitator, glycerol kinase, and glycerol-3-phosphate dehydrogenase. When grown aerobically on glycerol minimal medium, it produced 9.3 g L^{-1} of succinate with a volumetric productivity of 0.42 g L^{-1} h^{-1} (72). In another study, *B. subtilis* acetyl-coA synthetase was overexpressed on a plasmid in the base strain ZX1 to recycle the carbon in acetate. The ZX1 strain has deletions of *ldhA*, *pqo*, *cat*, and *pta* genes along with replacement of the native promoters of the *pyc* and *ppc* with the *sod* promoter. Strain ZX1(pEacsA) did not secrete any acetate, and the succinate yield was 0.5 mol mol^{-1} glucose. Overexpressing citrate synthase (encoded by *gltA*) led to an additional 22% increase in succinate yield. When strain ZX1(pEacsAgltA) was grown in fed-batch conditions on glucose it produced 28.4 g L^{-1} of succinate with a volumetric productivity of 0.41 g L^{-1} h^{-1} and a yield of 0.63 mol mol^{-1} glucose. This study, however, did use rich medium components in their production medium (73).

A recent study reported the development of an engineered strain of *C. glutamicum*, S071 (*ΔldhA Δpta-ackA ΔactA ΔpoxB* pyc^{P458S} P_{tuf}::*ppc Δpck*_P_{tuf}::*pckG ΔptsG*) able to produce 152.2 g L^{-1} of succinate with a yield of 1.67 mol mol^{-1} glucose by over-expressing the transcriptional regulator *NCgl0275*. Over-expressing *NCgl0275* allowed the release of an end-product inhibition due to succinate by increasing the glucose consumption rate (74).

1.4.2.3 Saccharomyces cerevisiae

S. cerevisiae is the only yeast that has been well studied for succinate production. Succinate is the main component that imparts flavor to Sake during fermentation (75). Arikawa et al.

studied the metabolic pathways leading to succinate production in *S. cerevisiae* and established that succinate could be produced by either n-ketoglutarate oxidation or fumarate reduction. In one of the first engineering studies of *S. cerevisiae*, inactivation of succinate dehydrogenase (*sdh* and *sdh1b* deletions) generated a strain with reduced malate productivity, and a succinate productivity about double that of the wild-type strain (76). Another group deleted *sdh1*, *sdh2*, *idh1*, and *idp1* to completely abolish isocitrate dehydrogenase and succinate dehydrogenase activities, diverting the TCA cycle flux into the glyoxylate pathway, and increasing succinate production. In glucose-shake flask cultures, this strain produced 3.62 g L^{-1} of succinate with a yield of 0.11 mol mol^{-1} glucose (77). Ito et al. (78) deleted *sdh1* and *sdh2* (encoding succinate dehydrogenase subunits) to allow for aerobic succinate production. To further increase succinate production they deleted the ethanol-producing genes, *adh1*, *adh2*, *adh3*, *adh4*, and *adh5* (encoding alcohol dehydrogenases). This strain, S149sdh12, produced 20-fold more succinate than the control strain, with a yield of 0.22 mol mol^{-1} glucose. Since succinate was accumulating intracellularly, they overexpressed the *Schizosaccharomyces pombe* malic acid transporter (*mae1*) in this engineered strain and further increased the succinate yield to 0.24 mol mol^{-1} glucose.

Yan et al. engineered *S. cerevisiae* for succinate production through the reductive branch of the TCA cycle. In *S. cerevisiae*, *fum1* encoding fumarate hydratase irreversibly converts fumarate to malate, which poses a problem for reductive succinate production. The authors deleted *fum1* and overexpressed *E. coli fumC* in its place, in a pyruvate decarboxylase-deficient (TAM) strain. The *pyc2*, *mdh3*, and *frd1* genes were also overexpressed to increase pyruvate decarboxylase, malate dehydrogenase, and fumarate reductase activity, respectively. GPD1 (glycerol-3-phosphate dehydrogenase) is responsible for NADH-dependent glycerol production

in aerobic conditions (79). Since production of 1 mol of succinate consumes 2 mol of NADH through the reductive TCA branch, the authors suspected that inactivating GPD1 would potentially increase succinate production. When grown in a bioreactor with optimal CO_2 (10%) and medium conditions, the final strain, PMCFfg, produced succinate with a titer of 13 g L^{-1} and a yield of 0.21 mol mol^{-1} glucose at a pH of 3.8 (79).

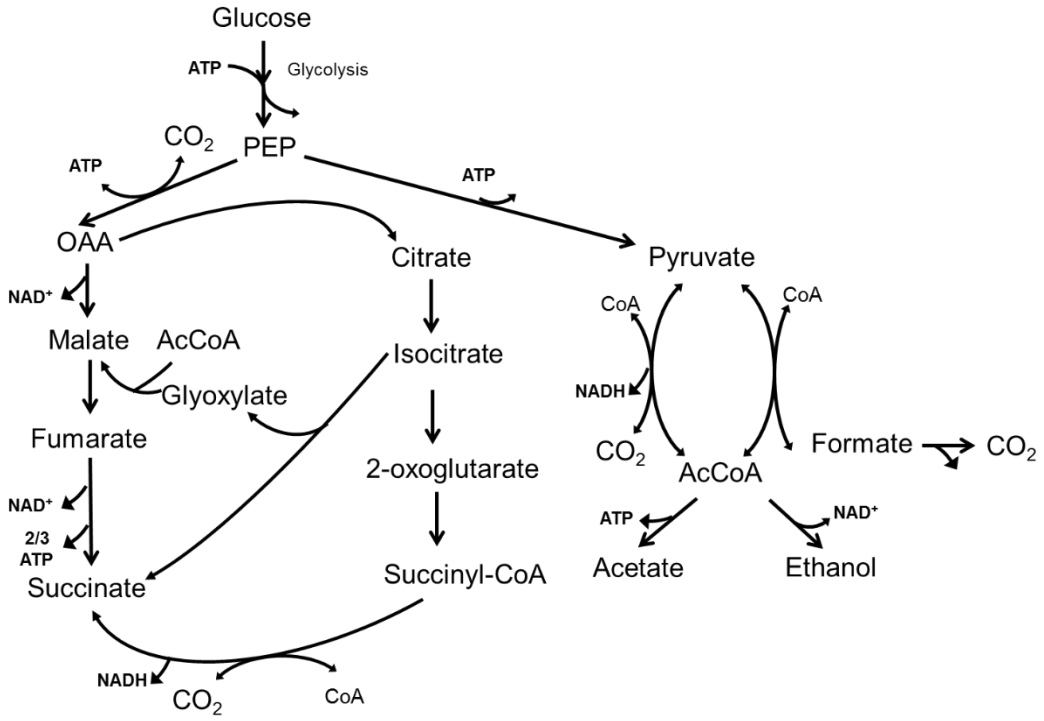

Figure 1.3 Pathways to succinate production by mixed acid metabolism in bacteria. Abbreviations are the same as in Figure 1.2.

Table 1.1 Succinate-producing strains with their yields, productivities, and titers

Strain	Conditions/Substrate	Succinate			Other products	Reference
		Titer ($g\ L^{-1}$)	Yield ($mol\ mol^{-1}$)	Productivity ($g\ L^{-1}\ h^{-1}$)		
Actinobacillus succinogenes						
130Z	D, CC, glycerol, Mi	31.7	0.75	0.14	A, F	(15)
130Z	D, glycerol, B, DMSO, An, YE	24.4	0.74	2.13	A, F	(32)
130Z	D, glycerol, FB, DMSO, An, YE	49.6	0.50	0.96	A, F	(32)
130Z	D, glucose, YE, CSL, An	48.5	1.39	ND	A, F	(34)
Mannheimia succiniciproducens						
PALK	D, FB, An, glucose + glycerol	90.7	1.15[a]	3.49	P, A	(46)
PALKF	D, FB, An, sucrose + glycerol, high cell concentration (Initial OD_{600} of 9.03)	78.4	1.64[a]	6.03	P, A	(47)
PALKF	D, MRCB (dilution rate of 2.93l h^{-1}), An, sucrose + glycerol	13.2	1.22[a]	38.6	P, A	(47)
Basfia succiniciproducens DD1						
LU15224	D, B, An, glycerol + maltose	69.8	1.70	2.91	A, P, L, M, F	(78, 80, 81)
LU15224 pJFF224 (*icl ms Y.m.)*	D, B, An, glucose, YE	46.3	1.33	ND	A	(80)
Escherichia coli						
AFP111(pTrc99A-*pyc*)	D, FB[c], YE, T	99.2	1.74	1.3	A, E	(59)
KJ060	D, B, glucose	86.6	1.41	0.9	A, M, L	(61)
	D, B, glycerol	11.7	0.89	ND	ND	(61)
KJ134	D, B, glucose	71.5	1.53	0.75	A, P, M	(62)
XZ465	D, B, glycerol	8.75	0.77	ND	ND	(65)
Corynebacterium glutamicum						
Δ*ldhA*-pCRA717	D[b], B, OD, glucose	146	1.4	3.2	A, L, M, P	(68)
BOL-3/pAN6-*gap* (Δ*cat*, Δ*pqo*, Δ*pta-ack*, Δ*ldh*, pyc^{P4585}, *fdh*)	D[b], FB, An, glucose + formate	134	1.67	2.48	K, M, A, F, P	(69)
Saccharomyces cerevisiae						
S149sdh12/pNV11-*mae1*	D, B, Ae, glucose	ND	0.24	ND	G, E, Fum	(78)
Δ*sdh1*Δ*sdh2*Δ*idh1*Δ*idp1*	D, B, Ae, glucose	3.62	0.11	ND	K, P, C	(77)
PMCFfg	D, B, Ae, glucose	13.0	0.21	ND	G, P	(79)

Abbreviations: A, acetate; Ae, aerobic; An, anaerobic; B, batch; C, Citrate; CC, continuous culture; CSL, corn steep liquor; D, defined; DMSO, dimethyl sulfoxide; E: ethanol; F, formate; Fum, fumarate; FB, fed-batch; G, Glycerol; K, α-ketoglutarate; L, lactate; M, Malate; Mi, microaerobic; MCRB, membrane cell recycling bioreactor; ND: Not determined; OD, oxygen deprivation (dissolved oxygen lower than 0.01 ppm); P, pyruvate; T, tryptone; VMD, vacuum membrane distillation; and YE, yeast extract.

[a] $mol\ mol^{-1}$ glucose equivalent

[b] aerobic cell propagation and transfer to anaerobic succinate production conditions

[c] dual phase fed-batch where cells were grown in aerobic conditions, then oxygen-free CO_2 was sparged at transition point

1.5 C4-dicarboxylate transporters

Engineering of microbial species for succinate production has mostly focused on targets within intracellular metabolic pathways. Very few engineering studies have focused on expressing succinate transporters to increase succinate production. *E. coli* and *C. glutamicum* are two commercially relevant strains in which succinate transporters have been characterized and used to increase succinate production. C4-dicarboxylates such as succinate, fumarate, and malate are intermediates in the TCA cycle and serve as carbon/energy sources during aerobic growth and therefore require uptake systems (82). During anaerobic growth, all C4-dicarboxylates other than succinate are used as electron acceptors in fumarate respiration. In contrast, succinate is exported via antiport against the uptake of malate or fumarate (82). Many different types of transporters are involved in C4-dicarboxylate transport and, with a few exceptions, most studies to identify these transporters have focused on *E. coli* and on C4-dicarboxylate uptake.

1.5.1 *E. coli* C4-dicarboxylate transporters

So far, *E. coli* has seven known C4-dicarboxylate transporters —DauA, DctA, DcuA, DcuB, DcuC, YjjPB, and CitT. Unidirectional uptake of fumarate, succinate, malate, and aspartate takes place under aerobic conditions, while exchange, uptake, and efflux of these C4-dicarboxylates take place in anaerobic conditions (83). DauA and DctA are primarily known to function in aerobic conditions while all others are active in anaerobic conditions. DauA (dicarboxylic acid uptake system A) has been characterized as a succinate uptake transporter active under acidic pH conditions (84). DctA, a dicarboxylate amino acid-cation symporter family (DAACS) transporter is used for aerobic growth in many bacteria such as *E. coli*, *Bacillus subtilis*, *Rhizobium leguminosarum*, and *C. glutamicum* (85-87). It mediates the uptake of

succinate, malate and fumarate in *E. coli* under aerobic conditions. DcuA, DcuB, and DcuC are antiporters found only in anaerobic and facultatively anaerobic bacteria and are involved in fumarate respiration. The DcuAB and DcuC families are independent from each other, but members of both families are capable of uptake, exchange, and efflux of C4-dicarboxylates. The main functions of DcuA, DcuB, and DcuC are still not completely elucidated. An early study determined that DcuAB carriers mostly functioned as uptake or exchange carriers whereas DcuC carriers likely mostly operated as efflux carriers (82). Still, an *E. coli* DcuC mutant grown anaerobically was not affected in succinate production, indicating that DcuC is not the main succinate exporter in anaerobic conditions (88). Another *E. coli* study showed that individual deletions of *dcuB* and *dcuC* decreased succinate titer by 15 and 11%, and a Δ*dcuB*Δ*dcuC* double mutant showed a 90% decrease in succinate titer in a glucose fermentation. Modulating the expression of these two transporters in *E. coli* also increased the succinate titer by 34% (89).

Another family of transporters involved in C4-dicarboxylate transport is the TRAP (tripartite-ATP independent periplasmic) family. In TRAP carriers, the driving force for solute accumulation is an electrochemical ion gradient (90). TRAP carriers consist of two membrane integral proteins (DctQM) and a periplasmic solute-binding protein (DctP). The first characterized TRAP transporter was the *R. capsulatus* DctPQM system and homologs of this system were found to be present in a wide range of bacteria and archaea (90). The TRAP family carriers are involved in solute uptake and are considered unidirectional (91). For this reason, it is unlikely that they could be involved in succinate efflux.

Another known *E. coli* C4-dicarboxylate transporter is CitT, however CitT is not restricted to only C4-dicarboxylate transport. *E. coli* CitT is involved in citrate uptake in exchange for succinate, fumarate, or tartarate under anaerobic conditions (82, 92). More recently, a study by

Fukui et al. (88) demonstrated that *E. coli* YjjP and YjjB form a single complex involved in succinate export under both aerobic and anaerobic conditions. A YjjPB mutant showed a 70% decrease in succinate production suggesting that YjjPB is involved in succinate export. YjjPB were found to contain different structures and conserved domains when compared to other known C4-dicarboxylate transporters (88).

1.5.2 C4-dicarboxylate transporters in other succinate producers

BlastP searches with *E. coli* YjjP in *A. succinogenes*, *M. succiniciproducens*, and *E. aerogenes* predicted homologs with 50% (87%), 54% (89%) and 88% (90%) identity (similarity), respectively (88). Similar BlastP searches with *E. coli* YjjB identified homologs in *A. succinogenes*, *M. succiniciproducens*, and *E. aerogenes* with 50% (81%), 52% (83%), and 83% (97%) identity (similarity) (88). Whether these homologs function in succinate export is unknown yet.

The *C. glutamicum* proteome does not contain any homologs of the *E. coli* Dcu family succinate transporters. In a study of *C. glutamicum* transmembrane proteins, 57 membrane proteins with at least four trans-membrane domains were found to be significantly similar to *A. succinogenes* and *M. succiniciproducens* proteins. Eliminating those with predicted or known functions brought the list of candidates down to twenty. Construction of deletion mutants of twelve of these candidates allowed the identification of SucE. The *sucE* mutant produced 70% of the wild-type levels of succinate and thus may not be the main succinate exporter in anaerobic conditions. SucE is a secondary carrier like the Dcu family carriers in *E. coli* (93). Another study identified SucE through transcriptomics studies for *C. glutamicum* grown in anaerobic and microaerobic conditions (94). Unlike the Dcu family carriers SucE is only involved in succinate

export (93). In another study, SucE-overexpressing *C. glutamicum* produced 109 g L^{-1} of succinate and a yield of 1.32 mol mol^{-1} glucose as mentioned above (70). SucE homologs were found in *E. coli*, *A. succinogenes*, and *M. succiniciproducens* (93), but their function has not been tested.

The only identified succinate transporter in *A. succinogenes* is Asuc_0304. It was identified by complementation studies done in a C4-dicarboxylate transport deficient mutant *E. coli* strain IMW213. Asuc_0304 was identified to be an uptake C4-dicarboxylate transporter capable of fumarate or malate uptake with Na^+ symport, although it may not be the sole or main transporter under aerobic conditions (95).

1.6 sRNAs as synthetic regulators for metabolic engineering

Industrially-relevant microorganisms have been subject to heavy modifications to optimize yield, productivity, and titers of their products. Most engineering strategies are aimed at identifying target genes and either knocking them out or over-expressing them. Many groups also take a rational approach for evolving strains for growth under certain conditions. However, it is often not easy to manipulate microorganisms and therefore conduct any type of metabolic engineering. In addition, these methods can only target a limited number of genes at a time and are therefore unsuitable for applications on a genomic scale (96). Another disadvantage of overexpressing genes is the limited control over expression level and/or stability in the particular organism. There is a growing need for strategies that enable researchers to modulate gene expression or repression at the translational level.

Synthetic sRNAs are gaining tremendous momentum in research, as using them would allow us to overcome many of the challenges listed above. One of the major advantages of

synthetic sRNAs is the possibility of designing them with a rational approach, making it unnecessary to build individual strains to test the intended phenotypes of the knockout or overexpression of genes. A recent explosion of studies has started to decipher the different ways in which sRNAs modulate different genes in nature. The increased understanding of how sRNAs function has led to an interest in using them as synthetic regulators by mimicking in vivo gene regulation to target genes of interest. Synthetic small RNAs have been designed for *E. coli* in several studies (42, 96-102). Desai et al. (103) demonstrated that expressing non-coding antisense sRNAs (asRNAs) complementary to the butyrate kinase, phosphotransbutyrylase, and phosphotransacetylase mRNAs inhibited these enzymes' activities and increased acetone and butanol production. Kim et al. designed antisense RNAs to lower the expression of acetate kinase and phosphotransacetylase. This strategy partially reduced the mRNA levels of the target genes and blocked the synthesis of the enzymes (97). Nakashima et al. designed an asRNA with paired-end termini for conditional silencing of acetate kinase and phosphotransacetylase. The paired-end termini increased the asRNA stability and improved the conditional silencing in *E. coli* (98, 99). Negrete et al. also demonstrated that overexpressing the GadY sRNA in *E. coli* decreased acetate production, which in turn decreased the inhibitory effect acetate had on the strain's growth (99).

Several of the early discovered sRNAs were found to be Hfq-dependent. Hfq is an RNA-binding protein that acts as a chaperone by enabling the binding of sRNAs to their target RNAs. Hfq is also known to stabilize sRNAs and their target mRNAs, thereby increasing their half-lives (104). *E. coli* Hfq-dependent sRNAs have an mRNA base-pairing region, Hfq-binding region, and a Rho-independent terminator (Figure 1.4). The Hfq-binding region contains a consensus secondary structure that provides a scaffold for Hfq to bind. In one of the first studies that

implemented sRNAs for gene repression, artificial sequences complementary to the ribosomal binding sites and start codons of *ipp*, *ompA*, and *ompC* were inserted in between the stem loops at the 5' and 3' termini of the natural hfq-dependent MicF sRNA. These synthetic sRNAs could inhibit the production of OmpC, OmpA, and Ipp (100). In another *E. coli* study the known sRNA RhyB (involved in regulation of iron metabolism) was overexpressed with an arabinose inducible promoter. Overexpressing RhyB during growth on glucose led to a 7-fold increase in succinate production without citrate production. RhyB was shown to regulate *sdhCDAB* and *acnB*, which encode succinate dehydrogenase and aconitase, respectively, in the TCA cycle (101). Overexpressing RhyB in the *E. coli* DALRA strain (engineered for 5-aminolevulinic acid production) allowed for 16% higher 5-aminolevulinic acid production than the parent strain (42). Na et al. designed 130 synthetic sRNAs by using two separate parts—the scaffold sequence and the mRNA or target-binding sequence. Using this approach the authors were able to substantially increase the production of cadaverine and tyrosine in *E. coli*. At high cell densities, the best engineered strain produced a tyrosine titer of 21.9 g L^{-1}, whereas cadaverine production went up to 12.6 g L^{-1}, compared to 9.6 g L^{-1} for the starting strain (96).

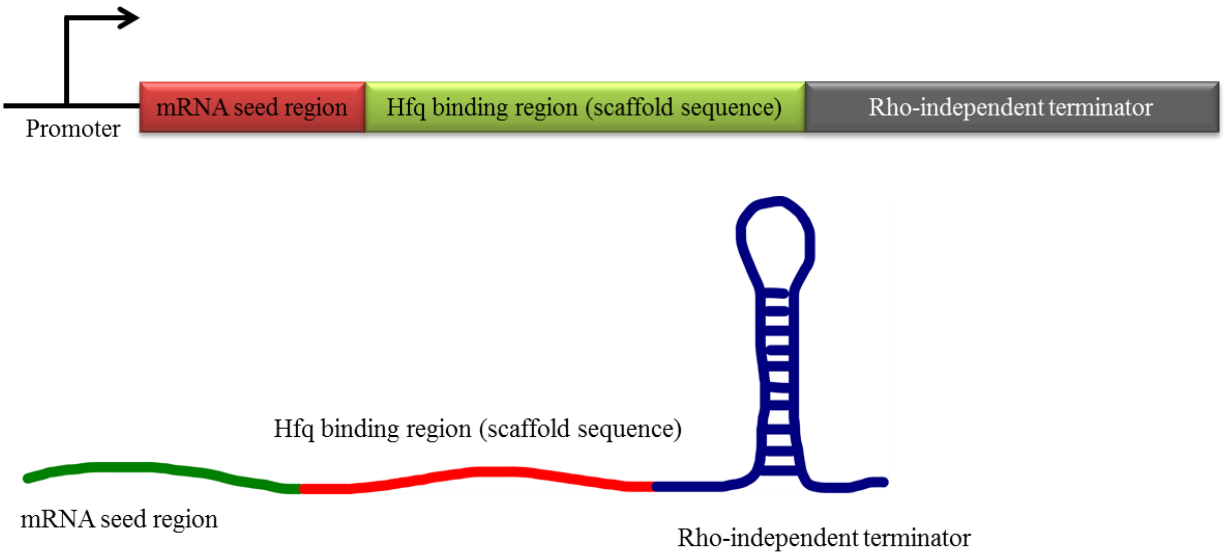

Figure 1.4 Hfq-dependent sRNA organization and structure in *E. coli*.

Designing sRNAs in *E. coli* is much more straightforward now with the vast amount of knowledge gained over the past decade. However, for other organisms where sRNAs are still being identified and the Hfq-binding scaffold sequence is not clearly understood or identified, designing synthetic, Hfq-dependent sRNAs is much more difficult. To engineer *B. subtilis*, Liu et al. (105) overcame these hurdles by designing synthetic sRNAs based on the *E. coli* MicC scaffold, and by coexpressing them with *E. coli* Hfq in *B. subtilis*. With this approach, the authors repressed the expression of *pfk* (encoding 6-phosphofructokinase) and *glmM* (encoding phosphoglucosamine mutase) to increase N-acetylglucosamine production. Cho et al. used the same approach in *Clostridium acetobutylicum* PJC4BK to repress phosphotransacetylase expression. Strain PJC4BK (pPta-Hfq_{Eco}) produced 16.9 g L^{-1} of butanol, which was higher than the starting strain as acetate production was inhibited (106). Research is underway to discover new sRNAs in many bacterial species.

1.7 Identification of small RNAs in Pasteurelleaceae

Pasteurellaceae are a family of Gram-negative bacteria that are mostly commensals on the mucosal surfaces of birds and mammals. The Pasteurellaceae family consists of eighteen genera, some of the major ones being *Haemophilus, Actinobacillus*, and *Pasteurella*. Only a handful of papers mention sRNAs in Pasteurelleaceae. Subashchandrabose et al. demonstrated the role of Hfq in pathogenesis of *Actinobacillus pleuropneumoniae*. A Δ*hfq* mutant of *A. pleuropneumoniae* was defective in biofilm formation and sensitive to oxidative stress; however not much more is known regarding the sRNAs involved (107). Three sRNAs have been identified in *Aggregatibacter actinomycetemcomitans* that are iron- and fur-regulated (108). Baddal *et al.* (109) identified seventeen novel putative sRNAs in intergenic regions, six of them homologous to known sRNA families in *Haemophilus influenzae*. HrrF, a fur-regulated sRNA was the first sRNA to be identified in any *Haemophilus* species (110).

More recently, the Bazzolli group conducted an extensive study to identify small RNAs in *A. pleuropneumoniae*, an organism known to cause porcine pleuropneumonia. The authors used four different algorithms to identify small RNAs and were able to experimentally confirm seventeen of the 23 sRNAs found by all four algorithms by Northern blotting, RT-PCR, and RNA sequencing. They also found that the sequences of these seventeen sRNAs were well conserved in the species that are evolutionarily close to *A. pleuropneumoniae* (111). These early studies have shed some much needed light in this unexplored area of sRNAs in Pasteurellaceae. However, much remains to be discovered.

1.8 Introduction to chapters

The main aims of this dissertation are to:

(i) provide insight into small RNAs and make progress in using small RNAs as a valuable tool for increasing succinate production in *A. succinogenes*.

(ii) make a promoter library for *A. succinogenes* for tunability of gene expression

(iii) carry out a proteomics- and transcriptomics-based search for succinate export transporters and see the effect of their overexpression on succinate production.

Chapter 2 describes a proteomics- and transcriptomics-based search for succinate exporters in *A. succinogenes*. Transporters with the most hits in our omics data were further tested by overexpression for succinate production. Tunability for expressing these transporters was required and so far no such system was available in *A. succinogenes*. For this purpose, a reporter assay system was developed in *A. succinogenes* to test promoter strength. Several promoters were tested for strength to generate a library of promoters of increasing strengths. The four candidate succinate exporters identified were expressed under control of a few of the weaker promoters to identify the transporter expression levels that would not interfere with *A. succinogenes* growth and test their effect on succinate production. .

Chapter 3 describes a sequencing study of *A. succinogenes* small RNAs conducted on cultures grown anaerobically on glucose and microaerobically on glycerol. The main aim of this study was to find Hfq-dependent small RNAs and compare their sequences to be able to design a synthetic small RNA for repression of mRNA in *A. succinogenes.*

Chapter 4 describes a method for making markerless knockout mutants in *A. succinogenes* using natural transformation.

Chapter 5 describes the work I did as a part of a previously published paper. I developed a continuous culture system for *A. succinogenes* growth on glycerol under microaerobic conditions.

In chapter 6, I summarize what was learned about succinate transporters and sRNAs in *A. succinogenes*. I also discuss future work for gaining more insight and further engineering of *A. succinogenes* for succinate production.

REFERENCES

REFERENCES

1. **Werpy TA**, **Petersen G**. 2004. Top value added chemicals from biomass for the U.S. Department of Energy (DOE) by National Renewable Energy Laboratory (NREL).

2. **Patel MK**, **Crank M**, **Dornburg V**, **Hermann B**, **Roes L**, **Hüsing B**, **Overbeek L**, **Terragni FR, E.** 2006. The BREW project: Medium and Long-term Opportunities and Risks of the Biotechnological Production of Bulk Chemicals from Renewable Resources–the Potential of White Biotechnology. Technology and Society (STS)/Copernicus Institute, Utrecht, the Netherlands.

3. **Bozell JJ**, **Petersen GR**. 2010. Technology development for the production of biobased products from biorefinery carbohydrates—the US Department of Energy's "top 10" revisited. Green Chem **12**:539-554.

4. WP 8.1. Determination of market potential for selected platform chemicals. Itaconic acid, succinic acid, 2,5-furandicarboxylic acid. Weastra, S. R. O.

5. **Biddy MJ**, **Scarlata C**, **Kinchin C**. 2016. Chemicals from Biomass: A Market Assessment of Bioproducts with Near-Term Potential. NREL (National Renewable Energy Laboratory, Golden, CO, USA.

6. www.myriant.com. Accessed

7. **Sawisit A**, **Jantama SS**, **Kanchanatawee S**, **Jantama K**. 2014. Efficient utilization of cassava pulp for succinate production by metabolically engineered *Escherichia coli* KJ122. Bioproc Biosyst Eng **38**:175-187.

8. **Chen C**, **Ding S**, **Wang D**, **Li Z**, **Ye Q**. 2014. Simultaneous saccharification and fermentation of cassava to succinic acid by *Escherichia coli* NZN111. Bioresour Technol **163**:100-105.

9. **Chatterjee R**, **Millard CS**, **Champion K**, **Clark DP**, **Donnelly MI**. 2001. Mutation of the *ptsG* gene results in increased production of succinate in fermentation of glucose by *Escherichia coli*. Appl Environ Microbiol **67**:148-154.

10. **Wang J**, **Zhu J**, **Bennett GN**, **San KY**. 2011. Succinate production from sucrose by metabolic engineered *Escherichia coli* strains under aerobic conditions. Biotechnol Prog **27**:1242-1247.

11. **Liu R**, **Liang L**, **Cao W**, **Wu M**, **Chen K**, **Ma J**, **Jiang M**, **Wei P**, **Ouyang P**. 2013. Succinate production by metabolically engineered *Escherichia coli* using sugarcane bagasse hydrolysate as the carbon source. Bioresour Technol **135**:574-577.

12. **Borges ER**, **Pereira Jr N**. 2011. Succinic acid production from sugarcane bagasse hemicellulose hydrolysate by *Actinobacillus succinogenes*. J Ind Microbiol Biotechnol **38**:1001-1011.

13. **Samuelov NS**, **Datta R**, **Jain MK**, **Zeikus JG**. 1999. Whey fermentation by *Anaerobiospirillum succiniciproducens* for production of a succinate-based animal feed additive. Appl Environ Microbiol **65**:2260-2263.

14. **Blankschien MD**, **Clomburg JM**, **Gonzalez R**. 2010. Metabolic engineering of *Escherichia coli* for the production of succinate from glycerol. Metab Eng **12**:409-419.

15. **Schindler BD**, **Joshi RJ**, **Vieille C**. 2014. Respiratory glycerol metabolism of *Actinobacillus succinogenes* 130Z for succinate production. J Ind Microbiol Biotechnol **41**:1339-1352.

16. www.bio-amber.com.

17. **Chen Y**, **Nielsen J**. 2016. Biobased organic acids production by metabolically engineered microorganisms. Current Opin Biotechnol **37**:165-172.

18. **Ahn J-H**, **Sang B-I**, **Um Y**. 2011. Butanol production from thin stillage using *Clostridium pasteurianum*. Bioresource Technol **102**:4934-4937.

19. **Cavani F**, **Albonetti S**, **Basile F**, **Gandini AA**. 2016. Chemicals and Fuels from Bio-Based Building Blocks. Wiley Online Library.

20. **Guettler MV**, **Rumler D**, **Jain MK**. 1999. *Actinobacillus succinogenes* sp. nov., a novel succinic-acid-producing strain from the bovine rumen. Int J Syst Bacteriol **49**:207–216.

21. **Kuhnert P**, **Scholten E**, **Haefner S**, **Mayor D**, **Frey J**. 2010. *Basfia succiniciproducens* gen. nov., sp. nov., a new member of the family Pasteurellaceae isolated from bovine rumen. Int J Syst Evol Microbiol **60**:44-50.

22. **Lee PC**, **Lee SY**, **Hong SH**, **Chang HN**. 2002. Isolation and characterization of a new succinic acid-producing bacterium, *Mannheimia succiniciproducens* MBEL55E, from bovine rumen. Appl Microbiol Biotechnol **58**:663–668.

23. **Samuelov NS**, **Lamed R**, **Lowe SE**, **Zeikus JG**. 1991. Influence of CO_2-HCO_3^- levels and pH on growth, succinate production, and enzyme activities of *Anaerobiospirillum succiniciproducens*. Appl Environ Microbiol **57**:3013–3019.

24. **Meynial‐Salles I**, **Dorotyn S**, **Soucaille P**. 2008. A new process for the continuous production of succinic acid from glucose at high yield, titer, and productivity. Biotechnol Bioengin **99**:129-135.

25. **Secchi C**, **Cantarelli VV**, **Pereira Fde S**, **Wolf HH**, **Brodt TC**, **Amaro MC**, **Inamine E**. 2005. Fatal bacteremia due to *Anaerobiospirillum succiniciproducens*: first description in Brazil. Braz J Infect Dis **9**:169-172.

26. **van der Werf MJ**, **Guettler MV**, **Jain MK**, **Zeikus JG**. 1997. Environmental and physiological factors affecting the succinate product ratio during carbohydrate fermentation by *Actinobacillus* sp. 130Z. Arch Microbiol **167**:332–342.

27. **McKinlay JB**, **Shachar-Hill Y**, **Zeikus JG**, **Vieille C**. 2007. Determining *Actinobacillus succinogenes* metabolic pathways and fluxes by NMR and GC-MS analyses of ^{13}C-labeled metabolic product isotopomers. Metab Eng **9**:177–192.

28. **McKinlay JB**, **Zeikus JG**, **Vieille C**. 2005. Insights into *Actinobacillus succinogenes* fermentative metabolism in a chemically defined growth medium. Appl Environ Microbiol **71**:6651-6656.

29. **McKinlay JB**, **Vieille C**. 2008. ^{13}C-metabolic flux analysis of *Actinobacillus succinogenes* fermentative metabolism at different $NaHCO_3$ and H_2 concentrations. Metab Eng **10**:55–68.

30. **McKinlay JB**, **Laivenieks M**, **Schindler BD**, **Mckinlay AA**, **Siddaramappa S**, **Challacombe JF**, **Lowry SR**, **Clum A**, **Lapidus AL**, **Burkhart KB**, **Harkins V**, **Vieille C**. 2010. A genomic perspective on the potential of *Actinobacillus succinogenes* for industrial succinate production. BMC Genomics **11**:680.

31. **Joshi RJ**, **Schindler BD**, **McPherson NR**, **Tiwari K**, **Vieille C**. 2014. Development of a markerless knockout method for *Actinobacillus succinogenes*. Appl Environ Microbiol **80**:3053-3061.

32. **Carvalho M**, **Matos M**, **Roca C**, **Reis MAM**. 2014. Succinic acid production from glycerol by *Actinobacillus succinogenes* using dimethylsulfoxide as electron acceptor. N Biotechnol **31**:133-139.

33. **Van Heerden CD**, **Nicol W**. 2013. Continuous succinic acid fermentation by *Actinobacillus succinogenes*. Biochem Engin J **73**:5-11.

34. **Bradfield MFA**, **Nicol W**. 2014. Continuous succinic acid production by *Actinobacillus succinogenes* in a biofilm reactor: steady-state metabolic flux variation. Biochem Engin J **85**:1-7.

35. **Guettler M**, **Hanchar R**, **Kleff S**, **Jadhav S**. 2015. Recombinant microorganisms for producing organic acids.

36. **Vlysidis A**, **Binns M**, **Webb C**, **Theodoropoulos C**. 2009. Utilisation of glycerol to platform chemicals within the biorefinery concept: a case for succinate production. Chem Eng Trans **18**:537-542.

37. **Vlysidis A**, **Binns M**, **Webb C**, **Theodoropoulos C**. 2011. Glycerol utilisation for the production of chemicals: conversion to succinic acid, a combined experimental and computational study. Biochem Eng J **58**:1-11.

38. **Pateraki C**, **Almqvist H**, **Ladakis D**, **Liden G**, **Koutinas AA**, **Vlysidis A**. 2016. Modelling succinic acid fermentation using a xylose based substrate. Biochem Eng J **114**:26-41.

39. **Salvachúa D**, **Mohagheghi A**, **Smith H**, **Bradfield MFA**, **Nicol W**, **Black BA**, **Biddy MJ**, **Dowe N**, **Beckham GT**. 2016. Succinic acid production on xylose-enriched biorefinery streams by *Actinobacillus succinogenes* in batch fermentation. Biotechnol Biofuels **9**:28.

40. **Guarnieri MT**, **Chou YC**, **Salvachua D**, **Mohagheghi A**, **St John PC**, **Peterson DJ**, **Bomble YJ**, **Beckham GT**. 2017. Metabolic Engineering of *Actinobacillus succinogenes* Provides Insights into Succinic Acid Biosynthesis. Appl Environ Microbiol 83.

41. **Zheng P**, **Dong J-J**, **Sun Z-H**, **Ni Y**, **Fang L**. 2009. Fermentative production of succinic acid from straw hydrolysate by *Actinobacillus succinogenes*. Bioresour Technol **100**:2425-2429.

42. **Li F**, **Wang Y**, **Gong K**, **Wang Q**, **Liang Q**, **Qi Q**. 2014. Constitutive expression of RyhB regulates the heme biosynthesis pathway and increases the 5-aminolevulinic acid accumulation in *Escherichia coli*. FEMS Microbiol Lett **350**:209-215.

43. **Zheng P**, **Fang L**, **Xu Y**, **Dong J-J**, **Ni Y**, **Sun Z-H**. 2010. Succinic acid production from corn stover by simultaneous saccharification and fermentation using *Actinobacillus succinogenes*. Bioresour Technol **101**:7889-7894.

44. **Wan C**, **Li Y**, **Shahbazi A**, **Xiu S**. 2008. Succinic acid production from cheese whey using *Actinobacillus succinogenes* 130 Z. Appl Biochem Biotech **145**:111-119.

45. **Lee SJ**, **Song H**, **Lee SY**. 2006. Genome-based metabolic engineering of *Mannheimia succiniciproducens* for succinic acid production. Appl Environ Microbiol **72**:1939–1948.

46. **Choi S**, **Song H**, **Lim SW**, **Kim TY**, **Ahn JH**, **Lee JW**, **Lee M-H**, **Lee SY**. 2016. Highly selective production of succinic acid by metabolically engineered *Mannheimia succiniciproducens* and its efficient purification. Biotechnol Bioeng **113**:2168-2177.

47. **Lee JW**, **Yi J**, **Kim TY**, **Choi S**, **Ahn JH**, **Song H**, **Lee M-H**, **Lee SY**. 2016. Homo-succinic acid production by metabolically engineered *Mannheimia succiniciproducens*. Metab Eng **38**:409-417.

48. **Pettigrew DW**, **Liu WZ**, **Holmes C**, **Meadow ND**, **Roseman S**. 1996. A single amino acid change in *Escherichia coli* glycerol kinase abolishes glucose control of glycerol utilization in vivo. J Bacteriol **178**:2846-2852.

49. **Lee KS**. 2004. Anaerobic hydrogen production with an efficient carrier-induced granular sludge bed bioreactor. Biotechnol Bioeng **87**:648-657.

50. **Scholten E**, **Dägele D**. 2008. Succinic acid production by a newly isolated bacterium. Biotechnol Lett **30**:2143-2146.

51. **Scholten E**, **Renz T**, **Thomas J**. 2009. Continuous cultivation approach for fermentative succinic acid production from crude glycerol by *Basfia succiniciproducens* DD1. Biotechnol Lett **31**:1947-1951.

52. **Becker J**, **Reinefeld J**, **Stellmacher R**, **Schäfer R**, **Lange A**, **Meyer H**, **Lalk M**, **Zelder O**, **von Abendroth G**, **Schröder H**. 2013. Systems-wide analysis and engineering of metabolic pathway fluxes in bio‐succinate producing *Basfia succiniciproducens*. Biotechnol Bioeng **110**:3013-3023.

53. **Chatterjee R**, **Millard CS**, **Champion K**, **Clark DP**, **Donnelly MI**. 2001. Mutation of the *ptsG* gene results in increased production of succinate in fermentation of glucose by *Escherichia coli*. Appl Environ Microbiol **67**:148–154.

54. **Thakker C**, **Martínez I**, **San K-Y**, **Bennett GN**. 2012. Succinate production in *Escherichia coli*. Biotechnol J **7**:213-224.

55. **Sawers RG**, **Clark DP**. 2004. Fermentative pyruvate and acetyl-coenzyme A metabolism. *In* Bock A (ed), *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. ASM Press.

56. **Millard CS**, **Chao YP**, **Liao JC**, **Donnelly MI**. 1996. Enhanced production of succinic acid by overexpression of phosphoenolpyruvate carboxylase in *Escherichia coli*. Appl Environ Microbiol **62**:1808–1810.

57. **Gokarn RR**, **Eiteman MA**, **Altman E**. 2000. Metabolic analysis of *Escherichia coli* in the presence and absence of the carboxylating enzymes phosphoenolpyruvate carboxylase and pyruvate carboxylase. Appl Environ Microbiol **66**:1844–1850.

58. **Chatterjee T**, **Datta AG**. 1973. Anaerobic formation of succinate from glucose and bicarbonate in resting cells of *Leishmania donovani*. Exp Parasitol **33**:138-146.

59. **Vemuri GN**, **Eiteman MA**, **Altman E**. 2002. Succinate production in dual-phase *Escherichia coli* fermentations depends on the time of transition from aerobic to anaerobic conditions. J Ind Microbiol Biotechnol **28**:325-332.

60. **Kim P**, **Laivenieks M**, **McKinlay J**, **Vieille C**, **Zeikus JG**. 2004. Construction of a shuttle vector for the overexpression of recombinant proteins in *Actinobacillus succinogenes*. Plasmid **51**:108–115.

61. **Jantama K**, **Haupt MJ**, **Svoronos SA**, **Zhang XL**, **Moore JC**, **Shanmugam KT**, **Ingram LO**. 2008. Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of *Escherichia coli* C that produce succinate and malate. Biotechnol Bioeng **99**:1140-1153.

62. **Jantama K**, **Zhang XL**, **Moore JC**, **Shanmugam KT**, **Svoronos SA**, **Ingram LO**. 2008. Eliminating side products and increasing succinate yields in engineered strains of *Escherichia coli* C. Biotechnol Bioeng **101**:881-893.

63. **Zhang XL**, **Jantama K**, **Moore JC**, **Jarboe LR**, **Shanmugam KT**, **Ingram LO**. 2009. Metabolic evolution of energy-conserving pathways for succinate production in *Escherichia coli*. Proc Natl Acad Sci USA **106**:20180-20185.

64. **Zhang XL**, **Jantama K**, **Shanmugam KT**, **Ingram LO**. 2009. Reengineering *Escherichia coli* for succinate production in mineral salts medium. Appl Environ Microbiol **75**:7807-7813.

65. **Zhang X**, **Shanmugam KT**, **Ingram LO**. 2010. Fermentation of glycerol to succinate by metabolically engineered strains of *Escherichia coli*. Appl Environ Microbiol **76**:2397-2401.

66. **Inui M**, **Murakami S**, **Okino S**, **Kawaguchi H**, **Vertès AA**, **Yukawa H**. 2004. Metabolic analysis of *Corynebacterium glutamicum* during lactate and succinate productions under oxygen deprivation conditions. J Mol Microbiol Biotechnol Bioeng **7**:182-196.

67. **Okino S**, **Inui M**, **Yukawa H**. 2005. Production of organic acids by *Corynebacterium glutamicum* under oxygen deprivation. Appl Microbiol Biotechnol **68**:475-480.

68. **Okino S**, **Noburyu R**, **Suda M**, **Jojima T**, **Inui M**, **Yukawa H**. 2008. An efficient succinic acid production process in a metabolically engineered *Corynebacterium glutamicum* strain. Appl Microbiol Biotechnol **81**:459-464.

69. **Litsanov B**, **Brocker M**, **Bott M**. 2012. Toward homosuccinate fermentation: metabolic engineering of *Corynebacterium glutamicum* for anaerobic production of succinate from glucose and formate. Appl Environ Microbiol **78**:3325-3337.

70. **Zhu N**, **Xia H**, **Yang J**, **Zhao X**, **Chen T**. 2014. Improved succinate production in *Corynebacterium glutamicum* by engineering glyoxylate pathway and succinate export system. Biotechnol Lett **36**:553-560.

71. **Litsanov B**, **Kabus A**, **Brocker M**, **Bott M**. 2012. Efficient aerobic succinate production from glucose in minimal medium with *Corynebacterium glutamicum*. Microb Biotechnol **5**:116-128.

72. **Litsanov B**, **Brocker M**, **Bott M**. 2013. Glycerol as a substrate for aerobic succinate production in minimal medium with *Corynebacterium glutamicum*. Microb Biotechnol **6**:189-195.

73. **Zhu N**, **Xia H**, **Wang Z**, **Zhao X**, **Chen T**. 2013. Engineering of acetate recycling and citrate synthase to improve aerobic succinate production in *Corynebacterium glutamicum*. PLoS One **8**:e60659.

74. **Chung SC**, **Park JS**, **Yun J**, **Park JH**. 2017. Improvement of succinate production by release of end-product inhibition in *Corynebacterium glutamicum*. Metab Eng **40**:157-164.

75. **Arikawa Y**, **Kobayashi M**, **Kodaira R**, **Shimosaka M**, **Muratsubaki H**, **Enomoto K**, **Okazaki M**. 1999. Isolation of sake yeast strains possessing various levels of succinate- and/or malate-producing abilities by gene disruption or mutation. J Biosci Bioeng **87**:333-339.

76. **Kubo Y**, **Takagi H**, **Nakamori S**. 2000. Effect of gene disruption of succinate dehydrogenase on succinate production in a sake yeast strain. J Biosci Bioeng **90**:619-624.

77. **Raab AM**, **Gebhardt G**, **Bolotina N**, **Weuster-Botz D**, **Lang C**. 2010. Metabolic engineering of *Saccharomyces cerevisiae* for the biotechnological production of succinic acid. Metab Eng **12**:518-525.

78. **Ito Y**, **Hirasawa T**, **Shimizu H**. 2014. Metabolic engineering of *Saccharomyces cerevisiae* to improve succinic acid production based on metabolic profiling. Biosci Biotechnol Biochem **78**:151-159.

79. **Yan D**, **Wang C**, **Zhou J**, **Liu Y**, **Yang M**, **Xing J**. 2014. Construction of reductive pathway in *Saccharomyces cerevisiae* for effective succinic acid fermentation at low pH value. Bioresour Technol **156**:232-239.

80. **Scholten E**, **Haefner S**, **Schröder H**. 2014. Bacterial cells having a glyoxylate shunt for the manufacture of succinic acid. US Patent 8,877,466 B2.

81. **Schroder H**, **Haefner S**, **Von Abendroth G**, **Hollmann R**, **Raddatz A**, **Ernst H**, **Gurski H**. 2014. Microbial succinic acid producers and purification of succinic acid.

82. **Janausch IG**, **Zientz E**, **Tran QH**, **Kroger A**, **Unden G**. 2002. C4-dicarboxylate carriers and sensors in bacteria. Biochim Biophys Acta **1553**:39-56.

83. **Zientz E**, **Six S**, **Unden G**. 1996. Identification of a third secondary carrier (DcuC) for anaerobic C4-dicarboxylate transport in *Escherichia coli*: roles of the three Dcu carriers in uptake and exchange. J Bacteriol **178**:7241-7247.

84. **Karinou E**, **Compton EL**, **Morel M**, **Javelle A**. 2013. The *Escherichia coli* SLC26 homologue YchM (DauA) is a C_4-dicarboxylic acid transporter. Mol Microbiol **87**:623-640.

85. **Six S**, **Andrews SC**, **Unden G**, **Guest JR**. **1994**. *Escherichia coli* possesses two homologous anaerobic C4-dicarboxylate membrane transporters (DcuA and DcuB) distinct from the aerobic dicarboxylate transport system (Dct). J Bacteriol **176**:6470-6478.

86. **Davies SJ**, **Golby P**, **Omrani D**, **Broad SA**, **Harrington VL**, **Guest JR**, **Kelly DJ**, **Andrews SC**. 1999. Inactivation and regulation of the aerobic C_4-dicarboxylate transport (dctA) gene of *Escherichia coli*. J Bacteriol **181**:5624-5635.

87. **Youn JW**, **Jolkver E**, **Kramer R**, **Marin K**, **Wendisch VF**. 2009. Characterization of the dicarboxylate transporter DctA in *Corynebacterium glutamicum*. J Bacteriol **191**:5480-5488.

88. **Fukui K**, **Nanatani K**, **Hara Y**, **Yamakami S**, **Yahagi D**, **Chinen A**, **Tokura M**, **Abe K**. 2017. Escherichia coli *yjjPB* genes encode a succinate transporter important for succinate production. Biosci Biotechnol Biochem **81**:1837-1844.

89. **Chen J**, **Zhu X**, **Tan Z**, **Xu H**, **Tang J**, **Xiao D**, **Zhang X**. 2014. Activating C4-dicarboxylate transporters DcuB and DcuC for improving succinate production. Appl Microbiol Biotechnol **98**:2197-2205.

90. **Kelly DJ**, **Thomas GH**. 2001. The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. FEMS Microbiol Rev **25**:405-424.

91. **Mulligan C**, **Fischer M**, **Thomas GH**. 2011. Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. FEMS Microbiol Rev **35**:68-86.

92. **Pos KM**, **Dimroth P**, **Bott M**. 1998. The *Escherichia coli* citrate carrier CitT: a member of a novel eubacterial transporter family related to the 2-oxoglutarate/malate translocator from spinach chloroplasts. J Bacteriol **180**:4160-4165.

93. **Huhn S**, **Jolkver E**, **Kramer R**, **Marin K**. 2011. Identification of the membrane protein SucE and its role in succinate transport in *Corynebacterium glutamicum*. Appl Microbiol Biotechnol **89**:327-335.

94. **Fukui K**, **Koseki C**, **Yamamoto Y**, **Nakamura J**, **Sasahara A**, **Yuji R**, **Hashiguchi K**, **Usuda Y**, **Matsui K**, **Kojima H**, **Abe K**. 2011. Identification of succinate exporter in *Corynebacterium glutamicum* and its physiological roles under anaerobic conditions. J Biotechnol **154**:25-34.

95. **Rhie MN**, **Yoon HE**, **Oh HY**, **Zedler S**, **Unden G**, **Kim OB**. 2014. A Na+-coupled C4-dicarboxylate transporter (Asuc_0304) and aerobic growth of *Actinobacillus succinogenes* on C4-dicarboxylates. Microbiology **160**:1533-1544.

96. **Na D**, **Yoo SM**, **Chung H**, **Park H**, **Park JH**, **Lee SY**. 2013. Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. Nat Biotechnol **31**:170-174.

97. **Kim JYH**, **Cha HJ**. 2003. Down-regulation of acetate pathway through antisense strategy in *Escherichia coli*: improved foreign protein production. Biotechnol Bioeng **83**:841-853.

98. **Nakashima N**, **Tamura T**, **Good L**. 2006. Paired termini stabilize antisense RNAs and enhance conditional gene silencing in *Escherichia coli*. Nucleic Acids Res **34**:e138.

99. **Negrete A**, **Majdalani N**, **Phue J-N**, **Shiloach J**. 2013. Reducing acetate excretion from *E. coli* K-12 by over-expressing the small RNA SgrS. N Biotechnol **30**:269-273.

100. **Coleman J**, **Green PJ**, **Inouye M**. 1984. The use of RNAs complementary to specific mRNAs to regulate the expression of individual bacterial genes. Cell **37**:429-436.

101. **Kang Z**, **Wang X**, **Li Y**, **Wang Q**, **Qi Q**. 2012. Small RNA RyhB as a potential tool used for metabolic engineering in *Escherichia coli*. Biotechnol Lett **34**:527-531.

102. **Ding W**, **Weng H**, **Du G**, **Chen J**, **Kang Z**. 2017. 5-Aminolevulinic acid production from inexpensive glucose by engineering the C4 pathway in *Escherichia coli*. J Ind Microbiol Biotechnol **44**:1127-1135.

103. **Desai RP**, **Papoutsakis ET**. 1999. Antisense RNA strategies for metabolic engineering of *Clostridium acetobutylicum*. Appl Environ Microbiol **65**:936-945.

104. **Boudry P**, **Gracia C**, **Monot M**, **Caillet J**, **Saujet L**, **Hajnsdorf E**, **Dupuy B**, **Martin-Verstraete I**, **Soutourina O**. 2014. Pleiotropic role of the RNA chaperone protein Hfq in the human pathogen *Clostridium difficile*. J Bacteriol **196**:3234-3248.

105. **Liu Y**, **Zhu Y**, **Li J**, **Shin H-D**, **Chen RR**, **Du G**, **Liu L**, **Chen J**. 2014. Modular pathway engineering of *Bacillus subtilis* for improved N-acetylglucosamine production. Metab Eng **23**:42-52.

106. **Cho C**, **Lee SY**. 2016. Efficient gene knockdown in *Clostridium acetobutylicum* by synthetic small regulatory RNAs. Biotechnol Bioeng **114**:374-383.

107. **Subashchandrabose S, Leveque RM, Kirkwood RN, Kiupel M, Mulks MH**. 2013. The RNA chaperone Hfq promotes fitness of *Actinobacillus pleuropneumoniae* during porcine pleuropneumonia. Infect Immun **81**:2952-2961.

108. **Amarasinghe JJ**, **Connell TD**, **Scannapieco FA**, **Haase EM**. 2012. Novel iron‐regulated and Fur‐regulated small regulatory RNAs in *Aggregatibacter actinomycetemcomitans*. Mol Oral Microbiol **27**:327-349.

109. **Baddal B**, **Muzzi A**, **Censini S**, **Calogero RA**, **Torricelli G**, **Guidotti S**, **Taddei AR**, **Covacci A**, **Pizza M**, **Rappuoli R**. 2015. Dual RNA-seq of nontypeable *Haemophilus influenzae* and host cell transcriptomes reveals novel insights into host-pathogen cross talk. mBio **6**:e01765-15.

110. **Santana EA**, **Harrison A**, **Zhang X**, **Baker BD**, **Kelly BJ**, **White P**, **Liu Y**, **Munson Jr RS**. 2014. HrrF is the Fur-regulated small RNA in nontypeable *Haemophilus influenzae*. PloS one **9**:e105644.

111. **Rossi CC**, **Bossé JT**, **Li Y**, **Witney AA**, **Gould KA**, **Langford PR**, **Bazzolli DMS**. 2016. A computational strategy for the search of regulatory small RNAs in *Actinobacillus pleuropneumoniae*. RNA **22**:1373-1385.

Chapter 2

Identification of succinate transporters in *Actinobacillus succinogenes*

2.1 Abstract

Succinate, listed as one of the top 12 building block chemicals by the U.S. Department of Energy has been produced from maleic anhydride, a petrochemical until recently. A few companies, including BioAmber, Myriant, and Reverdia, have started producing industrial scale amounts of bio-based succinate using yeast and *E. coli* strains. *A. succinogenes*, along with other succinate producing bacteria have been studied and engineered with the aim of increasing succinate production. Succinate exporters have been studied in very few of these bacteria, especially with the aim of using them as targets in engineering for increased succinate production. *E. coli* and *C. glutamicum* are the only strains where succinate transporters have been studied. Several succinate exporter candidates have been overexpressed in *E. coli* and in *C. glutamicum* and caused an increase in succinate efflux. So far, *A. succinogenes* has not been studied for succinate exporters.

In this study, we sought to identify succinate transporters using a combined proteomics and transcriptomics approach. The proteins with the most hits in our proteomics studies—Asuc_1999, Asuc_0142, Asuc_2058 and Asuc_1990-91 were selected for overexpression in the wild-type strain. Since we are limited by our only expression vector which harbors the native, strong, and constitutive promoter p_{pckA}, we developed a library of expression vectors with promoters weaker than p_{pckA} by truncating p_{pckA} and searching for additional promoter candidates from our RNA sequencing analysis. We looked for candidate genes that were expressed at a similar level across growth conditions tested and that had several fold lower transcript levels than *pckA*. Finally, we overexpressed our candidate succinate exporters under the control of promoter's $p_{pckA\text{-}92}$ and $p_{pckA\text{-}103}$, and observed an increase in succinate production in all the strains tested as compared to the control strain. The highest increase was seen when Asuc_1999

was expressed under the control of $p_{pckA\text{-}103}$. To the best of our knowledge this is the first study focusing on identifying succinate exporters in *A. succinogenes*. This is also the first report of a promoter library being constructed for *A. succinogenes*, further increasing the tunability of gene expression.

2.2 Introduction

Succinate is a dicarboxylate molecule with many possible applications. It can be used as a precursor to make products such as nylons, resins, paints, cosmetics, flavors, and adhesives (1). The conventional route to produce succinate has been by hydrogenation of maleic anhydride, a petroleum derivative, in the presence of a metal catalyst (2). Bio-based succinate production has gained momentum in recent years, though, as awareness about the benefits of renewable products and downsides of burning fossil fuels have increased. A number of companies, such as BioAmber, Myriant, and Succinity, have started selling bio-based succinate produced by engineered microorganisms (e.g., *Escherichia coli*, *Basfia succiniciproducens*, and yeast). Bio-based succinate is produced naturally by anaerobes such as *Actinobacillus succinogenes* and *Mannheimia succiniproducens* via the reductive pathway of the tricarboxylic acid cycle. Industrial workhorse microorganisms such as *E. coli*, *Saccharomyces cerevisiae*, and *Corynebacterium glutamicum* also have been or are being engineered to produce succinate (2-5). One of the major hurdles still lies in making bio-based succinate as cost–effective as conventionally produced succinate.

A. succinogenes is one of the best natural succinate producers, yet our knowledge of its uptake and efflux mechanisms remains limited. For example, the system or systems involved in glucose uptake as well as the one or ones involved in succinate export remain unknown, and the succinate export system has not yet been targeted for increasing succinate production.

Several succinate transporters have been identified and characterized in *E. coli,* but most studies have focused on uptake and antiport transporters. The dicarboxylate uptake (Dcu) carriers DcuA, B, and C are the best-characterized succinate transporters. DcuA and DcuB are involved in fumarate: succinate antiport during anaerobic respiration on fumarate. They are also capable of uptake and efflux of C_4-dicarboxylates (6, 7). Zeintz et al. (8) suggested that, even though DcuC can carry out the same transport functions as DcuA and DcuB, its major function likely is succinate export. *E. coli dcuD* shows similarity to *dcuC,* but DcuD mutants show no effect on *E. coli* growth under a number of conditions, and do not affect succinate uptake or efflux (9). DauA is the sole transporter for aerobic succinate uptake at acidic pH (10). Another transporter, CitT, catalyzes citrate/succinate antiport during citrate fermentation (11). DctA, a H^+/C_4-dicarboxylate symporter is responsible for C_4-dicarboxylate uptake in aerobic conditions (6, 12) at neutral pHs. In one report, a quintuple *dctA dcuA dcuB dcuC citT E. coli* mutant did not show any deficiency in succinate efflux as compared to the wild-type strain in glucose-grown anaerobic cultures (13). In another report, while deleting *dcuA* and *dcuD* did not have any effect on succinate efflux, individual *dcuB* and *dcuC* deletion mutants showed decreases in succinate titers of 15% and 11%, respectively (14). A double *dcuBC* deletion resulted in a 90% decrease in succinate titer, and modulating expression of both genes increased the succinate titer by 34% (14). Thus no clear picture of the involvement of Dcu transporters in succinate efflux is available. DauA is the sole characterized *E. coli* succinate transporter that has not been tested for succinate efflux during glucose fermentation.

Two studies using different approaches identified the same succinate exporter, SucE, of the aspartate:alanine exchanger family, in *Corynebacterium glutamicum* (15, 16). This transporter functioned in anaerobic and microaerobic conditions and was responsible for about

30% of the succinate export in microaerobic conditions, indicating that SucE is not the sole succinate exporter in *C. glutamicum*. In a recent study aimed at identifying the *E. coli* anaerobic succinate exporter, *C. glutamicum* strain AJ110655 Δ*sucE*, deficient in succinate export, was partially rescued for succinate export by expressing the *E. coli yjjPB* genes (17). Additionally, an *E. coli yjjPB* knockout mutant produced 70% less succinate than the parental strain, indicating that YjjPB may be one of the main *E. coli* anaerobic succinate exporters. There has been no report on whether or not the sole *E. coli* SucE homolog participates in succinate export.

In this study, we conducted transcriptomic studies of *A. succinogenes* grown on multiple carbon sources as well as proteomics studies on cytoplasmic membrane-enriched fractions of *A. succinogenes* grown anaerobically on glucose, xylose, and fructose (conditions favoring succinate production) or on glycerol-nitrate (less succinate is produced). Of the putative C_4 dicarboxylate transporters identified in the annotated genome, locus tags Asuc_1999, Asuc_1990-1991, Asuc_0142, and Asuc_2058 had the most hits in our proteomics and transcriptomics results, and we focused on these four transporters for further analysis. Currently, to overexpress *A. succinogenes* genes we only have one strong constitutive native promoter (p_{pckA}, (18) at our disposal, which is not optimal for the expression of membrane proteins. We constructed a library of promoters of varying strengths to overcome this problem and modulate the expression of these transporters. We show that overexpression of any of these four putative C4-dicarboxylate transporters significantly increases succinate production.

2.3 Materials and Methods

2.3.1 Strains, media, and culture conditions.

E. coli strains (Table 2.1) were cultivated in lysogeny broth (LB) and on LB agar plates and were supplemented with 100 µg mL^{-1} ampicillin or 50 µg mL^{-1} kanamycin when required for plasmid maintenance (19). *A. succinogenes* type strain 130Z was obtained from the American Type Culture Collection. *A. succinogenes* liquid cultures were grown in AM3 defined medium (20), Medium B (21), or Bacto brain heart infusion medium (BHI; Becton, Dickinson and Co., Franklin Lakes, NJ). Mutant strains of *A. succinogenes* were grown on AM2-isocitrate or AM16-isocitrate after natural transformations (21). Liquid cultures were grown anaerobically in 28-mL anaerobic tubes flushed with N_2 at 37°C with shaking at 250 rpm. Optical densities were measured at 660 nm using a Spectronic 20 (Bausch and Lomb, Rochester, NY) or a DU650 spectrophotometer (Beckman, Fullerton, CA). To isolate *A. succinogenes* colonies, strains were grown on LB plates containing 10 g L^{-1} glucose (22) and 10 µg mL^{-1} kanamycin (Km), 600 mg L^{-1} polyvinyl alcohol (PVA) and 10 mg L^{-1} $CaCl_2$. Agar plate cultures of *A. succinogenes* strains were grown under a CO_2-enriched atmosphere for 24-48 h. All strains used in this study are listed in Table 2.1.

2.3.2 Plasmids, DNA manipulations, and electroporations.

PCR products were cloned into pCR2.1 using the TA TOPO cloning kit (Invitrogen, Carlsbad, CA) and transformed into One Shot TOP10 chemically competent cells (Invitrogen) as per manufacturer's protocols. Shuttle plasmid pLGZ920 was used to express native or foreign genes in *A. succinogenes* (18). Electrocompetent *A. succinogenes* cells, prepared as described (21), were transformed with plasmid constructs and spread on LB Glucose PVA $CaCl_2$ Km plates.

DNA manipulations were carried out according to Ausubel et al. (19). PCR products were amplified using Phusion High-Fidelity DNA polymerase (New England Biolabs, Ipswich, MA). Restriction enzymes were from New England Biolabs. Plasmids and DNA fragments were purified using the Wizard SV miniprep kit and Wizard SV Gel and PCR Clean-Up system, respectively, (Promega, Madison, WI). Primers used in this work were synthesized by Integrated DNA Technologies (Coralville, IA) and are listed in Table S2.1 in supplementary material. DNA sequencing was performed by GENEWIZ, Inc. (South Plainfield, NJ). Colonies were screened by colony PCR using Taq polymerase.

Table 2.1 Strains and plasmids used in this study

	Description	**Source**
E. coli		
DH5α	F- φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17	Laboratory collection
TOP10	(rk-, mk+) *pho*A *sup*E44 λ- thi-1 *gyr*A96 *rel*A1 F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G λ-	Invitrogen
A. succinogenes		
130Z (ATCC55618)	Wild-type strain	ATCC
*Δlac*Z::*icd*	130Z derivative, contains the *AscI-FRT*-p_{pckA}-*icd-FRT-AscI* cassette in the *lacZ* deletion	This study
ΔlacZ	130Z derivative, *lacZ* deletion, contains one *FRT* site	This study
Plasmids		
pCR2.1 TOPO	Amp^R, Km^R, *lacZα*, cloning vector	Invitrogen
pCR2.1-Δ*lacZ*	pCR2.1 derivative, *lacZ* deletion (2.1-kb fusion product of frd_{up} and *frdCD*)	(21)
pCR2.1-*icd*	pCR2.1 derivative, *E. coli icd* under control of *A. succinogenes* p_{pckA}, flanked by *FRT* repeats and *Asc*I restriction sites	(21)
pCV933	pLGZ920 derivative, *S. cerevisiae flp* under control of p_{pckA}	(21)
pLGZ920-*lacZ*	pLGZ920 derivative containing *lacZ* under p_{pckA}	This study
pLGZ920-$p_{pckA\text{-}92}$-*lacZ*	pLGZ920 derivative containing *lacZ* under $p_{pckA\text{-}92}$	This study

Table 2.1 (cont'd)

pLGZ920-$p_{pckA-103}$-*lacZ*	pLGZ920 derivative containing *lacZ* under $p_{pckA-103}$	This study
pLGZ920-$p_{pckA-134}$-*lacZ*	pLGZ920 derivative containing *lacZ* under $p_{pckA-134}$	This study
pLGZ920-$p_{pckA-164}$-*lacZ*	pLGZ920 derivative containing *lacZ* under$p_{pckA-164}$	This study
pLGZ920*-$_{pAsuc\_0701}$	pLGZ920 derivative with the *Hin*dIII site 3' of ColE1 ori replaced with *Bam*HI and $p_{pckA-0701}$	This study
pLGZ920*-$_{pAsuc\_0701}$-*lacZ*	pLGZ920* derivative containing *lacZ* under p_{Asuc_0701}	This study
pLGZ920*-p_{Asuc_2109}-*lacZ*	pLGZ920* derivative containing *lacZ* under p_{Asuc_2109}	This study
pLGZ920*-p_{Asuc_0391}-*lacZ*	pLGZ920* derivative containing *lacZ* under p_{Asuc_0391}	This study
pLGZ920*-p_{Asuc_0289}-*lacZ*	pLGZ920* derivative containing *lacZ* under p_{Asuc_0289}	This study

2.3.3 Preparation of *A. succinogenes* cytoplasmic membrane-enriched fractions for proteomics analysis.

A. succinogenes cultures in AM3 glucose, fructose, and xylose were grown in 300-mL volumes and harvested in the middle of the exponential phase at an OD_{660} of 0.8 to 0.9. Because *A. succinogenes* does not grow fermentatively on glycerol, glycerol cultures were grown by nitrate respiration. The glycerol-nitrate cultures were grown in 600-mL volumes and harvested at an OD_{660} no higher than 0.4 to 0.5 to avoid succinate accumulation after nitrate depletion (Schindler et al, 2014). Cytoplasmic membrane-enriched fractions were prepared using a method adapted from (23). Cells were harvested by centrifugation at 12,000 × *g* for 10 min and washed twice with 100 mM Tris-HCl (pH 8.0). Bacterial pellets were resuspended in 15 mL of buffer A (100 mM Tris-HCl [pH 8.0] containing 20% sucrose and Complete Mini, EDTA-free, protease inhibitor cocktail [Sigma-Aldrich]). ReadyLyse lysozyme (EpiCentre, Madison, WI) was added

at a final concentration of 10^4 U mL^{-1}, and bacterial suspensions were stirred on a magnetic plate at the lowest setting for 8 min. Ethylenediaminetetraacetic acid (EDTA, 1.5 mL of 100 mM, pH 8.0) (1:10 v/v EDTA:cells) was added drop by drop over 2.5 min to avoid lysis, and suspensions were stirred for 8 more min. The outer membrane and periplasmic fractions were separated from the spheroplasts by centrifugation at 12,000 × *g* for 10 min.

The pellets, corresponding to the spheroplast-enriched fractions, were washed once with 15 mL buffer A, and resuspend in 15 mL buffer A containing 225,000 U DNase I and 0.2 mg RNase A. The spheroplasts were lysed by passing twice though a French press at 1,400 psi (high setting). The intact spheroplasts were removed by centrifugation at 12,000 × *g* for 10 min. The supernatants were then ultracentrifuged at 150,000 × *g* for 2 h at 4°C. The pellets were resuspended in buffer A and ultracentrifuged again to wash off the cytoplasmic proteins. The cytoplasmic membrane-enriched fractions were finally resuspended in 100 mM Tris-HCl (pH 8.0) containing 1 mM EDTA and protease inhibitors to a concentration of 4 mg mL^{-1}.

2.3.4 Proteomic analysis.

To submit the cytoplasmic membrane-enriched fractions for proteomic analysis, 50 µl of extracts (i.e., 200 µg protein) were loaded on sodium dodecyl sulfate-polyacrylamide gels (mini-PROTEAN TGX gel, 10% acrylamide, Bio-Rad, Hercules, CA) and subjected to electrophoresis until the samples had entered the stacking gel. The proteins were visualized by Coomassie blue staining. Gel bands were digested in-gel according to Shevchenko et al.[1] with modifications. Briefly, gel bands were dehydrated using 100% acetonitrile and incubated with 10 mM dithiothreitol in 100 mM ammonium bicarbonate (pH 8.0) at 56 °C for 45 min, dehydrated again and incubated in the dark with 50 mM iodoacetamide in 100 mM ammonium bicarbonate for 20

min. Gel bands were then washed with 100 mM ammonium bicarbonate and dehydrated again. Sequencing-grade modified trypsin (50 µL at 0.01 µg μL^{-1} in 50 mM ammonium bicarbonate) was added to each gel band to completely submerge the band. Bands were then incubated at 37 °C overnight. Peptides were extracted from the gel by water bath sonication in a solution of 60% acetonitrile/1% trichloroacetic acid and vacuum-dried to ~2µL.

Peptides were then re-suspended in 2% acetonitrile/0.1% trifluoroacetic acid to 25 µL. Five µL were automatically injected by a Thermo (www.thermo.com) EASYnLC 1000 onto a Thermo Acclaim PepMap RSLC 0.075 mm x 250 mm C18 column and eluted over 185 min with a 174-min gradient of 5% B to 25% B. The gradient was then raised to 100% B in 1 min and held at 100% B for the duration of the run (Buffer A = 99.9% water/0.1% formic acid, Buffer B = 99.9% acetonitrile/0.1% formic Acid).

Eluted peptides were sprayed into a Q-Exactive mass spectrometer (Thermo Fisher Scientific, Waltham, MA) using a FlexSpray spray ion source. Survey scans were taken in the Orbi trap (35,000 resolution, determined at m/z 200) and the top ten ions in each survey scan were subjected to automatic higher energy collision induced dissociation with fragment spectra acquired at 17,500 resolution.

The resulting MS/MS spectra were converted to peak lists using Mascot Distiller, v2.5.1 (www.matrixscience.com) and searched against a protein sequence database containing entries for *A. succinogenes* (downloaded on May 22, 2015 from NCBI, www.ncbi.nlm.nih.gov) and appended with common laboratory contaminants (downloaded from www.thegpm.org, cRAP project) using the Mascot searching algorithm, v 2.5. The following parameters were used: allow up to two missed tryptic sites, fixed modification of carbamidomethyl cysteines, variable modification of methionine oxidation and of asparagine and glutamine deamidation; +/- 5 ppm

peptide tolerance, 0.3 Da MS/MS tolerance, and false discovery rate (FDR) calculated using randomized database search. The Mascot output was then analyzed using Scaffold, v4.4.3 (www.proteomesoftware.com) to probabilistically validate protein identifications. Assignments validated using the Scaffold 1% FDR confidence filter were considered true. NSAF values (24) were calculated within Scaffold for protein quantitation.

2.3.5 Total RNA purification and RNA sequencing

Total RNA was purified from *A. succinogenes* cultures and quality-controlled as described (25). In short, cultures harvested in the middle of the exponential phase were immediately mixed with one volume of -20 °C methanol and stored at -20 °C for at least 30 min to quench all enzymatic activity. After centrifugation, RNA was purified from the cell pellets with the QIAGEN RNeasy Midi Kit (QIAGEN, Hilden, Germany), including the optional on-column DNase digestion. RNA samples were quantified on a Qubit Fluorometer (Life Technologies, Grand Island, NY), and RNA quality was validated visually on a 1% agarose gel. RNA was mixed 50% v/v with formaldehyde before loading on the gel.

Total RNA sequencing was performed by the Michigan State University Research Technology Support Facility as described (25). In short, libraries were prepared using the Illumina TruSeq Stranded Total RNA Library Preparation kit and rRNAs were depleted with the Illumina Ribo-Zero Gram-Negative Bacteria kit. Libraries were validated and quantified using the Qubit dsDNA assay, Caliper LabChipGX, and Kapa Illumina Library Quantification qPCR kits. The sequenced libraries (three biological replicates each) were from *A. succinogenes* 130Z fermenting glucose, xylose, and fructose, as well as *A. succinogenes* respiring glucose with nitrate as the electron acceptor, glycerol with nitrate as the electron acceptor, and glycerol in

microaerobic conditions. Pools of 12 libraries were loaded on separate lanes of an Illumina HiSeq 2500 High Output (v4) flow cell and sequenced in 1x50bp single end format using HiSeq SBS reagents. Base calling was done by Illumina Real Time Analysis v1.18.64. Output was demultiplexed and converted to FastQ format using Illumina Bcl2fastq v1.8.4. Sequencing results were analyzed using the SPARTA pipeline (26).

2.3.6 Identification of *A. succinogenes's pckA* transcription start site.

The transcriptional start site for *pckA* was identified with First Choice RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE) kit (Ambion, Inc.) following the manufacturer's instructions. DNase-treated total RNA (10 µg) extracted from microaerobic glycerol cultures (27) was treated with calf intestine phosphatase. The RNA was phenol-chloroform extracted and then treated with tobacco acid pyrophosphatase (TAP) to remove the cap from full-length mRNAs. TAP-treated RNA was then used for 5' RACE adapter ligation. The sample was reverse-transcribed and used as the template for nested PCR. Outer 5'RLM-RACE PCR was carried out using the 5' RACE outer primer and *pckA*-specific outer primer. The outer PCR product was used as the template for carrying out the inner 5'RLM-RACE PCR with the 5'RACE inner primer and *pckA*-specific inner primer. A-overhangs were added to the inner PCR product before cloning into pCR2.1. Colonies were screened by PCR using the inner 5'RLM-RACE PCR primers, and insert sequences were verified by sequencing.

2.3.7 Construction of truncated *pckA* promoters.

Plasmid pLGZ920-*lacZ* was constructed to use *lacZ* as the reporter gene to test the strength of truncated *pckA* promoters. Primers CV624 and CV625 were used to amplify *lacZ*

from *A. succinogenes* 130Z genomic DNA. The PCR product was cloned into pLGZ920 under control of the p_{pckA} promoter by Gibson Assembly (New England Biolabs, Ipswich, MA). The cloning reaction was transformed into cells provided with the kit. Colonies were screened by PCR with primers GZ1302 and GZ1303, and the insert sequence was verified by sequencing. Expression of *lacZ* from this plasmid construct was verified by β-galactosidase assay of recombinant strain Δ*lacZ* (pLGZ920-*lacZ*) (see below).

The ColE1 origin of replication (ori) was amplified from pLGZ920 using primers CV469 and CV471. Primer pairs CV473-CV443, CV474-CV443, CV475-CV443, and CV476-CV443 were used to amplify truncated *pckA* promoters $p_{pckA\text{-}92}$, $p_{pckA\text{-}103}$, $p_{pckA\text{-}134}$, and $p_{pckA\text{-}164}$ with pLGZ920-*lacZ* as the template. The ColE1 ori and truncated *pckA* promoters were fused by PCR using primers CV469 and CV470. The four fusion PCR products were gel purified and cloned into the pLGZ920-*lacZ* *Hin*dIII and *Xba*I sites by Gibson assembly (New England Biolabs, Ipswich, MA) as described (Gibson et al., 2009). Note that restriction by *Hin*dIII removes the ColE1 ori initially present in pLGZ920. The cloning reactions were transformed into TOP10 cells. Colonies were screened by PCR using primers CV565 and CV568. Positive plasmids were verified by sequencing. The constructs were then transformed into strain Δ*lacZ*.

2.3.8 Construction of an *A. succinogenes* Δ*lacZ* mutant strain.

The *A. succinogenes* Δ*lacZ* mutant strain was constructed as described (21). Briefly, the Δ*lacZ*::*icd* strain was built by natural transformation of *A. succinogenes* 130Z with a Δ*lacZ*::*icd* linear DNA fragment. Double recombinants were screened by colony PCR using primers CV376 and CV377. The *icd* marker was excised by the yeast flippase recombinase carried by plasmid pCV933. Transformants were screened for the loss of *icd* by colony PCR using primers CV376

and CV377. Plasmid pCV933 was cured from the Δ*lacZ*(pCV933) strain using acridine orange as described (21). Strain Δ*lacZ* was used as the host strain to test the promoter strength of the reporter promoter library constructs.

2.3.9 Preparation of crude extracts, β-galactosidase assays, and protein assays.

Ten-mL cultures were grown in AM3 supplemented with 50 mM glucose and 150 mM $NaHCO_3$. Cultures were harvested in exponential phase by centrifugation, washed, and resuspended in 0.5 mL of 50 mM sodium phosphate buffer (pH 7.0). Cells were lysed by sonication (Branson S-450A probe sonifier, Danbury, CT) with six 20-sec repetitions, 50% duty cycle, and a power level of 3. Cell extracts were centrifuged for 5 min (2,000 × *g*) at room temperature and the supernatant was stored on ice until used.

Beta-galactosidase assays were conducted in 96-well plates in a PowerWave HT microplate reader at room temperature (BioTek Instruments Inc., Winooski, VT). Reactions (200 µl) contained 20 µl cell extract mixed with 150 µl of Z-buffer (60 mM Na_2HPO_4 , 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 38 mM 2-mercaptoethanol) and 30 µL of 2-nitrophenyl β-D-galactopyranoside (4 mg mL^{-1}) in Z-buffer. Enzyme activity was calculated using the linear slope obtained from 2-nitrophenol production recorded at 420 nm. The extinction coefficient of 2-nitrophenol was 4.8 mM^{-1} cm^{-1}. Total cell protein was quantified using the Bio-Rad protein assay dye reagent using bovine serum albumin as the standard. Activities were reported as an average of three biological replicates.

2.3.10 Construction of a promoter library based on RNAseq results.

An initial promoter-*lacZ* reporter plasmid was assembled in pLGZ920 using the promoter

of Asuc_0701 (p_{Asuc_0701}). Other promoter-*lacZ* reporter plasmids were then constructed by substituting p_{Asuc_0701} with other promoters. First, ColE1 ori was amplified using primers CV615 and CV616 with pLGZ920 as the template. *A. succinogenes* 130Z genomic DNA was used as the template to amplify p_{Asuc_0701} with primers CV 617 and CV618. ColE1 ori and p_{pckA} were removed from pLGZ920 by digestion with *Hin*dIII and *Xba*I. ColE1 ori and p_{Asuc_0701} were then cloned into pLGZ920's *Hin*dIII and *Xba*I sites by Gibson Assembly to create plasmid pLGZ920*- p_{Asuc_0701}. In this plasmid, the *Hin*dIII site 3' of ColE1 ori is replaced with a *Bam*HI site to facilitate promoter replacement. Colonies were screened by PCR with primers CV615 and CV616 and positive plasmids were verified by sequencing. The *lacZ* gene was amplified from *A. succinogenes* 130Z genomic DNA using primers CV624 and CV625, and cloned into pLGZ920*-p_{Asuc_0701}'s *Xba*I and *Sac*I sites by Gibson cloning. Colonies were screened by PCR using primers CV636 and CV640. A plasmid with the correct *lacZ* sequence (sequencing with primers listed in Table S2.1) was called pLGZ920*-p_{Asuc_0701}-*lacZ*.

Promoters p_{Asuc_2109}, p_{Asuc_0391}, and p_{Asuc_0289} were amplified using primers CV 569 and CV570, CV571 and CV572, and CV578 and CV579, respectively (Table S2.1), cut with *Bam*HI and *Xba*I, and cloned into pLGZ920*-p_{Asuc_0701}-*lacZ*'s *Bam*HI and *Xba*I sites (these restrictions excise p_{Asuc_0701}) by ligation using T4 DNA ligase (New England Biolabs, Ipswich, MA). Colonies were screened using CV615 and the reverse primers used for amplifying each promoter fragment. Positive clones were verified by sequencing. All four plasmid constructs were then transformed into strain Δ*lacZ*.

2.3.11 Construction of strains expressing putative succinate transporters.

Asuc_0142, Asuc_1990-1991, Asuc_1999, and Asuc_2058 were amplified from *A. succinogenes* genomic DNA using primers CV579 and CV580, CV582 and CV586, CV587 and

CV588, and CV589 and CV590, respectively. The four PCR fragments were cloned in pLGZ920-$p_{pckA-92}$-*lacZ* and pLGZ920-$p_{pckA-103}$-*lacZ* cut with *Xba*I and *Sac*I (which remove *lacZ*) by Gibson cloning. Colonies were screened by PCR with GZ1302 and GZ1303 and positive clones were verified by sequencing. All constructs were then transformed into *A. succinogenes* 130Z.

2.3.12 High-performance liquid chromatography (HPLC) analysis of fermentation media from strains expressing transporters.

A. succinogenes growth in liquid cultures was monitored by optical density at 660 nm (OD_{660}) on a DU 650 spectrophotometer (Beckman, Fullerton, CA). OD_{660} values were used to calculate carbon balances. Glucose and fermentation products were quantified in culture supernatants by HPLC (Waters, Milford, MA) using an Aminex HPX-87H column (Bio-rad). Samples were run at room temperature with 4 mM H_2SO_4 as the eluent at a 0.6 mL min^{-1} flow rate. Glucose and ethanol were quantified on a Waters 410 differential refractometer. Organic acids were quantified on a Waters 2487 UV detector at 210 nm.

2.4 Results and Discussion

2.4.1 Potential C_4-dicarboxylate exporters in *A. succinogenes*

Succinate as a fermentation product is produced by fumarate reductase, whose active site faces the cytoplasm. Succinate has then to be exported out of the cell. *A. succinogenes* succinate exporters are completely unknown. Table 2.2 lists the putative C_4-dicarboxylate transporters identified in the *A. succinogenes* 130Z genome by BlastP using known C_4-dicarboxylate transporters from other bacterial species. *A. succinogenes* contains homologs of *E. coli* DcuA,

DcuB, DcuC, DauA, CitT, and YjjPB, as well as one homolog of *C. glutamicum* SucE, and multiple homologs of *C. glutamicum* DcsT. It also contains a protein of the tellurite-resistance/carboxylate transport (TDT) family, up to nine tripartite ATP-independent periplasmic (TRAP family) transporters, and up to five divalent anion:Na^+ symporters (DASS family).

Table 2.2 *A. succinogenes* putative C4-dicarboxylate transporters

Asuc locus	Putative function
0020	Anion transporter, ArsB/NhaD family
0023	Putative transporter, aspartate:alanine exchanger family, 32% identical to *C. glutamicum* SucE
0074	C_4-dicarboxylate transporter/malic acid transport protein
0142	Anaerobic C_4-dicarboxylate membrane transporter, 43% identical to *E. coli* DcuA
0146	Possible TRAP C_4-dicarboxylate transporter solute receptor, DctP
0147	Possible TRAP C_4-dicarboxylate transporter, small permease component, DctQ
0148	Possible TRAP C_4-dicarboxylate transporter, large permease component, DctM
0156	Possible TRAP C_4-dicarboxylate transporter, large permease component, DctM
0157	Conserved hypothetical protein (TRAP, small permease component, DctQ)
0158	Possible TRAP C_4-dicarboxylate transporter solute receptor, DctP
0183	Anion transporter; Anion transporter, ArsB/NhaD family; 33% identical to *E. coli* CitT
0270	Conserved hypothetical protein (TRAP, small permease component, DctQ)
0271	Possible TRAP C_4-dicarboxylate transporter, large permease component, DctM
0272	Possible TRAP C_4-dicarboxylate transporter solute receptor, DctP
0273	Possible TRAP C_4-dicarboxylate transporter solute receptor, DctP
0304	Anion transporter, DASS family, involved in fumarate uptake
0366	Possible TRAP C_4-dicarboxylate transporter solute receptor, DctP
0367	Conserved hypothetical protein (TRAP, small permease component, DctQ)
0368	Possible TRAP C_4-dicarboxylate transporter, large permease component, DctM
0715	Putative Thr/Ser exporter, 52% identical to *E. coli* YjjB
0716	Putative Thr/Ser exporter, 50% identical to *E. coli* YjjP
1063	Anaerobic C_4-dicarboxylate antiporter, DcuC family
1163	Possible TRAP C_4-dicarboxylate transporter solute receptor, DctP
1164	Possible TRAP C_4-dicarboxylate transporter, small permease component, DctQ
1165	Possible TRAP C_4-dicarboxylate transporter, large permease component, DctM
1482	Anion transporter, ArsB/NhaD family; 31% identical to *E. coli* CitT
1568	Anion permease ArsB/NhaD, 36% identical to *C. glutamicum* DcsT
1577	Possible TRAP C_4-dicarboxylate transporter, large permease component, DctM
1578	Conserved hypothetical protein (TRAP, small permease component, DctQ)
1579	Possible TRAP C_4-dicarboxylate transporter solute receptor, DctP
1781	Hypothetical protein; form of DctA or dicarboxylate transport protein found in many bacterial families
	Possible TRAP C_4-dicarboxylate transporter, large permease component, DctM
1921	Possible TRAP C_4-dicarboxylate transporter, small permease component, DctQ
1922	Possible TRAP C_4-dicarboxylate transporter solute receptor, DctP
1923	TRAP transporter, 4TM/12TM fusion protein , DctQM
1957	TRAP transporter solute receptor, TAXI family
1958	TRAP transporter solute receptor, TAXI family
1988	UspA domain protein, involved in stress response
1989	TRAP transporter, 4TM/12TM fusion permease protein, DctQM
1990	TRAP transporter solute receptor, TAXI family
1991	Putative anaerobic C_4-dicarboxylate transporter, 75% identical to *E. coli* DcuB (contains a 100-residue
1999	insertion)
2058	Sulfate permease (sulP) family, 58% identical to *E. coli* DauA

Some of the TRAP transporters could be involved in C_4-dicarboxylate transport, but TRAP transporters are involved in solute uptake and are considered mostly unidirectional (28), thus they are unlikely to be involved in succinate efflux. DASS transporters are also involved in solute uptake, and the *A. succinogenes* DASS transporter SdcA (Asuc_0304) was shown to restore growth on C_4-dicarboxylates in a C_4-dicarboxylate-transport-negative *E. coli* strain (29). Other putative C_4-dicarboxylate transporters tested (Asuc_0074, _0142, _1063, _1482, and _1999) did not (29). Because DASS transporters work in the direction of a decreasing Na^+ gradient, Asuc_0304 is unlikely to be involved in anion efflux. In the same study, qPCR reactions suggested that none of the genes tested (i.e., Asuc_0074, _0142, _0304, _1063, _1482, and _1999) were induced during anaerobic growth on glucose compared to other growth conditions. Asuc_1063 was upregulated 4 and 6 times during aerobic and anaerobic growth on fumarate, respectively, suggesting that this transporter is involved in C_4-dicarboxylate uptake as well (29).

2.4.2 Proteomic analysis of *A. succinogenes* cytoplasmic membrane-enriched fractions

Based solely on the putative C4-dicarboxylate transporters identified in the genome and on the functions of the different transporters already characterized in *E. coli* and *C. glutamicum*, it is impossible to predict which *A. succinogenes* putative transporters are the best candidates for succinate export. For this reason, we performed a proteomics analysis of cytoplasmic membrane-enriched fractions from *A. succinogenes* anaerobically grown on different carbon sources (i.e., glucose, fructose, xylose, and glycerol-nitrate). Note that our cytoplasmic membrane fractions were only enriched in cytoplasmic membrane. They did not consist of not entirely purified cytoplasmic membranes. For this reason the fractions still contained outer membrane,

cytoplasmic, and periplasmic proteins. Of the top 250 hits, though, 101 were associated with the cytoplasmic membrane, with 14 associated with the outer membrane, and 103 cytoplasmic proteins, indicating that these fractions could still provide good information on cytoplasmic membrane protein abundance.

The putative transporter candidates with the most hits (in NSAF format) are listed in Table 2.3. Asuc_1999, a homolog of *E. coli* DcuB, had the highest number of hits across all conditions tested, followed by Asuc_1990, Asuc_2058, Asuc_0142, and Asuc_1163. Asuc_1990 is the fused permease component (DctQM) of a TRAP-family transporter. Note that the putative solute receptor likely associated with Asuc_1990, Asuc_1991, was detected with high probability (rank 963) in all growth conditions tested, even though it is a periplasmic protein. Asuc_2058 belongs to the sulfate permease family, pfam00916. All characterized members of this family so far are sulfate uptake transporters (30). Asuc_0142 is a DcuA homolog. Asuc_1163 is the putative solute receptor of a TRAP C_4-dicarboxylate transporter. Surprisingly, this periplasmic solute receptor was detected in all growth conditions, with a high ranking, while its putative cytoplasmic membrane partners, Asuc_1164 and Asuc_1165, were not detected in any growth condition tested (Table 2.3). Peptides of other putative C4 dicarboxylate transporters were also detected in our samples, but typically at lower levels and not consistently across all growth conditions. Among those, DcuC homolog Asuc_1063 was detected at very low levels across all growth conditions. Asuc_0304, which had been shown to be involved in succinate uptake (29), was not detected in any growth condition. Neither were Asuc_0023, a *C. glutamicum* SucE homolog, nor Asuc_0715-16, homologs of *E. coli* YjjBP (Tables 2.3 and 2.4).

Table 2.3 Proteomics and transcriptomics data for succinate transporters in *A. succinogenes*

			Glucose anaerobic		Fructose anaerobic		Xylose anaerobic		Glycerol nitrate		Glycerol microaerobic
Rank[a]	**Protein putative function**	**Asuc_ locus**	**Protein[b]**	**RNA[c,d]**	**Protein[b]**	**RNA[c]**	**Protein[b]**	**RNA[c]**	**Protein[b]**	**RNA[c]**	**RNA[c]**
46	Anaerobic C4-dicarboxylate transporter, DcuB homolog	1999	0.164 ± 0.015	2698 ± 310	0.171 ± 0.015	1230 ± 98	0.199 ± 0.022	2980 ± 133	0.116 ± 0.017	919 ± 193	1407 ± 219
125	TRAP-TAXI transporter, 4TM/12TM fusion protein	1990	0.058 ± 0.023	104 ± 25	0.080 ± 0.007	195 ± 35	0.038 ± 7E-04	120 ± 2	0.020 ± 0.03	134 ± 47	88 ± 20
963	TRAP transporter solute receptor, TAXI family	1991	0.037 ± 0.007	752 ± 222	0.052 ± 5E-04	1094 ± 333	0.050 ± 0.014	702 ± 54	0.028 ± 0.004	887 ± 483	368 ± 85
216	C4-dicarboxylic acid transporter DauA	2058	0.023 ± 0.007	377 ± 17	0.025 ± 0.001	163 ± 9	0.028 ± 0.002	199 ± 9	0.024 ± 0.001	159 ± 19	210 ± 37
272	Anaerobic c4-dicarboxylate transporter, DcuA homolog	0142	0.031 ± 0.04	407 ± 84	0.026 ± 0.008	387 ± 48	0.054 ± 0.007	1025 ± 76	0.072 ± 0.009	377 ± 78	234 ± 113
342	TRAP C4-dicarboxylate transporter solute receptor, DctP	1163	0.009 ±0.003	12 ± 1	0.001 ± 0.002	6 ± 1	0.016 ± 0.002	14 ± 2	0.035 ± 0.007	19 ± 9	11 ± 4
554	TRAP C4-dicarboxylate transporter solute receptor, DctP	0366	0.002 ± 7E-04	21 ± 3	0.002 ± 2E-05	5 ± 1	0.004 ± 5E-04	31 ± 1	0.003 ± 0.000	18 ± 5	34 ± 2
672	C4-dicarboxylate ABC transporter	1781	0.004 ± 0.004	56 ± 5	0.005 ± 0.001	65 ± 2	0.012 ± 0.003	72 ± 1	0.000 ± 0.000	114 ± 36	63 ± 5
751	C4-dicarboxylate ABC transporter, DASS family	1482	0.002 ±2E-04	98 ± 9	0.003 ± 0.002	164 ± 44	0.009 ± 6E-04	236 ± 13	0.003 ± 0.001	507 ± 87	140 ± 28
918	C4-dicarboxylate ABC transporter, DcuC homolog	1063	0.002 ± 0.004	55 ± 2	0.007 ± 0.002	30 ± 2	0.000 ± 0.000	39 ± 2	0.000 ± 0.000	121 ± 7	44 ± 8
1122	Hypothetical protein	0272		93 ± 12		16 ± 1		47 ± 2	0.012 ± 0.002	92 ± 22	75 ± 23
1138	Anion permease ArsB/NhaD	0020	0.003 ± 0.002	87 ± 55	0.003 ± 0.001	43 ± 13		136 ± 23		115 ± 72	115 ± 72
1258	C4-dicarboxylate ABC transporter permease	1577		40 ± 3		36 ± 3	0.001 ± 0.002	87 ± 0		64 ± 9	57 ± 19
1289	TRAP family transporter membrane component, DctQ	0367		13 ± 2		5 ± 1	0.003 ± 0.004	18 ± 1	0.006 ± 0.001	13 ± 3	24 ± 1
1372	TRAP transporter, 4TM/12TM fusion protein, DctQM	1957	6E-04 ± 0.001	9 ± 1		13 ± 2		18 ± 0		60 ± 35	13 ± 3
ND	Corynebacterium glutamicum SucE homolog	0023		342 ± 38		146 ± 37		245 ± 12		326 ± 140	33 ± 3
ND	Dicarboxylate transporter/tellurite-resistance protein	0074		101 ± 11		96 ± 18		98 ± 3		84 ± 29	93 ± 3
ND	Anion transporter, DASS family	0304		236 ± 9		212 ± 17		289 ± 9		266 ± 26	309 ± 23
ND	TRAP transporter solute receptor, TAXI family	1988		140 ± 26		278 ± 61		189 ± 12		179 ± 63	121 ± 30
ND	Anion transporter, similar to Asuc_1482	0183		13 ± 2		11 ± 1		12 ± 1		39 ± 9	9 ± 2
ND	Anion permease ArsB/NhaD	1568		29 ± 6		23 ± 3		28 ± 3		103 ± 30	247 ± 96

[a]Rank is the overall rank of each protein in the entire normalized proteomics data set
ND: not detected
[b]Proteomics data are in NSAF format
[c]Transcriptome data are expressed per nt
[d]Glu RNA data are average ± standard deviations from three independent biological replicates.

2.4.3 Transcript levels of *A. succinogenes* putative C_4-dicarboxylate transporters

As seen in Table 2.3, transcript levels for the putative C4 dicarboxylate transporters do not always agree well with detected protein levels, but this observation is not particularly surprising due to post transcriptional regulation, the typically higher stability of proteins compared to RNA, and the large variability in protein detectability by mass spectrometry. This variability can explain why some transporter genes were clearly transcribed (e.g., Asuc_0023, Asuc_0304, or Asuc_1988) but the corresponding protein was not detected in any growth condition. For each putative transporter listed, though, trends in transcript levels were relatively uniform across growth conditions, and four of the top six transporters (including Asuc_1991) identified by proteomics, had the highest transcript levels in anaerobic glucose cultures (i.e., Asuc_1999, Asuc_2058, Asuc_0142, and Asuc_1991).

As observed in our proteomics results, Asuc_1999 has the highest transcript levels of all putative C4-dicarboxylate transporters in all growth conditions tested, even in the presence of nitrate. It also ranked very high relative to all transcript levels in the cell (e.g., rank of 101 on glucose). Succinate production during growth on glycerol-nitrate is much lower than during anaerobic growth on glucose (27). Interestingly, *asuc_1999*'s transcript levels decreased almost 3-fold and its ranking fell almost 4-fold in glycerol-nitrate compared to anaerobic glucose cultures (Table 2.3). Based on these results, Asuc_1999 is a good candidate as a transporter involved in succinate export.

Transcript levels for *asuc_1990* and *asuc_1991* were relatively steady across growth conditions, with transcript levels for *asuc_1991* always higher than those for *asuc_1990*, suggesting that the two genes are not cotranscribed. The facts that the intergenic region between

asuc_1990 and *asuc_1991* is 336-nt long and that very few RNA reads mapped to this region support this conclusion.

Transcripts levels for *asuc_2058* were higher in anaerobic glucose cultures than in any other conditions, but they remained at least 100-fold higher than the lowest transcript levels detected for other genes in any conditions. Transcripts levels for *asuc_0142* were among the highest for putative C4 dicarboxylate transporters across all growth conditions, in good agreement with the proteomics results.

While the Asuc_1163 protein was consistently detected by proteomics in the growth conditions tested, the corresponding transcript levels were very low across all growth conditions, and transcript levels for the putative associated membrane components Asuc_1164 and Asuc_1165 were extremely low as well.

Our combined proteomics and RNAseq data suggest that Asuc_1999, Asuc_2058, Asuc_0142, and Asuc_1990-91 could be involved in succinate export. Of these, only Asuc_1999 and Asuc_0142 are homologs of known succinate exporters, DcuB and DcuA, respectively (Table 2.2). Our proteomics and RNAseq data also seem to exclude Asuc_0304, DcuC homolog Asuc_1063, SucE homolog Asuc_0023, and *E. coli* YjjBP homologs Asuc_0715-16 as succinate exporters. Quantitative real-time PCR studies by Rhie et al (2014) in *A. succinogenes* grown anaerobically on glucose did not show any induction of genes *asuc_0142*, *asuc_1063*, *asuc_1482*, and *asuc_1999*. Our RNAseq results do not agree with these results for Asuc_1999 and Asuc_0142 since we saw high transcript levels for both of these genes.

2.4.4 Construction of an expression reporter system.

E. coli promoters are typically poorly or not functional at all in Pasteurellaceae species. At the beginning of this study, the only *A. succinogenes* promoter that had been tested was that of the phosphoenolpyruvate carboxykinase gene, *pckA*, which is a strong, constitutive promoter. No transcriptomic data were available for *A. succinogenes* to give us a sense of which promoters were strong or weak and when they were functional. Because recombinant membrane protein overexpression often requires a well-calibrated, weak-to-moderately strong promoter, and because this single available *A. succinogenes* native promoter did not allow for any tunability of gene expression, we sought to decrease the strength of p_{pckA} by first identifying the *pckA* transcription start site (TSS) and then constructing truncated versions of the promoter. To test these truncated promoters, though, we needed a reporter expression system.

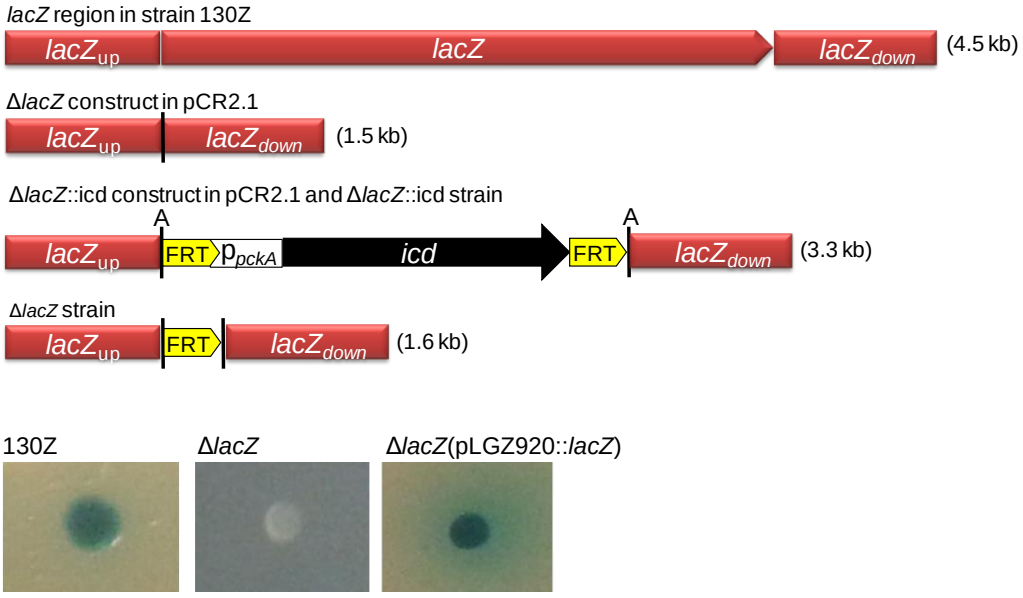

Figure 2.1 Construction of the Δ*lacZ* strain. (A) Schematic of the construction procedure. FRT, flippase recognition target site; A, *Asc*I restriction site. **(B)** *A. succinogenes* 130Z, Δ*lacZ*, and Δ*lacZ*(pLGZ920::*lacZ*) strains grown on LB-glucose supplemented with X-gal (40 µg mL^{-1}).

To test promoter strength in *A. succinogenes*, we used *lacZ* as the reporter gene. A *lacZ* deletion mutant of *A. succinogenes* was constructed (strategy described in Figure 2.1A) to be used as the background strain. The Δ*lacZ* strain was confirmed by testing *lacZ* expression on LB-glucose plates supplemented with 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal, 40 µg mL^{-1}) (Figure 2.1B). Plasmid pLGZ920::*lacZ*, with *lacZ* under control of p_{pckA}, was used as the positive control to test the reporter system (Figure 2.1B).

2.4.5 Identification of p_{pckA}'s TSS by RLM-RACE and construction of truncated p_{pckA} promoters

Out of the 14 clones sequenced, six clones indicated the TSS (+1) of *pckA* as a T nucleotide 32 nt upstream of the start codon. Sequences similar to *E. coli* promoter -10 and -35 consensus sequences were found 14 nt and 27 nt upstream of the TSS, respectively, with a 21-nt spacer sequence between the probable -10 and -35 sites (Figure 2.2).

ATTTGAAACGGATCACAAATCATGAAAAAAATACGTTCAAATTAGAACTAATTATCGAAAATTTGATCTA
GTTAACATTTTTTAGGTATAAATAGTTTTAAAATAGATCTAGTTTGGATTTTTAATTTTAAATTATCAAT
GAGGTGAAGT**ATGACTGACTTAAACAAACTCGTTAAAGAACT**

Figure 2.2 Identification of the *pckA* transcription start site. 5' region of the *A. succinogenes* *pckA* gene. Yellow highlight: probable -35 region; green highlight: probable -10 region; red and bold: transcription start site; underline: ribosome binding site; bold characters: start of the *pckA* coding region.

In our earlier studies, we routinely used a 231-nt sequence upstream of the *pckA* start codon to express genes under the control of p_{pckA}. We built truncated versions of this 231-nt sequence to decrease the strength of the promoter (Figure 2.3), and tested the strength of these truncated promoters using *lacZ* as the reporter gene. Activity of the truncated promoters decreased with the

length of the promoter region, and ranged from 1.4-fold ($p_{pckA-164}$ promoter) to 20-fold lower ($p_{pckA-92}$ promoter) than p_{pckA} (Figure 2.4). Almost no β-galactosidase activity was detected with promoter $p_{pckA-75}$ (data not shown), which suggests that the identified putative -35 region is indeed needed for promoter function.

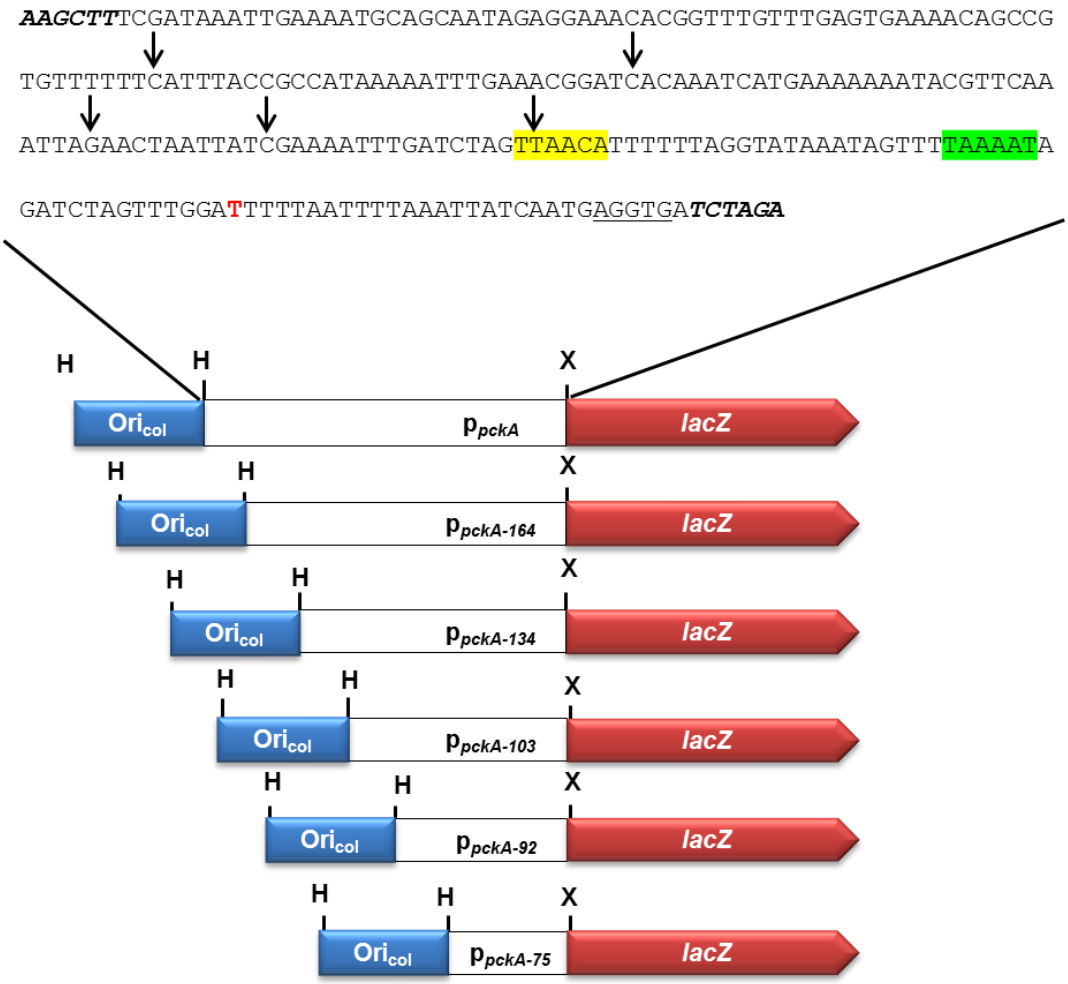

Figure 2.3 Reporter constructs with *lacZ* under control of truncated versions of *pckA*'s 5'UTR. H, *Hind*III; X, *Xba*I; yellow highlight: probable -35 region; green highlight: probable -10 region; red and bold: transcription start site; underline: ribosome binding site; italics: indicate restriction sites—*Hind*III (5' end) and *Xba*I (3' end); vertical arrows indicate start of truncated promoters. Drawings of DNA fragments are not to scale.

2.4.6 Identification of promoters of different strengths from the transcriptomics results.

To expand the expression range of our promoter library toward expression levels lower than from $p_{pckA\text{-}92}$, we used our RNAseq results. Candidate promoters were selected from the genes whose transcript levels did not vary between growth conditions using the RNAseq results from anaerobic cultures grown on glucose, xylose, fructose, and mannose, as well as glycerol microaerobic cultures, which all favor succinate production. Genes whose $\log_2$ fold-change was between -0.2 and 0.2 in all growth conditions compared to glucose were compiled and their transcript levels (using expression levels per nucleotide) (Table S2.2) were ranked, together with that of *pckA*, for further analysis. Candidates (individual genes or the first genes in an operon) were selected that covered a large range of decreasing strengths compared to p_{pckA}. Promoters of candidate genes *asuc*_0391, *asuc_0289*, *asuc_0701*, and *asuc_2109* were selected, whose transcript levels were 16-, 24-, 29-, and 62-fold lower than *pckA*'s. DNA fragments 247 bp upstream of candidate genes *asuc_0701*, *asuc_0289*, and *asuc_0391* and 261-bp upstream of *asuc_2109* were cloned in front of *lacZ* to be tested as promoters.

Promoter strength was tested by β-galactosidase assays using Δ*lacZ* as the expression strain, with *lacZ* under control of p_{pckA} as the positive control (Figure 4). Promoter p_{Asuc_0701} was the weakest promoter, 209-fold weaker than p_{pckA}, which was the strongest. To the best of our knowledge this is the first promoter library constructed for *A. succinogenes* 130Z, making it easier to test expression of different genes in this strain for succinate production.

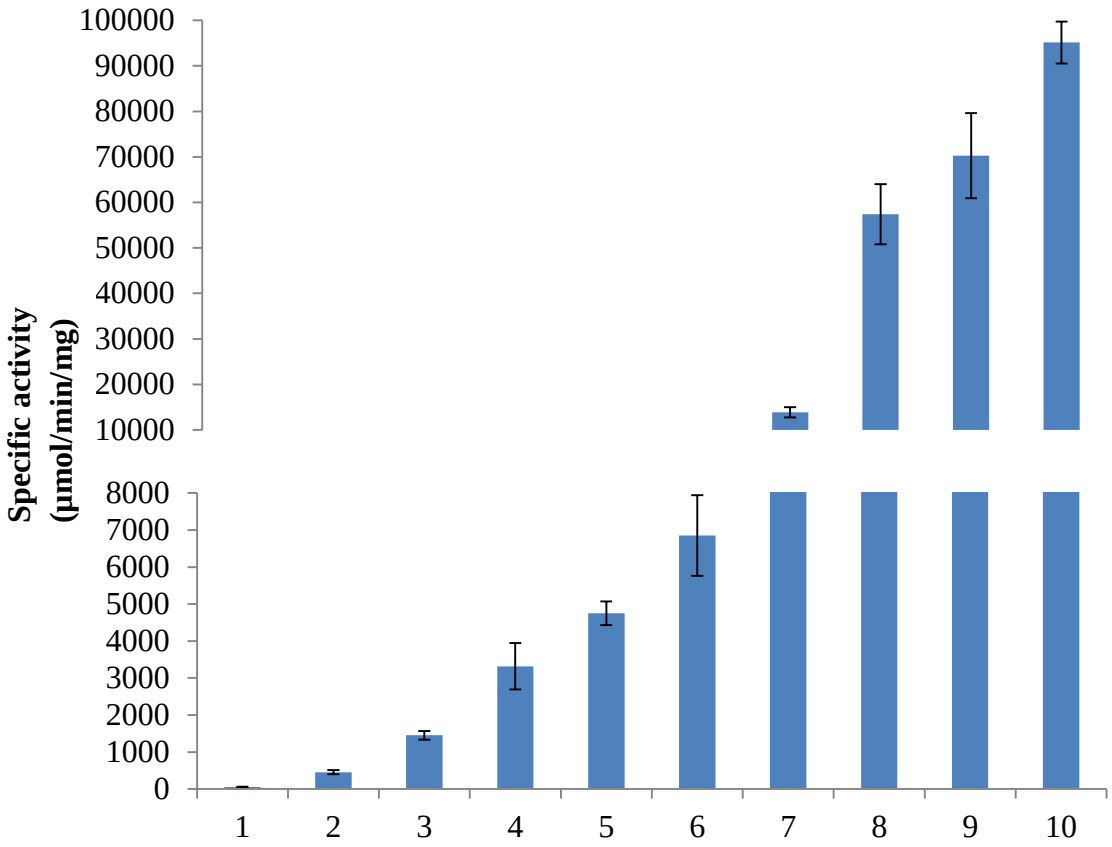

Figure 2.4 β-galactosidase activity of promoter reporter constructs. All promoter constructs were tested in the *ΔlacZ* strain. 1, *ΔlacZ*(pLGZ920); 2, *ΔlacZ*(pLGZ920*-p_{Asuc_0701}-*lacZ*); 3, *ΔlacZ*(pLGZ920*-p_{Asuc_2109}-*lacZ*); 4, *ΔlacZ*(pLGZ920*-p_{Asuc_0391}-*lacZ*); 5, *ΔlacZ*(pLGZ920-$p_{pckA-92}$-*lacZ*); 6, *ΔlacZ*(pLGZ920*-p_{Asuc_0289}-*lacZ*); 7, *ΔlacZ*(pLGZ920-$p_{pckA-103}$-*lacZ*); 8, *ΔlacZ*(pLGZ920-$p_{pckA-134}$-*lacZ*); 9, *ΔlacZ*(pLGZ920*-$p_{pckA-164}$-*lacZ*); 10, *ΔlacZ*(pLGZ920-*lacZ*). Results are the average activity ± standard deviation from three independent biological replicates.

2.4.7 Overexpression of putative succinate transporter candidates in *A. succinogenes* and succinate production.

Succinate transporter candidates with the most hits in the proteomics analysis and the highest transcript levels were over-expressed to determine their effects on succinate production. We focused on Asuc_1999, Asuc_1990-1991, Asuc_2058, and Asuc_0142. Even though

Asuc_1990 and Asuc_1991 seemed to be transcribed independently, they are likely part of the same transport system, so the two genes were expressed in a single construct. For each putative transporter, we tested several weak promoters to maximize protein expression while avoiding toxicity to the cells. Expression of all four transporters under control of $p_{pckA\text{-}134}$ was toxic to the cells, thus we tested expression of these proteins under control of $p_{pckA\text{-}92}$ and $p_{pckA\text{-}103}$. Asuc_2058 and Asuc_1990-1991 expressed under control of $p_{pckA\text{-}103}$ proved toxic to the cells as well, but expression of Asuc_0142 and Asuc_1999 under the same promoter did not inhibit cell growth. Promoter $p_{pckA\text{-}92}$ allowed expression of all four transporters without inhibiting growth of the 130Z recombinant strain.

Fermentation balances for *A. succinogenes* 130Z carrying the six expression constructs are shown in Table 2.4. All six expression constructs caused *A. succinogenes* 130Z's succinate yields to increase significantly, compared to the empty vector, in anaerobic glucose cultures. Much of the succinate yield increase seemed to be at the expense of biomass production. Note that expressing *Asuc_1999* and *Asuc_0142* genes under control of $p_{pckA\text{-}103}$ increased the succinate yield compared to $p_{pckA\text{-}92}$ for both transporters, although the increase was not statistically significant.

The strain with the highest succinate yield, 130Z (pLGZ920-$p_{pckA\text{-}103}$-*Asuc_1999*), also produced almost the highest acetate yield, but produced the least biomass. In contrast, the second best succinate producing strain, 130Z (pLGZ920-$p_{pckA\text{-}92}$-*Asuc_1990-1991*), produced almost as much succinate, plus it did not produce more acetate than the control strain, making 130Z (pLGZ920-$p_{pckA\text{-}92}$-*Asuc_1990-1991*) possibly a more suitable candidate for succinate production and further engineering.

Table 2.4 Fermentation balances of strain 130Z over-expressing candidate succinate transporters.

	Products (mmol/100 mmol glucose consumed)								
Strain	Succinate	Formate	Acetate	Ethanol	CO_2	Biomass[a]	Carbon recovery[b]	% increased succinate yield	Doubling time
130Z (pLGZ920)	47.3 ± 2.3	98.0 ± 5.4	65.5 ± 2.8	30.6 ± 2.8	0.0 ± 0.0	193 ± 10	104 ± 5	NA	2.04 ± 0.03
130Z (pLGZ920-$p_{pckA\text{-}92}$-*Asuc_1999*)	55.0 ± 0.8	103 ± 1	72.4 ± 1.1	28.5 ± 2.0	0.0 ± 0.0	157 ± 2	104 ± 1	16**	2.22 ± 0.05
130Z (pLGZ920-$p_{pckA\text{-}92}$-*Asuc_0142*)	53.7 ± 1.4	98.5 ± 2.4	64.6 ± 2.1	25.5 ± 4.2	0.0 ± 0.0	188 ± 6	104 ± 3	14*	2.27 ± 0.02
130Z (pLGZ920-$p_{pckA\text{-}103}$-*Asuc_1999*)	58.5 ± 2.8	101 ± 6	73.0 ± 3.3	25.2 ± 3.2	0. 0± 0.0	138 ± 5	101 ± 4	24**	2.74 ± 0.2
130Z (pLGZ920-$p_{pckA\text{-}103}$-*Asuc_0142*)	56.4 ± 1.1	101 ± 3	73.9 ± 1.6	29.8 ± 0.9	0.5 ± 0.3	152 ± 8	105 ± 3	19**	2.46 ± 0.01
130Z (pLGZ920-$p_{pckA\text{-}92}$-*Asuc_2058*)	55.7 ± 2.4	98.1 ± 5.2	68.4 ± 2.9	28.5 ± 0.9	0.0 ± 0.0	142 ± 2	100 ± 4	18*	2.49 ± 0.06
130Z (pLGZ920-$p_{pckA\text{-}92}$-*Asuc_1990-91*)	57.2 ± 2.2	96.9 ± 0.6	65.0 ± 0.8	26.8 ± 1.7	0.0 ± 0.0	154 ± 2	101 ± 1	21**	2.26 ± 0.05

Results are an average of three biological replicates ± standard deviations.

[a]Biomass was determined using assumed values of 567 mg dry cell weight/mL per OD_{660}(31) and a cell composition of $CH_2O_{0.5}N_{0.2}$ (24.967 g/mol) (32)

[b]Carbon balance is the carbon in products/carbon in glucose consumed. It is assumed that one CO_2 is fixed for each molecule of succinate produced.

CO_2 was calculated using the following formula:

CO_2 (in mM) = acetate (in mM) + ethanol (in mM) - formate (in mM)

* Significantly different from 130Z (pLGZ920) ($p < 0.05$, two-tailed student's t-test)

** Significantly different from 130Z (pLGZ920) ($p < 0.01$, two-tailed student's t –test)

When attempting to express all the constructs in the Δ*pflB* strain (a pyruvate-formate lyase deletion mutant) (21), which does not produce any formate, we were either unable to obtain any transformants or the strains grew extremely poorly, indicating that these constructs are toxic in Δ*pflB* and that weaker promoters are likely needed in this strain (data not shown).

2.4.8 The *A. succinogenes* succinate efflux transporters differ from those of *E. coli* and *C. glutamicum*

A. succinogenes contains homologs of many *E. coli* succinate transporters—DcuA, DcuB, DcuC, DauA, CitT, and YjjPB. It also contains one homolog of *C. glutamicum* SucE, along with multiple homologs of *C. glutamicum* DcsT.

E. coli DcuC can perform different functions, but, in particular, it has been shown to participate in succinate efflux under anaerobic conditions (9). Asuc_1063 is the *A. succinogenes* DcuC homolog, but it does not show high transcriptional levels or proteomics hits in our studies, eliminating it as a succinate efflux transporter. *E. coli* YjjP and YjjB constitute a major succinate efflux transporter (17) but Asuc_0715 and Asuc_0716, which are 52% and 50% identical to YjjB and YjjP, respectively, were not detected in our proteomics analysis.

Golby et al. (33) determined that *E. coli* DcuB is subject to catabolite repression, repressed by nitrate, and strongly induced by C4-dicarboxylates. These authors suggested that it is mostly involved in C4-dicarboxylate transport during fumarate respiration. A quintuple or sextuple deletion mutant, including Δ*dcuB,* did not show decreased succinate efflux (13), but then a single *dcuB* mutant affected succinate efflux during growth on glucose (9), thus DcuB could contribute to succinate efflux in *E. coli*. Asuc_1999, a putative anaerobic C_4-dicarboxylate transporter that is 75% identical to *E. coli* DcuB increases the succinate yield in *A. succinogenes* when

overexpressed. Asuc_1999 had the most hits in our proteomics analysis and had the highest transcript levels of all putative C4-dicarboxylate transporters across all conditions tested, suggesting that it plays a major role in succinate efflux in *A. succinogenes*.

Overexpressing *Asuc_1990-1991* in *A. succinogenes* caused an increase in succinate yield. This result is rather surprising, since Asuc_1990-1991 belongs to the TRAP family of transporters, which are primarily involved in uptake transport. Asuc_1990-1991 had the second highest hits in our proteomics analysis and had high transcriptional levels across all the conditions we tested.

Asuc_0142, an anaerobic C4-dicarboxylate membrane transporter, is 43% identical to *E. coli* DcuA. Overexpression of Asuc_0142 in *A. succinogenes* also caused an increase in succinate yield. *E. coli* DcuA is known to function as both a succinate uptake and efflux transporter, as confirmed using double mutants expressing only one of the DcuA, DcuB, or DcuC carriers at a time (8).

Asuc_2058, showing similarity to *E. coli* DauA, also led to an increase in succinate yield upon overexpression. This result suggests that Asuc_2058 has a function completely different from that of *E. coli* DauA, which has been shown to be the main succinate uptake transporter under acidic conditions and to be inactive at pH 7 (10).

SucE has been identified as one of the *C. glutamicum* transporters involved in succinate efflux under microaerobic and anaerobic conditions. Asuc_0023, a homolog of *C. glutamicum* SucE had moderate transcript levels but was not detected in our proteomics studies, indicating that the *C. glutamicum* and *A. succinogenes* homologs do not have the same function.

Thus succinate efflux transporters in *A. succinogenes* are different from the major succinate exporters of *E. coli* and *C. glutamicum,* and even include transporters (e.g., TRAP-family transporters Asuc_1990-91) that have so far never been shown to function in export.

2.5 Conclusion

This study is the first to focus on *A. succinogenes* dicarboxylate efflux transporters with the aim of increasing succinate production. Combining proteomics and transcriptomics approaches, we identified the four putative C4-dicarboxylate transporters with the highest expression levels. Overexpressing all four transporters individually increased succinate production. While, based on our expression data, Asuc_1999 is a likely candidate to be a major succinate exporter, we cannot identify a single major succinate exporter in *A. succinogenes*. Our results suggest instead that more than one transporter is involved in succinate export in *A. succinogenes,* similar to the situation observed in *E. coli*. Knockout mutants of these transporters (construction unsuccessful so far) would shed more light on the relative involvement of these transporters in succinate efflux in *A. succinogenes*.

We also developed a library of *A. succinogenes* promoters covering a large expression range across multiple growth conditions, which can be used for modulating protein expression in *A. succinogenes* as needed for strain engineering.

2.6 Acknowledgements

I thank Dr. Claire Vieille and Peggy Wolf for carrying out all work related to proteomics. I thank Dr. Nikolas McPherson for obtaining and analyzing all RNA sequencing data.

APPENDIX

Table A2.1 Primers used in this study

Primer	Sequence[a] (restriction site)	Specificity[b] (direction[c])
P1	CGTGGTTAACGTCCCTTTATCG	5'RLM-RACE *pckA*-specific outer primer (R)
P2	GGTTTCTTCCTCGAAAAGTTGTTC	5'RLM-RACE *pckA*-specific inner primer (R)
CV376	CTTGCCAAACCGACCGAAAG	*ΔlacZ*-fusion, nested (F)
CV377	TATTGATAATGAAAATCCGACCGCACTTGGCAGTACCGGCGTATTCCTC[b]	*ΔlacZ*-fusion, nested (R)
CV443	CACACCATTCCCAAACAAAAC	Truncated p_{pckA} constructs (R)
CV469	GACCATGATTACGCCAAGCTTTCTGCTAATCCTGTTACCAGTGG	colE1 ori (F)
CV470	TAGGTTGGCAGAATCATCTAGATCACCTCATTGATAATTTAAAATTAAAAATCCAAACT AGATCTATTTTAAAAC	p_{pckA} with *lacZ* overhangs (R)
CV471	AAGCTTTCTTCCGCTTCCTCGCTCACTG	ColE1 ori (R)
CV473	GAGGAAGCGGAAGAAAGCTTCGAAAATTTGATCTAGTTAACATTTTTTAGG	$p_{pckA\text{-}92}$ (F)
CV474	GAGGAAGCGGAAGAAAGCTTGAACTAATTATCGAAAATTTGATCTAG	$p_{pckA\text{-}103}$ (F)
CV475	GAGGAAGCGGAAGAAAGCTTCACAAATCATGAAAAAAATACGTTC	$p_{pckA\text{-}134}$ (F)
CV476	GAGGAAGCGGAAGAAAGCTTCATTTACCGCCATAAAAATTTGAAAC	$p_{pckA\text{-}164}$ (F)
CV569	TGAGCGAGGAAGCGGAAGAGGATCCGGTCAAACTCCTACGAATTTC	p_{Asuc_2109} (F)
CV570	AAGTAGGTTGGCAGAATCATCTAGAGCTCACCAACAGGCTTGA	p_{Asuc_2109} (R)
CV571	TGAGCGAGGAAGCGGAAGAGGATCCGAGAGAACAGGACAACAGTTTTATTG	p_{Asuc_0391} (F)
CV572	AAGTAGGTTGGCAGAATCATCTAGAAATAATAACTCTTAATATAGAAAAAAACGATTG	p_{Asuc_0391} (R)
CV577	TGAGCGAGGAAGCGGAAGAGGATCCGACGAGCTACTCCCTGGTTTG	p_{Asuc_0289} (F)
CV578	AAGTAGGTTGGCAGAATCATCTAGAGATCCTTTTAAAAAAAACGATGAAAAAAT	p_{Asuc_0289} (R)
CV579	TAAATTATCAATGAGGTGATCTAGATGACTGCAATGTTTATTATCC	*Asuc_0142* (F)
CV580	ACGGCCAGTGAATTCGAGCTCTTAGATAAACAGATTCGCAAATAAC	*Asuc_0142* (R)
CV586	TAAATTATCAATGAGGTGATCTAGATGAAAAAATTATTTAAACTTTCTCTTGTC	*Asuc_1990-1991* (F)
CV582	ACGGCCAGTGAATTCGAGCTCTATTCCGGACGACGACG	*Asuc_1990-1991* (R)
CV587	TAAATTATCAATGAGGTGATCTAGATGGATTTTTTGATGAATCTAAG	*Asuc_1999* (F)
CV588	ACGGCCAGTGAATTCGAGCTCTTAAAGATAACCGTATAAGCCTG	*Asuc_1999* (R)
CV589	TAAATTATCAATGAGGTGATCTAGATGCTAAATAAATGGTTTTTAACC	*Asuc_2058* (F)
CV590	ACGGCCAGTGAATTCGAGCTCTTAATGACGTAATTCCGATTC	*Asuc_2058* (R)
CV615	GACCATGATTACGCCAAGCTTTCTGCTAATCCTGTTACCAGTGGC	ColE1 ori (F)
CV616	ATCACCTTAGGATCCCCGCATCAGGCGCTCTTC	ColE1 ori (R), replaces ori's 3' *Hind*III site with *Bam*HI
CV617	TGATGCGGGGATCCTAAGGTGATTTATAGTCTGGACGG	p_{Asuc_0701} (F)
CV618	GTACCCGGGGATCCTCTAGATTCACCTCGAAACAGATAAAAAAATC	p_{Asuc_0701} (R)
CV624	AAATTATCAATGAGGTGATCTAGATGATTCTGCCAACCTACTTTGAAAATCC	*lacZ* Gibson cloning (F)
CV625	CCAGTGAATTCGAGCTCTTATTCAAAGTGAATATCGAAACGACAGTCAAATTTTG	*lacZ* Gibson cloning (R)
GZ1302	CGTTGTAAAACGACGGCC	pLGZ901 specific primer downstream of *Sac*I site (R)
GZ1303	AATTTTAAATTATCAATGAGGTG	pLG901 specific primer upstream of *Xba*I site (F)
CV431	GGTTTCGCCCACTCGTATTCC	*lacZ* sequencing primer-1

Table A2.1 (cont'd)

CV445	GCTCATATGAAATTACATAGCCTGTTTTCGGATCGC	*lacZ* sequencing primer-2
CV446	CTATGTAATTTCATATGAGCAGGCGTTAGTTGATTTG	*lacZ* sequencing primer-3
CV636	GGATATGGTGAAAACTTCGAAG	*lacZ* sequencing primer-4
CV637	CCATTGTTTGGTTTACTCTCC	*lacZ* sequencing primer-5

[a]F, forward primer; R, reverse primer

[b] Primer P4 contains the USS ACCGCACTT

Table A2.2 Transcript levels of candidate genes with low variation across several conditions, compared to *pckA*

Gene ID	Average expression per nucleotide	Expression compared to *pckA*	Fold lower expression than *pckA*
Asuc_2109	88	0.016	61.7
Asuc_0708	93	0.017	58.7
Asuc_1818	101	0.019	53.8
Asuc_1044	125	0.023	43.4
Asuc_1536	130	0.024	41.8
Asuc_1879	145	0.027	37.5
Asuc_1430	170	0.031	32.1
Asuc_0701	189	0.035	28.8
Asuc_0289	227	0.042	24.0
Asuc_1432	228	0.042	23.9
Asuc_2046	246	0.045	22.2
Asuc_1945	272	0.050	20.0
Asuc_1699	285	0.052	19.1
Asuc_1071	310	0.057	17.6
Asuc_0391	332	0.061	16.4
Asuc_1037	469	0.086	11.6
Asuc_0983	472	0.087	11.5
Asuc_1877	483	0.089	11.3
Asuc_2007	549	0.101	9.9
Asuc_2066	1423	0.261	3.8

REFERENCES

REFERENCES

1. **Zeikus JG**, **Jain MK**, **Elankovan P**. 1999. Biotechnology of succinic acid production and markets for derived industrial products. Appl Biochem Biotechnol **51**:545-552.

2. **Sanchez AM**, **Bennett GN**, **San KY**. 2005. Efficient succinic acid production from glucose through overexpression of pyruvate carboxylase in an *Escherichia coli* alcohol dehydrogenase and lactate dehydrogenase mutant. Biotechnol Prog **21**:358-365.

3. **Jantama K**, **Haupt MJ**, **Svoronos SA**, **Zhang XL**, **Moore JC**, **Shanmugam KT**, **Ingram LO**. 2008. Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of *Escherichia coli* C that produce succinate and malate. Biotechnol Bioeng **99**:1140-1153.

4. **Litsanov B**, **Brocker M**, **Bott M**. 2012. Toward homosuccinate fermentation: metabolic engineering of *Corynebacterium glutamicum* for anaerobic production of succinate from glucose and formate. Appl Environ Microbiol **78**:3325-3337.

5. **Ito Y**, **Hirasawa T**, **Shimizu H**. 2014. Metabolic engineering of *Saccharomyces cerevisiae* to improve succinic acid production based on metabolic profiling. Biosci Biotechnol Biochem **78**:151-159.

6. **Six S**, **Andrews SC**, **Unden G**, **Guest JR**. **1994**. *Escherichia coli* possesses two homologous anaerobic C4-dicarboxylate membrane transporters (DcuA and DcuB) distinct from the aerobic dicarboxylate transport system (Dct). J Bacteriol **176**:6470-6478.

7. **Engel P**, **Kramer R**, **Unden G**. 1994. Transport of C4-dicarboxylates by anaerobically grown *Escherichia coli.* Energetics and mechanism of exchange, uptake and efflux. Eur J Biochem **222**:605-614.

8. **Zientz E**, **Six S**, **Unden G**. 1996. Identification of a third secondary carrier (DcuC) for anaerobic C4-dicarboxylate transport in *Escherichia coli*: roles of the three Dcu carriers in uptake and exchange. J Bacteriol **178**:7241-7247.

9. **Janausch IG**, **Unden G**. 1999. The *dcuD* (former *yhcL*) gene product of *Escherichia coli* as a member of the DcuC family of C4-dicarboxylate carriers: lack of evident expression. Arch Microbiol **172**:219-226.

10. **Karinou E**, **Compton EL**, **Morel M**, **Javelle A**. 2013. The *Escherichia coli* SLC26 homologue YchM (DauA) is a C_4-dicarboxylic acid transporter. Mol Microbiol **87**:623-640.

11. **Pos KM**, **Dimroth P**, **Bott M**. 1998. The *Escherichia coli* citrate carrier CitT: a member of a novel eubacterial transporter family related to the 2-oxoglutarate/malate translocator from spinach chloroplasts. J Bacteriol **180**:4160-4165.

12. **Davies SJ**, **Golby P**, **Omrani D**, **Broad SA**, **Harrington VL**, **Guest JR**, **Kelly DJ**, **Andrews SC**. 1999. Inactivation and regulation of the aerobic C_4-dicarboxylate transport (dctA) gene of *Escherichia coli*. J Bacteriol **181**:5624-5635.

13. **Janausch IG**, **Kim OB**, **Unden G**. 2001. DctA- and Dcu-independent transport of succinate in *Escherichia coli*: contribution of diffusion and of alternative carriers. Arch Microbiol **176**:224-230.

14. **Chen J**, **Zhu X**, **Tan Z**, **Xu H**, **Tang J**, **Xiao D**, **Zhang X**. 2014. Activating C4-dicarboxylate transporters DcuB and DcuC for improving succinate production. Appl Microbiol Biotechnol **98**:2197-2205.

15. **Huhn S**, **Jolkver E**, **Kramer R**, **Marin K**. 2011. Identification of the membrane protein SucE and its role in succinate transport in *Corynebacterium glutamicum*. Appl Microbiol Biotechnol **89**:327-335.

16. **Fukui K**, **Koseki C**, **Yamamoto Y**, **Nakamura J**, **Sasahara A**, **Yuji R**, **Hashiguchi K**, **Usuda Y**, **Matsui K**, **Kojima H**, **Abe K**. 2011. Identification of succinate exporter in *Corynebacterium glutamicum* and its physiological roles under anaerobic conditions. J Biotechnol **154**:25-34.

17. **Fukui K**, **Nanatani K**, **Hara Y**, **Yamakami S**, **Yahagi D**, **Chinen A**, **Tokura M**, **Abe K**. 2017. Escherichia coli *yjjPB* genes encode a succinate transporter important for succinate production. Biosci Biotechnol Biochem **81**:1837-1844.

18. **Kim P**, **Laivenieks M**, **McKinlay J**, **Vieille C**, **Zeikus JG**. 2004. Construction of a shuttle vector for the overexpression of recombinant proteins in *Actinobacillus succinogenes*. Plasmid **51**:108–115.

19. **Ausubel FM**, **Brent R**, **Kingston RE**, **Moore DD**, **Seidman JG**, **Smith JA**, **Struhl K**. 1993. Current Protocols in Molecular Biology. Greene Publishing & Wiley-Interscience, New York.

20. **McKinlay JB**, **Zeikus JG**, **Vieille C**. 2005. Insights into *Actinobacillus succinogenes* fermentative metabolism in a chemically defined growth medium. Appl Environ Microbiol **71**:6651-6656.

21. **Joshi RJ**, **Schindler BD**, **McPherson NR**, **Tiwari K**, **Vieille C**. 2014. Development of a markerless knockout method for *Actinobacillus succinogenes*. Appl Environ Microbiol **80**:3053-3061.

22. **Lee SJ**, **Song H**, **Lee SY**. 2006. Genome-based metabolic engineering of *Mannheimia succiniciproducens* for succinic acid production. Appl Environ Microbiol **72**:1939–1948.

23. **Tai SP**, **Kaplan S**. 1985. Intracellular localization of phospholipid transfer activity in *Rhodopseudomonas sphaeroides* and a possible role in membrane biogenesis. J Bacteriol **164**:181-186.

24. **Zhang Y**, **Wen Z**, **Washburn MP**, **Florens L**. 2010. Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. Anal Chem **82**:2272-2281.

25. **McPherson N**. 2017. Novel Insights Into Sugar and Succinate Metabolism of *Actinobacillus succinogenes* from Strains Evolved for Improved Growth on Lignocellulose Hydrolysate Sugars. Doctoral. Michigan State University.

26. **Johnson BK**, **Scholz MB**, **Teal TK**, **Abramovitch RB**. 2016. SPARTA: Simple program for automated reference-based bacterial RNA-seq transcriptome analysis. BMC Bioinformatics **17**:66.

27. **Schindler BD**, **Joshi RJ**, **Vieille C**. 2014. Respiratory glycerol metabolism of *Actinobacillus succinogenes* 130Z for succinate production. J Ind Microbiol Biotechnol **41**:1339-1352.

28. **Mulligan C**, **Fischer M**, **Thomas GH**. 2011. Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. FEMS Microbiol Rev **35**:68-86.

29. **Rhie MN**, **Yoon HE**, **Oh HY**, **Zedler S**, **Unden G**, **Kim OB**. 2014. A Na+-coupled C4-dicarboxylate transporter (Asuc_0304) and aerobic growth of *Actinobacillus succinogenes* on C4-dicarboxylates. Microbiology **160**:1533-1544.

30. **Aguilar-Barajas E**, **Diaz-Perez C**, **Ramirez-Diaz MI**, **Riveros-Rosas H**, **Cervantes C**. 2011. Bacterial transport of sulfate, molybdate, and related oxyanions. Biometals **24**:687-707.

31. **McKinlay JB**, **Shachar-Hill Y**, **Zeikus JG**, **Vieille C**. 2007. Determining *Actinobacillus succinogenes* metabolic pathways and fluxes by NMR and GC-MS analyses of ^{13}C-labeled metabolic product isotopomers. Metab Eng **9**:177–192.

32. **van der Werf MJ**, **Guettler MV**, **Jain MK**, **Zeikus JG**. 1997. Environmental and physiological factors affecting the succinate product ratio during carbohydrate fermentation by *Actinobacillus* sp. 130Z. Arch Microbiol **167**:332–342.

33. **Golby P**, **Kelly DJ**, **Guest JR**, **Andrews SC**. 1998. Transcriptional regulation and organization of the *dcuA* and *dcuB* genes, encoding homologous anaerobic C4-dicarboxylate transporters in *Escherichia coli*. J Bacteriol **180**:6586-6596.

Chapter 3

Identification of sRNA in *Actinobacillus succinogenes* 130Z

3.1 Abstract

Small RNAs (sRNAs) are powerful tools in metabolic engineering and synthetic biology. Synthetic sRNAs are being designed using Hfq-dependent sRNA scaffolds. Hfq is a chaperone protein known to facilitate binding of sRNAs to their mRNA targets and also protecting them from degradation. Recent studies have focused on the identification of sRNAs in a few commercially-relevant organisms, particularly *Escherichia coli*, to develop them as tools for metabolic engineering. *Actinobacillus succinogenes* is one such industrially-relevant organism, known to be one the best natural succinate producers. In this study, we identified sRNAs in *A. succinogenes* to gain better insight into their sequence, size, structure, and function. Our goal was to gain enough information to be able to design synthetic sRNAs that would act as post-transcriptional regulators in our strain. We performed RNAseq analysis of the wild-type strain grown microaerobically on glycerol and anaerobically on glucose. RNAseq data were analyzed using Rockhopper and manually using the Integrated Genomics Viewer. Using ARNold, we identified sRNAs with Rho-independent terminators, a key characteristic of Hfq-dependent sRNAs. We also looked for other known features of Hfq-dependent sRNAs, which allowed us to narrow our focus down to only a few candidates for our synthetic sRNA design. Scaffolds from two sRNAs—smRNA8 and smRNA28, were used in synthetic sRNA constructs to test inhibition of β-galactosidase expression in *A. succinogenes*. A 32% decrease in β-galactosidase activity was seen with one of the constructs. The target binding region for *lacZ* was then replaced by the target binding regions for *ackA* and *pta*. One of four constructs we tested caused a 14% decrease in acetate yield. To our knowledge, this is the first study identifying sRNAs in *A. succinogenes*. We have also provided a proof of concept for using synthetic sRNAs in *A. succinogenes* as a

metabolic engineering tool. However, more optimization of these synthetic sRNAs in *A. succinogenes* is needed to develop a robust and tunable system for future engineering efforts.

3.1 Introduction

Small RNAs (~30-500 nucleotides) are regulatory RNAs known to regulate mRNA transcript levels and translation in bacteria. Since they were first discovered 30 years ago, a large number of small RNAs (sRNAs) have been identified in several bacterial species, in part due to advances in deep RNA sequencing and computational tools (1, 2). It is now known that sRNAs have a significant role in gene regulation, but understanding of their mechanisms of action lags behind, and many of their targets remain unknown. sRNAs are most commonly classified by their mechanism of action as reviewed by Gottesman et al. (3). Antisense sRNAs can be classified as cis-encoded sRNAs or trans-encoded sRNAs. Most commonly studied are trans-encoded antisense sRNAs (encoded by a locus different from their target[s]) that bind to their target mRNA(s) with limited and varying degrees of complementarity. Due to their limited binding capabilities these sRNAs may require the assistance of Sm-like protein Hfq (4) Cis-encoded antisense sRNAs are usually encoded by the same locus as the mRNA but on the opposite strand, making them entirely complementary to their target mRNA, in contrast to trans-encoded antisense sRNAs. Riboswitches are another group of cis-encoded and -acting RNA elements, which sense small molecules and change conformation to allow or block translation (5). In some cases they have been shown to act in trans (6). CRISPRs (clustered regularly interspaced short palindromic repeats) are a newly identified group of sRNAs playing an important role in bacterial immunity against foreign DNA (7), (3). Many small RNAs are also synthesized from the 3’ regions of mRNA. These small RNAs are produced by mRNA

processing or transcribed separately with a shared terminator (8). A couple of studies have shown that 3'-UTR derived sRNAs acted as trans regulators for mRNAs (9, 10). DapZ was the first example of a small RNA that was transcribed by an independent promoter from the sense 3'UTR of a gene, *dapB* (10). DapB is responsible for the catalyzing the second step of lysine biosynthesis (11). The function of DapZ was found to be regulation of major ABC transporters. Thus the gene and the small RNA derived from the sense 3'UTR of the gene had completely different functions. Another example; CpxQ is a sRNA that is generated by RNase E cleavage of CpxP, a stress chaperone. CpxQ, an Hfq-dependent sRNA was found to repress production of several inner membrane proteins (9). However, in this case, the gene and sRNA both were transcribed from the same promoter and therefore transcribed under the same conditions.

Most of the studies that have identified and characterized sRNAs have focused on bacterial members of the Enterobacteriaceae family, mainly in *Escherichia coli.* Sixty one sRNAs have been annotated in Ecogene for *E. coli*, but many more have been predicted. Many research groups are now investigating how these newly identified sRNAs function in gene regulation. A few studies have been conducted on sRNAs in other bacterial families. Most of them focus on human pathogens, such as *Clostridium difficile, Neisseria gonorrhoeae, Acinetobacter baumannii, Pseudomonas aeruginosa, Rickettsia prowazekii, Burkholderia cenocepacia, Streptococcus pneumoniae, Listeria monocytogenes,* and *Mycobacterium tuberculosis,* to identify virulence-related sRNAs (12-20). Very few studies have focused on the Pasteurellaceae family, even though many members of this family are known to be veterinary pathogens (e.g., *Actinobacillus pleuropneumoniae, Haemophilus parasuis, Mannheimia paralytica,* and *Pasteurella multocida*), and two, *Haemophilus influenzae* and *Aggregatibacter actinomycetemcomitans,* cause disease in humans (19). A majority of the Pasteurellaceae species

are difficult to culture outside their host environment, and they are known to have relatively small genomes possibly due to adaptation to their host environment (21). Very little is known about the sequence, structure, and function of sRNAs in Pasteurellaceae, apart from a few studies in *H. influenzae* and *A. pleuropneumoniae*. HrrF was the first sRNA identified in a *Haemophilus* species. It is expressed at high levels when iron availability is low, is regulated by Fur, and is conserved in several other Pasteurellaceae species (22). Eighteen sRNAs have been identified in *H. influenzae* by RNA sequencing (RNAseq), of which HrrF has been experimentally verified (22). Of these eighteen sRNAs, seven belong to known sRNA families in the Rfam database (23). Seven sRNAs possibly interacting with the carbon storage regulatory protein CsrA were computationally identified in *Haemophilus spp.* genomes (24). Three sRNAs were also identified in *A. actinomycetemcomitans* that are regulated by iron and Fur. Most recently, *A. pleuropneumoniae* sRNAs were identified using four prediction algorithms—BLAST/Rfam, SIPHT, Infernal, and RNAz. sRNAs predicted by two or more programs were considered to be candidates for further analysis. Seventeen of the 23 sRNAs identified using this approach were confirmed by northern blotting (25).

Actinobacillus succinogenes is a non-pathogenic member of the Pasteurellaceae family. This capnophilic organism, isolated from a cow's rumen, is among the best natural succinate producers (26). We are interested in engineering *A. succinogenes* for succinate production. Although knockout methods and overexpression vectors exist, they are not always the best means to engineer the organism for industrial level production. Many a times a more fine-tuned approach is needed. Synthetic sRNAs have being designed and used to modulate gene expression in a couple of industrially relevant microorganisms (27, 28). We sought to identify sRNAs in *A.*

succinogenes and gain further information on their sequences, structure, size, and function, to then be able to design synthetic sRNAs for engineering purposes.

In this study, we carried out deep sequencing of sRNAs from *A. succinogenes* grown on two different carbon sources—glucose and glycerol. We analyzed the RNAseq data using Rockhopper and by manually reviewing the data on the Integrative Genomics Viewer (IGV). We compared the data obtained from our sequencing results to data obtained from the predictions of a few computational programs. We also sought to identify Hfq-dependent sRNAs in *A. succinogenes*, and tested them as scaffolds in the design of synthetic regulatory sRNAs.

3.2 Materials and Methods

3.2.1 Bacterial strains and growth conditions.

E. coli strains (Table 3.1) were cultivated in lysogeny broth (LB) and plates and were supplemented with 100 µg mL^{-1} ampicillin for plasmid maintenance where required. *A. succinogenes* type strain 130Z (ATCC 55618) was grown in 60-mL AM3 medium containing 50 mM glucose and 150 mM $NaHCO_3$ under a nitrogen atmosphere in 150-mL anaerobic bottles. AM3 is a phosphate-based defined medium containing vitamins and the three amino acids cysteine, methionine, and glutamate (29). Pre-cultures were grown in the same medium in 10-mL volumes in 28-mL anaerobic tubes. Anaerobic tubes and bottles were flushed with O_2-free N_2 gas (Airgas, Independence, OH) for 10 min, stoppered with rubber bungs, flushed for another 10 min and sealed with aluminum crimps. For growth on glycerol, *A. succinogenes* was grown in AM3 supplemented with 150 mM glycerol and 150 mM $NaHCO_3$ in continuous culture conditions as described (30). Cultures were quenched by mixing 25-mL of culture at an OD_{660} of ~ 0.9-1.0 with 25-mL cold methanol and stored at -20 °C until RNA purification. Two independent

biological replicates were harvested for each condition. For strains expressing the synthetic sRNA constructs grown on lactose, cultures were supplemented with 50 mM lactose and 40 μg mL^{-1} ampicillin. All strains used in this study are listed in Table 3.1.

Table 3.1 Strains and plasmids used in this study

	Description	**Source**
E. coli		
DH5α	F- φ80*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *rec*A1 *end*A1 hsdR17	Laboratory collection
TOP10	(rk-, mk+) *pho*A *sup*E44 λ- thi-1 *gyr*A96 *rel*A1 F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lacZ*ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G λ-	Invitrogen
A. succinogenes		
130Z (ATCC55618)	Wild-type strain	ATCC
Plasmids		
pCR2.1 TOPO	Amp^R, Km^R, *lacZα*, cloning vector	Invitrogen
pLGZ920	*E. coli-A. succinogenes* shuttle vector; Amp^R; *A. succinogenes* p_{pckA}	(31)
pLGZ920*-$p_{pckA\text{-}92}$-smRNA_lacZ1	pLGZ920 derivative containing synthetic sRNA smRNA_lacZ1 downstream of $p_{pckA\text{-}92}$ promoter	This study
pLGZ920*-$p_{pckA\text{-}92}$-smRNA_lacZ2	pLGZ920 derivative containing synthetic sRNA smRNA_lacZ2 downstream of $p_{pckA\text{-}92}$ promoter	This study
pLGZ920*-$p_{pckA\text{-}92}$-smRNA_lacZ3	pLGZ920 derivative containing synthetic sRNA smRNA_lacZ3 downstream of $p_{pckA\text{-}92}$ promoter	This study
pLGZ920*-$p_{pckA\text{-}92}$-smRNA_lacZ4	pLGZ920 derivative containing synthetic sRNA smRNA_lacZ4 downstream of $p_{pckA\text{-}92}$ promoter	This study
pLGZ920*-$p_{pckA\text{-}92}$-smRNA_ackA1	pLGZ920 derivative containing synthetic sRNA smRNA_ackA1 downstream of $p_{pckA\text{-}92}$ promoter	This study
pLGZ920*-$p_{pckA\text{-}92}$-smRNA_ackA2	pLGZ920 derivative containing synthetic sRNA smRNA_ackA2 downstream of $p_{pckA\text{-}92}$ promoter	This study
pLGZ920*-$p_{pckA\text{-}92}$-smRNA_pta1	pLGZ920 derivative containing synthetic sRNA smRNA_pta1 downstream of $p_{pckA\text{-}92}$ promoter	This study
pLGZ920*-$p_{pckA\text{-}92}$-smRNA_pta2	pLGZ920 derivative containing synthetic sRNA smRNA_pta2 downstream of $p_{pckA\text{-}92}$ promoter	This study

3.2.2 RNA isolation, library preparation, and sequencing.

The 50-mL bacterial suspensions in methanol were centrifuged at 4,500×*g* for 15 min at 4 °C. The cell pellets were used to isolate the total RNA, including sRNAs, using the miRNeasy mini kit (Qiagen, Valencia, CA). Contaminating DNA was removed using DNase I (Ambion, Austin, TX). The RNA was tested for chromosomal DNA contamination by PCR and was then cleaned up using Zymo RNA Clean & Concentrator-5 (Zymo Research, Irvine, CA). RNA quantity and quality were analyzed on a Qubit fluorometer (Invitrogen, Carlsbad, CA) and 2100 Bioanalyzer (Agilent Technologies; Santa Clara, CA), respectively. Total RNAs that passed the quality check were submitted to the Michigan State University Research Technology Support Facility for sequencing. Small RNA libraries were created using the Illumina small RNA Sample Prep Kit and were purified from 6% native polyacrylamide gels by excising fragments in the 22-to-500 bp range. The sRNA libraries were quality controlled by validating them using Qubit dsDNA, Caliper LabChipGX, and Kapa qPCR assays. The libraries were pooled in equimolar amounts for multiplexed sequencing, and loaded onto one lane of an Illumina HiSeq 2500 Rapid Run flow cell (v2). Sequencing was performed in a 2×100bp paired-end format using HiSeq Rapid SBS reagents. Bases were called using Illumina Real Time Analysis (RTA) v 1.18.64. The RTA output was demultiplexed and converted to FASTQ format with Illumina Bcl2fastq v1.8.4.

3.2.3 Trimming of reads, mapping, and sRNA identification.

Demultiplexed raw FASTQ files were received in paired-end format (R1 files containing left reads and R2 files containing right reads). These paired FASTQ files were trimmed to remove low-quality bases and adapters using Trimmomatic (v0.33) (32). Illumina clip was run in palindrome mode and specified to keep both reads. The sliding window was set at 4 bp, and

reads were discarded when the quality dropped below 20 or when read length was below 15 bp. FastQC analysis was done before and after trimming of the reads. Reads that passed filtering were aligned with the genome (Genbank accession no. NC_009655) using Rockhopper (33, 34) in the reference-based transcript assembly mode. The aligned reads were then analyzed with Rockhopper twice with different expression parameters. In run 1, the maximum number of bases between paired-end mates and the minimum expression of UTRs and ncRNAs were set at 300 and 0.5, respectively (referred to as 300_0.5 in text ahead). In run 2, the maximum number of bases between paired end mates was 300, and the minimum expression of UTRs and ncRNAs was 0.1 (referred to as 300_0.1 in text ahead). Small RNAs were also manually identified by visually inspecting the Rockhopper alignment files in IGV. Transcript start and stop coordinates were manually determined by finding the maximum expression level and extending the transcript start and stop until the expression level dropped below 10% of the maximum.

3.2.4 Small RNA analysis.

All sRNA sequences were input into ARNold (Rho-independent terminator prediction program) to detect rho-independent terminators. If Rockhopper-determined sRNAs were shorter than the manually-annotated transcripts in IGV, the coordinates used were those observed in IGV. The program was run under default conditions. The free energy for the positive sequences was noted. IntaRNA (35, 36), run using default parameters, was used to identify target mRNAs for the sRNAs. RNAfold (37) was used to determine sRNA secondary structures wherever needed.

3.2.5 Computational prediction of sRNAs.

Small RNAs were downloaded from the web-based platform SIPHT (available at http://newbio.cs.wisc.edu/sRNA/). SIPHT can search for intergenic loci that contain putative Rho-independent terminators in any of the bacterial replicons in the NCBI database. The parameters used were those mentioned by the authors (38). The sRNAs predicted for *A. succinogenes* were also downloaded from the Bacterial Small Regulatory RNA Database (BSRD, http://www.bac-srna.org/BSRD/taxonomyIndexNew.jsp) (39). The BSRD parameters were a maximum *E*-value of 1×10^{-15}, TransTerm confidence value of 87%, maximum RNAMotif score of -9, FindTerm score of -10, and minimum and maximum lengths for predicted loci of 50 and 500, respectively.

INFERNAL (40), RNAz (41), and Blast against Rfam (blasts the genome against previously described sRNAs)(42)—all available through the platform RNAspace (www.rnaspace.org/) (43) were also run. INFERNAL, which searches for homologs of structural RNAs in sequence databases like the Rfam 10.0 database by building covariance models (CM), was run with the default parameters. RNAz is a program that can detect structurally conserved, functional, thermodynamically stable RNA in genome-wide screens or multiple alignments. RNAz was run using the default parameters, with a probability cutoff of 0.7, slice alignments longer than 300, window size of 200, and stepsize of 50. RNAz used the annotated genomes of *A. pleuropneumoniae* L20 and *H. influenzae* Rd KW20 for alignment with the *A. succinogenes* genome, using BLAST with default parameters in RNAspace. In the same RNAspace run, sequence aggregation was also done using CG-seq using default parameters (i.e., score lambda parameter of 1, minimal and maximal lengths of a conserved region of 30 and 500, minimum and

maximum identity thresholds of 60 and 100, respectively). CRISPRs were identified in the genome using CRISPRFinder (44) with default parameters.

3.2.6 RT-PCR validation of small RNAs.

First-strand cDNA was synthesized from 1 µg of total RNA with random hexamers and SuperScript II reverse transcriptase (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. The reverse transcriptase reaction was used as the template for PCR reactions using primers designed for each sRNA candidate to be verified. As a positive control, *A. succinogenes* genomic DNA was used as the template for PCR. To verify the absence of genomic DNA, DNase-treated total RNA (not subjected to reverse transcriptase) was used as the template. PCR reactions were done in a 20-µL volume containing Taq polymerase, 1X Taq buffer, 50 µM of each dNTP, 1.5 mM $MgCl_2$, and 1.25 µM of each primer. Thermocycler conditions were as follows: 95 °C for 3 min; 32 cycles of 95 °C for 30 sec, 60 °C for 30 sec and 72 °C for 1 min; 72 °C for 5 min. PCR products were visualized on a 3% agarose gel with a low molecular weight ladder (New England Biolabs, Ipswich, MA). All sRNAs were tested using cDNA synthesized from total RNA from glucose-grown cultures, except for smRNAs 47 and 126, which were tested with cDNA synthesized from RNA obtained from glucose and glycerol cultures.

3.2.7 Construction of synthetic small RNA constructs.

Synthetic sRNA constructs designed to inhibit *lacZ* expression—smRNA_lacZ1, smRNA_lacZ2, smRNA_lacZ3, and smRNA_lacZ4, were constructed by annealing oligonucleotides CV662 and CV663, CV664 and CV665, CV666 and CV667, and CV668 and CV669 respectively (Table S3.1). Annealed oligonucleotides were used as templates and

amplified using primers CV677 and CV678 (smRNA_lacZ1), and primers CV677 and CV 679 (smRNA_lacZ2, smRNA_lacZ3, and smRNA_lacZ4). PCR products were cloned into pLGZ920*-p_{Asuc_0701}-*lacZ's Bam*HI and *Sac*I sites using Gibson cloning (New England Biolabs, Ipswich, MA). pLGZ920*- p_{Asuc_0701}-*lacZ* is a derivative of pLGZ920 expressing *lacZ*, in which the *pckA* promoter was replaced by the promoter of *Asuc_0701*, and the Hind*III* restriction site has been removed (see chapter 2). Digesting pLGZ920*- p_{Asuc_0701}-*lacZ* with *Bam*HI and *Sac*I removes p_{Asuc_0701} and *lacZ*. Ligation mixtures were transformed into *E. coli* Top10 (Invitrogen) chemically competent cells and plated on LB agar plates containing 100 µg mL^{-1} ampicillin. Transformant colonies were tested for the presence of the insert by colony PCR with primers CV677 and CV640, followed by plasmid sequencing with primer CV640 (GENEWIZ, Inc., South Plainfield, NJ). Positive clones were transformed into *A. succinogenes* strain 130Z by electroporation as described (31).

Synthetic sRNA constructs targeting acetate kinase (AckA) and phosphotransacetylase (Pta) expression—smRNA_ackA1, smRNA_ackA2, smRNA_pta1, smRNA_pta2, were ordered as g-blocks from Integrated DNA Technologies (Coralville, IA) (Table S3.1). All g-blocks were digested with *Bam*HI and *Xba*I and ligated into pLGZ920*-p_{Asuc_0701}*'s Bam*HI and *Xba*I sites as described above. Colonies were screened by colony PCR using primers CV677 and CV640, and plasmid inserts were verified by sequencing (GENEWIZ, Inc.). Positive clones were transformed into *A. succinogenes* 130Z.

3.2.8 Beta-galactosidase assays.

Strain 130Z carrying constructs smRNA_lacZ1, smRNA_lacZ2, and smRNA_lacZ4 was grown in AM3 supplemented with 50 mM lactose, 150 mM $NaHCO_3$, and 40 µg mL^{-1}

ampicillin, and harvested in mid-exponential phase. The positive control was strain 130Z (pLGZ290). Cultures (200 µL) were transferred to 96-well plates and optical density at 600 nm (OD_{600}) was recorded on a PowerWave HT microplate reader (BioTek Instruments Inc., Winooski, VT). OD_{600} values were used as a measure of cell biomass. One hundred µL of cultures were transferred to a new 96-well plate and lysed by adding 100 µL of Bugbuster Protein Extraction Reagent (Millipore Sigma, Billerica, MA) and incubating for 5 min. The lysate was used as the crude extract for β-galactosidase assays. β-galactosidase assays were conducted in 96-well plates in the PowerWave HT microplate reader. Reactions (200 µL) contained 102.5 µL cell extract mixed with 67.5 µL of Z-buffer (60 mM Na_2HPO_4 , 40 mM NaH_2PO_4 , 10 mM KCl, 1 mM $MgSO_4$, 38 mM 2-mercaptoethanol, pH 7.0) and 30 µL of 2-nitrophenyl β-D-galactopyranoside (4 mg mL^{-1} in Z-buffer). Enzyme activity was calculated using the linear slope obtained from 2-nitrophenol production recorded at 420 nm for 20 min at room temperature. Specific activity values are reported in ΔA_{420} min^{-1} ml^{-1} OD_{600}^{-1} and are the average of three independent biological replicates.

3.2.9 Determination of fermentation balances.

Growth of *A. succinogenes* in liquid cultures was monitored by measuring OD_{660} on a DU 650 spectrophotometer (Beckman, Fullerton, CA). Samples collected early (OD_{660} ~ 0.5) and late (OD_{660} ~ 1.5) in the exponential phase were used to determine the carbon balance. Glucose and organic acids were quantified in the 10-fold diluted, filtered culture supernatants by HPLC (Waters, Milford, MA) using an Aminex HPX-87H column (Bio-rad). Samples were run with 4 mM H_2SO_4 as the eluent at a 0.6 mL min^{-1} flow rate at room temperature. Glucose and ethanol were quantified on a Waters 410 differential refractometer while organic acids were quantified

on a Waters 2487 UV detector at 210 nm. OD_{660} values were used to calculate the biomass as described (45).

3.3 Results and Discussion

3.3.1 Detection of sRNAs from RNAseq data using RockHopper.

The overall approach we took for identifying sRNAs is described in Figure 3.1. *A. succinogenes* sRNAs were sequenced in cultures grown anaerobically on glucose and microaerobically on glycerol, with two independent biological replicates for each condition. Over 95% of the read pairs remained after trimming and quality control, the details of which are mentioned in Table 3.2.

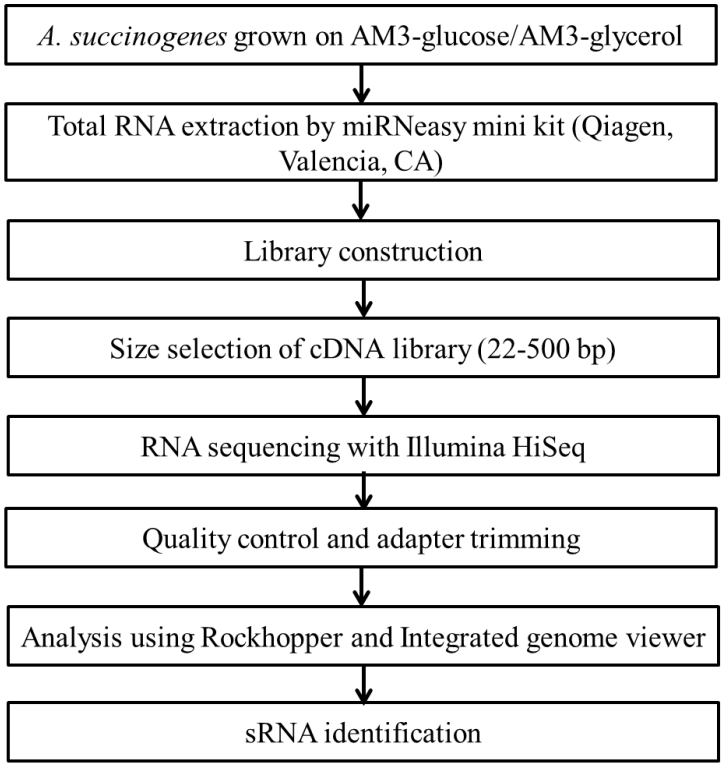

Figure 3.1 Flow chart of the approach used for sRNA identification in *A. succinogenes*.

Table 3.2 Statistics on read alignments by Rockhopper

Sample	No of read pairs before trimming	Read pairs remaining after trimming (% surviving)	Successful alignment by Rockhopper (% alignment)	Alignment to protein coding genes (sense) (%)	Alignment to protein coding genes (antisense) (%)	Alignment to rRNAs (%)	Alignment to tRNAs (%)	Alignment to unannotated regions (%)
Glucose-1	13,066,506	12,495,687 (95.6%)	85	16	1	65	5	13
Glucose-2	12,423,762	11,809,343 (95.0%)	85	16	1	66	5	12
Glycerol-1	10,842,533	10,452,076 (96.4%)	86	9	2	76	5	7
Glycerol-2	10,637,886	10,236,278 (96.2%)	85	11	2	73	5	10

The Rockhopper-predicted sRNAs were manually inspected by visualization in IGV. Three types of Rockhopper-predicted sRNAs were discarded or modified: i) sRNAs in sense direction located in 3'UTR or 5'UTR region of annotated mRNA, whose transcript levels did not differ much from the corresponding mRNA (sense) transcript levels were deleted; ii) sRNAs that were shorter than 30 nt were deleted, unless they appeared to be longer in IGV; and iii) two or more separate RNAs that were immediately consecutive but that, upon visual inspection, seemed to belong to one bigger sRNA transcript, were listed as a unique, larger transcript.

After analysis of the Rockhopper results (from 300_0.5 and 300_0.1), a total of 145 sRNAs were detected in glucose and glycerol cultures combined (Table 3. 3). Of these 145 sRNAs, five were found in glucose-grown cultures only, while 40 were found in glycerol-grown cultures only (Table 3.3). All others were found in both culture conditions. An additional 115 more sRNAs that had been missed by Rockhopper were found upon manual inspection in IGV (Table S3.1). In total, these 260 sRNAs were classified into six separate categories, as described in Figure 3.2.

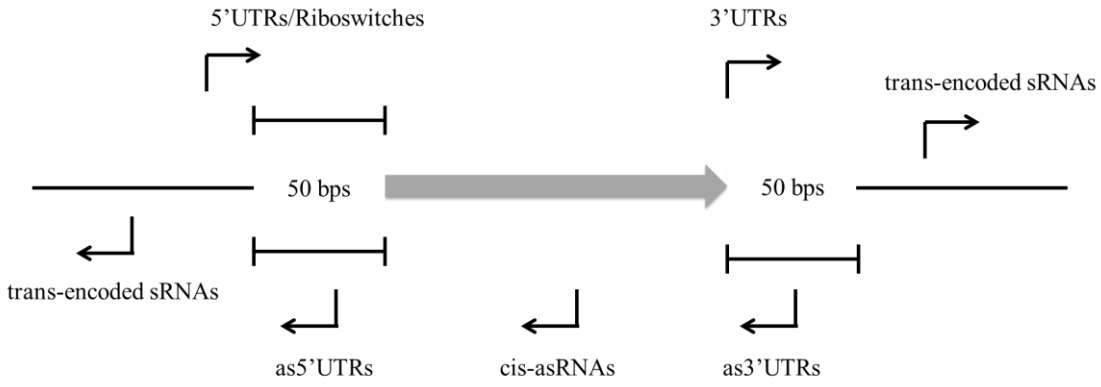

Figure 3.2 Classification of identified sRNAs.

Table 3.3 *A. succinogenes* sRNAs detected by RNAseq and Rockhopper analysis in anaerobically-grown glucose cultures and microaerobically-grown glycerol cultures

Rockhopper parameters; max bases between paired end mates=300 ; minimum expression of UTRs and nc RNAS=0.5

Candidates	Rfam Annotation	Transcription start	Transcription stop	Size	Strand	Neighborhood	Predicted by algorithms
smRNA1		134,044	134,063	19	+	ASUC_RS00615/ASUC_RS00620	
smRNA2		368,282	368,340	58	+	ASUC_RS01695/ASUC_RS01700	
smRNA3		369,316	369,371	55	+	ASUC_RS01700/ASUC_RS01705	
smRNA4	PyrR	431,715	431,814	99	-	ASUC_RS01960/ASUC_RS01965	I
smRNA5		517,122	517,162	40	+	ASUC_RS02365/ASUC_RS02370	Z
smRNA6		553,139	553,170	31	-	ASUC_RS02590/ASUC_RS02595	
smRNA7		609,640	609,664	24	+	ASUC_RS02865/ASUC_RS02870	
smRNA8		744,502	744,593	91	-	ASUC_RS03550/ASUC_RS03555	
smRNA9		785,724	785,705	19	-	ASUC_RS03760/ASUC_RS03765	Z
smRNA10		837,968	838,075	107	-	ASUC_RS04020/ASUC_RS04025	
smRNA11	MicF	981,582	981,628	46	+	ASUC_RS04740/ASUC_RS04745	I
smRNA12		1,000,490	1,000,464	26	-	ASUC_RS04845/ASUC_RS04850	
smRNA13		1,013,544	1,013,530	14	-	ASUC_RS04900/ASUC_RS04905	
smRNA14**		1,044,143	1,044,177	34	-	ASUC_RS05040/ASUC_RS05045	
smRNA15		1,045,717	1,045,764	47	-	ASUC_RS05045/ASUC_RS05050	
smRNA16		1,048,744	1,048,726	18	-	ASUC_RS05055/ASUC_RS05060	
smRNA17		1,136,086	1,136,035	51	-	ASUC_RS05410/ASUC_RS05415	
smRNA18	SRP_bact	1,215,624	1,215,735	111	+	ASUC_RS05760/ASUC_RS05765	B, I, Z, M
smRNA19		1,237,630	1,237,591	39	-	ASUC_RS05855/ASUC_RS0560	
smRNA20		1,276,490	1,276,511	21	+	ASUC_RS06025/ASUC_RS06030	Z
smRNA21		1,276,603	1,276,695	92	+	ASUC_RS06025/ASUC_RS06030	Z
smRNA22		1,367,934	1,367,964	30	+	ASUC_RS06370/ASUC_RS06375	
smRNA23		1,376,784	1,376,856	72	+	ASUC_RS06445/ASUC_RS11015	Z
smRNA24		1,441,632	1,441,660	28	+	ASUC_RS06770/ASUC_RS06775	S
smRNA25		1,525,386	1,525,407	21	-	ASUC_RS07160/ASUC_RS07165	
smRNA26		1,545,573	1,545,663	90	+	ASUC_RS07260/ASUC_RS07265	Z
smRNA27	GcvB	1,862,491	1,862,517	26	+	ASUC_RS08780/ASUC_RS08785	B, I, M

Table 3.3 (cont'd)

smRNA28		2,003,880	2,003,940	60	+	ASUC_RS09455/ASUC_RS09460	S
smRNA29		2,089,221	2,089,240	19	+	ASUC_RS09840/ASUC_RS09845	Z
smRNA30		2,204,631	2,204,672	41	-	ASUC_RS10330/ASUC_RS10335	
smRNA31		2,275,657	2,275,690	33	-	ASUC_RS10685/ASUC_RS10690	
Rockhopper parameters; max bases between paired end mates=300 ; minimum expression of UTRs and nc RNAS=0.1							
smRNA32		22	81	59	-	ASUC_RS10935/ASUC_RS00005	
smRNA33*		16,626	16,657	31	-	ASUC_RS00085/ASUC_RS00090	
smRNA34*		16,749	16,849	100	-	ASUC_RS00085/ASUC_RS00090	
smRNA35*		17,297	17,329	32	-	ASUC_RS00090/ASUC_RS00095	
smRNA36		44,059	44,089	30	+	ASUC_RS00210/ASUC_RS00215	
smRNA37		74,727	74,758	31	+	ASUC_RS00320/ASUC_RS00325	
smRNA38*		91,637	91,689	52	-	ASUC_RS00400/ASUC_RS00405	
smRNA39		132,913	132,945	32	-	ASUC_RS00610/ASUC_RS00615	
smRNA40		134,004	134,071	67	-	ASUC_RS00615/ASUC_RS00620	
smRNA41		278,628	278,675	47	-	ASUC_RS01270/ASUC_RS01275	
smRNA42		289,411	289,455	44	-	ASUC_RS01335/ASUC_RS01340	
smRNA43		313,893	314,010	117	+	ASUC_RS01445/ASUC_RS01450	
smRNA44*		330,475	330,511	36	-	ASUC_RS01530/ASUC_RS01535	
smRNA45		330,852	330,883	31	-	ASUC_RS01535/ASUC_RS01540	
smRNA46**		353,040	353,070	30	+	ASUC_RS01650/ASUC_RS01655	
smRNA47**		356,657	356,723	66	-	ASUC_RS01655/ASUC_RS01660	
smRNA48**		359,036	359,095	59	-	ASUC_RS01660/ASUC_RS01665	
smRNA49		412,230	412,339	109	+	ASUC_RS01890/ASUC_RS01895	
smRNA50		501,364	501,412	48	+	ASUC_RS02295/ASUC_RS02300	
smRNA51		508,217	508,265	48	-	ASUC_RS02325/ASUC_RS02330	
smRNA52		520,787	520,823	36	-	ASUC_RS02400/ASUC_RS02405	
smRNA53		522,418	522,483	65	-	ASUC_RS02420/ASUC_RS02425	
smRNA54		525,832	525,866	34	-	ASUC_RS02460/ASUC_RS02465	
smRNA55	Alpha_RBS	527,930	528,046	116	+	ASUC_RS02475/ASUC_RS02480	S, B, I, Z, M
smRNA56		596,463	596,584	121	-	ASUC_RS02810/ASUC_RS02825	
smRNA57	Parecho_CRE	599,496	599,569	73	+	ASUC_RS02835/ASUC_RS02840	I

Table 3.3 (cont'd)

smRNA58*		604,037	604,135	98	-	ASUC_RS02845/ASUC_RS02850	
smRNA59*		604,515	604,583	68	+	ASUC_RS02845/ASUC_RS02850	
smRNA60		620,650	620,681	31	-	ASUC_RS02915/ASUC_RS02920	
smRNA61*		632,299	632,398	99	+	ASUC_RS02980/ASUC_RS02985	
smRNA62*		632,547	632,583	36	+	ASUC_RS02980/ASUC_RS02985	
smRNA63		643,335	643,371	36	+	ASUC_RS03035/ASUC_RS03040	
smRNA64		643,644	643,675	31	+	ASUC_RS03035/ASUC_RS03040	
smRNA65**		656,185	656,204	19	+	ASUC_RS03080/ASUC_RS03085	
smRNA66*		657,344	657,404	60	-	ASUC_RS03085/ASUC_RS03090	
smRNA67		694,517	694,526	9	+	ASUC_RS03285/ASUC_RS03290	
smRNA68		705,794	705,810	16	+	ASUC_RS03340/ASUC_RS03345	
smRNA69*		743,678	743,708	30	-	ASUC_RS03545/ASUC_RS03550	
smRNA70*		744,115	744,089	26	-	ASUC_RS03545/ASUC_RS03550	
smRNA71*		744,307	744,380	73	-	ASUC_RS03550/ASUC_RS03555	
smRNA72*		744,450	744,438	12	-	ASUC_RS03550/ASUC_RS03555	
smRNA73		769,202	769,245	43	+	ASUC_RS03670/ASUC_RS03680	
smRNA74*		785,488	785,512	24	+	ASUC_RS03755/ASUC_RS03760	
smRNA75		787,773	787,816	43	+	ASUC_RS03770/ASUC_RS03775	
smRNA76	Glycine riboswitch	806,342	806,566	224	+	ASUC_RS03870/ASUC_RS03875	B, I
smRNA77	Thr_leader	858,732	858,775	43	+	ASUC_RS04140/ASUC_RS04145	M
smRNA78	Parecho_CRE	869,489	869,546	57	+	ASUC_RS04185/ASUC_RS04190	I
smRNA79		871,031	871,049	18	+	ASUC_RS04190/ASUC_RS04195	
smRNA80		883,325	883,345	20	+	ASUC_RS04250/ASUC_RS04255	
smRNA81		906,434	906,465	31	-	ASUC_RS04380/ASUC_RS04385	
smRNA82*		924,096	924,135	39	+	ASUC_RS04460/ASUC_RS04465	
smRNA83*		926,654	926,690	36	+	ASUC_RS04465/ASUC_RS04470	
smRNA84		936,865	936,906	41	+	ASUC_RS04530/ASUC_RS04535	
smRNA85		938,190	938,323	133	+	ASUC_RS04535/ASUC_RS04550	
smRNA86*		1,012,701	1,012,743	42	-	ASUC_RS04895/ASUC_RS04900	
smRNA87		1,014,494	1,014,571	77	+	ASUC_RS04905/ASUC_RS04910	

Table 3.3 (cont'd)

smRNA88	1,057,772	1,057,841	69	-	ASUC_RS05080/ASUC_RS05085	S
smRNA89*	1,067,837	1,067,870	33	+	ASUC_RS05120/ASUC_RS05125	
smRNA90*	1,069,547	1,069,772	225	+	ASUC_RS05125/ASUC_RS05130	
smRNA91	1,073,915	1,073,974	59	-	ASUC_RS05155/ASUC_RS05160	
smRNA92	1,083,430	1,083,410	20	-	ASUC_RS05215/ASUC_RS05220	
smRNA93	1,117,501	1,117,539	38	-	ASUC_RS05330/ASUC_RS05335	
smRNA94	1,119,016	1,119,057	41	+	ASUC_RS05330/ASUC_RS05340	
smRNA95*	1,146,049	1,146,091	42	+	ASUC_RS05450/ASUC_RS05455	
smRNA96	1,165,100	1,165,133	33	-	ASUC_RS05545/ASUC_RS05550	
smRNA97*	1,240,529	1,240,560	31	-	ASUC_RS05865/ASUC_RS05870	
smRNA98	1,253,225	1,253,175	50	-	ASUC_RS05925/ASUC_RS05930	
smRNA99*	1,250,832	1,250,866	34	+	ASUC_RS05915/ASUC_RS05920	
smRNA100*	1,252,103	1,252,198	95	+	ASUC_RS05920/ASUC_RS05925	
smRNA101*	1,252,811	1,252,842	31	+	ASUC_RS05925/ASUC_RS05930	
smRNA102*	1,264,237	1,264,348	111	-	ASUC_RS05980/ASUC_RS05985	
smRNA103	1,265,927	1,265,963	36	+	ASUC_RS05990/ASUC_RS05995	
smRNA104*	1,266,017	1,266,073	56	+	ASUC_RS05990/ASUC_RS05995	
smRNA105	1,266,795	1,266,826	31	+	ASUC_RS05990/ASUC_RS05995	
smRNA106*	1,267,226	1,267,256	30	+	ASUC_RS05990/ASUC_RS05995	
smRNA107*	1,267,779	1,267,810	31	+	ASUC_RS05995/ASUC_RS06000	
smRNA108	1,296,921	1,296,999	78	-	ASUC_RS06085/ASUC_RS06090	
smRNA109	1,298,854	1,298,937	83	-	ASUC_RS06085/ASUC_RS06090	
smRNA110	1,300,630	1,300,662	32	-	ASUC_RS06085/ASUC_RS06090	
smRNA111*	1,368,446	1,368,552	106	+	ASUC_RS06370/ASUC_RS06375	
smRNA112	1,370,455	1,370,499	44	-	ASUC_RS06390/ASUC_RS06395	
smRNA113	1,376,675	1,376,702	27	+	ASUC_RS06445/ASUC_RS11015	
smRNA114	1,377,939	1,377,967	28	+	ASUC_RS11015/ASUC_RS06455	
smRNA115	1,400,489	1,400,682	193	-	ASUC_RS06585/ASUC_RS06590	
smRNA116*	1,415,979	1,416,016	37	+	ASUC_RS06650/ASUC_RS06655	
smRNA117	1,424,398	1,424,495	97	-	ASUC_RS06685/ASUC_RS06690	
smRNA118	1,455,834	1,455,884	50	+	ASUC_RS06840/ASUC_RS06845	

Table 3.3 (cont'd)

smRNA119*	1,493,409	1,493,453	44	+	ASUC_RS07025/ASUC_RS07030	
smRNA120	1,499,794	1,499,829	35	-	ASUC_RS07050/ASUC_RS07055	
smRNA121	1,528,991	1,528,963	28	-	ASUC_RS07175/ASUC_RS07180	
smRNA122	1,543,327	1,543,374	47	+	ASUC_RS07245/ASUC_RS07250	
smRNA123	1,586,321	1,586,358	37	-	ASUC_RS07425/ASUC_RS07430	
smRNA124	1,623,706	1,623,739	33	+	ASUC_RS07620/ASUC_RS07625	
smRNA125	1,671,194	1,671,207	13	+	ASUC_RS07840/ASUC_RS07845	
smRNA126*	1,677,571	1,677,675	104	-	ASUC_RS07870/ASUC_RS07875	
smRNA127	1,700,597	1,700,770	173	-	ASUC_RS07990/ASUC_RS07995	
smRNA128	1,737,609	1,737,650	41	-	ASUC_RS08165/ASUC_RS08170	
smRNA129*	1,901,859	1,901,893	34	-	ASUC_RS08940/ASUC_RS08945	
smRNA130*	1,902,651	1,902,682	31	-	ASUC_RS08945/ASUC_RS08950	
smRNA131*	1,927,500	1,927,538	38	-	ASUC_RS09070/ASUC_RS09075	
smRNA132	1,974,287	1,974,272	15	-	ASUC_RS09300/ASUC_RS09305	
smRNA133	1,977,589	1,977,719	130	-	ASUC_RS09310/ASUC_RS09315	
smRNA134*	1,978,600	1,978,637	37	+	ASUC_RS09320/ASUC_RS09325	
smRNA135	1,981,750	1,981,830	80	-	ASUC_RS09340/ASUC_RS09345	Z
smRNA136	1,989,202	1,989,252	50	-	ASUC_RS09380/ASUC_RS09385	
smRNA137	2,099,513	2,099,588	75	+	ASUC_RS09885/ASUC_RS09890	
smRNA138	2,202,223	2,202,295	72	+	ASUC_RS10320/ASUC_RS10325	
smRNA139	2,220,187	2,220,167	20	-	ASUC_RS10410/ASUC_RS10415	
smRNA140	2,266,375	2,266,447	72	-	ASUC_RS10650/ASUC_RS10655	
smRNA141	2,291,971	2,291,996	25	+	ASUC_RS10785/ASUC_RS10790	
smRNA142*	2,300,851	2,300,963	112	+	ASUC_RS10825/ASUC_RS10830	
smRNA143*	2,307,624	2,307,693	69	-	ASUC_RS10870/ASUC_RS10875	
smRNA144	2,314,470	2,314,449	21	-	ASUC_RS10900/ASUC_RS10905	
smRNA145*	2,315,613	2,315,645	32	-	ASUC_RS10910/ASUC_RS10915	

**smRNA expressed in glucose cultures only
* smRNA expressed in glycerol cultures only.
A minimum cutoff of 50 for transcript level was used to determine whether or not a sRNA was expressed.
B, BSRD; I, INFERNAL; M, Rfam; S, SIPHT; Z, RNAz

5’UTRs were defined as transcripts whose 3’ end preceded a coding sequence by fewer than 50 nt. In some cases riboswitch transcripts did not end until well within the coding sequence. For ease of classification, these riboswitches were still grouped in the 5’UTR category. 3’UTRs were defined as transcripts whose 5’ end started within the coding region or immediately after the stop codon. Many transcripts started well within the coding sequence (sense direction) and extended beyond the coding sequence, usually accompanied by lower transcript levels of the gene itself (Figure 3.3). Antisense sRNAs were divided into three categories: as5’UTRs, cis-asRNAs, and as3’UTRs. Transcripts antisense to the 5’UTR of the coding sequence that started within 50 nt of the start codon were called as5’UTRs. Cis-asRNAs were defined as antisense and started within the coding sequence.

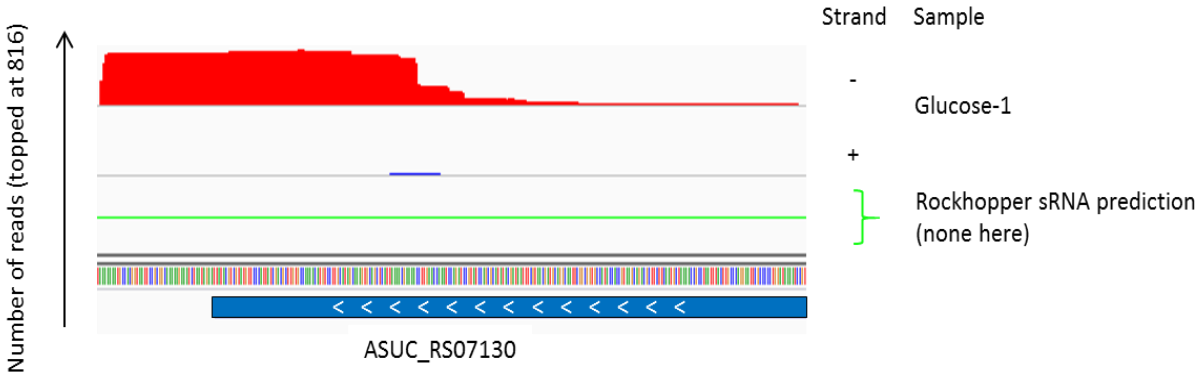

Figure 3.3 Example of 3’UTR sRNA overlapping with the mRNA transcript and downstream intergenic region with both mRNA and sRNA encoded on the reverse strand. -, reverse strand; +, forward strand; red: reads on the reverse strand; blue: reads on the forward strand; white arrowheads in gene indicate direction of the gene. Figure modified from screenshot in IGV.

Transcripts antisense to the coding sequence that started within 50 nt of the stop codon were called as3’UTRs. All others were called trans-encoded sRNAs. Using this classification,

the 260 sRNAs were catalogued as 28 5'UTRs, 71 3'UTRs, 2 as5'UTRs, 81 asRNAs, 10 as3'UTRs, and 68 trans-encoded sRNAs.

3.3.2 Comparison of RNAseq-identified to computationally predicted sRNAs.

SIPHT predicted 45 sRNAs, of which 40 were unknown sRNAs and five were known sRNAs or riboswitches annotated in Rfam. Eleven of the SIPHT-predicted sRNAs were also identified by sRNAseq, but some of them were combined since they looked like a single transcript in IGV, giving a total of six sRNAs shared between the two—smRNA24, smRNA28, smRNA55, smRNA88, smRNA175, and smRNA251 (Table S3.2). SIPHT predicted two sRNAs that were already annotated in *A. succinogenes* as rRNAs or tRNAs—asuc_R0067 and asuc_RS0071 (NCBI old locus tags). These were omitted in our analysis. For smRNA24, SIPHT predicted nine different transcripts, five of which started from the same coordinate but had different stop coordinates. The other four all had different start coordinates but shared the same stop coordinate. Because smRNA24 corresponded to a continuous transcript in IGV, it was considered to be a single transcript. SIPHT also predicted smRNA175 as two separate transcripts, but since IGV showed it as a continuous transcript, it was deemed to be one sRNA. Two of the sRNAs predicted by SIPHT and detected in RNAseq are already annotated in Rfam as Alpha_RBS (smRNA55) and CRISPR-DR34 (smRNA175) while the other four are unknown.

BSRD (Bacterial Small Regulatory RNA Database) lists eighteen sRNA candidates for *A. succinogenes* 130Z, all of which are orthologs of known sRNAs in other bacterial species. Twelve of these were detected in our RNAseq study, of which three (not listed in Table S3.3) are already annotated in *A. succinogenes*: asuc_R0067 (NCBI old locus tag) (annotated as tmRNA in Rfam), ASUC_RS11040 (annotated as Rnase P classA in Rfam), and ASUC_RS09385

(annotated as S15 in Rfam). These three sRNAs were not considered for further analysis. The remaining nine are listed in Table S3.3. BSRD lists two separate sRNAs spanning coordinates 806377–806453 and 806464–806566. Rockhopper predicted a single sRNA instead (smRNA76), with coordinates 806342–806566. smRNA76 was considered to be one single sRNA after visual confirmation in IGV (Figure 3.4).

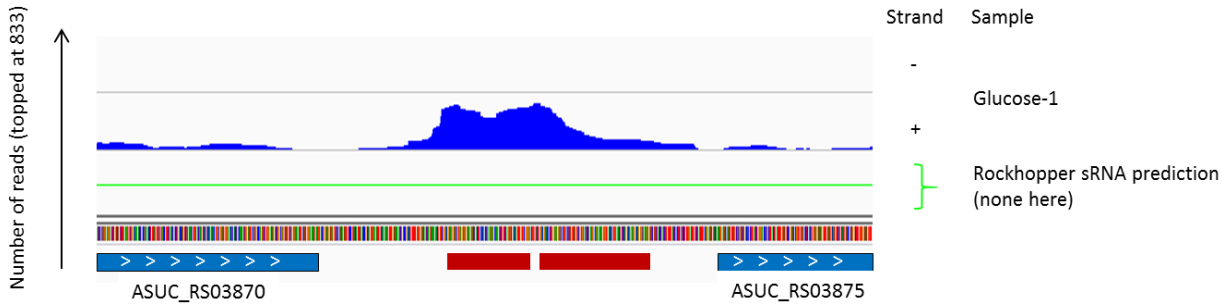

Figure 3.4 smRNA76 is seen as a single transcript in IGV. +, forward strand; -, reverse strand; blue: reads detected in RNAseq; green: Rockhopper-predicted sRNA in both glucose and glycerol libraries; red bars, two transcripts predicted by BSRD. Figure modified from screenshot in IGV.

INFERNAL, RNAz, and Blastn against Rfam predicted 167, 108, and 115 putative non coding RNAs (including tRNAs and rRNAs) respectively. Fifteen were common between INFERNAL and RNA-seq, all of which are annotated in Rfam (Table S3.4). INFERNAL predicted two separate transcripts for smRNA76, which we described above as a single transcript (Figure 3.4 and Table S3.4).

Twenty sRNAs were common between RNAz and RNA-seq, three of which are annotated in Rfam (Table S3.5). RNAseq and Blastn against Rfam had ten sRNAs in common, with annotations listed in Table S3.6. CRISPRFinder predicted two CRISPR elements in the *A.*

succinogenes genome, which were both detected in RNAseq. One of the CRISPR elements is annotated in Rfam as CRISPR –DR34 (smRNA175) (predicted as 7 separate trascripts in Rfam spanning coordinates 436,472 and 436,903), while the other is unannotated (smRNA24). smRNA24 had previously been described in our *A. succinogenes* genome paper (46) using CRISPRs web service. Overall, thirteen of small RNAs detected in RNAseq were also detected by at least one of the computational programs. Twenty-nine of the sRNAs detected by RNAseq were predicted by at least one of the computational programs, three were predicted by two computational programs, four were predicted by three of the programs, and two were predicted by four programs, while only one was predicted by all five computational programs.

3.3.3 sRNAs with Rho-independent terminators.

Many sRNAs depend on Hfq, a chaperone protein, to bind to their target mRNAs and for stability (4). Hfq typically binds to the U-rich region immediately preceding the poly(U) tail found in RNAs with Rho-independent terminators. Thus sRNAs with Rho-independent terminators are possibly Hfq-dependent. Twenty four Rho-independent terminators were identified among the 260 sRNAs detected in RNAseq (Rockhopper and manually identified) using ARNold (Table 3.4).

3.3.4 RT-PCR validation of small RNAs.

We selected fourteen sRNAs for validation with RT-PCR—smRNA4, 8, 11, 18, 21, 27, 28, 47, 49, 55, 76, 126, 135, and 187. smRNA4 was selected because it was predicted by a single computational algorithm, INFERNAL. smRNA8 and smRNA27 were not predicted by any of the computational algorithms. smRNA11 was selected because it is annotated as MicF in Rfam and was also predicted by INFERNAL. smRNA18 was selected because it was antisense to ffs

Table 3.4 Rho-independent smRNAs predicted by ARNold

sRNA candidate	Transcription start[a]	Transcription stop[a]	Size	Strand	Neighborhood	Free energy[b] (kcal/mol)
smRNA6	553,132	553,188	56	-	ASUC_RS02590/ASUC_RS02595	-5.06
smRNA8	744,498	744,593	95	-	ASUC_RS03550/ASUC_RS03555	-4.80
smRNA11	981,582	981,635	53	+	ASUC_RS04740/ASUC_RS04745	-12.7
smRNA13	1,013,483	1,013,544	61	-	ASUC_RS04900/ASUC_RS04905	-11.1
smRNA15	1,045,695	1,045,765	70	-	ASUC_RS05045/ASUC_RS05050	-8.45
smRNA16	1,048,685	1,048,748	63	-	ASUC_RS05055/ASUC_RS05060	-12.26
smRNA17	1,135,901	1,136,086	185	-	ASUC_RS05410/ASUC_RS05415	-11.2
smRNA25	1,525,370	1,525,415	45	-	ASUC_RS07160/ASUC_RS07165	-12.3
smRNA27	1,862,486	1,862,541	55	+	ASUC_RS08780/ASUC_RS08785	-12.86
smRNA28	2,003,880	2,003,947	67	+	ASUC_RS09455/ASUC_RS09460	-8.8
smRNA31	2,275,649	2,275,708	59	-	ASUC_RS10685/ASUC_RS10690	-10.67
smRNA77	858,732	858,865	133	+	ASUC_RS04140/ASUC_RS04145	-13.1
smRNA79	871,031	871,079	48	+	ASUC_RS04190/ASUC_RS04195	-9.3
smRNA88	1,057,768	1,057,841	73	-	ASUC_RS05080/ASUC_RS05085	-11.1
smRNA113	1,376,675	1,376,719	44	+	ASUC_RS06445/ASUC_RS11015	-3.4
smRNA132	1,974,240	1,974,280	40	-	ASUC_RS09300/ASUC_RS09305	-9.2
smRNA153	102,498	102,589	91	+	ASUC_RS003455/ASUC_RS00460	-11.64
smRNA175[c]	436,421	436,943	522	-	ASUC_RS01975/ASUC_RS01980	-10.4
smRNA185	691,158	691,339	181	+	ASUC_RS03270/ASUC_RS03275	-8.99
smRNA207	1,169,658	1,169,834	176	+	ASUC_RS05560/ASUC_RS05565	-11.1
smRNA209	1,185,693	1,185,896	203	-	ASUC_RS05625/ASUC_RS05630	-11.2
smRNA227	1,520,160	1,520,316	156	-	ASUC_RS07125/ASUC_RS07130	-14.2
smRNA238	1,724,668	1,724,828	160	-	ASUC_RS08120/ASUC_RS08125	-10.3
smRNA240	1,825,558	1,825,646	88	-	ASUC_RS08605/ASUC_RS08610	-7.5

[a] sRNA sizes predicted by Rockhopper were often shorter than the transcripts seen in IGV and did not include the rho-independent terminators. Rockhopper coordinates were extended at the 3'end until they included the rho-independent terminator, but only until the transcript ends seen in IGV. These extended corodinates were used as input in ARNold and are listed in this table.

[b] Free energy of the predicted terminator stem-loop structure using RNAfold (47), can be used as a confidence value for predicted terminators.

[c] Rho-independent terminator is also predicted about 265 nts into the transcript of smRNA175, which is annotated as CRISPR-DR34 with a free energy of stem loop region of -6.10 kcal/mol.

(encoding the signal recognition particle 4.5S RNA), was predicted by most algorithms, and is annotated as SRP_bact in Rfam (Table 3.2). smRNA21 and smRNA137 were only predicted by RNAz. smRNA27 is annotated as GcvB in Rfam and was selected because it was identified as a

regulatory ncRNA in *E.coli* and other enteric bacteria. It regulates several protein-encoding genes involved in metabolic pathways (48). smRNA47 was detected only in the glucose-grown cultures in our RNAseq analysis (any sRNA expression level below 50 was considered as not expressed). smRNA49 was selected as it was classified as a 3'UTR sRNA in our analysis. smRNA55 was classified as a 5'UTR and annotated in Rfam as Alpha_RBS. smRNA76 was selected because BSRD predicts it as 2 separate transcripts, but it looks like a single transcript on IGV (Figure 3.3). smRNA126 was classified as an as3'UTR and was detected only in the glycerol cultures. smRNA187 was selected as it is annotated as 6S RNA in Rfam.

All sRNAs tested by RT-PCR showed a PCR product of the expected size (Figure 3.5). smRNA47 was tested with cDNA synthesized from RNA extracted from both growth conditions. Even though smRNA47 was not detected in our RNAseq data of glycerol cultures, it produced a PCR band in both conditions. This may be due to a small amount of transcript being expressed in AM3-glycerol that may have been missed by RNAseq, or the expression level was below our cutoff of 50. Note that RT-PCR is not quantitative and that even trace amounts of DNA can be amplified to produce intense bands. It is not due to contaminating genomic DNA in the sample, as the negative control with the RNA sample does not produce a PCR band (lane 11).
(Note that the band seen at ~50 bp in lane 11 [DNase-treated total RNA used as template] is likely due to primer dimers, as the same band is not seen with another primer set tested with the same template [lane 17]). A similar result (a PCR product of the expected size for both glucose and glycerol samples, lanes 15-16), with a similar interpretation, was obtained for smRNA126, for which we expected a PCR product only with the cDNA originating from the glycerol sample. For smRNA76, the primers were designed to amplify a product only if the two sRNAs predicted

by BSRD corresponded to a single transcript. The 240 bp PCR product seen for smRNA76 (lane 14) confirms our RNAseq results.

As described above, we chose the 14 small RNAs to cover a wide variety of sRNAs belonging to different categories in our analysis. We were able to confirm all these sRNAs by RT-PCR. Based on these results, most other sRNAs detected by RNAseq are likely to be valid sRNAs.

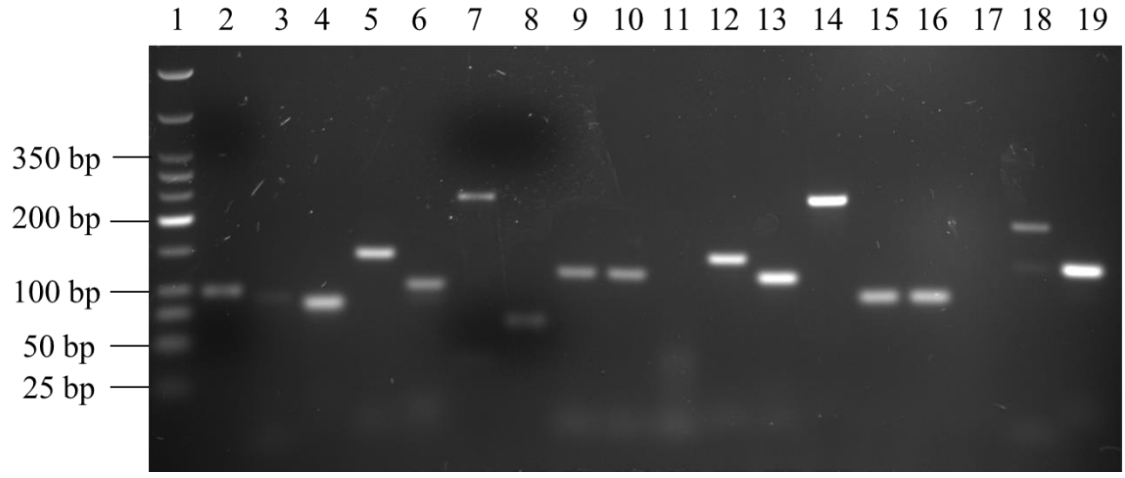

Figure 3.5 RT-PCR validation of small RNAs. Lane 1, low molecular weight ladder; Lane 2, smRNA4; Lane 3, smRNA8; Lane 4, smRNA11; Lane 5, smRNA18; Lane 6, smRNA21; Lane 7, smRNA27, Lane 8, smRNA28; Lane 9, smRNA47 (cDNA from Glucose-1); Lane 10, smRNA47 (cDNA from Glycerol-1), Lane 11, smRNA47 (DNAse treated RNA control from Glycerol-1); Lane 12, smRNA49; Lane 13, smRNA55; Lane 14, smRNA76; Lane 15, smRNA126 (cDNA from Glucose-1); Lane 16, smRNA126 (cDNA from Glycerol-1); Lane 17, smRNA126 (DNase treated RNA control from Glucose-1); Lane 18, smRNA135; Lane 19, smRNA187.

3.3.5 Scaffold selection for synthetic sRNA design based on predicted Hfq-dependent small RNAs.

Among the twenty four Hfq-dependent small RNAs predicted by ARNold, we short-listed a few candidates with other characteristic features of Hfq-dependent small RNAs known to be important for binding to Hfq: (i) Rho-independent terminator with a long poly(U) tail, (ii) internal hairpin, and (iii) U-rich sequence preceding the hairpin (49, 50). Seven candidates (Table 3.5) met the three criteria. Out of these, smRNA8 and smRNA28 were selected as scaffolds for synthetic small RNA design. In *E. coli*, sRNA MicA (51) is located near the 5' end of the *luxS* gene in the antisense orientation. In *A. succinogenes*, smRNA8 is similarly positioned upstream of *luxS* (ASUC_RS03555) in the opposite orientation (Figure 3.6). Although these two sRNAs do not share any sequence identity, they may be functionally similar. When we searched for target genes for smRNA8, OmpA was one of the top predicted targets (data not shown). In *E. coli*, MicA has been confirmed to be an antisense OmpA regulator and Hfq-dependent. We chose the smRNA8 scaffold for testing our synthetic small RNA for these reasons. We chose smRNA28 as a second scaffold, because it is conserved in at least three other Pasteurellaceae species—*Mannheimia succiniciproducens*, *Haemophilus somnus* 2336, and *Haemophilus somnus* 129PT.

Table 3.5 smRNA sequences that fit the criteria for Hfq-dependent sRNAs

sRNA candidates	Sequences
smRNA8 (-)	GCAGUUGUGAUUAAUAAUAAAAAAAUUGGUUCCUUAGGUUAUAUUCACCGCUCAAUUCCGCA AGGAAAAGAGCGG*UUUUUUUU*
smRNA11 (+)	UUAUUCAUAAAACCCCUUUUUACGCCGGAUUCCCUUAGUCCGGC*UUUUUU*
smRNA13 (-)	UGGCUAAACCGAUUUCACCCAGUAAUUCGCCGGCAUUAGUACCGACAAUCGCACCACCUAAA AUACGGUG*CGAUUCUUUGUC*
smRNA16 (-)	UCCUAAAUAAAAUAAUCAUAAAUAAAUAGCUAAACUACUUUCCUACCGUCCUUUUGGACGG*U UUUUUUUC*
smRNA27 (+)	UUUCUAGUUUGUCCGCUCUGCUUUCUUUUUCUACAAUACGCGCAUACUUAAUGACUGGUAAU UCCUUAAUUGAUUAAGAGUUGAAUCUUUUAGUUAAGUAUUAUGUUGUGUUUGCAUAUUGUUU GGGUAACCAAACAAAAGUAAUUAAUCCUUCUAUUUAAUUACUUAUUAACUUCCUGUAUAUUU ACUACCUAAUUUUAGGUUAUUGGCACCGCGCUUAAACUCCAAAAAAGUGCGGUG*UUUUUU*
smRNA28 (+)	AUUCAAACAAUAAUAGAUAAUCACUCCAACUUUCGGCGUUUCUCUCCCCACAAGGAAACGCC *UUUUUCU*
smRNA77 (+)	AUUCGUUUUUAACGGAAAAAACACCAUGAAAUCCGACCGCACUUUUACCAUGAUGACGAUUA CCACCAUUAUGACCUUUAUAAUGGCGGGGUAGUGCGAACGAAAGAUAUUCAAUUAAGCCCGC AACCUGAAAAGUGCGGGC*UUUUUU*

+ and – signs indicate the coding strand. Blue: sequences that may be complementary to target mRNA(s); red: inverted repeats in the hairpin; underline: AU-rich regions before the inverted repeats; and black and italics: poly(U) tails.

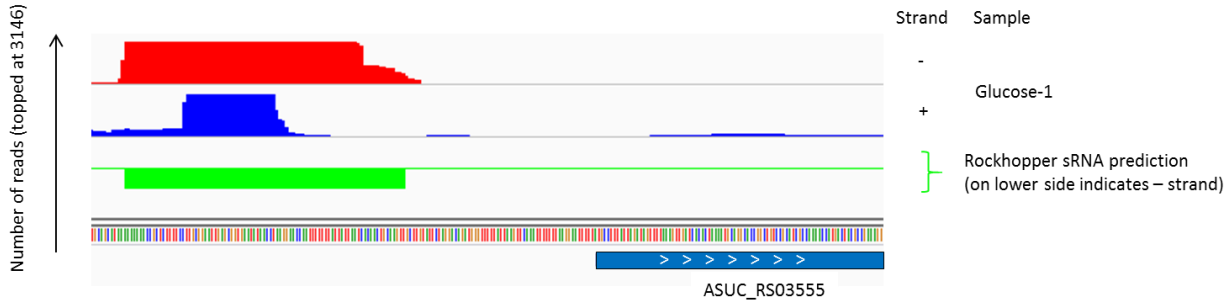

Figure 3.6 *A. succinogenes* smRNA8's genomic locus (Glucose-1 sample shown as a representative sample). -, reverse strand; red, transcript level from RNAseq; green, sRNA predicted by Rockhopper (in both glucose and glycerol libraries).

3.3.6 Testing of synthetic sRNA designs with *lacZ* as target gene.

Yoo et al. (52) designed a synthetic sRNA in *E. coli* using the scaffold from MicC, which is a well-characterized *E. coli* sRNA. Yoo et al. simply changed the complementary target gene binding regions. Addition of a terminator sequence was optional since the scaffold sequence has

a hairpin at the 3' end. In that study, a target binding sequence of 20-30 nt with binding energy between -30 and -40 kcal mol^{-1} worked best (52). Using these guidelines, we constructed four synthetic sRNA constructs targeting *lacZ* using smRNA8 and smRNA28 as scaffolds (Figure 3.7). smRNA_lacZ1 contains the smRNA8 scaffold sequence. smRNA_lacZ2 contains the smRNA28 scaffold sequence with a region complementary to the initial target mRNA kept intact (in blue) (Figure 3.7). smRNA_lacZ3 contains the smRNA28 scaffold sequence without any region complementary to the initial target mRNA. smRNA_lacZ1, smRNA_lacZ2, and smRNA_lacZ3 contain the same lacZ target sequence. smRNA_lacZ4 has the same scaffold as smRNA_lacZ3, but has a different *lacZ* target binding region. All constructs were cloned under control of a promoter of moderate strength (constructed in chapter 2).

We were unable to obtain *A. succinogenes* transformants with pLGZ920-$p_{pckA\text{-}92}$-smRNA_lacZ3. Strains 130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_lacZ1), 130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_lacZ2), and 130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_lacZ4) were grown on lactose to determine whether or not *lacZ* was expressed. Although all three strains were able to grow on AM3-lactose, 130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_lacZ1) and 130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_lacZ2) grew slower than 130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_lacZ4) and 130Z (pLGZ920), the positive control. Crude extracts of strain 130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_lacZ1) had a 32% decrease in β-galactosidase activity compared to the positive control ($p = 0.007$) (Figure 3.8). Crude extracts of strain 130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_lacZ2) showed a slight decrease in β-galactosidase activity that was not significant ($p = 0.12$), and β-galactosidase activity of strain 130Z pLGZ920-$p_{pckA\text{-}92}$-smRNA_lacZ4) was not affected.

CGAAAAUUUGAUCUAGUUAACAUUUUUUUAGGUAUAAAUAGUUUUAAAAUAGAUCUAGUUUGGAUUCAAAGUAGGUUGGCAGAAUCA
UUCAAUUCCGCAAGGAAAAGAGCGGUUUUUUUU (A)

CGAAAAUUUGAUCUAGUUAACAUUUUUUUAGGUAUAAAUAGUUUUAAAAUAGAUCUAGUUUGGAUUCAAAGUAGGUUGGCAGAAUCA
UAAUCACUCCAACUUUCGGCGUUUCUCUCCCCACAAGGAAACGCCUUUUUCU (B)

CGAAAAUUUGAUCUAGUUAACAUUUUUUUAGGUAUAAAUAGUUUUAAAAUAGAUCUAGUUUGGAUUCAAAGUAGGUUGGCAGAAUCA
UAACUUUCGGCGUUUCUCUCCCCACAAGGAAACGCCUUUUUCU (C)

CGAAAAUUUGAUCUAGUUAACAUUUUUUUAGGUAUAAAUAGUUUUAAAAUAGAUCUAGUUUGGAUAGAAUCAUGCUGAACUCCUUAA
ACUUUCGGCGUUUCUCUCCCCACAAGGAAACGCCUUUUUCU (D)

***lacZ* target binding regions**

5′ **AUG**AUUCUGCCAACCUACUUUGAA 3′ (E)

5′ UAAGGAGUUCAGC**AUG**AUUCU 3′ (F)

Figure 3.7 Synthetic sRNAs targeting *lacZ* expression. (A) smRNA_lacZ1; (B) smRNA_lacZ2; (C) smRNA_lacZ3; (D) smRNA_lacZ4; (E) Target binding region in synthetic small RNAs in constructs A, B, and C; (F) Target binding region in synthetic sRNA construct D. Underlined:$p_{pckA\text{-}92}$ promoter; green highlight: sequence complementary to *lacZ* mRNA; italics: scaffold sequence from smRNA8 (construct A) and smRNA28 (constructs B, C, and D); blue: sequences that may be complementary to initial target mRNA(s); red: inverted repeats preceded by AU-rich region and followed by poly(U) tail; bold: start codons in panels E and F. Binding energy for complementary regions is -36 kcal mol^{-1} for constructs A, B, and C, and -32.4 kcal mol^{-1} for construct D.

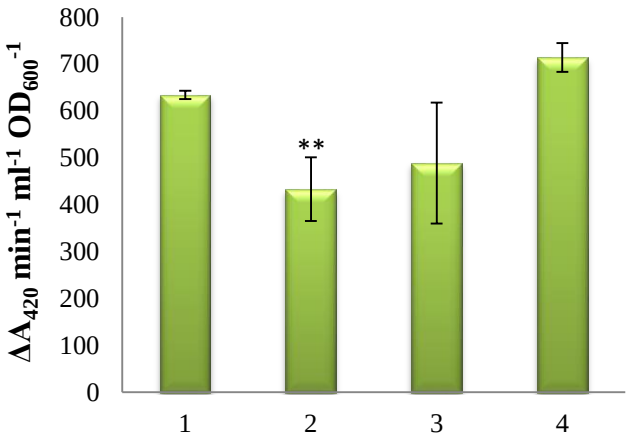

Figure 3.8 β-galactosidase activity of strain 130Z expressing synthetic sRNAs targeting *lacZ* expression. 1, 130Z (pLGZ920), positive control; 2, 130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_lacZ1); 3, pLGZ920-$p_{pckA\text{-}92}$-smRNA_lacZ2; and 4, pLGZ920-$p_{pckA\text{-}92}$-smRNA_lacZ4. Results are the average ± standard deviation based on three independent biological replicates.

3.3.7 Synthetic sRNAs for inhibiting acetate production.

Since a 32% decrease in β-galactosidase activity was observed using a synthetic sRNA based on the smRNA8 scaffold sequence, we used the same scaffold to produce synthetic sRNAs and decrease expression of the *ackA-pta* genes. The *ackA* and *pta* genes encode acetate kinase and phosphate acetyltransferase, respectively, which are responsible for acetate production from acetyl-CoA. We also used the smRNA28 scaffold of the smRNA_lacZ2 construct, which also caused a slight decrease in β-galactosidase activity. In both scaffolds, we replaced the *lacZ* target binding region specific for *lacZ* with *ackA* and *pta* target binding regions (Figure 3.9) to generate four constructs. Fermentation balances of strains 130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_ackA1), 130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_ackA2), 130Z (pLGZ920- $p_{pckA\text{-}92}$-smRNA_pta1) and 130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_pta2) grown anaerobically on glucose are shown in Table 3.6. Only strain 130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_ackA2) showed a significant decrease, 14%, in acetate production compared to 130Z (pLGZ920). Surprisingly, the two sRNA constructs based on the smRNA28 scaffold caused significant increases in acetate and succinate production at the expense of ethanol and biomass production.

CGAAAAUUUGAUCUAGUUAACAUUUUUUUAGGUAUAAAUAGUUUUAAAAUAGAUCUAGUUUGGACAGUUAAGAAUUAAAACUAAU
UUGGACAU*UCAAUUCCGCAAGGAAAAGAGCGGUUUUUUUU* (A)

CGAAAAUUUGAUCUAGUUAACAUUUUUUUAGGUAUAAAUAGUUUUAAAAUAGAUCUAGUUUGGACAGUUAAGAAUUAAAACUAAU
UUGGACAU*AAUCACUCCAACUUUCGGCGUUUCUCUCCCCACAAGGAAACGCCUUUUUCU* (B)

CGAAAAUUUGAUCUAGUUAACAUUUUUUUAGGUAUAAAUAGUUUUAAAAUAGAUCUAGUUUGGAGGGAUAAGAAUAAAUGUACGAGA
CAU*UCAAUUCCGCAAGGAAAAGAGCGGUUUUUUUU* (C)

CGAAAAUUUGAUCUAGUUAACAUUUUUUUAGGUAUAAAUAGUUUUAAAAUAGAUCUAGUUUGGAGGGAUAAGAAUAAAUGUACGAGA
CAU*AAUCACUCCAACUUUCGGCGUUUCUCUCCCCACAAGGAAACGCCUUUUUCU* (D)

ackA target binding region
5′ **AUG**UCCAAAUUAGUUUUAAUUCUUAACUG 3′ (E)

pta target binding region

5′ **AUG**UCUCGUACAUUUAUUCUUAUCCC 3′ (F)

Figure 3.9 Synthetic sRNA designs targeting expression of *ackA* and *pta*. (A) smRNA_ackA1; (B) smRNA_ackA2; (C) smRNA_pta1; (D) smRNApta2; (E) *ackA* target binding regions in synthetic sRNAs A and B; (F) *pta* target binding regions in synthetic sRNAs C and D. Underlined: $p_{pckA\text{-}92}$ promoter; green highlight: sequences complementary to *ackA* mRNA (A and B) and to *pta* mRNA (C and D) (binding energy is -35.2 kcal mol^{-1} [A and B] and -36.7 kcal mol^{-1} [C and D]); italics: scaffold sequence from smRNA_lacZ1 (A and C) and smRNA_lacZ2 (B and D); red: inverted repeats preceded by AU-rich region and followed by poly(U) tail; bold: start codons.

Table 3.6 Fermentation balances of strain 130Z expressing synthetic sRNA constructs.

	Products (mmol/100 mmol glucose consumed)							
Strain	Succinate	Formate	Acetate	Ethanol	CO_2	Biomass[a]	Carbon recovery[b]	Doubling time
130Z (pLGZ920)	47.9 ± 1.4	100 ± 2	65.5 ± 1.6	36 ± 3.4	2.06 ± 1.87	193 ± 10	104 ± 6	2.05 ± 0.03
130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_ackA1)	58.9 ± 4.2*	106 ± 10	76.8 ± 6.5*	25.3 ± 3.6	0.0±0.0	165 ± 8	104 ± 1	2.45 ± 0.15
130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_ackA2)	46.5 ± 1.4	81.1 ± 4.2	56.1 ± 2.2**	36.5 ± 16.9	0.0±0.0	186 ± 3	105 ± 3	2.17 ± 0.05
130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_pta1)	60.8 ± 4.2*	103 ± 10	73.6 ± 3.8*	25.5 ± 12.5	0.0±0.0	146 ± 3	102 ± 2	2.51 ± 0.01
130Z (pLGZ920-$p_{pckA\text{-}92}$-smRNA_pta2)	56.6 ± 5.4	97.2 ± 8.8	68.2 ± 5.4	22.9 ± 7.6	0.0 ±0.0	149 ± 8	100 ± 9	2.40 ± 0.09

Results are an average of three biological replicates ± standard deviations.

[a]Biomass was determined using assumed values of 567 mg dry cell weight/mL per OD_{660} (45) and a cell composition of $CH_2O_{0.5}N_{0.2}$ (24.967 g/mol) (53)

[b]Carbon balance is the carbon in products/carbon in glucose consumed. It is assumed that one CO_2 is fixed for each molecule of succinate produced.

CO_2 is calculated using the following formula: CO_2 (in mM)= Ethanol (in mM)+ Acetate (in mM) - Formate (in mM)

* Significantly different from 130Z (pLGZ920) ($p < 0.05$, two-tailed student's t-test)

** Significantly different from 130Z (pLGZ920) ($p < 0.01$, two-tailed student's t-test)

3.4 Conclusion

In this study, we have successfully identified sRNAs in *A. succinogenes* and verified select ones using RT-PCR. In all, 260 sRNAs were identified using Rockhopper and manual annotation using IGV, in *A. succinogenes* grown anaerobically on glucose and/or microaerobically on glycerol. Combined with RNAseq, we used five computational programs to identify additional small RNAs in *A. succinogenes.* Thirty-nine of the sRNAs detected by RNAseq were also predicted by at least one, if not more, of the computational programs we used. We validated 14 small RNAs from different categories in our analysis using RT-PCR.

We have also shown that synthetic sRNAs can be used in *A. succinogenes* to decrease gene expression. While our first attempt at designing and using a synthetic sRNA was successful as a proof of concept, much optimization is needed in the synthetic sRNA design and promoter strength as well. Further optimization of these synthetic sRNAs would likely entail changing the length and binding energy of the target regions or even testing several different scaffold sequences. One experiment that is likely to give us more insight into Hfq-binding sRNAs is a co-purification experiment using Hfq-his_6. We plan to sequence the Hfq-binding RNAs obtained from this experiment to confirm the Hfq-binding sRNAs identified in this study, maybe identify more Hfq-binding sRNAs, and identify some of the target mRNAs.

Our study is a great first step towards understanding sRNAs and realizing their potential as metabolic tools for engineering *A. succinogenes* and studying other pasteurellaceae species.

3.5 Acknowledgements

I thank Justin Beauchamp for his hands on help with the initial troubleshooting for analysis of the RNAseq. Rhiannon LeVeque for sharing the Mulk's laboratory RNA purification protocol.

I thank Dr. Martha Mulks for her helpful discussions in the initial planning phase of this study. I thank Dr. Ben Johnson and John Johnson for helpful discussions regarding RNA analysis. I am grateful to Jeff Landgraf of the MSU Genomics Core for making the small RNA libraries and sequencing.

APPENDIX

Table A3.1 Manually identified smRNAs using alignment files created by Rockhopper

Candidate	Rfam annotation	Transcription start	Transcription stop	Size (nt)	Coding strand	Neighborhood	Predicted by algorithms
smRNA146	tRNA	30,584	30,640	56	+	ASUC_RS00135/ASUC_RS00140	M
smRNA147		31,859	31,907	48	-	ASUC_RS00140/ASUC_RS00145	
smRNA148		39,498	39,529	31	-	ASUC_RS00185/ASUC_RS00190	
smRNA149		45,565	45,614	49	-	ASUC_RS00220/ASUC_RS00225	
smRNA150		48,319	48,364	45	+	ASUC_RS00230/ASUC_RS00235	
smRNA151		83,116	83,215	99	-	ASUC_RS00360/ASUC_RS00365	
smRNA152		102,424	102,470	46	+	ASUC_RS00455/ASUC_RS00460	
smRNA153		102,498	102,582	84	+	ASUC_RS00455/ASUC_RS00460	
smRNA154		102,542	102,588	46	-	ASUC_RS00455/ASUC_RS00460	
smRNA155		113,439	113,530	91	-	ASUC_RS00530/ASUC_RS00535	
smRNA156		115,771	115,918	147	-	ASUC_RS00535/ASUC_RS00540	
smRNA157		130,180	130,330	150	-	ASUC_RS00600/ASUC_RS00605	
smRNA158		138,915	139,034	119	+	ASUC_RS00630/ASUC_RS00635	
smRNA159		180,930	181,103	173	-	ASUC_RS00810/ASUC_RS00815	
smRNA160		194,098	194,153	55	+	ASUC_RS00875/ASUC_RS00880	
smRNA161		195,459	195,540	81	+	ASUC_RS00880/ASUC_RS00885	
smRNA162		199,493	199,572	79	+	ASUC_RS00895/ASUC_RS00900	
smRNA163		199,497	199,570	73	-	ASUC_RS00895/ASUC_RS00900	
smRNA164		214,454	214,513	59	-	ASUC_RS00965/ASUC_RS00970	
smRNA165		214,473	214,508	35	+	ASUC_RS00965/ASUC_RS00970	
smRNA166		218,099	218,172	73	+	ASUC_RS00975/ASUC_RS00980	
smRNA167		248,074	248,108	34	+	ASUC_RS01110/ASUC_RS01115	
smRNA168		293,615	293,774	159	-	ASUC_RS01350/ASUC_RS01355	
smRNA169		298,474	298,517	43	-	ASUC_RS01375/ASUC_RS01380	
smRNA170		306,273	306,369	96	+	ASUC_RS01415/ASUC_RS01420	
smRNA171		332,486	332,548	62	-	ASUC_RS01540/ASUC_RS01545	
smRNA172		339,352	339,413	61	+	ASUC_RS01585/ASUC_RS01590	

Table A3.1 (cont'd)

smRNA173		420,735	420,767	32	+	ASUC_RS01910/ASUC_RS01915	
smRNA174		429,830	429,921	91	+	ASUC_RS01955/ASUC_RS01960	
smRNA175		436,423	436,943	520	-	ASUC_RS01975/ASUC_RS01980	S, M
smRNA176		442,689	442,725	36	-	ASUC_RS02010/ASUC_RS02015	
smRNA177		445,536	445,584	48	-	ASUC_RS02030/ASUC_RS02035	
smRNA178		445,542	445,585	43	+	ASUC_RS02030/ASUC_RS02035	
smRNA179		449,450	449,585	135	+	ASUC_RS02055/ASUC_RS02060	
smRNA180		553,135	553,183	48	+	ASUC_RS02590/ASUC_RS02595	
smRNA181		606,527	606,776	249	-	ASUC_RS02850/ASUC_RS02855	
smRNA182		635,596	635,688	92	+	ASUC_RS02995/ASUC_RS03000	
smRNA183		662,062	662,237	175	-	ASUC_RS03105/ASUC_RS03110	
smRNA184		663,380	663,492	112	+	ASUC_RS03115/ASUC_RS03120	
smRNA185	His_leader	691,158	691,339	181	+	ASUC_RS03270/ASUC_RS03275	B, I
smRNA186		701,823	701,942	119	+	ASUC_RS03320/ASUC_RS03325	
smRNA187	6S RNA	718,766	718,947	181	-	ASUC_RS03405/ASUC_RS03410	B, I, Z, M
smRNA188		731,105	731,329	224	+	ASUC_RS03475/ASUC_RS03480	
smRNA189		744,521	744,551	30	+	ASUC_RS03550/ASUC_RS03555	
smRNA190		785,398	785,589	191	-	ASUC_RS03755/ASUC_RS03760	Z
smRNA191		831,957	832,102	145	+	ASUC_RS04000/ASUC_RS010990	
smRNA192		837,804	838,075	271	-	ASUC_RS04020/ASUC_RS04025	
smRNA193		845,242	845,452	210	+	ASUC_RS04055/ASUC_RS04060	
smRNA194		866,148	866,268	120	+	ASUC_RS04165/ASUC_RS04170	
smRNA195		868,810	868,875	65	+	ASUC_RS04180/ASUC_RS04185	
smRNA196		936,289	936,468	179	+	ASUC_RS04530/ASUC_RS04535	
smRNA197		967,596	967,728	132	+	ASUC_RS04675/ASUC_RS04680	
smRNA198		978,135	978,207	72	+	ASUC_RS04730/ASUC_RS04735	
smRNA199		979,854	980,036	182	+	ASUC_RS04735/ASUC_RS04740	Z
smRNA200		1,000,431	1,000,539	108	-	ASUC_RS04845/ASUC_RS04850	Z
smRNA201		1,052,438	1,052,472	34	-	ASUC_RS05065/ASUC_RS05070	Z
smRNA202	Parecho_CRE	1,059,515	1,059,731	216	-	ASUC_RS05085/ASUC_RS05090	I
smRNA203		1,064,546	1,064,610	64	+	ASUC_RS05105/ASUC_RS05110	

Table A3.1 (cont'd)

smRNA204		1,101,346	1,101,403	57	+	ASUC_RS05270/ASUC_RS05275	
smRNA205		1,136,235	1,136,327	92	-	ASUC_RS05410/ASUC_RS05415	
smRNA206		1,148,049	1,148,185	136	+	ASUC_RS05455/ASUC_RS05460	
smRNA207		1,169,658	1,169,830	172	+	ASUC_RS05560/ASUC_RS05565	
smRNA208		1,178,550	1,178,645	95	-	ASUC_RS05600/ASUC_RS05605	
smRNA209		1,185,698	1,185,896	198	-	ASUC_RS05625/ASUC_RS05630	
smRNA210		1,192,341	1,192,408	67	-	ASUC_RS05655/ASUC_RS05660	
smRNA211		1,192,343	1,192,388	45	+	ASUC_RS05655/ASUC_RS05660	
smRNA212		1,226,566	1,226,635	69	+	ASUC_RS05800/ASUC_RS05805	
smRNA213		1,249,629	1,249,785	156	+	ASUC_RS05910/ASUC_RS05915	
smRNA214	TPP	1,254,045	1,254,220	175	+	ASUC_RS05930/ASUC_RS05935	I
smRNA215		1,258,601	1,258,640	39	+	ASUC_RS05950/ASUC_RS05955	
smRNA216		1,259,001	1,259,063	62	-	ASUC_RS05955/ASUC_RS05960	
smRNA217		1,350,512	1,350,633	121	-	ASUC_RS06255/ASUC_RS06260	
smRNA218		1,356,570	1,356,750	180	+	ASUC_RS06285/ASUC_RS06290	
smRNA219		1,377,923	1,377,967	44	+	ASUC_RS06450/ASUC_RS06455	Z
smRNA220		1,378,631	1,378,725	94	+	ASUC_RS06455/ASUC_RS06460	
smRNA221		1,379,749	1,379,827	78	+	ASUC_RS06465/ASUC_RS06470	Z
smRNA222	TPP riboswitch	1,412,500	1,412,583	83	-	ASUC_RS06640/ASUC_RS06645	B, I, M
smRNA223		1,430,292	1,430,344	52	+	ASUC_RS06720/ASUC_RS06725	
smRNA224		1,438,223	1,438,385	162	+	ASUC_RS06755/ASUC_RS06760	
smRNA225		1,440,698	1,440,777	79	-	ASUC_RS06770/ASUC_RS06775	
smRNA226		1,498,787	1,498,840	53	+	ASUC_RS07045/ASUC_RS07050	
smRNA227		1,520,162	1,520,316	154	-	ASUC_RS07125/ASUC_RS07130	
smRNA228		1,543,389	1,543,436	47	+	ASUC_RS07245/ASUC_RS07250	
smRNA229		1,607,502	1,607,619	117	+	ASUC_RS07530/ASUC_RS07535	
smRNA230		1,625,036	1,625,112	76	-	ASUC_RS07625/ASUC_RS7630	
smRNA231		1,628,813	1,628,940	127	-	ASUC_RS07635/ASUC_RS07640	
smRNA232		1,644,162	1,644,218	56	+	ASUC_RS07700/ASUC_RS07705	
smRNA233		1,654,562	1,654,858	296	+	ASUC_RS07745/ASUC_RS07750	
smRNA234		1,665,202	1,665,386	184	-	ASUC_RS07805/ASUC_RS07810	

Table A3.1 (cont'd)

smRNA235	TPP riboswitch	1,668,651	1,668,974	323	+	ASUC_RS07825/ASUC_RS07830	B, I, M
smRNA236		1,707,974	1,708,024	50	-	ASUC_RS08025/ASUC_RS08030	
smRNA237		1,712,810	1,712,890	80	-	ASUC_RS08055/ASUC_RS08060	
smRNA238		1,724,670	1,724,828	158	-	ASUC_RS08120/ASUC_RS08125	
smRNA239		1,741,704	1,741,752	48	-	ASUC_RS08185/ASUC_RS08190	
smRNA240		1,825,558	1,825,646	88	-	ASUC_RS08605/ASUC_RS08610	
smRNA241		1,837,348	1,837,609	261	-	ASUC_RS08675/ASUC_RS08680	
smRNA242	MOCO_RNA_motif	1,838,604	1,838,746	142	-	ASUC_RS08680/ASUC_RS08685	B, I, M
smRNA243		1,851,093	1,851,234	141	-	ASUC_RS08730/ASUC_RS08735	Z
smRNA244		1,862,282	1,862,350	68	-	ASUC_RS08780/ASUC_RS08785	Z
smRNA245		1,881,063	1,881,151	88	+	ASUC_RS08865/ASUC_RS08870	
smRNA246		1,968,657	1,968,878	221	-	ASUC_RS09250/ASUC_RS09255	
smRNA247		1,985,767	1,985,884	117	+	ASUC_RS09360/ASUC_RS09365	
smRNA248		2,001,685	2,001,800	115	-	ASUC_RS09445/ASUC_RS09450	
smRNA249		2,013,568	2,013,612	44	+	ASUC_RS09495/ASUC_RS09500	
smRNA250		2,038,722	2,038,821	99	+	ASUC_RS09605/ASUC_RS09610	
smRNA251		2,089,369	2,089,514	145	+	ASUC_RS09840/ASUC_RS09845	S
smRNA252		2,144,262	2,144,302	40	-	ASUC_RS10070/ASUC_RS10075	Z
smRNA253		2,144,312	2,144,415	103	-	ASUC_RS10070/ASUC_RS10075	
smRNA254		2,177,961	2,178,043	82	+	ASUC_RS10215/ASUC_RS10220	
smRNA255		2,181,502	2,181,720	218	-	ASUC_RS10325/ASUC_RS10240	
smRNA256		2,222,240	2,222,496	256	+	ASUC_RS10420/ASUC_RS10425	
smRNA257		2,258,917	2,259,021	104	-	ASUC_RS10620/ASUC_RS10625	
smRNA258		2,271,792	2,271,888	96	+	ASUC_RS10665/ASUC_RS10670	
smRNA259		2,282,160	2,282,232	72	+	ASUC_RS10725/ASUC_RS10730	
smRNA260		2,286,315	2,286,393	78	+	ASUC_RS10750/ASUC_S10755	

B, BSRD; I, INFERNAL; M, Rfam; S, SIPHT; Z, RNAz

Table A3.2 smRNAs common between SIPHT and RNAseq

Candidate	Rfam annotation	Predicted by SIPHT			
		Transcription start	Transcription stop	Size (nt)	Coding strand
smRNA24	-	1,441,909	1,441,669	465	+
smRNA28	-	2,003,809	2,003,941	132	+
smRNA55	Alpha_RBS	527,920	527,984	64	+
smRNA88	-	1,057,774	1,057,957	183	-
smRNA175	CRISPR_DR34	436,427	436,904	477	-
smRNA251	-	2,089,126	2,089,365	239	+

Table A3.3 smRNAs common between BSRD and RNAseq

Candidate	Rfam annotation	Listed in BSRD			Coding strand
		Transcription Start	Transcription Stop	Size (nt)	
smRNA18	SRP_bact	1,215,630	1,215,727	98	+
smRNA27	GcvB	1,862,334	1,862,537	204	+
smRNA55	Alpha_RBS	527,966	528,079	114	+
smRNA76	Glycine riboswitch	806,377	806,566	180	+
smRNA185	His_leader	691,203	691,334	132	+
smRNA187	6S RNA	718,765	718,947	183	-
smRNA222	TPP riboswitch	1,412,495	1,412,588	94	-
smRNA235	TPP riboswitch	1,668656	1,668,748	221	+
smRNA242	MOCO_RNA_motif	1,838,596	1,838,749	154	-

Table A3.4 smRNAs common between INFERNAL and RNAseq

Candidate	Rfam annotation	Predicted in INFERNAL			Coding strand
		Transcription start	Transcription stop	Size (nt)	
smRNA4	PyrR	431,716	431,819	104	-
smRNA11	MicF	981,537	981,628	92	+
smRNA18	SRP_bact	1,215,630	1,215,730	101	+
smRNA27	GcvB	1,862,334	1,862,537	204	+
smRNA55	Alpha_RBS	527,966	528,079	114	+
smRNA57	Parecho_CRE	599,496	599,595	100	+
smRNA76	Glycine	806,377	806,566	189	+
smRNA78	Parecho_CRE	869,491	869,593	103	+
smRNA185	His_leader	691,203	691,334	132	+
smRNA187	6S	718,765	718,947	183	-
smRNA202	Parecho_CRE	1,059,499	1,059,618	120	+
smRNA214	TPP	1,254,038	1,254,175	138	+
smRNA222	TPP	1,412,495	1,412,588	94	-
smRNA235	TPP	1,668,656	1,668,748	93	+
smRNA242	MOCO_RNA_motif	1,838,596	1,838,749	154	-

Table A3.5 smRNAs common between RNAz and RNAseq

Candidate	Rfam annotation	Predicted by RNAz			
		Transcription start	Transcription stop	Size (nt)	Coding strand
smRNA5		517,020	51,7211	192	+
smRNA9		785,646	785,837	192	-
smRNA18	SRP_bact	1,215,619	1,215,737	119	+
smRNA20		1,276,338	1,276,512	175	+
smRNA21		1,276,596	1,276,762	167	+
smRNA23		1,376,647	1,376,896	250	+
smRNA26		1,545,554	1,545,663	110	+
smRNA29		2,089,122	2,089,319	198	+
smRNA55	Alpha_RBS	527,919	528,046	128	+
smRNA135		1,981,636	1,981,823	188	-
smRNA187	6S	718,771	718,946	176	-
smRNA190		785,408	785,590	183	-
smRNA199		979,919	980,050	132	+
smRNA200		1,000,433	1,000,492	60	-
smRNA201		1,052,422	1,052,593	172	-
smRNA219		1,377,838	1,377,966	129	+
smRNA221		1,379,748	1,379,829	82	+
smRNA243		1,851,057	1,851,225	169	-
smRNA244		1,862,313	1,862,473	161	-
smRNA252		2,144,253	2,144,349	97	-

Table A3.6 smRNAs common between Rfam and RNAseq

Candidate	Rfam annotation	Predicted by BLAST against Rfam			Coding strand
		Transcription start	Transcription stop	Size (nt)	
smRNA18	SRP_bact	1,215,630	1,215,730	101	+
smRNA27	GcvB	1,862,318	1,862,528	211	+
smRNA55	Alpha_RBS	5,279,66	5,28,079	114	+
smRNA77	Thr_leader	8,58,744	8,58,862	119	+
smRNA146	tRNA	30,593	30,634	42	-
smRNA175	CRISPR-DR34	4,36,472	4,36,903	36	-
smRNA187	6S	718,765	718,947	183	-
smRNA222	TPP	1,412,505	1,412,583	79	-
smRNA235	TPP	1,668,661	1,668,738	78	+
smRNA242	MOCO_RNA_motif	1,838,596	1,838,749	154	-

Table A3.7 Primers used for RT-PCR validation of sRNAs

Primer	Sequence	Expected product size (n)
smRNA4_F	ATTGTAATGGCACTGCGAAATG	99
smRNA4_R	AAAAAACGATGCCCCCTGAG	
smRNA8_F	AAGCAACTTGCAGTTGTG	91
smRNA8_R	AAAAAAAACCGCTCTTTTCCTTG	
smRNA11_F	AACGTTGTGATGATGTTAAACG	85
smRNA11_R	AAAAAAAGCCGGACTAAGGGAAT	
smRNA18_F	ACTAAGCCGGTGTGCGAAAG	146
smRNA18_R	GAAACCTCCCCAGTGATTC	
smRNA21_F	AAAATTAACCGCACTTTATGTTC	105
smRNA21_R	ACAAGTAAACTCGCTACGC	
smRNA27_F	TATTTCTAGTTTGTCCGCTCT	244
smRNA27_R	AACACCGCACTTTTTTGGAG	
smRNA28_F	ATTCAAACAATAATAGATAATCACTCC	70
smRNA28_R	AAGAAAAAGGCGTTTCCTTGTG	
smRNA47_F	TGCAATTCGTCGGAGATTTG	120
smRNA47_R	ATTGAACAATCCGAAAGTTCC	
smRNA49_F	ATGCCGACTTTAGGACAAAAG	141
smRNA49_R	CTTACGGCGAGGGTATTC	
smRNA55_F	ACAGGTTGAGCAGTTATACTG	114
smRNA55_R	GCACTCCTATATTTTAACTAATTTG	
smRNA76_F	CAATAAAAAACTGTTCGGACGAAGG	240
smRNA76_R	GGCAGAAGAGATAAAAATGATAGG	
smRNA126_F	GAATTTGGCAGAGAAGTAGTAC	93
smRNA126_R	AATACCCAAGCGGTCCTC	
smRNA135_F	TTTAGCAAGTGCGGTCAG	193
smRNA135_R	AAAGGCAACCCTGCTTTAC	
smRNA187_F	ATTACCTGAGATGCTCGCCAGC	128
smRNA187_R	GGAGTCAGTTGTAACCGTTTTTAGG	

Table A3.8 Primers and g-blocks used in this study

Primer	Sequence	Specificity (direction)
CV662[a]	CATACGGATCCCGAAAATTTGATCTAGTTAACATTTTTTAGGTATAAATAGTTTT AAAATAGATCTAGTTTGGAAGAATCATGCTGAACTCCTTAAACTTTCGGCGTTTC TCTCCCCACAAGGAAACGCCTTTTTCTTCTAGAATCGA	smRNA_lacZ1 (F)
CV663[a]	TCGATTCTAGAAAAAAAAACCGCTCTTTTCCTTGCGGAATTGAATGATTCTGCCA ACCTACTTTGAATCCAAACTAGATCTATTTTAAAACTATTTATACCTAAAAAATG TTAACTAGATCAAATTTTCGGGATCCGTATG	smRNA_lacZ1 (R)
CV664[a]	CATACGGATCCCGAAAATTTGATCTAGTTAACATTTTTTAGGTATAAATAGTTTT AAAATAGATCTAGTTTGGATTCAAAGTAGGTTGGCAGAATCATAATCACTCCAAC TTTCGGCGTTTCTCTCCCCACAAGGAAACGCCTTTTTCTTCTAGAATCGA	smRNA_lacZ2 (F)
CV665[a]	TCGATTCTAGAAGAAAAAGGCGTTTCCTTGTGGGGAGAGAAACGCCGAAAGTTGG AGTGATTATGATTCTGCCAACCTACTTTGAATCCAAACTAGATCTATTTTAAAAC TATTTATACCTAAAAAATGTTAACTAGATCAAATTTTCGGGATCCGTATG	smRNA_lacZ2 (R)
CV666[a]	CATACGGATCCCGAAAATTTGATCTAGTTAACATTTTTTAGGTATAAATAGTTTT AAAATAGATCTAGTTTGGATTCAAAGTAGGTTGGCAGAATCATAACTTTCGGCGT TTCTCTCCCCACA AGGAAACGCCTTTTTCTTCTAGAATCGA	smRNA_lacZ3 (F)
CV667[a]	TCGATTCTAGAAGAAAAAGGCGTTTCCTTGTGGGGAGAGAAACGCCGAAAGTTAT GATTCTGCCAACCTACTTTGAATCCAAACTAGATCTATTTTAAAACTATTTATAC CTAAAAAATGTTAACTAGATCAAATTTTCGGGATCCGTATG	smRNA_lacZ3 (R)
CV668[a]	CATACGGATCCCGAAAATTTGATCTAGTTAACATTTTTTAGGTATAAATAGTTTT AAAATAGATCTAGTTTGGATAGAATCATGCTGAACTCCTTAAACTTTCGGCGTTT CTCTCCCCACAAG GAAACGCCTTTTTCTTCTAGAATCGA	smRNA_lacZ4 (F)
CV669[a]	TCGATTCTAGAAGAAAAAGGCGTTTCCTTGTGGGGAGAGAAACGCCGAAAGTTTA AGGAGTTCAGCATGATTCTATCCAAACTAGATCTATTTTAAAACTATTTATACCT AAAAAATGTTAACTAGATCAAATTTTCGGGATCCGTATG	smRNA_lacZ4 (R)
CV677	AGCGCCTGATGCGGGGATCCCGAAAATTTGATCTAGTTAACATTTTTTAG	smRNA_lacZ1 (F)
CV678	CCAGTGAATTCGAGCTCAAAAAAAACCGCTCTTTTC	smRNA_lacZ1 (R)
CV679	CCAGTGAATTCGAGCTCAGAAAAAGGCGTTTCCTTG	smRNA_lacZ2, smRNA_lacZ3, smRNA_lacZ4 (R)
CV640	CTTCGCTATTACGCCAGCTG	Sequencing primer for pLGZ920 (90 bp downstream of *Sac*I site)
CV695[a,b]	CATACGGATCCCGAAAATTTGATCTAGTTAACATTTTTTAGGTATAAATAGTTTT AAAATAGATCTAGTTTGGACAGTTAAGAATTAAAACTAATTTGGACATTCAATTC CGCAAGGAAAAGAGCGGTTTTTTTTTCTAGAATCGA	smRNA_ackA1
CV696[a,b]	CATACGGATCCCGAAAATTTGATCTAGTTAACATTTTTTAGGTATAAATAGTTTT AAAATAGATCTAGTTTGGACAGTTAAGAATTAAAACTAATTTGGACATAATCACT CCAACTTTCGGCGTTTCTCTCCCCACAAGGAAACGCCTTTTTCTTCTAGAATCGA	smRNA_ackA2
CV697[a,b]	CATACGGATCCCGAAAATTTGATCTAGTTAACATTTTTTAGGTATAAATAGTTTT AAAATAGATCTAGTTTGGAGGGATAAGAATAAATGTACGAGACATTCAATTCCGC AAGGAAAAGAGCGGTTTTTTTTTCTAGAATCGA	smRNA_pta1
CV698[a,b]	CATACGGATCCCGAAAATTTGATCTAGTTAACATTTTTTAGGTATAAATAGTTTT AAAATAGATCTAGTTTGGAGGGATAAGAATAAATGTACGAGACATAATCACTCCA ACTTTCGGCGTTTCTCTCCCCACAAGGAAACGCCTTTTTCTTCTAGAATCGA	smRNA_pta2

[a] Restriction sites *Bam*HI and *Xba*I underlined

[b] Sequences for g-blocks.

REFERENCES

REFERENCES

1. **Itoh T**, **Tomizawa J-I**. 1980. Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. Proc Natl Acad Sci USA **77**:2450-2454.

2. **Stougaard P**, **Molin S**, **Nordström K**. 1981. RNAs involved in copy-number control and incompatibility of plasmid R1. Proc Natl Acad Sci USA **78**:6008-6012.

3. **Gottesman S**, **Storz G**. 2011. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Biol **3**:1-16.

4. **Moller T**, **Franch T**, **Hojrup P**, **Keene DR**, **Bachinger HP**, **Brennan RG**, **Valentin-Hansen P**. 2002. Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. Mol Cell **9**:23-30.

5. **Mandal M**, **Boese B**, **Barrick JE**, **Winkler WC**, **Breaker RR**. 2003. Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. Cell **113**:577-586.

6. **Loh E**, **Dussurget O**, **Gripenland J**, **Vaitkevicius K**, **Tiensuu T**, **Mandin P**, **Repoila F**, **Buchrieser C**, **Cossart P**, **Johansson J**. 2009. A trans-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. Cell **139**:770-779.

7. **Brouns SJJ**, **Jore MM**, **Lundgren M**, **Westra ER**, **Slijkhuis RJH**, **Snijders APL**, **Dickman MJ**, **Makarova KS**, **Koonin EV**, **Van Der Oost J**. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science **321**:960-964.

8. **Chao Y**, **Papenfort K**, **Reinhardt R**, **Sharma CM**, **Vogel J**. 2012. An atlas of Hfq-bound transcripts reveals 3′ UTRs as a genomic reservoir of regulatory small RNAs. EMBO J **31**:4005-4019.

9. **Chao Y**, **Vogel J**. 2016. A 3' UTR-derived small RNA provides the regulatory noncoding arm of the inner membrane stress response. Mol Cell **61**:352-363.

10. **Chao Y**, **Papenfort K**, **Reinhardt R**, **Sharma CM**, **Vogel J**. 2012. An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. EMBO J **31**:4005-4019.

11. **Bouvier J**, **Stragier P**, **Morales V**, **Remy E**, **Gutierrez C**. 2008. Lysine represses transcription of the *Escherichia coli* dapB gene by preventing its activation by the ArgP activator. J Bacteriol **190**:5224-5229.

12. **Soutourina OA**, **Monot M**, **Boudry P**, **Saujet L**, **Pichon C**, **Sismeiro O**, **Semenova E**, **Severinov K**, **Le Bouguenec C**, **Coppee JY**, **Dupuy B**, **Martin-Verstraete I**. 2013.

Genome-wide identification of regulatory RNAs in the human pathogen *Clostridium difficile*. PLoS Genet **9**:e1003493.

13. **McClure R**, **Tjaden B**, **Genco C**. 2014. Identification of sRNAs expressed by the human pathogen *Neisseria gonorrhoeae* under disparate growth conditions. Front Microbiol 5:456-456.

14. **Sharma R**, **Arya S**, **Patil SD**, **Sharma A**, **Jain PK**, **Navani NK**, **Pathania R**. 2014. Identification of novel regulatory small RNAs in *Acinetobacter baumannii*. PloS one **9**:e93833.

15. **Livny J**, **Brencic A**, **Lory S**, **Waldor MK**. 2006. Identification of 17 *Pseudomonas aeruginosa* sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. Nucleic Acids Res **34**:3484-3493.

16. **Schroeder CL**, **Narra HP**, **Sahni A**, **Rojas M**, **Khanipov K**, **Patel J**, **Shah R**, **Fofanov Y**, **Sahni SK**. 2016. Identification and characterization of novel small RNAs in *Rickettsia prowazekii*. Front Microbiol **7**:859.

17. **Ramos CG**, **Grilo AM**, **da Costa PJ**, **Leitao JH**. 2013. Experimental identification of small non-coding regulatory RNAs in the opportunistic human pathogen *Burkholderia cenocepacia* J2315. Genomics **101**:139-48.

18. **Mann B**, **van Opijnen T**, **Wang J**, **Obert C**, **Wang YD**, **Carter R**, **McGoldrick DJ**, **Ridout G**, **Camilli A**, **Tuomanen EI**, **Rosch JW**. 2012. Control of virulence by small RNAs in *Streptococcus pneumoniae*. PLoS Pathog **8**:e1002788.

19. **Christiansen JK**, **Nielsen JS**, **Ebersbach T**, **Valentin-Hansen P**, **Sogaard-Andersen L**, **Kallipolitis BH**. 2006. Identification of small Hfq-binding RNAs in *Listeria monocytogenes*. RNA **12**:1383-96.

20. **Arnvig KB**, **Young DB**. 2009. Identification of small RNAs in *Mycobacterium tuberculosis*. Mol Microbiol **73**:397-408.

21. **Kuhnert P**, **Christensen H**. 2008. Pasteurellaceae: biology, genomics and molecular aspects. Horizon Scientific Press.

22. **Santana EA**, **Harrison A**, **Zhang X**, **Baker BD**, **Kelly BJ**, **White P**, **Liu Y**, **Munson Jr RS**. 2014. HrrF is the Fur-regulated small RNA in nontypeable *Haemophilus influenzae*. PloS one **9**:e105644.

23. **Baddal B**, **Muzzi A**, **Censini S**, **Calogero RA**, **Torricelli G**, **Guidotti S**, **Taddei AR**, **Covacci A**, **Pizza M**, **Rappuoli R**. 2015. Dual RNA-seq of nontypeable *Haemophilus influenzae* and host cell transcriptomes reveals novel insights into host-pathogen cross talk. mBio **6**:e01765-15.

24. **Sridhar J**, **Sekar K**, **Rafi ZA**. 2009. CsrA interacting small RNAs in *Haemophilus* spp genomes: a theoretical analysis. Arch Microbiol **191**:451-459.

25. **Rossi CC**, **Bossé JT**, **Li Y**, **Witney AA**, **Gould KA**, **Langford PR**, **Bazzolli DMS**. 2016. A computational strategy for the search of regulatory small RNAs in *Actinobacillus pleuropneumoniae*. RNA **22**:1373-1385.

26. **Guettler MV**, **Rumler D**, **Jain MK**. 1999. *Actinobacillus succinogenes* sp. nov., a novel succinic-acid-producing strain from the bovine rumen. Int J Syst Bacteriol **49**:207–216.

27. **Cho C**, **Lee SY**. 2016. Efficient gene knockdown in *Clostridium acetobutylicum* by synthetic small regulatory RNAs. Biotechnol Bioeng **114**:374-383.

28. **Na D**, **Yoo SM**, **Chung H**, **Park H**, **Park JH**, **Lee SY**. 2013. Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. Nat Biotechnol **31**:170-174.

29. **McKinlay JB**, **Zeikus JG**, **Vieille C**. 2005. Insights into *Actinobacillus succinogenes* fermentative metabolism in a chemically defined growth medium. Appl Environ Microbiol **71**:6651-6656.

30. **Schindler BD**, **Joshi RJ**, **Vieille C**. 2014. Respiratory glycerol metabolism of *Actinobacillus succinogenes* 130Z for succinate production. J Ind Microbiol Biotechnol **41**:1339-1352.

31. **Kim P**, **Laivenieks M**, **McKinlay J**, **Vieille C**, **Zeikus JG**. 2004. Construction of a shuttle vector for the overexpression of recombinant proteins in *Actinobacillus succinogenes*. Plasmid **51**:108–115.

32. **Bolger AM**, **Lohse M**, **Usadel B**. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**:2114-20.

33. **Tjaden B**. 2015. De novo assembly of bacterial transcriptomes from RNA-seq data. Genome biology **16**:1.

34. **McClure R**, **Balasubramanian D**, **Sun Y**, **Bobrovskyy M**, **Sumby P**, **Genco CA**, **Vanderpool CK**, **Tjaden B**. 2013. Computational analysis of bacterial RNA-Seq data. Nucleic Acids Res **41**:e140-e140.

35. **Busch A**, **Richter AS**, **Backofen R**. 2008. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. Bioinformatics 24:2849-2856.

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

37. **Gruber AR**, **Lorenz R**, **Bernhart SH**, **Neuböck R**, **Hofacker IL**. 2008. The Vienna RNA Websuite. Nucleic Acids Res **36**:W70-W74.

38. **Livny J**, **Teonadi H**, **Livny M**, **Waldor MK**. 2008. High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. PLoS One **3**:e3197.

39. **Li L**, **Huang D**, **Cheung MK**, **Nong W**, **Huang Q**, **Kwan HS**. 2013. BSRD: a repository for bacterial small regulatory RNA. Nucleic Acids Res **41**:D233-D238.

40. **Nawrocki EP**, **Eddy SR**. 2013. Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics **29**:2933-2935.

41. **Gruber AR**, **Findeiss S**, **Washietl S**, **Hofacker IL**, **Stadler PF**. 2010. RNAz 2.0: improved noncoding RNA detection. Pac Symp Biocomput:69-79.

42. **Nawrocki EP**, **Burge SW**, **Bateman A**, **Daub J**, **Eberhardt RY**, **Eddy SR**, **Floden EW**, **Gardner PP**, **Jones TA**, **Tate J**, **Finn RD**. 2015. Rfam 12.0: updates to the RNA families database. Nucleic Acids Res **43**:D130-D137.

43. **Cros MJ**, **de Monte A**, **Mariette J**, **Bardou P**, **Grenier-Boley B**, **Gautheret D**, **Touzet H**, **Gaspin C**. 2011. RNAspace.org: An integrated environment for the prediction, annotation, and analysis of ncRNA. RNA **17**:1947-1956.

44. **Grissa I**, **Vergnaud G**, **Pourcel C**. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res **35**:W52-W57.

45. **McKinlay JB**, **Shachar-Hill Y**, **Zeikus JG**, **Vieille C**. 2007. Determining *Actinobacillus succinogenes* metabolic pathways and fluxes by NMR and GC-MS analyses of ^{13}C-labeled metabolic product isotopomers. Metab Eng **9**:177–192.

46. **McKinlay JB**, **Laivenieks M**, **Schindler BD**, **Mckinlay AA**, **Siddaramappa S**, **Challacombe JF**, **Lowry SR**, **Clum A**, **Lapidus AL**, **Burkhart KB**, **Harkins V**, **Vieille C**. 2010. A genomic perspective on the potential of *Actinobacillus succinogenes* for industrial succinate production. BMC Genomics **11**:680.

47. **Hofacker IL**, **Fontana W**, **Stadler PF**, **Bonhoeffer LS**, **Tacker M**, **P. S**. 1994. Fast folding and comparison of RNA secondary structures. Monatshefte für Chemie/Chemical Monthly **125**:167-188.

48. **Pulvermacher SC**, **Stauffer LT**, **Stauffer GV**. 2009. Role of the *Escherichia coli* Hfq protein in GcvB regulation of *oppA* and *dppA* mRNAs. Microbiology **155**:115-123.

49. **Ishikawa H**, **Otaka H**, **Maki K**, **Morita T**, **Aiba H**. 2012. The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich sequence and followed by a 3' poly(U) tail. RNA **18**:1062-1074.

50. **Otaka H**, **Ishikawa H**, **Morita T**, **Aiba H**. 2011. PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. Proc Natl Acad Sci U S A **108**:13059-13064.

51. **Udekwu KI**, **Darfeuille F**, **Vogel J**, **Reimegard J**, **Holmqvist E**, **Wagner EG**. 2005. Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. Genes Dev **19**:2355-2366.

52. **Yoo SM**, **Na D**, **Lee SY**. 2013. Design and use of synthetic regulatory small RNAs to control gene expression in *Escherichia coli*. Nat Protocols **8**:1694-1707.

53. **van der Werf MJ**, **Guettler MV**, **Jain MK**, **Zeikus JG**. 1997. Environmental and physiological factors affecting the succinate product ratio during carbohydrate fermentation by *Actinobacillus* sp. 130Z. Arch Microbiol **167**:332–342.

Chapter 4

Development of a markerless knockout method for *Actinobacillus succinogenes*

This work is previously published in Applied and Environmental Microbiology as:

Joshi RV, Schindler BD, McPherson N, Tiwari K, Vieille C (2014) Development of a Markerless Knockout Method for *Actinobacillus succinogenes*. Appl Environ Microbiol **80**:3053–3061. doi:10.1128/AEM.00492-14

The article can be viewed at the following URL:

http://aem.asm.org/content/80/10/3053

The work described in this chapter was partly completed as a requirement for a professional master's degree and was submitted as a form of a report titled, "Development of a gene knockout method for *A. succinogenes* strain 130Z and engineering of *A. succinogenes* for aspartate production". My contributions to this paper during my PhD are listed as below:

- Construction of double knockouts Δ*pflB* Δ*lacZ*::*icd*, Δ*pflB* Δ*cit*::*icd*, Δ*pflB* Δ*acn*::*icd*.

- Training and mentoring Kanupriya Tiwari during her research project in the summer of 2013. She worked on determining the effects of the DNA construct length and of the length of the homologous regions on the efficiency of homologous recombination in *A. succinogenes*.

- Drafting parts, reading, and editing manuscript.

4.1 Abstract

Actinobacillus succinogenes is one of the best natural succinate-producing organisms, but it still needs engineering to further increase succinate yield and productivity. In this study we developed a markerless knockout method for *A. succinogenes* using natural transformation or electroporation. The *Escherichia coli* isocitrate dehydrogenase gene with flanking flippase recognition target sites was used as the positive selection marker, making use of *A. succinogenes*'s auxotrophy for glutamate to select for growth on isocitrate. The *Saccharomyces cerevisiae* flippase recombinase (Flp) was used to remove the selection marker, allowing its re-use. Finally, the plasmid expressing *flp* was cured using acridine orange. We demonstrate that at least two consecutive deletions can be introduced into the same strain using this approach, that no more than a total of 1 kb of DNA is needed on each side of the selection cassette to protect from endonucleases activity during transformation, and that no more than 200 bp of homologous DNA is needed on each side for efficient recombination. We also demonstrate that electroporation can be used as an alternative transformation method to obtain knockout mutants, and that an enriched defined medium can be used for direct selection of knockout mutants on agar plates with high efficiency. Knockout mutants of the fumarate reductase and the citrate lyase operons, as well as of the pyruvate formate lyase, β-galactosidase, and aconitase genes were obtained using this knockout strategy.

4.2 Introduction

Dicarboxylic acids are among the top of the US's Department of Energy list of value-added chemicals from biomass (1). In particular, if produced economically by fermentation, succinic acid could replace maleic anhydride as the precursor to many bulk and commodity chemicals, with a potential market of 25 billion tons per year (2). *Actinobacillus succinogenes* strain 130Z (ATCC 55618) is a facultative anaerobe and a member of the *Pasteurellaceae*. It is the highest natural succinate producer known (3-5). *A. succinogenes* produces succinate, acetate, and formate as its major fermentative products with ethanol as a minor by-product (6). Under optimized conditions, wild-type *A. succinogenes* produces 80 g L^{-1} succinate, while chemically induced mutant strains resistant to fluoroacetate produce up to 110 g L^{-1} succinate (3), suggesting that even higher succinate yields can be obtained with genetically engineered strains. Bio-based succinate will only be price-competitive with petroleum-based maleic anhydride if alternative fermentation products can be eliminated. Achieving a homosuccinate fermentation would drastically reduce the cost of downstream succinate purification (7).

Until recently, genetic tools for *A. succinogenes* were limited to the expression vector pLGZ920 and electroporation (8). Plasmid pLGZ920 confers ampicillin (Amp) resistance, replicates in *A. succinogenes* and *Escherichia coli*, and allows high-level expression of foreign genes in *A. succinogenes* from the strong, constitutive *A. succinogenes pckA* promoter (p_{pckA}) (8). Electroporation of *A. succinogenes* with pLGZ920 yields transformants with an efficiency of 10^4–10^6 CFU/μg of plasmid, depending on the electroporation parameters (8) (Schindler and Vieille, unpublished results). Early attempts to construct knockout mutants of *A. succinogenes* by allelic exchange used electroporation with suicide vectors containing gene knockout constructs interrupted by antibiotic resistance genes. The high frequency of spontaneous antibiotic resistant

mutants masked the low frequency of double recombination events, and no knockout mutants were isolated in these studies (McKinlay and Vieille, unpublished results). Another selection method and possibly other means to introduce DNA into cells are thus needed to develop a knockout method for *A. succinogenes*.

In other *Pasteurellaceae* species conjugation and natural transformation have been commonly used to construct knockout mutants. Conjugation has been used in *A. pleuropneumoniae* (9), *Mannheimia haemolytica*, *Pasteurella multocida*, and *Haemophilus somnus* (10). Deletions were typically selected for with an antibiotic resistance marker, and SacB-based sucrose counterselection was used to select for double recombination. The conjugated plasmids were either suicide vectors or temperature-sensitive shuttle plasmids (10). One study used the Cre-*lox* system to remove the selection marker (11). Natural transformation has been used in many studies of *H. influenzae* and other naturally competent *Pasteurellaceae*. Natural competence is induced by starvation stress. Transformation frequencies can be as high as 10^{-2} (12). Natural transformation is well-suited for strain engineering since it works best with linear DNA (e.g., PCR products), which requires double recombination events for complete chromosome integration.

Naturally competent *Pasteurellaceae* preferentially take up DNA from their own species over unrelated DNA (13, 14). They do so with a membrane-bound DNA uptake machinery that specifically recognizes and binds a conserved uptake signal sequence (USS) (15). Low transformation efficiencies of *A. actinomycetemcomitans* were observed with DNA fragments not containing a USS (16). All *Pasteurellaceae* genomes sequenced contain USS repeats, in numbers ranging from 41 to 1,760 (1,690 in *A. succinogenes*) (17), even though not all these species are naturally competent (12). *A. succinogenes*'s genome contains twenty three of the

twenty five genes in *H. influenzae*'s natural competence regulon (17, 18), including those encoding the competence regulatory proteins Sxy and CRP (cyclic AMP receptor protein). The missing two genes encode hypothetical proteins with unknown roles in natural competence (12). The frequency of USS repeats in *A. succinogenes*'s genome and the likely presence of a complete natural competence machinery suggest that *A. succinogenes* is naturally competent.

In this study, we demonstrate that natural competence can be used to create knockout mutants of *A. succinogenes*, that the *E. coli* isocitrate dehydrogenase gene can be used as a positive selection marker, that the *Saccharomyces cerevisiae* Flp recombinase (encoded by *flp*) (19) can be used in *A. succinogenes* to remove the positive selection marker flanked by two direct Flp recognition target sites (*FRT*) (20), and that more than one deletion can be introduced in the same strain. We also demonstrate that no more than a total of 1 kb of DNA is needed on each side of the selection cassette to protect from exonucleases activity during transformation, and that no more than 200 bp of homologous DNA is needed for efficient recombination. The genes encoding fumarate reductase and pyruvate formate lyase were chosen as initial targets for knockout constructs. Fumarate reductase (encoded by *frdABCD*) converts fumarate to succinate. Knocking out fumarate reductase would likely be the first step needed to engineer *A. succinogenes* into a fumarate or aspartate producer. During *A. succinogenes*'s fermentative growth on glucose, pyruvate formate lyase (PFL, encoded by *pflB*) is the main enzyme converting pyruvate into acetyl-CoA, with the concomitant production of formate (21, 22). Acetyl-CoA is then the precursor of acetate and ethanol (23). Fluoroacetate-resistant mutants devoid of PFL activity are not affected in their growth and produce increased amounts of succinate (3). A *pflB* knockout mutation is likely to be the first mutation needed to engineer a homosuccinate-producing *A. succinogenes* strain (24-27). Three other sets of genes (encoding β-

galactosidase, citrate lyase, and aconitase) deemed non-essential for fermentative growth on glucose were targeted for deletion as well.

4.3 Materials and Methods

4.3.1 Strains, media, culture conditions, and chemicals.

Strains used in this study are listed in Table 4.1. *E. coli* strains were cultivated in lysogeny broth (LB) and on LB agar plates (28). *A. succinogenes* strains were cultivated in medium B (in g L^{-1}: yeast extract, 5; bactotryptone, 10; $NaH_2PO_4 \cdot H_2O$, 8.5; K_2HPO_4, 15.5; $NaHCO_3$, 2.1; and glucose, 9), AM3 defined medium (29), or Bacto brain heart infusion (BHI; Becton-Dickinson and Co., Franklin Lakes, NJ). AM3 is a phosphate-based medium containing glucose, NH_4^+, ten vitamins, and the amino acids glutamate, cysteine, and methionine. After natural transformations, *A. succinogenes* was grown in liquid AM2 (AM3 minus glutamate) in the absence or presence of isocitrate (30 mM; AM2-isocitrate) or on AM16-isocitrate or AM17 agar plates. AM17 is modified AM3 containing twice the amount of vitamins, cysteine, and methionine normally present in AM3, plus an amino acid supplement mix (in mg L^{-1}, final concentrations in AM17: alanine, 120; aspartate, 105; asparagine$\cdot H_2O$, 120; glycine, 115; histidine, 21; isoleucine, 58; leucine, 90; lysine$(HCl)_2$, 113; phenylalanine, 45; serine, 44; threonine, 56; tryptophan, 13; tyrosine, 28; and valine, 82). AM16 is AM17 without glutamate. *A. succinogenes* liquid cultures were grown in N_2-flushed 28-mL anaerobic tubes at 37°C, with shaking at 250 rpm. OD_{660} of *A. succinogenes* cultures was monitored using a Spectronic 20 (Bausch and Lomb, Rochester, NY). To isolate single colonies, *A. succinogenes* strains were grown on LB plates containing 10 g L^{-1} glucose (27) and 10 μg mL^{-1} kanamycin (Km, *A. succinogenes* is naturally resistant to Km). Amp was added to all growth media for plasmid maintenance when required (40 μg mL^{-1} for *A.*

succinogenes and 100 μg mL^{-1} for *E. coli*). All cultures of *A. succinogenes* on agar plates were grown under a CO_2-enriched atmosphere. Unless stated otherwise, chemicals were from Sigma-Aldrich (St. Louis, MO).

Table 4.1 Strains and plasmids used in this study

	Description	Source
E. coli		
K-12 (ATCC19020)	Wild-type strain	Laboratory collection
DH5α	F- φ80*lacZ*ΔM15 Δ(l*acZYA-argF*) U169 *rec*A1 *end*A1 hsdR17	Laboratory collection
TOP10	(rk-, mk+) *pho*A *sup*E44 λ- thi-1 *gyr*A96 *rel*A1 F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lacZ*ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G λ-	Invitrogen
A. succinogenes		
130Z (ATCC55618)	Wild-type strain	ATCC
Δfrd::icd	130Z derivative, contains the *AscI-FRT*-p_{pckA}-*icd-FRT-AscI* cassette in the *frdAB* deletion	This study
Δfrd	130Z derivative, *frdAB* deletion, contains one *FRT* site	This study
ΔpflB::*icd*	130Z derivative, contains the *AscI-FRT*-p_{pckA}-*icd-FRT-AscI* cassette in the *pflB* deletion	This study
ΔpflB	130Z derivative, *pflB* deletion, contains one *FRT* site	This study
ΔpflB-ΔlacZ::icd	*ΔpflB* derivative, *lacZ* deletion, still contains the *icd* cassette	This study
ΔpflB-ΔcitDEF	*ΔpflB* derivative, *citDEF* deletion	This study
ΔpflB-Δacn	*ΔpflB* derivative, *acn* deletion	This study
Plasmids		
pCR2.1	Amp^R, Km^R, *lacZα*, cloning vector	Invitrogen
pLGZ920	*E. coli-A. succinogenes* shuttle vector; Amp^R; *A. succinogenes* p_{pckA}	(8)
pCP20	pSC101 derivative, *flp+*, λ cI857+, λ p_R Rep^{ts}, Amp^R, Cm^R	(36)
pCR2.1-*icd*	pCR2.1 derivative, *E. coli icd* under control of *A. succinogenes* p_{pckA}, flanked by *FRT* repeats and *Asc*I restriction sites	This study
pCR2.1-Δ*frd*::*icd*	pCR2.1 derivative, *A. succinogenes* Δ*frd*::*FRT*-p_{pckA}-*icd-FRT*	This study
pCR2.1-Δ*frd*	pCR2.1 derivative, *frdAB* deletion (2.1-kb fusion product of frd_{up} and *frdCD*)	This study

Table 4.1 (cont'd)

pCR2.1-Δ*pflB*	pCR2.1 derivative, *pflB* deletion (2-kb fusion product of $pflB_{up}$ and $pflB_{down}$)	This study
pCR2.1-Δ*pflB*::*icd*	pCR2.1 derivative, *A. succinogenes* Δ*pflB*::*FRT*-p_{pckA}-*icd-FRT*	This study
pLGZ924	pLGZ920 derivative, *E. coli icd* under control of *A. succinogenes* p_{pckA} promoter	This study
pCV933	pLGZ920 derivative, *S. cerevisiae flp* under control of *A. succinogenes* p_{pckA}	This study
pCR2.1-Δ*lacZ*	pCR2.1 derivative, *lacZ* deletion (fusion product of $lacZ_{up}$ and $lacZ_{down}$)	This study
pCR2.1-Δ*lacZ*::*icd*	pCR2.1 derivative, *A. succinogenes* Δ*lacZ*::*FRT*-p_{pckA}-*icd-FRT*	This study
pCR2.1-Δ*cit*	pCR2.1 derivative, *citDEF* deletion (fusion product of $citDEF_{up}$ and $citDEF_{down}$)	This study
pCR2.1-Δ*cit*::*icd*	pCR2.1 derivative, *A. succinogenes* Δ*citDEF*::*FRT*-p_{pckA}-*icd-FRT*	This study
pCR2.1-Δ*acn*	pCR2.1 derivative, *lacZ* deletion (fusion product of acn_{up} and acn_{down})	This study
pCR2.1-Δ*acn*::*icd*	pCR2.1 derivative, *A. succinogenes* Δ*acn*::*FRT*-p_{pckA}-*icd-FRT*	This study

4.3.2 Plasmids, DNA manipulations, and electroporations.

Plasmids used in this study are listed in Table 4.1. PCR products were cloned into pCR2.1 using the TOPO-TA cloning kit (Invitrogen, Carlsbad, CA). pLGZ920 was used to express foreign genes in *A. succinogenes* under control of p_{pckA} (8). One Shot TOP10 chemically competent *E. coli* cells (Invitrogen) were transformed as described by the manufacturer. Electrocompetent *A. succinogenes* cells used for transformation of linear DNA were prepared as described (28). Electrocompetent *A. succinogenes* cells used for transformation of circular DNA were prepared using a method modified from (30) as follows. A 10-mL culture of actively growing *A. succinogenes* (OD_{660} 0.5 ± 0.2) in medium B or AM3 was incubated on ice for 10 min. Six mL of culture were harvested by centrifugation (3 min, 4,500 × *g*) at 4°C. The pellet was washed three times with 1 mL of cold 272 mM sucrose and finally resuspended in 100 µL of 272 mM sucrose. Electroporation was performed with 1 µL DNA (100−300 ng for circular DNA,

700−800 ng for linear DNA) and 100 μL electrocompetent cells in a 2-mm gap width electroporation cuvette (Bio-rad, Hercules, CA). After electroporation (settings: 25 μF, 400 Ω for plasmids, 600 Ω for linear DNA, and 2.5 kV on a Bio-Rad GenePulser™), electroporation mixtures were incubated with 0.5 mL of super optimal broth with catabolite repression (SOC) outgrowth medium (New England Biolabs, Ipswich, MA) for 1 h at 37°C. The cells were then centrifuged (3 min, 4,500 × *g*), resuspended in 100 μL supernatant, and spread on a single plate.

DNA manipulations used standard protocols (28). PCR reactions were performed with the Advantage HD polymerase kit (Clontech, Mountain View, CA) unless otherwise stated. Restriction enzymes were from New England Biolabs. Genomic DNA extractions were performed using the Wizard genomic DNA purification kit, plasmid DNA was purified using the Wizard SV miniprep kit, and DNA fragments were recovered from PCR mixtures and agarose gels using the Wizard SV gel and PCR Clean-Up System (Wizard kits from Promega, Madison, WI). Primers used in this work are listed in Table S4.1. Primers were synthesized by the Michigan State University Research Technology Support Facility (MSU RTSF) or by Integrated DNA Technologies (Coralville, IA). PCR and cloning accuracy were confirmed by DNA sequencing performed by the MSU RTSF or by GENEWIZ, Inc. (South Plainfield, NJ). Colony PCR was performed with Taq polymerase to confirm cloning steps.

4.3.3 Construction of plasmid pLGZ924.

The isocitrate dehydrogenase gene (*icd*) was amplified from *E. coli* K-12 genomic DNA using primers P1 and P2. The PCR product was cloned into pCR2.1, and then subcloned into the *Xba*I and *Sac*I sites of pLGZ920, downstream of p_{pckA}, yielding plasmid pLGZ924.

4.3.4 Construction of Δ*frd*::*icd* and Δ*pflB*::*icd* mutants.

The *E. coli icd* gene under control of p_{pckA} in pLGZ924 was amplified with primers P3 and P4. The primers added *S. cerevisiae FRT* sequences (5'-GAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTC-3') (20) in the same orientation plus *Asc*I restriction sites on each side of the p_{pckA}-*icd* cassette to yield a 1.6 kb, *AscI-FRT*-p_{pckA}-*icd-FRT-AscI,* cassette (Figure 4.1). After the PCR reaction, the 100-µL mixture was heated at 95°C for 20 min. 3' A-overhangs were added to the PCR product using Taq polymerase as described in the TOPO-TA cloning kit, before cloning into pCR2.1 to yield pCR2.1-*icd*.

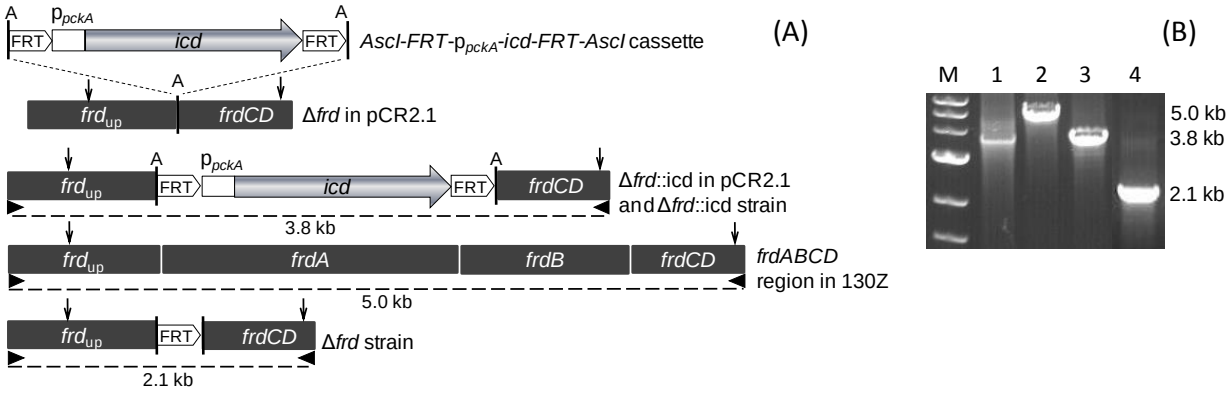

Figure 4.1 Construction of the *icd* selection cassette and construction of strain Δ*frd*. (A) Physical maps of the intermediary constructs and strains involved. A: *Asc*I; Black triangles: primers P9 and P10 used to amplify the PCR products shown in (B). (B) PCR verification of intermediate and final constructs using primers P9-P10. PCR product amplified from the 5.7-kb *Δfrd::icd* PCR fragment used for natural transformation (lane 1), from *A. succinogenes* 130Z (lane 2), from *A. succinogenes* Δ*frdAB::icd* (lane 3), and from *A. succinogenes* Δ*frd* (lane 4). Lane M: DNA molecular markers.

A Δ*frd* DNA construct was assembled in two steps. First, the 1.1-kb region upstream of *frdABCD* (frd_{up}) was amplified with primers P5 and P6, and a 1-kb fragment containing the *frdCD* genes was amplified with primers P7 and P8. Primers P6 and P7 overlapped by 19 nt,

including an *Asc*I site. Second, the 1.1- and 1.0-kb purified fragments were fused by PCR using nested primers P9 and P10. The 2.1-kb frd_{up}-*Asc*I-*frdCD* product (Figure 4.1) was cloned into pCR2.1, yielding pCR2.1-Δ*frd*. The *Asc*I-*FRT*-p_{pckA}-*icd*-*FRT*-*Asc*I fragment from pCR2.1-*icd* was cloned into the *Asc*I site of pCR2.1-Δ*frd*, yielding pCR2.1-Δ*frd::icd* (Figure 4.1). This plasmid was used as a template with primers P11 and P12 to amplify a 5.4-kb fragment that was naturally transformed into *A. succinogenes* 130Z to create strain *Δfrd::icd*.

A Δ*pflB* construct was assembled in pCR2.1 in two steps as well. Two 1.4-kb DNA fragments upstream ($pflB_{up}$) and downstream ($pflB_{down}$) of *A. succinogenes pflB* were amplified separately using primer pairs P13-P14 and P15-P16, respectively. Primers P14 and P15 overlapped by 21 nt, including an *Asc*I site. The $pflB_{up}$ and $pflB_{down}$ purified PCR fragments were then fused by PCR using nested primers P17 and P18. The 2.1-kb $pflB_{up}$-*Asc*I-$pflB_{down}$ product was cloned into pCR2.1, yielding plasmid pCR2.1-Δ*pflB*. Finally, the *Asc*I-*FRT*-p_{pckA}-*icd*-*FRT*-*Asc*I cassette from pCR2.1-*icd* was cloned into the *Asc*I site of pCR2.1-Δ*pflB*, yielding pCR2.1-Δ*pflB*::*icd*. After restriction digest with *Xmn*I, the 7.6-kb linear plasmid DNA was naturally transformed into *A. succinogenes* to create strain Δ*pflB::icd*.

4.3.5 Construction of the Δ*pflB* Δ*lacZ*::*icd*, Δ*pflB* Δ*cit*::*icd*, and Δ*pflB* Δ*acn*::*icd* double mutants.

Strain Δ*pflB* was used as the host to delete *lacZ*, the genes encoding citrate lyase (*citDEF*), and the aconitase gene (*acn*). The strategy used was identical to that used to build Δ*frd* and Δ*pflB*. Construction of plasmids pCR2.1-Δ*lacZ*::*icd*, pCR2.1-Δ*cit*::*icd*, and pCR2.1-Δ*acn*::*icd* used primers P35 to P40 (Δ*lacZ*), p41 to p46 (Δ*cit*), and P47 to P52 (Δ*acn*) (Table S4.1). Δ*cit* was flanked by 0.5 kb of its upstream and downstream regions in pCR2.1-Δ*cit*::*icd*. Δ*acn* was

flanked by 0.4 kb and 0.45 kb of its upstream and downstream regions, respectively, in pCR2.1-Δ*acn*::*icd*. Δ*lacZ* was flanked by 0.75 kb and 0.85 kb of its upstream and downstream regions, respectively, in pCR2.1-Δ*lacZ*::*icd*. These plasmids were used as the templates with primer pairs P11-P12 (add 1 kb of vector DNA on each side of the regions flanking the deletion) and P53-P54 (add 0.1 kb of vector DNA on each side of the regions flanking the deletion) to amplify the fragments used to naturally transform strain *ΔpflB.* Double recombinations in strain Δ*pflB* were screened by colony PCR using primer pairs P39-P40, P39-P26, and P25-P40 (Δ*lacZ*::*icd*); P45-P46, P45-P25, and P26-P46 (Δ*cit*::*icd*), and P51-P52, P51-P26, and P25-P52 (Δ*acn*::*icd*). Δ*pflB* Δ*lacZ*::*icd* double mutant candidates were further confirmed by patching the colonies on LB glucose plates containing 40 μL 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal).

4.3.6 Natural transformation.

The natural transformation protocol for *A. succinogenes* closely resembles those for *H. influenzae* (31) and *A. pleuropneumoniae* (32). To prepare competent cells, 400 μL of an overnight culture of *A. succinogenes* in BHI were transferred to 35 mL fresh BHI and grown to an OD_{600} of 0.2–0.25. Cells were washed twice with AM3 phosphate buffer (per liter: $NaH_2PO_4 \cdot H_2O$, 8.5 g; K_2HPO_4, 15.5 g), re-suspended in 10 mL of anaerobic MIV medium (31) containing 2 mM cAMP, and incubated at 37°C for 100 min with shaking at 100 rpm to induce competence. For natural transformation, a mixture of 1 μg DNA and 1 mL competent cells was incubated in a 37°C water bath for 25 min. A negative control without DNA was included. Two volumes of BHI were added to the transformation mixtures and incubated at 37°C with shaking (250 rpm) for 100 min. Cells were then harvested (4,500 × *g*, 4°C, 15 min), washed twice with 0.75 mL AM3 phosphate buffer, and resuspended in 0.75 mL AM3 phosphate buffer. Finally,

0.25 mL of cell suspension was used to inoculate one tube each of AM3 (positive control), AM2 (negative control), and AM2-isocitrate (selective medium). Tubes were incubated at 37°C with shaking until growth was observed in AM2-isocitrate. AM2-isocitrate cultures (0.25 mL) were streaked onto LB-glucose-Km plates to isolate putative recombinants. A method for plate selection is evaluated in the Results section.

4.3.7 Removal of the *icd* marker.

The *S. cerevisiae flp* gene was cloned into pLGZ920 for expression in *A. succinogenes.* The *flp* gene was amplified from plasmid pCP20 using primers P19 and P20. The purified PCR product was cloned into the *Xba*I and *Sac*I sites of pLGZ920 using the In-Fusion cloning system (Clontech). In the resulting plasmid, pCV933, *flp* is expressed constitutively under control of *A. succinogenes* p_{pckA}. pCV933 was purified and electroporated into *A. succinogenes* strains Δ*frd::icd,* Δ*pflB::icd*, Δ*pflB* Δ*cit::icd*, and Δ*pflB* Δ*acn::icd.* Transformants were screened for excision of the *icd* marker by colony PCR using several primer pairs. Excision of *icd* yielded strains Δ*pflB*(pCV933), Δ*frd*(pCV933), Δ*pflB* Δ*cit*(pCV933), and Δ*pflB* Δ*acn*(pCV933), each of which contains a single *FRT* scar in the chromosome. The deletions and the presence of the 48-nt FRT sequence framed by two *Asc*I sites were confirmed by sequencing these regions in the genomes of strains Δ*frd* and Δ*pflB.*

4.3.8 Curing of plasmid pCV933.

An overnight culture of *A. succinogenes* Δ*frd*(pCV933) or Δ*pflB*(pCV933) in medium B was used to inoculate a series of medium B tubes containing 0, 10, 50, 100, and 200 μg mL^{-1} acridine orange (AO). After 6.5 h of growth (37°C), cultures from each tube were streaked onto

LB-glucose-Km plates. The resulting colonies were patched onto LB-glucose-Amp plates to screen for the loss of pCV933. Amp-sensitive (Amp^s) colonies were screened for plasmid curing by colony PCR using primers P21 and P22, specific for sequences upstream and downstream of the pLGZ920 multiple cloning site.

4.3.9 Enzyme assays.

Enzyme assays were performed with cell extracts prepared from exponential phase cultures (OD_{660} between 0.6 and 1.0) grown in AM3. Cells from 10-mL cultures were harvested by centrifugation (4,500 × *g*, 15 min, 4˚C) and washed three times with 50 mM Tris-HCl (pH 7.4) at room temperature. For ICD and citrate lyase assays, pellets were resuspended in 0.5 mL of 50 mM Tris-HCl (pH 7.4). For fumarate reductase assays pellets were resuspended in 1 mL of 50 mM potassium phosphate (pH 7.0) containing 0.1 mM dithiothreitol. Cells were sonicated with a Branson Sonifier 450 (Danbury, CT; 50% duty cycle, level 2, 20 sec, 4 times). The lysate was centrifuged and the supernatant used for the assays as crude cell extract. Total proteins were quantified using the Bio-Rad (Hercules, CA) Protein Assay Dye Reagent Concentrate with bovine serum albumin as the standard. One unit of enzyme activity was defined as the amount of enzyme needed to convert 1.0 μmol of substrate into product per min. ICD and citrate lyase activities were assayed using a Cary 300 UV/vis spectrophotometer (Varian Instruments, Walnut Creek, CA) equipped with a Peltier system. ICD activity was assayed by monitoring $NADP^+$ reduction at 340 nm in a 1-mL reaction mixture containing 50 mM Tris-HCl (pH 7.4), 5 mM DL-isocitric acid trisodium salt, 5 mM $MgCl_2$, 1 mM dithiothreitol, 0.3 mM $NADP^+$, and 90 mM NaCl. The reaction was started by adding 50 μL crude cell extract. The molar extinction coefficient of NADPH was 6,200 $M^{-1}cm^{-1}$. Fumarate reductase activity was assayed in 96-well

plates in a Sunrise microplate reader (Tecan, Durham, NC) in an anaerobic glove bag. The 200-μL reaction contained 50 mM potassium phosphate (pH 7.0), 5 mM benzyl viologen, 0.4 mM sodium dithionite, and 5 mM sodium fumarate. The reaction was started by adding 10 μL cell extract. The activity was measured at 595 nm as the difference between the slopes before and after adding the cell extract. Measurements were taken every 30 sec with shaking between measurements. All enzyme assays were performed at 37˚C on three biological replicates.

4.3.10 Analysis of fermentation products.

Overnight cultures of strains *130Z* and Δ*pflB* in AM3-25 mM $NaHCO_3$ were inoculated (0.25 mL) into 10 mL of fresh AM3-25 mM $NaHCO_3$. Fermentation products in culture supernatants were quantified by high-performance liquid chromatography (HPLC) (Waters, Milford, MA) using a 300 × 7.8 mm Aminex HPX-87H column (Bio-Rad, Hercules, CA) at 30°C with 4 mM H_2SO_4 as the mobile phase (0.6 mL/min flow rate). Organic acids were detected with a Waters 2478 UV detector at 210 nm or 254 nm.

4.4 Results

4.4.1 A positive selection method for recombination events in *A. succinogenes*.

To avoid using antibiotic resistance genes as selection markers, a positive selection method was developed based on *A. succinogenes*'s metabolism. Because it is missing citrate synthase and ICD in the citric acid cycle, *A. succinogenes* is auxotrophic for glutamate (29). *A. succinogenes* has glutamate dehydrogenase activity and is able to grow on α-ketoglutarate instead of glutamate (29). We proposed that *A. succinogenes* expressing ICD would be able to grow in minimal medium with isocitrate substituted for glutamate.

A. succinogenes 130Z(pLGZ924) could grow in AM2-isocitrate, whereas 130Z could not (data not shown). Strain 130Z(pLGZ924) grew slower in AM3 and AM2-isocitrate, (generation times of 2.08 ± 0.03 h and 5.7 ± 0.1 h, respectively) than 130Z did in AM3 (generation time of 1.66 ± 0.01 h). AM3 contains 1.4 mM glutamate. Even with 5.7 mM L-isocitrate (4-fold more than glutamate in AM3), 130Z(pLGZ924) grew in AM2-isocitrate to a maximum OD_{660} 40% lower than in AM3. In contrast to glutamate, which can only be a precursor of biomass in 130Z, isocitrate can also be metabolized to oxaloacetate and acetyl-CoA through aconitase and citrate lyase, and it seems to be partially diverted to acetate, explaining why 130Z(pLGZ924) grows slower and to a lower final OD in AM2-isocitrate than in AM3. While the *icd* gene looks to be a promising positive selection marker for *A. succinogenes,* its introduction in *A. succinogenes* increases the complexity of the central metabolism, suggesting that the *icd* gene should be excised from knockout mutant strains.

4.4.2 Natural transformation of *A. succinogenes* using a Δ*frd*::*icd* construct.

Two USSs were present in the Δ*frd*::*icd* construct, 650 nt upstream (in frd_{up}) and 900 nt downstream (in *frdCD*) of the *FRT*-p_{pckA}-*icd*-*FRT* cassette, respectively (Figure 4.1). *Asc*I restriction sites were created on both ends of the *FRT*-p_{pckA}-*icd*-*FRT* cassette and in the center of the Δ*frd* construct for easy insertion of the cassette into the Δ*frd* construct. The *A. succinogenes* chromosome contains only 33 *Asc*I sites, allowing our selection cassette to be used for almost any gene deletion in the genome. DNA starts being degraded by exonucleases as soon as it enters the cell. To protect the *A. succinogenes* sequences (~1 kb on each side of the selection marker) from degradation and maximize the chances of recombination, the PCR product used to construct strain *Δfrd::icd* contained 1 kb of pCR2.1 DNA on each side of the knockout construct.

To determine whether and when cAMP is needed to induce *A. succinogenes* competence, different growth and competence induction conditions were tested. cAMP was completely omitted in the first experiment. In the second experiment, 2 mM cAMP was added to BHI only. In the third experiment, 2 mM cAMP was added to MIV only. In each of these experiments the transformation mixtures were finally incubated in AM3, AM2-isocitrate, and AM2 liquid media. Only the third experiment led to growth in AM2-isocitrate after four days. Cultures grew overnight in AM3 (positive control) but no growth was observed after four days in AM2 (negative control). The four-day AM2-isocitrate culture (0.25 mL) was used to inoculate 10 mL of fresh AM2-isocitrate to enrich for recombinants able to grow on isocitrate. This second culture was streaked onto LB-glucose-Km plates to isolate putative recombinants.

4.4.3 Confirmation of the Δ*frd*::*icd* knockout strain.

Double recombination of the *Δfrd::icd* construct in the *A. succinogenes* chromosome was confirmed by colony PCR. Of six putative recombinant colonies screened by PCR with primers P9 and P10, one yielded the expected 3.8-kb fragment (Figure 4.1). This putative mutant was confirmed using colony PCR screens with other primers. PCR with primers P11 and P12 confirmed that this colony did not contain any pCR2.1 sequences, while PCRs with primer pairs P23-P24 (internal to *frdB*) and P25-P26 (internal to *E. coli icd*) confirmed that this colony no longer contained the *frdB* gene but that it contained the *icd* gene (not shown). This mutant strain was called Δ*frd::icd.*

Construction of the Δ*frd::icd* mutant was confirmed with *in vitro* ICD and fumarate reductase activity assays. Strain 130Z(pLGZ924) was the positive control for ICD activity. *A. succinogenes* 130Z showed almost no detectable ICD activity (0.30 ± 0.49 mU mg^{-1}). With an

activity level of 28 ± 7 mU mg^{-1} Δ*frd::icd* showed over 90-fold higher ICD activity than 130Z, and 13-fold lower activity than 130Z(pLGZ924) (360 ± 130 U/mg). Higher ICD activity in 130Z(pLGZ924) reflects the fact that this strain contains multiple copies of *icd.* Strain Δ*frd::icd* lacks the *frdAB* genes and should be devoid of fumarate reductase activity. Indeed, Δ*frd::icd* showed over 500-fold lower fumarate reductase activity (0.0078 ± 0.0011 U mg^{-1}) than 130Z (3.97 ± 0.08 U mg^{-1}).

4.4.4 Excision of the selection marker.

Strain Δ*frd*::*icd* was transformed with pCV933 (expressing the *S. cerevisiae* FLP recombinase) and plated on LB-glucose-Amp. One transformant colony was re-isolated on LB-glucose-Amp and excision of the *icd* marker was tested by colony PCR using primers P25-P26. Out of the twenty potential Δ*frd* colonies tested, one yielded no PCR product. Colony PCR of this strain using primers P9-P10 yielded a 2.1-kb product instead of the 3.8-kb product obtained with Δ*frd*::*icd*, confirming that the *icd* gene was no longer present in that strain (Figure 4.1). The new strain was called Δ*frd*(pCV933).

4.4.5 Curing pCV933 from *A. succinogenes* Δ*frd*(pCV933).

AO is commonly used to inhibit plasmid replication (33). After AO treatment, colonies isolated on LB-glucose-Km were replica-plated onto LB-glucose-Km and LB-glucose-Amp plates to identify Amp^s colonies. Twenty, ten, and ten Amp^s colonies from the Δ*frd*(pCV933) cultures grown with 50 μg ml^{-1}, 75 μg ml^{-1}, and 200 μg ml^{-1} AO, respectively, were screened by colony PCR with primers P21-P22. A single colony originating from the 50 μg ml^{-1} AO culture was confirmed to have lost the plasmid (not shown), and was called strain Δ*frd*.

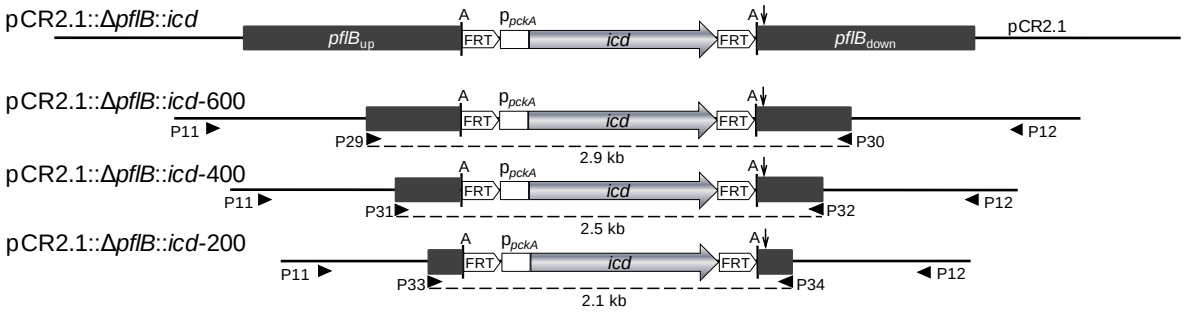

Figure 4.2 Physical maps of pCR2.1-Δ*pflB*::*icd* and its truncated constructs
pCR2.1::Δ*pflB*::*icd*-600, pCR2.1::Δ*pflB*::*icd*-400, and pCR2.1::Δ*pflB*::*icd*-200. A: *Asc*I; black triangles: primers; vertical arrows: USS sequences.

4.4.6 Construction of strain Δ*pflB*.

The Δ*pflB*::*icd* construct contained a single USS in the $frdB_{down}$ fragment, 33 nt downstream of the *FRT*-p_{pckA}-*icd*-*FRT* cassette (Figure 4.2, top construct). *Xmn*I-linearized pCR2.1-*ΔpflB::icd*, containing 2 kb of pCR2.1 DNA on each side of the knockout construct, was used to construct strain *ΔpflB::icd.*. Construction of the *A. succinogenes ΔpflB::icd* mutant was confirmed by colony PCR with primers P27-P28, which match sequences flanking the *pflB* deletion. Thirty-five of the thirty-six colonies tested yielded the expected 2-kb PCR fragment (Figure 4.3). PCR with primers P25-P26 confirmed that the putative mutants contained the *E. coli icd* gene (Figure 4.3). One isolated mutant was called *ΔpflB::icd.* The *icd* marker was excised by transforming strain *ΔpflB::icd* with pCV933. Of the thirty-eight colonies tested by PCR with primers P25-P26, eighteen yielded no PCR product. Excision of the *icd* gene was confirmed by colony PCR using primers P27-P28 (Figure 4.3), and one confirmed isolate was called Δ*pflB*(pCV933).

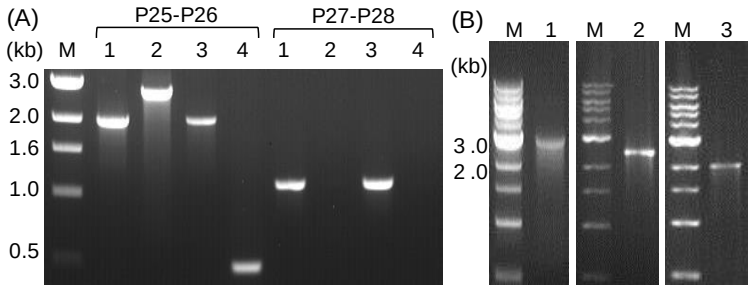

Figure 4.3 Construction of strain Δ*pflB* (A) and natural transformation with the truncated products Δ*pflB*::*icd*-600, Δ*pflB*::*icd*-400, and Δ*pflB*::*icd*-200 (B). (A) Gel electrophoresis of the PCR products amplified from the Δ*pflB::icd* cassette (lane 1), from *A. succinogenes* 130Z (lane 2), from an *A. succinogenes* Δ*pflB::icd* derivative (lane 3), and from a Δ*pflB*(pCV933) derivative (lane 4) using primers P25-P26 (internal to *icd*) and P27-P28 (framing the *pflB* deletion). Lane M: DNA molecular markers. (B) Examples of colony PCRs after transformation with the truncated constructs Δ*pflB*::*icd*-600 (lane 1, PCR with primers P29-P30), Δ*pflB*::*icd*-400 (lane 2, PCR with primers P31-P32), and Δ*pflB*::*icd*-200 (PCR with primers P33-P34). Lane M: DNA molecular markers.

Strain Δ*pflB*(pCV933) was cured of pCV933 by AO treatment. Ten and fifty μg mL^{-1} AO produced four and ten Amp^s colonies, respectively. The 100 μg mL^{-1} and 200 μg mL^{-1} AO treatments produced no Amp^s colonies. Of the fourteen Amp^s colonies, one from the 10 μg mL^{-1} AO and two from the 50 μg mL^{-1} AO treatments were shown to have lost the plasmid after two replicate colony PCR screens with primers P21 and P22 (not shown). One cured derivative from the 50 μg mL^{-1} AO treatment was called strain Δ*pflB*.

During growth in AM3-25 mM $NaHCO_3$, conditions that normally favor formate production (6, 29), Δ*pflB* did not produce any formate (Figure 4.4), confirming that this strain is devoid of pyruvate formate lyase activity.

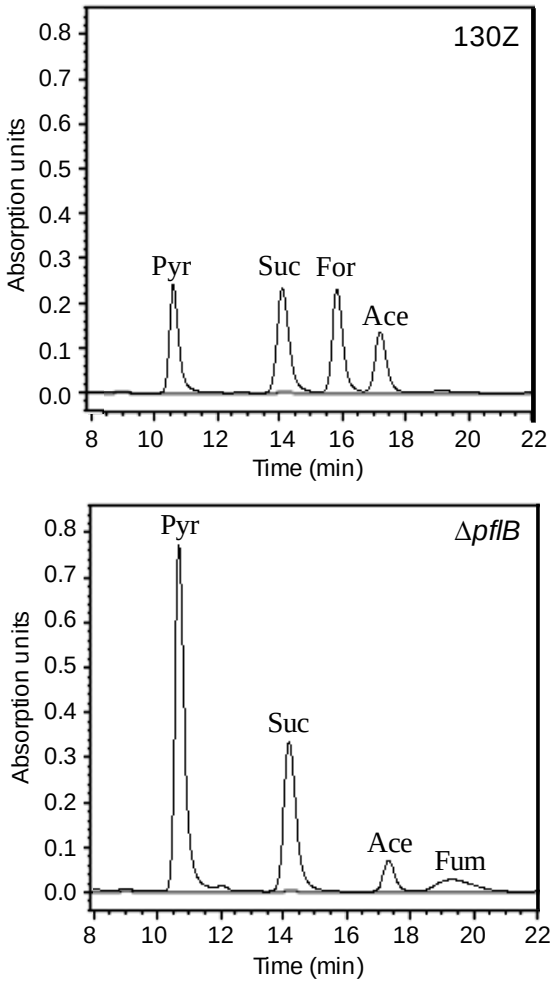

Figure 4.4 UV spectra of the HPLC profiles of fermentation supernatants of *A. succinogenes* 130Z, and Δ*pflB* grown on AM3 in the presence of 25 mM NaHCO3. Supernatant samples were collected immediately after inoculation (baseline) and after 12 h growth. Ace: acetate; For: formate; Fum: fumarate; Pyr: pyruvate.

4.4.7 Effects of the DNA construct length and of the length of the homologous regions on the efficiency of homologous recombination.

The DNA fragments used in natural transformation so far contained approx. 1 kb of *A. succinogenes* DNA on each side of the selection marker for efficient recombination plus at least 1 kb of non-homologous plasmid DNA on each side to protect the homologous DNA from degradation by exonucleases. To test whether the non-homologous plasmid DNA on each side of

the recombination cassette could be omittted, the 3.8-kb Δ*frd::icd* cassette (amplified from pCR2.1-Δ*frd::icd* with primers P9-P10, Figure 4.1) was used in natural transformation. In this experiment, the selection marker was flanked by approx. 1 kb of *A. succinogenes* DNA on each side. Transformants grew in AM2-isocitrate. Forty percent of recombinant colonies screened by PCR with primers P9 and P10 yielded the expected 3.8-kb fragment (Table 4.2). This result indicates that no more than a total of 1 kb of DNA is needed on each side of the marker for recombination to take place.

To determine what minimum length of homologous DNA is needed on each side of the selection marker to allow double recombination to take place, homologous recombination cassettes were generated that contained 200 bp, 400 bp, and 600 bp homologous DNA on each side of the selection marker. pCR2.1-Δ*pflB*::*icd* was used as the template because truncating the frd_{up} and *frdCD* fragments would have deleted both USS sequences (Figure 4.1). The Δ*pflB*::*icd*-600, -400 and -200 fragments were amplified from pCR2.1-Δ*pflB*::*icd* with primers P29-P30, P31-P32, and P33-P34, respectively, and cloned back into pCR2.1 (Figure 4.2). The fragments used for transformation were generated by PCR with primers P11-P12, including 1 kb of vector DNA on each side of the recombination cassette. Three independent natural transformations were performed with each fragment, in which 15 colonies were screened by PCR. While the standard deviations were large in all cases, averages of 47%, 57%, and 60% colonies contained the Δ*pflB*::*icd* deletion for the transformations with fragments Δ*pflB*::*icd*-600, -400, and -200, respectively. These results suggest that 200 bp of homologous DNA on each side of the selection marker are still enough to allow double recombination. Conversely, all transformations using the Δ*pflB*::*icd*-600 cassette amplified with primers P30-P31, the Δ*pflB*::*icd*-400 amplified with

primers P31-P32, or the Δ*pflB*::*icd*-200 cassette amplified with primers P33-P34 yielded no growth in AM2-isocitrate and no Δ*pflB*::*icd* deletions.

4.4.8 Construction of double knockout mutants.

To determine whether the same knockout strategy could be used to introduce two successive deletions in the same strain, three genes deemed non-essential for growth in AM3 glucose were targeted for deletion in strain Δ*pflB.* The genes encoding citrate lyase and aconitase were chosen as targets, because the tricarboxylic acid cycle is incomplete in *A. succinogenes*. The third target gene was *lacZ*. The strategy used to build the three deletions was identical to that used to build the Δ*frd* and Δ*pflB* strains. Citrate lyase is encoded by three genes organized in an operon, *citDEF* (Asuc_1194-1196). CitE encodes the subunit with lyase activity. The citrate lyase deletion, Δ*cit*, encompassed over half of *citD*, the entire *citE*, and over a third of *citF* (not shown). One USS was present in the upstream fragment. The 2.6-kb aconitase gene (*acn*, Asuc_0185) is immediately preceded by the only USS in that area. To preserve that USS and to preserve the promoter of the gene downstream of *acn*, the *acn* deletion encompassed only 1.6 kb inside the *acn* gene. The *lacZ* deletion encompassed the entire 3.0-kb *lacZ* gene, including the only four USSs present in the area. For this reason, a USS was introduced into the reverse nested primer used to fuse the *lacZ* up and down regions (P40, Table S4.1).

The three double knockout mutant strains were obtained with frequencies ranging between 25% and 100% (Table 4.2). Note that double recombinants were obtained with frequencies of 25% or higher for Δ*cit*::*icd* in strains Δ*pflB* and 130Z and for Δ*acn*::*icd* in Δ*pflB* with only 0.52 kb to 0.6 kb of total flanking DNA on each side of the selection marker (Table 4.2).

Table 4.2 Frequencies of knockout mutations introduced by natural transformation in *A. succinogenes* strains 130Z and Δ*pflB*

Transformation cassette	Host strain	Homologous DNA on each side of the selection cassette (kb)	Vector DNA on each side of the recombination cassette (kb)	Knockout frequency (positive colonies/colonies screened)	Knockout frequency (%)
Δfrd::icd	130Z	1	0.8	1/6	17
Δfrd::icd	130Z	1	0	6/15	40
Δfrd::icd	130Z	1	0.8	No growth[a]	–
ΔpflB::icd	130Z	1	1.9	35/36	97
Δ*pflB*::*icd*-600	130Z	0.6	1.0	6/16, 15/15, 6/15	60 ± 35
Δ*pflB*::*icd*-400	130Z	0.4	1.0	4/15, 7/15, 15/15	57 ± 38
Δ*pflB*::*icd*-200	130Z	0.2	1.0	3/15, 4/15, 14/15	47 ± 41
Δ*pflB*::*icd*-600	130Z	0.6	0	No growth	–
Δ*pflB*::*icd*-400	130Z	0.4	0	No growth	–
Δ*pflB*::*icd*-200	130Z	0.2	0	No growth	–
Δ*cit*::*icd*	130Z	0.5	1.0	5/12	42
Δ*cit*::*icd*	130Z	0.5	0.1	3/12	25
Δ*lacZ*::*icd*	*ΔpflB*	0.8	1.0	17/19	89
Δ*cit*::*icd*	*ΔpflB*	0.5	1.0	30/30	100
Δ*cit*::*icd*	*ΔpflB*	0.5	0.1	6/12	50
Δ*acn*::*icd*	*ΔpflB*	0.42	0.1	4/16	25

[a] Experiment repeated five times with the same results.

Figure 4.5 confirms that Δ*pflB* Δ*lacZ*::*icd* transformants no longer have β-galactosidase activity. Δ*pflB* Δ*cit* showed 27-fold less in vitro citrate lyase activity than Δ*pflB* (0.02 μmol min^{-1} mg^{-1} in Δ*pflB* Δ*cit* vs. 0.55 μmol min^{-1} mg^{-1} in Δ*pflB*), confirming the double deletion.

Δ*pflB* Δ*pflB*-Δ*lacZ*::*icd*

Figure 4.5 Phenotypic characterization of double knockout mutant Δ*pflB*-Δ*lacZ*::*icd*. One colony each of Δ*pflB* and one Δ*pflB*-Δ*lacZ*::*icd* were patched on LB X-gal Amp plates.

4.4.9 Development of a selective solid medium.

A. succinogenes 130Z grows slowly on AM3 plates, with colonies visible only after 4 days. Even though it contains multiple copies of the *E. coli icd* gene, 130Z(pLGZ924) did not form visible colonies on AM2-isocitrate plates. To develop a selective agar-isocitrate medium that allows strains containing a single copy of *E. coli icd* to form colonies, all amino acids but glutamate, glutamine, arginine, and proline (i.e., amino acids derived from glutamate) were added to the medium. Glutamate was replaced by 30 mM DL-isocitrate. 130Z(pLGZ924) formed colonies on this medium after two days, but Δ*frd::icd* did not, even after six days. The same medium containing additional vitamins, cysteine, and methionine (i.e., medium AM16-isocitrate) supported growth of Δ*frd::icd,* with colonies forming after five days. AM16-isocitrate plates were tested as selective solid medium in two replicate natural transformations of strain 130Z with 1 μg *Nco*I-linearized pCR2.1-Δ*frd::icd*. Transformation mixtures were inoculated into liquid AM3 and AM2-isocitrate (250 μL in each), and spread on AM16-isocitrate plates (225 μL spread at 10^0, 10^{-1}, and 10^{-2} dilutions). As expected, transformants grew overnight in AM3 and after four days in liquid AM2-isocitrate. Colonies grew on AM16-isocitrate plates after five days, with an efficiency of approximately 140 colonies/μg DNA. Eighteen colonies grown on AM16-isocitrate plates were tested by colony PCR using primers P9-P2. All were Δ*frd::icd*

mutants, confirming that AM16-isocitrate plates can be used to directly select *icd*-containing knockout mutant constructs in *A. succinogenes*.

All attempts to measure transformation and double recombination efficiencies failed because of a yet unexplained problem with serial dilutions and plating. Starting from fresh dense precultures, all *A. succinogenes* strains tested, including 130Z, grew as lawns when plated from the undiluted preculture, they grew as numerous compact colonies when plated from the 10^{-1} dilution, but they produced only a few colonies when plated from the 10^{-2} dilution and no colonies at higher dilutions. Similar results were observed when the serial dilutions were prepared with AM3 phosphate buffer, AM3 with and without glucose, LB, or with the supernatant of the preculture, and when the bacterial suspensions were plated on LB-glucose, AM3, AM17, or AM16-isocitrate.

4.4.10 Knockout mutants via electroporation.

The knockout mutants obtained using natural transformation showed that linear DNA fragments could recombine with the *A. succinogenes* chromosome. Despite previous failed attempts at constructing knockout mutants using electroporation and antibiotic resistance genes as selection markers, electroporation was tested again as an alternative transformation method to create knockout mutants in *A. succinogenes*.

Electroporations of strain 130Z with *Nco*I-linearized pCR2.1-Δ*frd::icd* were performed using 0.1–1 μg DNA in 100 ng increments. After recovery, electroporation mixtures were washed once with MIV and resuspended in MIV. One third of each electroporation mixture was inoculated in liquid AM3, one third was inoculated in liquid AM2-isocitrate, and one third was spread on a single AM16-isocitrate plate. Electroporations with 300 ng or more linearized pCR2.1-Δ*frd::icd*

consistently yielded colonies on AM16-isocitrate plates. All colonies growing on AM16-isocitrate plates that were tested by colony PCR with primers P9-P2 were confirmed as Δ*frd::icd* knockouts, showing that electroporation can be used as an alternative to natural transformation to introduce linear DNA in *A. succinogenes* for homologous recombination.

4.5 Discussion

The fermentative metabolism of *A. succinogenes* has been studied in detail using fermentation balances, *in vitro* enzyme activity assays, genome sequencing, and metabolic flux analysis, and how ATP is produced and NAD^+ is regenerated in the different fermentative branches is well understood (6, 17, 21, 22, 29). To enhance the production of a given fermentative metabolite (e.g., fumarate or succinate), we needed a means to produce knockout mutants. Here, we demonstrated that natural transformation and electroporation can be used to introduce DNA in *A. succinogenes* for recombination, that the *E. coli icd* gene can be used as a positive selection marker for recombination events, that the yeast Flp/*FRT* recombination system can be used to excise the selection marker, that plasmids can be cured from *A. succinogenes* using AO, and that the selection marker can be reused to introduce at least two consecutive deletions into the same strain. The 64-nt scar (FRT sequence flanked by *Asc*I sites) interrupting the deletion does not by itself leave the remaining sequences in frame, but if an in-frame deletion is desired, the deletion itself can be designed with a frameshift to place the remaining sequences back in frame.

Except for our first transformation to construct the *Δfrd::icd* strain, at least 25% of colonies obtained after all other natural transformations were knockout mutants (Table 4.2). The frequency of knockout colonies increased to 100% when transformants (obtained by

electroporation) were selected directly on the enriched defined medium, AM16-isocitrate. Note that we could not repeat the natural transformation with the Δ*frd::icd* cassette in at least five more attempts, suggesting that the efficiency of the method is somewhat gene-specific. The sequences of the two USS repeats present in the Δ*frd::icd* cassette were closer to the consensus USS for *A. succinogenes* than the single USS in the Δ*pflB::icd* cassette was (not shown). Therefore poor recognition of the USSs flanking the Δ*frd::icd* cassette by the DNA uptake machinery cannot explain these results. We have also tried repeatedly to knockout other genes (i.e., *ackA*, *pta*, *pykA*, and *zwf*) in 130Z and Δ*pflB* without success, but these genes might be essential for growth, or we have not yet found conditions in which a knockout strain can grow.

With the *ΔpflB::icd*, Δ*lacZ*::*icd*, Δ*cit*::*icd*, and Δ*acn*::*icd* constructs, we demonstrated that a single USS in the knockout construct is enough to allow DNA uptake by *A. succinogenes*. Introducing a USS in the Δ*lacZ*::*icd* construct—which would otherwise not have contained any USS, also allowed double recombination to take place. Linear constructs of varying lengths could be used for transformation, from the entire plasmid carrying the knockout construct (e.g., Δ*pflB::icd* in 130Z) down to recombination cassettes containing 0.52 kb to 0.6 kb of DNA on each side of the selection marker (e.g., Δ*acn*::*icd* in Δ*pflB*) (Table 4.2). While flanking regions of 0.5 kb to 0.6 kb on each side of the cassette yielded knockouts with frequencies as low as 25% (or even 0% for Δ*pflB::icd*-600), 1 kb of flanking DNA on each side was enough to protect against exonucleases prior to recombination, routinely yielding knockout frequencies between 42% and 100% (Table 4.2). As little as 200 bp of homologous DNA on each side of the selection marker was enough for efficient double recombination, as long as the homologous DNA was itself flanked by non-homologous DNA on each side.

The development of a markerless knockout method for *A. succinogenes* will greatly facilitate future genetic engineering of *A. succinogenes*. Strain Δ*pflB* represents an important first step in engineering a strain that can produce succinate at near maximum theoretical yields.

4.6 Acknowledgements

This work was supported by a grant from the Michigan Economic Development Corporation, by MSU startup funds, and by grant # 2010-04061 from the USDA National Institute for Food and Agriculture's Sustainable Bioenergy Research Program to CV. BS was supported in part by a research fellowship from the MSU Quantitative Biology Initiative. NM was supported in part by a recruiting fellowship from the MSU Plant Research Laboratory. KT was supported by a Khorana scholarship. We thank Reena Jain, Jean Kim, and Maeva Bottex for their technical assistance, as well as Dr. Michael Bagdasarian for his critical reading of the manuscript.

APPENDIX

Table A4.1 Oligonucleotide primers used in this study

Primer	Sequence[a] (restriction site)	Specificity[b] (direction[c])
P1	GTTCTAGAGATGGAAAGTAAAGTAGTTG (*Xba*I)	*E. coli* K-12 *icd* gene (F)
P2	CTGAATTCATTACATGTTTTCGATGATC (*Eco*RI)	*E. coli* K-12 *icd* gene (R)
P3	GAGGCGCGCCGAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAA AGTATAGGAACTTCTCGATAAATTGAAAATGCAGCAA (*Asc*I)	*Asc*I-*FRT*- p_{pckA} (F)
P4	CTGGCGCGCCGAAGTTCCTATACTTTCTAGAGAATAGGAACTTC GGAATAGGAACTTCTTACATGTTTTCGATGATCGC (*Asc*I)	*E. coli icd*-*FRT*-*Asc*I (R)
P5	TTAGGTACCGGCAACAAAGG	frd_{up} (F)
P6	CCCCTATGCTTGGCGCGCCCGTAAACCTAAACGCGCAAT (*Asc*I)	frd_{up} (R)
P7	GGCGCGCCAAGCATAGGGGGAAAGCAAT (*Asc*I)	*frdCD* (F)
P8	GTCCGATTTGGGTTTTGCTA	*frdCD* (R)
P9	GTAAGCTTACGGCAAACACGATCACATA	Δ*frd*-fusion, nested (F)
P10	CACTCGAGGCACCGCCTGTCACTAAAAT	Δ*frd*-fusion, nested (R)
P11	CCGGATCAAGAGCTACCAAC	pCR2.1 backbone (F)
P12	CGAAACGATCCTCATCCTGT	pCR2.1 backbone (R)
P13	CGTTAACCGTGGGAATCAGT	$pflB_{up}$ (F)
P14	TTACGTTACCCCAGGCGCGCCCTTCCTTTTGCTAGTATTGATAAT GA (*Asc*I)	$pflB_{up}$ (R)
P15	GGCGCGCCTGGGGTAACGTAATAAAAATGTAATG (*Asc*I)	$pflB_{down}$ (F)
P16	TCTCTCCTTCGCGGAATAAA	$pflB_{down}$ (R)
P17	TGAGCCTGACTGGTAAATCCA	Δ*pflB*-fusion, nested (F)
P18	CACATCGACCCCGATAACTT	Δ*pflB*-fusion, nested (R)
P19	ATGAGGTGATCTAGATGCCACAATTTGGTATATTATGTAAA (*Xba*I)	*S. cerevisiae flp* (F)
P20	CGGCCAGTGAATTCGAGCTCTTATATGCGTCTATTTATGTAGG (*Sac*I)	*S. cerevisiae flp* (R)
P21	CGTTGTAAAACGACGGCC	pLGZ920 (F)
P22	AATTTTAAATTATCAATGAGGTG	pLGZ920 (R)
P23	TGCGTTACAACCCTGAAACA	*frdB* (F)
P24	TCTTTCGCACTTTCCAGCTT	*frdB* (R)
P25	GAAGGTGATGGAATCGGTGT	Inside *E. coli* K-12 *icd* (F)
P26	GTTACGGTTTTCGCGTTGAT	Inside *E. coli* K-12 *icd* (R)
P27	AATTTGCGGATCTGGATGTC	217 nt upstream of *pflB* (F)
P28	TGGTGTATTCCGTCAGTTCG	146 nt downstream of *pflB* (R)
P29	GTTGCGAACTGTTCACCTC	Δ*pflB*::*icd*-600 (F)
P30	CAATATCGCGGCGTTCCAGC	Δ*pflB*::*icd*-600 (R)
P31	GCATTGGGCATTTTGTGTAAC	Δ*pflB*::*icd*-400 (F)
P32	CTAGTAGAAATGCGCGTTTTA	Δ*pflB*::*icd*-400 (R)
P33	GAGTTTTGGCAGGCAATC	Δ*pflB*::*icd*-200 (F)
P34	CTTTCCCAGCGTTCCACAAAAAC	Δ*pflB*::*icd*-200 (R)
P35	GGATAAACTCGGCGTACG	$lacZ_{up}$ (F)
P36	GCGGTTATGGGCGCGCCGCAGAATCATGCTGAACTCC (*Asc*I)	$lacZ_{up}$ (R)
P37	GGCGCGCCCATAACCGCATAAAAATACAGGGCAG (*Asc*I)	$lacZ_{down}$ (F)
P38	CATGAAACCGAACAGTATTGTGC	$lacZ_{down}$ (R)
P39	CTTGCCAAACCGACCGAAAG	Δ*lacZ*-fusion, nested (F)

Table A4.1 (cont'd)

P40	TATTGATAATGAAAATCCG**ACCGCACTT**GGCAGTACCGGCGTAT TCCTC[d]	Δ*lacZ*-fusion, nested (R)
P41	CTTGACGGTAGCAGTGATC	*citDEF*$_{up}$ (F)
P42	GGCGCGCCGAATTGCTTTCCGAC (*Asc*I)	*citDEF*$_{up}$ (R)
P43	GAAAGCAATTCGGCGCGCCAATCAGTATCGGCCAGG (*Asc*I)	*citDEF*$_{down}$ (F)
P44	CCGTATCGCAGTGACCGC	*citDEF*$_{down}$ (R)
P45	CTGGTGGTGACGATCATAC	Δ*citDEF*-fusion, nested (F)
P46	GGTATCGATTTCCAATGCGG	Δ*citDEF*-fusion, nested (R)
P47	CGGCGAAGTGTTCTACGATG	*acn*$_{up}$ (F)
P48	GGCGCGCCGTGGTATCCACCTCAAC (*Asc*I)	*acn*$_{up}$ (R)
P49	GGATACCACGGCGCGCCATATGCCACAGCACGG (*Asc*I)	*acn*$_{down}$ (F)
P50	CATGCGGTAGAATTACTCGG	*acn*$_{down}$ (R)
P51	ACGATTATCCGAGTCATGAG	Δ*acn*-fusion, nested (F)
P52	GCCAGATTCGCTCCGCTTG	Δ*acn*-fusion, nested (R)
P53	TGTAAAACGACGGCCAGT	M13 forward
P54	CAGGAAACAGCTATGACC	M13 reverse

[a] 5' to 3' direction; restriction sites are underlined and named at the end of the primer sequences

[b] If unspecified, gene names refer to *A. succinogenes*

[c] F, forward primer; R, reverse primer

[d] Primer P41 contains the USS ACCGCACTT

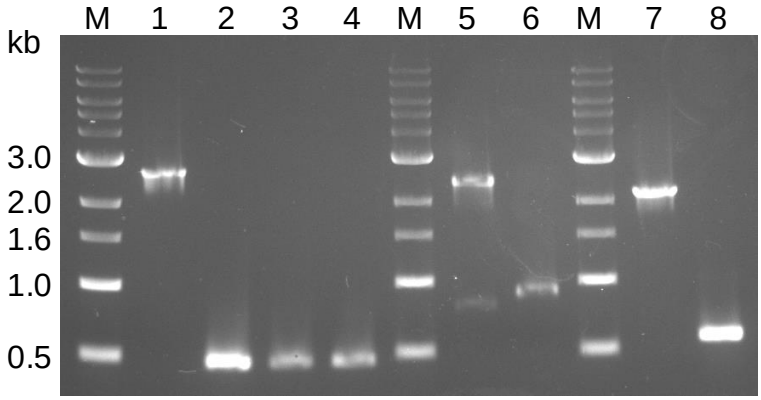

Figure A4.1 Verification of the construction of strains Δ*pflB* Δ*cit* and Δ*pflB* Δ*acn* by PCR. Lanes M: DNA molecular markers. PCR products amplified from 130Z genomic DNA (Lane 1), Δ*pflB* genomic DNA (Lanes 2, 5, and 7), a Δ*pflB* Δ*acn* colony (Lanes 3 and 6), and a Δ*pflB* Δ*cit* colony (Lanes 4 and 8). Lanes 1 to 4: verification that the host for the double mutations is strain *pflB*; primers P27-P28 (flanking *pflB*) amplify a 2.7-kb fragment in 130Z, and a 434-bp fragment in Δ*pflB*, Δ*pflB* Δ*acn*, and Δ*pflB* Δ*cit*). Lanes 5 and 6: verification that *acn* is deleted in strain Δ*pflB* Δ*acn*; primers P51-P52 amplify a 2.4-kb fragment in Δ*pflB* and a 855-bp fragment in Δ*pflB* Δ*acn*. Lanes 7 and 8: verification that *citDEF* are deleted in strain Δ*pflB* Δ*cit*; primers P55-P56 amplify a 2.2-kb fragment in Δ*pflB* and a 524-bp fragment in Δ*pflB* Δ*cit*. The PCR products in lanes 6 and 8 would be 2.35 kb (Δ*acn*) and 2.0 kb (Δ*cit*), if the deletions still contained the *icd* marker.

REFERENCES

REFERENCES

1. **Werpy TA**, **Petersen G**. 2004. Top value added chemicals from biomass for the U.S. Department of Energy (DOE) by National Renewable Energy Laboratory (NREL).

2. **Bozell JJ, Petersen G**. 2010. Technology development for the production of biobased products from biorefinery carbohydrates—the US Department of Energy's Top 10 revisited. Green Chemistry **12**:539–554.

3. **Guettler MV, Jain MK, Rumler D**. 1996. Method for making succinic acid, bacterial variants for use in the process, and methods for obtaining variants. U.S. patent 5,573,931.

4. **Guettler MV, Jain MK, Soni BK**. 1996. Process for making succinic acid, microorganisms for use in the process and methods of obtaining the microorganisms. U.S. patent 5,504,004.

5. **Guettler MV**, **Rumler D**, **Jain MK**. 1999. *Actinobacillus succinogenes* sp. nov., a novel succinic-acid-producing strain from the bovine rumen. Int J Syst Bacteriol **49**:207–216.

6. **van der Werf MJ**, **Guettler MV**, **Jain MK**, **Zeikus JG**. 1997. Environmental and physiological factors affecting the succinate product ratio during carbohydrate fermentation by *Actinobacillus* sp. 130Z. Arch Microbiol **167**:332–342.

7. **Kurzrock T, Weuster-Botz D**. 2010. Recovery of succinic acid from fermentation broth. Biotechnol Lett **32**:331–339.

8. **Kim P**, **Laivenieks M**, **McKinlay J**, **Vieille C**, **Zeikus JG**. 2004. Construction of a shuttle vector for the overexpression of recombinant proteins in *Actinobacillus succinogenes*. Plasmid **51**:108–115.

9. **Oswald W, Tonpitak W, Ohrt G, Gerlach G-F**. 1999. A single-step transconjugation system for the introduction of unmarked deletions into *Actinobacillus pleuropneumoniae* serotype 7 using a sucrose sensitivity marker. FEMS Microbiol Lett **179**:153–160.

10. **Tatum FM, Briggs RE**. 2005. Construction of in-frame *aroA* deletion mutants of *Mannheimia haemolytica*, *Pasteurella multocida*, and *Haemophilus somnus* by using a new temperature-sensitive plasmid. Appl Environ Microbiol **71**:7196–7202.

11. **Kim JM, Lee KH, Lee SY**. 2008. Development of a markerless gene knock-out system for *Mannheimia succiniciproducens* using a temperature-sensitive plasmid. FEMS Microbiol Lett **278**:78–85.

12. **Maughan H, Sinha S, Wilson L, Redfield R**. 2008. Competence, DNA uptake, and transformation in *Pasteurellaceae*, p 79–98. *In* Kuhnert P, Christensen H (ed), *Pasteurellaceae* Biology, Genomics and Molecular Aspects. Caister Academic Press.

13. **Redfield RJ, Findlay WA, Bossé J, Kroll JS, Cameron AD, Nash JH**. 2006. Evolution of competence and DNA uptake specificity in the Pasteurellaceae. BMC Evol Biol **6**:82.

14. **Sisco KL, Smith HO**. 1979. Sequence-specific DNA uptake in *Haemophilus* transformation. Proc Natl Acad Sci USA **76**:972–976.

15. **Danner DB, Deich RA, Sisco KL, Smith HO**. 1980. An 11-base pair sequence determines the specificity of DNA uptake in *Haemophilus* transformation. Gene **11(3–4)**:311–318.

16. **Wang Y, Orvis J, Dyer D, Chen C**. 2006. Genomic distribution and functions of uptake signal sequences in *Actinobacillus actinomycetemcomitans*. Microbiology **152**:3319–3325.

17. **McKinlay JB, Laivenieks M, Schindler BD, McKinlay AA, Siddaramappa S, Challacombe JF, Lowry SR, Clum A, Lapidus AL, Burkhart KB, Harkins V, Vieille C**. 2010. A genomic perspective on the potential of *Actinobacillus succinogenes* for industrial succinate production. BMC Genomics **11**:680.

18. **Redfield RJ, Cameron AD, Qian Q, Hinds J, Ali TR, Kroll JS, Langford PR**. 2005. A novel CRP-dependent regulon controls expression of competence genes in *Haemophilus influenzae*. J Mol Biol **347**:735–747.

19. **Cox MM**. 1983. The FLP protein of the yeast 2-microns plasmid: expression of a eukaryotic genetic recombination system in *Escherichia coli*. Proc Natl Acad Sci USA **80**:4223–4227.

20. **Zhu XD, Pan G, Luetke K, Sadowski PD**. 1995. Cleavage-dependent ligation by the FLP recombinase. Characterization of a mutant FLP protein with an alteration in a catalytic amino acid. J Biol Biol **270**:11646–11653.

21. **McKinlay JB, Shachar-Hill Y, Zeikus JG, Vieille C**. 2007. Determining *Actinobacillus succinogenes* metabolic pathways and fluxes by NMR and GC-MS analyses of ^{13}C-labeled metabolic product isotopomers. Metab Engin **9**:177–192.

22. **McKinlay JB, Vieille C**. 2008. ^{13}C-metabolic flux analysis of *Actinobacillus succinogenes* fermentative metabolism at different $NaHCO_3$ and H_2 concentrations. Metab Engin **10**:55–68.

23. **Sawers RG, Clark DP**. 2004. Chapter 3.5.3, Fermentative pyruvate and acetyl-coenzyme A metabolism. *In* Böck A, Curtiss III R, Kaper JB, Karp PD, Neidhardt FC, Nyström T, Slauch JM, Squires CL, Ussery D (ed), *EcoSal—Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. ASM Press, Washington, DC.

24. **Stols L, Donnelly MI**. 1997. Production of succinic acid through overexpression of NAD^+-dependent malic enzyme in an *Escherichia coli* mutant. Appl Environ Microbiol **63**:2695–2701.

25. **Jantama K, Haupt MJ, Svoronos SA, Zhang XL, Moore JC, Shanmugam KT, Ingram LO**. 2008. Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of *Escherichia coli* C that produce succinate and malate. Biotechnol Bioengin **99**:1140–1153.

26. **Zhang X, Jantama K, Shanmugam KT, Ingram LO**. 2009. Reengineering *Escherichia coli* for succinate production in mineral salts medium. Appl Environ Microbiol **75**:7807–7813.

27. **Lee SJ**, **Song H**, **Lee SY**. 2006. Genome-based metabolic engineering of *Mannheimia succiniciproducens* for succinic acid production. Appl Environ Microbiol **72**:1939–1948.

28. **Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (ed)**. 1993. Current Protocols in Molecular Biology. Greene Publishing & Wiley-Interscience, New York.

29. **McKinlay JB, Zeikus JG, Vieille C**. 2005. Insights into the *Actinobacillus succinogenes* fermentative metabolism in a chemically defined growth medium. Appl Environ Microbiol **71**:6651–6656.

30. **Choi KH, Kumar A, Schweizer HP**. 2006. A 10-min method for preparation of highly electrocompetent *Pseudomonas aeruginosa* cells: application for DNA fragment transfer between chromosomes and plasmid transformation. J Microbiol Methods **64**:391–397.

31. **Poje G, Redfield RJ**. 2003. Transformation of *Haemophilus influenzae*. Methods Mol Med **71**:57–70.

32. **Bosse JT, Sunita S, Schippers T, Kroll JS, Redfield RJ, Langford PR**. 2009. Natural competence in strains of *Actinobacillus pleuropneumoniae*. FEMS Microbiol Lett **298**:124–130.

33. **Trevors JT**. 1986. Plasmid curing in bacteria. FEMS Microbiol Rev **32**:149–157.

Chapter 5

Respiratory glycerol metabolism of *Actinobacillus succinogenes* 130Z for succinate production

The work completed in this chapter is part of a collaboration with Dr. Bryan Schindler (primary author, a former member of the Vieille lab) and Dr. Claire Vieille, published as:

Schindler BD, **Joshi RJ**, **Vieille C**. 2014. Respiratory glycerol metabolism of *Actinobacillus succinogenes* 130Z for succinate production. J Ind Microbiol Biotechnol **41**:1339-1352. doi: 10.1007/s10295-014-1480-x

Article can be viewed at the following URL:

http://rdcu.be/v6Wy

My contribution to this report included:

- Developing and optimizing a microaerobic continuous culture system for *A. succinogenes* (chemostats) for growth on glycerol.
- Conducting continuous culture experiments with *A. succinogenes* 130Z and carrying out fermentation balances.
- Drafting portions of the manuscript for above bullet points

Reading and reviewing entire manuscript.

Chapter 6

Conclusions and future directions

6.1 Introduction

A. succinogenes is one of the best natural succinate producers known. Many advances have been carried out with respect to its metabolism on substrates like glucose, glycerol, and lignocellulosic hydrolysates (1-4). Microorganisms such as *E. coli*, *S. cerevisiae*, and *C. glutamicum* are being studied and engineered for succinate production as well. Many of the succinate production increases obtained in *E. coli* were obtained by mimicking *A. succinogenes* metabolism. Despite much insight into *A. succinogenes*'s metabolism, we still know very little about other aspects, such as succinate efflux transporters or small RNAs (sRNAs) in *A. succinogenes*. In this dissertation, we have focused on identifying succinate exporters, and overexpressing four individual transporters increased succinate production. Additionally, we have also identified sRNAs in *A. succinogenes* and shown that synthetic sRNAs can be used in *A. succinogenes* for inhibiting expression of unwanted pathways. The knowledge gained about succinate transporters and sRNAs in this dissertation can be used for future metabolic engineering of *A. succinogenes* for succinate production.

6.2 Succinate transporters in *A. succinogenes*

We have identified succinate transporters in *A. succinogenes* using a combination of proteomics and RNA sequencing (RNAseq). The transporters with the most hits in our proteomics studies and high transcripts levels were different from the succinate exporters identified in *E. coli* and *C. glutamicum*. Although homologs of the *E. coli* and *C. glutamicum* succinate exporters exist in *A. succinogenes*, none of them were among the top hits in our proteomics and transcriptomics studies. Asuc_1999, Asuc_0142, Asuc_2058, and Asuc_1990-91 were the transporter candidates that topped our proteomics list. Since expressing the transporters

under the control of a strong promoter was toxic for *A. succinogenes*, we developed a promoter library with a large span of expression strength. Expressing Asuc_1999 under the control of weaker promoter $p_{pckA\text{-}103}$ increased the succinate yield in *A. succinogenes* by 24%, mainly at the cost of biomass production.

6.3 Identification of sRNAs in *A. succinogenes*

The purpose of this study was to gain some insight into the size, structure, and function of sRNAs in *A. succinogenes*, especially of Hfq-dependent sRNAs. We identified 260 sRNAs in *A. succinogenes* in glucose- and glycerol-grown cultures, out of which 24 had predicted Rho-independent terminators, a characteristic feature of Hfq-dependent sRNAs. We were also able to validate 14 randomly selected sRNAs by RT-PCR. One of the major goals of this study was to use the knowledge gained from the RNAseq studies of sRNAs and design a synthetic sRNA to inhibit the translation of select mRNAs in *A. succinogenes*. We were most interested in this approach since knockout mutants maybe impossible to construct for certain essential genes, and this approach offers a much more tunable way to lower gene expression. In this study, we successfully designed synthetic sRNAs targeting β-galactosidase and acetate kinase expression. We were able to decrease β-galactosidase activity by 32% and to decrease the acetate yield by 14%.

6.4 Future directions

6.4.1 Experimental identification of Hfq-binding sRNAs

In this study we used computational and manual approaches to identify possible Hfq-binding sRNAs. We still need to experimentally verify these Hfq-binding sRNAs, though. One approach to identifying Hfq-binding sRNAs as well as their target mRNAs would be to compare transcriptomics data between a Δ*hfq* strain and the wild-type strain. I built a Δ*hfq*::Cm^R strain, but that strain was only transiently culturable, not long enough to allow transcriptomic analysis. Another approach is copurification using a strain expressing His-tagged Hfq, which would allow us to isolate and sequence Hfq-binding sRNAs. This approach would allow us to confirm a number of the predicted Hfq-binding sRNAs in our RNAseq studies, as well as identify some mRNAs regulated by Hfq-dependent sRNA. Because we could not complement our Δ*hfq*::Cm^R strain with a plasmid-borne His-tagged *hfq* gene, we are copurifying His^6-tagged Hfq with its binding RNAs after expression in wild-type *A. succinogenes* instead. Results from this experiment will yield additional scaffold sequences from Hfq-binding sRNAs.

6.4.2 Continuing the development of synthetic sRNAs in *A. succinogenes*

Once we confirm Hfq-binding sRNAs experimentally, it will be possible to compare many scaffolds sequences and build additional synthetic sRNAs. Additionally, we need more studies with synthetic sRNAs designed with differing scaffold and targeting sequence lengths to modulate the binding with and regulation of the target mRNA. A more in depth look at the secondary structures of the native Hfq-binding sRNAs is required. Mimicking the secondary structure of native Hfq-binding sRNAs in our synthetic sRNAs may increase stability and

improve effectiveness of the designed sRNAs. All these studies will help in building a library of synthetic sRNA scaffolds, where simply replacing the target binding regions would be possible.

6.4.3 Identifying the major succinate transporter(s) in *A. succinogenes*

In our proteomics and RNAseq studies we were able to identify a few transporter candidates that may be responsible for exporting succinate out of the cell. However, we were unable to identify one major succinate transporter that functions under anaerobic conditions. As seen in *E. coli* and *C. glutamicum*, several succinate exporters may contribute to succinate efflux under anaerobic conditions. Our attempts to knockout any of the four succinate exporters identified in this study have been unsuccessful so far. But once we have these knockout mutants, we will be able to test the knockout strains under anaerobic and microaerobic conditions to shed more light on succinate exporters active under different conditions in *A. succinogenes*.

6.4.4 Construction of a Δ*ackA* knockout mutant

All our attempts at making a Δ*pflB*Δ*ackA* mutant have been unsuccessful so far. Recently, the Bechkam group was able to make a Δ*pflB*Δ*ackA* mutant strain using homologous recombination (5). The difference in approach was the use of a different selection marker and growth medium. The authors used an antibiotic resistance gene as their selection marker and tryptic soy broth supplemented with glucose and the antibiotic as the selection medium. In the same study, the authors overexpressed phosphoenolpyruvate carboxykinase, malate dehydrogenase, and fumarase independently and in the same strain. The succinate yields and titers of strains expressing malate dehydrogenase and all the enzymes together were higher than

that of the wild-type strain. However, the same constructs in the Δ*pflB*Δ*ackA* background did not cause any increase in succinate titer or yield.

Our lack of success in constructing a Δ*ackA* strain suggests that the acetate pathway is essential for *A. succinogenes'*s growth on a minimal medium such as AM3. We have previously constructed Δ*lacZ*::Cm^R and Δ*hfq*::Cm^R strains by selecting colonies on LB medium supplemented with chloramphenicol, so we should be able to obtain a Δ*ackA* mutant using the same approach.

6.4.5 Combination of different approaches to increase succinate production

Evolved strains of *A. succinogenes* that produce more succinate than the wild-type strain on glucose and xylose have been isolated in our lab. The X strains, evolved for fast growth on xylose, have mutations in the *xylE* and *fusA* genes, encoding the xylose permease and elongation factor G, respectively. Introducing the *xylE* mutation in the wild-type strain increased the succinate yield by 40%. We have a Δ*pflB* mutant (devoid of pyruvate formate lyase) that no longer produces formate and produces 69% more succinate than the wild-type strain on glucose (6). We also have identified succinate exporters that, upon overexpression, increase succinate production as well. And we now have a library of promoters of a range of strengths,

Combining different mutations, expressing succinate exporter(s) as well as possibly some of the enzymes in the succinate pathway should allow us to engineer strains that produce record amounts of succinate with minimum byproducts.

REFERENCES

REFERENCES

1. **McKinlay JB**, **Shachar-Hill Y**, **Zeikus JG**, **Vieille C**. 2007. Determining *Actinobacillus succinogenes* metabolic pathways and fluxes by NMR and GC-MS analyses of ^{13}C-labeled metabolic product isotopomers. Metab Eng **9**:177–192.

2. **McKinlay JB**, **Vieille C**. 2008. ^{13}C-metabolic flux analysis of *Actinobacillus succinogenes* fermentative metabolism at different $NaHCO_3$ and H_2 concentrations. Metab Eng **10**:55–68.

3. **Schindler BD**, **Joshi RJ**, **Vieille C**. 2014. Respiratory glycerol metabolism of *Actinobacillus succinogenes* 130Z for succinate production. J Ind Microbiol Biotechnol **41**:1339-1352.

4. **McPherson N**. 2017. Novel Insights Into Sugar and Succinate Metabolism of *Actinobacillus succinogenes* from Strains Evolved for Improved Growth on Lignocellulose Hydrolysate Sugars. Doctoral. Michigan State University.

5. **Guarnieri MT**, **Chou YC**, **Salvachua D**, **Mohagheghi A**, **St John PC**, **Peterson DJ**, **Bomble YJ**, **Beckham GT**. 2017. Metabolic Engineering of *Actinobacillus succinogenes* Provides Insights into Succinic Acid Biosynthesis. Appl Environ Microbiol 83: e00996-17.

6. **Schindler B**. 2011. Understanding and improving respiratory succinate production from glycerol by *Actinobacillus succinogenes*. Doctoral Dissertation. Michigan State University.