MECHANISTIC INSIGHTS INTO *ESCHERICHIA COLI* UTILIZATION OF ACETYLENEDICARBOXYLATE AS A SUSTAINABLE FEEDSTOCK

By

Bismarck Kyei Amaniampong

A DISSERTATION

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

Chemistry – Doctor of Philosophy

2024

ABSTRACT

Raw materials for production of chemicals and energy are classified as either non-renewable or renewable. Non-renewable feedstocks, such as petroleum, natural gas, and coal, account for over 85% of global energy use. However, these sources cannot be readily replenished, and the production of chemicals from this route results in the emission of greenhouse gases. The development of sustainable strategies is critical to addressing challenges associated with industrial chemical production from fossil fuels. Such a strategy would prioritize the use of abundant and renewable feedstocks while minimizing greenhouse gas emissions and environmental impacts. Renewable feedstocks derived from plants represent a sustainable and lower carbon emission alternative. However, this approach places the food industry in direct competition with the chemical industry for feedstock, in addition to requiring intensive preprocessing steps in order to utilize every component of this biomass for chemical production. Direct microbial utilization of CH_4 and CO_2 presents a more sustainable and green approach for commercial chemical production. However, microbes metabolizing methane and carbon dioxide require specific conditions for optimal growth and metabolic activity. C4-Dicarboxylate compounds, such as fumarate, succinate, malate, and aspartate, present another avenue for microbial growth but are not readily available. Their structural similarity to citric acid cycle-related metabolites positions them as prime candidates for microbial assimilation. This study presents acetylenedicarboxylate (ADCA), an unnatural C4-dicarboxylate derivable from methane and carbon dioxide, as a potential alternate feedstock. This research provides insight into microbial growth on acetylenedicarboxylate as the sole carbon source for the bioproduction of chemicals. We employ a multi-omic approach to elucidate the adaptation mechanism to the acetylenic carbon source, integrating comparative genomics, transcriptomics, and the analysis of excreted metabolites. These findings highlight the capacity of *E. coli* to assimilate and utilize acetylenedicarboxylate as the sole carbon source and identify putative genes responsible for ADCA utilization. We further demonstrate the bioconversion of ADCA to lactic acid (a commodity chemical) and shikimic acid (precursor for a pharmaceutical agent), offering a promising proof of concept for this novel feedstock as a sustainable decarbonization approach in microbial biotechnology.

This dissertation is dedicated to the loving memory and honor of my late brother,
Dominic Gyasi Brannor.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude and admiration to my research advisor, Dr. Karen Draths, for her unwavering support and guidance throughout my academic journey. Your encouragement and mentorship have been instrumental in shaping me into the scientist I am today and what I hope to achieve. Your dedication, expertise, and kindness have inspired a great deal of confidence in me. I can't thank you a lot for giving me the opportunity to be creative and explore new territories beyond our lab's capabilities. This challenged me in lots of ways and pushed me outside my comfort zone to learn several skills useful in my research field. Thank you for believing in me and for being a source of both academic and emotional support. I am truly grateful for all you have done and continue to do. A huge thanks to my committee, Dr. James Geiger, Dr. Babak Borhan, and Dr. Robert Maleczka, for their continued guidance and feedback throughout my academic journey.

A special thanks to Konika Konika for being a fantastic partner and my support system. Thanks for being there since day one! Thanks to Solomon Attionu, Kofi Yebooah Simpa (Papa Vee), Oliver Aggrey, Dare George, and Chia-Wei Yang for also being amazing friends and study/travel squad since our first year.

My sincere thanks to Dr. Hadi Nayebi, Dr. Yasheen Jadidi, Dr. Amaya Sirinimal Mathes Hewage, Dr. Katie Kwiatkowski, and Kate Silva for their friendship, support, and advice throughout this journey. I'll always be grateful for the impact you've had on my life. Big thanks to Draths Lab alumni Dr. Megan Gruenberg and brilliant mentees I had the opportunity to work with, Jada Harvey, Elizabeth Courthright, Greg Battley, Noor Saber, Chrissy Lee, and Nada Al-Ahmad.

To all the friends I made in Michigan, the list can never be complete; Daniel, Charles, Sunanda, Souvik, Enoch, Haritha, Ishita, George, Shukurah, Michael, Divine, East Lansing soccer club, and to everyone, I say ‘medaase’.

My biggest thanks to my family. To my MOM and DAD, thanks for supporting my dreams even though these dreams took me further away from you. I’ll keep making you proud. To my siblings, Bright and Anna, you are the best! Most times. I hope to be the best role model for you.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ADCA	Acetylenedicarboxylic acid
ACA	Acetylenecarboxylic acid
ALDH	Aldehyde dehydrogenase
Ap	Ampicillin
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
M9	Minimal media
M9A	Minimal media containing ADCA
M9G	Minimal media containing glucose
M9F	Minimal media containing fumarate
M9GA	Minimal media containing glucose and ADCA
M9GF	Minimal media containing glucose and fumarate
M9FA	Minimal media containing fumarate and ADCA
h	Hour (s)
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
Kan	Kanamycin
kb	Kilobase pair
kDa	Kilodalton
LB	Luria-Bertani
M	Molar
mg	Milligram
min	Minute

mL	Milliliter
mM	Millimolar
NAD$^+$	Nicotinamide adenine dinucleotide, reduced form
NADH	Nicotinamide adenine dinucleotide, oxidized form
NADP$^+$	Nicotinamide adenine dinucleotide phosphate, oxidized form
NADPH	Nicotinamide adenine dinucleotide phosphate, reduced form
NMR	Nuclear magnetic resonance
OD	Optical Density
PCR	Polymerase Chain Reaction
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
ser	L-serine
TAE	Tris acetate EDTA
TCA	Tricarboxylic acid
Th	Thiamine
U	Unit; µmol/min/mg

CHAPTER ONE: Current Strategies for the Bioproduction of Chemicals

1.1 Non-renewable and renewable feedstocks

Feedstocks for production of chemicals and energy are generally classified as either non-renewable or renewable.[1,2] Non-renewable feedstocks, which account for over 85% of global energy use, are those resources that cannot be readily replenished, including fossil-based resources such as petroleum, natural gas, and coal (Figure 1.1).[3] Most developed nations rely heavily on these non-renewable sources for energy generation.[4]

Petroleum is the most common non-renewable resource exploited in processes such as cracking and reforming to derive a variety of chemicals.[5] As a significant energy source, crude oil or liquid petroleum is refined into different products, including gasoline, diesel fuel, jet fuel, and heating oil. According to the Annual Energy Outlook 2023 report by the U.S. Energy Information Administration, petroleum was the primary source of energy in the U.S. in 2022, accounting for 35.85 quadrillion British thermal units, followed closely by natural gas with 33.41 quadrillion British thermal units (Figure 1.2). Natural gas, predominantly used for producing methanol and ammonia, is essential for synthesizing other chemicals and fertilizers.[6,7]

Significant oil reserves are concentrated in the Middle East, with considerable reserves also in North America.[8] However, the majority of known oil reserves are being exploited at a rate that exceeds discoveries. Due to this overuse, it has been projected that oil supplies might be exhausted in approximately 30 years if current consumption rates continue and no significant new sources are found.[3,9] Despite these projections, oil remains the preferred fuel source over coal due to its higher energy output and relatively cleaner burn, producing approximately 50% less sulfur dioxide. Oil combustion for energy also releases pollutants such as sulfur dioxide, nitrogen oxides, carbon dioxide, and carbon monoxide that pose environmental and health threats.[10] These emissions contribute to air pollution and, consequently, global warming, leading to climate change.

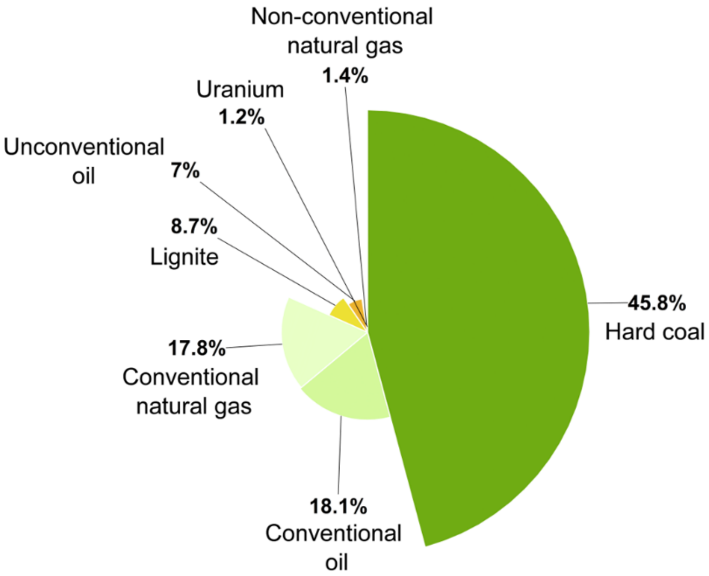


Figure 1.1. Distribution of non-renewable feedstocks for fuel worldwide reserves in 2022 by source. (Data source: BGR – Energiedaten 2023, tab A-1; published by Statista 2024).

Oil drilling, especially offshore, may result in oil spills that pose risks to marine life and birds, and terrestrial drilling can severely impact wildlife and aquatic ecosystems.[11,12] The extraction, refining, and combustion of fossil fuels, predominantly from drilling and mining, have resulted in nearly three-fourths of human-caused emissions over the past 20 years.[13] Coal, the most abundant fossil fuel, has an estimated global reserve of one trillion metric tons. Its composition varies due to the conditions of its formation. The energy output and pollution levels of coal vary according to its type, ranging from lignite to anthracite.[14] The most commonly found bituminous coal in the U.S. has a high sulfur content, which results in the release of significant sulfur dioxide when burned.[15]

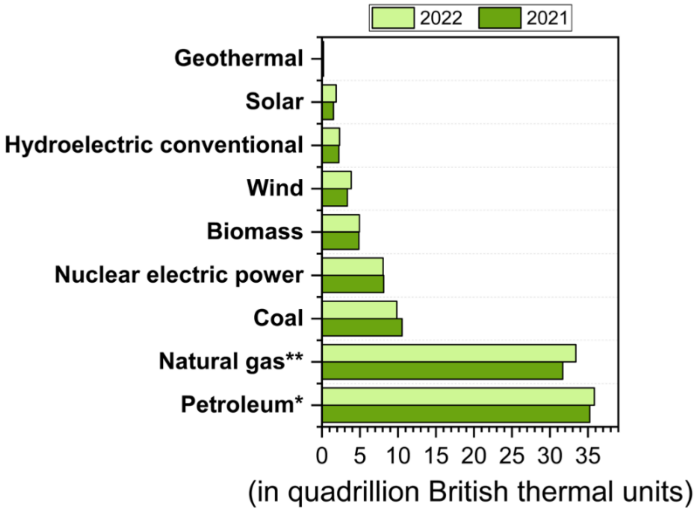


Figure 1.2. Primary energy consumption by source in the U.S. 2021 – 2022. (Data source: U.S. Energy Information Administration (EIA), Annual Energy Outlook (AEO2023), June 2023).

Due to the environmental challenges and regulatory pressures from fluctuating crude oil prices and carbon emission mandates, there is a high interest in transitioning towards renewable feedstocks. The use of renewable energy sources has gradually increased over the past decade, with approximately 7 quadrillion BTU consumed in the U.S. in 2020.[16,17] By 2022, renewable electricity accounted for approximately 22.5% of the nation's total electricity generation.[18] Renewable energy sources are derived from naturally replenishable sources such as biomass, hydro, wind, and solar.[2] These sources include agricultural residues such as corn stover and sugarcane bagasse, energy crops like switchgrass and miscanthus, and various forms of algae that can be cultivated on non-arable land.[4] Renewable feedstocks derived from plants, which absorb CO_2, represent a sustainable and lower-carbon footprint alternative to non-renewable sources.[19] In order for these biomasses to be considered truly renewable, it is essential to implement sustainable practices such as replanting. These practices include ensuring that for every tree that is cut down for firewood, another is planted.

Renewable feedstocks are critical for developing a sustainable and circular economy, offering reduced environmental impacts and lower carbon emissions.[19–22] Technological advancements are pivotal in this transition, focusing on developing new chemical and biochemical pathways, efficient processes, and catalysts to convert renewable resources into valuable chemical products sustainably.[21,23] Of particular interest is the bioproduction of chemicals from biomass, which has the potential to lower GHG emissions. Moreover, the bioproduction of chemicals from methane and carbon dioxide using microbes offers a potential avenue for utilizing these gases as feedstocks.[24–29]

1.2 Role of microorganisms as cell factories

Microbial cell factories (MCFs) are engineered microorganisms that have been genetically modified to produce chemicals, fuels, and other valuable products through streamlined biosynthetic pathways.[26,30,31] MCFs can utilize renewable carbon sources as their primary source of carbon, offering a sustainable alternative to traditional methods of chemical synthesis and extraction from plants and animals. MCFs have emerged as a significant area of focus within applied microbiology and biotechnology. These cell factories provide a sustainable platform for producing a variety of chemicals and materials traditionally obtained from fossil resources. The construction and design of MCFs involve integrating synthetic biology and metabolic engineering to modify and optimize host microorganisms such as *Escherichia coli* and *Saccharomyces cerevisiae.*[32–35]

The applications of MCFs are diverse and impactful, particularly in the production of bulk chemicals and high-value compounds such as amino acids, alcohols, and organic acids, which are essential in various industries (Figure 1.3).[36–38] In the pharmaceutical industry, engineered microbial pathways enable the synthesis of complex molecules and precursors, offering a more sustainable and economical solution than chemical catalysis approaches.[33,39,40] Furthermore, MCFs are instrumental in the production of biofuels, such as bioethanol and biodiesel, from agricultural residues and other non-edible biomass. Hence, they are crucial for reducing reliance on fossil fuels.[41,42]

Despite the promising applications of microbial cell factories, the broader implementation of these technologies faces several challenges.[32,43] These include the development of robust microbial strains that can endure industrial bioprocessing conditions, the optimization of metabolic pathways to boost yield and productivity, and the scaling of production processes from laboratory to industrial scales.

The microbial production of chemicals has attracted significant industrial research attention, driven by environmental concerns and the depletion of petroleum resources.[33,44] Using fermenters, microbes provide an efficient bioprocessing method for industrial applications. Technological advances have refined methods for the isolation and genetic modification of production strains, expanding the potential of MCFs beyond traditional, easy-to-handle cell factories to a diverse range of microorganisms. Metabolic engineering of MCFs involves modifying cellular metabolism to enhance or introduce the production of desired metabolites, optimizing metabolic routes, removing by-product pathways, and adjusting cellular energy balance to favor target product synthesis. The advent of genome sequencing, metabolic modeling, gene expression profiling, and other genetic tools, such as CRISPR, has greatly facilitated microbial cell factory engineering, making it more efficient and cost-effective. These techniques have revolutionized the rapid and precise modification of microbial genomes, allowing for the fine-tuning of biosynthetic pathways to improve performance.[24,31]

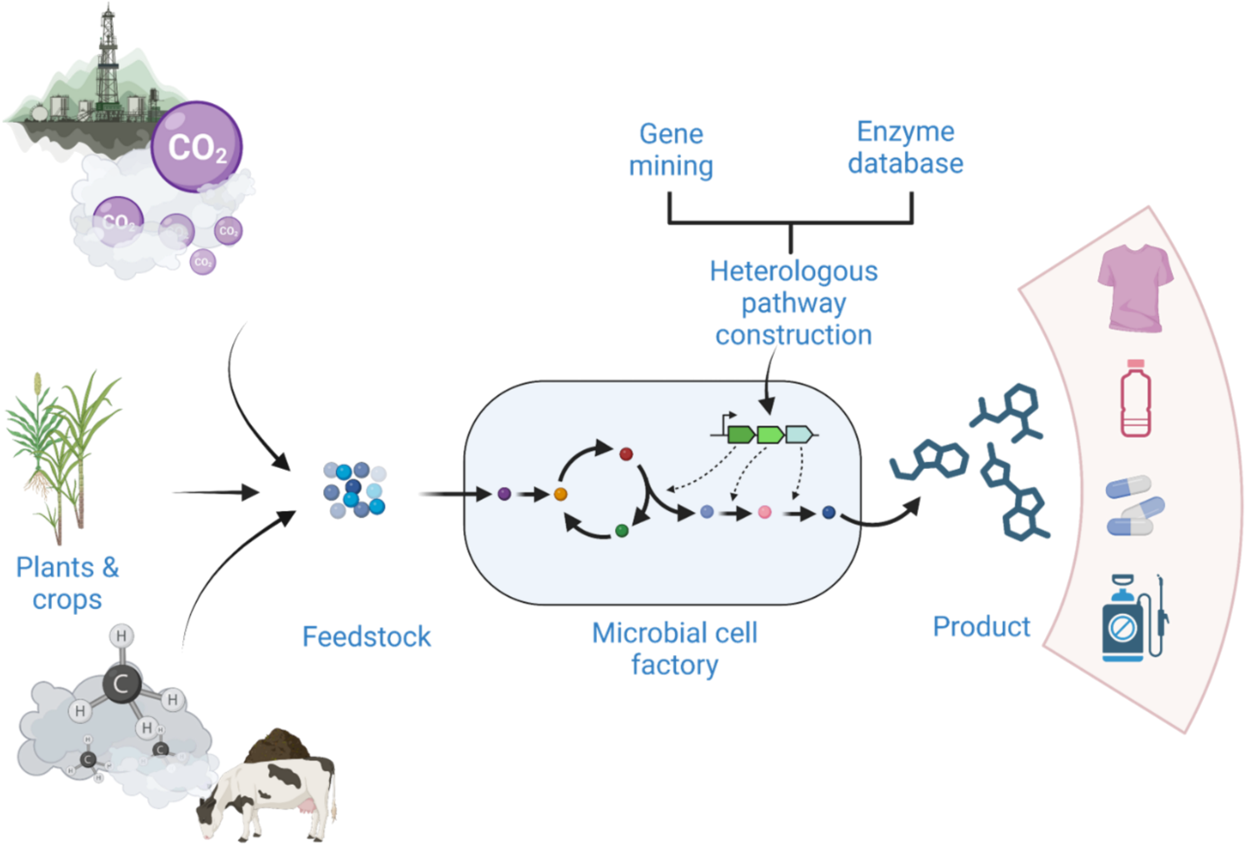


Figure 1.3. Application of microbial cell factories (MCFs) for the manufacturing of high-value chemicals from renewable feedstocks.

1.3 Current challenges in feedstock sources for the bioproduction of chemicals

The use of biomass for the bioproduction of chemicals has increased over the last 10 to 15 years.[45,46] The transition to using biomass for biofuels, like bioethanol and biodiesel, offers a renewable alternative to fossil-derived transport fuels, providing carbon from a renewable source. Historically, bioethanol and biodiesel produced from food crops such as sugar, starch, and oils from palm, rapeseed, soy, beets, and cereals such as corn and wheat have constituted the majority of this sector (Figures 1.4 and 1.5). However, as the European Commission has outlined, there has been a shift towards advanced biofuels - both second and third-generation - which are made from feedstocks that do not compete directly with food and feed crops.[46,47] These include wastes and agricultural residues like wheat straw and municipal waste, non-food crops such as miscanthus and short rotation coppice, and algae. Currently, advanced biofuels account for only about 2% of biofuel production.[47]

The U.S. Department of Energy (DOE) and the European Commission have enacted measures to mitigate the negative impacts of using food crops for bioenergy.[48,49]

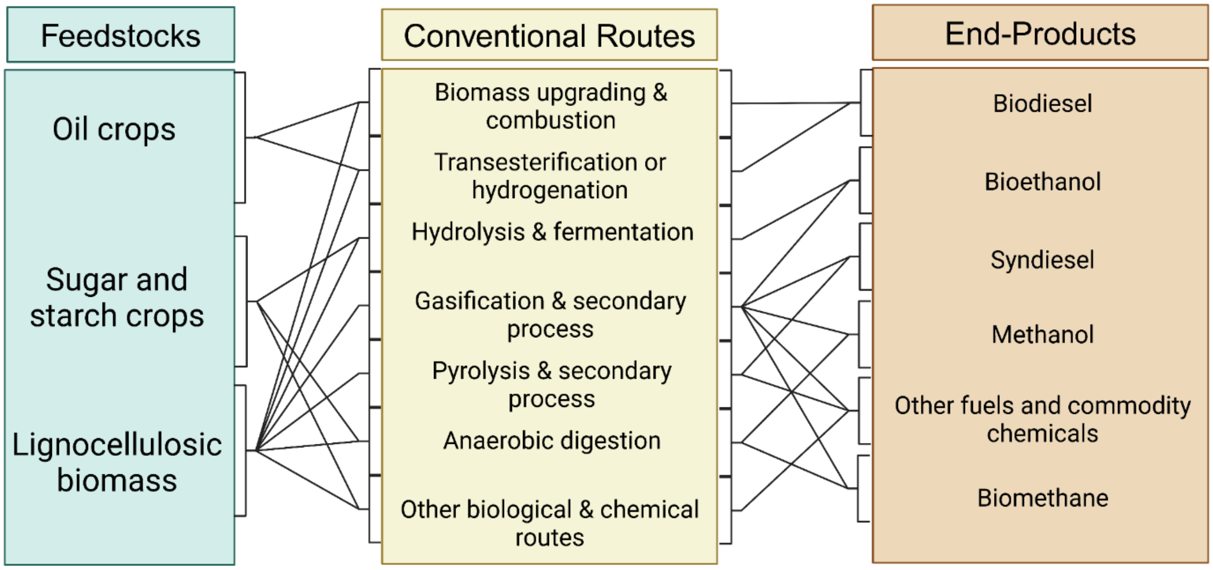


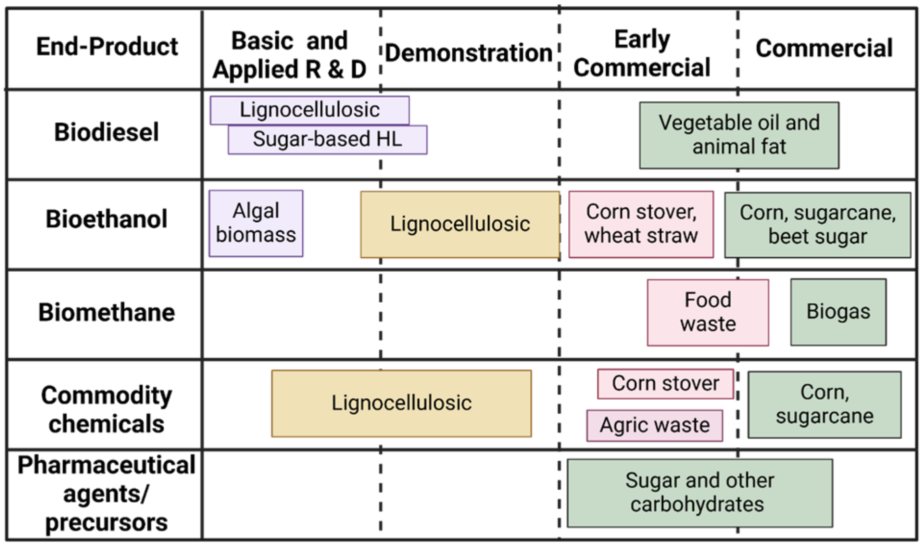
Figure 1.4. Conventional routes connecting renewable feedstocks to commercial products.

These measures include the restriction of state aid for conventional biofuels using food crops, the proposal of a cap of 7% on biofuels derived from "food crops" in transport energy use, and the encouragement of the use of biofuels produced from wastes and residues through double counting.[50] This strategy aims to promote a more diverse and sustainable range of feedstocks and necessitates the commercial development of innovative, albeit more complex and costly, conversion technologies.

The term "advanced biofuels" is commonly used to refer to those derived from lignocellulosic biomass, which includes non-food "woody" parts of plants.[44,51] This biomass exhibits considerable variation in the composition of cellulose, hemicellulose, and lignin, as well as in chain length and degrees of polymerization. Some lignocellulosic materials, such as citrus peel, can be readily broken down into substrates for fuel molecules (Table 1.1). In contrast, other materials, particularly those with higher lignin content, such as hardwood, require pretreatment and present more significant production challenges and costs.[52,53]

The competition for biomass is a significant challenge across sectors of the economy, affecting bioenergy development. This competition arises from the overlapping demands for biomass as food and as feedstock for making materials (such as bioplastics and textiles), as well as an energy source.[44] Factors influencing this competition include raw material prices, product prices, policy, land availability, and technological constraints. Current challenges in feedstock sourcing for the bioproduction of chemicals are multifaceted, covering technical, economic, environmental, and sustainability issues.[44] These include the availability and affordability of quality feedstock at a large scale, the sustainable production of feedstocks, and the technological challenges in feedstock conversion.

In addition, the logistics of the supply chain, including the collection, transportation, and storage of biomass, are of great importance for the consistent and cost-effective delivery of feedstocks to bio-refineries.[5] In addition, policy and regulatory frameworks play a pivotal role.[49] Supportive policies can enhance the growth of the bio-based economy by promoting sustainable biomass production, supporting technological innovation, and fostering market development for bio-based products. Consequently, a biorefinery approach is necessary to valorize all biomass components.[53,54] To address these challenges, a multidisciplinary approach is required to bolster agricultural practices, biotechnology, chemical engineering, and policy development. This approach is essential for ensuring the continued expansion and sustainability of bio-based chemical production.



End-Product	Basic and Applied R & D	Demonstration	Early Commercial	Commercial
Biodiesel	Lignocellulosic Sugar-based HL		Vegetable oil and animal fat	
Bioethanol	Algal biomass	Lignocellulosic	Corn stover, wheat straw	Corn, sugarcane, beet sugar
Biomethane			Food waste	Biogas
Commodity chemicals		Lignocellulosic	Corn stover Agric waste	Corn, sugarcane
Pharmaceutical agents/ precursors			Sugar and other carbohydrates	

Figure 1.5. Current stages in the use of various renewable feedstocks for commercial products.

Table 1.1. Examples of strategies and chemicals produced from lignocellulose and food waste as renewable sources.

Carbon source	Organism/Strains	Target compound	Yield g/g	Reference
L-Arabinose	*S. cerevisiae* AH22	L-Arabitol	0.620	Jacob et. al 2023
Xylose	*O. polymorpha*	3-Hydroxypropionic acid	0.200	Gao et. al 2023
Technical lignin	*C. cellulolyticum* and ADP1_undA	1-Undecene	0.003	Salmela et. al 2020
Pure cellulose	*C. cellulolyticum* and ADP1_undA	1-Undecene	0.004	Salmela et. al 2020
Corn stover	Microbial consortium DUT47	Lactic acid	0.500	Sun et. al 2021
Food waste	*A. succinogenes* Z130	Succinic acid	0.730	Lad et. al 2021
Durian peel waste	*R. mucilaginosa* KKUSY14	Lipids	0.055	Lad et. al 2021

1.4 Exploring methane and carbon dioxide as alternative feedstocks

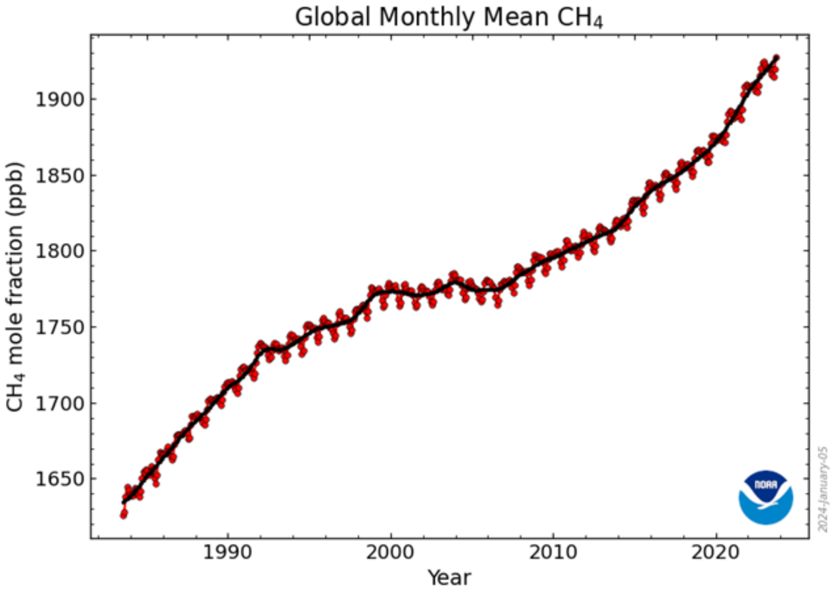
Most industrial fermentation processes utilize carbohydrate substrates, such as glucose and sucrose.[36] However, there is growing interest in using C1 substrates (carbon dioxide, carbon monoxide, methane, methanol, and formate) as these materials are inexpensive, naturally abundant, easy to produce, or available as industrial wastes and by-products.[24,27] This shift is driven by a global movement towards sustainability and the exploration of alternative feedstocks that do not compete with food sources, support a sustainable economy, and reduce carbon emissions.

Carbon dioxide (CO_2) and methane (CH_4) are greenhouse gases that significantly contribute to global warming. CO_2, comprising 0.04% of Earth's atmosphere, is on the rise due to human activities (Figure 1.6, NOAA[55]).[56] Synthetic biology and metabolic engineering present opportunities to capture and convert these gases into valuable products, leveraging natural and biological processes.[1,19,23] Autotrophic organisms can fix CO_2 from the environment using it as a carbon source, and several natural CO_2-fixing pathways have been identified. These include the Calvin-Benson-Bassham (CBB) cycle (Figure 1.7), which is the most prevalent and accounts for 90% of global CO_2 fixation, the reductive TCA (rTCA) cycle, the dicarboxylate/4-hydroxybutyrate cycle, the Wood-Ljungdahl pathway (WLP) (Figure 1.8), and the reductive glycine pathway. These pathways are present in both photosynthetic and non-photosynthetic organisms, thereby demonstrating the broad applicability of these biological systems.

Despite the advantages, the main challenge with next-generation feedstocks (NGFs) like CO_2 and CH_4 lies in the inefficiency of microbes to assimilate them into biomass and bioproducts. These gases, which are potent greenhouse contributors, hold significant potential when repurposed as raw materials in biotechnological and chemical processes. This approach helps mitigate their environmental impact while generating value-added products.[25,57]

Methane, the primary component of natural gas, has been identified as a valuable feedstock for microbial production of chemicals due to its abundance and relatively low cost.[24] Methanotrophic bacteria, which metabolize methane as their sole carbon and energy source, utilize the enzyme methane monooxygenase to convert methane into methanol, a versatile chemical that serves as a building block for various chemical syntheses.[28,57–59] Recent advances have focused on enhancing the efficiency and scalability of methane bioconversion processes, with the goal of developing robust microbial cell factories for the production of a wide array of chemicals.[24,28]

Similarly, carbon dioxide has attracted interest as a feedstock for the production of fuels and chemicals through biological and electrochemical means.[5,28,60,61] The biotechnological approach usually involves the use of autotrophic microorganisms that are capable of fixing CO_2 through photosynthesis or chemosynthesis, converting it into organic compounds.[62]



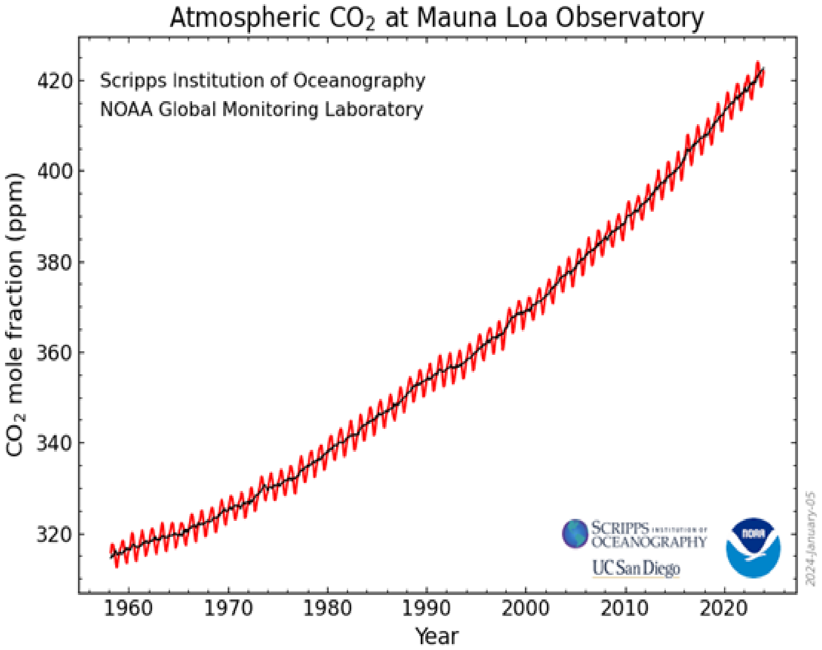


Figure 1.6. Trends in globally averaged methane (top) and carbon dioxide (bottom) levels determined from NOAA Global Monitoring Laboratory measurements.

Recent studies have highlighted the potential of using CO_2 for the bioproduction of high-value products such as methanol, formate, and even biomass, leveraging processes such as gas fermentation and photoautotrophic carbon fixation.[63,64]

The electrochemical conversion of CO_2 into methane and methanol has also been explored, with advancements in catalysis offering pathways to transform CO_2 into valuable C1 and C2+ products. Despite these promising prospects, several challenges remain in utilizing methane and CO_2 as feedstocks.[24,64] These challenges include the development of microbial strains with enhanced carbon utilization capabilities, solubilities of these gases in aqueous systems, and the integration of these processes into existing industrial infrastructures. The exploration of methane and CO_2 as alternative feedstocks are aligned with global efforts to transition towards a more sustainable and circular bioeconomy. By converting these greenhouse gases into valuable chemicals and fuels, it is possible to address the dual challenges of climate change mitigation and the sustainable production of chemicals.[65] Continued research and development in this field are crucial for the realization of the full potential of these feedstocks.

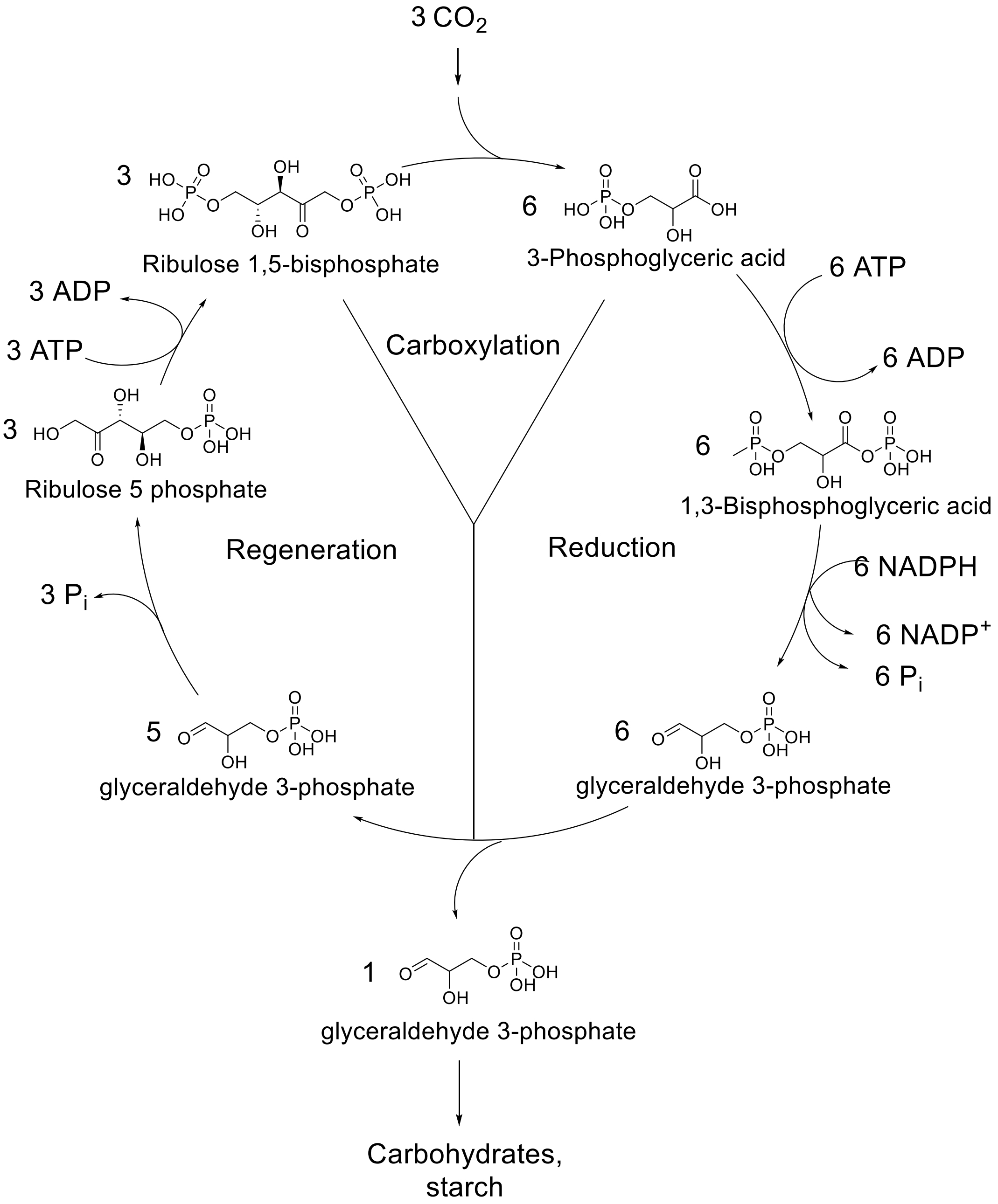


Figure 1.7. The three phases of the Calvin-Benson-Bassham (CBB) cycle. The carboxylation phase is involved in the fixation of CO_2 into 3-phosphoglyceric acid (3-PG). Reduction phase reduces 3-PG to the level of carbohydrates. The final stage (regeneration) involves recycling the CO_2 acceptor Ribulose 1,5-biphosphate.

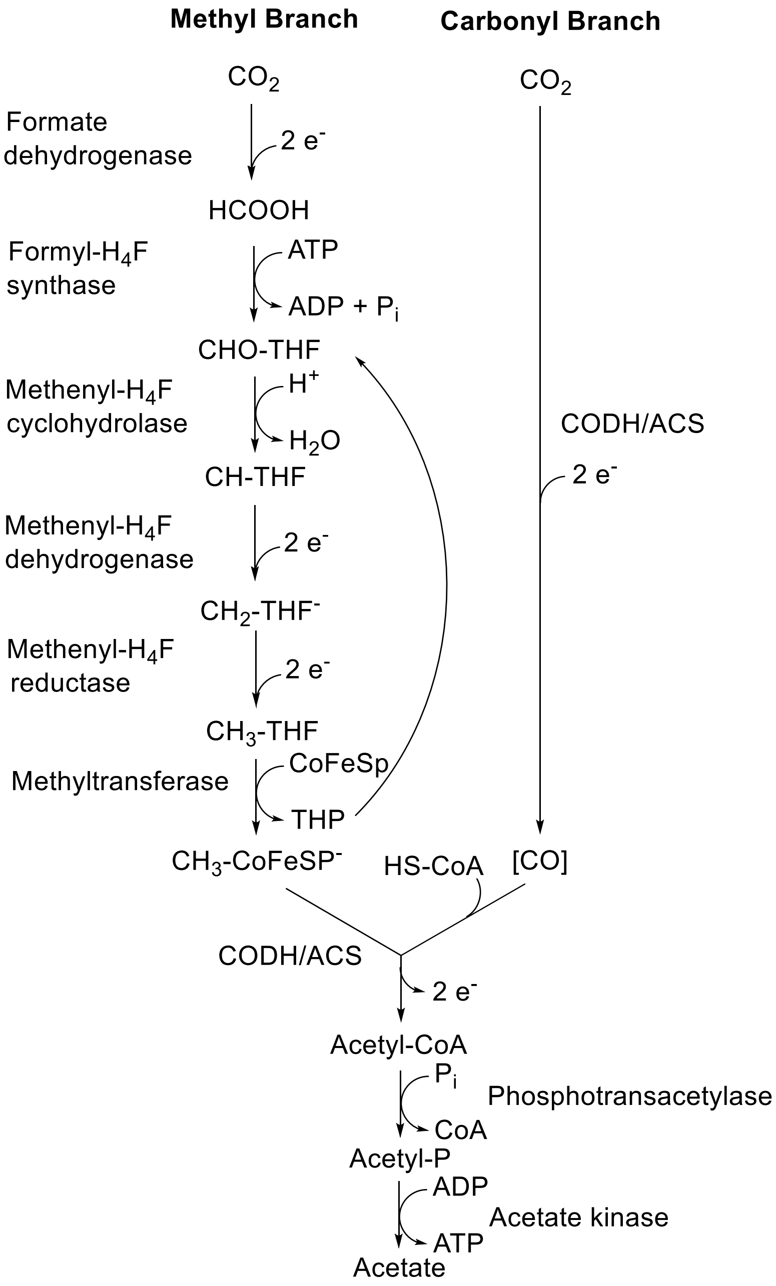


Figure 1.8. The Wood-Ljundahl pathway. This pathway consists of two arms, the methyl and the carbonyl branch. They involve a series of reactions resulting in the reduction of two molecules of carbon dioxide leading to the production of acetate. No net ATP is generated, and energy production is dependent on chemiosmotic processes couples to the translocation of protons or sodium ions. This figure was redrafted from Westerholm et al., 2016.

1.5 C4-dicarboxylates as potential substrates for microbial growth and bioproduction of chemicals

The exploration of C4-dicarboxylates, including succinate, malate, fumarate, and aspartate, as potential substrates for microbial growth and chemical production represents a promising area of research in industrial biotechnology.[66] These compounds are readily integrated into microbial pathways due to their central roles in microbial metabolism. Most of these C4-dicarboxylates serve as key intermediates in the citric acid cycle (TCA cycle), which is pivotal for aerobic metabolism. They provide energy and precursors for biosynthesis. Under anaerobic conditions, these compounds can also participate in fermentation processes.[67] For example, fumarate can act as an electron acceptor in anaerobic respiration. This dual role makes C4-dicarboxylates versatile substrates that support the growth of various microorganisms under diverse conditions. Microbes such as *Escherichia coli* and *Salmonella enterica* can utilize these compounds under both aerobic and anaerobic conditions, facilitated by specific transport systems.[66,68,69] One such system is the dicarboxylate uptake (DCU) system in *E. coli*, which includes transporters such as DcuB and DctA.

The utilization of C4-dicarboxylates for the bioproduction of chemicals offers significant environmental benefits by reducing reliance on fossil fuels and decreasing carbon emissions. Economically, these processes can lead to cost-effective production methods for high-demand chemicals, thereby contributing to more sustainable industrial practices. Several microorganisms have been identified to grow on C4-dicarboxylates,[69,70] utilizing these compounds as carbon sources for growth and metabolism. *P. syringae* and *P. aeruginosa* have been shown to prefer C4-dicarboxylates over other carbon sources.[71,72]

In 1958, Yamada et al. reported the discovery of an enzyme system capable of converting acetylenedicarboxylic acid (ADCA) to pyruvate.[73] A *Pseudomonas sp.* was isolated from the soil and found to be capable of growth on ADCA (the unnatural acetylenic compound) as a sole carbon source. An enzyme that possessed both hydratase and decarboxylase activities was partially purified from this strain, which directly converted ADCA to pyruvate while generating CO_2. The authors concluded that there was no evidence to suggest the involvement of more than one enzyme and that no cofactor was required. Additionally, it was observed that oxaloacetic acid, hydroxyfumaric acid, and acetylenecarboxylic acid are not free intermediates in the reaction. While these findings demonstrated a link between ADCA and microbial metabolism, neither the enzyme nor the strain was characterized nor archived.

Subsequently, in 1994, Flint et al. characterized an *E. coli* fumarase, FumA, which is capable of hydrating ADCA to oxaloacetic acid.[74,75] Oxaloacetic acid is a TCA-cycle intermediate that can be integrated directly into the TCA cycle. Given that *E. coli* possesses a sophisticated C4-dicarboxylate transport system, we hypothesized that ADCA will be an excellent carbon source provided it is assimilated by the cells.

1.6 Acetylenedicarboxylate from methane and carbon dioxide

The low solubility of CH_4 in water presents a significant challenge for its bioconversion in aqueous systems.[76] This requires the use of advanced process engineering to facilitate effective contact between methane and microbial cells in large-scale operations. Additionally, the use of methane, especially in mixtures with oxygen, introduces risks due to its flammability. Therefore, strict safety measures and process controls are required in industrial settings to mitigate these risks. Similarly, while certain microorganisms, such as algae and specific bacteria, are capable of fixing CO_2 through photosynthesis or other metabolic pathways, the efficiency of these biological processes is typically low, resulting in insufficient yields of desired products and hampering commercial viability.[36,77]

To solve these challenges, there has been a strategic shift to the use of methane or carbon dioxide-derived feedstocks, such as methanol, acetate, and formic acid.[24,78,79] CH_4 and CO_2 have been converted to C1-feedstocks like formic acid and methanol for microbial utilization. The integration of C1-utilization pathways into heterologous strains often results in metabolite imbalance, presenting further challenges in effectively harnessing these pathways for C1 metabolism. Continued research and development are essential to address these issues.[24,78] This should focus on the discovery and engineering of more efficient microbial strains, the development of improved enzymes, and the optimization of process conditions to enhance yield and reduce energy consumption.

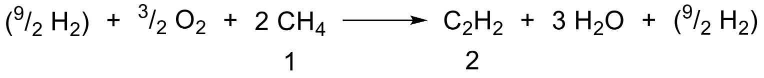
This biotechnological shift paves the way for the development of next-generation feedstocks, which are stable and can be easily metabolized by microbes. This thesis exploits acetylenedicarboxylic acid (ADCA), which can be produced from methane and carbon dioxide for the bioproduction of chemicals. This study opens new possibilities in the field and illustrates the potential of integrating advanced biotechnological strategies with traditional chemical processes to address environmental challenges and create sustainable solutions.

The proposed microbial utilization of acetylenedicarboxylic acid (ADCA) presents an innovative approach to exploiting methane and carbon dioxide as feedstocks for the bioproduction of chemicals. The chemical transformation of methane and carbon dioxide to ADCA involves two steps: first, acetylene is synthesized from methane via partial combustion of methane (Figure 1.9)[80–82], and then the unstable acetylene can be carboxylated to form acetylenedicarboxylate (Figure 1.10).[83–86] One method for the direct conversion of methane to acetylene involves a rotating arc driven by alternating current (A.C.) electrical power. By feeding a high concentration of methane (greater than 43%) diluted in hydrogen through the arc column, this process achieves a low specific energy requirement (SER) of 10.2 kW h kg^{-1} C_2H_2.[87] The use of hydrogen not only suppresses the formation of soot during the methane conversion process but also ensures a carbon balance of greater than 95% and a C_2H_2 selectivity of greater than 90%, with a methane conversion rate exceeding 70% under optimal conditions.

Historically, acetylene production from methane has evolved significantly since the 1930s, a period when acetylene became a critical feedstock for the rubber and fuel industries.[80] Two primary routes exist for the conversion of methane to acetylene: dehydrogenation and oxidative coupling. Dehydrogenation of methane to C2 hydrocarbons, including acetylene, is an endothermic process that requires temperatures above 800 °C to maintain reaction equilibrium. Techniques such as pyrolysis, arc processes, and microwave plasma have been employed in these dehydrogenation processes.[80] To enhance energy efficiency, oxygen is occasionally introduced into the feed, initiating an exothermic reaction that facilitates oxidative methane dehydrogenation.

The oxidative processes, partial combustion or oxidative coupling of methane (OCM), have also been developed. For example, the BASF partial combustion process involves preheating methane to 600 °C before mixing with oxygen and passing it through a burner block. This method produces acetylene with high selectivity, although it also generates significant amounts of soot - approximately 1.25 metric tons daily. Hydrogen is introduced to boost the yield of acetylene to 17% without coking, with modifications by Vedrenne and colleagues further minimizing soot production.[80] These processes not only achieve considerable efficiency in the conversion of methane but also attempt to control the environmental impact by managing by-product formation.

Continued research is essential to optimize these processes further, reduce energy consumption, and enhance the yield and selectivity of acetylene.



$$(^{9}/_{2}\,H_2) + {}^{3}/_{2}\,O_2 + \underset{1}{2\,CH_4} \longrightarrow \underset{2}{C_2H_2} + 3\,H_2O + (^{9}/_{2}\,H_2)$$

Figure 1.9 Reaction for the conversion of methane (1) to acetylene (2) as performed by Vedrenne.

Direct carboxylation of terminal alkynes is thermodynamically unfavorable, and the reaction is driven by the presence of a base that forms stable carboxylate salts.[83,88] The carboxylation of acetylene can be performed with 5 bar CO_2 using (4,7-diphenyl phenanthroline) bis(triphenylphosphine) copper (I), and cesium carbonate. However, this affords a mixture of acetylenemonocarboxylate (ACA) and acetylenedicarboxylate (ADCA). The direct carboxylation of acetylene to solely acetylenedicarboxylic acids has proven to be a challenging reaction, primarily due to the low solubility of acetylene in organic solvents and carbon dioxide, which serves as a co-reactant. This is further complicated by the considerably higher solubility of carbon dioxide in organic solvents, which makes it difficult for acetylene to accumulate in the reaction medium. The salts of acetylenemonocarboxylic acid initially formed also have only a low solubility, which hinders the second carboxylation step from generating the acetylenedicarboxylic acid.

Consequently, product mixtures of acetylenemonocarboxylic acid and acetylenedicarboxylic acid are formed. Moreover, both acids, particularly acetylenedicarboxylic acid, tend to undergo decarboxylation in solution, and this process is further catalyzed by silver and copper complexes. Additionally, acetylene is susceptible to polymerization or vinylation reactions, which can further reduce the yield. Another disadvantage is the accumulation of salts when stoichiometric amounts of base are used on an industrial scale.

An efficient method for producing ADCA from acetylene and carbon dioxide has been developed and patented by BASF.[89] The developed carboxylation reaction takes place in the presence of silver or copper catalysts and an amine base, resulting in the carboxylation of acetylene to acetylenedicarboxylate within 16 h. Suitable silver catalysts maybe elemental silver, colloidal silver particles, which may optionally comprise stabilizing additives like phosphine ligands, dimethyl sulfoxide, or polyvinylpyrrolidone, silver (I) halides such as AgF, AgCl, AgBr, and AgI, and silver nitrate, silver tetrafluoroborate, silver trifluoromethane sulfonate, silver carboxylate (silver acetate), silver hexafluorophosphate, silver oxide, silver sulfate, silver hexafluoroantimonate, silver p-toluenesulfonate, and silver carbonate. Preference is given to silver (I) iodide AgI, silver (I) nitrate AgNO, and silver tetrafluoroborate $AgBF_4$. For the copper catalysts, elemental copper and colloidal copper particles and copper salts chosen from the group consisting of copper (I) halides such as CuF, CuCl, CuBr and Cul, copper (I) cyanide, copper tetrafluoroborate, copper trifluoromethanesulfonate, copper acetate, copper hexafluorophosphate, copper oxide, copper sulfate, copper hexafluoroantimonate, copper p-toluenesulfonate and copper carbonate.

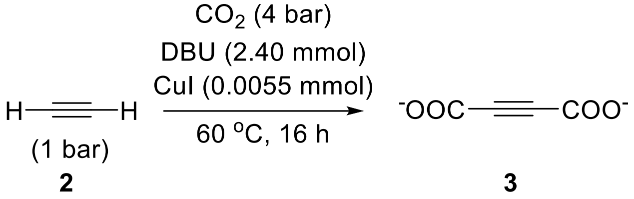


Figure 1.10. Reaction conditions for carboxylation of acetylene (2) to produce acetylenedicarboxylate (3).

CuI. They observed acetylenedicarboxylate formation in the absence of inorganic bases, such as cesium carbonate, when the reaction was performed in the presence of amine bases. Suitable amine bases include alkylamines, particularly tri-C_3-C_6-alkylamines such as tripropylamine and tributylamine, alkanolamines, mainly mono-, di- and tri C_2-C4-alkanolamines such as mono-, diand triethanolamine, and particularly heterocyclic amine bases such as N-methylpiperidine, N-methylpiperidone, N-methylmorpholine, N-methyl-2-pyrrolidone, particularly preferred amine bases are 1,5-diazabicyclo[4.3.0]non-5-ene (DBN) and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU). Surprisingly, the acetylenedicarboxylate produced in this manner does not decarboxylate.

1.7 Adaptive laboratory evolution for improving substrate utilization by microbes

Adaptive Laboratory Evolution (ALE) is a powerful strategy for improving substrate utilization and microbial growth by leveraging the natural process of evolution under controlled laboratory conditions.[90,91] This approach involves subjecting microbial populations to specific environmental conditions or selective pressures over extended periods, allowing for the natural selection of genetic mutations that confer advantageous traits. The ALE process begins with the identification of the desired trait to be improved, such as enhanced substrate utilization[78,92], tolerance to toxic compounds[93,94], or increased growth rate.[95] Microorganisms are then cultivated under conditions that apply selective pressure, favoring the evolution of the desired trait. The selection is typically achieved through continuous cultivation methods such as serial batch cultures or chemostats, which maintain the selective pressure and allow for the accumulation of beneficial mutations over generations.

Throughout the ALE process, the evolution of the microbial population is characterized by assessing growth rates, substrate consumption, metabolomics, transcriptomics, and other relevant phenotypic characteristics. Genomic analyses are also performed to identify mutations that arise during evolution.[91] Fast-growing mutants are isolated, and their genomes are sequenced to pinpoint mutations responsible for improved traits. Further characterization involves studying the metabolic pathways and regulatory networks affected by these mutations.

ALE has been instrumental in enhancing the capacity of microorganisms to utilize non-native or less preferred substrates. For instance, *E. coli* and *S. cerevisiae* have evolved to efficiently metabolize xylose, a significant sugar found in lignocellulosic biomass, and methanol, respectively.[96–98] This has led to an improvement in the feasibility of biofuel production from renewable resources. Furthermore, ALE has enabled microorganisms to tolerate higher concentrations of inhibitory compounds present in feedstocks, such as furans and phenolics derived from lignocellulosic hydrolysates. This approach enhances the robustness of microbial strains in industrial bioprocesses. Furthermore, ALE can facilitate the emergence of mutations that enhance the growth rates of microorganisms under specific conditions, thereby improving the overall productivity of bioproduction processes.

Despite the significant achievements that have been made, there are still obstacles to be overcome in fully comprehending the complex genetic and metabolic alterations that occur during ALE.[90,91] Future research may involve integrating ALE with advanced genomic, transcriptomic, and metabolomic analyses in order to gain deeper insights into evolutionary mechanisms. Furthermore, the integration of ALE with rational genetic engineering approaches, such as CRISPR-Cas systems, could facilitate the accelerated development of strains and expand their applications in industrial biotechnology. This integrative approach will continue to expand the frontiers of what is feasible with microbial production systems, thereby facilitating the development of more sustainable and efficient industrial practices.

REFERENCES

(1) Palkovits, R.; Delidovich, I. Efficient Utilization of Renewable Feedstocks: The Role of Catalysis and Process Design. *Philos. Trans. A Math. Phys. Eng. Sci.* **2018**, *376* (2110). https://doi.org/10.1098/rsta.2017.0064.

(2) Chapter 1.14: Renewable Energy Sources https://bio.libretexts.org/Bookshelves/Ecology/AP_Environmental_Science/01%3A_Chapters/1.14%3A_Renewable_Energy_Sources (accessed May 6, 2024).

(3) Miller, R. G.; Sorrell, S. R. The Future of Oil Supply. *Philos. Trans. A Math. Phys. Eng. Sci.* **2014**, *372* (2006), 20130179. https://doi.org/10.1098/rsta.2013.0179.

(4) Lv, Y. Transitioning to Sustainable Energy: Opportunities, Challenges, and the Potential of Blockchain Technology. *Frontiers in Energy Research* **2023**, *11*. https://doi.org/10.3389/fenrg.2023.1258044.

(5) Yadav, V. G.; Yadav, G. D.; Patankar, S. C. The Production of Fuels and Chemicals in the New World: Critical Analysis of the Choice between Crude Oil and Biomass Vis-à-Vis Sustainability and the Environment. *Clean Technol. Environ. Policy* **2020**, 18–1 ,() . https://doi.org/10.1007/s10098-020-01945-5.

(6) Yüzbaşıoğlu, A. E.; Tatarhan, A. H.; Gezerman, A. O. Decarbonization in Ammonia Production, New Technological Methods in Industrial Scale Ammonia Production and Critical Evaluations. *Heliyon* **2021**, *7* (10), e08257. https://doi.org/10.1016/j.heliyon.2021.e08257.

(7) Ghavam, S.; Vahdati, M.; Wilson, I. A. G.; Styring, P. Sustainable Ammonia Production Processes. *Frontiers in Energy Research* **2021**, *9*. https://doi.org/10.3389/fenrg.2021.580808.

(8) Petroleum - Global, Reserves, Production. *Encyclopedia Britannica*.

(9) Petroleum - Global Supply, Demand, Prices. *Encyclopedia Britannica*.

(10) Perera, F. Pollution from Fossil-Fuel Combustion Is the Leading Environmental Threat to Global Pediatric Health and Equity: Solutions Exist. *Int. J. Environ. Res. Public Health* **2017**, *15* (1). https://doi.org/10.3390/ijerph15010016.

(11) Asif, Z.; Chen, Z.; An, C.; Dong, J. Environmental Impacts and Challenges Associated with Oil Spills on Shorelines. *J. Mar. Sci. Eng.* **2022**, *10* (6), 762. https://doi.org/10.3390/jmse10060762.

(12) National Research Council (US) Committee on Oil; Fates; Effects. *Biological Effects of Oil Releases*; National Academies Press (US), 2003.

(13) Fossil https://www.energy.gov/fossil (accessed May 6, 2024).

(14) Ritchie, H.; Rosado, P.; Roser, M. Fossil Fuels. *Our World in Data* **2024**.

(15) Chapter 15.6: Coal - The Carbon Rock of Ages https://chem.libretexts.org/Bookshelves/Introductory_Chemistry/Chemistry_for_Changing_Times_(Hill_and_McCreary)/15%3A_Energy/15.06%3A_Coal_-_The_Carbon_Rock_of_Ages (accessed May 6, 2024).

(16) The United States consumed a record amount of renewable energy in 2020 - U.S. Energy Information Administration (EIA) https://www.eia.gov/todayinenergy/detail.php?id=48396 (accessed May 6, 2024).

(17) Strielkowski, W.; Civín, L.; Tarkhanova, E.; Tvaronavičienė, M.; Petrenko, Y. Renewable Energy in the Sustainable Development of Electrical Power Sector: A Review. *Energies* **2021**, *14* (24), 8240. https://doi.org/10.3390/en14248240.

(18) In the first half of 2022, 24% of U.S. electricity generation came from renewable sources - U.S. Energy Information Administration (EIA) https://www.eia.gov/todayinenergy/detail.php?id=53779 (accessed May 6, 2024).

(19) Lobus, N. V.; Knyazeva, M. A.; Popova, A. F.; Kulikovskiy, M. S. Carbon Footprint Reduction and Climate Change Mitigation: A Review of the Approaches, Technologies, and Implementation Challenges. *Nuovo Cimento: C: Geophys. Space Phys.* **2023**, *9* (4), 120. https://doi.org/10.3390/c9040120.

(20) Ragazzi, I.; Farley, M.; Jeffery, K.; Butnar, I. Using Life Cycle Assessments to Guide Reduction in the Carbon Footprint of Single-Use Lab Consumables. *PLOS Sustainability and Transformation* **2023**, *2* (9), e0000080. https://doi.org/10.1371/journal.pstr.0000080.

(21) Centi, G.; Iaquaniello, G.; Perathoner, S. Chemical Engineering Role in the Use of Renewable Energy and Alternative Carbon Sources in Chemical Production. *BMC Chemical Engineering* **2019**, *1* (1), 1–16. https://doi.org/10.1186/s42480-019-0006-8.

(22) Ulber, R.; Muffler, K.; Tippkötter, N.; Hirth, T.; Sell, D. Introduction to Renewable Resources in the Chemical Industry. In *Renewable Raw Materials*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2011; pp 1–5. https://doi.org/10.1002/9783527634194.ch1.

(23) Alok, A.; Shrestha, R.; Ban, S.; Devkota, S.; Uprety, B.; Joshi, R. Technological Advances in the Transformative Utilization of CO2 to Value-Added Products. *Journal of Environmental Chemical Engineering* **2022**, *10* (1), 106922. https://doi.org/10.1016/j.jece.2021.106922.

(24) Jiang, W.; Hernández Villamor, D.; Peng, H.; Chen, J.; Liu, L.; Haritos, V.; Ledesma-Amaro, R. Metabolic Engineering Strategies to Enable Microbial Utilization of C1 Feedstocks. *Nat. Chem. Biol.* **2021**, *17* (8), 845–855. https://doi.org/10.1038/s41589-021-00836-0.

(25) He, L.; Groom, J. D.; Wilson, E. H.; Fernandez, J.; Konopka, M. C.; Beck, D. A. C.; Lidstrom, M. E. A Methanotrophic Bacterium to Enable Methane Removal for Climate Mitigation. *Proc. Natl. Acad. Sci. U. S. A.* **2023**, *120* (35), e2310046120. https://doi.org/10.1073/pnas.2310046120.

(26) Jeon, Y. C.; Nguyen, A. D.; Lee, E. Y. Bioproduction of Isoprenoids and Other Secondary Metabolites Using Methanotrophic Bacteria as an Alternative Microbial Cell Factory Option: Current Stage and Future Aspects. *Catalysts* **2019**, *9* (11), 883. https://doi.org/10.3390/catal9110883.

(27) Zhang, C.; Ottenheim, C.; Weingarten, M.; Ji, L. Microbial Utilization of Next-Generation Feedstocks for the Biomanufacturing of Value-Added Chemicals and Food Ingredients. *Front Bioeng Biotechnol* **2022**, *10*, 874612. https://doi.org/10.3389/fbioe.2022.874612.

(28) Sahoo, K. K.; Goswami, G.; Das, D. Biotransformation of Methane and Carbon Dioxide Into High-Value Products by Methanotrophs: Current State of Art and Future Prospects. *Front. Microbiol.* **2021**, *12*, 636486. https://doi.org/10.3389/fmicb.2021.636486.

(29) Almomani, F.; Abdelbar, A.; Ghanimeh, S. A Review of the Recent Advancement of Bioconversion of Carbon Dioxide to Added Value Products: A State of the Art. *Sustain. Sci. Pract. Policy* **2023**, *15* (13), 10438. https://doi.org/10.3390/su151310438.

(30) Wang, J.; Ma, W.; Wang, X. Insights into the Structure of Escherichia Coli Outer Membrane as the Target for Engineering Microbial Cell Factories. *Microb. Cell Fact.* **2021**, *20* (1), 73. https://doi.org/10.1186/s12934-021-01565-8.

(31) Ko, Y.-S.; Kim, J. W.; Lee, J. A.; Han, T.; Kim, G. B.; Park, J. E.; Lee, S. Y. Tools and Strategies of Systems Metabolic Engineering for the Development of Microbial Cell Factories for Chemical Production. *Chem. Soc. Rev.* **2020**, *49* (14), 4615–4636. https://doi.org/10.1039/d0cs00155d.

(32) Cho, J. S.; Kim, G. B.; Eun, H.; Moon, C. W.; Lee, S. Y. Designing Microbial Cell Factories for the Production of Chemicals. *JACS Au* **2022**, *2* (8), 1781–1799. https://doi.org/10.1021/jacsau.2c00344.

(33) Chaudhary, R.; Nawaz, A.; Fouillaud, M.; Dufossé, L.; Haq, I. ul; Mukhtar, H. Microbial Cell Factories: Biodiversity, Pathway Construction, Robustness, and Industrial Applicability. *Microbiol. Res.* **2024**, *15* (1), 247–272. https://doi.org/10.3390/microbiolres15010018.

(34) Kang, M.-K.; Yoon, S.-H.; Kwon, M.; Kim, S.-W. Microbial Cell Factories for Bio-Based Isoprenoid Production to Replace Fossil Resources. *Current Opinion in Systems Biology* **2024**, *37* (4), 100502. https://doi.org/10.1016/j.coisb.2023.100502.

(35) Hong, K.-K.; Nielsen, J. Metabolic Engineering of Saccharomyces Cerevisiae: A Key Cell Factory Platform for Future Biorefineries. *Cell. Mol. Life Sci.* **2012**, *69* (16), 2671–2690. https://doi.org/10.1007/s00018-012-0945-1.

(36) Sivaramakrishnan, R.; Suresh, S.; Kanwal, S.; Ramadoss, G.; Ramprakash, B.; Incharoensakdi, A. Microalgal Biorefinery Concepts' Developments for Biofuel and Bioproducts: Current Perspective and Bottlenecks. *Int. J. Mol. Sci.* **2022**, *23* (5). https://doi.org/10.3390/ijms23052623.

(37) Tuo, J.; Nawab, S.; Ma, X.; Huo, Y.-X. Recent Advances in Screening Amino Acid Overproducers. *Engineering Microbiology* **2023**, *3* (1), 100066. https://doi.org/10.1016/j.engmic.2022.100066.

(38) Wang, S.; Cui, J.; Bilal, M.; Hu, H.; Wang, W.; Zhang, X. Pseudomonas Spp. as Cell Factories (MCFs) for Value-Added Products: From Rational Design to Industrial Applications. *Crit. Rev. Biotechnol.* **2020**, *40* (8), 1232–1249. https://doi.org/10.1080/07388551.2020.1809990.

(39) Lee, S. Y.; Kim, H. U.; Park, J. H.; Park, J. M.; Kim, T. Y. Metabolic Engineering of Microorganisms: General Strategies and Drug Production. *Drug Discov. Today* **2009**, *14* (1–2), 78–88. https://doi.org/10.1016/j.drudis.2008.08.004.

(40) Han, H.; Yu, C.; Qi, J.; Wang, P.; Zhao, P.; Gong, W.; Xie, C.; Xia, X.; Liu, C. High-Efficient Production of Mushroom Polyketide Compounds in a Platform Host *Aspergillus oryzae*. *Microb. Cell Fact.* **2023**, *22* (1), 60. https://doi.org/10.1186/s12934-023-02071-9.

(41) Mubarak, M.; Shaija, A.; Suchithra, T. V. A Review on the Extraction of Lipid from Microalgae for Biodiesel Production. *Algal Research* **2014**, *7*. https://doi.org/10.1016/j.algal.2014.10.008.

(42) Gao, Z.; Zhao, H.; Li, Z.; Tan, X.; Lu, X. Photosynthetic Production of Ethanol from Carbon Dioxide in Genetically Engineered Cyanobacteria. *Energy Environ. Sci.* **2012**, *5* (12), 9857–9865. https://doi.org/10.1039/C2EE22675H.

(43) Ismail, W. A.; Van Hamme, J. D.; Kilbane, J. J.; Gu, J.-D. Editorial: Petroleum Microbial Biotechnology: Challenges and Prospects. *Front. Microbiol.* **2017**, *8*, 833. https://doi.org/10.3389/fmicb.2017.00833.

(44) Singh, A.; Prajapati, P.; Vyas, S.; Gaur, V. K.; Sindhu, R.; Binod, P.; Kumar, V.; Singhania, R. R.; Awasthi, M. K.; Zhang, Z.; Varjani, S. A Comprehensive Review of Feedstocks as Sustainable Substrates for Next-Generation Biofuels. *Bioenergy Res.* **2023**, *16* (1), 105–122. https://doi.org/10.1007/s12155-022-10440-2.

(45) Osman, A. I.; Mehta, N.; Elgarahy, A. M.; Al-Hinai, A.; Al-Muhtaseb, A. H.; Rooney, D. W. Conversion of Biomass to Biofuels and Life Cycle Assessment: A Review. *Environ. Chem. Lett.* **2021**, *19* (6), 4075–4118. https://doi.org/10.1007/s10311-021-01273-0.

(46) Overview https://www.etipbioenergy.eu/value-chains/feedstocks/biofuels-feedstocks-an-overview (accessed May 4, 2024).

(47) Mahapatra, S.; Kumar, D.; Singh, B.; Sachan, P. K. Biofuels and Their Sources of Production: A Review on Cleaner Sustainable Alternative against Conventional Fuel, in the

Framework of the Food and Energy Nexus. *Energy Nexus* **2021**, *4*, 100036. https://doi.org/10.1016/j.nexus.2021.100036.

(48) Bentsen, N. S.; Felby, C. Biomass for Energy in the European Union - a Review of Bioenergy Resource Assessments. *Biotechnol. Biofuels* **2012**, *5* (1), 25. https://doi.org/10.1186/1754-6834-5-25.

(49) Bozell, J. J.; Petersen, G. R. Technology Development for the Production of Biobased Products from Biorefinery Carbohydrates-the US Department of Energy's "Top 10" Revisited. *Green Chem.* **2010**, *12* (4), 539–554. https://doi.org/10.1039/b922014c.

(50) McGarrity, J. EU Council backs 7% crop cap in Renewable Energy Directive https://www.agricensus.com/Article/EU-Council-backs-7-crop-cap-in-Renewable-Energy-Directive-23021.html (accessed May 8, 2024).

(51) Escobar, N.; Laibach, N. Sustainability Check for Bio-Based Technologies: A Review of Process-Based and Life Cycle Approaches. *Renewable Sustainable Energy Rev.* **2021**, *135*, 110213. https://doi.org/10.1016/j.rser.2020.110213.

(52) Liu, Y.; Tang, Y.; Gao, H.; Zhang, W.; Jiang, Y.; Xin, F.; Jiang, M. Challenges and Future Perspectives of Promising Biotechnologies for Lignocellulosic Biorefinery. *Molecules* **2021**, *26* (17). https://doi.org/10.3390/molecules26175411.

(53) Ragauskas, A. J.; Beckham, G. T.; Biddy, M. J.; Chandra, R.; Chen, F.; Davis, M. F.; Davison, B. H.; Dixon, R. A.; Gilna, P.; Keller, M.; Langan, P.; Naskar, A. K.; Saddler, J. N.; Tschaplinski, T. J.; Tuskan, G. A.; Wyman, C. E. Lignin Valorization: Improving Lignin Processing in the Biorefinery. *Science* **2014**, *344* (6185), 1246843. https://doi.org/10.1126/science.1246843.

(54) Linger, J. G.; Vardon, D. R.; Guarnieri, M. T.; Karp, E. M.; Hunsinger, G. B.; Franden, M. A.; Johnson, C. W.; Chupka, G.; Strathmann, T. J.; Pienkos, P. T.; Beckham, G. T. Lignin Valorization through Integrated Biological Funneling and Chemical Catalysis. *Proceedings of the National Academy of Sciences* **2014**, *111* (33), 12013–12018. https://doi.org/10.1073/pnas.1410657111.

(55) US Department of Commerce; NOAA; Global Monitoring Laboratory. Global Monitoring Laboratory - Carbon Cycle Greenhouse Gases. **2005**.

(56) Science X. Atmospheric research provides clear evidence of human-caused climate change signal associated with CO2 increases https://phys.org/news/2023-05-atmospheric-evidence-human-caused-climate-co2.html (accessed May 8, 2024).

(57) Bennett, R. K.; Steinberg, L. M.; Chen, W.; Papoutsakis, E. T. Engineering the Bioconversion of Methane and Methanol to Fuels and Chemicals in Native and Synthetic Methylotrophs. *Curr. Opin. Biotechnol.* **2018**, *50*, 81–93. https://doi.org/10.1016/j.copbio.2017.11.010.

(58) Henard, C. A.; Akberdin, I. R.; Kalyuzhnaya, M. G.; Guarnieri, M. T. Muconic Acid Production from Methane Using Rationally-Engineered Methanotrophic Biocatalysts. *Green Chem.* **2019**, *21* (24), 6731–6737. https://doi.org/10.1039/c9gc03722e.

(59) Wang, Y.; Fan, L.; Tuyishime, P.; Liu, J.; Zhang, K.; Gao, N.; Zhang, Z.; Ni, X.; Feng, J.; Yuan, Q.; Ma, H.; Zheng, P.; Sun, J.; Ma, Y. Adaptive Laboratory Evolution Enhances Methanol Tolerance and Conversion in Engineered *Corynebacterium glutamicum*. *Commun Biol* **2020**, *3* (1), 217. https://doi.org/10.1038/s42003-020-0954-9.

(60) Harary, I.; Korey, S. R.; Ochoa, S. Biosynthesis of Dicarboxylic Acids by Carbon Dioxide Fixation. VII. Equilibrium of Malic Enzyme Reaction. *J. Biol. Chem.* **1953**, *203* (2), 595–604.

(61) National Research Council (US) Chemical Sciences Roundtable. *Carbon Dioxide as a Feedstock*; National Academies Press (US), 2001.

(62) Gleizer, S.; Ben-Nissan, R.; Bar-On, Y. M.; Antonovsky, N.; Noor, E.; Zohar, Y.; Jona, G.; Krieger, E.; Shamshoum, M.; Bar-Even, A.; Milo, R. Conversion of *Escherichia coli* to Generate All Biomass Carbon from CO_2. *Cell* **2019**, *179* (6), 1255-1263.e12. https://doi.org/10.1016/j.cell.2019.11.009.

(63) Liang, F.; Lindberg, P.; Lindblad, P. Engineering Photoautotrophic Carbon Fixation for Enhanced Growth and Productivity. *Sustainable Energy Fuels* **2018**, *2* (12), 2583–2600. https://doi.org/10.1039/C8SE00281A.

(64) Kurt, E.; Qin, J.; Williams, A.; Zhao, Y.; Xie, D. Perspectives for Using CO_2 as a Feedstock for Biomanufacturing of Fuels and Chemicals. *Bioengineering (Basel)* **2023**, *10* (12). https://doi.org/10.3390/bioengineering10121357.

(65) Burkart, M. D.; Hazari, N.; Tway, C. L.; Zeitler, E. L. Opportunities and Challenges for Catalysis in Carbon Dioxide Utilization. *ACS Catal.* **2019**, *9* (9), 7937–7956. https://doi.org/10.1021/acscatal.9b02113.

(66) Schubert, C.; Unden, G. C4-Dicarboxylates as Growth Substrates and Signaling Molecules for Commensal and Pathogenic Enteric Bacteria in Mammalian Intestine. *J. Bacteriol.* **2022**, *204* (4), e0054521. https://doi.org/10.1128/JB.00545-21.

(67) Engel, P.; Krämer, R.; Unden, G. Anaerobic Fumarate Transport in *Escherichia coli* by an Fnr-Dependent Dicarboxylate Uptake System Which Is Different from the Aerobic Dicarboxylate Uptake System. *J. Bacteriol.* **1992**, *174* (17), 5533–5539. https://doi.org/10.1128/jb.174.17.5533-5539.1992.

(68) Kay, W. W.; Kornberg, H. L. The Uptake of C4-Dicarboxylic Acids by *Escherichia coli*. *Eur. J. Biochem.* **1971**, *18* (2), 274–281. https://doi.org/10.1111/j.1432-1033.1971.tb01240.x.

(69) Janausch, I. G.; Zientz, E.; Tran, Q. H.; Kröger, A.; Unden, G. C4-Dicarboxylate Carriers and Sensors in Bacteria. *Biochim. Biophys. Acta* **2002**, *1553* (1–2), 39–56. https://doi.org/10.1016/s0005-2728(01)00233-x.

(70) Rosa, L. T.; Dix, S. R.; Rafferty, J. B.; Kelly, D. J. A New Mechanism for High-Affinity Uptake of C4-Dicarboxylates in Bacteria Revealed by the Structure of *Rhodopseudomonas palustris* MatC (RPA3494), a Periplasmic Binding Protein of the Tripartite Tricarboxylate Transporter (TTT) Family. *J. Mol. Biol.* **2019**, *431* (2), 351–367. https://doi.org/10.1016/j.jmb.2018.11.016.

(71) Valentini, M.; Storelli, N.; Lapouge, K. Identification of C(4)-Dicarboxylate Transport Systems in *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* **2011**, *193* (17), 4307–4316. https://doi.org/10.1128/JB.05074-11.

(72) Martín-Mora, D.; Ortega, Á.; Pérez-Maldonado, F. J.; Krell, T.; Matilla, M. A. The Activity of the C4-Dicarboxylic Acid Chemoreceptor of *Pseudomonas aeruginosa* Is Controlled by Chemoattractants and Antagonists. *Sci. Rep.* **2018**, *8* (1), 2102. https://doi.org/10.1038/s41598-018-20283-7.

(73) Yamada, E. W.; Jakoby, W. B. Enzymatic Utilization Of An Acetylenic Compound. *J. Am. Chem. Soc.* **1958**, *80* (9), 2343–2344. https://doi.org/10.1021/ja01542a089.

(74) Flint, D. H.; Emptage, M. H.; Guest, J. R. Fumarase a from *Escherichia coli*: Purification and Characterization as an Iron-Sulfur Cluster Containing Enzyme. *Biochemistry* **1992**, *31* (42), 10331–10337. https://doi.org/10.1021/bi00157a022.

(75) Flint, D. H.; McKay, R. G. *Escherichia coli* Fumarase A Catalyzed Transfer of 18O from C-2 and 2H from C-3 of Malate to Acetylene Dicarboxylate To Form 18O and 2H Labeled Oxalacetate. *J. Am. Chem. Soc.* **1994**, *116* (13), 5534–5539. https://doi.org/10.1021/ja00092a002.

(76) Grabowska, J.; Blazquez, S.; Sanz, E.; Zerón, I. M.; Algaba, J.; Míguez, J. M.; Blas, F. J.; Vega, C. Solubility of Methane in Water: Some Useful Results for Hydrate Nucleation. *J. Phys. Chem. B* **2022**, *126* (42), 8553–8570. https://doi.org/10.1021/acs.jpcb.2c04867.

(77) Burlacot, A.; Dao, O.; Auroy, P.; Cuiné, S.; Li-Beisson, Y.; Peltier, G. Alternative Photosynthesis Pathways Drive the Algal CO_2-Concentrating Mechanism. *Nature* **2022**, *605* (7909), 366–371. https://doi.org/10.1038/s41586-022-04662-9.

(78) Calvey, C. H.; Sànchez I Nogué, V.; White, A. M.; Kneucker, C. M.; Woodworth, S. P.; Alt, H. M.; Eckert, C. A.; Johnson, C. W. Improving Growth of *Cupriavidus necator* H16 on Formate Using Adaptive Laboratory Evolution-Informed Engineering. *Metab. Eng.* **2023**, *75*, 78–90. https://doi.org/10.1016/j.ymben.2022.10.016.

(79) Burgstaller, L.; Löffler, S.; De Marcellis, L.; Ghassemi, K.; Neureiter, M. The Influence of Different Carbon Sources on Growth and Single Cell Oil Production in Oleaginous Yeasts *Apiotrichum brassicae* and Pichia Kudriavzevii. *N. Biotechnol.* **2022**, *69*, 1–7. https://doi.org/10.1016/j.nbt.2022.02.003.

(80) Jiang, C.; Hu, J. Conversion of Methane to Acetylene; 2023. https://doi.org/10.1039/9781839160257-00093.

(81) Ma, Y.; Han, X.; Xu, S.; Li, Z.; Lu, W.; An, B.; Lee, D.; Chansai, S.; Sheveleva, A. M.; Wang, Z.; Chen, Y.; Li, J.; Li, W.; Cai, R.; da Silva, I.; Cheng, Y.; Daemen, L. L.; Tuna, F.; McInnes, E. J. L.; Hughes, L.; Manuel, P.; Ramirez-Cuesta, A. J.; Haigh, S. J.; Hardacre, C.; Schröder, M.; Yang, S. Direct Conversion of Methane to Ethylene and Acetylene over an Iron-Based Metal-Organic Framework. *J. Am. Chem. Soc.* **2023**, *145* (38), 20792–20800. https://doi.org/10.1021/jacs.3c03935.

(82) Wang, S.; Uwakwe, K.; Yu, L.; Ye, J.; Zhu, Y.; Hu, J.; Chen, R.; Zhang, Z.; Zhou, Z.; Li, J.; Xie, Z.; Deng, D. Highly Efficient Ethylene Production via Electrocatalytic Hydrogenation of Acetylene under Mild Conditions. *Nat. Commun.* **2021**, *12* (1), 7072. https://doi.org/10.1038/s41467-021-27372-8.

(83) Manjolinho, F.; Arndt, M.; Goossen, K.; Goossen, L. J. Catalytic C-H Carboxylation of Terminal Alkynes with Carbon Dioxide. *ACS Catal.* **2012**, *2* (9), 2014–2021. https://doi.org/10.1021/cs300448v.

(84) Zhang, X.; Zhang, W.-Z.; Ren, X.; Zhang, L.-L.; Lu, X.-B. Ligand-Free Ag(I)-Catalyzed Carboxylation of Terminal Alkynes with CO_2. *Org. Lett.* **2011**, *13* (9), 2402–2405. https://doi.org/10.1021/ol200638z.

(85) Yu, D.; Zhang, Y. Copper- and Copper-N-Heterocyclic Carbene-Catalyzed C-H Activating Carboxylation of Terminal Alkynes with CO_2 at Ambient Conditions. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (47), 20184–20189. https://doi.org/10.1073/pnas.1010962107.

(86) Cheng, H.; Zhao, B.; Yao, Y.; Lu, C. Carboxylation of Terminal Alkynes with CO_2 Catalyzed by Bis(Amidate) Rare-Earth Metal Amides. *Green Chem.* **2015**, *17* (3), 1675–1682. https://doi.org/10.1039/c4gc02200a.

(87) Dinh, D. K.; Lee, D. H.; Song, Y.-H.; Jo, S.; Kim, K.-T.; Iqbal, M.; Kang, H. Efficient Methane-to-Acetylene Conversion Using Low-Current Arcs. *RSC Adv.* **2019**, *9* (56), 32403–32413. https://doi.org/10.1039/c9ra05964d.

(88) Wang, X.; Lim, Y. N.; Lee, C.; Jang, H.-Y.; Lee, B. Y. 1,5,7-Triazabicyclo[4.4.0]Dec-1-Ene-Mediated Acetylene Dicarboxylation and Alkyne Carboxylation Using Carbon Dioxide. *Eur. J. Org. Chem.* **2013**, *2013* (10), 1867–1871. https://doi.org/10.1002/ejoc.201201608.

(89) Arndt, M.; Goossen, L. J. Method for Producing Acetylenedicarboxylic Acid from Acetylene and Carbon Dioxide. 20140058132:A1, February 27, 2014.

(90) Long, C. P.; Antoniewicz, M. R. How Adaptive Evolution Reshapes Metabolism to Improve Fitness: Recent Advances and Future Outlook. *Curr. Opin. Chem. Eng.* **2018**, *22*, 209–215. https://doi.org/10.1016/j.coche.2018.11.001.

(91) Dragosits, M.; Mattanovich, D. Adaptive Laboratory Evolution -- Principles and Applications for Biotechnology. *Microb. Cell Fact.* **2013**, *12*, 64. https://doi.org/10.1186/1475-2859-12-64.

(92) Ueki, T. Key Enzymes for Anaerobic Lactate Metabolism in Geobacter Sulfurreducens. *Appl. Environ. Microbiol.* **2021**, *87* (2). https://doi.org/10.1128/AEM.01968-20.

(93) Matson, M. M.; Cepeda, M. M.; Zhang, A.; Case, A. E.; Kavvas, E. S.; Wang, X.; Carroll, A. L.; Tagkopoulos, I.; Atsumi, S. Adaptive Laboratory Evolution for Improved Tolerance of Isobutyl Acetate in *Escherichia coli*. *Metab. Eng.* **2021**, *69*, 50–58. https://doi.org/10.1016/j.ymben.2021.11.002.

(94) Mohamed, E. T.; Werner, A. Z.; Salvachúa, D.; Singer, C. A.; Szostkiewicz, K.; Rafael Jiménez-Díaz, M.; Eng, T.; Radi, M. S.; Simmons, B. A.; Mukhopadhyay, A.; Herrgård, M. J.; Singer, S. W.; Beckham, G. T.; Feist, A. M. Adaptive Laboratory Evolution of *Pseudomonas putida* KT2440 Improves P-Coumaric and Ferulic Acid Catabolism and Tolerance. *Metab Eng Commun* **2020**, *11*, e00143. https://doi.org/10.1016/j.mec.2020.e00143.

(95) Choe, D.; Lee, J. H.; Yoo, M.; Hwang, S.; Sung, B. H.; Cho, S.; Palsson, B.; Kim, S. C.; Cho, B.-K. Adaptive Laboratory Evolution of a Genome-Reduced Escherichia Coli. *Nat. Commun.* **2019**, *10* (1), 935. https://doi.org/10.1038/s41467-019-08888-6.

(96) Choudhury, D.; Saini, S. Evolution of Escherichia Coli in Different Carbon Environments for 2,000 Generations. *J. Evol. Biol.* **2019**, *32* (12), 1331–1341. https://doi.org/10.1111/jeb.13517.

(97) Balderas-Hernández, V. E.; Hernández-Montalvo, V.; Bolívar, F.; Gosset, G.; Martínez, A. Adaptive Evolution of *Escherichia coli* Inactivated in the Phosphotransferase System Operon Improves Co-Utilization of Xylose and Glucose Under Anaerobic Conditions. *Appl. Biochem. Biotechnol.* **2011**, *163* (4), 485–496. https://doi.org/10.1007/s12010-010-9056-3.

(98) Espinosa, M. I.; Gonzalez-Garcia, R. A.; Valgepea, K.; Plan, M. R.; Scott, C.; Pretorius, I. S.; Marcellin, E.; Paulsen, I. T.; Williams, T. C. Adaptive Laboratory Evolution of Native Methanol Assimilation in *Saccharomyces cerevisiae*. *Nat. Commun.* **2020**, *11* (1), 5564. https://doi.org/10.1038/s41467-020-19390-9.

CHAPTER TWO: Short-term adaptive laboratory evolution of *E. coli* to growth on acetylenedicarboxylate

2.1 Introduction

Microorganisms have long been recognized as microbial cell factories in biotechnological applications, providing platforms for producing various chemicals ranging from pharmaceuticals to commodity chemicals and biofuels.[1–3] Notable among these platform organisms are *E. coli*, Pseudomonads, and yeasts, which can utilize a broad range of organic compounds as carbon and energy sources.[4–6] These organisms, however, exhibit a preferential consumption of sugars and often require the activation of other specific natural pathways to exploit non-carbohydrate carbon sources.[7–10]

Adaptive laboratory evolution (ALE) approaches have been used to activate and optimize strains to utilize unnatural carbon sources.[11] Conventionally, the feedstock for microbial production of chemicals is predominantly sugars and plant-derived biomass, a renewable but challenging resource due to the need for the deconstruction of complex components, such as cellulose, lignin, and carbohydrates, to monomers through time and energy-consuming processes like pyrolysis[12,13] and liquefaction.[14]

In response to these challenges, there is a growing interest in exploring alternative feedstocks for chemical production, emphasizing the development of sustainable and net-zero carbon balance or decarbonization strategies. Global production of methane (CH_4) and carbon dioxide (CO_2) and their environmental implications offer potential use as sustainable feedstocks. Direct microbial utilization of CH_4 and CO_2 could usher in a more sustainable and green approach to commercial chemical production.[15] Direct microbial utilization of CH_4 and CO_2, however, poses multiple challenges. Microbes metabolizing methane and carbon dioxide require specific conditions and pathways for optimal growth and metabolic activity.[16–18] Methanotrophs typically exhibit slow growth rates, which limit the overall productivity of the microbial conversion process. Reports have shown that more tractable organisms like *E. coli* have been engineered for growth exclusively on CH_4 or CO_2.[15,19] Current research primarily focuses on understanding how microbes can utilize these gases and the potential biotechnological application in the global carbon cycle. Despite the challenges, microbial utilization of CH_4 and CO_2 presents a promising approach for a sustainable and decarbonized chemical industry. C4-Dicarboxylates, including fumarate, succinate, malate, and aspartate, are additional carbon and energy sources for microbial growth.[20–23] Their structural similarity to citric acid cycle-related metabolites positions them as prime candidates for microbial assimilation. This is mainly due to the presence of transport systems as well as enzymes that incorporate these compounds into metabolic pathways to produce energy and cell biomass.

However, while most C4-dicarboxylates are primarily converted to fumarate, which is used as an electron acceptor in the fumarate respiratory pathway to produce succinate,[20] the latter cannot be metabolized efficiently in the presence of other C4-dicarboxylates and is thus excreted. Hence, microbial growth tends to be limited in the presence of C4-dicarboxylates as the sole carbon source due to accumulation of succinate. This suggests that the metabolism of C4-dicarboxylates is not straightforward and may vary depending on the specific dicarboxylate and the metabolic context. This metabolic peculiarity, however, does not preclude the potential of C4-dicarboxylates as substrates for microbial growth and chemical production. One such substrate is acetylenedicarboxylate, which can be produced from methane and carbon dioxide in a two-step process. The two steps involve plasma-mediated methane dehydrodimerization to generate acetylene,[24,25] followed by acetylene carboxylation to produce acetylenecarboxylate and acetylenedicarboxylate.[26] Acetylene can also be produced through partial burning of CH_4 (patented by BASF).

This chapter presents acetylenedicarboxylate (ADCA), a compound not typically found in nature but could be synthesized from CH_4 and CO_2, as an alternate feedstock. Despite previous identification of Pseudomonads soil isolates capable of growing on ADCA, the identified strain was not archived, and ADCA's potential as a feedstock remains unexplored.[27,28] This investigation provides insight into improved growth of *E. coli* on ADCA after a Short-Term Adaptive Laboratory Evolution (STALE). We employ a multi-omics approach to elucidate the adaptation mechanism to the acetylenic carbon source, exploiting comparative genomics, transcriptomics, and the analysis of excreted metabolites. Our findings demonstrate the capacity of *E. coli* to assimilate and utilize acetylenedicarboxylate as the sole carbon source and identify putative genes responsible for its utilization. We further demonstrate the bioconversion of ADCA to shikimic acid and lactic acid, offering a promising proof of concept for this novel feedstock as a sustainable alternative to the use of biomass for microbial synthesis of chemicals.

2.2 Growth of *E. coli* MG1655 and RB791 on glucose, fumarate, and ADCA

The identification of feedstocks for the bioproduction of chemicals, particularly microbial fermentation, is hindered by the capacity of unnatural carbon sources to support the growth of microbes or the absence of enzymes to utilize these carbon sources as substrate. To overcome this challenge, synthetic biologists and metabolic engineers have developed strategies to either evolve enzymes and microbes or engineer strains to utilize wide range of substrates.

Table 2.1. Growth of *E. coli* on glucose (M9G), fumarate (M9F), and acetylenedicarboxylate (M9A) solid agar.

Plate	Number of colonies [dilutions]			
	MG1655		RB791	
	[10^{-5}]	[10^{-6}]	[10^{-5}]	[10^{-6}]
M9 salts	0	0	0	0
M9/ 20 mM Glucose (M9G)	3000	1000	1800	500
M9/ 40 mM Fumarate (M9F)	2200	800	500	200
M9/ 40 mM ADCA (M9A)*	2000	800	300	150

Note: For M9G and M9F, colonies were counted at 24 h after plating. For M9A, colonies were not apparent after 24 h and were therefore counted 48 h after plating (indicated by *).

These strategies usually require tractable laboratory strains like *E. coli*, Pseudomonads, or *S. cerevisiae*, amongst others. While organisms that are currently in use utilize carbohydrate substrates like glucose and xylose, some microbes can utilize carbon sources such as methane, methanol, formate, fumarate, carbon dioxide, and other carbon sources not occurring in nature. The metabolic pathways of *E. coli* are among the most thoroughly examined which makes this organism a good choice for metabolic engineering. To investigate the ability of *E. coli* to utilize ADCA as a sole source of carbon and energy, a seed culture of *E. coli* MG1655 was first grown in minimal media containing glucose (M9G) to log phase (OD_{600} = 0.6). Following two rounds of washing the cells in minimal media to remove glucose, dilutions (1:10^5 and 1:10^6) were plated onto minimal media agar containing either glucose (M9G), fumarate (M9F), or ADCA (M9A).

Fumarate was selected as a C4-dicarboxylate control due to its structural similarity to ADCA and previous reports of the ability of *E. coli* to metabolize fumarate. We observed that about 3000 and 2200 colonies appeared on M9G and M9F agar plates, respectively, after 24 h of incubation for 1:10^5, and about 2000 colonies appeared on ADCA media after 48 h of incubation (Table 2.1).

A similar pattern was observed for 1:10^6 dilution with colony numbers between 1000 and 800. *E. coli* RB791 shared a similar phenotype albeit significantly lower colony numbers.

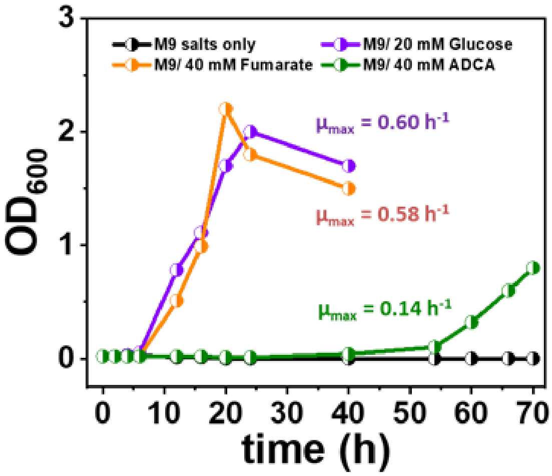


Figure 2.1. Growth of *E. coli* in liquid M9 media containing no carbon source (black), glucose (violet), fumarate (orange), or acetylenedicarboxylate (green).

A similar growth phenotype was observed when OD_{600} and growth rates were monitored in liquid minimal media containing the same carbon sources: glucose (M9G), fumarate (M9F), and ADCA (M9A). A prolonged lag phase (approximately 50 h) and a significantly lower maximum growth rate (0.14 h^{-1}) (Figure 2.1) were observed in M9A, in contrast to MG1655 in M9G and M9F with similar lag phase time and growth rate. This extended lag phase and lower maximum growth rate in the M9A medium highlight the metabolic challenge ADCA presents to *E. coli*. To test if suboptimal growth in M9A was due to ADCA concentration or metabolic limits, MG1655 was cultured in varying concentrations of glucose, fumarate, and ADCA (5, 20, 50, and 100 mM) (Figure 2.2). Cells were also cultivated in media containing 50 mM of two carbon sources including glucose, fumarate, and ADCA. Cell growth was monitored using a microplate reader (see Experimental section).

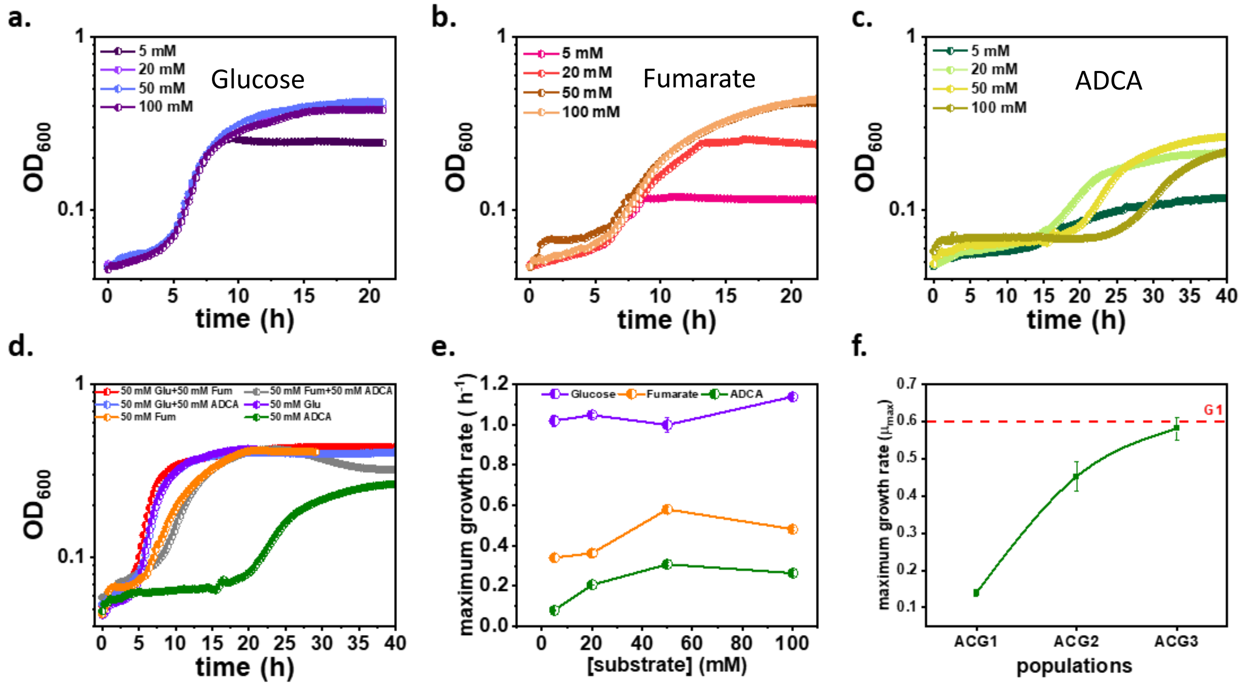


Figure 2.2. Growth profiles of *E. coli* in varying concentrations of (a) glucose, (b) fumarate, and (c) ADCA. (d) Growth profiles of *E. coli* in media containing multiple carbon sources. (e) Effect of substrate concentration on growth rate. Cell growth was monitored using a microplate reader with orbital shaking at 37 °C. (f) Improved growth rate across ACG series obtained during ALE in 50 mL shake flask experiments. The plots are averages of three replicates; error bars for have been removed to facilitate visualization of data.

Although growth rates were slightly lower when grown on 5 mM ADCA growth rates appeared approximately equal on 20, 50, and 100 mM ADCA(Figure 2.2 and Figure 2.3). The duration of lag phase varied significantly with ADCA concentrations, with higher concentrations showing longer lag phases. We observed that cultures could be grouped into three clusters based on carbon source rather than substrate concentrations. Cultures grown in mixed feedstocks (M9GA, M9GF, and M9FA) were most similar to cells growing on the preferred carbon source, demonstrating growth phenotype comparable to those observed in 50 mM single-substrate media of the preferred substrate. These observations suggest that ADCA is nontoxic to *E. coli* at lower concentrations but longest lag phase was observed for 100 mM ADCA which may indicate toxicity. Interestingly, whereas varying substrate concentrations do not affect lag phase time and growth rate with glucose, growth rates generally improve with increasing fumarate concentrations. However, corresponding lag phase times remained unaffected.

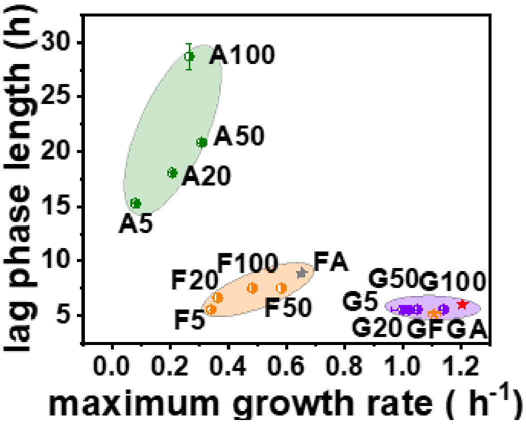


Figure 2.3. Bacterial elasticity plot highlighting growth on glucose (violet), fumarate (orange), or acetylenedicarboxylate (green). G = glucose F = fumarate, A = ADCA; the numbers represent concentrations in mM; GF = 50 mM glucose and 50 mM fumarate, GA = 50 mM glucose and 50 mM ADCA, FA = 50 mM fumarate and 50 mM ADCA.

2.3 Improving *E. coli* utilization of acetylenedicarboxylate via short-term adaptive laboratory evolution (ACG series).

We hypothesized that metabolic rewiring during the extended lag phase might cause adaptation for growth on ADCA. Hence, an adaptive laboratory evolution approach may help restore short lag phase time and improve the growth rate. The first round of cultures growing in ADCA was termed ADCA-Consuming Generations 1 (ACG1) (Figure 2.4 and 2.5). At the late log phase of ACG1, an aliquot of the culture was used to inoculate fresh M9A media and cultured as before to obtain ACG2. This strategy was repeated to afford ACG3. The fitness of the strains were measured as the ratio of the maximum growth rate of strains in ADCA to maximum growth rate of wild-type in glucose. ACG1 showed about 80% fitness loss compared to MG1655 in M9G (Figure 2.6). Significant reductions in lag phase times and increase in maximum growth rates across the ACG series (ACG1, ACG2, ACG3) were observed, with growth rate and fitness of ACG3 similar to

MG1655 cultured in glucose (0.60 h^{-1}) (Figure 2.6), indicating that the fitness of the strains was regained after ALE.

These findings were intriguing because ALE experiments typpically have long timescales, usually months. Achieving an ACG3 population with growth rates on ADCA similar to growth on glucose takes eight days. ACG2 attained a maximum growth rate of 0.45 h^{-1} with a 12 h lag-phase time in M9A. We noted that ACG1 performed better than MG1655 WT in fumarate in media. This rapid adaptation and change in fitness in the M9G medium after ALE raised questions about the underlying mechanisms enabling *E. coli* to metabolize ADCA efficiently. A multi-omics approach was used to elucidate these mechanisms, integrating genomic, transcriptomic, and metabolomic analyses. This comprehensive strategy aimed to identify specific genetic modifications, differential gene expression patterns, and metabolic pathway alterations that collectively contribute to the enhanced utilization of ADCA

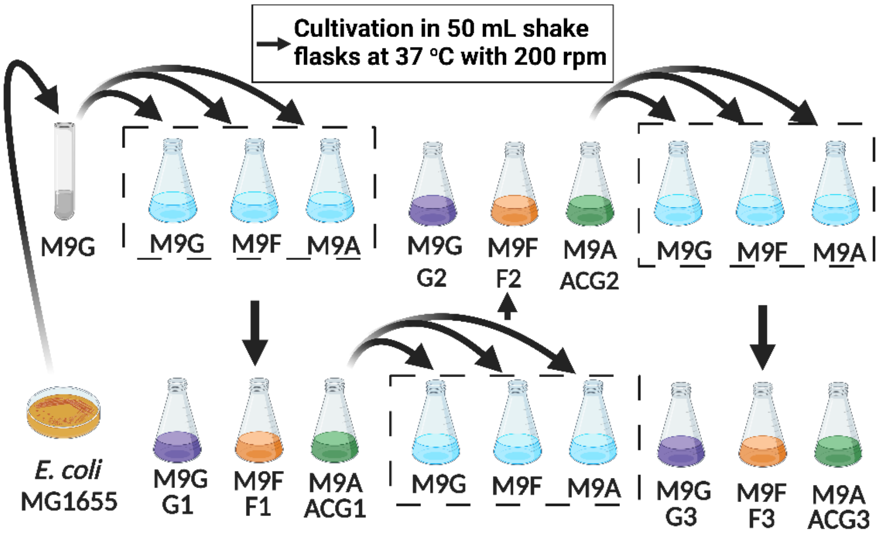


Figure 2.4. Experimental scheme depicting the adaptive laboratory evolution of *E. coli* MG1655. M9G contained 20 mM glucose; M9F contained 40 mM fumarate; M9A contained 40 mM ADCA. Curved arrows indicate inoculation into fresh media. Straight arrows indicate cultivation in corresponding medium in shake flasks at 37 °C with 200 rpm agitation. G1, F1, and ACG1 represents cultures from first round of cultivation in glucose, fumarate, and ADCA, respectively. G2, F2, and ACG2 represents cultures from second round of cultivation. G3, F3, and ACG3 represents cultures after three rounds of cultivation.

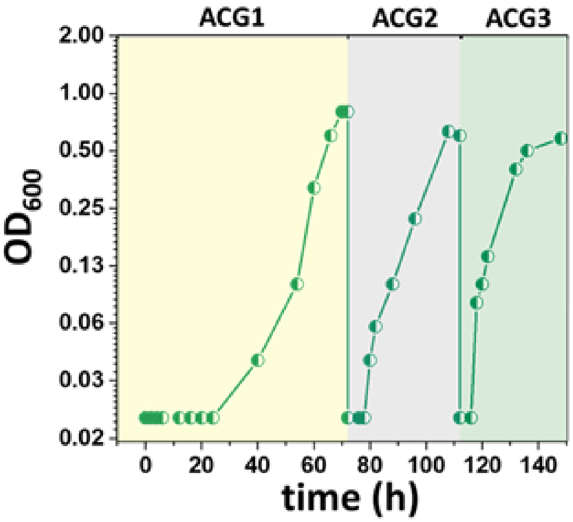


Figure 2.5. Growth profiles of evolved strains ACG1 (round 1), ACG2 (round 2), and ACG3 (round 3) show shortened lag phases and improved growth rates across the ACG series. ALE was performed in 50 mL shake flasks containing 40 mM ADCA.

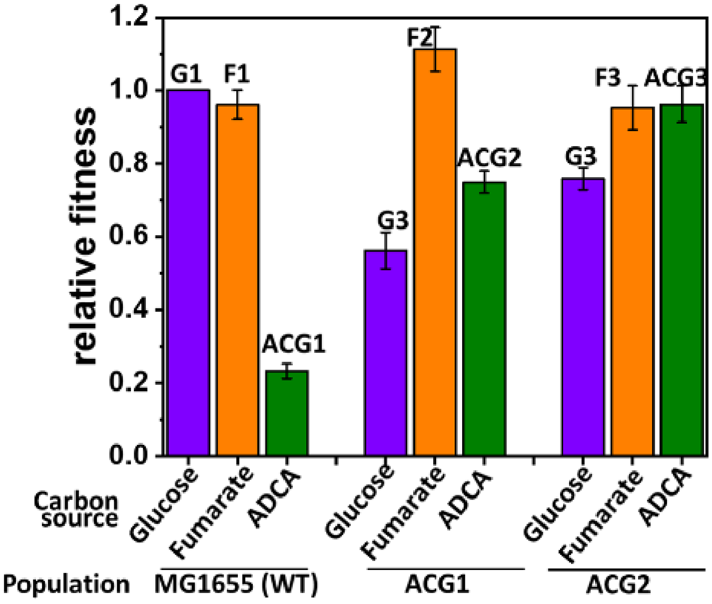


Figure 2.6. Comparative fitness assessment of strains in M9G, M9F, and M9A. Relative fitness was determined as the ratio of the maximum growth rate of a strain in a specific carbon source to the maximum growth rate of MG1655 in glucose. Error bars were calculated from three replicates.

2.4 Comparative genomics of ACG populations and clones

Whole genome sequence analysis is essential for identifying the adaptation mechanism after an ALE experiment. To elucidate the genetic basis for the improved growth of the evolved populations, three lineages of independently evolved populations of ACG1, ACG2, and ACG3 (Figure 2.7), as well as nine selected colonies from three additional lineages of ACG3 were sequenced along with the MG1655 ancestor strain cultured in glucose. These colonies were selected strategically based on their growth performance in ADCA (as shown in Figures 2.7 and 2.8).

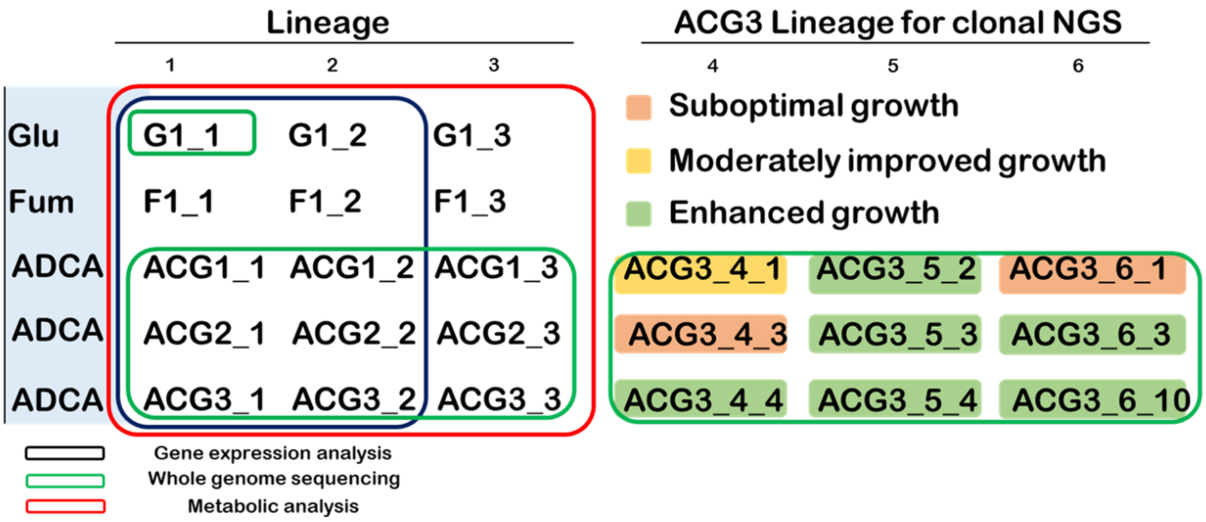


Figure 2.7. Venn Diagram indicating ALE populations for various multi-omics analyses (left) and clonal sequencing (right). For clonal samples, highlights indicate the growth phenotype in ADCA media, (red = suboptimal growth; yellow = moderately improved growth; and green = more enhanced growth phenotype).

This selection included a spectrum of growth phenotypes: ACG3_4_3 and ACG3_6_1, which exhibited suboptimal growth phenotypes; ACG3_4_1, which demonstrated moderately improved growth; and ACG3_4_4, ACG3_5_2, ACG3_5_3, ACG3_5_4, ACG3_6_3, and ACG3_6_10, all of which showed significantly enhanced growth on ADCA.

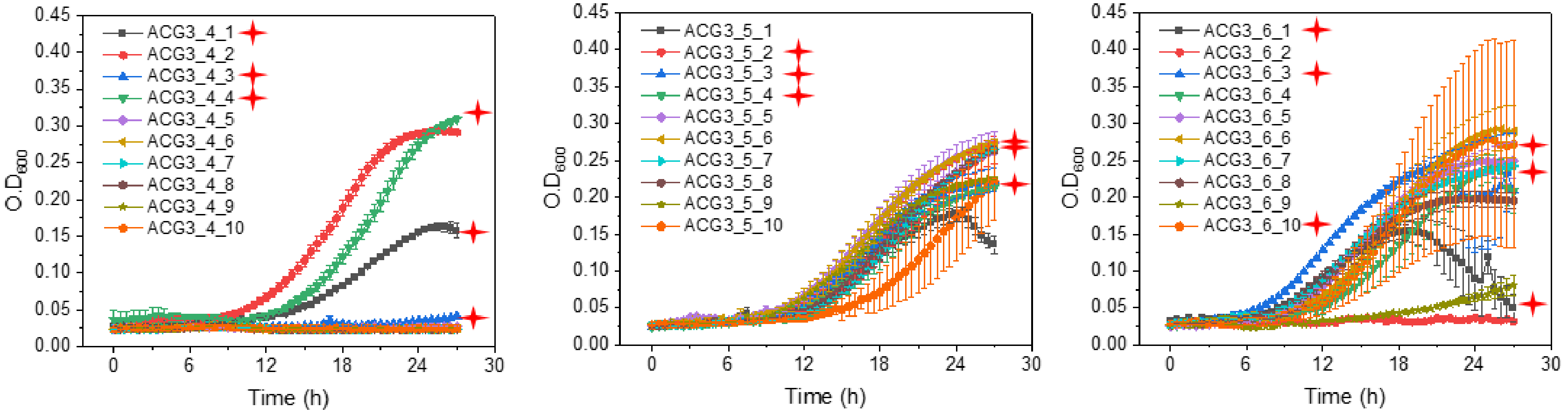


Figure 2.8. Growth profiles of ACG3 clones of lineages 4, 5, and 6 in M9A media. Growth was monitored using a plate reader. Colonies selected for whole genome sequencing analysis are indicated by a red cross.

The rationale for this selection approach was based on the hypothesis that colonies exhibiting the best growth would be more likely to harbor beneficial mutations conferring a growth advantage in ADCA, while mutations occurring in poorly growing strains are less likely to be beneficial.

By identifying and analyzing these mutations, we aimed to pinpoint the genetic modifications that contribute most significantly to the metabolic rewiring and adaptation processes. Sequencing of clonal samples were performed using Oxford Nanopore Technology R10.4.1 PromethION flow cell and population samples were sequenced using Illumina NovaSeq 6000 SP flow cell (Figure 2.9). Variant calling was performed using the breseq pipeline. The workflow and summary statistics for the WGS analysis are shown in Figure 2.9. The whole genome sequence analyses identified random mutations in all ACG1 lineages. ACG2 populations possessed more mutations with higher frequencies in all three lineages (Figure 2.10). The number of mutations further increased for ACG3 populations (Figure 2.11). Notably, a significant number of mutations in ACG2 and ACG3 were shared across all three lineages, suggesting a pattern of convergent evolution. Phylogenetic analysis of ACG populations revealed that all three lineages of ACG3 belong to a single clade (Figure 2.12), reinforcing the repeatability of the observed evolutionary adaptations and the convergence of evolutionary trajectories.

This evidence suggests that it is unlikely for the observed mutations being stochastic events, suggesting instead a selection strategy favoring the enrichment of populations with strains harboring mutations beneficial for enhanced growth on ADCA. This intricate pattern of genetic adaptation reflects the complex interplay of natural selection and evolutionary convergence in shaping the genomic landscape of microbes in response to the utilization of unnatural carbon sources. The analysis of population samples provided insightful observations into the genetic diversity within the cultures across the ACG series.

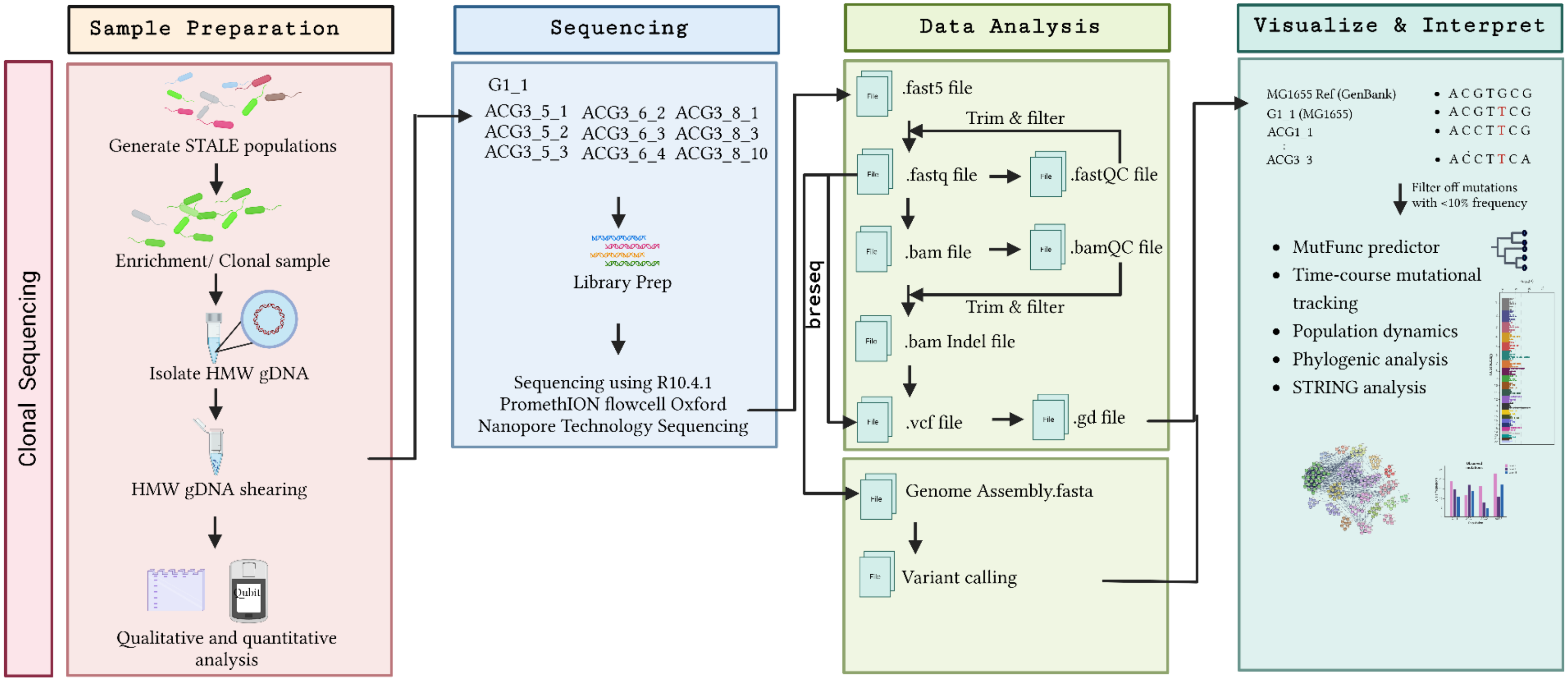


Figure 2.9. Schematic representation of whole genome sequencing analysis. The process is divided into four main stages: Sample Preparation, Sequencing, Data Analysis, and Visualization. Sample Preparation: Adaptive Laboratory Evolution (ALE) was conducted to derive acetylenedicarboxylate (ACG) adapted populations. The population samples were plated onto M9A plates and incubated as described in the experimental section. Ten colonies from each replicate (lineage) were selected and characterized to determine growth phenotype on ADCA. Genomic DNA of selected colonies were extracted and sheared to obtain fragments between 3 – 12 kbp using a Covaris® g-TUBE. Sequencing was by collaborators at Oak Ridge National Laboratory using an Oxford Nanopore Technology R10.4.1 PromethION flow cell. The breseq pipeline was used for variant calling and annotation. Visualization: Mutations identified were further explored and interpreted via phylogenetic analysis, MutFunc, STRING, among other tools. This approach allows for an in-depth understanding of the genomic variations that occur within ACG populations.

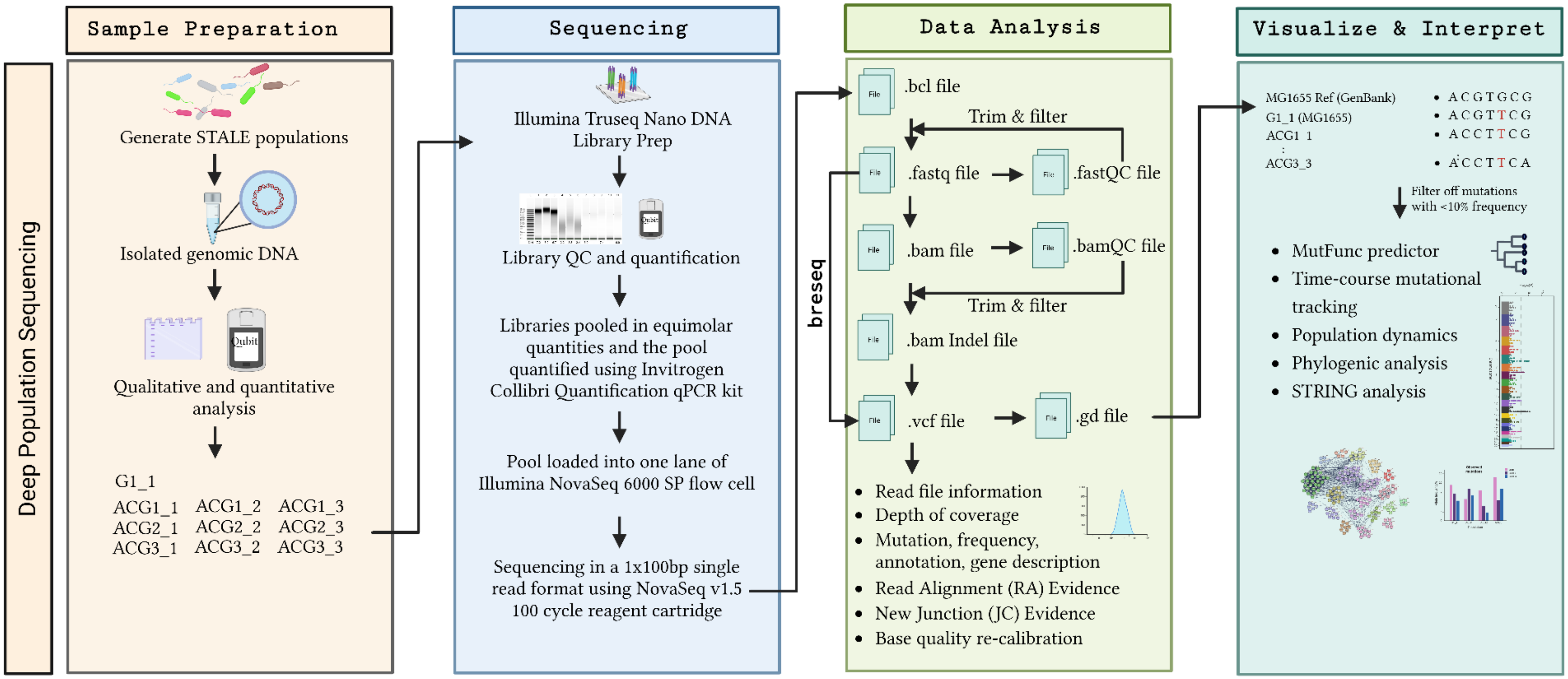


Figure 2.9. CONTD. Schematic representation of whole genome sequencing analysis. The process is divided into four main stages: Sample Preparation, Sequencing, Data Analysis, and Visualization. Sample Preparation: Adaptive Laboratory Evolution (ALE) was conducted to derive acetylenedicarboxylate (ACG) adapted populations, followed by DNA isolation. Sequencing: Library preparation was executed using an Illumina Truseq Nano DNA Library Prep. The resultant library pool was subjected to quality control and quantification. Sequencing was performed in a 1 x 100 bp single-read format utilizing a NovaSeq v1.5 100 cycle reagent cartridge. Data Analysis: The bcl file generated post-sequencing was converted to a fastq format. Read quality was assessed, and the breseq pipeline was employed for variant calling and annotation. Visualization: Mutations identified were further explored and interpreted via phylogenetic analysis, MutFunc, STRING, among other tools. This approach allows for an in-depth understanding of the genomic variations that occur within ACG populations.

These observed genetic variations indicate the coexistence of dual population dynamics: a subset of the population undergoes rapid mutation (mutators), potentially as an adaptive strategy to optimize growth on ADCA, while a subgroup appears to adopt an opportunistic approach.[11,29]

This claim is supported by the number of mutations observed in the clones, specifically ACG3_4_4 (mutator; 286 mutations) and ACG3_4_3 (opportunist; 7 mutations). Both clones shared the same ancestor (lineage 4), but while the mutator ACG3_4_4 showed enhanced growth, ACG3_4_1 and ACG3_4_3 (opportunists) showed moderate and suboptimal growths, respectively. These opportunistic cells may benefit indirectly from the adaptive mutations of the former group, a complex adaptive landscape within the population, where mutator phenotype facilitates community-wide adaptation to ADCA, underscoring the multifaceted nature of microbial evolution in response to the unnatural substrate. Another explanation could be that beneficial mutations arise in the mutator background and sweep through the population due to positive selection.

This phenomenon, called mutation-selection balance or mutator hitchhiking, was previously reported in ALE experiments.[29–31] Non-mutators can also coexist with mutators if the adaptive landscape allows for multiple fitness peaks or if the cost of the higher mutation rate in mutators balances out the benefit of the adaptive mutation, allowing non-mutators to remain competitive. For lineage 4 clones, we observed a positive correlation between the number of mutations and improvement in growth on ADCA, where ACG3_4_4 was a mutator with significantly improved growth, while ACG3_4_3 showed significantly fewer mutations and poor growth. However, all clones of lineage 5 showed very few mutations but significantly improved growth. We highlight 29 commonly mutated genes across multiple lineages from population and clonal sequencing, covering both mutators and non-mutators based on frequency of occurrence (Figure 2.13).

The relationship between all mutated genes occurring in at least two lineages was determined using the STRING database. This analysis allowed us to generate 20 clusters, each containing proteins implicated in similar pathways or cellular processes (Figure 2.14 and Figure 2.15).

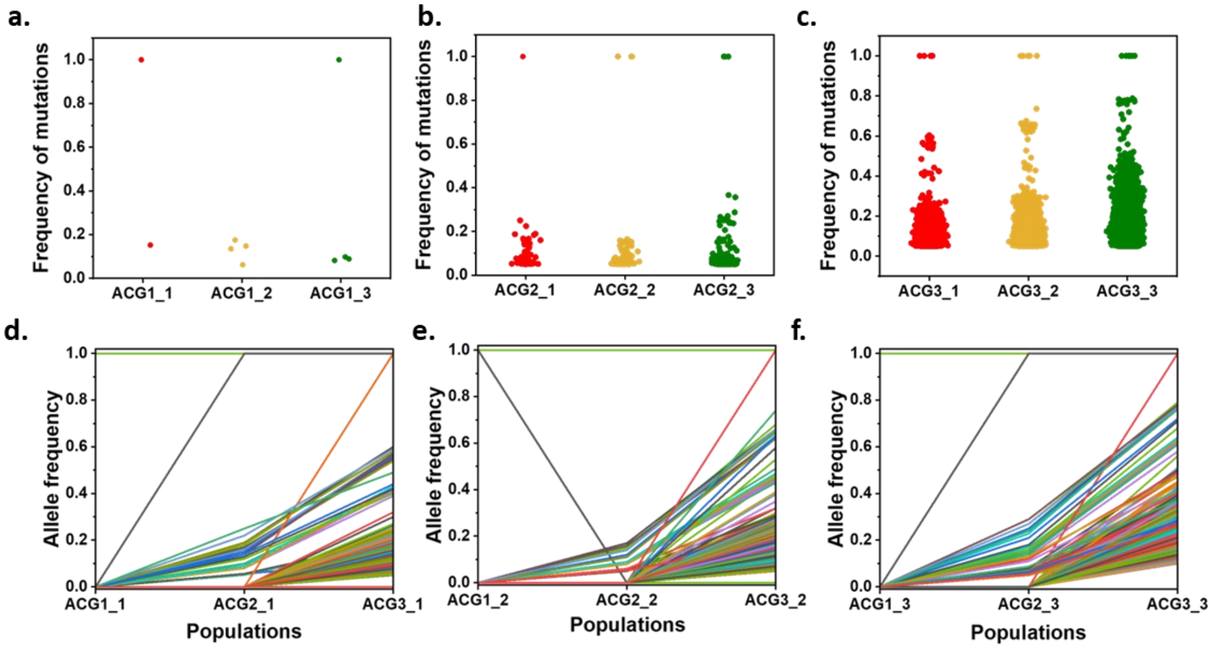


Figure 2.10. Population sequencing analysis of evolved ACG series. (a – c) Mutational frequencies of observed mutations in the 3 lineages of (a) ACG1, (b) ACG2, and (c) ACG3. Trajectory of mutations for lineages (d) 1, (e) 2, and (f) 3. Each line represents a single mutation. Figures d – f were obtained after filtering (see Methods).

The STRING network statistics analysis obtained a PPI enrichment p-value $< 10^{-16}$ which suggests that the identified clusters have more interactions among themselves than would be expected for a random set of proteins of the same sample size and degree distribution drawn from the genome.

We further estimated the effect of mutations using the mutfunc resource.[32] The impacts of these predicted mutations were identified as influencing protein stability, occurring at protein interaction interfaces, within a conserved region, or at transcription factor binding sites.

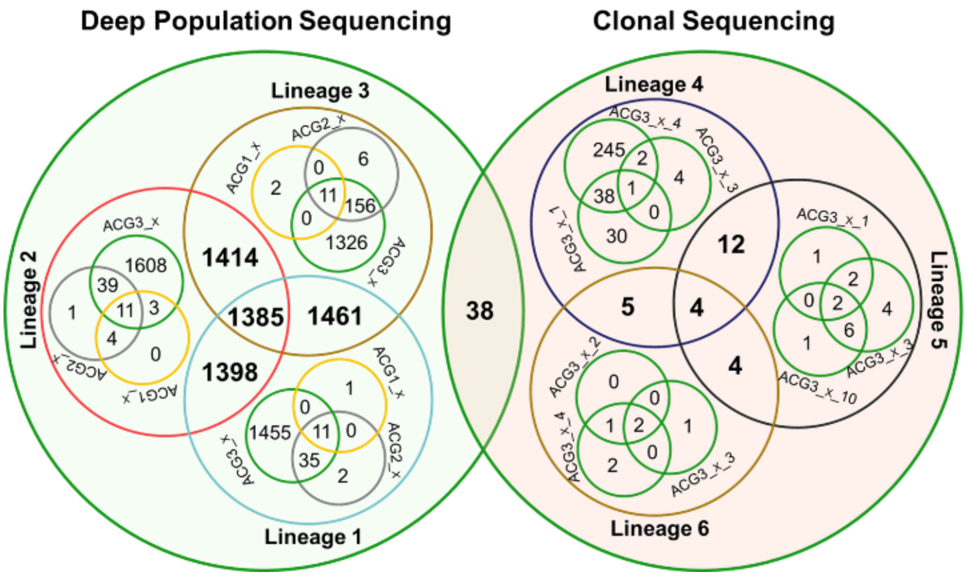


Figure 2.11. Venn diagram showing numbers of mutations observed for population and clonal samples. Population sequencing was performed using Illumina NovaSeq 6000 with variant calling done with the breseq pipeline in population mode. Clonal samples were sequenced using Long-read Oxford Nanopore Technology sequencer, and variant calling was done using Medaka.

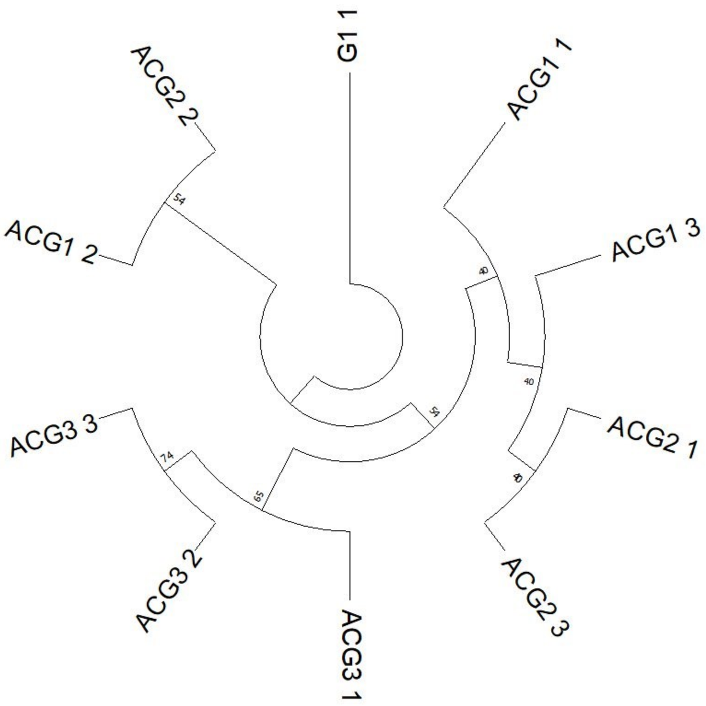
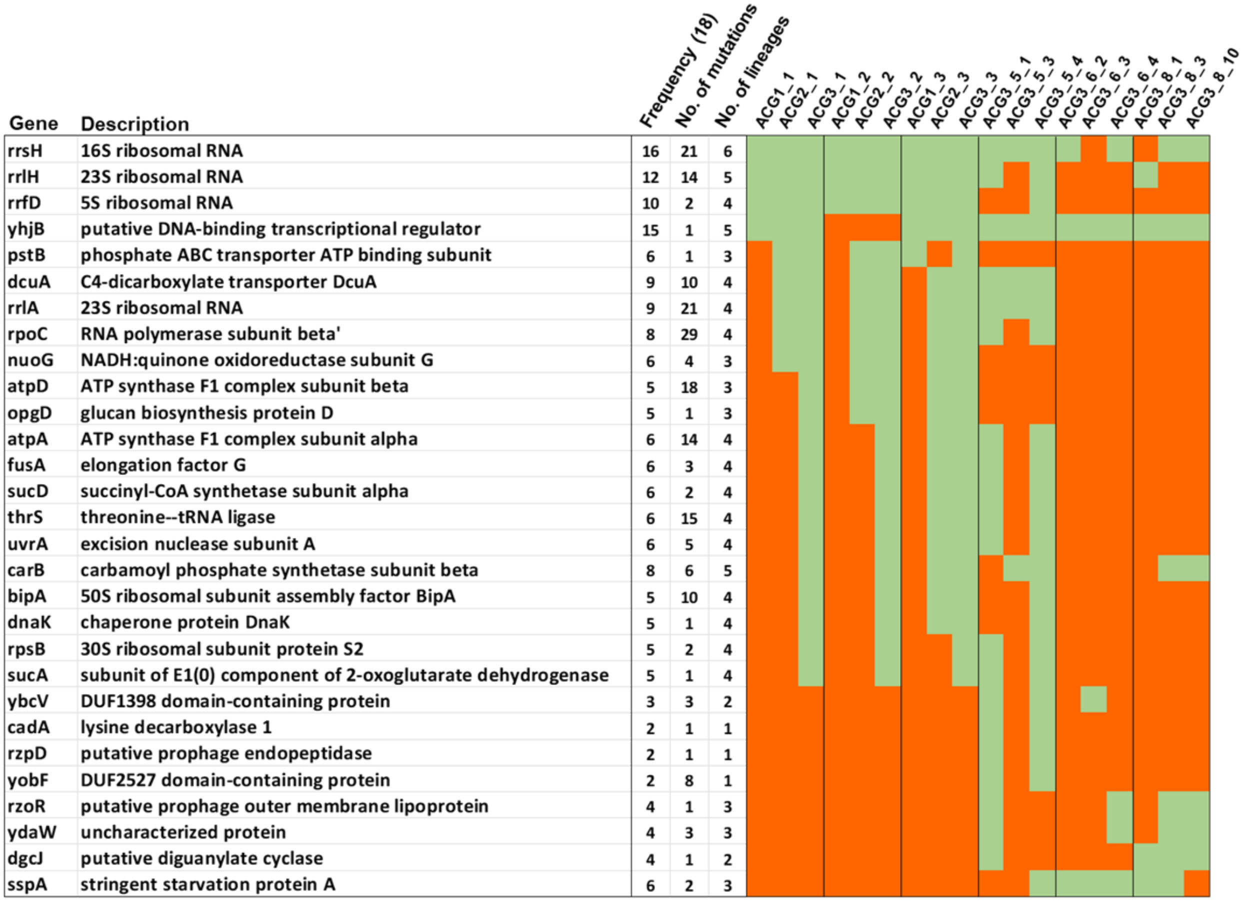


Figure 2.12. Phylogenetic analysis positioning of the ACG series in relation to the ancestor *E. coli* MG1655 cultivated in M9G.

The ability to predict the effects of genetic variations offers a time- and cost-effective alternative to experimental approaches and enables researchers to prioritize variants of interest and accelerate and elucidate the mechanisms driving these evolutionary changes. The impacts of these mutations are discussed in the sections that follow.



Gene	Description	Frequency (18)	No. of mutations	No. of lineages	ACG1_1	ACG2_1	ACG3_1	ACG1_2	ACG2_2	ACG3_2	ACG1_3	ACG2_3	ACG3_3	ACG3_5_1	ACG3_5_3	ACG3_5_4	ACG3_6_2	ACG3_6_3	ACG3_6_4	ACG3_8_1	ACG3_8_3	ACG3_8_10
rrsH	16S ribosomal RNA	16	21	6	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟩	🟥	🟩	🟩
rrlH	23S ribosomal RNA	12	14	5	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟩	🟥	🟥	🟥	🟩	🟥	🟥
rrfD	5S ribosomal RNA	10	2	4	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
yhjB	putative DNA-binding transcriptional regulator	15	1	5	🟩	🟩	🟩	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
pstB	phosphate ABC transporter ATP binding subunit	6	1	3	🟥	🟩	🟩	🟥	🟩	🟩	🟩	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
dcuA	C4-dicarboxylate transporter DcuA	9	10	4	🟥	🟩	🟩	🟥	🟩	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟥
rrlA	23S ribosomal RNA	9	21	4	🟥	🟩	🟩	🟥	🟩	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟥
rpoC	RNA polymerase subunit beta'	8	29	4	🟥	🟩	🟩	🟥	🟩	🟩	🟥	🟩	🟩	🟩	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
nuoG	NADH:quinone oxidoreductase subunit G	6	4	3	🟥	🟩	🟩	🟥	🟩	🟩	🟥	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
atpD	ATP synthase F1 complex subunit beta	5	18	3	🟥	🟥	🟩	🟥	🟩	🟩	🟥	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
opgD	glucan biosynthesis protein D	5	1	3	🟥	🟥	🟩	🟥	🟩	🟩	🟥	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
atpA	ATP synthase F1 complex subunit alpha	6	14	4	🟥	🟥	🟩	🟥	🟥	🟩	🟥	🟩	🟩	🟩	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
fusA	elongation factor G	6	3	4	🟥	🟥	🟩	🟥	🟥	🟩	🟥	🟩	🟩	🟩	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
sucD	succinyl-CoA synthetase subunit alpha	6	2	4	🟥	🟥	🟩	🟥	🟥	🟩	🟥	🟩	🟩	🟩	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
thrS	threonine--tRNA ligase	6	15	4	🟥	🟥	🟩	🟥	🟥	🟩	🟥	🟩	🟩	🟩	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
uvrA	excision nuclease subunit A	6	5	4	🟥	🟥	🟩	🟥	🟥	🟩	🟥	🟩	🟩	🟩	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
carB	carbamoyl phosphate synthetase subunit beta	8	6	5	🟥	🟥	🟩	🟥	🟥	🟩	🟥	🟩	🟩	🟥	🟩	🟩	🟥	🟥	🟥	🟥	🟩	🟩
bipA	50S ribosomal subunit assembly factor BipA	5	10	4	🟥	🟥	🟩	🟥	🟥	🟩	🟥	🟩	🟩	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
dnaK	chaperone protein DnaK	5	1	4	🟥	🟥	🟩	🟥	🟥	🟩	🟥	🟩	🟩	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
rpsB	30S ribosomal subunit protein S2	5	2	4	🟥	🟥	🟩	🟥	🟥	🟩	🟥	🟥	🟩	🟩	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
sucA	subunit of E1(0) component of 2-oxoglutarate dehydrogenase	5	1	4	🟥	🟥	🟩	🟥	🟥	🟩	🟥	🟥	🟩	🟩	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
ybcV	DUF1398 domain-containing protein	3	3	2	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟩	🟥	🟩	🟥	🟥	🟥	🟥
cadA	lysine decarboxylase 1	2	1	1	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
rzpD	putative prophage endopeptidase	2	1	1	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
yobF	DUF2527 domain-containing protein	2	8	1	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
rzoR	putative prophage outer membrane lipoprotein	4	1	3	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟩	🟥	🟩	🟩
ydaW	uncharacterized protein	4	3	3	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟩	🟥	🟩	🟩
dgcJ	putative diguanylate cyclase	4	1	2	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩
sspA	stringent starvation protein A	6	2	3	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟥

Figure 2.13. Mutations occurring across ACG populations and clones. Green and red indicates presence or absence of mutations in the sample, respectively.

2.5 Activation of the stringent response modulates adaptation.

In response to nutrient scarcity or adverse environments, *E. coli* initiates the Stringent Response (SR), an adaptation mechanism activated by the presence of high levels of uncharged tRNAs, which is signaled by the accumulation of guanosine tetraphosphate or pentaphosphate ((p)ppGpp) alarmones.[33–35] While the WGS analysis revealed no mutations in alarmone metabolism genes, mutations were observed in tRNA-related genes. These mutations might be triggering SR by increasing uncharged tRNA. Mutations in rRNA genes were observed early in evolution across all ACG populations (Figures 2.13 – 2.15) and were classified as Clade I and II mutations (Figures 2.16 and 2.17).

Some of these rRNA mutations were also observed in the Long-read Oxford Nanopore Technology (ONT) sequencing analyses of the clones. These mutations, especially in genes encoding 16S rRNA (*rrsC, rrsG, rrsH*) and 23S rRNA (*rrlA*), could disrupt ribosome function but correlate with improved growth on ADCA, possibly by reducing rRNA production costs. These alterations in the protein synthesis machinery suggest a complex adaptation strategy, where ribosomal stalling induced by mutations in tRNA-related genes may activate the SR, explaining the extended lag phases in ACG1 and ACG2.

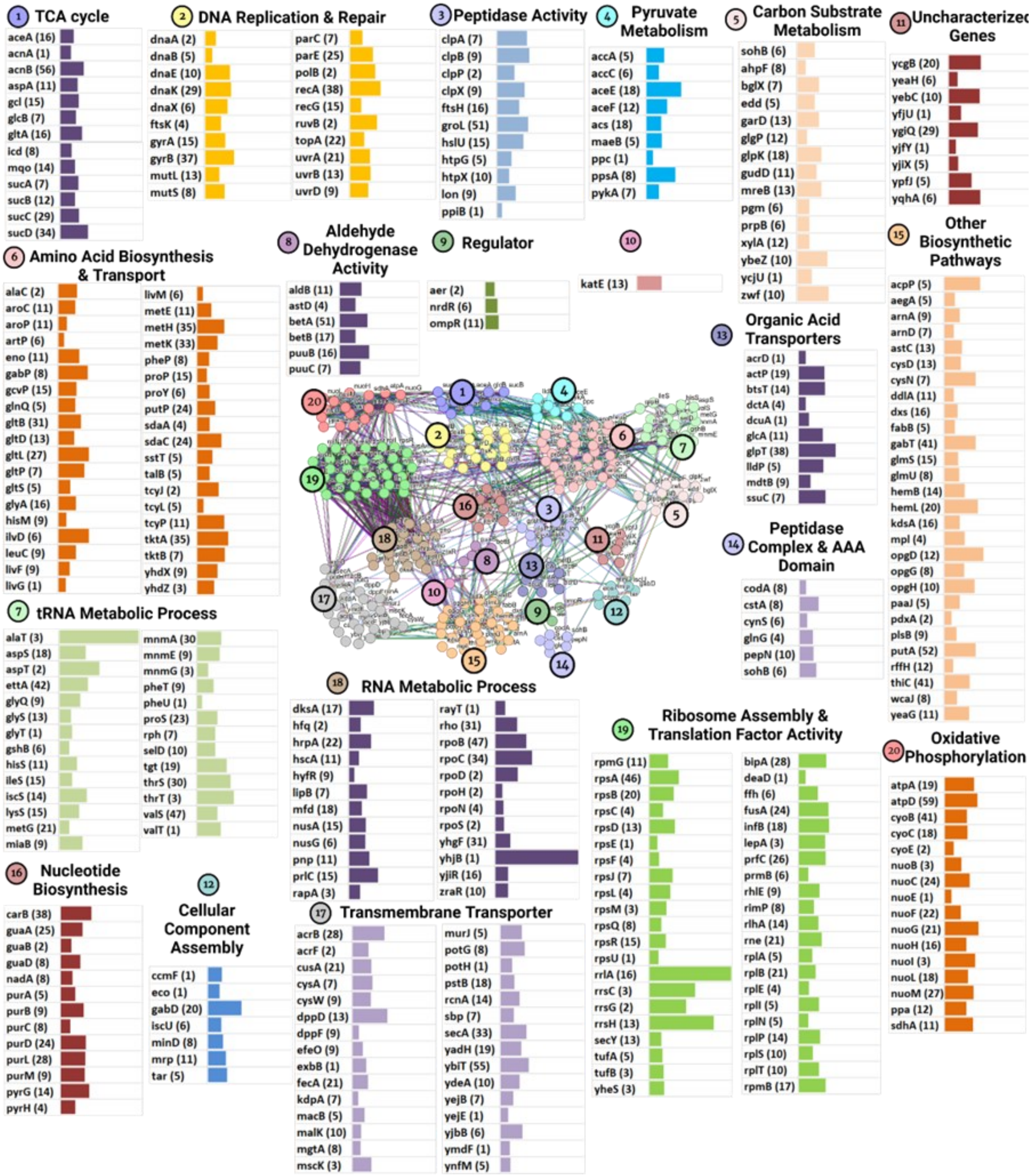


Figure 2.14. Network-based clustering of mutated genes using STRING analysis. Genes were clustered into 20 groups based on function. Figures in parenthesis for each gene represents the number of mutations for the specific gene. Bar charts represent the highest observed mutational frequencies. There are 319 nodes and 1058 edges with an average node degree of 6.63. The PPI enrichment p-value < 10^{-16}. These network statistics suggest that these proteins have more interactions among themselves than what is expected for a random set of proteins of the same size and degree distribution drawn from the genome.

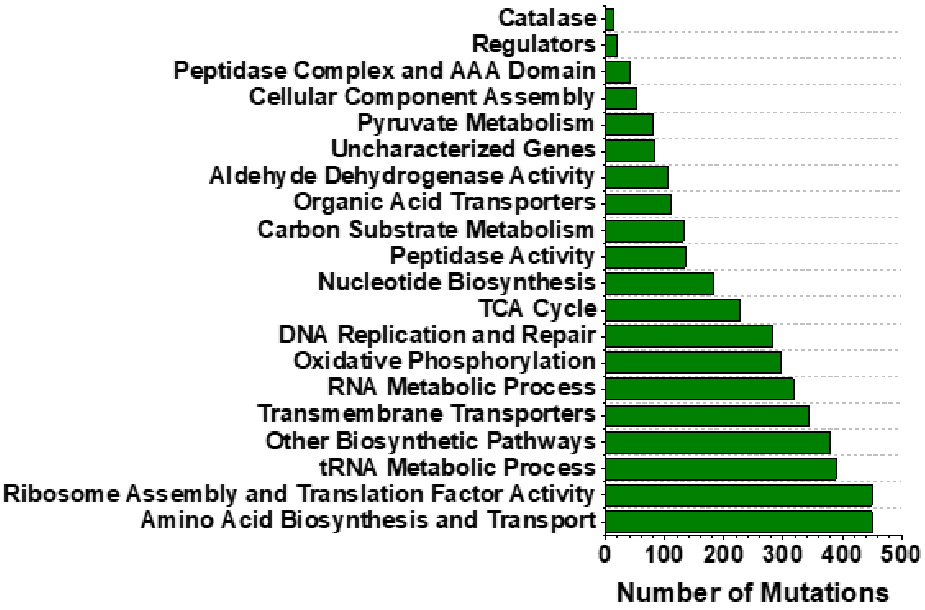


Figure 2.15. Distribution of mutations across various functional pathways and cellular processes.

2.6 Transcriptome analysis

Substrate transport into the cell is the crucial first step in utilizing carbon sources for cell biomass and energy generation. The prolonged lag phase during growth on ADCA can partly be explained by the necessity to synthesize specific enzymes required for the transport and metabolism of C4-diacids.[21] RNA sequencing (RNA-seq) analysis is a tool for transcriptomic profiling that allows the quantification of gene expression changes on a genome-wide scale, making it an indispensable tool for ALE studies. In this study, RNA-seq was used to investigate the transcriptional responses of each ACG population. This analysis provided insight into the mechanism by which *E. coli* optimizes the metabolism of ADCA. For this analysis, total RNA was isolated from MG1655 (wt) and evolved ACG3 populations (ACG3_1, ACG3_2, and ACG3_3) (triplicates) and submitted to SeqCoast Genomics (Oakhurst Biosciences, Inc.) for RNA-seq and differential expression analysis using Deseq2. The output of the Deseq2 analysis was used to generate a volcano plot (Figure 2.18) highlighting differentially expressed genes. We observed significantly more up-regulated genes than down-regulated genes. This pattern suggests a robust activation of different pathways as part of the adaptive response. The differentially expressed genes were also examined to determine which metabolic pathways are affected by the adaptation process. This will be useful for future studies as these genes could potentially be manipulated to achieve improved growth on ADCA without the need for ALE. We observed that the amino acid metabolism pathway showed the highest number (51) of up-regulated genes (Figure 2.19). Other pathways with high upregulation include transporters (50), bacterial motility proteins (28), oxidative phosphorylation (26), citric acid (TCA) cycle and carbon fixation pathways (22), ribosome biogenesis (22), flagellar assembly (20), glyoxylate and dicarboxylate metabolism (11), and biofilm formation (11). The highest downregulation was observed in genes of unknown function (13) and amino acid metabolism (9) (Figure 2.19).

We examined and verified the expression levels of 22 target genes previously implicated in the metabolism of dicarboxylates by RT-qPCR.[20,21,36] These genes are involved in four key pathways; non-specific transporters (*ompC, ompF, phoE*); C4-dicarboxylate transporters (*dauA, dctA, dcuA, dcuB, dcuC, dcuD, dcuR, dcuS*); TCA cycle-related genes (*fumA, fumB, fumC, frdA, frdB, frdC, frdD, gltA, mdh, prpC*); and pyruvate metabolism-related genes (*eda, maeA, maeB, pck*).

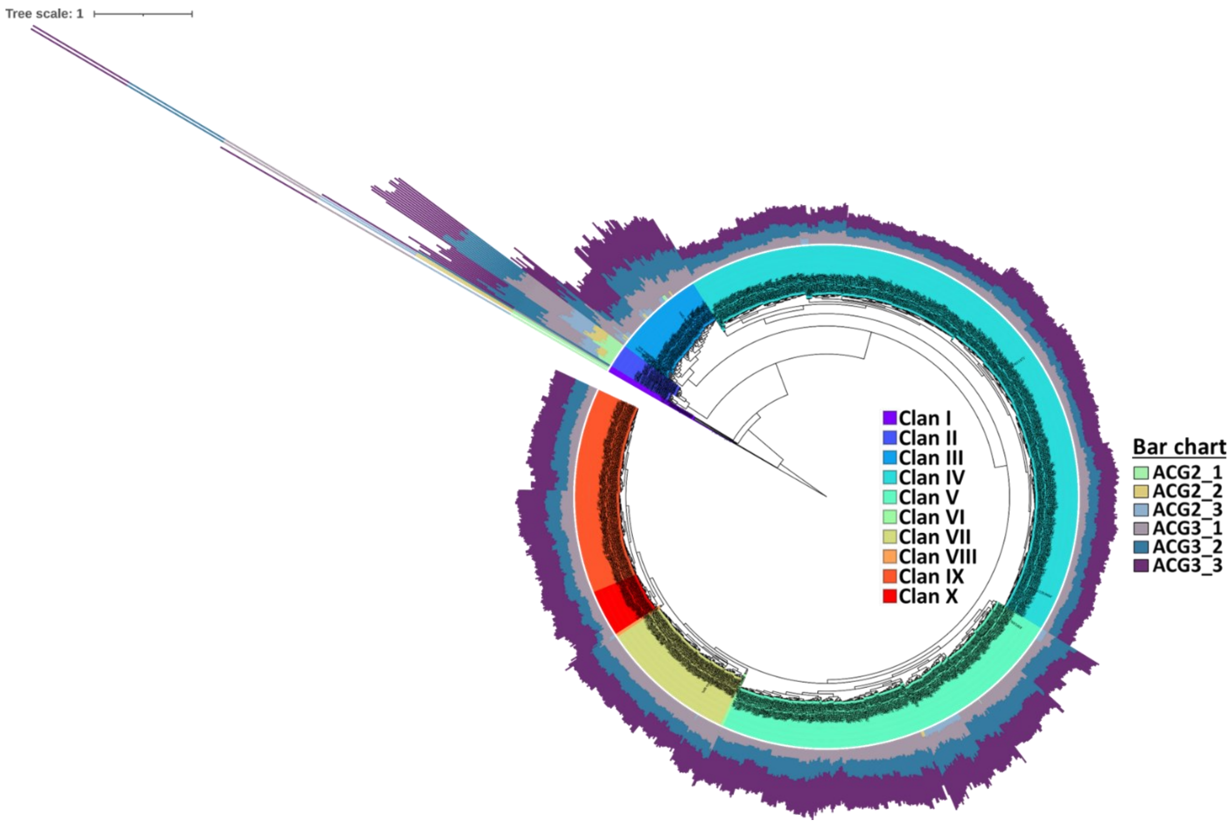


Figure 2.16. The dendrogram categorizes observed mutations into ten distinct clans based on frequencies, with the circumferential bar plot denoting individual mutation frequencies. Mutations with higher frequencies are grouped into earlier clans to indicate they may have occurred earlier during the evolutionary process. For instance, mutations observed in clan II.

We analyzed the transcriptomes of populations from G1, F1, and ACG series (ACG1, ACG2, and ACG3) to determine differentially expressed target genes facilitating growth on ADCA. Expression levels of target genes in strains cultured in M9F (F1) and M9A (ACG1, ACG2, ACG3) media were determined relative to strains in glucose media (G1), using *gyrA* as the housekeeping gene. Among the investigated non-specific outer membrane proteins, *phoE* and *ompC* expression levels remained unchanged across all carbon sources, whereas *ompF* was up-regulated about 32-fold for F1 and ACG series (Figure 2.20), suggesting OmpF plays a crucial role in facilitating the entry of ADCA and fumarate entry into the periplasm (Figure 2.22). The upregulation of *ompF* was confirmed by the RNA-seq analysis. The reciprocal regulation of *ompC* and *ompF* by the

EnvZ/OmpR two-component signal transduction system, as observed in *E. coli,* is a complex mechanism that adjusts membrane permeability in response to environmental cues.[37,38] The phosphorylation of EnvZ, induced by the detection of C4-dicarboxylates, leads to the activation of OmpR through phosphorylation, resulting in the differential expression of *ompC* and *ompF* (Figure 2.21 and 2.22). This regulatory mechanism is consistent with our observations of distinct expression patterns for these transporters, where *ompF* is preferentially overexpressed. Interestingly, we observed several mutations, mainly in the DNA-binding domain of the regulatory gene *ompR*, specifically Y194W, S195D, M197L, V212M, E214P, A217S, and H218K (Figure 2.21). These mutations may alter OmpR's ability to bind DNA and regulate the expression of *ompF* and *ompC*. Previous studies have examined the impact of various mutations within the OmpR DNA-binding domain on the expression of *ompC* and *ompF*.[39,40] These studies reported a phenotypic correlation to OmpR residues, allowing for the construction of a detailed functional map of the protein. Their study revealed that mutations within the same region can exert opposite effects on regulating *ompC* and *ompF*.

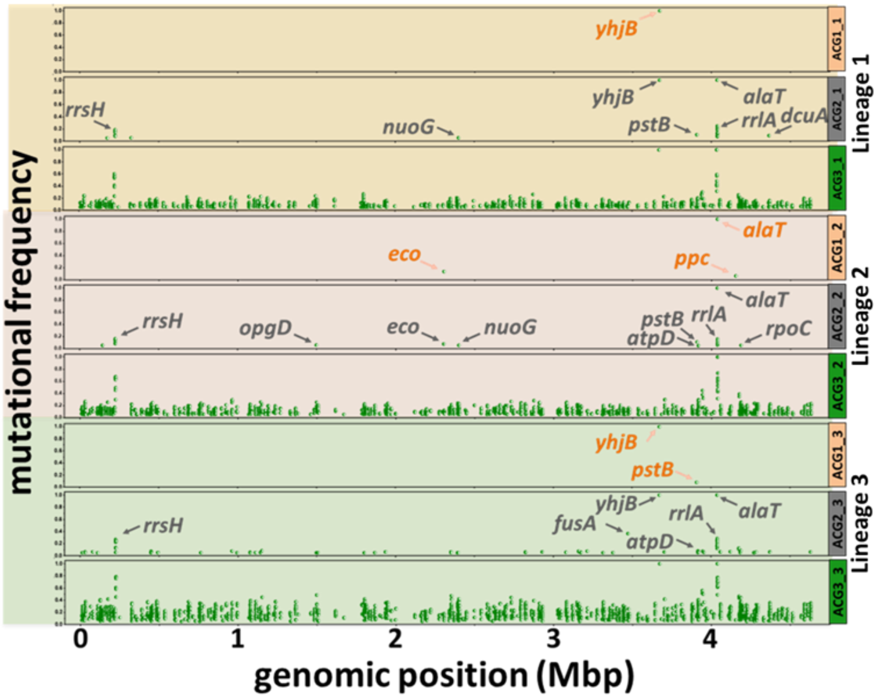


Figure 2.17. Mutation trajectory across three distinct lineages showing common mutations across ACG populations.

OmpF, characterized by a larger pore and faster flow rate than OmpC, is regulated by four OmpR binding sites, while OmpC is regulated by three.[40] Despite the high sequence homology shared by these 20 bp binding sites, computational analysis of the OmpR variants' binding affinities to the promoter regions of *ompC* and *ompF* revealed significant disparities (Figure 2.23). Binding affinity was determined using the HDOCK server, a platform for protein-protein and protein-DNA/RNA docking based on a hybrid algorithm of template-based modeling and *ab initio* free docking.[41,42] The 20 bp binding sites were constructed using Discovery Studio 2021. We observed that wild-type OmpR exhibited the highest binding affinity across all OmpC promoter sites, with a marked preference for the site furthest from the gene.

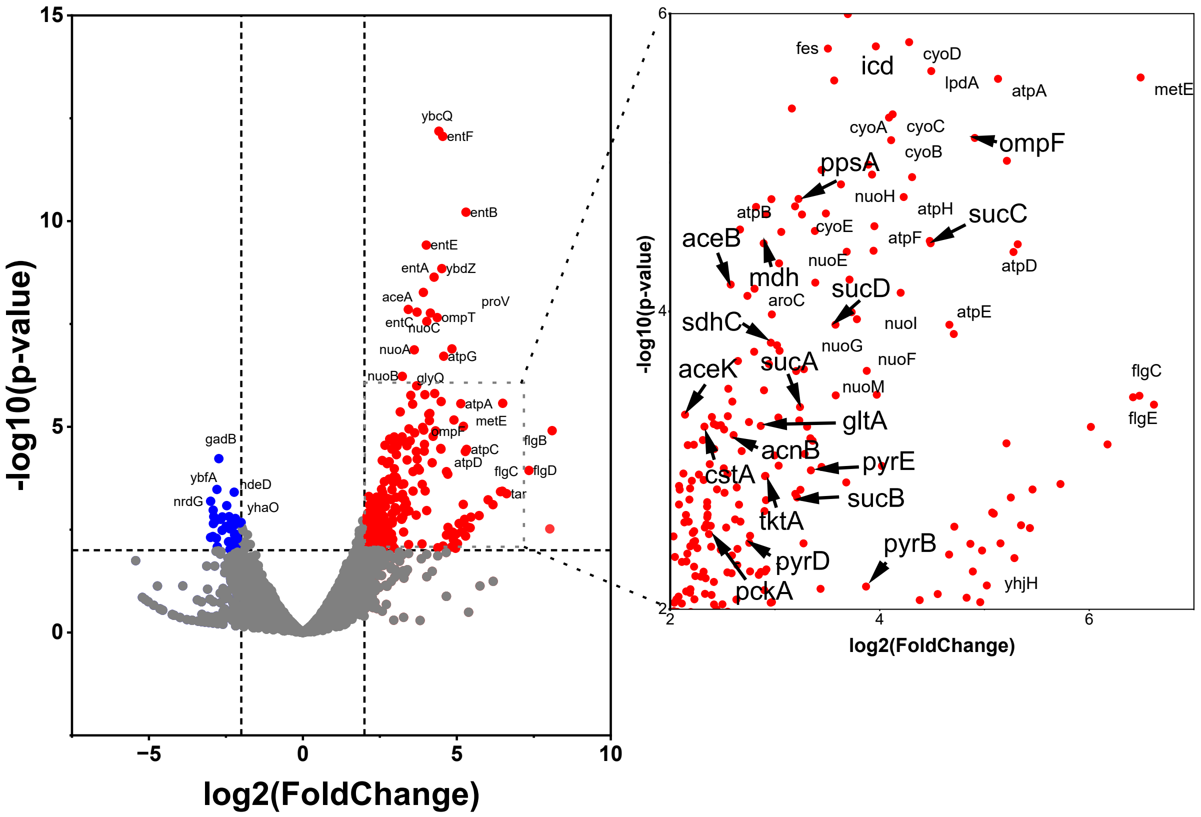


Figure 2.18. Volcano plot for RNA-seq analysis of ACG3 populations compared to G1. Log2 fold changes is shown on x-axis and the statiscal significance (-log10 p-value) on y-axis. The dotted horizontal line indicates significance threshold value p-value < 0.01. The dotted vertical lines indicate mininal fold change threshold for differentially expressed genes ($\log_2$ (foldchange) = ±2). Data were obtained from analyzing triplicates.

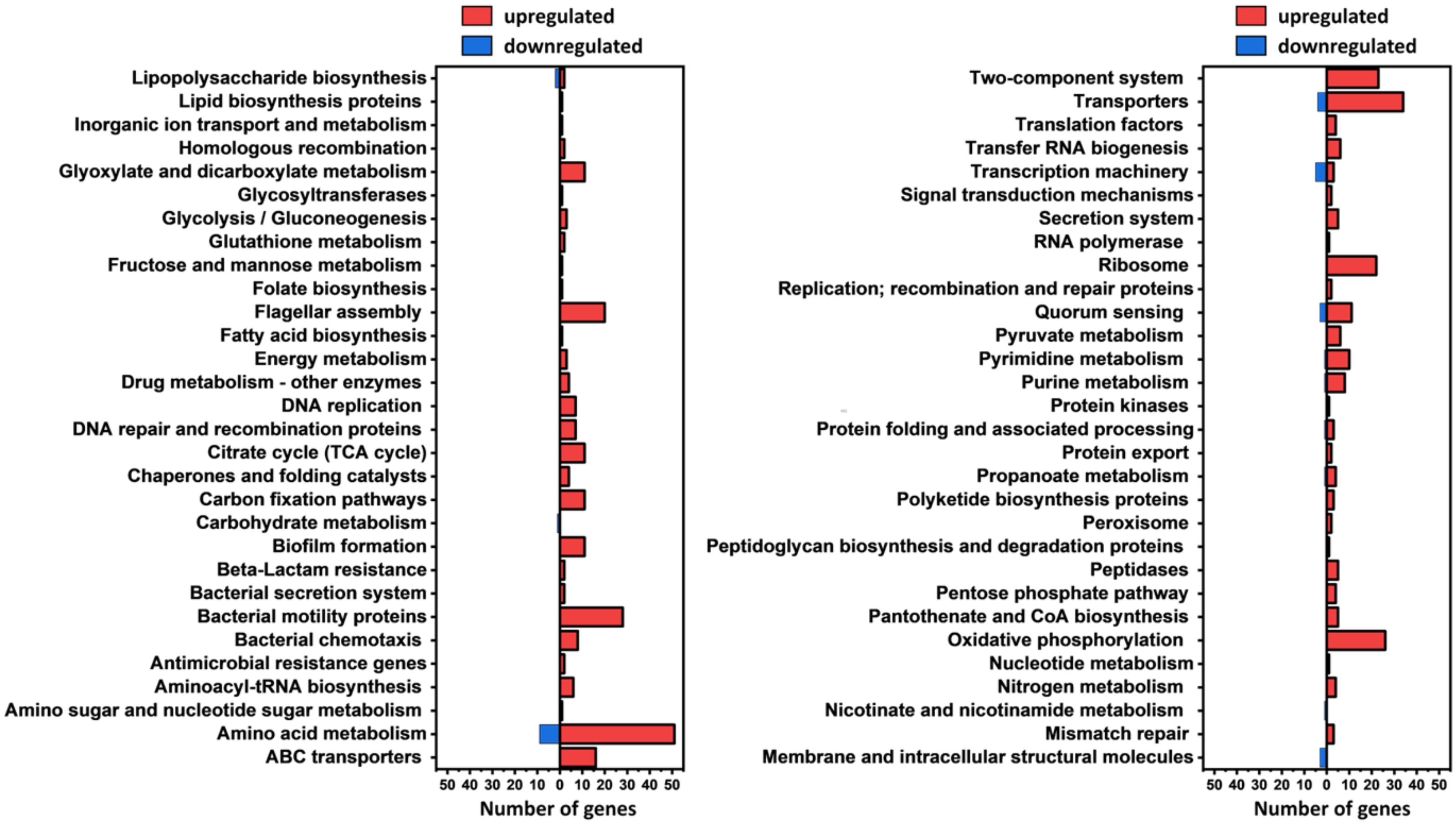


Figure 2.19. KEGG pathway analysis of differentially expressed genes (n = 482 genes)

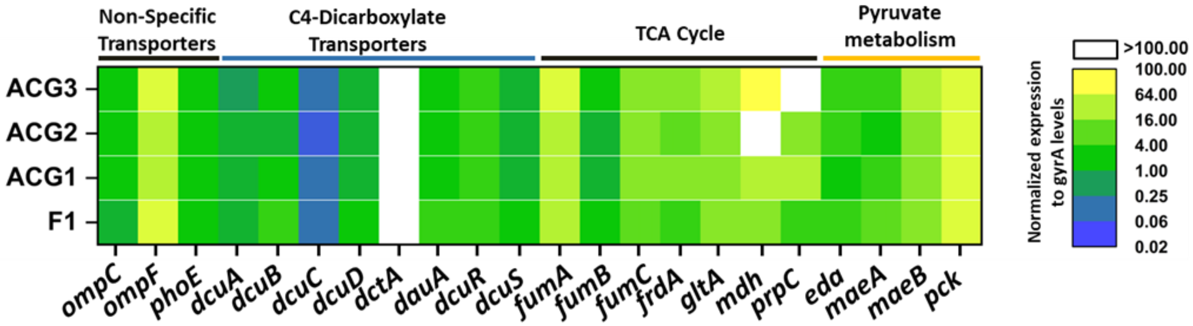


Figure 2.20. RT-qPCR analysis of genes involved in non-specific transport, C4-dicarboxylate transport, TCA cycle related metabolism, and pyruvate metabolism. The averages of four replicates for each sample was used for this plot.

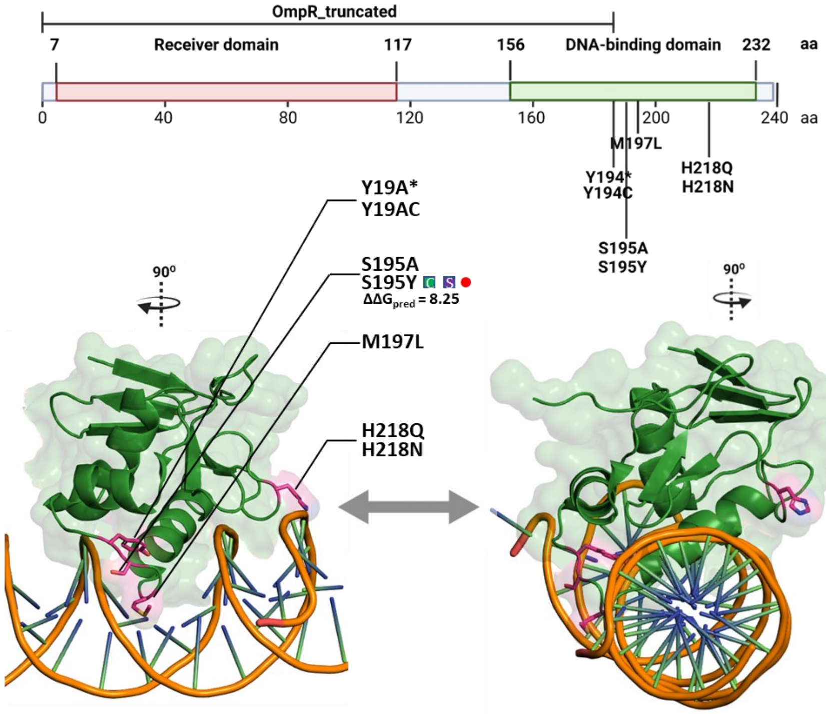


Figure 2.21. Mutations in the DNA binding domain of OmpR observed after STALE of MG1655 grown on ADCA as sole carbon source.

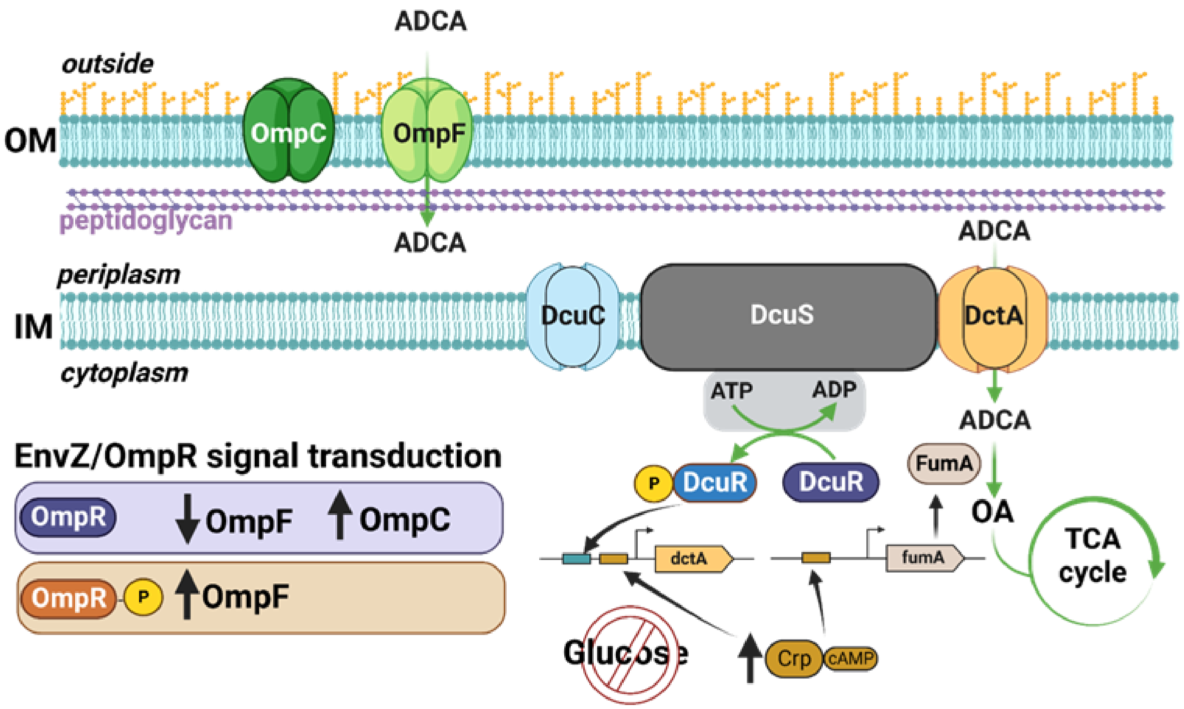
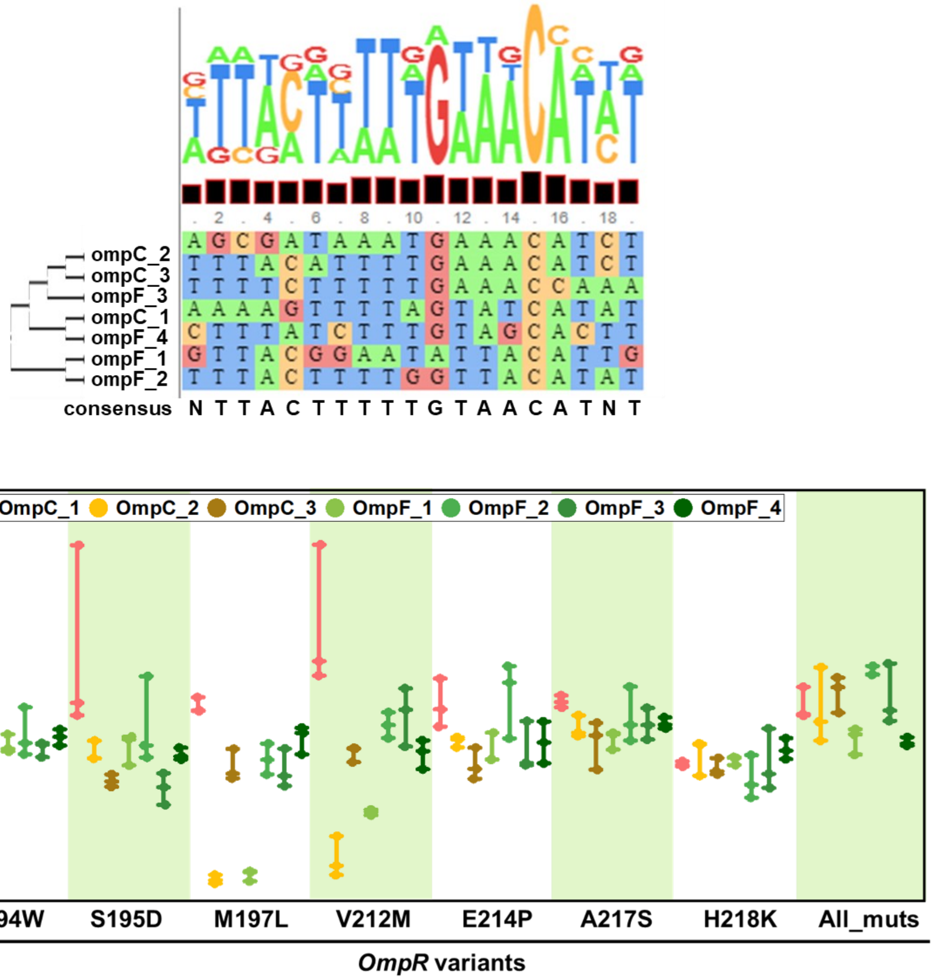


Figure 2.22. Proposed uptake of ADCA across the inner and outer membrane, and subsequent integration into the TCA cycle.

The Y194W, E214P, and A217S variants demonstrated reduced binding to the OmpC_1 site while slightly enhancing binding to OmpF_2. Except for the M197L and H218K variants, all other mutations showed an increased affinity for OmpF binding sites, particularly highlighting a trend towards reduced binding affinity for the OmpC_1 site.

These observations suggest a mutation-driven adaptation to optimize the uptake of ADCA through the larger and faster OmpF channel. Notably, *dctA* was up-regulated in all ACG populations, whereas expression of some C4-dicarboxylates inner-membrane transporters (DauA, DcuB, and DcuD) remained unchanged. *DcuC* was downregulated in F1, ACG1, ACG2, and ACG3 by about 15-fold (Figure 2.20). The observed upregulation of the TCA cycle and pyruvate metabolism-related genes, except for *fumB*, suggests a strategic metabolic shift in *E. coli* to optimize the utilization of ADCA as a carbon source. Among the fumarases, the gene encoding FumA was the most up-regulated, highlighting its role in the metabolic assimilation of ADCA.

FumA has been previously reported to convert ADCA to oxaloacetate.[43] We hypothesized that FumA hydrates cytoplasmic ADCA to produce oxaloacetate, which is then metabolized by the reactions of the TCA cycle and other vital pathways for cell biomass and energy.



	.	2	.	4	.	6	.	8	.	10	.	12	.	14	.	16	.	18	.
ompC_2	A	G	C	G	A	T	A	A	A	T	G	A	A	A	C	A	T	C	T
ompC_3	T	T	T	A	C	A	T	T	T	T	G	A	A	A	C	A	T	C	T
ompF_3	T	T	T	T	C	T	T	T	T	T	G	A	A	A	C	C	A	A	A
ompC_1	A	A	A	A	G	T	T	T	T	A	G	T	A	T	C	A	T	A	T
ompF_4	C	T	T	T	A	T	C	T	T	T	G	T	A	G	C	A	C	T	T
ompF_1	G	T	T	A	C	G	G	A	A	T	A	T	T	A	C	A	T	T	G
ompF_2	T	T	T	A	C	T	T	T	T	G	G	T	T	A	C	A	T	A	T
consensus	N	T	T	A	C	T	T	T	T	T	G	T	A	A	C	A	T	N	T

Figure 2.23. Multiple Sequence Alignments (MSA) of the 20 bp OmpR binding site sequences of *ompC* and *ompF* using clustal omega (version 1.2.4) (top). Docking analysis investigating the effect of mutations in OmpR on binding affinity for various OmpR binding sites upstream of *ompC* and *ompF* (bottom).

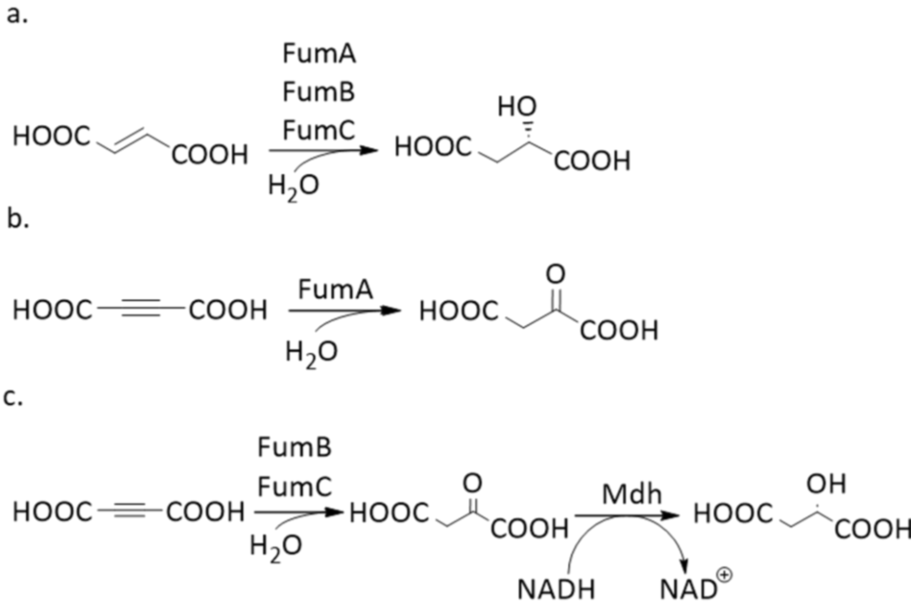


Figure 2.24. Assay of *E. coli* fumarases (FumA, FumB, and FumC) with (a) fumarate and (b) ADCA. Assays for hydration of fumarate (a) were monitored at 250 nm corresponding to loss of fumarate consumption. (b) Reaction was monitored at 275 nm corresponding to production of oxaloacetate. (c) Due to low activity, a coupled enzyme assay was employed for the hydration of ADCA by FumB and FumC.

E. coli possesses three different fumarase isoenzymes (FumA, FumB, and FumC).[44,45] Each of these enzymes is encoded by a distinct gene *fumA*, *fumB*, and *fumC*, respectively. These isoenzymes are integral to the TCA cycle, catalyzing the reversible conversion of fumarate to malate. Despite their shared enzymatic function, the isoenzymes have distinct regulatory mechanisms and functional roles. The isoenzymes also exhibit varying degrees of homology. FumA and FumB share about 80% sequence homology while FumC is significantly different. This suggests that FumA and FumB might have evolved from a common ancestral gene that adapted to different metabolic conditions. An important feature of the fumarases, particularly FumA and FumB, is the presence of iron-sulfur ([4Fe-4S]) clusters. These clusters are critical for the catalytic activity of the enzymes. They facilitate electron transfer during the catalytic process and stabilizing the transition state. FumA and FumB contain these iron-sulfur clusters, contributing to their optimal function in the TCA cycle.[46] The presence of the iron-sulfer clusters make FumA and

FumB unstable in the presence of oxygen. FumA is primarily in the TCA cycle during aerobic respiration. Its expression is regulated by oxygen levels, being more active under aerobic conditions.

To elucidate the contributions of the fumarases to the metabolism of acetylenedicarboxylate, we determined kinetic parameters, k_{cat} (turnover number), K_m (Michaelis constant), and k_{cat}/K_m (catalytic efficiency) for FumA, FumB, and FumC with fumarate (native substrate) and ADCA as substrates (Figure 2.24 and 2.25). We also determine ADCA specificity which is calculated as the ratio of catalytic efficiency (k_{cat}/K_m) for ADCA and fumarate.

FumA showed a higher K_m value for ADCA (2.1 mM) than fumarate (0.7 mM) but better k_{cat} for ADCA (160 $M^{-1}s^{-1}$) than fumarate (89 $M^{-1}s^{-1}$). FumB and FumC have a much higher K_m for ADCA, 28 mM and 17.9 mM, respectively. FumC also showed the lowest k_{cat} for ADCA (0.074 $M^{-1}s^{-1}$). The kinetics confirmed that the fumarases hydrate ADCA to oxaloacetate with a specificity (preference of ADCA over fumarate) of 0.7 for FumA, 0.04 for FumB, and 0.0001 for FumC (Table 2.2). Overall, FumA showed the best catalytic efficiency and specificity for ADCA (Figure 2.26), suggesting it could be the most suitable for hydrating ADCA to oxaloacetate among the fumarases.

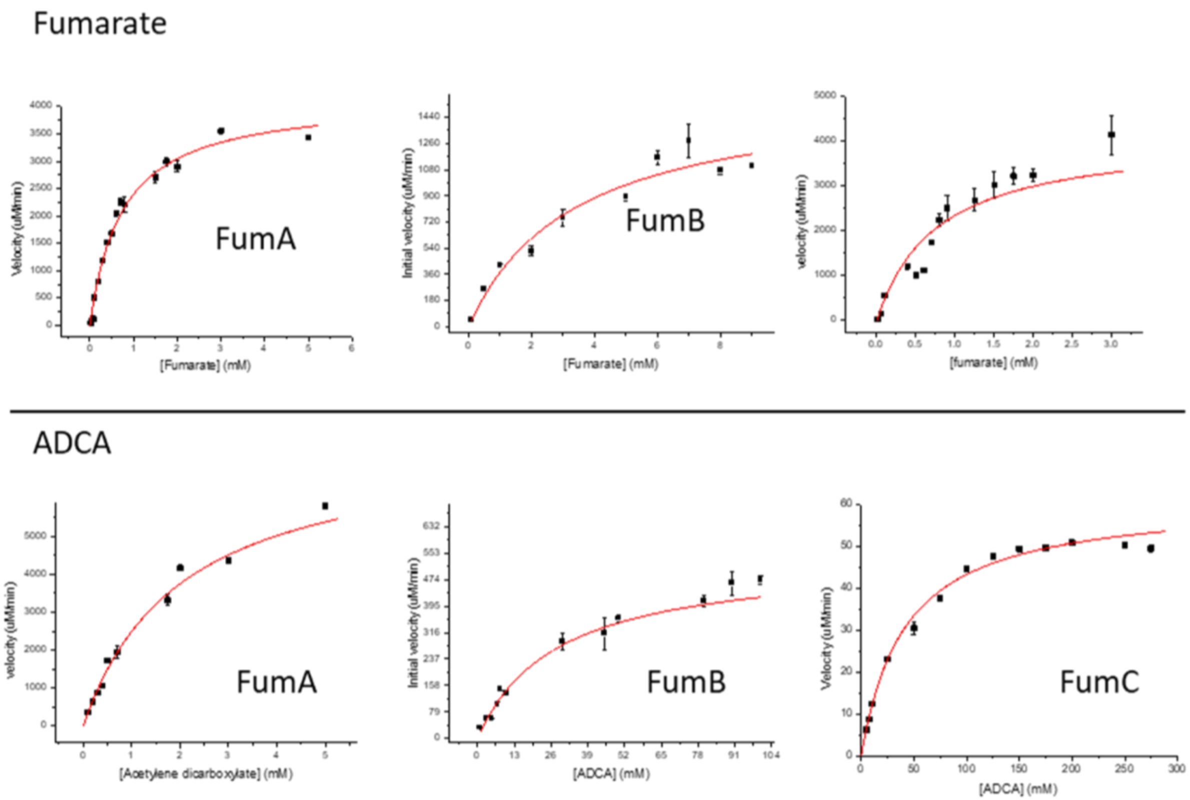


Figure 2.25. Michaelis-Menteen plots for fumarases with fumarate and ADCA substrates

To further investigate the role of the fumarases and other differentially expressed genes in ADCA utilization, we generated knockout mutants (Δ*fumA*, Δ*fumB*, Δ*fumA*Δ*fumB*, Δ*dcuA*, Δ*dcuB*, Δ*dcuC*, Δ*dcuD*, Δ*dcuR*, Δ*dcuS*, Δ*dctA*, Δ*pck*, Δ*gltP*, and Δ*frdA*) by P1 transduction.[47] We then assessed their growth rates and fitness across different liquid media to understand the roles of these genes in ADCA metabolism.

Table 2.2. Kinetic parameters of *E. coli* fumarases with fumarate and acetylenedicarboxylate substrate

Enzyme	Substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)	Enzyme specificity for ADCA
FumA	Fumarate	0.7	8.9×10^1	1.2×10^5	0.7
	ADCA	2.1	1.6×10^2	7.8×10^4	
FumB	Fumarate	3.4	6.8×10^0	2.0×10^3	0.04
	ADCA	28	2.2×10^0	7.8×10^1	
FumC	Fumarate	0.8	2.9×10^1	3.7×10^4	0.0001
	ADCA	17.9	7.4×10^{-2}	4.1×10^0	

We observed that all knockout mutants studied maintained similar fitness to wild-type when cultured in glucose, indicating a lack of general growth impairment of mutants during . Distinct growth profiles were observed in ADCA media (Figure 2.27). Mutants *fumB*, *fumC*, *dcuB*, *dcuC*, *dcuD*, *gltP*, and *frdA* showed slightly improved growth in ADCA media, suggesting these genes may not be critical for ADCA utilization. Conversely, *fumA*, *dcuA*, *dcuR*, *dcuS*, *dctA*, and *pck* mutants did not grow in ADCA media, highlighting their essential roles in this metabolic process. The growth deficiency observed in *dctA*, *dcuR*, and *dcuS* mutants, coupled with the upregulation of *ompR*, *dctA*, and *dcuR*, underscores the significance of these genes in ADCA utilization. These observations reveal adaptation to utilize ADCA efficiently, facilitated by the two dual-component signal transduction systems: the EnvZ/OmpR system, mediating transport from the medium to the periplasm, and the DcuS/DcuR system, mediating transport from the periplasm to the cytoplasm (Figure 2.22).

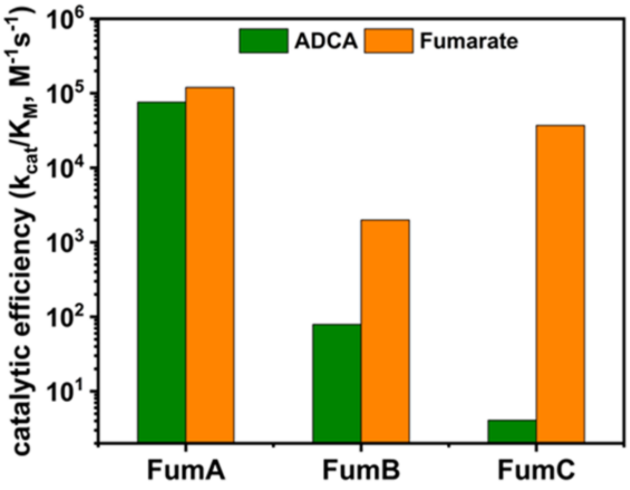


Figure 2.26. Catalytic efficiencies of FumA, FumB, and FumC with ADCA and fumarate substrates.

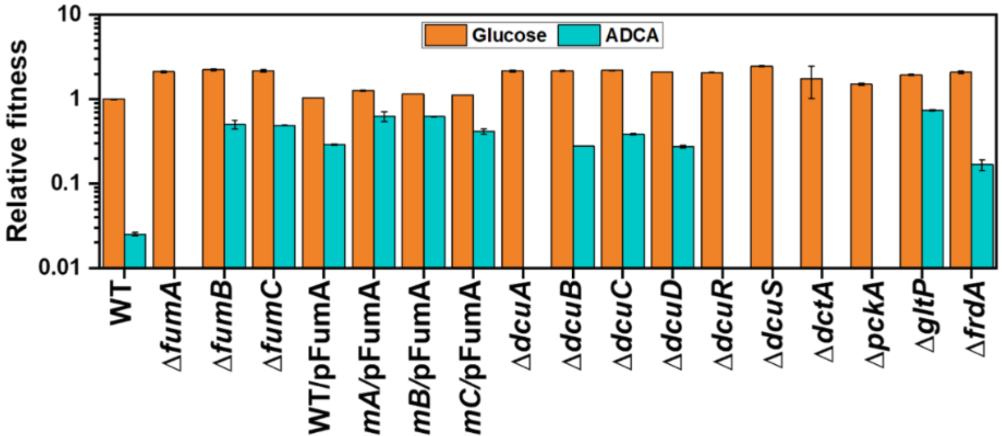


Figure 2.27. Effect of gene knockouts on growth fitness in ADCA and glucose media.

The loss of growth phenotype by the *fumA* mutant was restored when complemented with a plasmid (pFumA) encoding FumA. These findings confirm our hypothesis on the requirement for FumA for ADCA utilization in *E. coli*. The Δ*fumA*Δ*fumB* mutant showed no growth in minimal media tested (supplemented with either glucose, fumarate, or ADCA) but grew in LB-rich media.

To broaden our understanding of microbial adaptation to acetylenedicarboxylate as a carbon source, we extended our growth characterization to include the following microorganisms: *B. subtilis* and *B. circulans* (gram-positive rods), *C. braakii*, *K. oxytoca*, *K. pneumoniae*, *P. putida*, and *P. denitrificans* (gram-negative rods). The Bacillus species and *P. putida* could not utilize ADCA for growth (Figure 2.28). In contrast, *C. braakii* and *K. oxytoca* demonstrated high fitness on ADCA (Figures 2.29 and 2.30), with *K. pneumoniae* and *P. denitrificans* also showing poor growth.

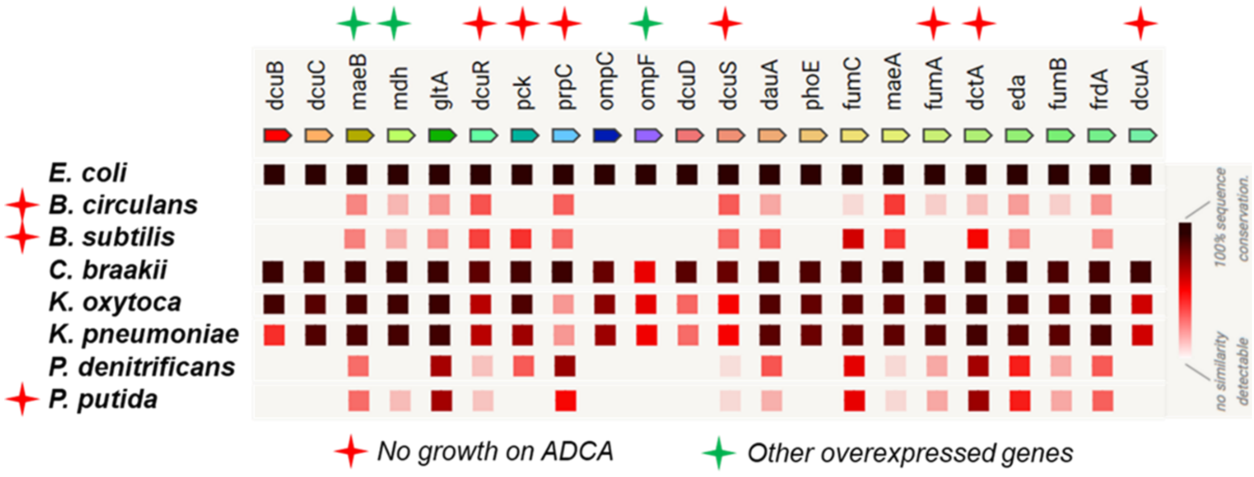


Figure 2.28. Comparison of sequences of various microbes studied for growth on ADCA. Red star on strains indicates no growth on ADCA. Red and green star on genes indicate essential and non-essential genes, respectively. Knocking out essential genes results in no growth on ADCA.

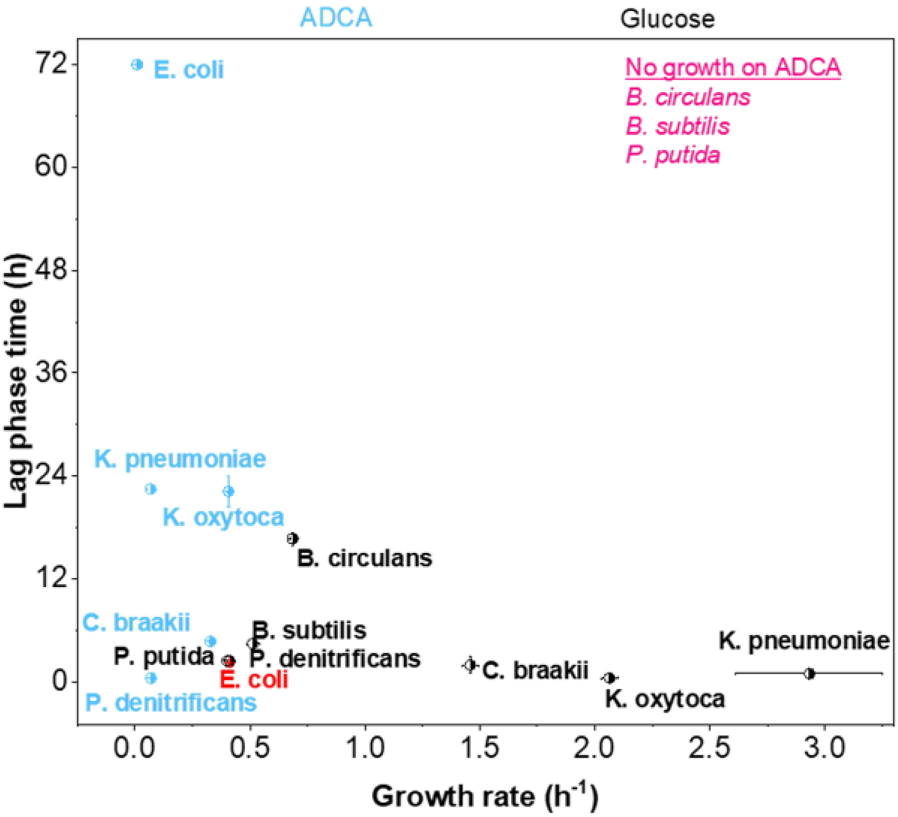
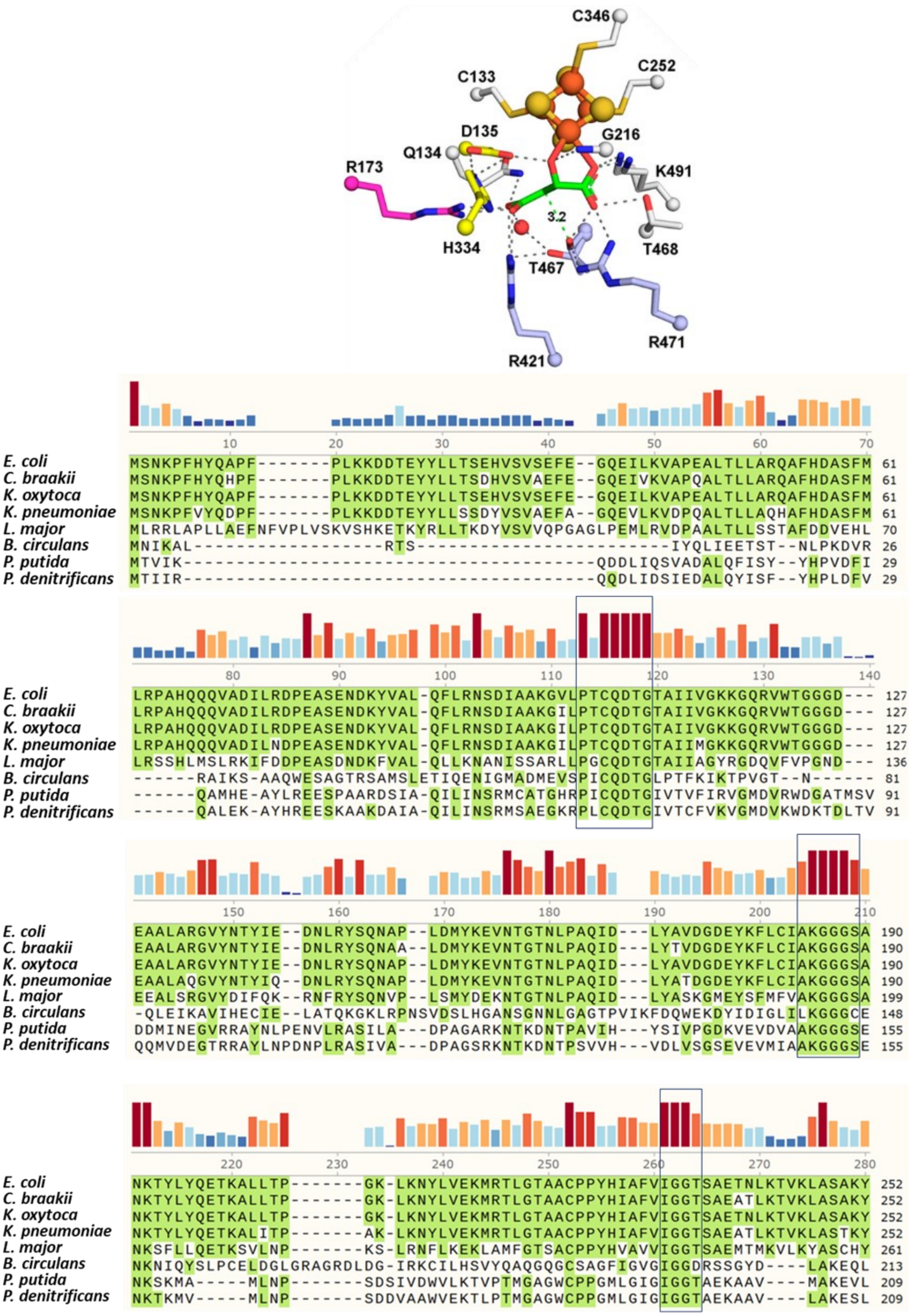


Figure 2.29. Bacterial elasticity plot of microbes comparing growth in glucose and ADCA media. Strains shown in blue were cultivated in ADCA media. Strains shown in black were cultured in glucose. The *E. coli* in red was cultivated in glucose and is highlighted as a control.

We further analyzed the sequence similarity of the 22 genes (from Figure 2.20) previously identified for the gene expression analysis, comparing them across the studied strains (Figure 2.29 and 2.31). We observed that strains capable of growing on ADCA shared a high degree of sequence similarity in these genes, except for *P. denitrificans*. *C. braakii* and *K. oxytoca* significantly improved growth compared to *E. coli* on ADCA. These findings highlight the potential of leveraging strains like *C. braakii* that exhibit better growth on ADCA.



```
E. coli          MSNKPFHYQAPF-------PLKKDDTEYYLLTSEHVSVSEFE--GQEILKVAPEALTLLARQAFHDASFM 61
C. braakii       MSNKPFHYQHPF-------PLKKDDTEYYLLTSDHVSVAEFE--GQEIVKVAPQALTLLARQAFHDASFM 61
K. oxytoca       MSNKPFHYQAPF-------PLKKDDTEYYLLTSEHVSVSEFE--GQEILKVAPEALTLLARQAFHDASFM 61
K. pneumoniae    MSNKPFVYQDPF-------PLKKDDTEYYLLSSDYVSVAEFA--GQEVLKVDPQALTLLAQHAFHDASFM 61
L. major         MLRRLAPLLAEFNFVPLVSKVSHKETKYRLLTKDYVSVVQPGAGLPEMLRVDPAALTLLSSTAFDDVEHL 70
B. circulans     MNIKAL-------------------RTS-----------------------IYQLIEETST--NLPKDVR 26
P. putida        MTVIK---------------------------------------QDDLIQSVADALQFISY--YHPVDFI 29
P. denitrificans MTIIR---------------------------------------QQDLIDSIEDALQYISF--YHPLDFV 29

E. coli          LRPAHQQQVADILRDPEASENDKYVAL-QFLRNSDIAAKGVLPTCQDTGTAIIVGKKGQRVWTGGGD--- 127
C. braakii       LRPAHQQQVADILRDPEASENDKYVAL-QFLRNSDIAAKGILPTCQDTGTAIIVGKKGQRVWTGGGD--- 127
K. oxytoca       LRPAHQQQVADILRDPEASENDKYVAL-QFLRNSDIAAKGILPTCQDTGTAIIVGKKGQRVWTGGGD--- 127
K. pneumoniae    LRPAHQQQVADILNDPEASENDKYVAL-QFLRNSDIAAKGILPTCQDTGTAIIMGKKGQRVWTGGGD--- 127
L. major         LRSSHLMSLRKIFDDPEASDNDKFVAL-QLLKNANISSARLLPGCQDTGTAIIAGYRGDQVFVPGND--- 136
B. circulans     ------RAIKS-AAQWESAGTRSAMSLETIQENIGMADMEVSPICQDTGLPTFKIKTPVGT--N------ 81
P. putida        ------QAMHE-AYLREESPAARDSIA-QILINSRMCATGHRPICQDTGIVTVFIRVGMDVRWDGATMSV 91
P. denitrificans ------QALEK-AYHREESKAAKDAIA-QILINSRMSAEGKRPLCQDTGIVTCFVKVGMDVKWDKTDLTV 91

E. coli          EAALARGVYNTYIE--DNLRYSQNAP--LDMYKEVNTGTNLPAQID---LYAVDGDEYKFLCIAKGGGSA 190
C. braakii       EAALARGVYNTYIE--DNLRYSQNAA--LDMYKEVNTGTNLPAQID---LYTVDGDEYKFLCIAKGGGSA 190
K. oxytoca       EAALARGVYNTYIE--DNLRYSQNAP--LDMYKEVNTGTNLPAQID---LYAVDGDEYKFLCIAKGGGSA 190
K. pneumoniae    EAALAQGVYNTYIQ--DNLRYSQNAP--LDMYKEVNTGTNLPAQID---LYATDGDEYKFLCIAKGGGSA 190
L. major         EEALSRGVYDIFQK--RNFRYSQNVP--LSMYDEKNTGTNLPAQID---LYASKGMEYSFMFVAKGGGSA 199
B. circulans     -QLEIKAVIHECIE--LATQKGKLRPNSVDSLHGANSGNNLGAGTPVIKFDQWEKDYIDIGLILKGGGCE 148
P. putida        DDMINEGVRRAYNLPENVLRASILA---DPAGARKNTKDNTPAVIH---YSIVPGDKVEVDVAAKGGGSE 155
P. denitrificans QQMVDEGTRRAYLNPDNPLRASIVA---DPAGSRKNTKDNTPSVVH---VDLVSGSEVEVMIAAKGGGSE 155

E. coli          NKTYLYQETKALLTP-------GK-LKNYLVEKMRTLGTAACPPYHIAFVIGGTSAETNLKTVKLASAKY 252
C. braakii       NKTYLYQETKALLTP-------GK-LKNYLVEKMRTLGTAACPPYHIAFVIGGTSAEATLKTVKLASAKY 252
K. oxytoca       NKTYLYQETKALLTP-------GK-LKNYLVEKMRTLGTAACPPYHIAFVIGGTSAETNLKTVKLASAKY 252
K. pneumoniae    NKTYLYQETKALITP-------AK-LKNYLVEKMRTLGTAACPPYHIAFVIGGTSAEATLKTVKLASTKY 252
L. major         NKSFLLQETKSVLNP-------KS-LRNFLKEKLAMFGTSACPPYHVAVVIGGTSAEMTMKVLKYASCHY 261
B. circulans     NKNIQYSLPCELDGLGRAGRDLDG-IRKCILHSVYQAQGQGCSAGFIGVGIGGDRSSGYD----LAKEQL 213
P. putida        NKSKMA-----MLNP-------SDSIVDWVLKTVPTMGAGWCPPGMLGIGIGGTAEKAAV----MAKEVL 209
P. denitrificans NKTKMV-----MLNP-------SDDVAAWVEKTLPTMGAGWCPPGMLGIGIGGTAEKAAV----LAKESL 209
```

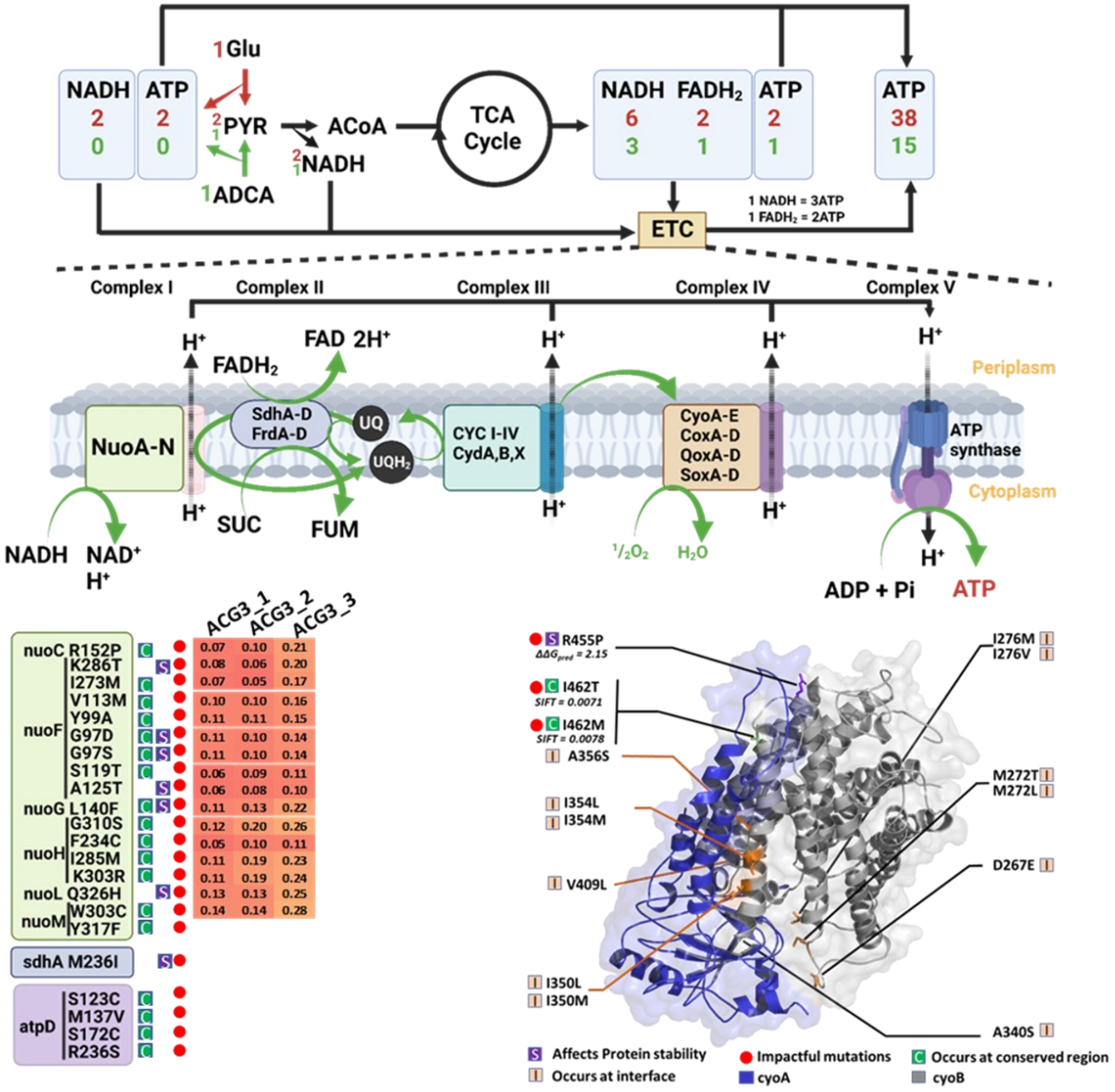
Figure 2.30. Multiple Sequence Alignment of Class I-type fumarase with T-Coffee and Clustal Omega. Top structure shows active site residues of *L. major* Fumarase class I. This is the only known solved structure for a Class I-type fumarase showing iron-sulfur cluster as at the time of this write-up.

We performed multiple sequence analyses comparing the fumarases of strains capable of growth on ADCA (Figure 2.30) to a Class I-type fumarase from *L. major* whose structure has been solved. We observed that strains capable of growth on ADCA shared higher sequence homology with the Class I-type fumarase from *L. major*.

2.7 Regulation of ATP production and bioenergetics in ACG populations.

Under most conditions, glucose is the preferred carbon and energy source for *E coli*.[48] Theoretically, for each molecule of glucose metabolized, 38 molecules of ATP are produced,[49] compared to 15 ATP molecules per ADCA molecule (Figure 2.31). *E. coli* generates energy via substrate-level and oxidative phosphorylation (electron transport chain (ETC) and subsequently utilizes this energy for biosynthetic purposes.[50] The ETC functions to successively transfer electrons from electron donors to electron acceptors while translocating protons from the cytoplasm to the periplasmic space.[51,52] The resulting proton gradient may be used for ATP synthesis.

We observed mutations in nuoC, nuoF, nuoG, nuoH, nuoL, and nuoM**,** which forms the NDH complex, and mutations in sdhA, cyoB, cyoC, cyoE, and atpAD of the ETC (Figure 2.31). The predicted effects of these mutations by the mutFunc suggest most of these mutations are impactful (Figure 2.31 and 2.32) by either affecting protein stability as shown by $\Delta\Delta G_{pred} > 2$ or occurring in conserved regions. These findings suggest that components of the ETC are knocked out during the ALE. There is, however, significant variability in the ETC system of *E. coli,* with various components differing in reactant scope, allowing easy adaptation due to a change in energy budget. Hence, isoenzymes can be used in the absence of others. However, the inability to use specific components affects the efficiency of ATP production since most enzymes in the ETC differ in catalytic efficiencies.[50] We hypothesize that the strains supplement ATP generation with ATP production from substrate-level phosphorylation. The use of ADCA bypasses the need for glycolysis and the Entner-Doudoroff pathway, which are the primary pathways for glucose metabolism, leading to substrate-level phosphorylation. Cytoplasmic ADCA is hydrated directly to oxaloacetate, a TCA cycle intermediate, by FumA.



Gene	Mutation				ACG3_1	ACG3_2	ACG3_3
nuoC	R152P	C		●	0.07	0.10	0.21
nuoF	K286T		S	●	0.08	0.06	0.20
	I273M	C		●	0.07	0.05	0.17
	V113M	C		●	0.10	0.10	0.16
	Y99A	C		●	0.11	0.11	0.15
	G97D	C	S	●	0.11	0.10	0.14
	G97S	C	S	●	0.11	0.10	0.14
	S119T	C		●	0.06	0.09	0.11
	A125T		S	●	0.06	0.08	0.10
nuoG	L140F	C	S	●	0.11	0.13	0.22
nuoH	G310S	C		●	0.12	0.20	0.26
	F234C	C		●	0.05	0.10	0.11
	I285M	C		●	0.11	0.19	0.23
	K303R	C		●	0.11	0.19	0.24
nuoL	Q326H		S	●	0.13	0.13	0.25
nuoM	W303C	C		●	0.14	0.14	0.28
	Y317F	C		●			
sdhA	M236I		S	●			
atpD	S123C	C		●			
	M137V	C		●			
	S172C	C		●			
	R236S	C		●			

Figure 2.31. Generation of ATP via electron transport chain (ETC) from ADCA compared to glucose. Mutations observed in various components of the ETC in evolved ACG populations.

The direct generation of ATP via substrate-level phosphorylation from oxaloacetate is limited because oxaloacetate is primarily metabolized in the TCA cycle, which does not generate ATP directly through substrate-level phosphorylation except in the conversion of succinyl-CoA to succinate. To maximize the yield of ATP from available pathways, the flux through the TCA cycle may increase to enhance succinyl-CoA production and, subsequently, ATP via substrate-level phosphorylation. We observed succinate accumulation during ADCA growth, as discussed in the next section.

To investigate the bioenergetics during growth on ADCA, we determined NAD^+ and NADH ratios, total protein, and ATP concentrations for G1, F1, ACG1, ACG2, and ACG3 (Figure 2.33). The analyses revealed that all the populations produced similar amounts of ATP. The ACG series, however, had higher NAD^+/NADH ratios. The total protein, cofactor amounts, and intracellular redox state have a major impact on growth performance. The NAD^+/NADH ratio has been reported to vary across organisms and is also dependent on growth conditions.[53] The observed increase in the NAD^+/NADH ratio supports the hypothesis of a shift towards more oxidative metabolic processes. A higher NAD^+/NADH ratio indicates that cells are oxidizing NADH to NAD^+ more efficiently. NAD^+ is required for the dehydrogenases that catalyze the oxidation of substrates in the TCA cycle, supporting the claim of carbon flux to succinate for substrate-level ATP production via the TCA cycle. The change in NAD^+/NADH ratio reflects an altered redox state of the cell, which can affect many cellular processes, including regulating metabolic pathways and gene expression.

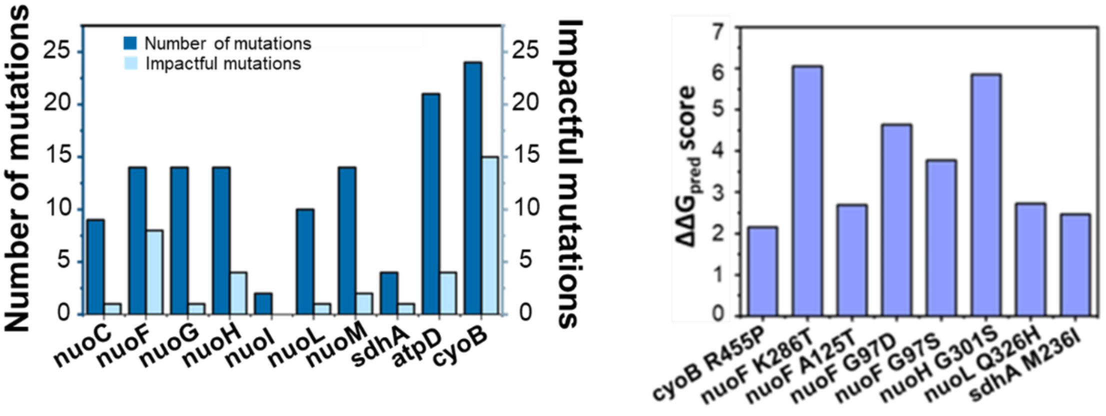


Figure 2.32. Number of mutants occurring in various components of the ETC. Effect of mutations on ΔG_{pred} of genes in the ETC. The difference between wild-type and mutants $\Delta\Delta G_{pred}$ indicates whether or not a mutation will impact protein stability. $\Delta\Delta G_{pred} > 2$ suggests mutation makes protein unstable.

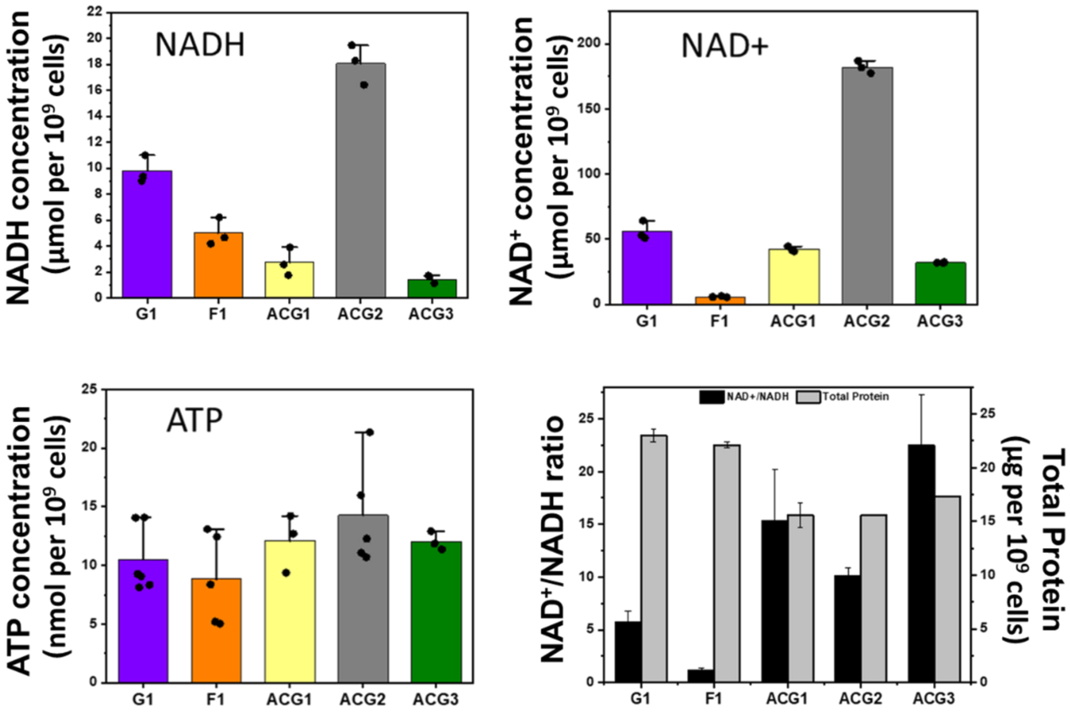


Figure 2.33. Quantification of intracellular NADH, NAD$^+$, ATP, and total protein concentrations in cells cultured in glucose (G1) (purple), fumarate (F1) (orange), and ADCA (ACG1 (yellow), ACG2 (gray), and ACG3(green)).

The reduction in protein levels (Figure 2.33), suggests that the metabolism of ADCA provides less energy, leading to the downregulation of protein synthesis. The occurrence of mutations in protein synthesis machinery could explain this. The cells may prioritize specific proteins necessary for the metabolism of ADCA over others, leading to an overall decrease in protein synthesis. Proteomics will be an excellent approach to determining relative protein levels during the growth on ADCA. Metabolomic analysis could also provide insights into pathways that are more active during the evolution process.

2.8 Differential metabolomics of evolved populations.

Metabolic profiling provides insight into the cellular mechanisms and pathways that underscore substrate utilization.[54–56] This analysis enables the identification of metabolic changes, revealing how differential gene expression and genetic modification translate into functional alterations in metabolic networks. By examining excreted metabolites, we aim to uncover the biochemical strategies employed to optimize microbial growth on ADCA. Metabolomic analysis was conducted using GC-MS and HPLC, revealing significant differences in the metabolic profiles of the ancestral MG1655 and ACG series. To facilitate growth in ADCA media, we also strategically overexpressed the FumA enzyme via a plasmid (pFumA), as hinted by the upregulation of *fumA* detected via transcriptomics during growth on ADCA. Concurrently, a significant reduction in the expression of *dcuC* was observed, suggesting that a *dcuC* knockout could potentially enhance growth on ADCA. We analyzed the metabolomic profiles of various strains, including MG1655 (wild type), MG1655/pFumA (FumA overexpressed), MG1655 Δ*dcuC*/pFumA, and the evolved populations ACG1, ACG2, and ACG3 in the presence of three carbon sources: glucose (20 mM), fumarate (40 mM), and ADCA (40 mM) (Figure 2.34 and 2.35). The metabolic model (Figure 2.34) shows carbon flux from ADCA to acetate, pyruvate, and succinate. All strains exhibited a significant shift in these metabolites during growth in all three carbon sources.

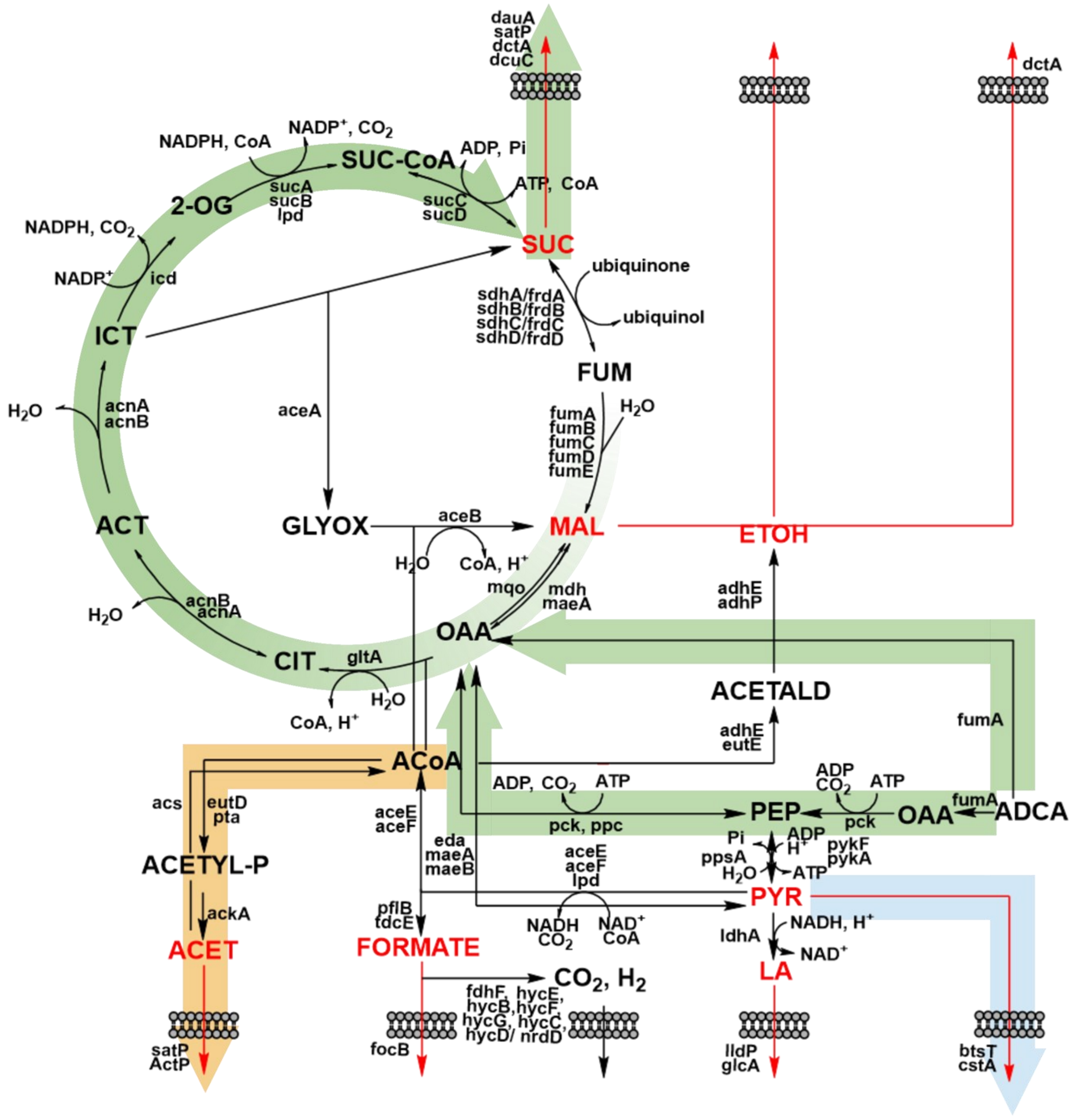


Figure 2.34. Metabolic model constructed using Ecocyc database showing route for excretion of succinate (green), acetate (yellow), and pyruvate (blue).

We observed that wild-type, MG1655/pFumA, and MG1655 Δ*dcuC*/pFumA accumulated about 9 mM, 13 mM, and 18 mM acetate, respectively, within 15 h of culturing in glucose media, which was subsequently consumed after 48 h of culturing (Figure 2.35). In M9F media, fumarate was completely consumed after 24 h for wild-type and MG1655/pFumA, while the Δ*dcuC*/pFumA strain showed relatively poor consumption of fumarate.

MG1655 wild-type showed poor ADCA consumption as expected (ACG1), while MG1655/pFumA and MG1655 Δ*dcuC*/pFumA demonstrated improved consumption of ADCA and accumulated a significant amount of acetate as observed during growth in glucose media. A closer look at the ACG series revealed a progressive improvement in ADCA consumption from ACG1 through to ACG3. About 0.2 mM of acetate was accumulated in ACG1, while ACG2 saw increased acetate accumulation and significant succinate production. The ACG3 strain exhibited the highest rate of ADCA consumption, mirroring patterns observed during the growth on glucose and accumulated acetate. The growth performance of the strains positively correlates with substrate uptake.

Interestingly, mutations were observed along pathways that directed carbon flux from ADCA to succinate via the TCA cycle. This shift in carbon flux resulting in the accumulation of succinate indicates a mechanism to supplement ATP production via substrate-level phosphorylation, specifically by converting succinyl-CoA to succinate. These alterations in metabolite concentrations align consistently with the genetic modifications observed in our evolved strains, reinforcing the correlation between genetic adaptation and metabolic reprogramming.

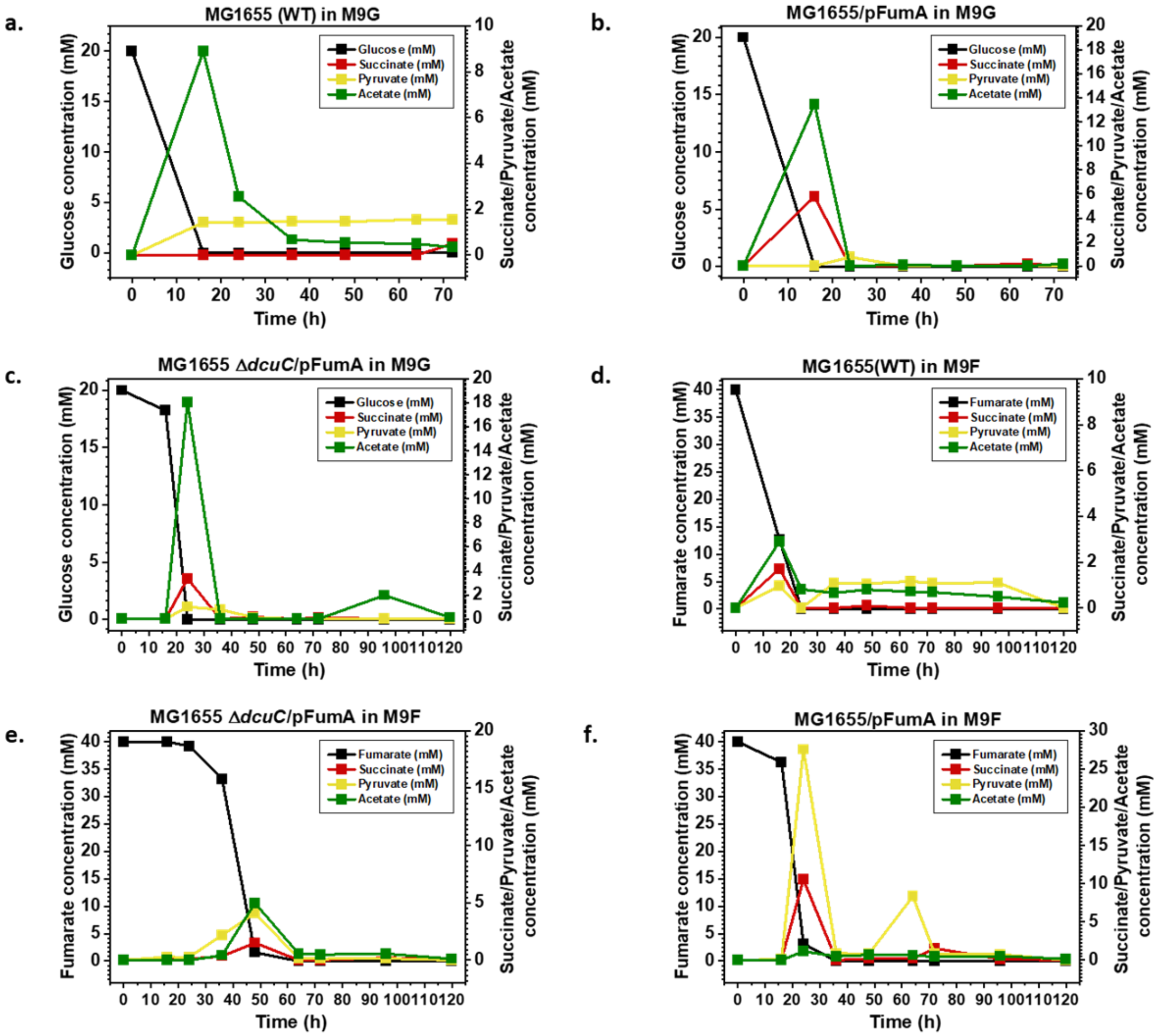


Figure 2.35. Metabolic profiling of evolved ACG series and engineered strains. (a) MG1655 (wt) in M9G (b) MG1655/pFumA in M9G (c) Δ*dcuC*/pFumA in M9G (d) MG1655(wt) in M9F (e) MG1655/pFumA in M9F (f) Δ*dcuC*/pFumA in M9F. (g) MG1655(wt) in M9A (h) MG1655/pFumA in M9A (i) Δ*dcuC*/pFumA in M9A (j) ACG1 in M9A (ACG2). (k) ACG2 in M9A (ACG3) (l) Growth profile of strains in glucose, fumarate, and ADCA.

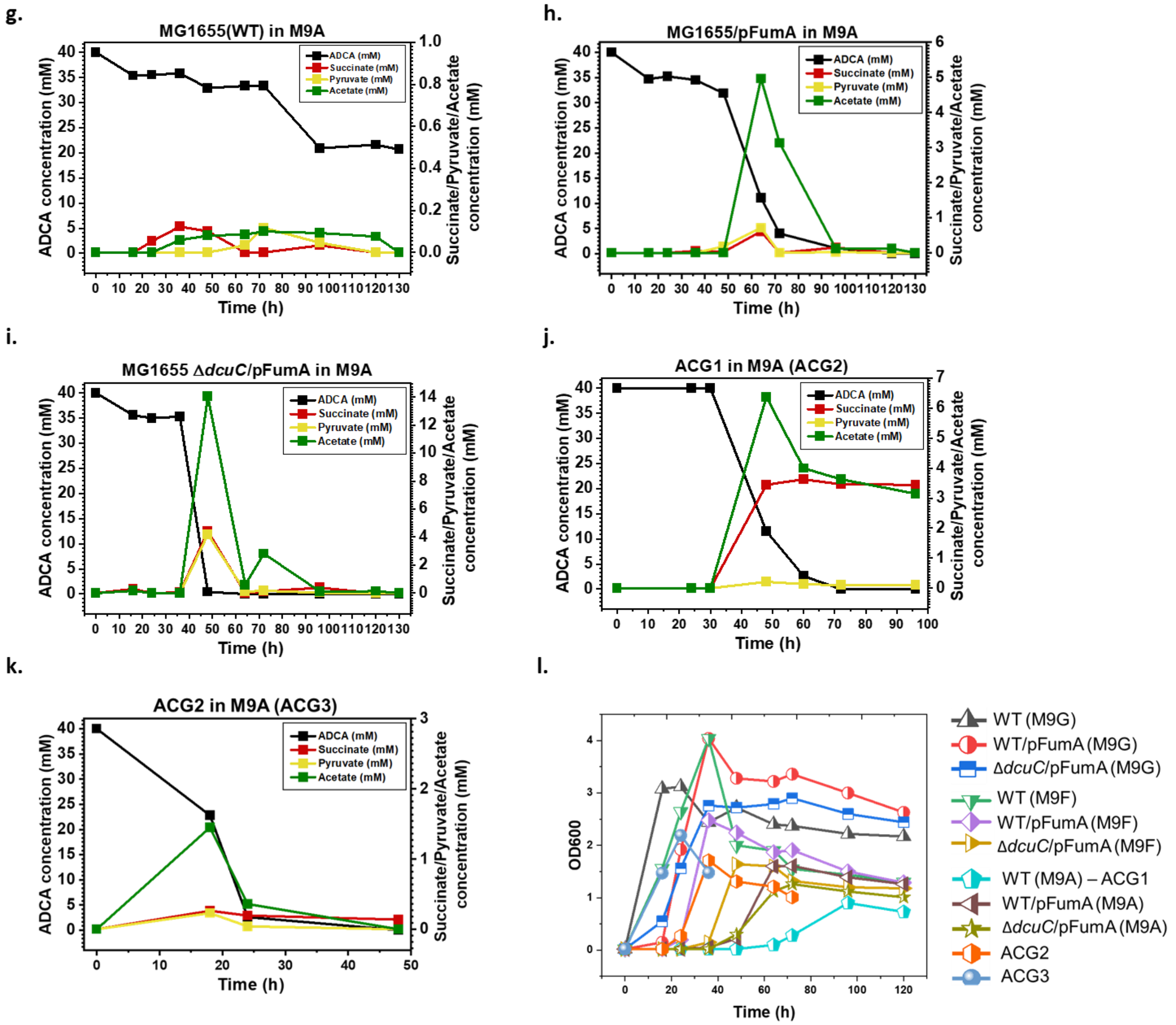


Figure 2.35. CONTD. Metabolic profiling of evolved ACG series and engineered strains. (a) MG1655 (wt) in M9G **(b)** MG1655/pFumA in M9G **(c)** Δ*dcuC*/pFumA in M9G **(d)** MG1655(wt) in M9F **(e)** MG1655/pFumA in M9F **(f)** Δ*dcuC*/pFumA in M9F. **(g)** MG1655(wt) in M9A **(h)** MG1655/pFumA in M9A **(i)** Δ*dcuC*/pFumA in M9A **(j)** ACG1 in M9A (ACG2). **(k)** ACG2 in M9A (ACG3) **(l)** Growth profile of strains in glucose, fumarate, and ADCA.

2.9 Variability in cell dimensions of ACG populations

Understanding how cell characteristics, such as cell size and morphology, affect bioproduction processes is central to the optimization of microbial cell factories. In this section, a comparative analysis of cell dimensional variability between MG1655 (wt) and evolved ACG series is presented. The changes in cell morphology as a result of optimized growth on the unnatural substrate, ADCA, were investigated by scanning electron microscopy. The implications of these changes are significant, affecting substrate uptake, metabolic flux, and the overall efficiency of the microbial production process. Scanning electron microscopy (SEM) images provide an opportunity to snapshot the adaptive morphological responses during the ALE experiment.

We observed a significant variations in cell length and width of ACG populations (Figures 2.37 and 2.38). Interestingly, we observed a wide range of sizes for ACG1 and ACG2. This is evidence of the dynamic nature of the culture and provides evidence for the occurrence of mutators and non-mutators. The reduced cell length and width of the evolved population can significantly affect fermentations in several ways. Cells that are smaller in size have a higher surface area-to-volume ratio. This change can increase the rate of substrate uptake and metabolite excretion, potentially increasing the efficiency of biochemical pathways involved in the production of target chemicals. Smaller cells often exhibit faster growth rates because the reduced size can reduce the time required for cells to duplicate their contents and divide.

The morphological changes observed are consistent with mutations identified in genes associated with cell wall biosynthesis and cell division, namely *murJ* (lipid II flippase), *mreB* (dynamic cytoskeletal protein), *ftsK* (cell division DNA translocase), and *ftsH* (ATP-dependent zinc metalloprotease).

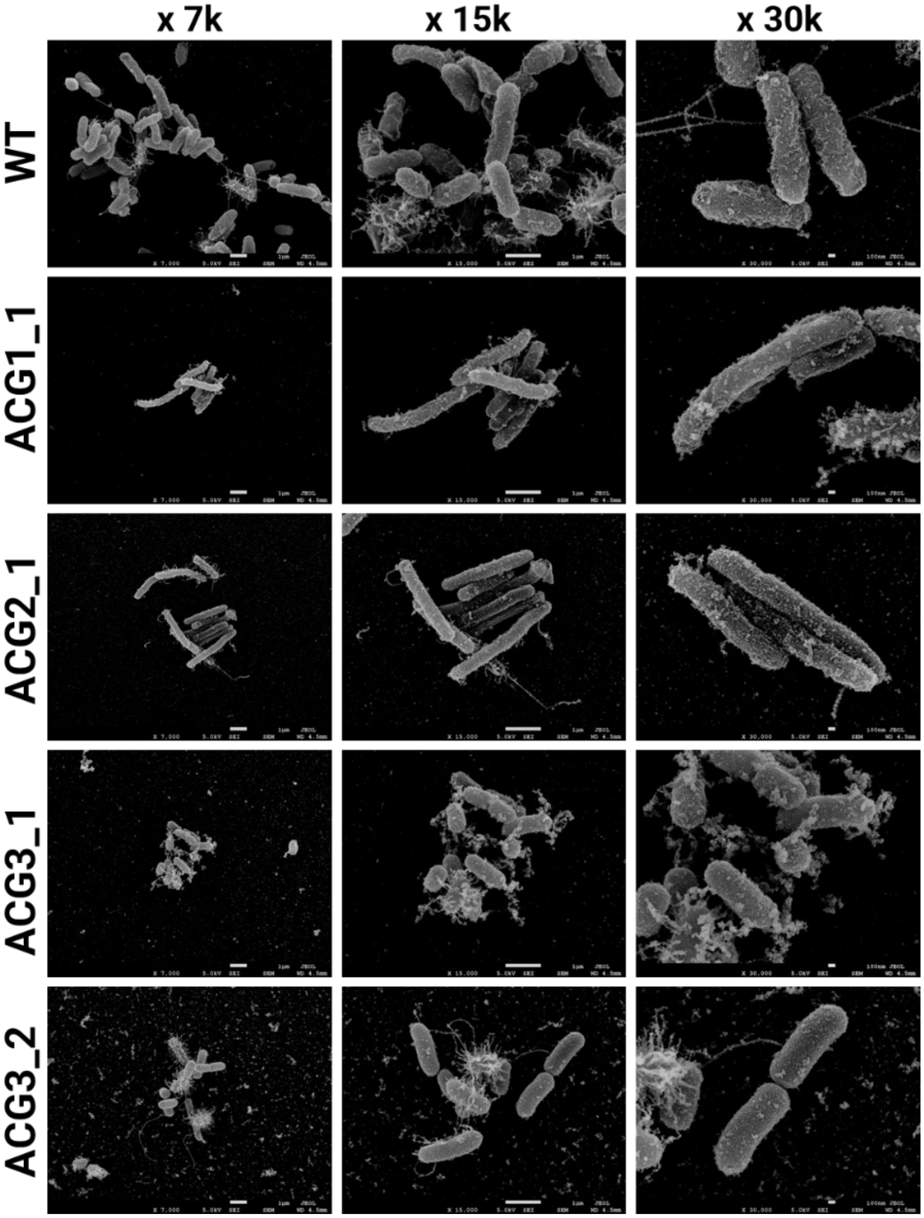


Figure 2.36. SEM images comparing the morphology of *E. coli* MG1655 cultivated in glucose (WT) and the ACG populations grown in ADCA.

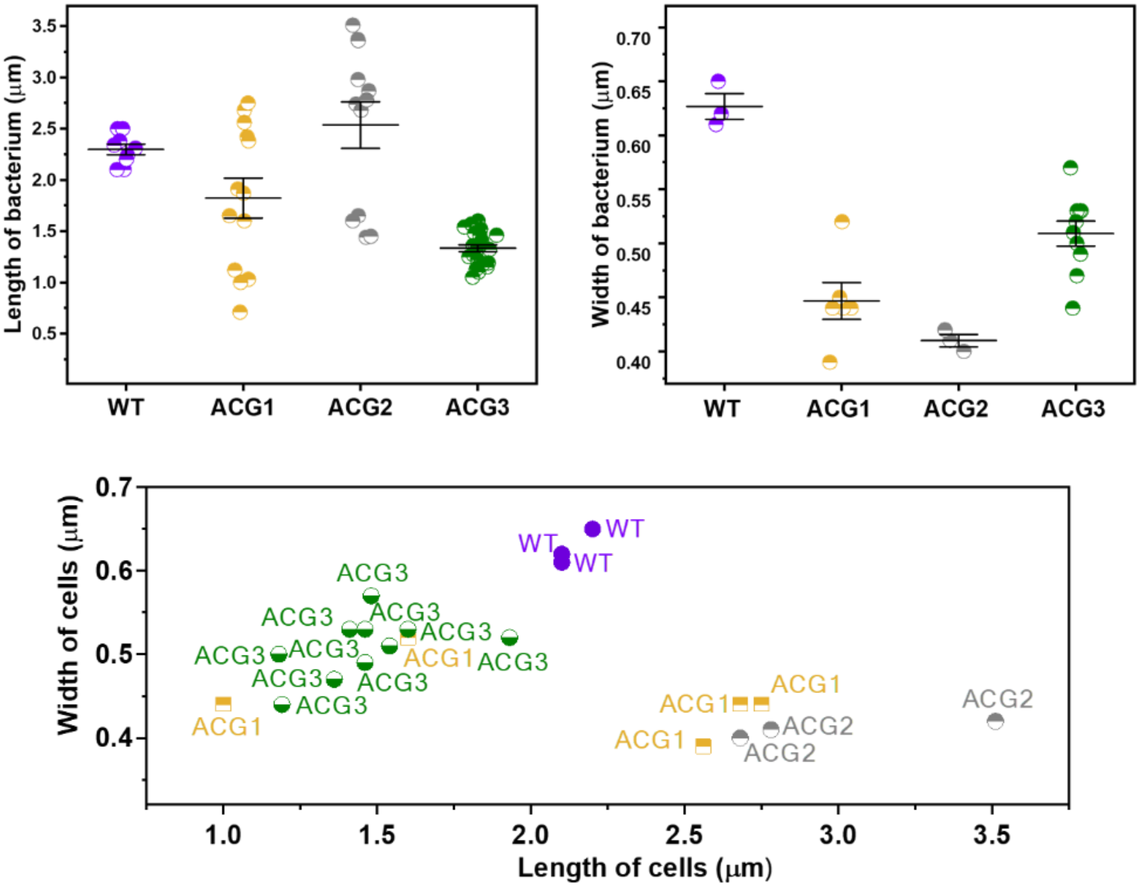


Figure 2.37. Changes in cell sizes after ALE.

Some of these mutations were predicted by the mutfunc database to be impactful, as indicated by SIFT scores and $\Delta\Delta G_{pred}$ (Figure 2.38). MurJ is critical for lipid II flippase activity, which is essential for peptidoglycan biosynthesis in bacteria.

Mutations in *murJ* could disrupt the normal synthesis and maintenance of the cell wall, potentially leading to altered cell shape and size due to defective peptidoglycan assembly. The occurrence of *murJ* mutations in the populations suggests a disruption of the lipid II cycle across the membrane, affecting the overall stability and structure of the cell wall. MreB, a significant component of the bacterial cytoskeleton, plays a central role in determining cell shape by directing the synthesis of peptidoglycan cell walls. Typically, mutations in *mreB* result in cells that are less rod-shaped and more spherical, consistent with the observed reduction in cell length.

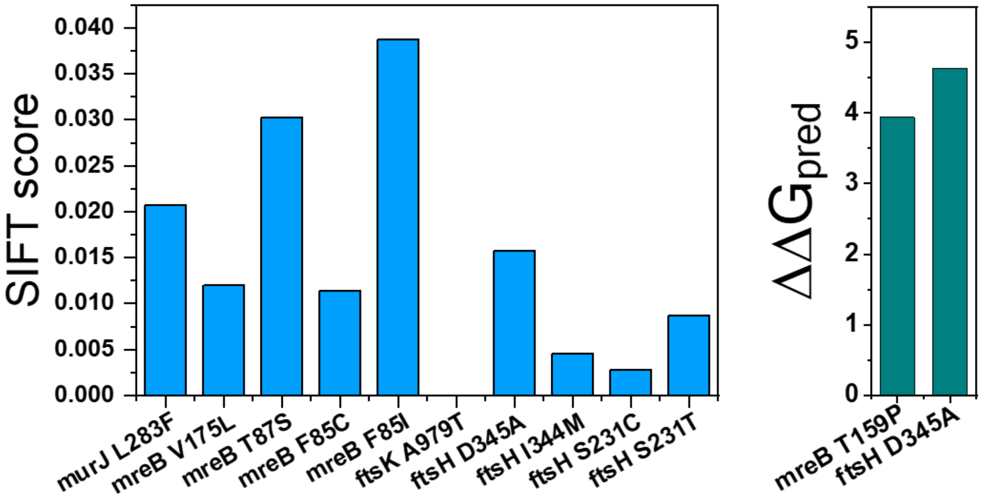


Figure 2.38. Mutfunc predicted effect of mutations occurring in genes associated with *E. coli* cell shape and division.

Observed mutations in genes *ftsK* and *ftsH*, both involved in cell division, also support the observed morphological changes. FtsK is a DNA translocase that coordinates the segregation of chromosomes during cell division and is associated with the division septum, suggesting that mutations in this gene could lead to abnormalities in cell division and, consequently, cell size and shape. Meanwhile, *ftsH* encodes a membrane-bound protease involved in both protein quality control and regulation of the heat shock response; mutations in this gene could indirectly affect cell division through altered regulatory pathways and stress responses.

These findings support the hypothesis that evolutionary pressures during the ALE select for phenotypic variation that confers some adaptive advantage. Future studies should aim to characterize the individual and combined effects of these mutations on cell morphology and fitness. This will improve our understanding of the cellular mechanisms that control bacterial morphology.

2.10 Conclusion

In this chapter we identified OmpF and DctA as key transporters of ADCA from the media into the cytoplasm of *E. coli*. We found that FumA is essential for ADCA metabolism, as it links ADCA to key pathways including the TCA cycle. ADCA is first hydrated by FumA to produce oxaloacetate which may be integrated into the TCA cycle via malate shunt or acetyl CoA. We observed that ACG3 population possessed lots of mutations making them genetically unstable. Hence, further ALE is needed to obtain more stable strains. Another way to obtain stable strains with improved growth on ADCA is to introduce beneficial mutations into a wild-type strain to confer the growth phenotype. We are currently exploiting an automated recombineering using CRISPR-enabled trackable genome engineering (CREATE) to identify beneficial mutations for recapitulation into a wild-type strain.

REFERENCES

(1) Cho, J. S.; Kim, G. B.; Eun, H.; Moon, C. W.; Lee, S. Y. Designing Microbial Cell Factories for the Production of Chemicals. *JACS Au* **2022**, *2* (8), 1781–1799.

(2) Lee, S. Y.; Kim, H. U. Systems Strategies for Developing Industrial Microbial Strains. *Nat. Biotechnol.* **2015**, *33* (10), 1061–1072.

(3) Xue, C.; Zhao, J.; Chen, L.; Yang, S.-T.; Bai, F. Recent Advances and State-of-the-Art Strategies in Strain and Process Engineering for Biobutanol Production by Clostridium Acetobutylicum. *Biotechnol. Adv.* **2017**, *35* (2), 310–322.

(4) Weimer, A.; Kohlstedt, M.; Volke, D. C.; Nikel, P. I.; Wittmann, C. Industrial Biotechnology of *Pseudomonas Putida*: Advances and Prospects. *Appl. Microbiol. Biotechnol.* **2020**, *104* (18), 7745–7766.

(5) Burgstaller, L.; Löffler, S.; De Marcellis, L.; Ghassemi, K.; Neureiter, M. The Influence of Different Carbon Sources on Growth and Single Cell Oil Production in Oleaginous Yeasts *Apiotrichum brassicae* and *Pichia kudriavzevii*. *N. Biotechnol.* **2022**, *69*, 1–7.

(6) Martínez-Gómez, K.; Flores, N.; Castañeda, H. M.; Martínez-Batallar, G.; Hernández-Chávez, G.; Ramírez, O. T.; Gosset, G.; Encarnación, S.; Bolivar, F. New Insights into *Escherichia coli* Metabolism: Carbon Scavenging, Acetate Metabolism and Carbon Recycling Responses during Growth on Glycerol. *Microb. Cell Fact.* **2012**, *11* (1), 46.

(7) Rojo, F. Carbon Catabolite Repression in Pseudomonas : Optimizing Metabolic Versatility and Interactions with the Environment. *FEMS Microbiol. Rev.* **2010**, *34* (5), 658–684.

(8) Gancedo, J. M. Yeast Carbon Catabolite Repression. *Microbiol. Mol. Biol. Rev.* **1998**, *62* (2), 334–361.

(9) Görke, B.; Stülke, J. Carbon Catabolite Repression in Bacteria: Many Ways to Make the Most out of Nutrients. *Nat. Rev. Microbiol.* **2008**, *6* (8), 613–624.

(10) Deutscher, J.; Francke, C.; Postma, P. W. How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria. *Microbiol. Mol. Biol. Rev.* **2006**, *70* (4), 939–1031.

(11) Choudhury, D.; Saini, S. Evolution of *Escherichia coli* in Different Carbon Environments for 2,000 Generations. *J. Evol. Biol.* **2019**, *32* (12), 1331–1341.

(12) Ragauskas, A. J.; Beckham, G. T.; Biddy, M. J.; Chandra, R.; Chen, F.; Davis, M. F.; Davison, B. H.; Dixon, R. A.; Gilna, P.; Keller, M.; Langan, P.; Naskar, A. K.; Saddler, J. N.; Tschaplinski, T. J.; Tuskan, G. A.; Wyman, C. E. Lignin Valorization: Improving Lignin Processing in the Biorefinery. *Science* **2014**, *344* (6185), 1246843.

(13) Rishikesh, M. S.; Harish, S.; Mahendran Prasanth, S.; Gnana Prakash, D. A Comprehensive Review on Lignin Obtained from Agro-Residues: Potential Source of Useful Chemicals. *Biomass Conversion and Biorefinery* **2021**. https://doi.org/10.1007/s13399-021-01637-0.

(14) Zhang, S.; Yang, X.; Zhang, H.; Chu, C.; Zheng, K.; Ju, M.; Liu, L. Liquefaction of Biomass and Upgrading of Bio-Oil: A Review. *Molecules* **2019**, *24* (12). https://doi.org/10.3390/molecules24122250.

(15) Zhang, C.; Ottenheim, C.; Weingarten, M.; Ji, L. Microbial Utilization of Next-Generation Feedstocks for the Biomanufacturing of Value-Added Chemicals and Food Ingredients. *Front Bioeng Biotechnol* **2022**, *10*, 874612.

(16) He, L.; Groom, J. D.; Wilson, E. H.; Fernandez, J.; Konopka, M. C.; Beck, D. A. C.; Lidstrom, M. E. A Methanotrophic Bacterium to Enable Methane Removal for Climate Mitigation. *Proc. Natl. Acad. Sci. U. S. A.* **2023**, *120* (35), e2310046120.

(17) Sahoo, K. K.; Goswami, G.; Das, D. Biotransformation of Methane and Carbon Dioxide Into High-Value Products by Methanotrophs: Current State of Art and Future Prospects. *Front. Microbiol.* **2021**, *12*, 636486.

(18) Nguyen, D. T. N.; Lee, O. K.; Nguyen, T. T.; Lee, E. Y. Type II Methanotrophs: A Promising Microbial Cell-Factory Platform for Bioconversion of Methane to Chemicals. *Biotechnol. Adv.* **2021**, *47*, 107700.

(19) Gleizer, S.; Ben-Nissan, R.; Bar-On, Y. M.; Antonovsky, N.; Noor, E.; Zohar, Y.; Jona, G.; Krieger, E.; Shamshoum, M.; Bar-Even, A.; Milo, R. Conversion of *Escherichia coli* to Generate All Biomass Carbon from CO2. *Cell* **2019**, *179* (6), 1255-1263.e12.

(20) Unden, G.; Strecker, A.; Kleefeld, A.; Kim, O. B. C4-Dicarboxylate Utilization in Aerobic and Anaerobic Growth. *EcoSal Plus* **2016**, *7* (1). https://doi.org/10.1128/ecosalplus.ESP-0021-2015.

(21) Kay, W. W.; Kornberg, H. L. The Uptake of C4-Dicarboxylic Acids by Escherichia Coli. *Eur. J. Biochem.* **1971**, *18* (2), 274–281.

(22) Rhie, M. N.; Cho, Y. B.; Lee, Y. J.; Kim, O. B. High-Affinity l-Malate Transporter DcuE of *Actinobacillus succinogenes* Catalyses Reversible Exchange of C4-Dicarboxylates. *Environ. Microbiol. Rep.* **2019**, *11* (2), 129–139.

(23) Schubert, C.; Unden, G. C4-Dicarboxylates as Growth Substrates and Signaling Molecules for Commensal and Pathogenic Enteric Bacteria in Mammalian Intestine. *J. Bacteriol.* **2022**, *204* (4), e0054521.

(24) Dinh, D. K.; Lee, D. H.; Song, Y.-H.; Jo, S.; Kim, K.-T.; Iqbal, M.; Kang, H. Efficient Methane-to-Acetylene Conversion Using Low-Current Arcs. *RSC Adv.* **2019**, *9* (56), 32403–32413.

(25) Fincke, J. R.; Anderson, R. P.; Hyde, T.; Detering, B. A.; Wright, R.; Bewley, R. L.; Haggard, D. C.; Swank, W. D. Plasma Thermal Conversion of Methane to Acetylene. *Plasma Chem. Plasma Process.* **2002**, *22* (1), 105–136.

(26) Wang, X.; Lim, Y. N.; Lee, C.; Jang, H.-Y.; Lee, B. Y. 1,5,7-Triazabicyclo[4.4.0]Dec-1-Ene-Mediated Acetylene Dicarboxylation and Alkyne Carboxylation Using Carbon Dioxide. *Eur. J. Org. Chem.* **2013**, *2013* (10), 1867–1871.

(27) Yamada, E. W.; Jakoby, W. B. Enzymatic Utilization of an Acetylenic Compound. *J. Am. Chem. Soc.* **1958**, *80* (9), 2343–2344.

(28) Yamada, E. W.; Jakoby, W. B. Enzymatic Utilization of Acetylenic Compounds. I. An Enzyme Converting Acetylenedicarboxylic Acid to Pyruvate. *J. Biol. Chem.* **1958**, *233* (3), 706–711.

(29) Maddamsetti, R.; Grant, N. A. Divergent Evolution of Mutation Rates and Biases in the Long-Term Evolution Experiment with *Escherichia coli*. *Genome Biol. Evol.* **2020**, *12* (9), 1591–1603.

(30) Barrick, J. E.; Lenski, R. E. Genome-Wide Mutational Diversity in an Evolving Population of Escherichia Coli. *Cold Spring Harb. Symp. Quant. Biol.* **2009**, *74*, 119–129.

(31) Van den Bergh, B.; Swings, T.; Fauvart, M.; Michiels, J. Experimental Design, Population Dynamics, and Diversity in Microbial Experimental Evolution. *Microbiol. Mol. Biol. Rev.* **2018**, *82* (3). https://doi.org/10.1128/MMBR.00008-18.

(32) Wagih, O.; Galardini, M.; Busby, B. P.; Memon, D.; Typas, A.; Beltrao, P. A Resource of Variant Effect Predictions of Single Nucleotide Variants in Model Organisms. *Mol. Syst. Biol.* **2018**, *14* (12), e8430.

(33) Boutte, C. C.; Crosson, S. Bacterial Lifestyle Shapes Stringent Response Activation. *Trends Microbiol.* **2013**, *21* (4), 174–180.

(34) Kraemer, J. A.; Sanderlin, A. G.; Laub, M. T. The Stringent Response Inhibits DNA Replication Initiation in *E. coli* by Modulating Supercoiling of OriC. *MBio* **2019**, *10* (4). https://doi.org/10.1128/mBio.01330-19.

(35) Potrykus, K.; Cashel, M. (P)PpGpp: Still Magical? *Annu. Rev. Microbiol.* **2008**, *62*, 35–51.

(36) Schubert, C.; Kim, N. Y.; Unden, G.; Kim, O. B. C4-Dicarboxylate Metabolons: Interaction of C4-Dicarboxylate Transporters of *Escherichia coli* with Cytosolic Enzymes. *FEMS Microbiol. Lett.* **2022**, *369* (1). https://doi.org/10.1093/femsle/fnac078.

(37) Kenney, L. J.; Anand, G. S. EnvZ/OmpR Two-Component Signaling: An Archetype System That Can Function Noncanonically. *EcoSal Plus* **2020**, *9* (1). https://doi.org/10.1128/ecosalplus.ESP-0001-2019.

(38) Gerken Henri; Vuong Phu; Soparkar Ketaki; Misra Rajeev. Roles of the EnvZ/OmpR Two-Component System and Porins in Iron Acquisition in *Escherichia coli*. *MBio* **2020**, *11* (3), e01192-20.

(39) Russo, F. D.; Slauch, J. M.; Silhavy, T. J. Mutations That Affect Separate Functions of OmpR the Phosphorylated Regulator of Porin Transcription in *Escherichia coli*. *J. Mol. Biol.* **1993**, *231* (2), 261–273.

(40) Mattison, K.; Oropeza, R.; Byers, N.; Kenney, L. J. A Phosphorylation Site Mutant of OmpR Reveals Different Binding Conformations at OmpF and OmpC. *J. Mol. Biol.* **2002**, *315* (4), 497–511.

(41) Remmert, M.; Biegert, A.; Hauser, A.; Söding, J. HHblits: Lightning-Fast Iterative Protein Sequence Searching by HMM-HMM Alignment. *Nat. Methods* **2011**, *9* (2), 173–175.

(42) Pearson, W. R.; Lipman, D. J. Improved Tools for Biological Sequence Comparison. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85* (8), 2444–2448.

(43) Flint, D. H. Initial Kinetic and Mechanistic Characterization of *Escherichia coli* Fumarase A. *Arch. Biochem. Biophys.* **1994**, *311* (2), 509–516.

(44) Woods, S. A.; Schwartzbach, S. D.; Guest, J. R. Two Biochemically Distinct Classes of Fumarase in *Escherichia coli*. *Biochim. Biophys. Acta* **1988**, *954* (1), 14–26.

(45) Bock, R. M.; Alberty, R. A. Studies of the Enzyme Fumarase. I. Kinetics and Equilibrium. *J. Am. Chem. Soc.* **1953**, *75* (8), 1921–1925.

(46) Flint, D. H.; Emptage, M. H.; Guest, J. R. Fumarase a from *Escherichia coli*: Purification and Characterization as an Iron-Sulfur Cluster Containing Enzyme. *Biochemistry* **1992**, *31* (42), 10331–10337.

(47) Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H. Construction of *Escherichia coli* K-12 in-Frame, Single-Gene Knockout Mutants: The Keio Collection. *Mol. Syst. Biol.* **2006**, *2*, 2006.0008.

(48) Bren, A.; Park, J. O.; Towbin, B. D.; Dekel, E.; Rabinowitz, J. D.; Alon, U. Glucose Becomes One of the Worst Carbon Sources for *E. coli* on Poor Nitrogen Sources Due to Suboptimal Levels of CAMP. *Sci. Rep.* **2016**, *6*, 24834.

(49) Gottschalk, G. Biosynthesis of *Escherichia coli* Cells from Glucose. *Bacterial Metabolism*; Gottschalk, G., Ed.; Springer US: New York, NY, 1979; pp 34–80.

(50) Henkel, S. G.; Ter Beek, A.; Steinsiek, S.; Stagge, S.; Bettenbrock, K.; de Mattos, M. J. T.; Sauter, T.; Sawodny, O.; Ederer, M. Basic Regulatory Principles of *Escherichia coli's* Electron Transport Chain for Varying Oxygen Conditions. *PLoS One* **2014**, *9* (9), e107640.

(51) Bonora, M.; Patergnani, S.; Rimessi, A.; De Marchi, E.; Suski, J. M.; Bononi, A.; Giorgi, C.; Marchi, S.; Missiroli, S.; Poletti, F.; Wieckowski, M. R.; Pinton, P. ATP Synthesis and Storage. *Purinergic Signal.* **2012**, *8* (3), 343–357.

(52) Reguera, G. Biological Electron Transport Goes the Extra Mile. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (22), 5632–5634.

(53) Zhou, Y. J.; Yang, W.; Wang, L.; Zhu, Z.; Zhang, S.; Zhao, Z. K. Engineering NAD^+ Availability for *Escherichia coli* Whole-Cell Biocatalysis: A Case Study for Dihydroxyacetone Production. *Microb. Cell Fact.* **2013**, *12*, 103.

(54) Miguez, A. M.; McNerney, M. P.; Styczynski, M. P. Metabolic Profiling of *Escherichia coli*-Based Cell-Free Expression Systems for Process Optimization. *Ind. Eng. Chem. Res.* **2019**, *58* (50), 22472–22482.

(55) Chaudhuri, R. R.; Sebaihia, M.; Hobman, J. L.; Webber, M. A.; Leyton, D. L.; Goldberg, M. D.; Cunningham, A. F.; Scott-Tucker, A.; Ferguson, P. R.; Thomas, C. M.; Frankel, G.; Tang, C. M.; Dudley, E. G.; Roberts, I. S.; Rasko, D. A.; Pallen, M. J.; Parkhill, J.; Nataro,

J. P.; Thomson, N. R.; Henderson, I. R. Complete Genome Sequence and Comparative Metabolic Profiling of the Prototypical Enteroaggregative *Escherichia coli* Strain 042. *PLoS One* **2010**, *5* (1), e8801.

(56) Kayser, A.; Weber, J.; Hecht, V.; Rinas, U. Metabolic Flux Analysis of *Escherichia coli* in Glucose-Limited Continuous Culture. I. Growth-Rate-Dependent Metabolic Efficiency at Steady State. *Microbiology* **2005**, *151* (Pt 3), 693–706.

CHAPTER THREE: Characterization of Malic Enzyme B and ALE-Induced Variants

3.1 Introduction

The central metabolic pathways in most chemotrophic, aerobic, and facultatively anaerobic bacteria are primarily the Embden-Meyerhof-Parnas (EMP) pathway (glycolysis) or the Entner-Doudoroff (ED) pathway and the tricarboxylic acid (TCA) cycle.[1–4] The EMP pathway enables the production of ATP, NADH, and precursors like pyruvate or 3-phosphoglycerate from glucose metabolism. It is the most common pathway for glucose breakdown in many types of cells or organisms. The ED pathway is an alternate pathway used by some bacteria and archaea that additionally allows the breakdown of sugar acids like gluconate. These pathways are critical for the breakdown of carbohydrates to phospho*enol*pyruvate (PEP), pyruvate, and acetyl-CoA, providing both energy and essential building blocks for the synthesis of cellular components. While the EMP and ED pathways are crucial for the initial breakdown and utilization of glucose and sugar acids, the intermediates or end-products from these pathways are shuttled through the TCA cycle. The TCA cycle plays a role in both catabolism and anabolism, enabling the oxidation of acetyl-CoA to CO_2 for respiratory ATP formation and providing carbon precursor metabolites and NAD(P)H for biosynthetic processes.[5]

The PEP-pyruvate-oxaloacetate (PPO) node, also known as the anaplerotic node, is a critical metabolic link between EMP/ED/gluconeogenesis and the TCA cycle.[4,6] This node directs carbon flux in appropriate directions and acts as a highly relevant switching point for carbon flux distribution within central metabolism. The PPO node involves a series of reactions that are essential for flexible responses to cellular demands, thereby adjusting carbon flux to meet both energetic and anabolic requirements under different growth conditions, such as during growth on C4-dicarboxylates. For instance, when growing microbes on TCA cycle intermediates or substrates that are integrated into central metabolism via acetyl-CoA, such as acetate, malate, oxaloacetate, or ADCA, as proposed in Chapter Two, these carboxylates must be converted to pyruvate and PEP by the activities of the PPO node (Figure 3.1). This conversion is critical for the synthesis of glycolytic intermediates through gluconeogenesis, which is accomplished by the reversible reactions of glycolysis.[4,7]

During the growth of *E. coli* on ADCA, fumarase A (FumA) catalyzes the hydration of ADCA to produce oxaloacetate, which can subsequently enter the TCA cycle directly. Alternatively, oxaloacetate can be converted to pyruvate and then to PEP, providing another entry point via acetyl-CoA back to oxaloacetate under gluconeogenic conditions (Figure 3.1). This flexibility in the carbon flux through the PPO node is essential as it allows cells to tightly regulate these fluxes based on the specific demands of the culture condition and available carbon source.

This regulatory role ensures optimal metabolic responses to environmental nutrient changes, demonstrating the PPO node's pivotal role at the junction of glycolysis, the TCA cycle, and other metabolic pathways. This central role is further highlighted by the functions of the intermediates: PEP, which possesses the highest-energy phosphate bond among all natural organo-phosphates,[8] serves as a precursor for aromatic amino acids. Pyruvate, on the other hand, is a vital precursor of several amino acids, like alanine, valine, leucine, and isoleucine, and as the primary entry point for many metabolic pathways that re-oxidize NAD(P)H.[9] Oxaloacetate, which is part of the TCA cycle, functions as a precursor for aspartate, leading to the biosynthesis of many amino acids and nucleotides.[10]

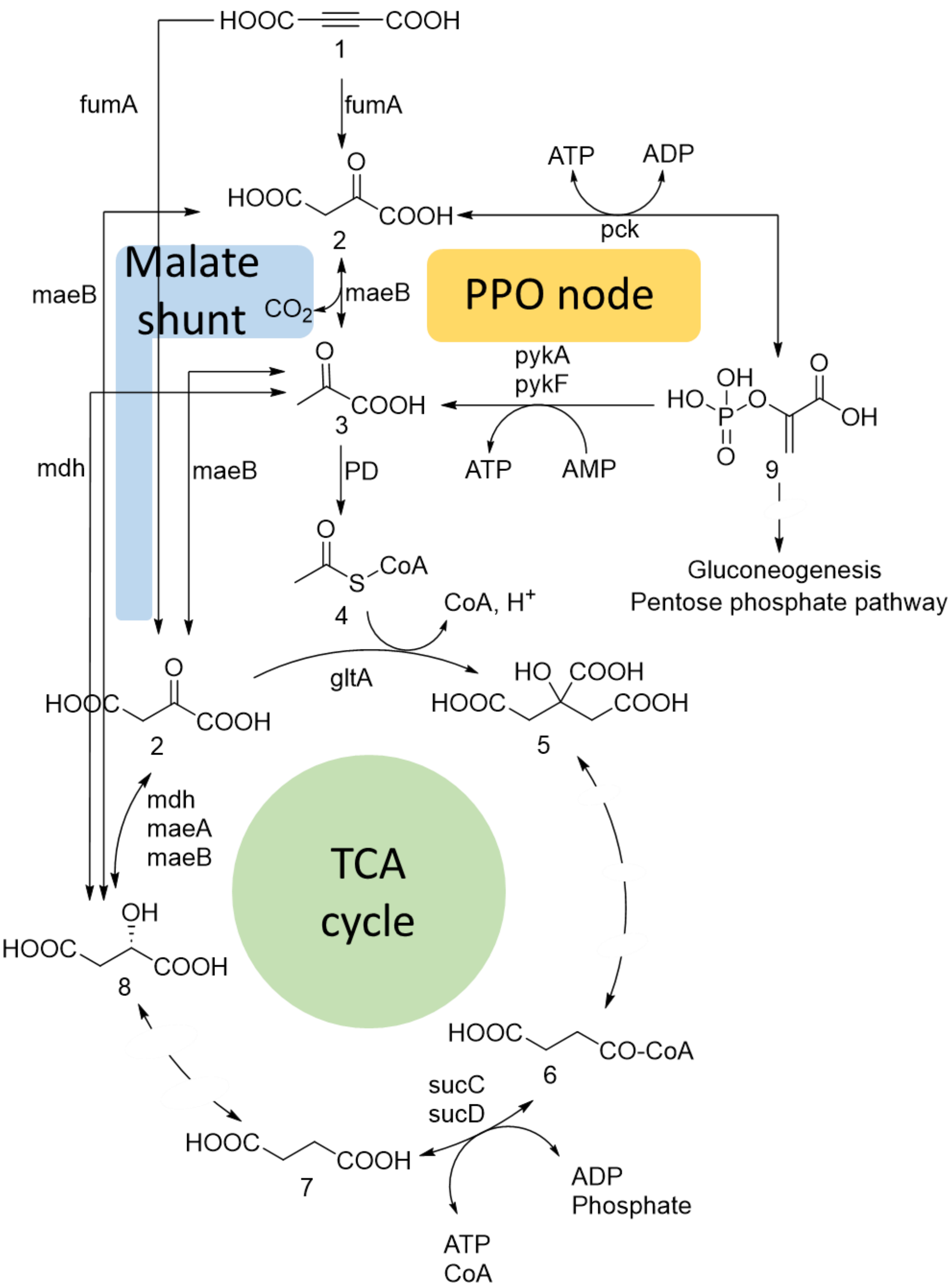


Figure 3.1. Metabolic model showing assimilation of ADCA into phospho*enol*pyruvate-pyruvate-oxaloacetate (PPO) node, malate shunt, and TCA cycle. First ADCA (1) is converted to oxaloacetate (2) by Fumarase A (FumA). The PPO node is highlighted in yellow consisting of oxaloacetate (2), pyruvate (3), and phospho*enol*pyruvate (9). The malate shunt is highlighted in blue. The TCA cycle is highlighted in green, showing selected intermediates acetyl CoA (4), citrate (5), succinyl CoA (6), succinate (7), (*S*)-malate (8), and oxaloacetate (2). The enzymes shown are; maeA, malic enzyme A; maeB, malic enzyme B; pck, phospho*enol*pyruvate carboxykinase; ppsA, phospho*enol*pyruvate synthetase; PD, pyruvate dehydrogenase; mdh, malate dehydrogenase; gltA, citrate synthase; sucC, succinyl-CoA synthetase subunit β ; sucD, succinyl-CoA synthetase subunit α.

As discussed in Chapter Two, the conversion of ADCA to oxaloacetate is essential for the utilization of ADCA as a sole source of carbon and energy. The cytosolic oxaloacetate produced can be acted upon by several enzymes, including malic enzymes (MaeA and MaeB), malate dehydrogenase (Mdh), and phospho*enol*pyruvate carboxykinase (Pck), which are involved in the PPO node (Figure 3.1). During the transcriptomic analysis of the evolved ACG populations, we observed upregulation of genes encoding Pck, Mdh, and MaeB. This differential expression suggests that the PPO node is more active during growth on ADCA to facilitate the integration of oxaloacetate into other key pathways for cell biomass and energy generation.

E. coli expresses two malic enzymes, MaeA [EC 1.1.1.38] and MaeB [EC 1.1.1.40], that primarily catalyze the oxidative decarboxylation of malate to pyruvate and CO_2 accompanied by reduction of $NAD(P)^+$. MaeA is NAD^+-dependent, while MaeB requires $NADP^+$.[11] MaeB is a multifunctional enzyme that possesses activities including reductase activity - conversion of oxaloacetate to (*S*)-malate in the presence of NADPH, and decarboxylase activity - oxaloacetate to pyruvate in the absence of NADPH.[12] Malic enzymes play critical roles in the PPO node and malate shunt, providing an influence over major carbon fluxes and linking gluconeogenesis to the TCA cycle. Prokaryotic MaeB has also been reported to be required for optimal growth on TCA cycle intermediates.[12] Some prokaryotic MaeB has been reported to have a multimodular structure and possess two domains: the malic enzyme domain responsible for catalysis and an appended phosphotransacetylase (PTA) domain that forms a hexameric sensor system which relays acetyl-CoA occupancy to the active site of the malic enzyme domain for feedback inhibition through an allosteric rotary mechanism (Figure 3.2).[13] While the crystal structure of *E. coli* MaeA variants has been elucidated, *E. coli* MaeB structure has not been reported yet.

Interestingly, we observed ten mutations in MaeB in ACG3 populations, some of them occurring across all three lineages of populations investigated. These mutations were Y145E, Q148K, K149A, R151I, R153Q, M154C, N155D, G335R, T337S, and A338V. It was more interesting that the mutations were localized in two regions, between residues 145 to 155 and 335 to 338, all in the ME domain. The positions of mutated residues are indicated in a MaeB model structure from an AlphaFold-predicted structure later in this chapter.[14]

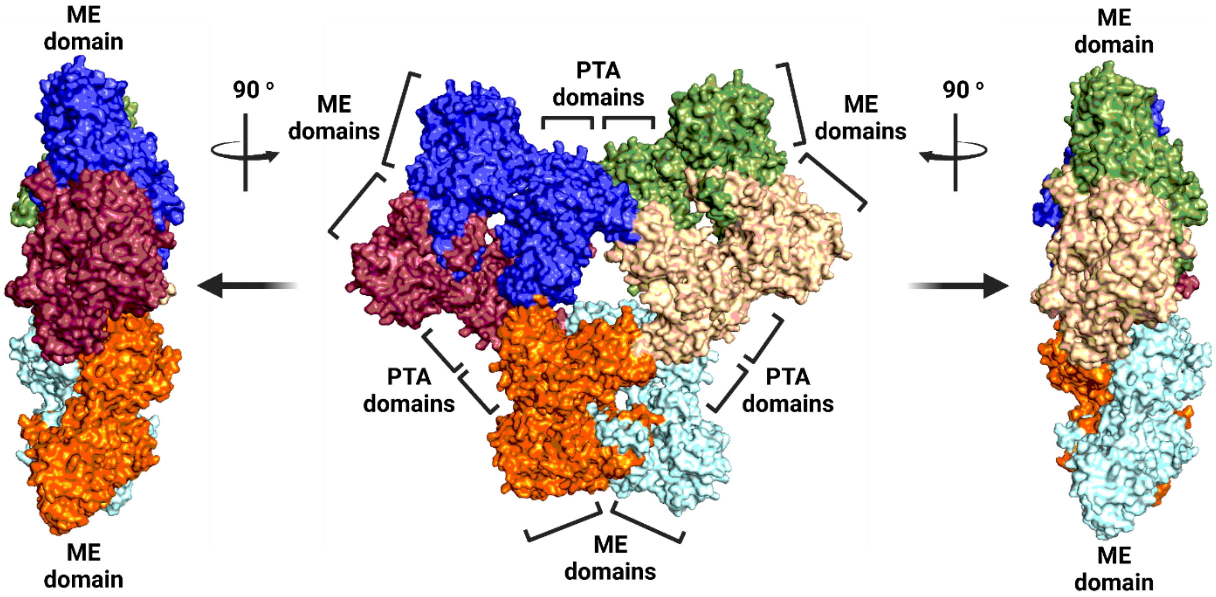
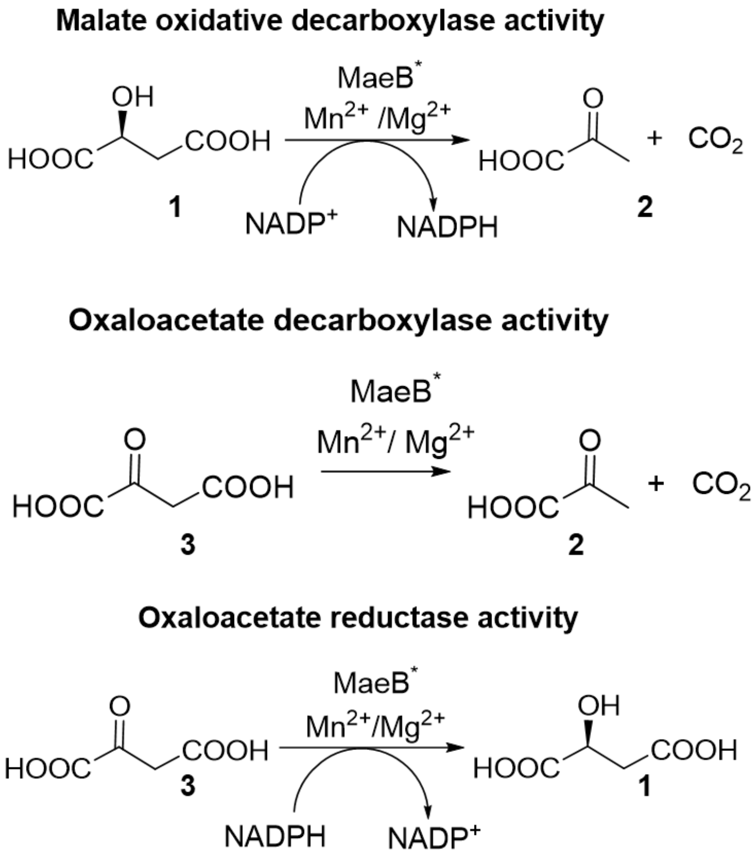


Figure 3.2. Structural features of homohexameric NAD(P)$^+$ -dependent MaeB from *B. bacteriovorus* HD100 showing malic enzyme (ME) and phosphotransacetylase domains. Left and right images represent 90 ° rotation along the y-axis counterclockwise and clockwise. PDB ID 6ZNJ.


Malate oxidative decarboxylase activity

Oxaloacetate decarboxylase activity

Oxaloacetate reductase activity

Figure 3.3. Enzyme assays for determining kinetic parameters for the three type of reactions catalyzed by MaeB variants (MaeB*).

In this chapter, the impact of ALE-induced mutations (K149A, R151I, and N155D) on the malate oxidative decarboxylase, oxaloacetate decarboxylase, and oxaloacetate reductase activities of MaeB was investigated (Figure 3.3). The kinetic parameters, including the Michaelis constant (K_m), maximum velocity (V_{max}), and catalytic turnover number (k_{cat}), were determined for the substrates malate and oxaloacetate, as well as the cofactor $NADP^+$ in the presence of divalent metal activator Mn^{2+}. Significant differences in the kinetics of the reactions across variants will provide insights into the roles of these mutants in ensuring a more efficient PPO node or malate shunt. Cofactor specificity plays a critical role in regulating metabolic processes to achieve cellular homeostasis. Previous reports have demonstrated that even single mutations from ALE experiments can alter the cofactor specificity of an enzyme.[15] This previous report used ALE to induce NADP$^+$ cofactor utilization by MaeA in an engineered NADPH-auxotrophic *E. coli* strain. The strain was deficient in four of the five genes known to support NADPH regeneration in *E. coli*, with the exception of 6-phosphogluconate dehydrogenase (Gnd). The genes mentioned above encode glucose 6-phosphate dehydrogenase (Zwf), NADP$^+$-dependent malic enzyme (MaeB), isocitrate dehydrogenase (Icd), and membrane-bound transhydrogenase (PntAB). These findings indicate that ALE may induce mutations that could result in different cofactor or metal specificities, particularly when the NAD(P)$^+$ /NAD(P)H ratio is altered, as observed during growth on ADCA.

3.2 Determination of kinetic parameters

The *E. coli maeB* gene was amplified by PCR from the genome of W3110 strain and cloned into a pET-15b vector to generate pBKA2.202 (see Chapter Five: Experimental, section 5.9.2). This vector enables the induction of protein overexpression by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG). Plasmid pBKA2.202 was used as a template to generate plasmids for K149A (pBKA2.204), R151I (pBKA2.205), and N155D (pBKA2.206) using the Q5 site-directed mutagenesis (see Chapter Five: Experimental, section 5.9.3). The sequences of the original and derivative plasmids were confirmed by Sanger sequencing. The confirmed plasmids were transformed into the expression host BL21(DE3), and overexpressed MaeB and variants were purified using a Ni^{2+} His-trap column via FPLC (see Chapter Five: Experimental, section 5.9.4) SDS-PAGE of MaeB wt and K149A variant is shown in Figure 3.4. All other purified variants showed a similar band at 83 kDa. All enzyme assays were conducted at 25 °C in 100 mM Tris-HCl (pH 8). The initial velocities of MaeB wild-type (wt) and variants were determined as a function of (*S*)-malate, oxaloacetate, and NADP$^+$ concentrations for their respective reactions, malate oxidative decarboxylase, oxaloacetate decarboxylase, and oxaloacetate reductase activities by fitting initial rates to the Hill equation (Figure 3.5).

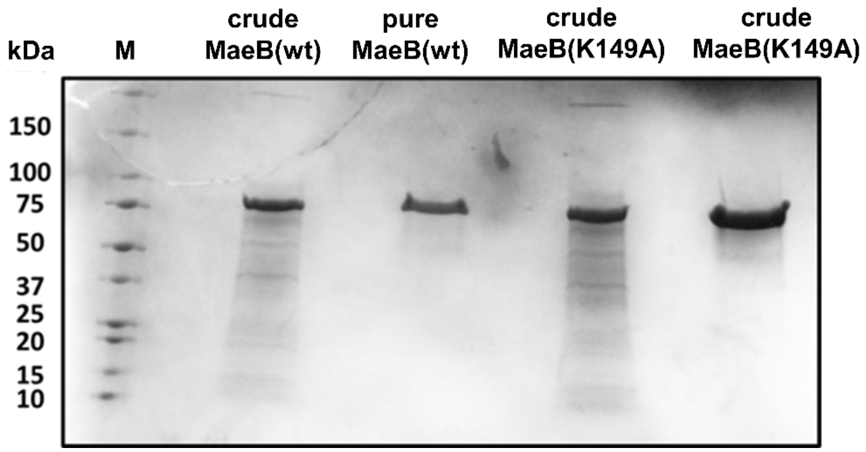
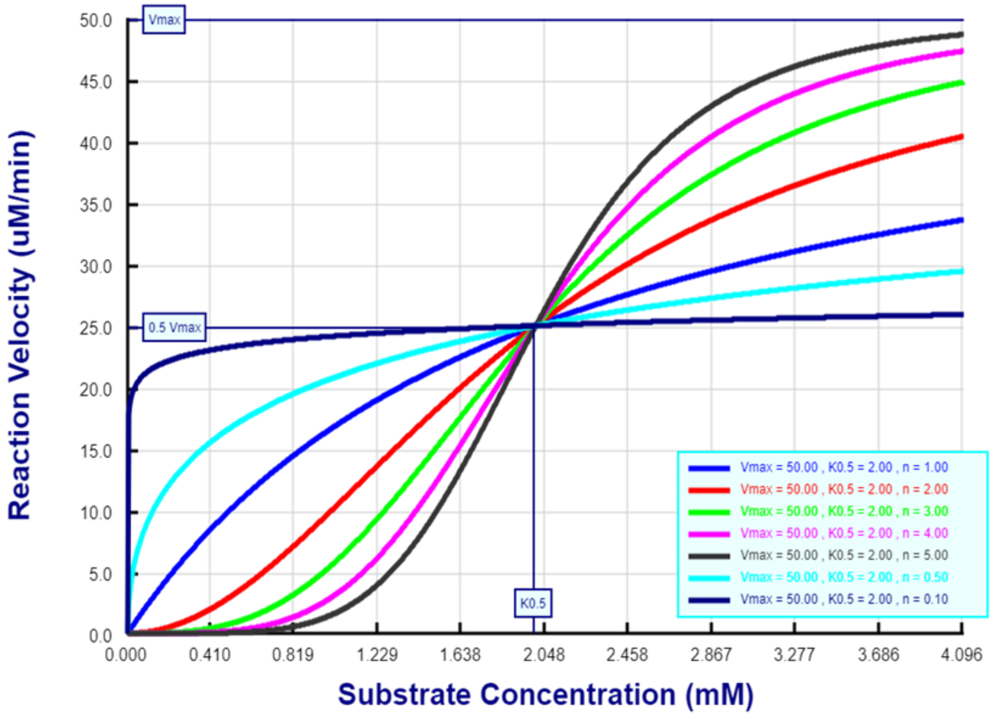


Figure 3.4. SDS-PAGE analysis of crude and purified MaeB. The gel shows the separation of MaeB before and after purification. Lane 1: Molecular weight marker (M); Lane 2: crude MaeB extract from BL21(DE3)/pBKA2.202; Lane 3: purified MaeB; Lane 4: crude MaeB (K149A) extract; Lane 5: purified MaeB K149A. The darkest band on the gel between 75 and 100 kDa corresponds to MaeB.

3.2.1 Malate oxidative decarboxylase activity

MaeB wt and three ALE-induced variants, K149A, R151I, and N155D, were investigated for their oxidative decarboxylase activities for conversion of (*S*)-malate to pyruvate and CO_2 with concomitant reduction of $NADP^+$ (Fig 3.3). This reaction requires the presence of a divalent metal ions, Mg^{2+} or Mn^{2+}.[16] Initial velocities of enzyme assays across varying malate or $NADP^+$ concentrations were fitted to the non-linear Hill equation to determine kinetic parameters (Figure 3.5 – 3.9). Table 3.1 shows the kinetic characteristics of MaeB wt, K149A, R151I, and N155D.

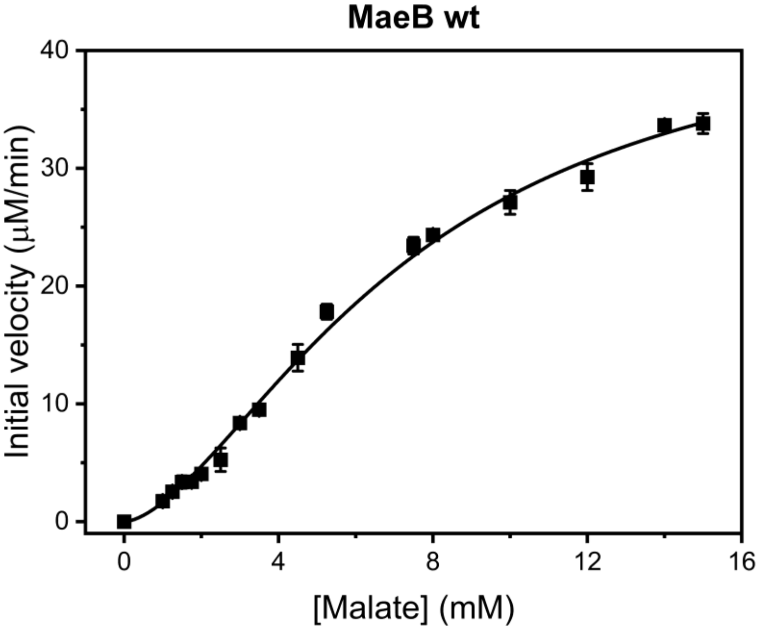


Hill equation:

$$V = \frac{V_{max}[S]^n}{(K_{0.5})^n + [S]^n}$$

V	is the reaction velocity
$V_{\max}$	is the maximum velocity of the reaction
$[S]$	is the substrate concentration
$K_{0.5}$	is the half-maximal concentration

Figure 3.5. Hill equation and simulated plots for various Hill coefficient (n). The plot (top) was obtained from PhysiologyWeb.



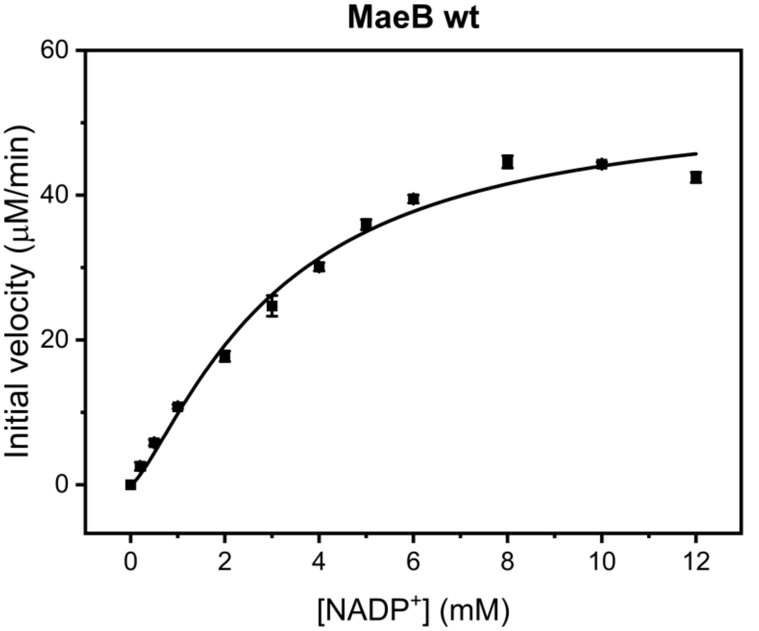
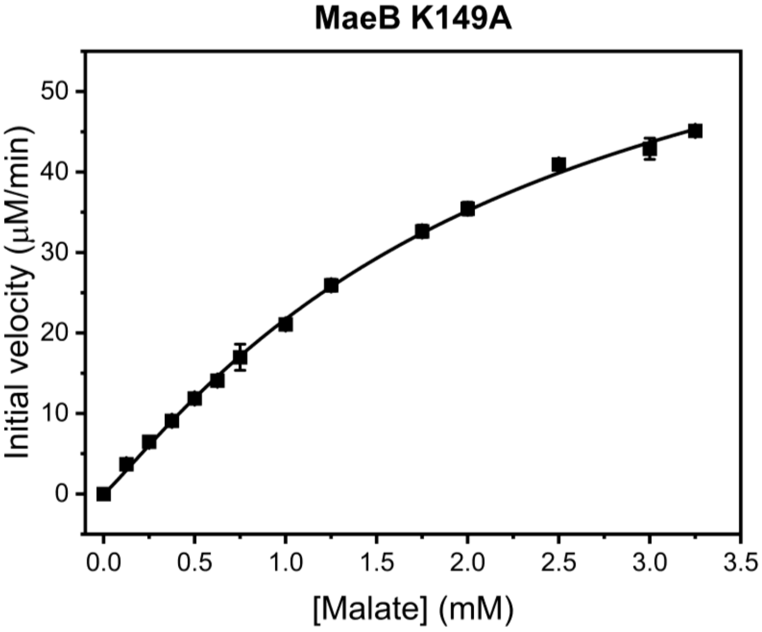


Figure 3.6. Hill plot for oxidative decarboxylase activity of MaeB wt with varying concentrations of malate (top) and NADP$^+$ (bottom). The assay was run on a plate reader using 60 nM of MaeB in 0.1 M Tris-HCl buffer (pH 8). Absorption at 340 nm was monitored, which corresponds to conversion of NADP$^+$ to NADPH. Initial velocities for each concentration were fitted to the Hill equation to determine kinetic parameters. Error bars indicate standard errors for three measurements.



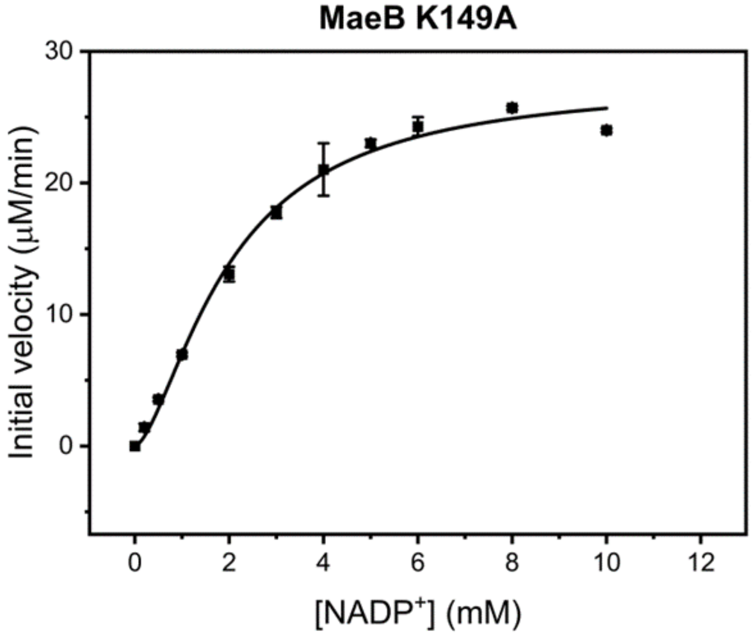


Figure 3.7. Hill plot for oxidative decarboxylase activity of MaeB K149A with varying concentrations of malate (top) and NADP$^+$ (bottom). The assay was run on a plate reader using 290 nM (for malate assays) or 60 nM (for NADP$^+$ assays) of MaeB K149A in 0.1 M Tris-HCl buffer (pH 8). Absorption at 340 nm was monitored, which corresponds to conversion of NADP$^+$ to NADPH. Initial velocities for each concentration were fitted to the Hill equation to determine kinetic parameters. Error bars indicate standard errors for three measurements.

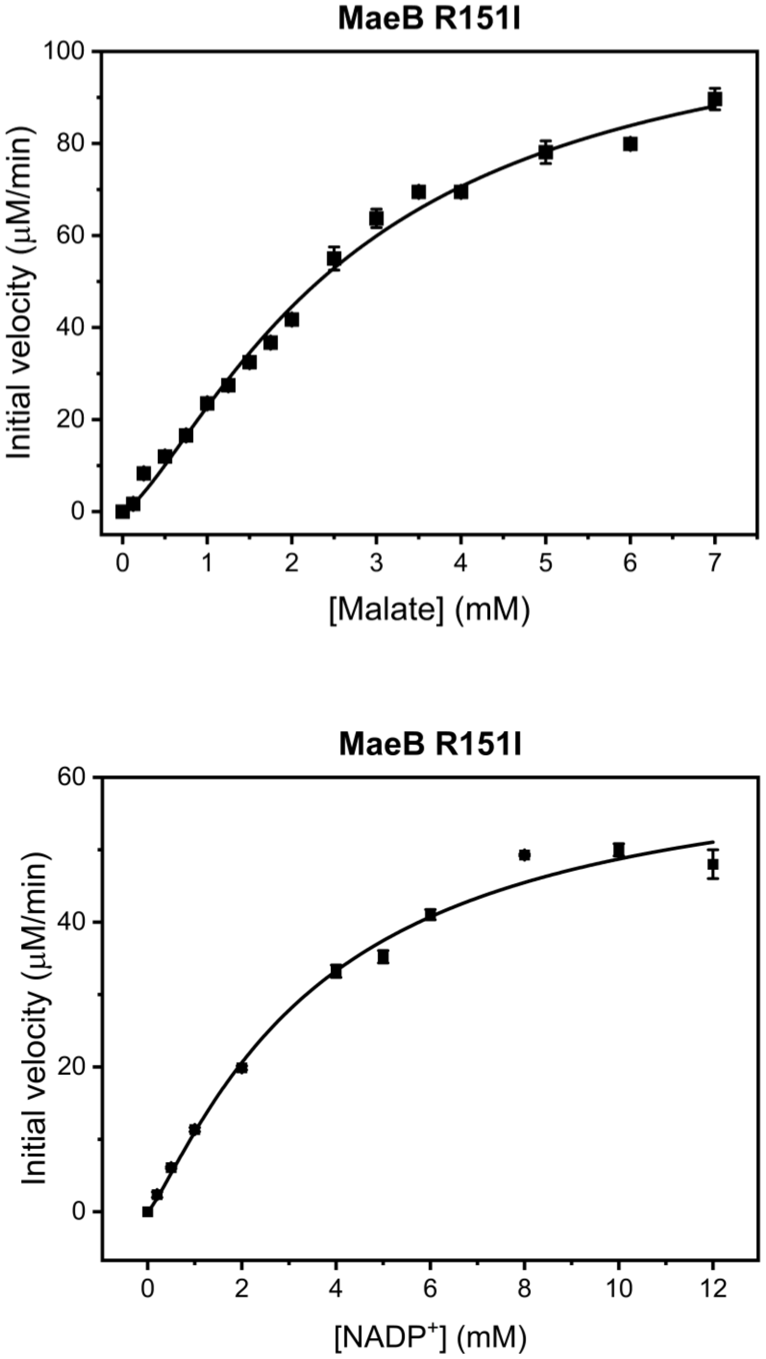


Figure 3.8. Hill plot for oxidative decarboxylase activity of MaeB R151I with varying concentrations of malate (top) and NADP$^+$ (down). The assay was run on a plate reader using 150 nM (for malate assays) or 60 nM (for NADP$^+$ assays) of MaeB R151I in 0.1 M Tris-HCl buffer (pH 8). The reduction of NADP$^+$ was monitored at 340 nm. Initial velocities for each concentration were fitted to the Hill equation to determine kinetic parameters. Error bars indicate standard errors for three measurements.

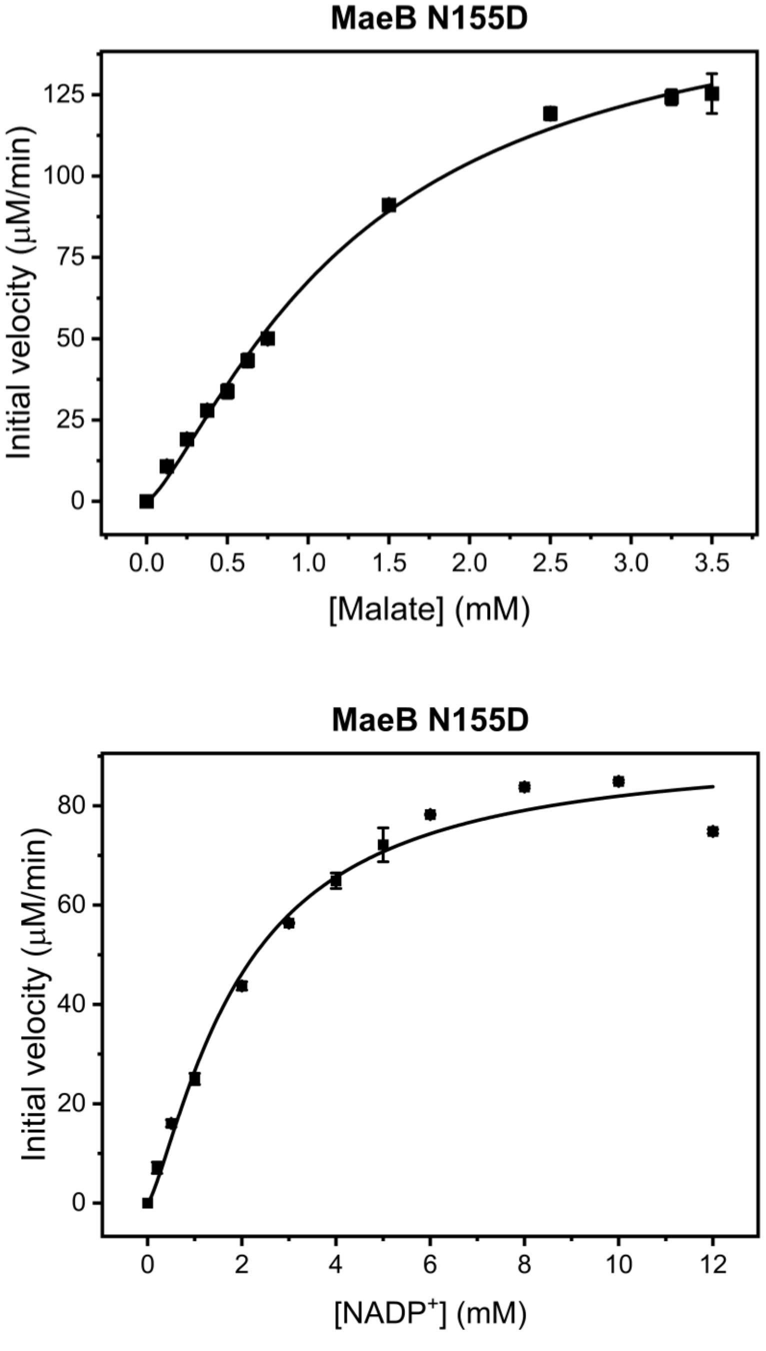
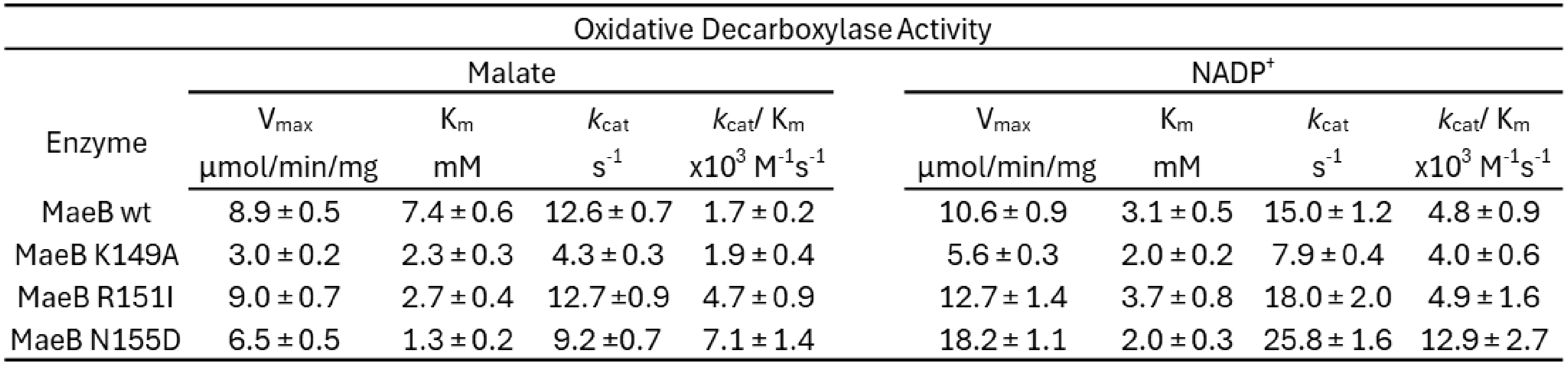
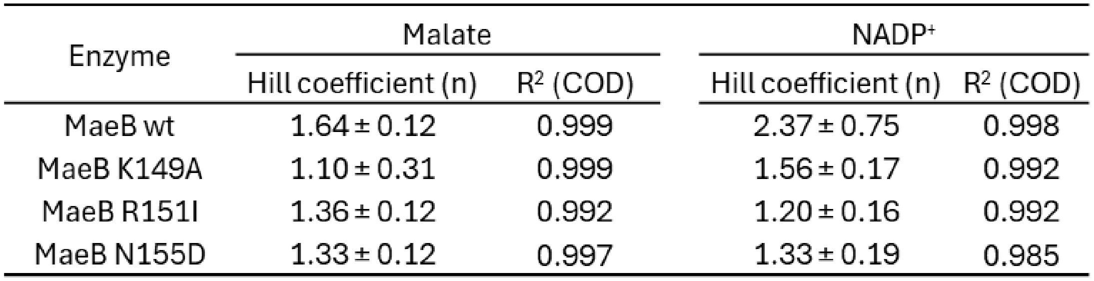


Figure 3.9. Hill plot for oxidative decarboxylase activity of MaeBN155D with varying concentrations of malate (top) and NADP$^+$ (down). The assay was run on a plate reader using 295 nM (for malate assays) or 60 nM (for NADP$^+$ assays) of MaeBN155D in 0.1 M Tris-HCl buffer (pH 8). The reduction of NADP$^+$ was monitored at 340 nm. Initial velocities for each concentration were fitted to the Hill equation to determine kinetic parameters. Error bars indicate standard errors for three measurements.

Table 3.1. Kinetic parameters of MaeB wt and variants for oxidative decarboxylase activity. Bottom table shows statistical values for kinetic plots.



Oxidative Decarboxylase Activity								
Enzyme	Malate				$NADP^+$			
	V_{max} µmol/min/mg	K_m mM	k_{cat} s^{-1}	k_{cat}/K_m $x10^3$ $M^{-1}s^{-1}$	V_{max} µmol/min/mg	K_m mM	k_{cat} s^{-1}	k_{cat}/K_m $x10^3$ $M^{-1}s^{-1}$
MaeB wt	8.9 ± 0.5	7.4 ± 0.6	12.6 ± 0.7	1.7 ± 0.2	10.6 ± 0.9	3.1 ± 0.5	15.0 ± 1.2	4.8 ± 0.9
MaeB K149A	3.0 ± 0.2	2.3 ± 0.3	4.3 ± 0.3	1.9 ± 0.4	5.6 ± 0.3	2.0 ± 0.2	7.9 ± 0.4	4.0 ± 0.6
MaeB R151I	9.0 ± 0.7	2.7 ± 0.4	12.7 ±0.9	4.7 ± 0.9	12.7 ± 1.4	3.7 ± 0.8	18.0 ± 2.0	4.9 ± 1.6
MaeB N155D	6.5 ± 0.5	1.3 ± 0.2	9.2 ±0.7	7.1 ± 1.4	18.2 ± 1.1	2.0 ± 0.3	25.8 ± 1.6	12.9 ± 2.7



Enzyme	Malate		$NADP^+$	
	Hill coefficient (n)	R^2 (COD)	Hill coefficient (n)	R^2 (COD)
MaeB wt	1.64 ± 0.12	0.999	2.37 ± 0.75	0.998
MaeB K149A	1.10 ± 0.31	0.999	1.56 ± 0.17	0.992
MaeB R151I	1.36 ± 0.12	0.992	1.20 ± 0.16	0.992
MaeB N155D	1.33 ± 0.12	0.997	1.33 ± 0.19	0.985

The wild-type (wt) MaeB exhibited the highest K_m value of 7.4 mM for malate, indicating a lower affinity for malate compared to its variants. In contrast, the N155D variant showed significantly enhanced malate binding, as shown by a much lower K_m of 1.3 mM. The other variants, K149A and R151I, also showed improved K_m values of 2.3 mM and 2.7 mM, respectively. For the catalytic turnover rate (k_{cat}), both the wild-type and the R151I variant showed the highest values, around 12.7 s^{-1}. Conversely, the K149A variant showed a much lower k_{cat} of 4.3 s^{-1}. When comparing catalytic efficiencies (k_{cat}/K_m), the wild-type and K149A variants showed similar efficiencies with values of 1.7×10^3 $M^{-1}s^{-1}$ and 1.9×10^3 $M^{-1}s^{-1}$, respectively. However, a significant improvement was observed with the N155D variant, which showed a catalytic efficiency of 7.1×10^3 $M^{-1}s^{-1}$.

The kinetics with varying $NADP^+$ concentrations showed that the K_m values of the wild-type and the variants were not significantly different, although both the K149A and N155D variants showed slightly better K_m values at 2.0 mM. Notably, the N155D variant showed not only an improvement in K_m but also an enhanced catalytic efficiency in the presence of NADP+, highlighting its potential as a more effective enzyme variant for biochemical applications.

A previous study by Federico et al. (2007) provides a comparative analysis of the oxidative decarboxylase activity of MaeB wt under similar assay conditions.[12] Their research involved the use of MaeB without a His-tag in a 100 mM Tris-HCl buffer at pH 7.5, which is similar to the reaction conditions used in our study, albeit with differences in protein concentration and pH values. Although Federico et al. did not report the k_{cat} value for malate, they did report a higher catalytic efficiency for $NADP^+$, calculated to be 1.6×10^6 $M^{-1}s^{-1}$, with a slightly better K_m value for malate at 3.41 mM. The Hill coefficients they reported were similar to what we obtained in our analysis.

3.2.2 Oxaloacetate decarboxylase activity

The decarboxylating activity of MaeB and variants K149A, R151I, and N155D, was investigated with regard to their ability to catalyze the conversion of oxaloacetate to pyruvate and CO_2. This reaction proceeds without the need for cofactors such as $NADP^+$ or NADPH, although it did require the presence of either Mn^{2+} or Mg^{2+}. To monitor the progress of the reaction, the consumption of oxaloacetate was measured at 275 nm using a plate reader. To analyze the enzyme kinetics, the initial rates observed at varying concentrations of oxaloacetate were fitted to the Hill equation (Figure 3.10 - 3.13). The results of these kinetic analyses are summarized in Table 3.2.2.1. 19.2 ± 3.3 mM

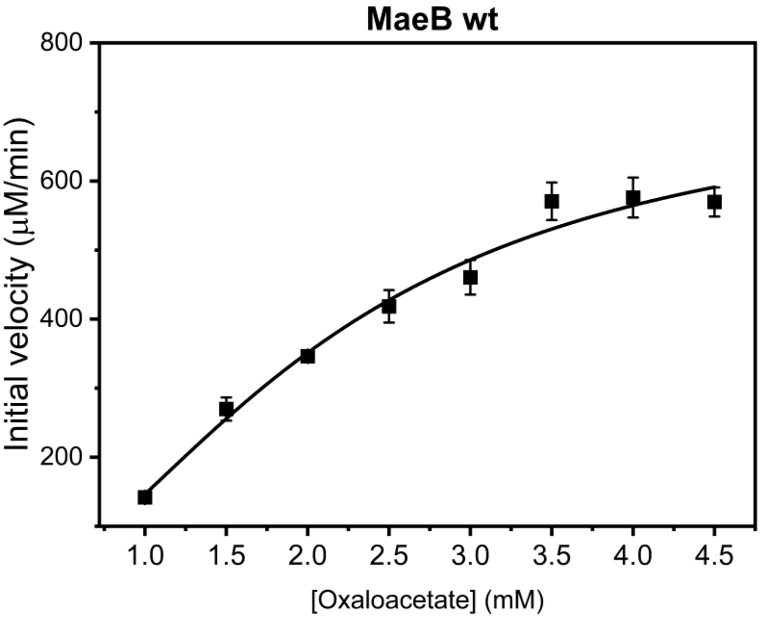


Figure 3.10. Hill plot for decarboxylase activity of MaeB wt with varying concentrations of oxaloacetate. The assay was run on a plate reader using 6 μM of MaeB wt in 0.1 M Tris-HCl buffer (pH 8). The consumption of oxaloacetate was monitored at 275 nm. Initial velocities for each concentration were fitted to the Hill equation to determine kinetic parameters. Error bars indicate standard errors for three measurements.

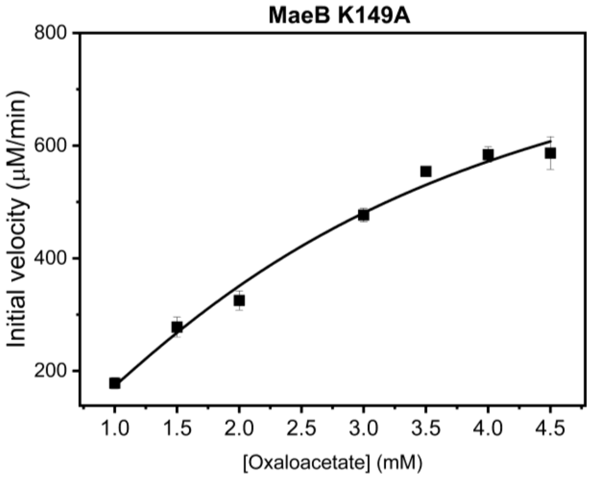


Figure 3.11. Hill plot for decarboxylase activity of MaeB K149A with varying concentrations of oxaloacetate. The assay was run on a plate reader using 6 μM of MaeB K149A in 0.1 M Tris-HCl buffer (pH 8). The consumption of oxaloacetate was monitored at 275 nm. Initial velocities for each concentration were fitted to the Hill equation to determine kinetic parameters. Error bars indicate standard errors for three measurements.

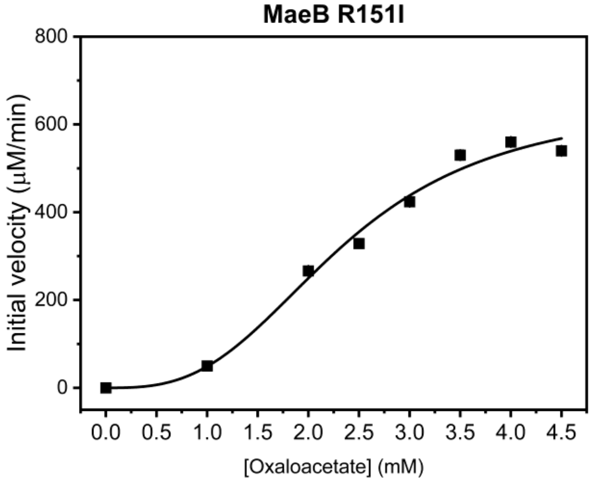


Figure 3.12. Hill plot for decarboxylase activity of MaeB R151I with varying concentrations of oxaloacetate. The assay was run on a plate reader using 6 μM of MaeB R151I in 0.1 M Tris-HCl buffer (pH 8). The consumption of oxaloacetate was monitored at 275 nm. Initial velocities for each concentration were fitted to the Hill equation to determine kinetic parameters. Error bars indicate standard errors for three measurements.

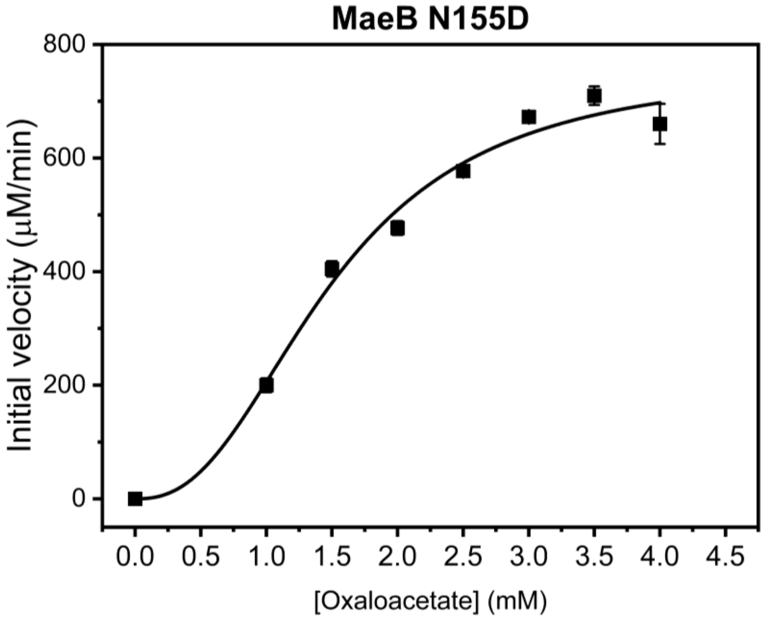
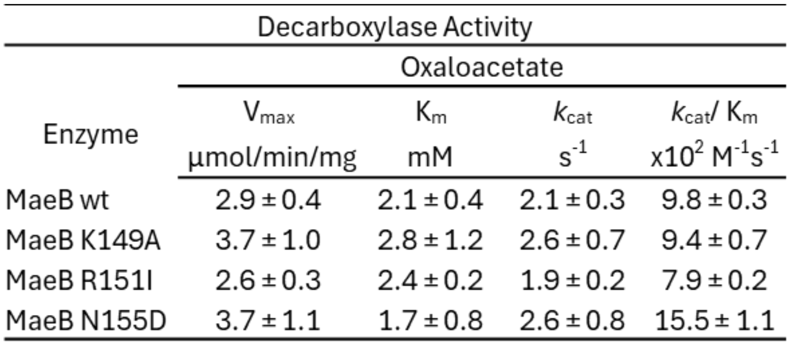


Figure 3.14. Hill plot for decarboxylase activity of MaeBN155D with varying concentrations of oxaloacetate. The assay was run on a plate reader using 6 µM of MaeBN155D in 0.1 M Tris-HCl buffer (pH 8). The consumption of oxaloacetate was monitored at 275 nm. Initial velocities for each concentration were fitted to the Hill equation to determine kinetic parameters. Error bars indicate standard errors for three measurements.

Table 3.2. Kinetic parameters for decarboxylase activity of MaeB wt and variants



Decarboxylase Activity				
	Oxaloacetate			
Enzyme	V_{max} µmol/min/mg	K_m mM	k_{cat} s^{-1}	k_{cat}/ K_m x10^2 $M^{-1}s^{-1}$
MaeB wt	2.9 ± 0.4	2.1 ± 0.4	2.1 ± 0.3	9.8 ± 0.3
MaeB K149A	3.7 ± 1.0	2.8 ± 1.2	2.6 ± 0.7	9.4 ± 0.7
MaeB R151I	2.6 ± 0.3	2.4 ± 0.2	1.9 ± 0.2	7.9 ± 0.2
MaeB N155D	3.7 ± 1.1	1.7 ± 0.8	2.6 ± 0.8	15.5 ± 1.1

The analysis of oxaloacetate decarboxylation kinetics by wild-type MaeB, K149A, and R151I revealed that these exhibited similar kinetic profiles. Specifically, the K_m values recorded were 2.1 mM for MaeB wt, 2.8 mM for K149A, and 2.4 mM for R151I, while their catalytic efficiencies were noted as 9.8 x 10^2 $M^{-1}s^{-1}$, 9.4 x 10^2 $M^{-1}s^{-1}$, and 7.9 x 10^2 $M^{-1}s^{-1}$, respectively. In contrast, the

N155D variant demonstrated improved catalytic efficiency of 15.5 x 10^2 $M^{-1}s^{-1}$. It is noteworthy that the kinetic parameters for oxaloacetate decarboxylation by MaeB have not been previously reported in the literature.

3.2.3 Oxaloacetate reductase activity

The NADP-dependent malic enzymes, like MaeB, have been predicted to exhibit reductase activity, specifically in the reduction of oxaloacetate to (*S*)-malate in the presence of NADPH.[12] We investigated the ability of wild-type MaeB, K149A, R151I, and N155D variants of MaeB to catalyze this reaction. The progress of the reaction was monitored by measuring the rate of NADPH oxidation at 340 nm. The specific activities of the enzymes were determined using the rates obtained from oxaloacetate-saturated reactions. Notably, all variants demonstrated an improvement in specific activity for this conversion, with the N155D variant exhibiting the highest activity (Figure 3.14). This suggests that the mutations may have led to an improvement in their enzymatic efficiency for the reduction of oxaloacetate to (*S*)-malate

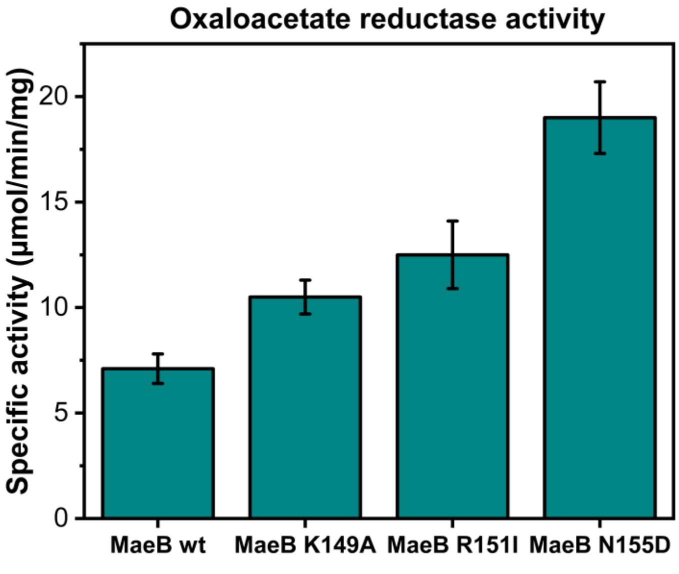


Figure 3.14. Specific activities of MaeB wt and variants for oxaloacetate reductase activity. The assay was run on a plate reader using 6 µM of MaeB variants in 0.1 M Tris-HCl buffer (pH 8). The consumption of oxaloacetate was monitored at 275 nm. Error bars indicate standard errors for five measurements.

3.3 Structural basis for improved kinetics of MaeBN155D

The lack of a crystal structure for *E. coli* MaeB presents a challenge for structural analysis of roles of active site residues and determination of the mechanism of all three kinds of reactions investigated. A predictive approach was employed to address this limitation. The predicted structure of *E. coli* MaeB was obtained using AlphaFold, an AI-driven protein structure prediction tool. The predicted model was evaluated for accuracy by comparing its predicted secondary elements and overall fold with those of closely related enzymes whose structures have been resolved experimentally. Docking analysis was conducted with the predicted MaeB structure to explore the interactions between MaeB and its substrates: malate, oxaloacetate, and NADPH. A flexible blind docking approach was applied using AutoDock Vina, which allows for small conformational changes within the active site during the docking process. The objective of the docking simulations was to elucidate the most probable binding poses and to estimate the binding affinities associated with each substrate.

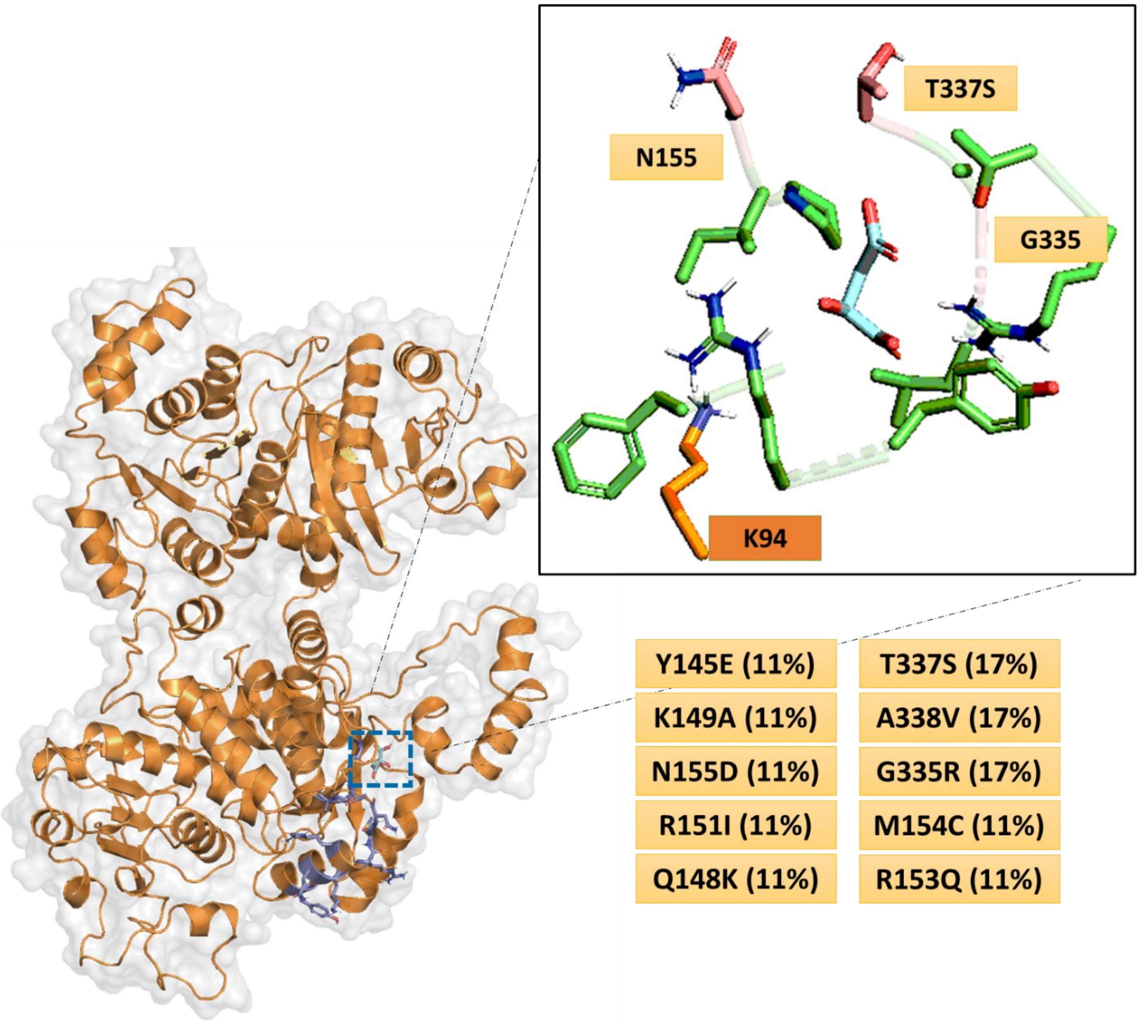


Figure 3.15. Predicted Binding Interactions of MaeB with Malate. These structures show a docking pose of malate within the active site of MaeB, as predicted by AlphaFold2 and AutoDock Vina. The residues in the propose malate binding pocket are zoomed in and shown in green. The predicted residue involve in binding to substrate, K94, is shown in brown. Mutated residues that are closest to malate binding site (N155D, G335R, and T337S) are shown in pink. All other mutated residues are shown in the full structure in violet. A list of mutations with their frequencies are shown in the bottom right. The structure of MaeB was modeled using AlphaFold2, and docking was performed using AutoDock Vina.

We observed that the observed mutations were localized at or around the proposed substrate binding site. Mainly, N155D, G335R, and T337S were within 5 Å of the bound malate (Figure 3.15). This placement could explain the improved N155D binding affinities for malate and their catalytic efficiencies for all three kinds of reactions investigated. It is unclear why a switch from asparagine to aspartic acid significantly improved these reactions. Future studies shall focus on investigating these residues closest to the proposed malate binding site. The other mutations around the site might be playing a role in ushering the substrates to the binding pocket. It will be interesting to study the effect of combining some or all these mutations on the malate oxidative decarboxylase, oxaloacetate decarboxylase, and reductase activities.

3.4 Conclusion

The characterization of *E. coli* MaeB and ALE-induced variants (K149A, R151I, and N155D) provided significant insights into the enzyme's catalytic properties. These findings are the first reports of kinetic parameters for the oxaloacetate decarboxylating activity of *E. coli* MaeB and the second report of the malate oxidative decarboxylation activity. The results demonstrated that MaeB wt and the variants K149A and R151I exhibited similar kinetics in the decarboxylation of oxaloacetate to pyruvate, with modest differences in K_m and catalytic efficiencies. The N155D variant exhibited markedly improved binding affinities and an increase in catalytic efficiency for all three kinds of reactions investigated, which may be due to the structural placement of the residue. It is still unclear how a swap from asparagine to aspartic acid might affect the mechanism of the reactions at the active site.

The study of the reductase activity of these enzymes, particularly the conversion of oxaloacetate to (*S*)-malate, revealed that all variants exhibited enhanced specific activities compared to the wild-type, with N155D once again demonstrating the highest activity.

The enhanced activities observed in the variants, especially N155D, provide a promising basis for further development of enzymatic processes that require better efficiency in the catalysis of specific reactions within metabolic pathways like the production of lactic acid from ADCA, which proceeds via oxaloacetate-pyruvate route as discussed in Chapter Four. It will be interesting to investigate the effect of combining the mutants. Further studies could be conducted to include other MaeB mutations observed in the ALE experiment.

REFERENCES

(1) He, L.; Groom, J. D.; Lidstrom, M. E. The Entner-Doudoroff Pathway Is an Essential Metabolic Route for *Methylotuvimicrobium buryatense* 5GB1C. *Appl. Environ. Microbiol.* **2021**, *87* (3). https://doi.org/10.1128/AEM.02481-20.

(2) Peretó, J. Embden-Meyerhof-Parnas Pathway. In *Encyclopedia of Astrobiology*; Gargaud, M., Amils, R., Quintanilla, J. C., Cleaves, H. J. (jim), Irvine, W. M., Pinti, D. L., Viso, M., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; pp 485–485.

(3) Peretó, J. Entner–Doudoroff Pathway. In *Encyclopedia of Astrobiology*; Gargaud, M., Amils, R., Quintanilla, J. C., Cleaves, H. J. (jim), Irvine, W. M., Pinti, D. L., Viso, M., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; pp 498–498.

(4) Sauer, U.; Eikmanns, B. J. The PEP-Pyruvate-Oxaloacetate Node as the Switch Point for Carbon Flux Distribution in Bacteria. *FEMS Microbiol. Rev.* **2005**, *29* (4), 765–794.

(5) The Editors of Encyclopedia Britannica. Tricarboxylic Acid Cycle. *Encyclopedia Britannica*; 2024.

(6) Koendjbiharie, J. G.; van Kranenburg, R.; Kengen, S. W. M. The PEP-Pyruvate-Oxaloacetate Node: Variation at the Heart of Metabolism. *FEMS Microbiol. Rev.* **2021**, *45* (3). https://doi.org/10.1093/femsre/fuaa061.

(7) Peretó, J. Gluconeogenesis. In *Encyclopedia of Astrobiology*; Gargaud, M., Amils, R., Quintanilla, J. C., Cleaves, H. J. (jim), Irvine, W. M., Pinti, D. L., Viso, M., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; pp 680–680.

(8) Bowman, E.; McQueney, M.; Barry, R. J.; Dunaway-Mariano, D. Catalysis and Thermodynamics of the Phosphoenolpyruvate/Phosphonopyruvate Rearrangement. Entry into the Phosphonate Class of Naturally Occurring Organophosphorus Compounds. *J. Am. Chem. Soc.* **1988**, *110* (16), 5575–5576.

(9) O'Daly, A. Encyclopedia of Life Sciences. *(No Title)* **1996**.

(10) Park, J. H.; Lee, S. Y. Metabolic Pathways and Fermentative Production of L-Aspartate Family Amino Acids. *Biotechnol. J.* **2010**, *5* (6), 560–577.

(11) Takeo, K. Existence and Properties of Two Malic Enzymes in *Escherichia coli* Especially of NAD-Linked Enzyme. *J. Biochem.* **1969**, *66* (3), 379–387.

(12) Bologna, F. P.; Andreo, C. S.; Drincovich, M. F. *Escherichia coli* Malic Enzymes: Two Isoforms with Substantial Differences in Kinetic Properties, Metabolic Regulation, and Structure. *J. Bacteriol.* **2007**, *189* (16), 5937–5946.

(13) Harding, C. J.; Cadby, I. T.; Moynihan, P. J.; Lovering, A. L. A Rotary Mechanism for Allostery in Bacterial Hybrid Malic Enzymes. *Nat. Commun.* **2021**, *12* (1), 1228.

(14) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska,

M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589.

(15) Bouzon, M.; Döring, V.; Dubois, I.; Berger, A.; Stoffel, G. M. M.; Calzadiaz Ramirez, L.; Meyer, S. N.; Fouré, M.; Roche, D.; Perret, A.; Erb, T. J.; Bar-Even, A.; Lindner, S. N. Change in Cofactor Specificity of Oxidoreductases by Adaptive Evolution of an *Escherichia coli* NADPH-Auxotrophic Strain. *MBio* **2021**, *12* (4), e0032921.

(16) Chang, G.-G.; Tong, L. Structure and Function of Malic Enzymes, a New Class of Oxidative Decarboxylases. *Biochemistry* **2003**, *42* (44), 12721–12733.

CHAPTER FOUR: Bioconversion of acetylenedicarboxylate to chemicals

4.1 Introduction

Developing strategies that are both sustainable and decarbonizing is crucial to addressing challenges like carbon dioxide (CO_2) associated with industrial chemical production from fossil fuels. Such strategies would focus on exploiting abundant and renewable feedstocks while minimizing greenhouse gas emissions and environmental impacts.[1–3] Various global initiatives, including the U.S Department of Energy's Carbon Dioxide Removal (CDR) program, aim to advance diverse technologies to enable the removal of CO_2 on the gigatons scale per year by 2050.[4] One promising approach is using alternative feedstocks such as methane and carbon dioxide for large scale production of commodity chemicals.[5,6] CH_4 and CO_2 are two greenhouse gases produced from various sources, including biogas, landfill gas, and industrial emissions. Methane, the primary component of natural gas, is available in vast reserves worldwide and is also produced as a product of anaerobic digestion.[7,8] Utilizing these gases as carbon sources for commodity chemical synthesis could help reduce their emissions while providing a sustainable and economically viable circular pathway for their conversion into value-added products.

However, direct utilization of methane and carbon dioxide as carbon and energy sources by microorganisms is challenging due to their limited solubility in aqueous systems.[5] Microbes typically require dissolved substrates for growth and metabolism, and gas utilization requires specific microbial uptake mechanisms and metabolic pathways to integrate intermediates into microbial processes. This is usually a challenge when exploiting unnatural carbon sources. Overcoming these challenges is critical to unlocking the potential of methane and carbon dioxide as sustainable carbon sources for biomanufacturing. One approach to overcome the challenges associated with the direct utilization of methane and carbon dioxide is the conversion of these gases to stable molecules that microbes could utilize, including formate[9], methanol[10], or acetate[11], among others. Methane can be converted to acetylene[12–15] (see Chapter One), followed by carboxylation to produce acetylene monocarboxylate and acetylenedicarboxylate.[16,17] This two-step process provides a potential pathway for converting CH_4 and CO_2 to compounds that can serve as precursors for synthesizing various commodity chemicals, including 3-hydroxypropionic acid[18] and lactic acid (this study).

We have shown that *Escherichia coli*, a thoroughly studied microorganism, can grow on ADCA as the exclusive carbon source with growth rates comparable to growth on glucose after an adaptive laboratory evolution (Chapter Two). This metabolic capability of *E. coli* and other reported microbes opens exciting possibilities for using ADCA, which could be derived from methane and carbon dioxide, as a sustainable feedstock for microbial production of lactic acid and other commodity chemicals.

Lactic acid (2-hydroxypropanoic acid) is a valuable commodity chemical with broad applications in the manufacture of food additives and pharmaceuticals.[19] It is also a precursor for the synthesis of lactide, which is used to produce polylactic acid (PLA), a biodegradable and renewable polymer.[20,21] Other commodity chemicals, such as acetic acid, ethanol, and butanol, are also critical in many industries serving as key raw materials in the production of plastics, solvents, and fuels. Most of these chemicals are traditionally produced from carbohydrate-rich feedstocks derived from crops such as corn and sugar cane.[22–25] While these routes offer advantages, such as high production volumes due to raw material availability and established infrastructure, they also raise sustainability concerns, including using food crops for industrial purposes introducing competition between food and chemical industries.[26,27]

Shikimic acid, an intermediate in the shikimate pathway, is essential for the biosynthesis of aromatic amino acids and secondary metabolites[28–30] and has extensive pharmacological importance, including as precursors for antiviral[31], anticancer[28,32], and anti-inflammatory drugs.[33] Its critical role in the synthesis of the influenza drug oseltamivir (Tamiflu) highlights the need for its sustainable and scalable production without reliance on traditional extraction from plant sources. The bioproduction of shikimic acid and other pharmaceuticals through microbial fermentation has become an effective alternative, offering a sustainable approach that aligns with the principles of green chemistry.[29,34] A seminal report by Draths et al. (1999) demonstrated the development of recombinant *E. coli* capable of producing about 80 g L^{-1} shikimic acid from glucose.[34]

As proof of concept, we demonstrate the bioproduction of lactic and shikimic acid from acetylenedicarboxylate. Taking advantage of the metabolic adaptability of *E. coli*, we develop a pathway to exploit ADCA as an alternative to fossil fuel-based and carbohydrate-rich feedstocks. This work presents acetylenedicarboxylate as a sustainable and decarbonization strategy for chemical synthesis. By utilizing abundant feedstocks derived from methane and carbon dioxide and employing enzymatic and microbial biomanufacturing approaches, we can develop environmentally friendly and economically viable routes to produce a variety of commodity chemicals with broad applications.

4.2 Bioconversion of ADCA to lactate

4.2.1 Proposed route for biosynthesis of lactate

This study utilizes both cell-free enzymatic (*in vitro*) and microbe-catalyzed (*in vivo*) methods to produce lactic acid from ADCA. Both approaches take advantage of enzymes' inherent selectivity and efficiency.

In addition to stereospecific control of enzyme reactions, an additional advantage of using enzymes for chemical synthesis is their cost-effectiveness. Contrary to the perception that biocatalysts might introduce additional expenses to the process, enzymes may present a more economical alternative owing to their catalytic efficiency and selectivity.[35,36] By constructing plasmids which include genes that encode the desired enzymes, it is possible to facilitate their overexpression in suitable host organisms, followed by purification for subsequent applications.

The proposed *in vitro* route for lactate production involves a cascade of crude and purified enzymes to achieve the conversion of ADCA to lactate through a series of steps (Figure 4.1). First, fumarase A (FumA) catalyzes the hydration of ADCA to oxaloacetate, followed by the conversion of oxaloacetate to pyruvate mediated by a variant of malic enzyme B (MaeBN155D).

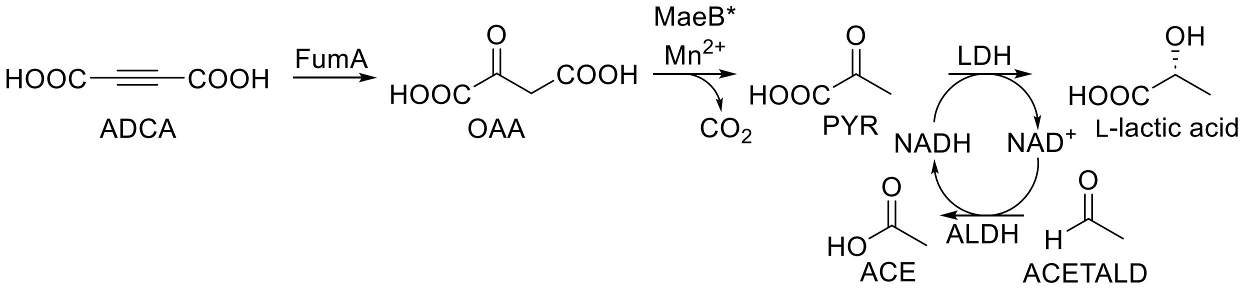


Figure 4.1. Proposed route for *in vitro* biocatalytic synthesis of lactic acid from ADCA. The pathway involves the hydration of ADCA to oxaloacetate (OAA) by crude fumarase A (FumA), followed by decarboxylation of OAA to pyruvate (PYR) by malic enzyme B variant (MaeBN155D). Lactate dehydrogenase (LDH) then reduces PYR to lactic acid with equivalents of NADH cofactor. Aldehyde dehydrogenase (ALDH) was used as a cofactor regeneration system with acetaldehyde as a sacrificial substrate.

The final step involves the reduction of pyruvate to lactate catalyzed by lactate dehydrogenase (LDH), a reaction that requires NADH cofactor. Recognizing the requirement of cofactor availability in maintaining reaction efficiency, we integrated a cofactor recycling system into the biocatalytic route to avoid stoichiometric use of NADH, an expensive cofactor. This system used the enzyme aldehyde dehydrogenase (ALDH) and acetaldehyde as a sacrificial substrate to regenerate NADH, thereby ensuring a continuous supply of the cofactor required for the LDH-mediated reduction step.

4.2.2 Determination of kinetic parameters of LDH and ALDH

To develop an efficient biocatalytic route for lactate production, we determined the kinetic parameters of the enzymes involved in the proposed route. This investigation was essential to understanding the efficiencies of these enzymes in catalyzing their respective reactions. We purified MaeBN155D and used crude FumA for these experiments. Crude FumA was used because previous experiments demonstrated that purified FumA is less stable than the crude extract with overexpressed FumA. LDH and ALDH enzymes were purchased from Sigma Aldrich. The pLBA110 plasmid harboring the *E. coli fumA* under the control of the IPTG-inducible T7 expression system and the pBKA2.202 plasmid encoding *E. coli* MaeB variant (N155D) under the same expression system were used for enzyme overexpression. Previous chapters have discussed the kinetic parameters of FumA (Chapter 2) and MaeBN155D (Chapter 3). In this chapter, we focus on determining kinetic parameters for ALDH and LDH, as these enzymes play critical roles in the final stages of our biocatalytic route. Both assays were conducted in 0.1 M Tris-HCl buffer (pH 8) with varying substrate concentration, using a plate reader (see experimental section) and monitoring consumption or production of NADH at 340 nm (Figure 4.2).

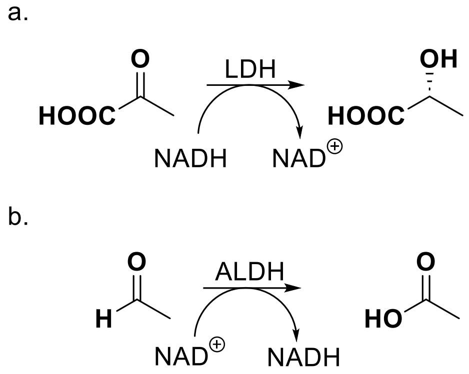


Figure 4.2. Assays for determining kinetic parameters for (a) lactate dehydrogenase (LDH) and (b) aldehyde dehydrogenase (ALDH). Assays were conducted in 0.1 M Tris-HCl buffer (pH 8) using a plate reader to measure the consumption or production of NADH at 340 nm.

The initial velocities at various concentrations of pyruvic acid were fitted to the Hill equation using Origin software to obtain plots (Figure 4.3 – 4.6) and determine kinetic parameters, V_{max} and K_m. Other parameters were calculated using the equation discussed in the corresponding experimental section 5.10.3. For the ALDH-catalyzed reaction with acetaldehyde as substrate, the Michaelis-Menten constant (K_m) was determined to be 0.8 mM, the catalytic turnover number (k_{cat}) was 0.12 s^{-1}, and the specificity constant (k_{cat}/K_m) was 150 M^{-1} s^{-1}. When NAD^+ concentration was varied, K_m was 0.42 mM, k_{cat} was 0.081 s^{-1}, and k_{cat}/K_m reached 192 M^{-1} s^{-1} (Table 4.1).

For LDH assays, with pyruvate as substrate, the K_m was determined to be 0.24 mM, reflecting a high affinity for pyruvate. The enzyme exhibited a remarkable k_{cat} of 1000 s^{-1}, and the k_{cat}/K_m ratio was determined to be 4.2 x 10^6 M^{-1} s^{-1}, indicating a high catalytic efficiency. For NADH substrate, the K_m was low at 0.0023 mM, k_{cat} remained at 1000 s^{-1}, and the k_{cat}/K_m ratio was 4.4 x10^8 M^{-1} s^{-1} (Table 4.1).

The high catalytic efficiencies (k_{cat}/K_m), especially for LDH, suggest that the enzymes are highly efficient in their respective catalytic roles, making them great candidates for incorporation into the biocatalytic pathway. Determination of these parameters was critical for fine-tuning enzyme concentrations and reaction conditions, ultimately leading to improved overall process efficiency. At the time of this report, there have been no documented studies detailing the kinetic parameters of lactate dehydrogenase (LDH) from rabbit muscle or aldehyde dehydrogenase (ALDH) from *Saccharomyces cerevisiae* in the existing literature.

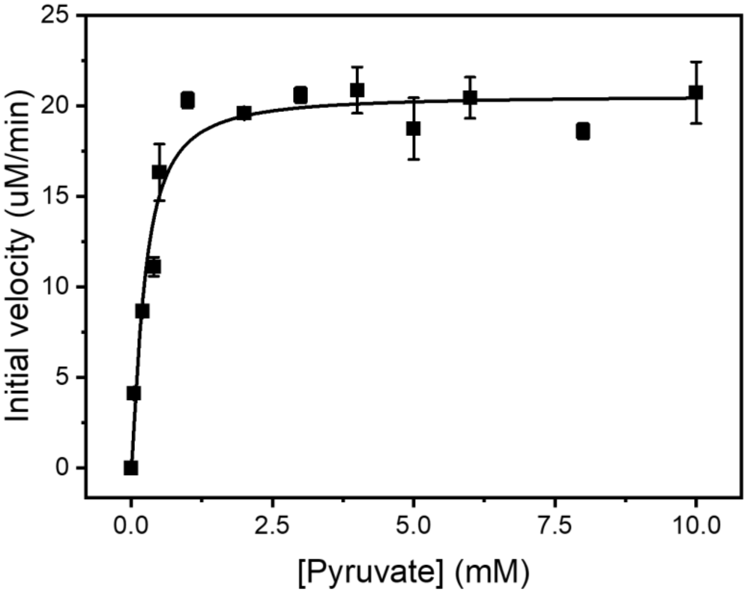


Figure 4.3. Michaelis-Menten plot of LDH with varying pyruvate concentrations. The assay was run on a plate reader using 0.05 µg mL^{-1} of LDH in 0.1 M Tris-HCl buffer (pH 8). Consumption of NADH was monitored at 340 nm. Initial velocities for each pyruvate concentration were fitted to the Hill equation to determine kinetic parameters. Error bars indicate standard errors for three measurements.

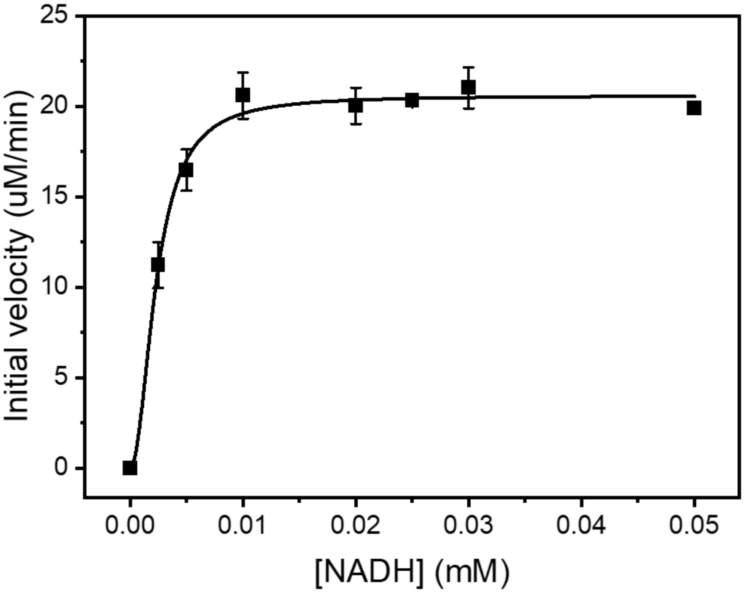


Figure 4.4. Michaelis-Menten plot of LDH with varying NADH concentrations. The assay was run on a plate reader using 0.05 µg mL^{-1} of LDH in 0.1 M Tris-HCl buffer (pH 8). Consumption of NADH was monitored at 340 nm. Initial velocities for each NADH concentration were fitted to the Hill equation to determine kinetic parameters. Error bars indicate standard errors for three measurements.

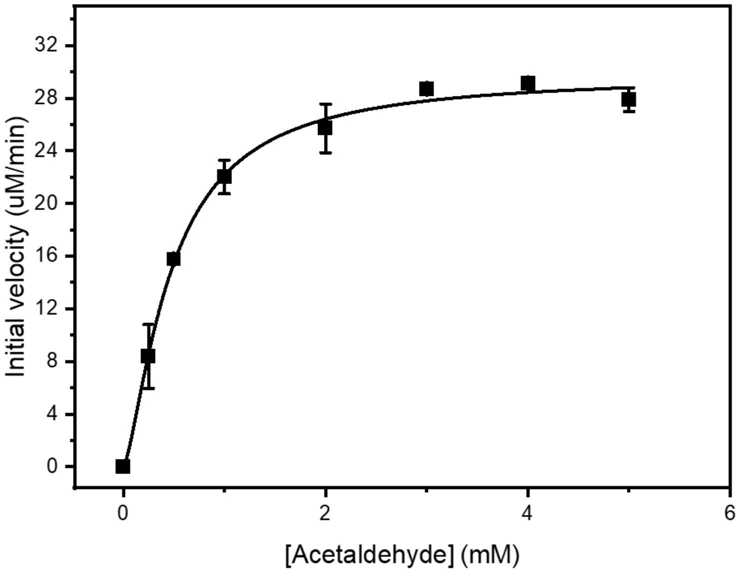


Figure 4.5. Michaelis-Menten plot of ALDH with varying acetaldehyde concentrations. The assay was run on a plate reader using 250 µg mL^{-1} of ALDH in 0.1 M Tris-HCl buffer (pH 8). Production of NADH was monitored at 340 nm. Initial velocities for each acetaldehyde concentration were fitted to the Hill equation to determine kinetic parameters. Error bars indicate standard errors for three measurements.

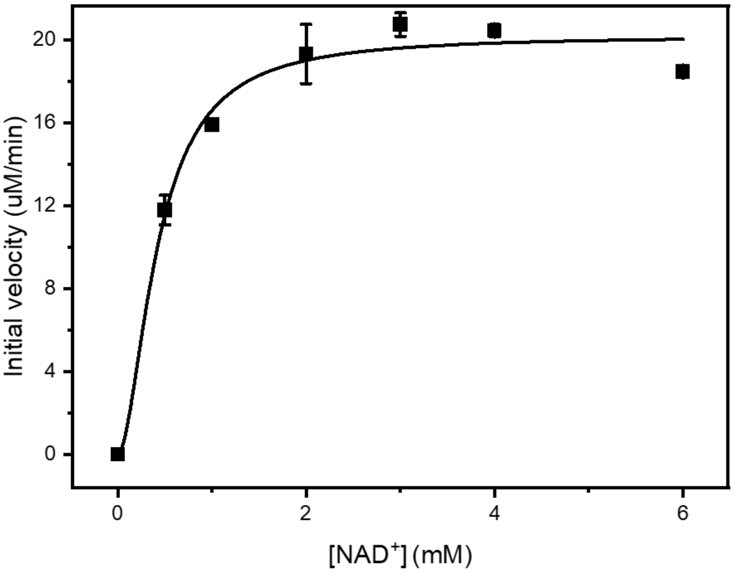


Figure 4.6. Michaelis-Menten plot of ALDH with varying NAD$^+$ concentrations. The assay was run on a plate reader using 250 µg mL^{-1} of ALDH in 0.1 M Tris-HCl buffer (pH 8). Production of NADH was monitored at 340 nm. Initial velocities for each NAD$^+$ concentration were fitted to the Hill equation to determine kinetic parameters. Error bars indicate standard errors for three measurements.

Table 4.1. Kinetic parameters for LDH and ALDH with various substrates and cofactors.

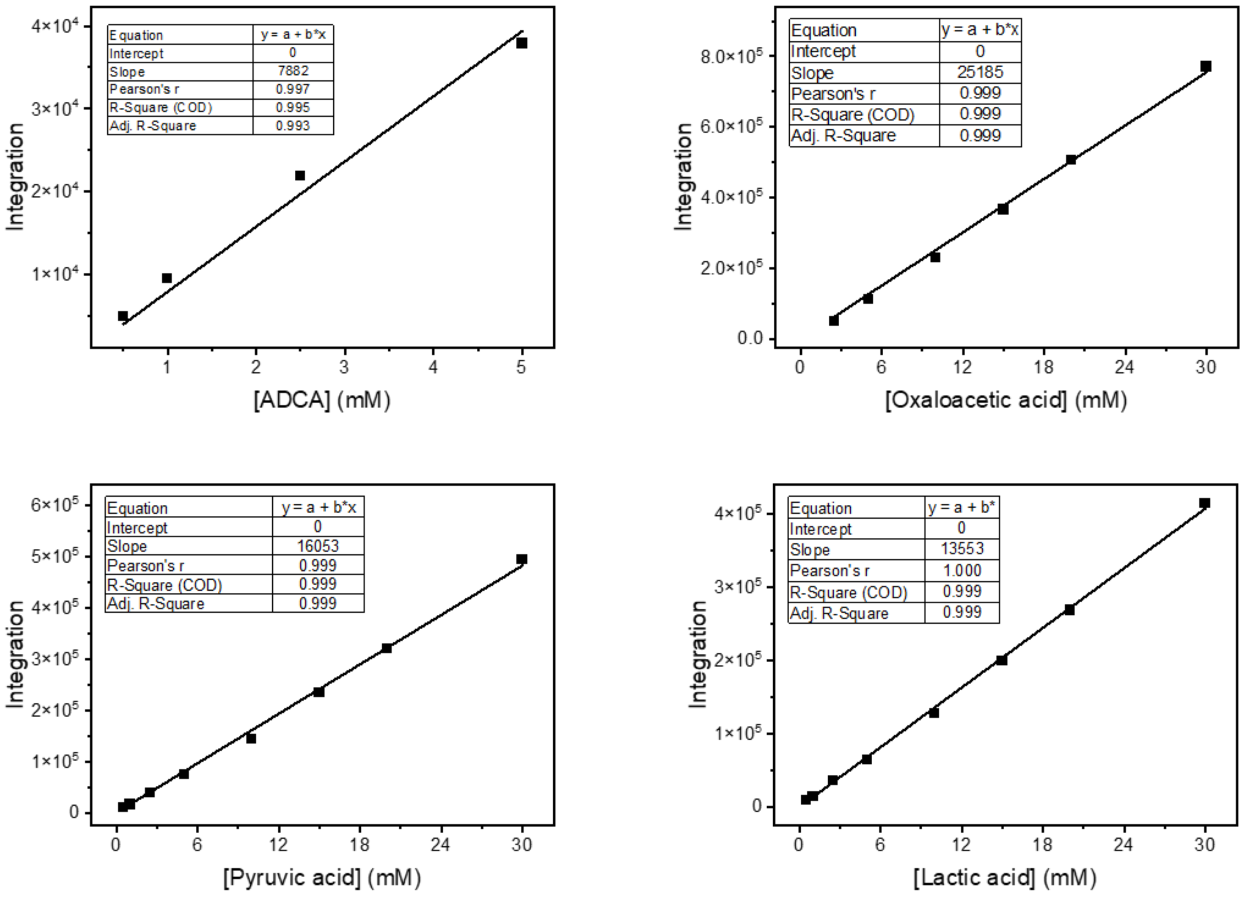
Enzyme	**Substrate/ Cofactor**	V_{max}	$\mathbf{K_m}$	k_{cat}	$k_{cat}/\mathbf{K_m}$
		µmol/min/mg	mM	s^{-1}	$M^{-1}s^{-1}$
LDH	Pyruvate	412 ± 16	0.24 ± 0.04	1000 ± 39	$(4.2 \pm 0.9) \times 10^6$
	NADH	412 ± 8	0.0023 ± 0.0002	1000 ± 19	$(4.4 \pm 0.5) \times 10^8$
ALDH	Acetaldehyde	0.120 ± 0.003	0.80 ± 0.03	0.120 ± 0.003	$(1.5 \pm 0.9) \times 10^3$
	NAD$^+$	0.080 ± 0.004	0.42 ± 0.06	192 ± 37	$(8.1 \pm 0.01) \times 10^{-2}$

4.2.3 *In vitro* production of lactate

In vitro reactions were set up for lactate production using 25 mM ADCA, 0.1 U (µmol/min) of each enzyme (crude FumA, MaeBN155D, and LDH), and 0.01 U of ALDH with different molar equivalents of NADH cofactor (1, 0.1, 0.01, and 0.001) relative to ADCA concentration. Different concentrations of NADH were investigated to determine the effect of cofactor concentrations on the reaction efficiency and identify appropriate working concentrations. The reaction progress was monitored using ^{1}H NMR, and the yields of products and reaction intermediates were determined at various time points using HPLC. Aliquots of samples were taken every 30 min for 3 h, followed by quenching with 5 M HCl for subsequent analysis HPLC analysis. The HPLC instrument was equipped with refractive index detection (RID) and variable wavelength detection (VWD) to identify and quantify the reaction intermediates and products for kinetics and yield analyses. Calibration curves used for the study are shown in Figure 4.7. HPLC analysis was conducted at the 15 and 24 h time points.

The results revealed that varying NADH concentration significantly affects the hydration of ADCA by FumA. At lower NADH molar equivalents (0.001 (0.025 mM) and 0.01 (0.25 mM)), we observed complete consumption of ADCA within about 30 min, suggesting better enzymatic activity under these conditions (Figure 4.8). However, increasing the NADH molar equivalents to 0.1 and 1 significantly slowed the rate of ADCA consumption, with negligible consumption of ADCA for the reaction containing 25 mM (1 equiv.) of NADH, indicating a possible inhibitory effect at higher NADH concentrations on FumA activity. This observation informed our selection of the appropriate equivalents of NADH (0.001) in subsequent optimization and scale-up efforts.

The reaction was then scaled up to 100 mM and 500 mM ADCA concentrations to further investigate the reaction kinetics and titers. ^{1}H NMR spectra for the reactions (100 mM and 500 mM initial ADCA concentrations) after 15 h are shown in Figure 4.9 – 4.11.The concentrations of enzymes were deliberately not changed to adequately probe the kinetics of the reaction by saturating with substrates. The 100 mM scale reaction (Reaction B) produced 0.6 mol mol^{-1} lactate yield with complete ADCA consumption within 180 min, although with significant accumulation of pyruvate (40 mM) (Figure 4.12). For the 500 mM reaction (Reaction C), about 55% ADCA was consumed after 24 h, with significant pyruvate accumulation (50 mM after 24 h), suggesting either an insufficient enzyme concentration or an inhibitory effect of elevated NADH levels on FumA activity. About 225 mM of ADCA remained after 24 h with 0.47 mol mol^{-1} lactate yield (Figure 4.13).



Equation	y = a + b*x
Intercept	0
Slope	7882
Pearson's r	0.997
R-Square (COD)	0.995
Adj. R-Square	0.993

Equation	y = a + b*x
Intercept	0
Slope	25185
Pearson's r	0.999
R-Square (COD)	0.999
Adj. R-Square	0.999

Equation	y = a + b*x
Intercept	0
Slope	16053
Pearson's r	0.999
R-Square (COD)	0.999
Adj. R-Square	0.999

Equation	y = a + b*
Intercept	0
Slope	13553
Pearson's r	1.000
R-Square (COD)	0.999
Adj. R-Square	0.999

Figure 4.7. Calibration curves for ADCA (top left) and oxaloacetic acid (top right) pyruvic acid (bottom left) and lactic acid (bottom right). Samples were run at 0.6 mL min^{-1} with 0.01 N H_2SO_4 mobile phase for 18 min.

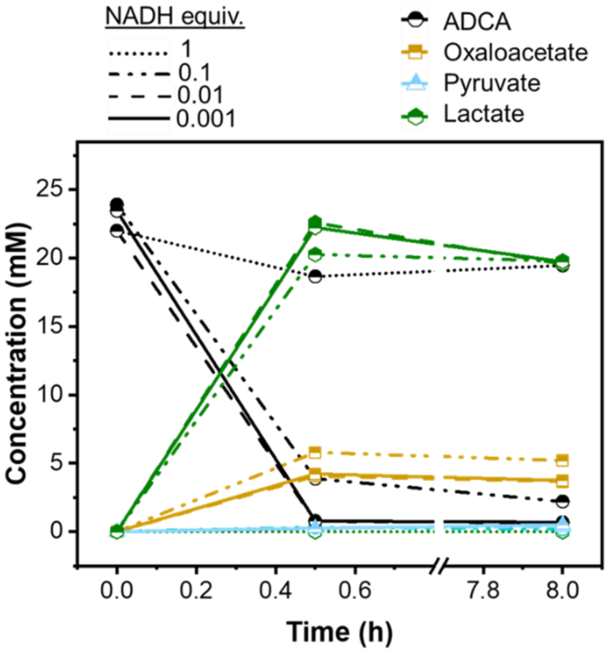


Figure 4.8 ***In vitro*** **synthesis of lactic acid with 25 mM ADCA and various equivalents of NADH.** The reaction was initiated by adding crude FumA, MaeBN155D, LDH, ALDH, and 1 equivalents of acetaldehyde to 25 mM ADCA in a 2 mL reaction. The 1 equivalents reaction did not contain ALDH and acetaldehyde. Concentrations were determined using HPLC.

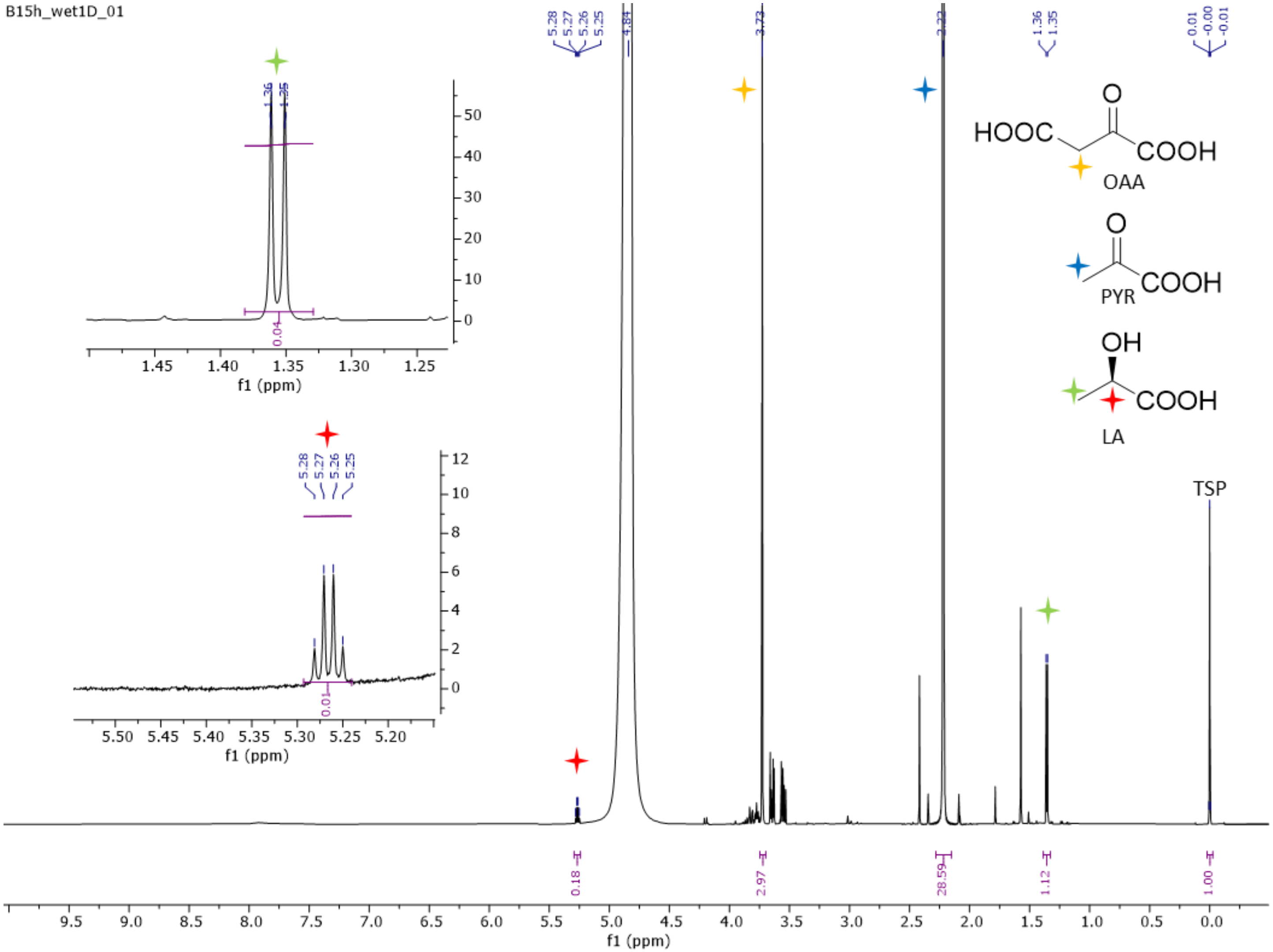


Figure 4.9. ^{1}H NMR for *in vitro* synthesis of lactic acid with 100 mM ADCA and 0.1 mM (0.001 equivalents) of NADH. This spectrum was taken after 15 h. The zoomed panels indicate the peak signals for lactic acid.

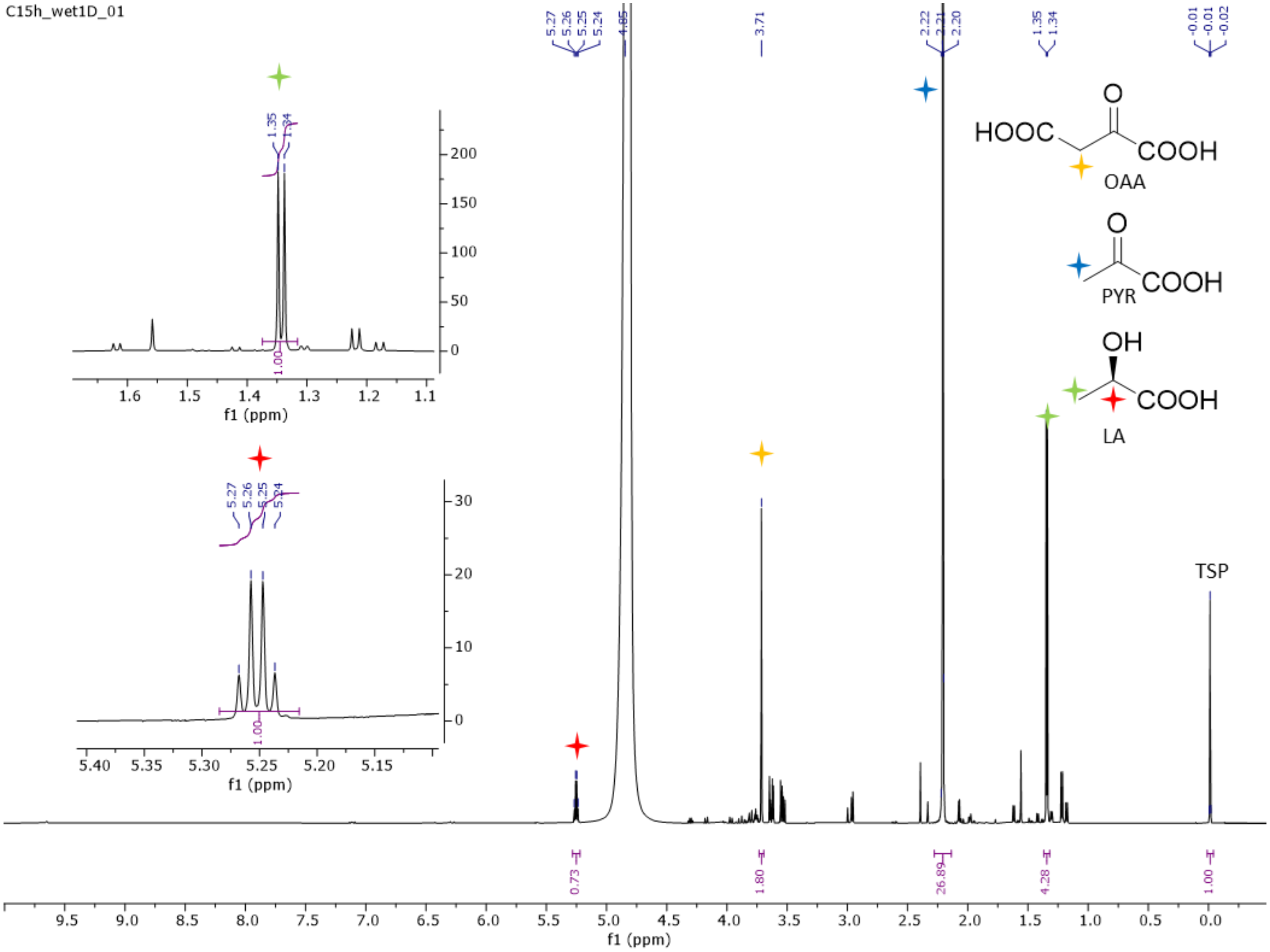


Figure 4.10. ^{1}H NMR for *in vitro* synthesis of lactic acid with 500 mM ADCA and 0.5 mM (0.001 equivalents) of NADH. This reaction was performed in one step. This spectrum was taken after 15 h. The zoomed panels indicate the peak signals for lactic acid.

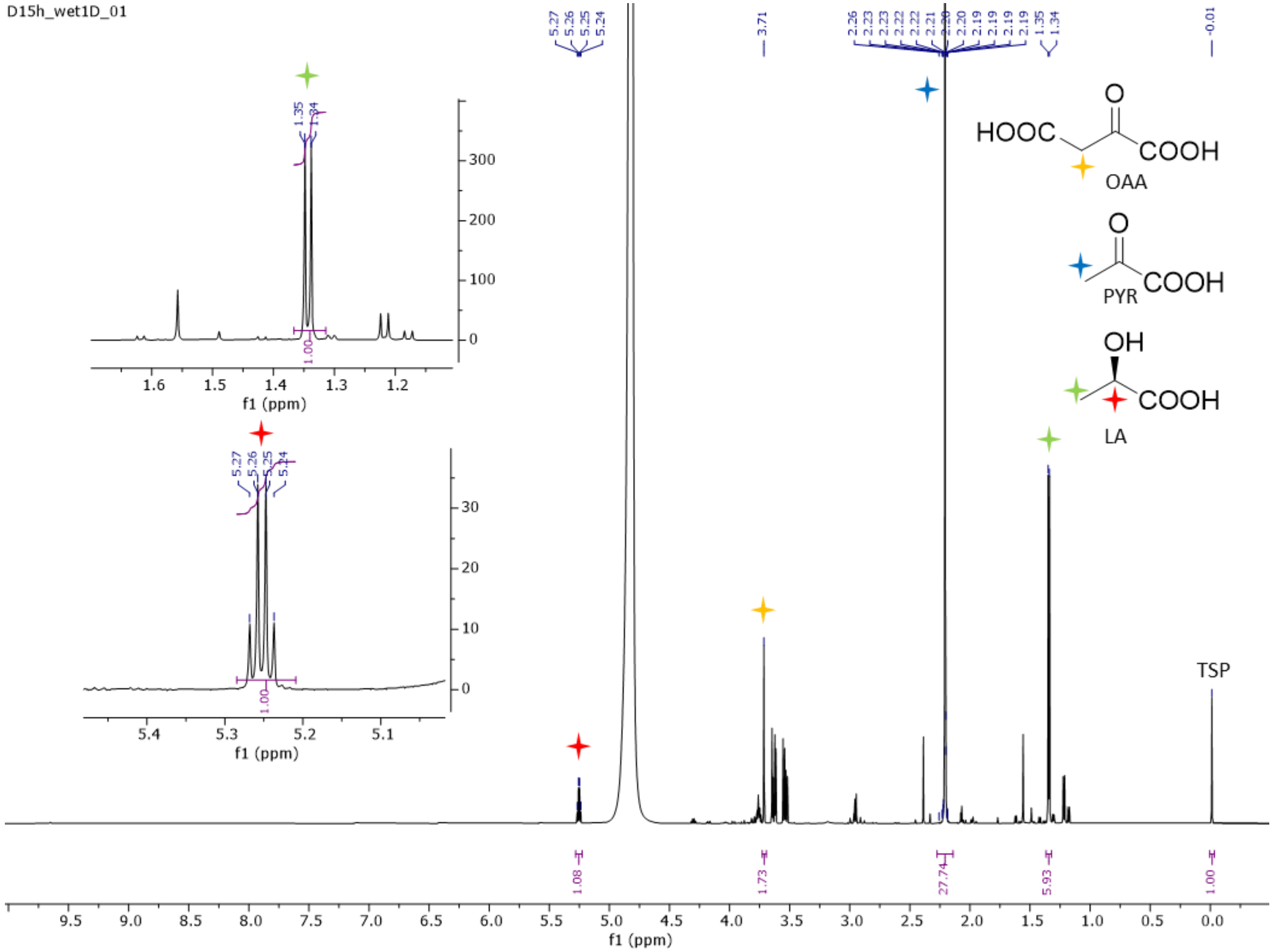


Figure 4.11. ^{1}H NMR of *in vitro* synthesis of lactic acid with 500 mM ADCA and 0.5 mM (0.001 equivalents) of NADH. This reaction was performed in multiple steps. This spectrum was taken after 15 h. The zoomed panels indicate the peak signals for lactic acid.

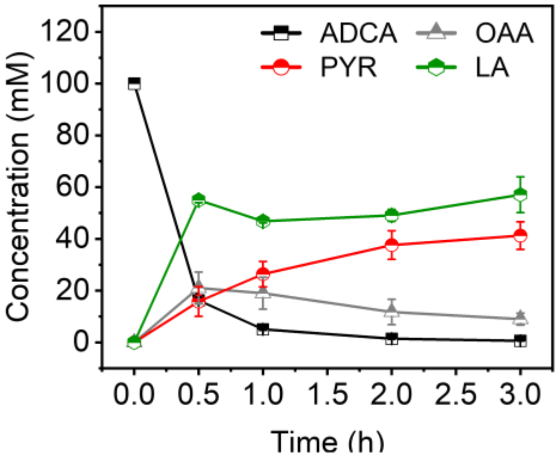


Figure 4.12. *In vitro* synthesis of lactic acid with 100 mM ADCA and 0.001 equivalents of NADH (Reaction B). The reaction was initiated by adding crude FumA, MaeBN155D, LDH, ALDH, and 1 equivalents of acetaldehyde to 100 mM ADCA in a 2 mL reaction. Concentrations were determined using HPLC. Error bars represent duplicates.

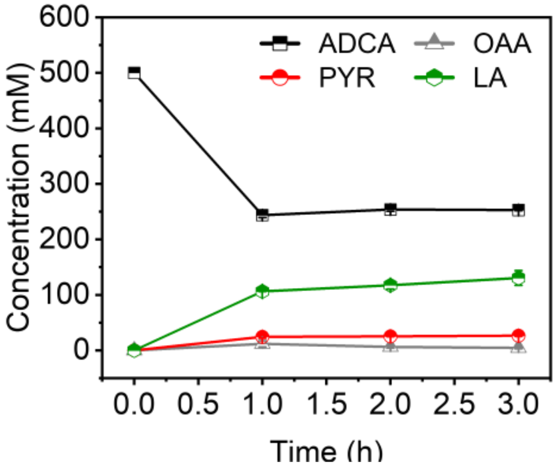


Figure 4.13. *In vitro* synthesis of lactic acid with 500 mM ADCA and 0.001 equivalents of NADH (Reaction C). The reaction was initiated by adding crude FumA, MaeBN155D, LDH, ALDH, and 1 equivalents of acetaldehyde to 500 mM ADCA in a 2 mL reaction. Concentrations were determined using HPLC. Error bars represent duplicates.

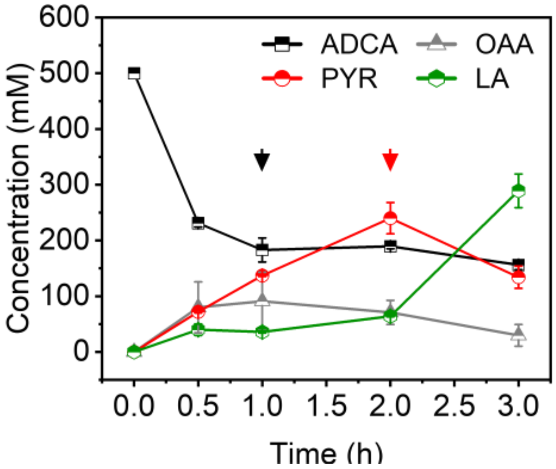
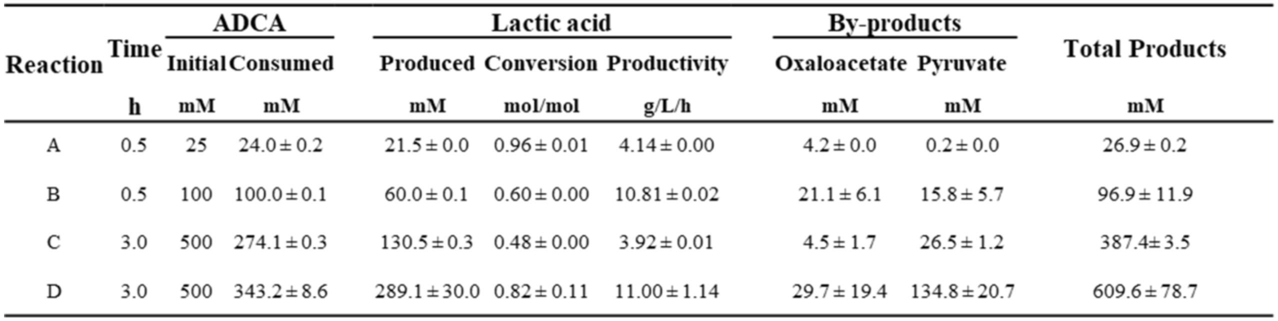


Figure 4.14. *In vitro* synthesis of lactic acid with 500 mM ADCA and 0.001 equivalents of NADH (Reaction D). The reaction was carried out in multiple steps by adding crude FumA to 500 mM ADCA in a 2 mL reaction, then MaeBN155D after an h (t = 1 h), LDH, ALDH, and 1 equivalents of acetaldehyde were added after another h (t = 2 h). Concentrations were determined using HPLC. Error bars represent duplicates.

Table 4.2. Summary table for *in vitro* production of lactic acid.



Reaction	Time	ADCA		Lactic acid			By-products		Total Products
		Initial	Consumed	Produced	Conversion	Productivity	Oxaloacetate	Pyruvate	
	h	mM	mM	mM	mol/mol	g/L/h	mM	mM	mM
A	0.5	25	24.0 ± 0.2	21.5 ± 0.0	0.96 ± 0.01	4.14 ± 0.00	4.2 ± 0.0	0.2 ± 0.0	26.9 ± 0.2
B	0.5	100	100.0 ± 0.1	60.0 ± 0.1	0.60 ± 0.00	10.81 ± 0.02	21.1 ± 6.1	15.8 ± 5.7	96.9 ± 11.9
C	3.0	500	274.1 ± 0.3	130.5 ± 0.3	0.48 ± 0.00	3.92 ± 0.01	4.5 ± 1.7	26.5 ± 1.2	387.4± 3.5
D	3.0	500	343.2 ± 8.6	289.1 ± 30.0	0.82 ± 0.11	11.00 ± 1.14	29.7 ± 19.4	134.8 ± 20.7	609.6 ± 78.7

To further investigate FumA inhibition or turnover for optimization of lactate yields, the 500 mM reaction was run in multiple steps. First, the hydration of ADCA to oxaloacetate by FumA for an hour, followed by the addition of MaeBN155D and Mn^{2+} to catalyze decarboxylation of oxaloacetate to pyruvate. After another 1 h, all remaining enzymes (LDH and ALDH), acetaldehyde, and cofactors were added to the reaction mixture. These stepwise additions resulted in an improved ADCA conversion rate of approximately 82% mol mol^{-1} within 3 h. This observation strengthens the hypothesis of inhibition of FumA activity in the presence of high concentrations of NADH. The summary for the *in vitro* production of lactate is shown in Table 4.2.

4.2.4 Challenges with *in vitro* lactate production route

The proposed method for *in vitro* biosynthesis of lactate uses four enzymes: FumA, MaeB, LDH, and ALDH. Although each of the enzymes exhibit good catalytic efficiencies for their specific reactions (except for ALDH), there are several challenges associated with this method, particularly the stability of the FumA enzyme. FumA, a class I fumarase containing a 4Fe-4S iron-sulfur cluster, has an iron atom that is not bound to any amino acid and actively participates in catalysis. This unbound iron is highly susceptible to oxidation when exposed to air; approximately 70% of the enzymatic activity of FumA has been reported to be lost after only 30 seconds of exposure to air, and only 10% remains after two min.[37] FumA activity can be restored after incubation with ferrous ammonium sulfate and 2-mercaptoethanol.

To evaluate the stability and identify optimal storage conditions, we overexpressed, purified, and stored aliquots of FumA in 20% ethylene glycol in 100 mM sodium phosphate buffer and 0.5 M ammonium sulfate prepared in degassed and deionized water.

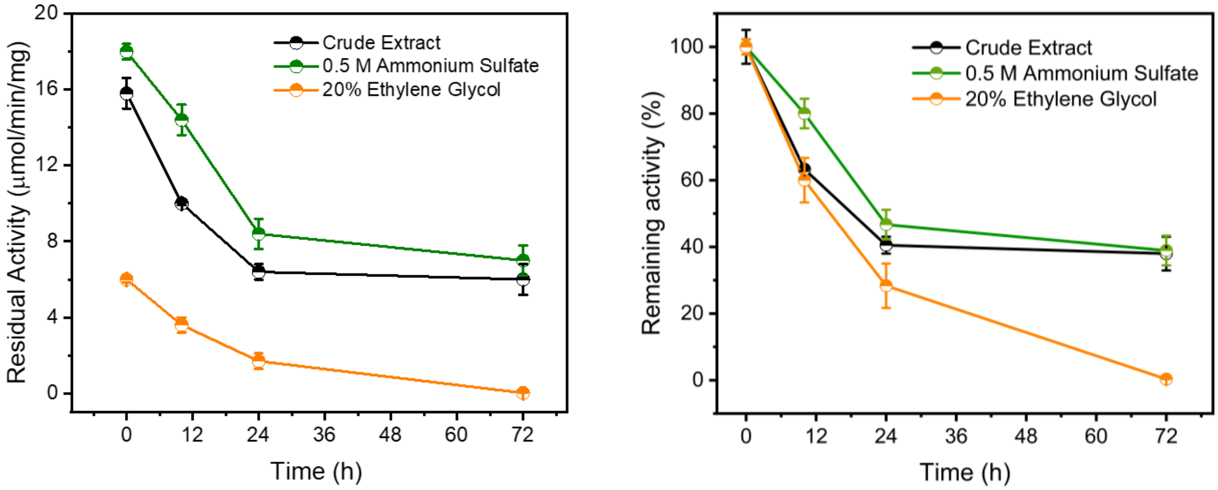


Figure 4.15. Stability of FumA. Residual activity of crude FumA (black), purified FumA stored in 0.5 M ammonium sulfate buffer (green), and 100 mM sodium phosphate buffer (pH 8) with 20% ethylene glycol. Activity was determined using the plate reader and initiated by adding 5 mM of ADCA to buffer containing enzymes. The rate of formation of oxaloacetate was monitored at 275 nm. Both plots were obtained from same assays reported as residual specific activity (left) and remaining activity (right).

The residual activities of FumA in these two storage solutions were determined over time alongside crude extract. It was found that FumA stored in 20% ethylene glycol showed a significant loss of activity from the purification time to assay. This is shown by the specific activities at t = 0, where crude FumA and pure FumA in ammonium sulfate showed 15.8 and 18 µmol/min/mg activities, respectively, while FumA in 20% ethylene glycol showed 6 µmol/min/mg. FumA in ammonium sulfate retained over 80% of its activity 10 h after purification, providing a reliable timeframe for enzymatic assays and reactions (Figure 4.15). We observed that crude FumA was more stable than the purified enzyme stored in 20% ethylene glycol. Besides the challenge of dealing with an unstable FumA enzyme, the *in vitro* pathway also includes decarboxylation step. Theoretically, the lost CO_2 could be captured and recycled to the carboxylation cycle, potentially providing a closed-loop system with net-zero CO_2 emissions.

An approach to address both the stability issues of FumA and the loss of CO_2 could be achieved using microbial cell factories since no FumA purification and production of lactic acid is coupled to cell growth on ADCA. This strategy is discussed in more detail in the following section.

4.2.5 Production of lactic acid from ADCA using evolved and engineered *E. coli*

Chapter Two examined the capability of *E. coli* to utilize ADCA exclusively for carbon and energy. The results revealed that the *E. coli* enzyme fumarase A (FumA) is essential for the growth on ADCA. Mutants deficient in FumA lost growth phenotype on this substrate, whereas mutants lacking the anaerobic dicarboxylate transporter (DcuC) show significantly improved growth. The critical function of FumA, as supported by the enzyme assays of Flint et al.[38,39] and further validated by enzymatic analysis in this study, is to convert cytosolic ADCA to oxaloacetate. This conversion is critical because oxaloacetate can enter the tricarboxylic acid (TCA) cycle, and lead to pyruvate formation which enables cell growth and biomass production.

Within the *E. coli* Dcu (anaerobic dicarboxylate carrier) protein family, DcuC acts as a secondary dicarboxylate carrier and is activated when DcuA and DcuB are insufficient.[40,41] We observed that adapted ACG populations tend to upregulate *dctA* (aerobic dicarboxylate transporter) and downregulate *dcuC*, optimizing ADCA uptake.

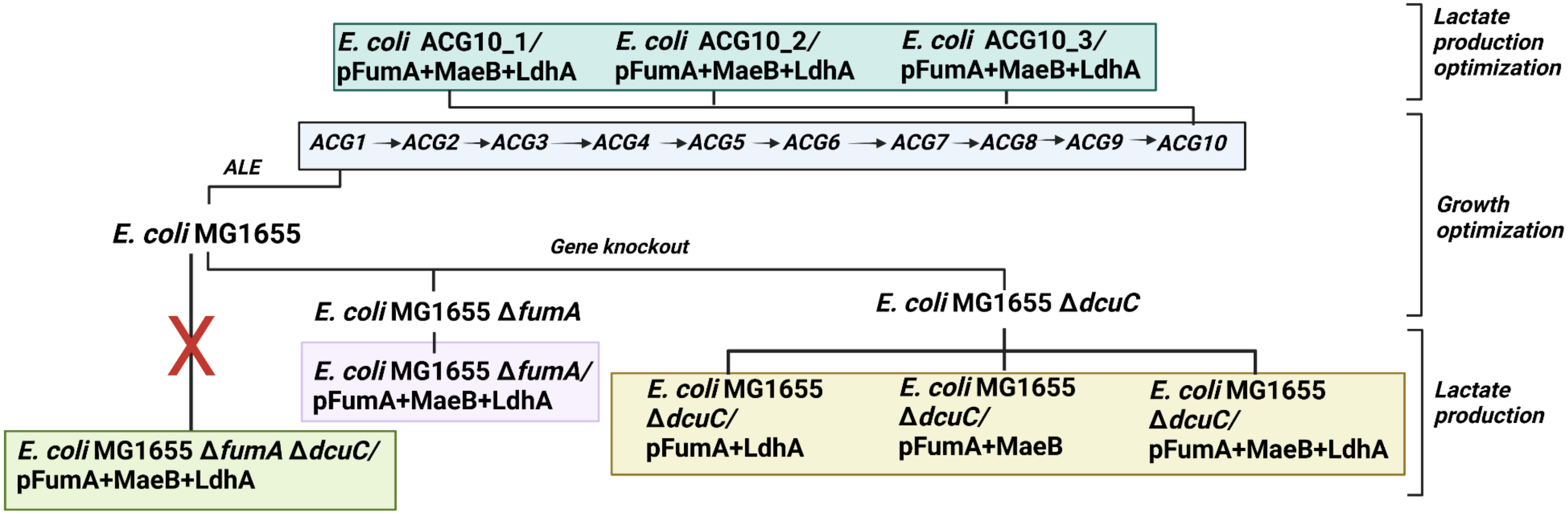


Figure 4.16. Strategies for *in vitro* production of lactic acid. Evolved clones from ACG10 were transformed with plasmid encoding FumA, MaeB, and LdhA. Knock out strains without evolution were also transformed with same plasmid. A strategy to use double knockout mutant Δ*fumA* Δ*dcuC* was attempted but the strain was not stable.

Four approaches were attempted for the *in vivo* production of lactate (Figure 4.16); the first strategy involved using a *dcuC* knockout strain transformed with plasmids to overexpress *E. coli* fumarase A (FumA), malic enzyme B (MaeB), and D-lactate dehydrogenase (LdhA). The engineered pathway for the conversion of ADCA to D-lactate involves these three enzymatic reactions as discussed for the *in vitro* route (Figure 4.17). Evaluation of the growth fitness of wild-type (W) and *dcuC* knockout (C) strains harboring the plasmids pFumA (2), pMaeB+FumA (3), pFumA+LdhA (4), and pMaeB+FumA+LdhA (5) in ADCA media confirmed that *dcuC* knockout strains generally exhibited improved growth fitness compared to their wild-type counterparts (Figure 4.18). In particular, strains combining *dcuC* knockout with *fumA* overexpression showed reduced lag phase times and increased growth rates.

The best-growing strains, C2, C3, C4, and C5, were cultivated in ADCA for lactic acid production using two conditions. Condition A was a continuous culture in which seed cultures were inoculated into 50 mL media containing 11.4 g L^{-1} ADCA and incubated at 37 °C for 100 h, with 200 rpm shaking (Figure 4.19). Condition B followed a batch process, starting with 50 mL media containing 6 g L^{-1} ADCA for 60 – 78 h (depending on strain), then adjusted with 6 mL of 34 g L^{-1} ADCA to adjust the total concentration to 11.4 g L^{-1} (Figure 4.20). Aliquots of cultures were taken every 12 h for HPLC analysis. Under both conditions, ADCA consumption was suboptimal, with cells consuming a maximum of 14.4 mM and 45 mM per 100 mM ADCA for conditions A (continuous) and B (batch), respectively (Figure 4.19 and 4.20, Table 4.3 entry 1 – 8). However, condition B resulted in better substrate utilization for all strains, with pyruvate and lactate production ratios consistent with the presence of MaeB and LdhA on the plasmid. Under condition A, more lactate was produced across all strains except for C5A (Δ*dcuC*/ pMaeB+FumA+LdhA: condition A). Strain C4A (Δ*dcuC*/ pFumA+LdhA: condition A) produced about 0.11 g L^{-1} (0.11 mol mol^{-1}) lactate and 0.013 g L^{-1} pyruvate. Strain C3A (Δ*dcuC*/ pMaeB+FumA: condition A) produced more lactate than pyruvate, although this strain overexpressed FumA and MaeB but not LdhA. with maximum O.D$_{600}$ between 0.8 and 1.5 and lag phase times of at least 48 h.

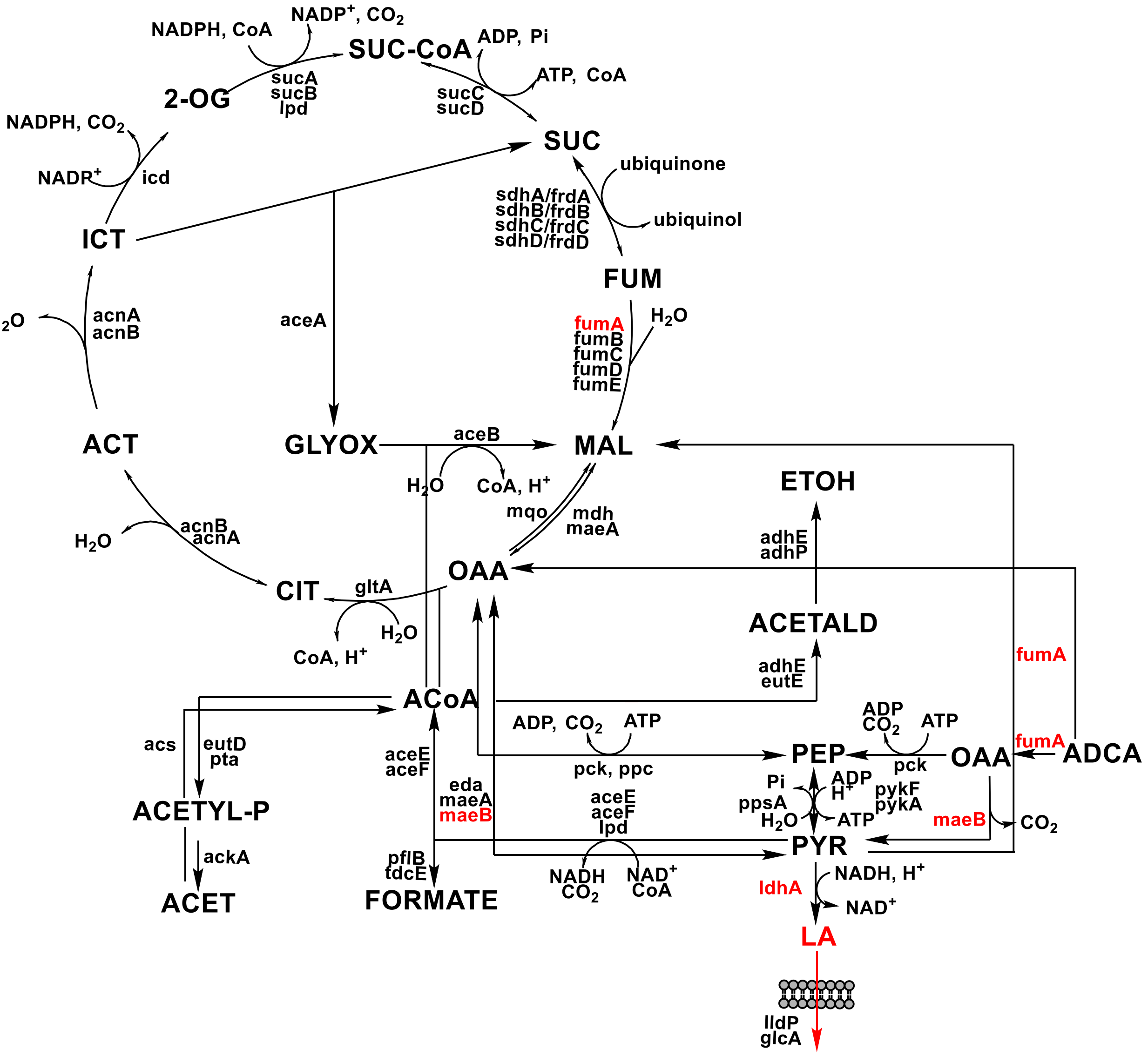


Figure 4.17. Route for *in vivo* production of lactate from ADCA. Enzymes in red were overexpressed from a plasmid. The pathway involves hydration of ADCA to oxaloacetate (OAA) by fumarase A (FumA), then decarboxylation of OAA to pyruvate (PYR) by malic enzyme B (MaeB) followed by reduction to lactic acid (LA) by lactate dehydrogenase (LdhA).

This could suggest automatic carbon flux to lactate under high ADCA concentrations, albeit only between (8 – 15) mM/ 100 mM ADCA was consumed (Table 4.3), and cells showed poor growth However, strain C5A overexpressing FumA, MaeB, and LdhA showed the poorest growth (long lag phase time) with the lowest lactate titers using condition A. This poor growth could be attributed to the combined metabolic burden of assimilating a new carbon source and overexpressing the enzymes from the plasmid. As previously discussed in Chapter Two, we observed that the wild-type experienced long lag phase times during growth in 100 mM ADCA.

Condition B showed better ADCA consumption across all strains, between 24 – 45 mM / 100 mM ADCA. The ratio of pyruvate and lactate titers correlated with the presence of MaeB and LdhA on the plasmid. C2B overexpressing only FumA produced 0.3 g L^{-1} pyruvate and about 0.13 g L^{-1} (0.03 mol mol^{-1}) lactate. C3B overexpressing FumA and MaeB produced more pyruvate (0.21 g L^{-1}) than lactate (0.04 g L^{-1}, 0.01 mol mol^{-1}). Overexpression of LdhA (C4B and C5B) resulted in higher titers of lactate, 0.12 g L^{-1} (0.03 mol mol^{-1}) and 0.25 g L^{-1} (0.07 mol mol^{-1}), respectively (Figure 4.21). See summary table (Table 4.3).

During growth on ADCA, some *E. coli* populations can develop mutations and exhibit differential gene expression, potentially leading to antibiotic resistance, as discussed in Chapter Two. Since ampicillin is used to maintain plasmids, the emergence of resistance could lead to plasmid loss. To further investigate this hypothesis, the susceptibility of ACG populations to various antibiotics was examined and discussed in the next section.

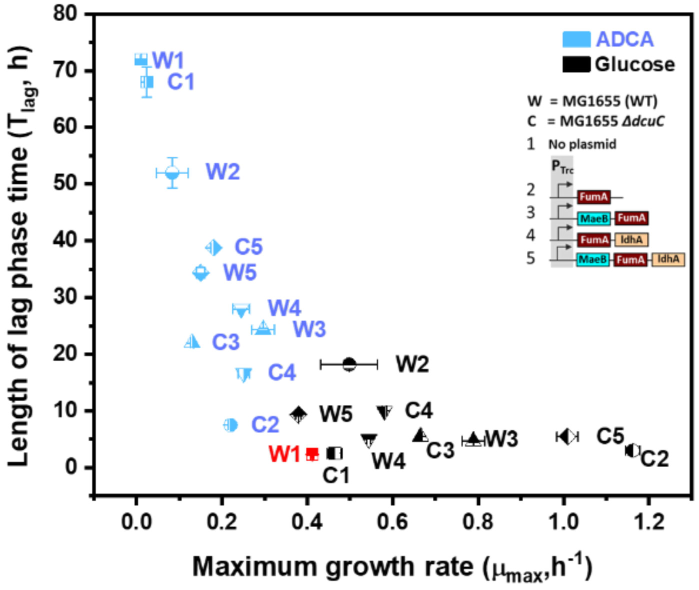


Figure 4.18. Bacterial elasticity plot comparing growth performance of wild-type MG1655 (label W) and *dcuC* (label C) derivatives in glucose and ADCA media. Strains cultivated in glucose (black) performed better than cells in ADCA (blue). The red datapoint represents MG1655 (wt) in glucose as reference. Each strain is represented by a different symbol.

4.2.6 Plasmid curing due to gain of antibiotic resistance during ALE

While the ALE approach was initially intended to enhance bacterial utilization of ADCA, the experiment may induce the development of antibiotic resistance in the adapted populations. This phenomenon could occur because the environmental stress applied during ALE, such as nutrient limitations or toxic compounds, can induce gene mutations, as discussed in Chapter Two. These mutations may not only improve the metabolic capabilities of bacteria but also confer resistance to some antibiotics. There are few reports of non-targeted gain of antibiotic resistance during ALE.[42,43] However, these reports identify cross-resistance among various antibiotics after treating cells with specific antibiotics in an ALE approach. The intersection of metabolic optimization and antibiotic resistance provides a unique opportunity to explore the genetic and environmental mechanisms that drive dual adaptation in microbial systems. We hypothesize that the selective pressures imposed by the new carbon source and the presence of antibiotics could synergistically influence the evolutionary trajectories of these bacteria, resulting in optimized growth capabilities alongside increased antibiotic resistance. This is more likely for lineages that become mutators.

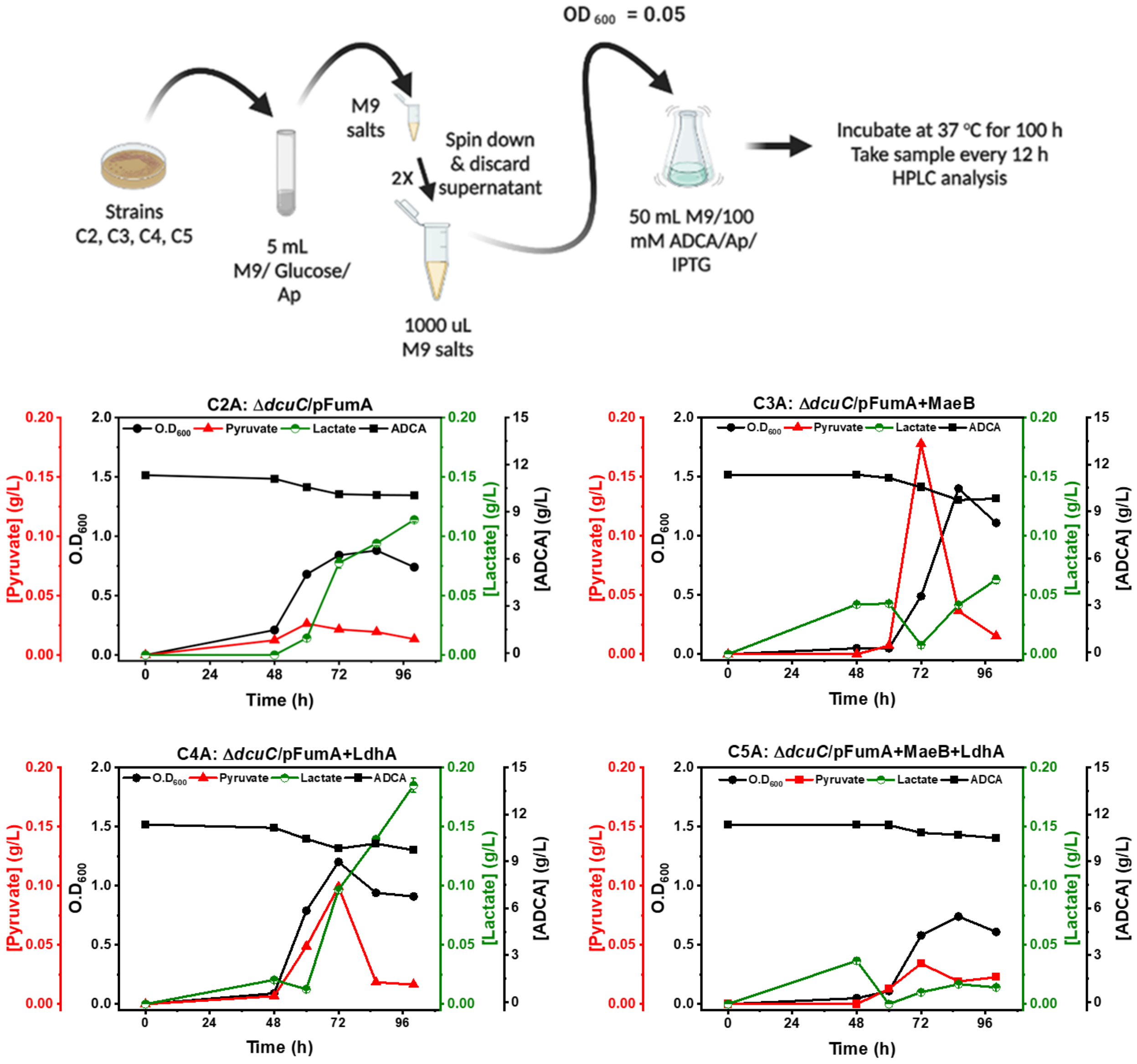


Figure 4.19. *In vivo* production of lactic acid (condition A- continuous method). Seed cultures of strains in M9G medium were used to inoculate 50 mL M9A media containing 100 mM ADCA. Samples were analyzed by HPLC for intermediate and product yields. Error bars indicate standard deviations for triplicate HPLC runs.

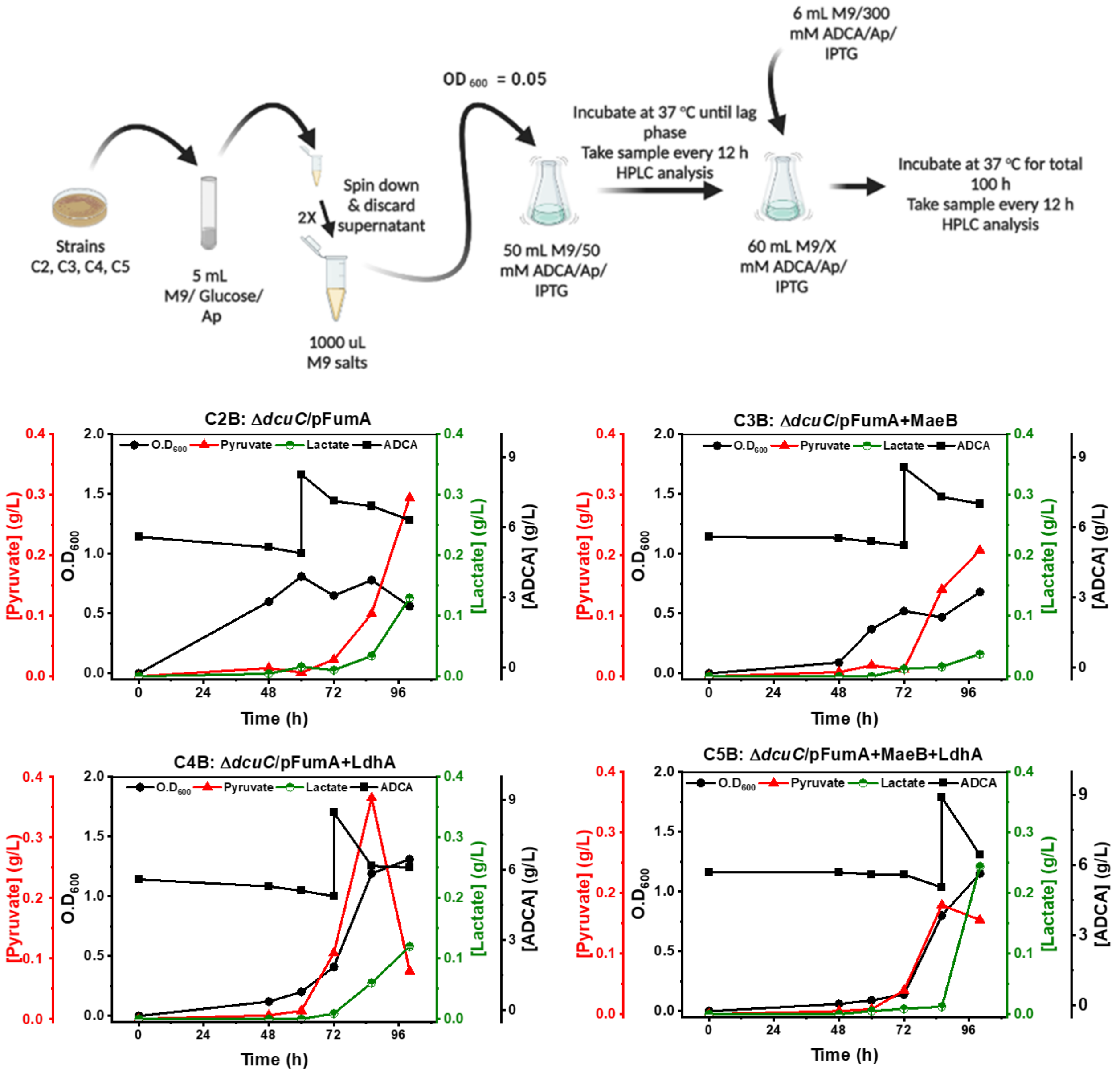


Figure 4.20. *In vivo* production of lactic acid (condition B- batch method). Seed cultures of strains in M9G medium were used to inoculate 50 mL M9A media containing 50 mM ADCA. ADCA (6 mL, 34 g L^{-1}) was added to the culture after 72 h. Samples were analyzed by HPLC ADCA, oxaloacetate, pyruvate and lactate. Error bars indicate standard deviations for triplicate HPLC runs.

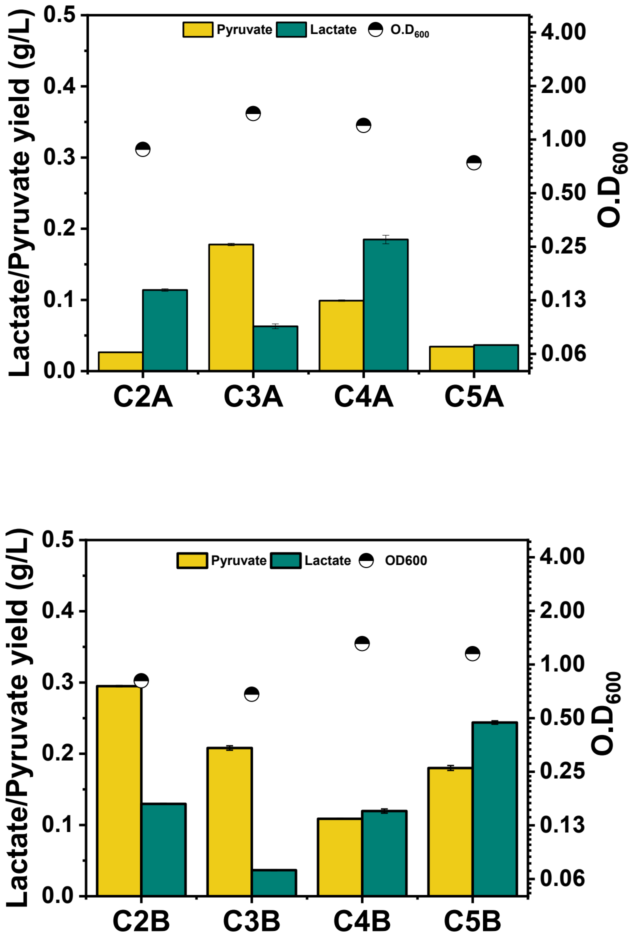


Figure 4.21. Lactic acid titers for conditions A (top) and B (down). The titers shown are the maximum observed lactate concentrations for each strain and condition. Strain abbreviation; C2, Δ*dcuC*/ pFumA; C3, Δ*dcuC*/ pMaeB+FumA; C4, Δ*dcuC*/ pFumA+LdhA; C5, Δ*dcuC*/ pMaeB+FumA+LdhA. Labels A and B represent conditions A and B as shown in Figure 4.19 and 4.20. Error bars indicate standard deviations for triplicate HPLC runs.

Disk diffusion assays, also known as the Kirby-Bauer antibiotic susceptibility test, is a widely used technique for evaluating the antibiotic susceptibility of bacteria. It involves placing antibiotic-impregnated paper disks on an agar plate with streaked bacterial culture.[44] The principle behind the assay is that antibiotics diffuse radially from the disk into the agar and inhibit bacterial growth, forming distinct inhibition zones around the disks. The size of these zones is measured to determine the susceptibility of the bacteria to the antibiotics.

The disk diffusion assay can be used to study the time course of antibiotic resistance due to adaptation to a new carbon source when studying *E. coli* populations from the ALE experiment. Before the ALE experiment, a disk diffusion assay can be performed on the wild-type, non-evolved *E. coli* strains to establish a baseline for antibiotic susceptibility. This baseline serves as a reference for later comparisons. After the ALE experiment, the same assay can be repeated with the evolved strains. This allows a direct comparison between the original and evolved strains, highlighting any changes in antibiotic resistance patterns. The sizes of the inhibition zones can be precisely measured. Changes in these measurements provide a clear visual representation of changes in resistance. Decreased zone sizes indicate increased resistance, while increased zones indicate sensitivity. Performing the disk diffusion assay at different time points during the ALE experiment can provide insight into the dynamics of resistance development. This analysis helps to understand how quickly resistance emerges and stabilizes in the population.

We tested the susceptibility of evolved ACG populations against seven antibiotics that are commonly used to maintain plasmids (working concentrations shown in parenthesis): ampicillin (50 µg mL^{-1}), chloramphenicol (35 µg mL^{-1}), erythromycin (20 µg mL^{-1}), kanamycin (50 µg mL^{-1}), spectinomycin (100 µg mL^{-1}), streptomycin (100 µg mL^{-1}), and tetracycline (15 µg mL^{-1}). The

ALE experiment was conducted for ten rounds to obtain ACG10 using *E. coli* MG1655 and shikimic acid producer RB791 *serA*::*aroB*Δ*aroK*Δ*aroL* as ancestor strains.

We observed that all evolved populations from both MG1655 and RB791 *serA*::*aroB*Δ*aroK*Δ*aroL* ancestors were still susceptible to spectinomycin, streptomycin, and tetracycline at the tested concentrations (Figure 4.22 – 4.23). However, some ACG populations from the MG1655 ancestor acquired resistance to ampicillin, chloramphenicol, and erythromycin. Erythromycin resistance was more prominent across all populations. Interestingly, in addition to erythromycin, some evolved populations from RB791 *serA*::*aroB*Δ*aroK*Δ*aroL* ancestor acquired resistance to chloramphenicol and kanamycin but not ampicillin. This observation suggests a difference in adaptation mechanism during the evolution of these two strains.

We observed that the acquired resistance could be lost and later regained during the course of evolution. For instance, MG1655 ACG3 acquires resistance to chloramphenicol, which is lost in ACG4 through ACG9. The resistance in chloramphenicol reappears in ACG10. A similar phenomenon was observed for MG1655 ACG populations with ampicillin.

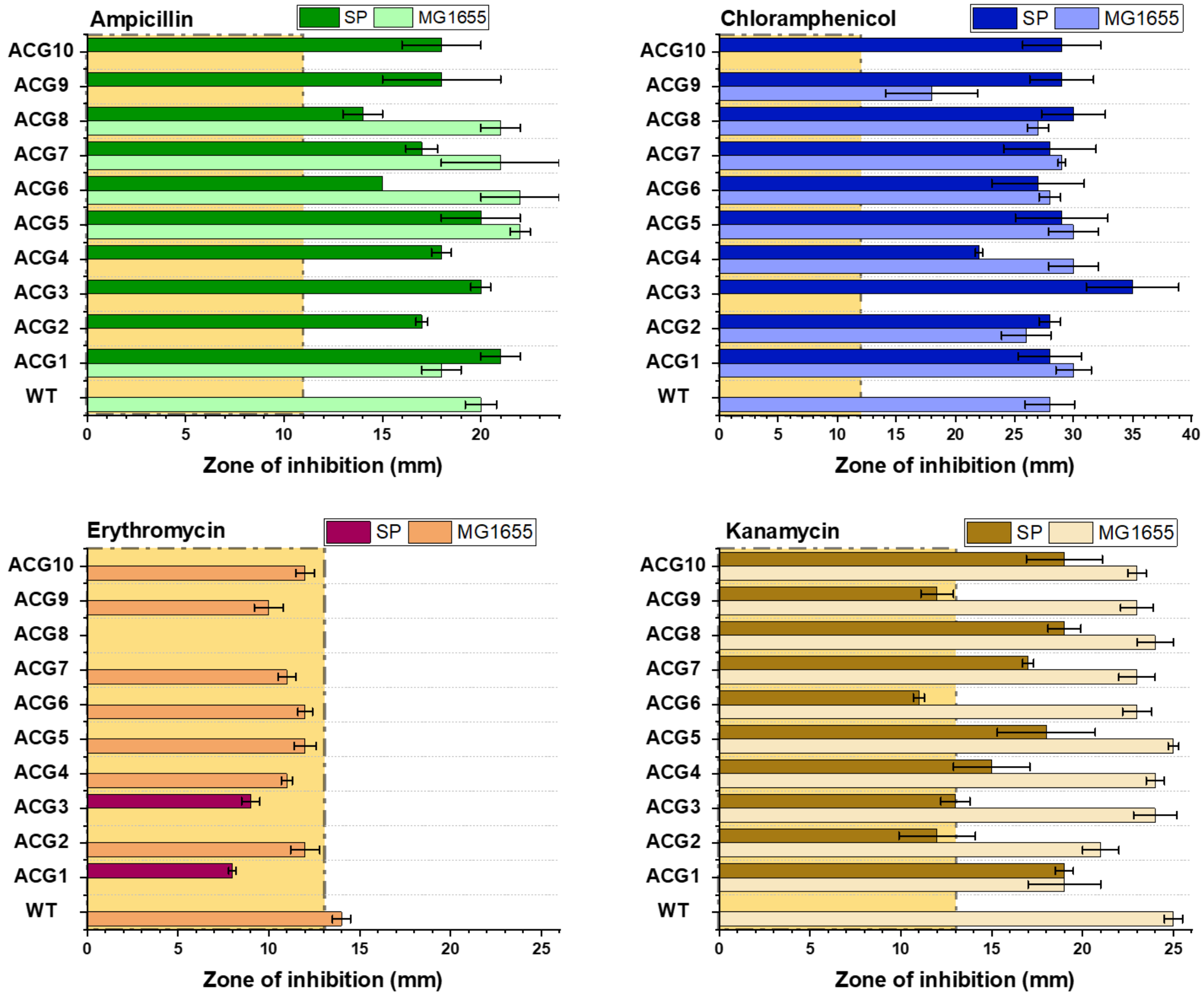


Figure 4.22 Disk diffusion assay testing susceptibility of ACG populations to various antibiotics; ampicillin (top left), chloramphenicol (top right), erythromycin (down left), and kanamycin (down right). Samples of ACG populations were plated onto LB agar plates. Disks impregnated with appropriate concentrations of various antibiotics were placed on plated samples and incubated at 37 °C for 24 h. Zones of inhibition were measured as diameter of clear zones around disks after incubation. MG1655 and *RB791 ser::aroB*Δ*aroK*Δ*aroL* (SP) were used as ancestor strains for the ALE. The shaded zones within plot are 'zones of resistance'. Populations with bars within this zone have acquired resistance. Error bars represents standard deviations for triplicates.

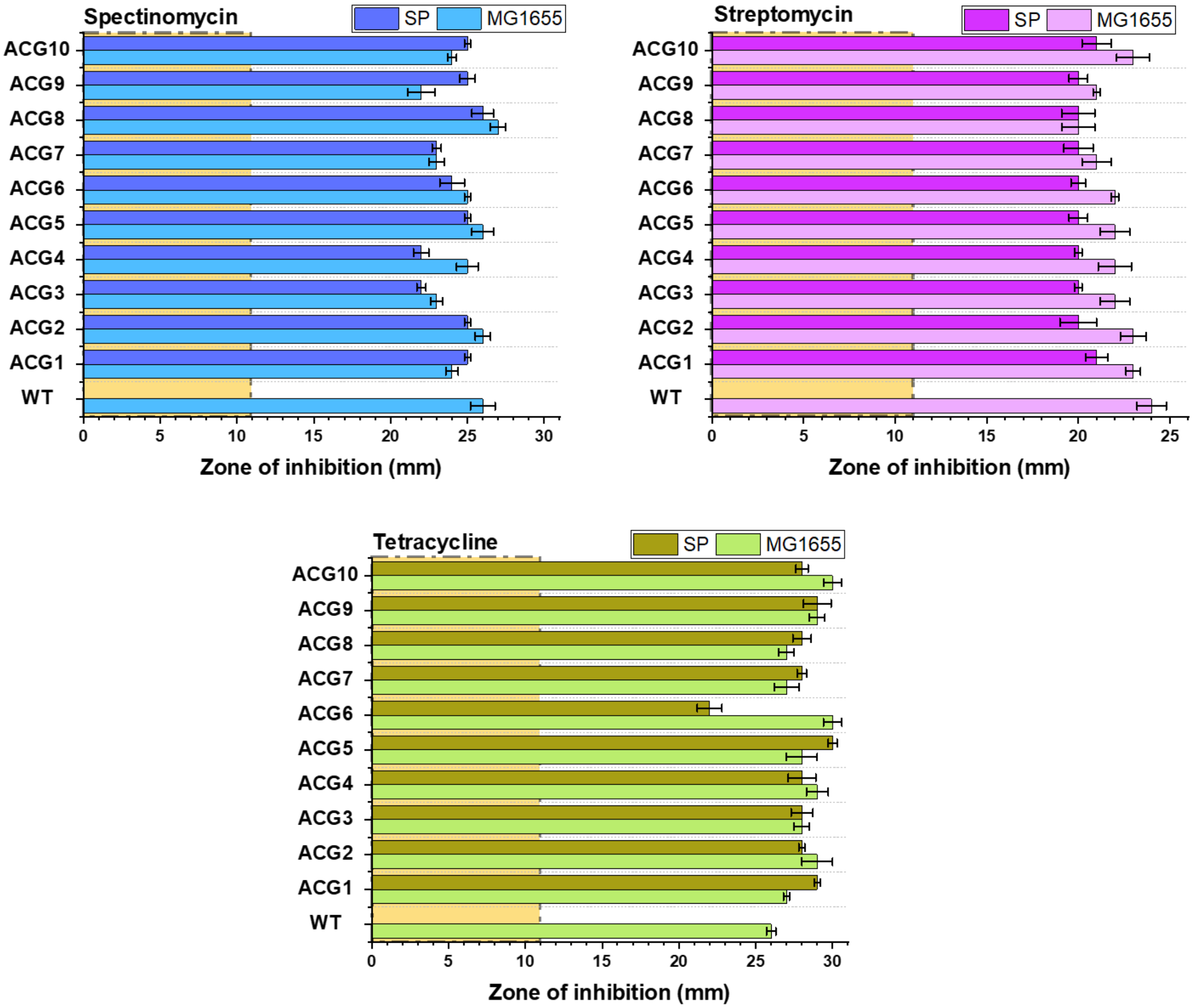


Figure 4.23. Disk diffusion assay testing susceptibility of ACG populations to various antibiotics; spectinomycin (left), streptomycin (right), and tetracycline (down). Samples of ACG populations were plated onto LB agar plates. Disks impregnated with appropriate concentrations of various antibiotics were placed on plated samples and incubated at 37 °C for 24 h. Zones of inhibition were measured as diameter of clear zones around disks after incubation. MG1655 and *RB791 ser::aroBΔaroKΔaroL* (SP) were used as ancestor strains for the ALE. The shaded zones within plot are 'zones of resistance'. Populations with bars within this zone have acquired resistance. Error bars represents standard deviations for triplicates.

Multidrug resistance in *E. coli* can arise from the action of several genes that enable the bacterium to resist the effects of multiple antibiotics. These genes often involve efflux pump systems, regulatory mechanisms, and other resistance pathways. Efflux pump genes encode proteins that form efflux systems capable of pumping a wide range of antibiotics out of the cell, reducing their effective intracellular concentrations. While antibiotics were not a selection pressure during the ALE experiment, mutations were observed in seven genes involved in efflux pump systems: multidrug efflux pump RND permeases (*acrBF*), putative arginine: ornithine antiporter (*acrD*), fosmidomycin efflux pump (*fsr*), ABC-type tripartite efflux pump ATP binding subunit (*macB*), and multidrug efflux pump RND permease subunits BC (*mdtBC*). The observed mutations could improve the efficiency of these efflux pumps, thereby resulting in a gain in multidrug resistance.

4.2.7 Optimization of *in vivo* lactic acid production

Antibiotics are typically used to force the retention of plasmids by bacterial strain. The inclusion of the antibiotic resistance gene on the plasmid and addition of the antibiotic to the culture media enables stable maintenance of the plasmid. Maintaining plasmid stability with an antibiotic like ampicillin is a common strategy; however, plasmid curing may occur if the cells acquire resistance to ampicillin during culturing, particularly during ALE. The addition of ampicillin ensures that only plasmid-containing cells grow. If cells become resistant to ampicillin, this breaks the selective pressure to maintain the plasmid, especially when the plasmids impose a fitness cost in the presence of antibiotic pressure. Hence, resistance may lead to plasmid loss in some cells, especially if the resistance provides a competitive advantage over maintaining the plasmid. This can result in a population of cells where some no longer produce the target chemical because they have lost the plasmid that encodes genes necessary for the biocatalytic pathway.

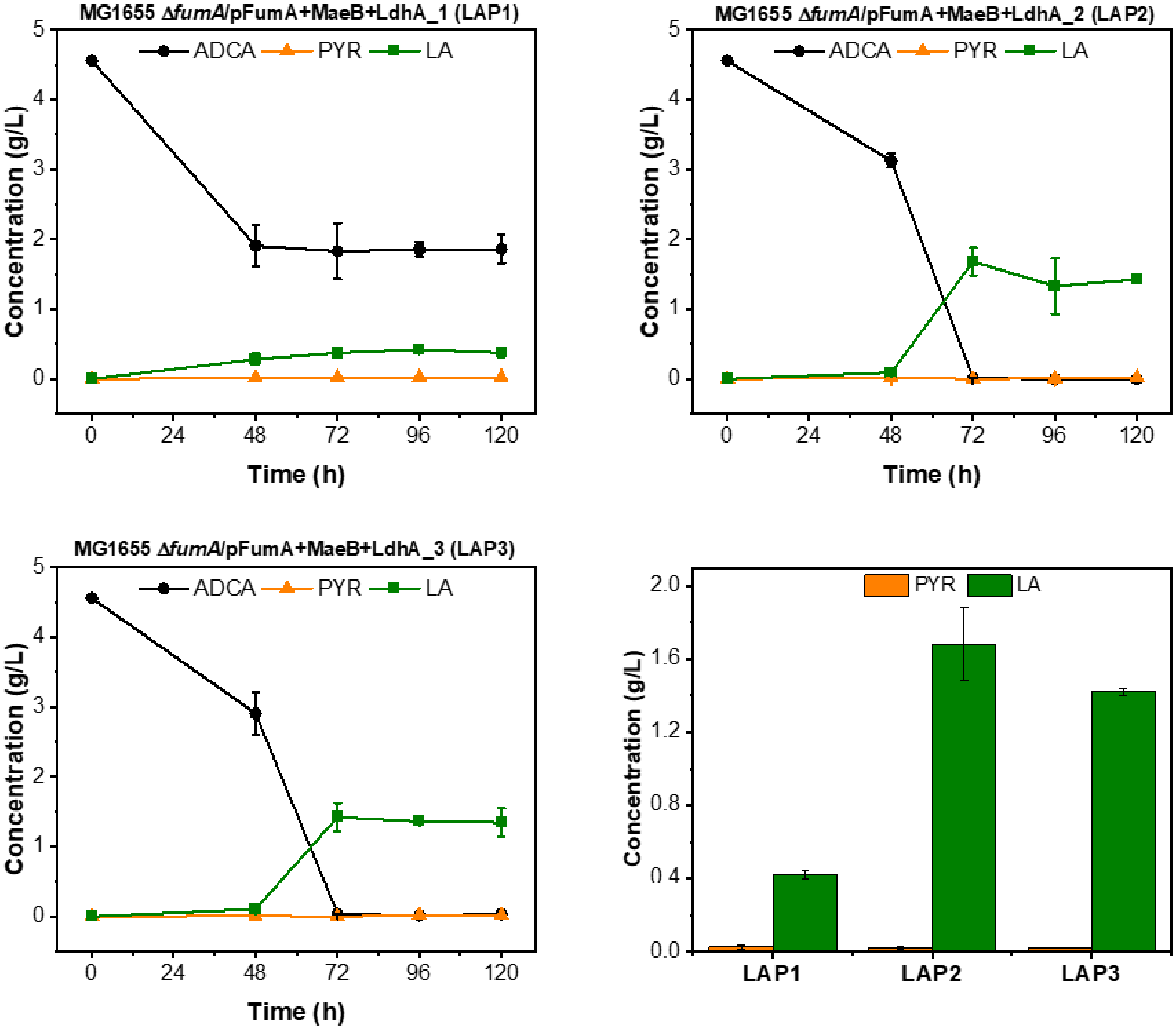


Figure 4.24. *In vivo* production of lactic acid using three MG1655Δ*fumA*/pFumA+MaeB+LdhA clones. Top and bottom left panels are HPLC analysis of fermentation supernatant during 120 h of culturing. Bottom right panel shows maximum pyruvate and lactate titers for three lactic acid producing (LAP123) clones. Error bars show standard deviations of duplicates of fermentations.

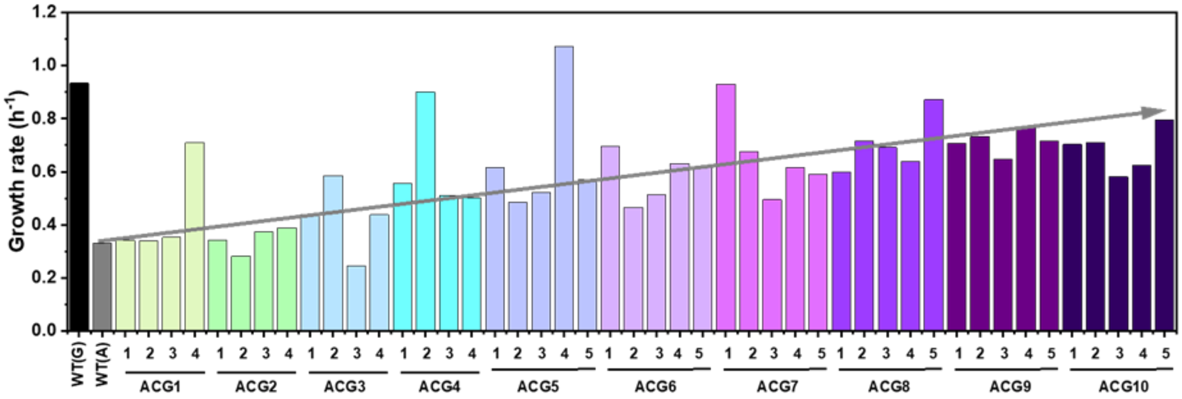


Figure 4.25. Growth fitness of single clones from ACG populations. Single colonies of evolved populations (ACG1 – 10) were selected from LB plates to inoculate 2 mL LB media, which was incubated at 37 °C with 200 rpm to obtain seed cultures. The seed cultures were used to inoculate 200 µL M9A media and growth was monitored using the plate reader. Growth rates were determined using the GrowthCurver

A strategy to maintain plasmid without addition of antibiotics to the media was implemented using MG1655Δ*fumA* as the parental strain. As described in Chapter Two, the Δ*fumA* mutant does not grow on ADCA. Complementing this mutant with a plasmid encoding the *fumA* gene can restore this growth on ADCA. Our strategy was to generate and test two strains, MG1655Δ*fumA*/pFumA+MaeB+LdhA and MG1655Δ*dcuC*Δ*fumA*/pFumA+MaeB+LdhA for lactate production. Several attempts were made to generate MG1655Δ*dcuC*Δ*fumA,* but this strain was unstable because we observed the loss growth phenotype even in rich media after rounds of cultivation. The MG1655Δ*fumA*/pFumA+MaeB+LdhA strain was prepared, and three clones were cultured in 50 mL media containing 4.5 g L^{-1}ADCA for lactate accumulation experiment (Figure 4.24). The need for FumA expression to support cell survival eliminates the need for antibiotics in this cultivation method. This approach increased lactate production significantly from 0.25 g L^{-1} (MG1655Δ*fumA*/pFumA+MaeB+LdhA) to approximately 1.7 g L^{-1} (0.47 mol mol^{-1}) (Figure 4.24, LAP2) with a productivity of 0.023 g L^{-1} h^{-1}. LAP2 and LAP3 strains metabolized all added ADCA within 72 h, during which maximum lactate titers were observed. The increase in lactate production supports the hypothesis that plasmid loss occurred during the growth on ADCA.

4.2.8 Lactate production using ACG populations

The MG1655Δ*fumA*/pFumA+MaeB+LdhA strain significantly improved lactate titers, although growth and ADCA consumption rate remained slow. This slow growth could be attributed to poor ADCA assimilation and the metabolic burden of expressing additional enzymes encoded by the plasmid since this strain was not adapted to grow on ADCA. To further improve lactate production, the MG1655 was subjected to ten rounds of adaptive laboratory evolution (ALE) to produce the ACG10 population. This extended evolution was necessary due to the instability observed in the ACG3 population, which exhibited a mix of mutator and non-mutator phenotypes, indicating significant heterogeneity. Not all clones within the ACG3 population could grow on ADCA, as confirmed by monitoring individual clones from lines 4, 5, and 6, which exhibited different growth efficiencies (Chapter Two).

It was crucial to select individual clones that were optimally adapted for growth on ADCA in order to advance the use of ACG strains for lactate production. Samples from ACG1 to ACG10 populations were diluted and plated onto LB and M9A media. Five clones from each population were used to generate seed cultures, which were used to inoculate M9A media containing 4.5 g L^{-1} ADCA. Growth profiles were monitored using a plate reader, revealing variability within populations. Early rounds of evolution (ACG1 to ACG4) showed inconsistent growth rates, with at least one strain failing to grow in each of these rounds (Figure 4.25). However, later rounds (ACG8, ACG9, and ACG10) showed more consistent and improved growth phenotypes, suggesting greater stability and reduced likelihood of further mutations during growth on ADCA.

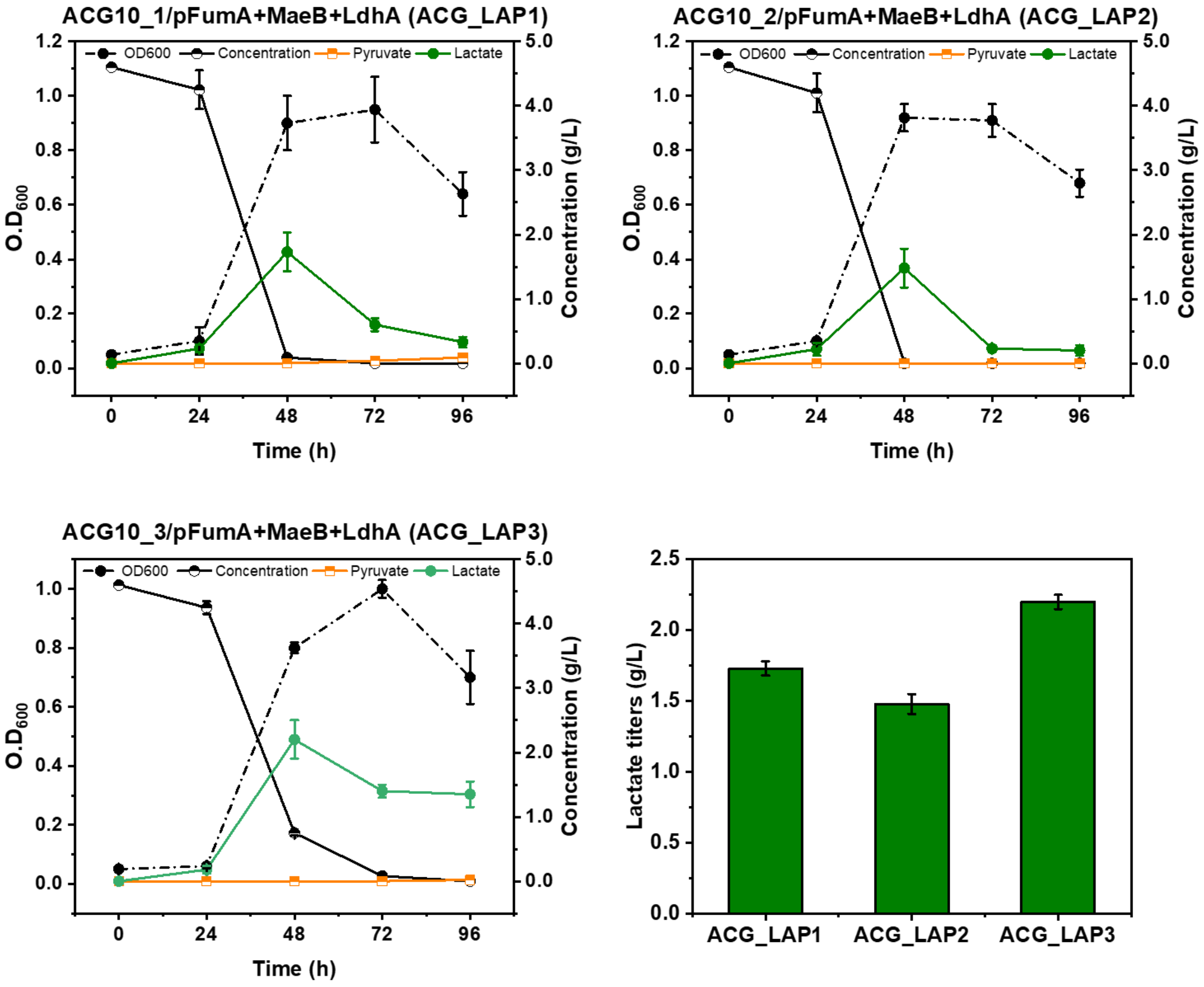


Figure 4.26. *In vivo* production of lactic acid using three ACG10/pFumA+MaeB+LdhA clones. Top and bottom left panels are HPLC analysis of fermentation supernatant. Bottom right shows lactate titers for the three clones after 48 h of culturing. Bottom right panel shows maximum lactate titers for three ACG10_LAP clones. Error bars show standard deviations of duplicates of fermentations.

Table 4.3. **Summary table for *in vivo* production of lactate from ADCA**

Entry	Strain	Condition	Time	[ADCA]		[Lactate]			By-products	Growth rate
				Initial	Consumed	Produced	Conversion	Productivity	Pyruvate	
			h	mM	mM	mM	mol/mol	$x10^{-3}$ g/L/h	mM	h^{-1}
1	Δ*dcuC*/pFumA	Continuous	100	100	11.3	1.26	0.11	1.14	0.15	0.058
2	Δ*dcuC*/pFumA+MaeB	Continuous	100	100	13.4	0.70	0.05	0.63	0.17	0.030
3	Δ*dcuC*/pFumA+LdhA	Continuous	100	100	14.4	2.05	0.14	1.85	0.19	0.062
4	Δ*dcuC*/pFumA+MaeB+LdhA	Continuous	100	100	7.5	0.16	0.02	0.14	0.26	0.038
5	Δ*dcuC*/pFumA	Batch	100	100	42.8	1.44	0.03	1.30	3.35	0.060
6	Δ*dcuC*/pFumA+MaeB	Batch	100	100	36.7	0.41	0.01	0.37	2.36	0.046
7	Δ*dcuC*/pFumA+LdhA	Batch	100	100	44.6	1.33	0.03	1.20	0.90	0.029
8	Δ*dcuC*/pFumA+MaeB+LdhA	Batch	100	100	41.7	2.71	0.06	2.44	1.77	0.018
9	Δ*fumA*/pFumA+MaeB+LdhA_1	Continuous	96	40	23.8	4.64	0.20	4.35	0.28	0.048
10	Δ*fumA*/pFumA+MaeB+LdhA_2	Continuous	72	40	40.0	18.68	0.47	23.37	0.27	0.054
11	Δ*fumA*/pFumA+MaeB+LdhA_3	Continuous	72	40	40.0	15.77	0.39	19.73	0.22	0.066
12	ACG10_1/pFumA+MaeB+LdhA	Continuous	48	40	40.0	19.21	0.48	36.04	0.00	0.086
13	ACG10_2/pFumA+MaeB+LdhA	Continuous	48	40	40.0	16.43	0.41	30.83	0.00	0.088
14	ACG10_3/pFumA+MaeB+LdhA	Continuous	48	40	40.0	24.42	0.61	45.83	0.00	0.083

Three colonies from ACG10 were selected and transformed with the pFumA+MaeB+LdhA+TetR plasmid to generate evolved ACG10 lactic acid producing (LAP) strains, ACG10_LAP1, ACG10_LAP2, and ACG10_LAP3. These strains were cultivated in shake flasks containing 50 mL M9A media with 4.5 g L^{-1} ADCA, and samples were analyzed by HPLC for lactate and pyruvate levels every 24 h. ACG10_LAP1 and ACG10_LAP2 rapidly consumed ADCA within 48 h, reaching maximum lactate titers of approximately 1.6 g L^{-1} (0.48 mol mol^{-1}) and 1.5 g L^{-1} (0.41 mol mol^{-1}), respectively, without accumulating pyruvate, suggesting efficient pyruvate utilization, possibly enhanced by the evolution and overexpressed MaeB (Figure 4.26).

Interestingly, these strains also showed lactate consumption between 48 and 72 h after depletion of ADCA. In contrast, ACG10_LAP3 exhibited slower ADCA consumption but reached the highest lactate titer of approximately 2.2 g L^{-1} (0.61 mol mol^{-1}) and productivity of 0.046 g L^{-1} h^{-1}, with no significant lactate consumption observed after ADCA depletion. The ability of these strains to switch to lactate utilization after ADCA depletion is likely due to the high NAD^{+}/NADH ratio, which promotes the conversion of lactate to pyruvate by LdhA. This pyruvate enters the TCA cycle to support energy production under nutrient-limited conditions. This observation is consistent with metabolic profiles indicating optimized growth on alternative substrates such as succinate, as noted in earlier generations. Chapter Two discusses how early generations, such as ACG1, accumulated significant amounts of succinate, while ACG2 and ACG3 did not, suggesting evolutionary adaptations to optimize succinate consumption.

4.3 Bioconversion of ADCA to shikimic acid

The biosynthesis of lactic acid from ADCA demonstrates a promising strategy for potentially converting C1 feedstocks such as methane and carbon dioxide to a chiral C3 compound via an acetylenedicarboxylate intermediate. Bioconversion to shikimic acid could demonstrate ADCA as a feedstock for synthesizing more complex molecules that are further downstream in the biosynthetic pathway. The Draths laboratory has previously demonstrated the production of shikimic acid from glucose using modified strains of *E. coli* RB791. These strains include *E. coli* RB791 *serA*::*aroB*Δ*aroK*Δ*aroL* and *E. coli* RB791 *serA*::*aroB*Δ*aroK*Δ*aroL*/pKD12.138A, in which relevant genes in the shikimate pathway have been knocked out (Figure 4.27 – 4.29). Specifically, shikimate kinase *aroK* and *aroL* were deleted to prevent the consumption of shikimate, thereby resulting in shikimic acid accumulation. In addition, the gene encoding 3-dehydroquinate synthase (aroB) was inserted into the locus of the phosphoglycerate dehydrogenase (serA) gene in *E. coli* RB791, effectively disrupting *serA* and facilitating plasmid maintenance through *serA* expression.

Plasmid pKD12.138A includes genes that encode a tyrosine-insensitive (feedback-resistant) 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase (aroF^{FBR}), shikimate dehydrogenase (aroE), 3-dehydroquinate synthase (aroB), transketolase 1 (tktA), and serA. The expression of *serA* allows pKD12.138A to be maintained in cells in minimal media. Under conditions lacking this plasmid, serine supplementation is required for cell growth.

DAHP synthase initiates the chorismate pathway by catalyzing the reaction between phosphoenolpyruvate (PEP) and D-erythrose 4-phosphate (E4P) to form 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP). In subsequent steps, dehydroquinate (DHQ) synthase catalyzes the conversion of DAHP to DHQ, followed by its conversion to 3-dehydroshikimate (DHS) and finally to shikimate, facilitated by DHQ dehydratase and shikimate dehydrogenase, respectively. Overexpression of *tktA* increases the carbon flux towards the shikimate pathway by improving the availability of the precursor E4P. Deletion of *aroK* and *aroL* is crucial to stopping the consumption pathways of shikimic acid and maximizing its accumulation. Various approaches were investigated for the *in vivo* production of shikimic acid from ADCA (Figure 4.27 – 4.29). First, the *E. coli* RB791 *serA*::*aroB*Δ*aroK*Δ*aroL* strain was cultured in 50 mL of M9A media with 4.5 g L^{-1} ADCA. This strain had not been adapted to utilize ADCA; the experiment aimed to assess whether shikimate could be produced, indicating potential redirection of carbon to the shikimate pathway, which will be crucial for strain enhancement.

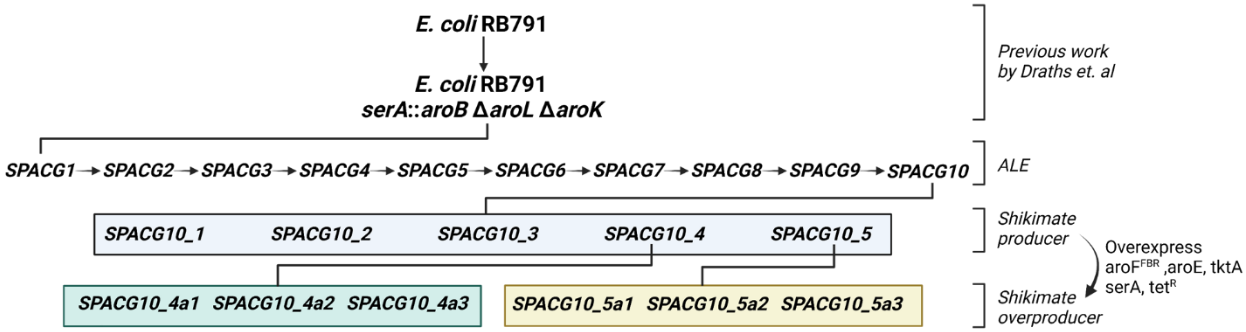


Figure 4.27. Strategies for in vitro shikimic acid production. *E. coli* RB791 *serA*::*aroB*Δ*aroK*Δ*aroL* is evolved after ten rounds to obtain SPACG1 – 10. Five colonies from SPACG10 were cultivated in M9A and monitored for shikimic acid production. Best shikimic acid producers were transformed with pKD12.138A encoding *aroF*^{FBR}, *aroE*, *tktA*, and *serA* in order to improve shikimate titers.

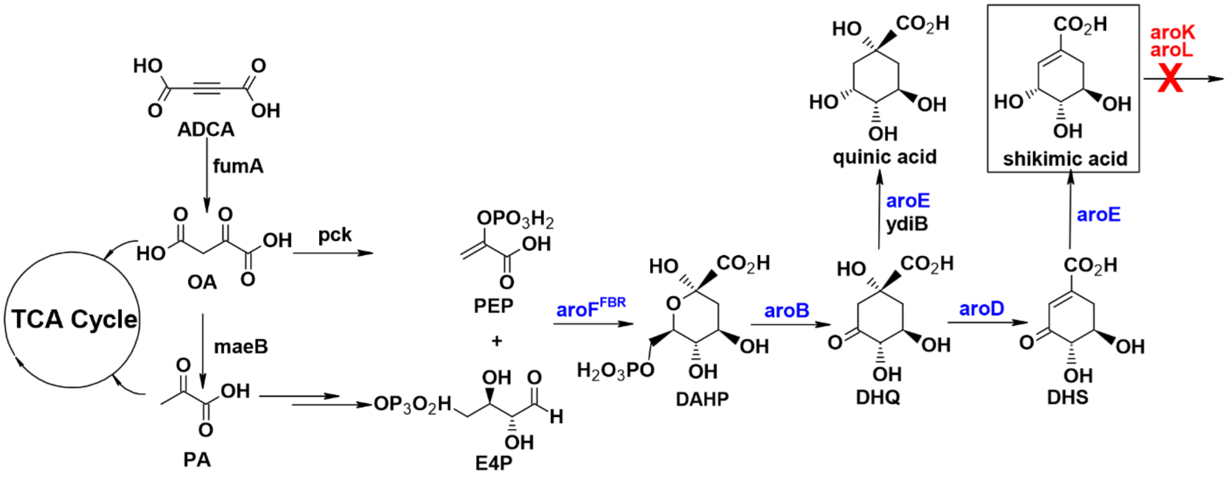


Figure 4.28. Route for shikimic acid synthesis from ADCA. FumA catalyzes hydration of ADCA to oxaloacetate, which is converted to PEP by Pck. E4P can be made available through series of steps from oxaloacetate beginning with conversion to pyruvate by MaeB

HPLC analysis of the culture supernatant after 200 h showed shikimic acid production at approximately 0.042 g L^{-1} when the cells reached the mid-stationary phase of growth. Additionally, quinic acid and DHS were detected as by-products at concentrations of 0.017 g L^{-1} and 0.025 g L^{-1}, respectively. Given the strain's limited growth and ADCA utilization, the low shikimate titers were anticipated. To enhance growth and ADCA utilization, the strain underwent ten rounds of evolution, resulting in the shikimate-producing variant ACG10 (SPACG10) (Figure 4.27). The selection of five colonies (SPACG10_1 to SPACG10_5) from the SPACG10 populations plated on M9A agar supplemented with aromatic amino acids and vitamins allowed further testing to identify better growing strains. Among these, two clones did not grow on ADCA; however, the others exhibited robust growth and complete ADCA consumption within 48 h. These strains also did not accumulate by-products quinic acid or DHS and exhibited enhanced shikimate accumulation, reaching 0.59 g L^{-1}, 0.34 g L^{-1}, and 0.60 g L^{-1} for SPACG10_3, SPACG10_4, and SPACG10_5, respectively.

The strain exhibiting the most promising performance, SPACG10_5, was selected for further optimization. Subsequent transformation with the pKD12.138A plasmid afforded evolved *E. coli* RB791 *serA*::*aroB*Δ*aroK*Δ*aroL*/pKD12.138A (SPACG10_5a). Cultivation under identical conditions with 3.5 g L^{-1} ADCA led to a slight increase in shikimate production of approximately 0.65 g L^{-1} by SPACG10_5a, demonstrating the feasibility of directing ADCA into the shikimate pathway for complex biomolecule synthesis.

Further optimization might involve overexpressing upstream enzymes of the shikimate pathway, and FumA to enhance ADCA conversion to oxaloacetate and subsequently to phosphoenolpyruvate (PEP). Additionally, increasing the availability of E4P, essential for growth on ADCA, might be achieved by overexpressing enolase (Eno) and phosphoglycerate kinase (Pgk).

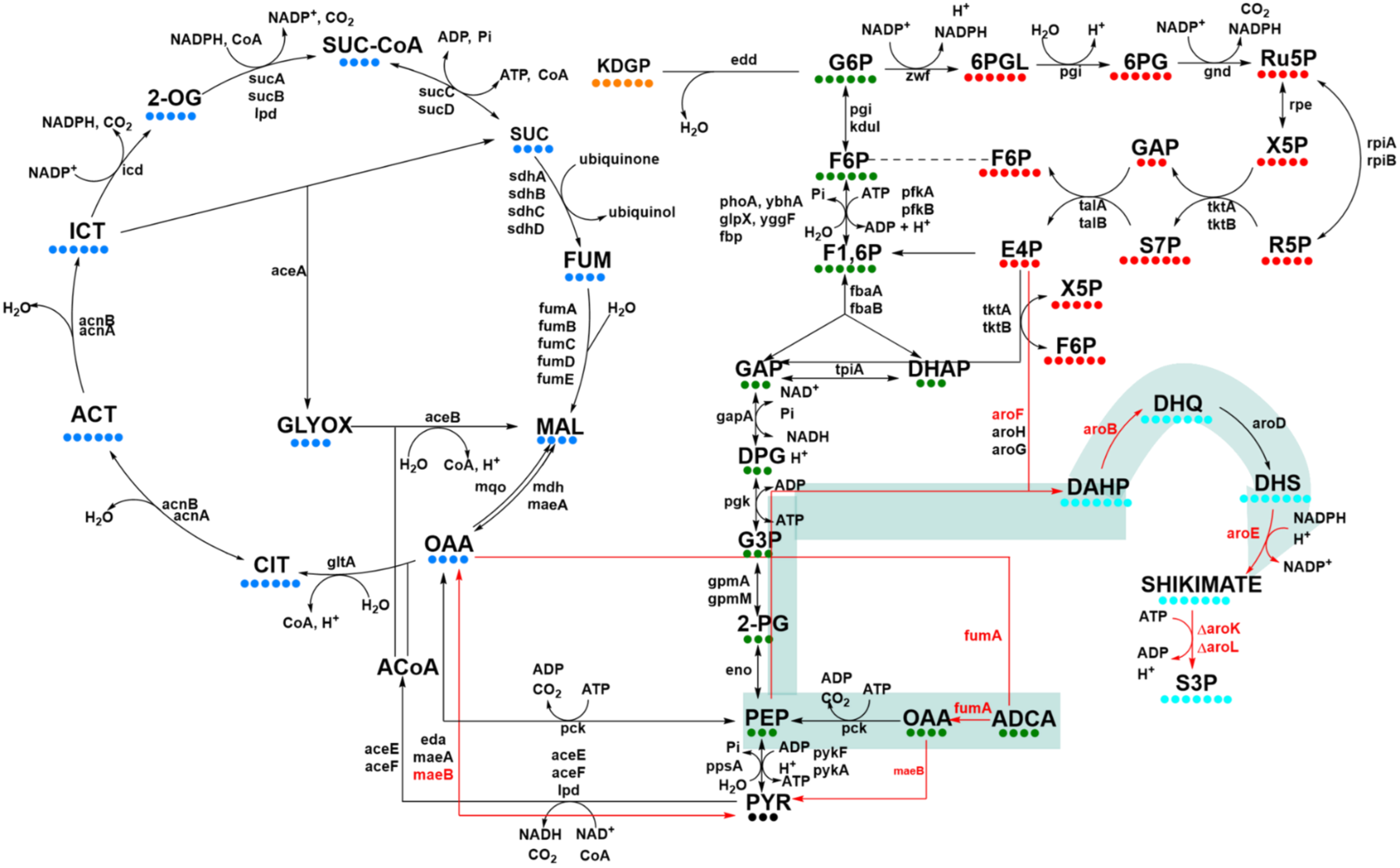


Figure 4.29. Metabolic model showing carbon flux from ADCA to shikimic acid. Number of colored dots indicates number of carbon atoms. The pathway shows flux to TCA cycle (blue), gluconeogenesis (green), and pentose phosphate pathway (red).

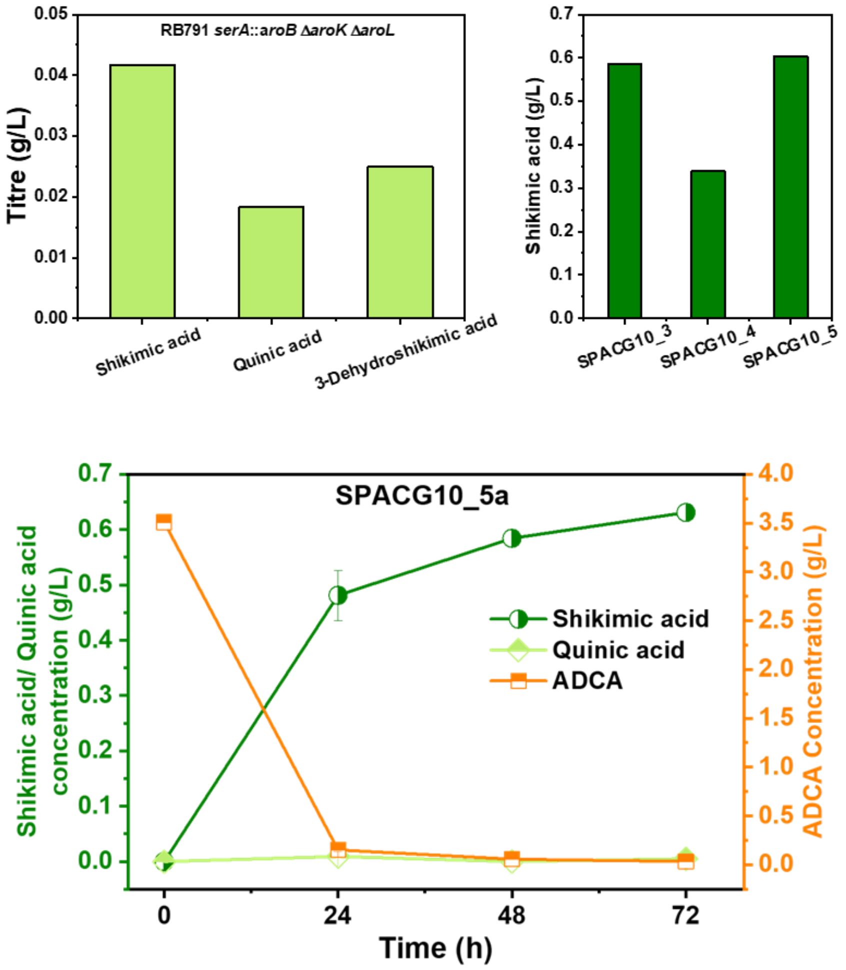


Figure 4.30. Shikimate titers from various approaches. Panels indicate titers for *E. coli* RB791 *serA*::*aroB*Δ*aroK*Δ*aroL* before evolution (top left), three clones of *E. coli* RB791 *serA*::*aroB*Δ*aroK*Δ*aroL* after evolution (top right), and evolved clone transformed with pKD12.138A.

Table 4.4. Summary table for *in vivo* production of shikimic acid from ADCA.

Entry	Strains	Time	ADCA		Shikimic acid			By-products	
			Initial	Consumed	Produced	Conversion	Productivity	Quinic acid	3-Dehydroshikimic acid
		h	mM	mM	mM	mol/mol	g/L/h	mM	mM
1	RB791 *serA*::*aroB* Δ*aroL* Δ*aroK*	200	40	16	0.24 ± 0.00	0.015	0.21	0.09	0.15
2	SPACG10	72	40	40	2.96 ± 0.63	0.074	7.13	n.d	n.d
3	SPACG10_5a	72	40	40	3.62 ± 0.00	0.091	8.75	n.d	n.d

4.4 Conclusion

In conclusion, this chapter highlights the potential of acetylenedicarboxylic acid (ADCA) as a carbon source for *E. coli* bioproduction of chemicals using metabolic engineering and adaptive evolution. Successful generation or demonstration of *E. coli* strains capable of utilizing ADCA to produce relevant compounds such as lactic acid and shikimic acid demonstrates the practical applicability and flexibility of microbial metabolic systems and their potential to utilize unnatural carbon sources like ADCA. The relevance of FumA and DcuC, identified through multi-omic studies, enabled their incorporation into the lactic acid biosynthetic strategy. The production of lactate serves as an important proof of concept even though further optimization of production strains could be achieved by targeting competing pathways such as phosphoenolpyruvate carboxylase (Ppc) and phosphate acetyltransferase (PTA). This approach is expected to increase intracellular pyruvate concentrations and thus increase lactate synthesis, although the impact of these modifications on microbial growth remains uncertain due to the essential role of specific genes in ADCA metabolism.

This work demonstrates the proof of concept for synthesizing two significant molecules: a chiral commodity chemical, which is a target of the biorefinery concept, and a chiral precursor for an antiviral drug. The next step involves identifying key mutations that enable growth on ADCA and transferring those mutations into the MG1655 strain. This will yield a strain capable of growing on ADCA without the additional mutations present in ACG strains.

Additionally, the ability to produce shikimic acid from ADCA opens up new possibilities for the biosynthesis of more complex molecules. Optimizing shikimate titers may involve overexpressing FumA and other upstream genes in the shikimate pathway to more effectively direct carbon flux toward this pathway. An unexpected finding was the plasmid curing observed in resistance adaptations during growth on ADCA. Using FumA as a selective marker significantly increased production titers. This novel insight could be leveraged in future fermentation processes to enhance plasmid stability.

Overall, this chapter highlights the innovative use of ADCA as a sustainable feedstock for producing relevant chemicals and suggests a way forward that involves careful strain and pathway optimization to fully exploit the metabolic capabilities of methane and carbon dioxide-derived feedstock ADCA. The results presented here not only advance our understanding of microbial metabolism in the context of unconventional carbon sources, but also lay the groundwork for future studies aimed at maximizing the economic and practical benefits of microbial biotechnology.

REFERENCES

(1) Singh, A.; Prajapati, P.; Vyas, S.; Gaur, V. K.; Sindhu, R.; Binod, P.; Kumar, V.; Singhania, R. R.; Awasthi, M. K.; Zhang, Z.; Varjani, S. A Comprehensive Review of Feedstocks as Sustainable Substrates for Next-Generation Biofuels. *Bioenergy Res.* **2023**, *16* (1), 105–122.

(2) Osman, A. I.; Mehta, N.; Elgarahy, A. M.; Al-Hinai, A.; Al-Muhtaseb, A. H.; Rooney, D. W. Conversion of Biomass to Biofuels and Life Cycle Assessment: A Review. *Environ. Chem. Lett.* **2021**, *19* (6), 4075–4118.

(3) Tang, C.; Zheng, Y.; Jaroniec, M.; Qiao, S.-Z. Electrocatalytic Refinery for Sustainable Production of Fuels and Chemicals. *Angew. Chem. Int. Ed Engl.* **2021**, *60* (36), 19572–19590.

(4) Carbon Dioxide Removal https://www.energy.gov/fecm/carbon-dioxide-removal-0 (accessed May 14, 2024).

(5) Sahoo, K. K.; Goswami, G.; Das, D. Biotransformation of Methane and Carbon Dioxide Into High-Value Products by Methanotrophs: Current State of Art and Future Prospects. *Front. Microbiol.* **2021**, *12*, 636486.

(6) Jeon, Y. C.; Nguyen, A. D.; Lee, E. Y. Bioproduction of Isoprenoids and Other Secondary Metabolites Using Methanotrophic Bacteria as an Alternative Microbial Cell Factory Option: Current Stage and Future Aspects. *Catalysts* **2019**, *9* (11), 883.

(7) Overview – Global Methane Tracker 2022 – Analysis https://www.iea.org/reports/global-methane-tracker-2022/overview (accessed Nov 5, 2022).

(8) Teixeira, L. V.; Moutinho, L. F.; Romão-Dumaresq, A. S. Gas Fermentation of C1 Feedstocks: Commercialization Status and Future Prospects. *Biofuels Bioprod. Biorefin.* **2018**, *12* (6), 1103–1117.

(9) Calvey, C. H.; Sànchez I Nogué, V.; White, A. M.; Kneucker, C. M.; Woodworth, S. P.; Alt, H. M.; Eckert, C. A.; Johnson, C. W. Improving Growth of *Cupriavidus Necator* H16 on Formate Using Adaptive Laboratory Evolution-Informed Engineering. *Metab. Eng.* **2023**, *75*, 78–90.

(10) Sarwar, A.; Lee, E. Y. Methanol-Based Biomanufacturing of Fuels and Chemicals Using Native and Synthetic Methylotrophs. *Synth Syst Biotechnol* **2023**, *8* (3), 396–415.

(11) Kim, Y.; Lama, S.; Agrawal, D.; Kumar, V.; Park, S. Acetate as a Potential Feedstock for the Production of Value-Added Chemicals: Metabolism and Applications. *Biotechnol. Adv.* **2021**, *49*, 107736.

(12) Fincke, J. R.; Anderson, R. P.; Hyde, T.; Detering, B. A.; Wright, R.; Bewley, R. L.; Haggard, D. C.; Swank, W. D. Plasma Thermal Conversion of Methane to Acetylene. *Plasma Chem. Plasma Process.* **2002**, *22* (1), 105–136.

(13) Dinh, D. K.; Lee, D. H.; Song, Y.-H.; Jo, S.; Kim, K.-T.; Iqbal, M.; Kang, H. Efficient Methane-to-Acetylene Conversion Using Low-Current Arcs. *RSC Adv.* **2019**, *9* (56), 32403–32413.

(14) Heintze, M.; Magureanu, M. Methane Conversion into Acetylene in a Microwave Plasma: Optimization of the Operating Parameters. *J. Appl. Phys.* **2002**, *92* (5), 2276–2283.

(15) Shen ChangSheng; Sun YongZhi; Sun DeKun. A Study on Methane Coupling to Acetylene under the Microwave Plasma. *Science China-Chemistry* **2010**, *53* (1), 231–237.

(16) Manjolinho, F.; Arndt, M.; Goossen, K.; Goossen, L. J. Catalytic C-H Carboxylation of Terminal Alkynes with Carbon Dioxide. *ACS Catal.* **2012**, *2* (9), 2014–2021.

(17) Wang, X.; Lim, Y. N.; Lee, C.; Jang, H.-Y.; Lee, B. Y. 1,5,7-Triazabicyclo[4.4.0]Dec-1-Ene-Mediated Acetylene Dicarboxylation and Alkyne Carboxylation Using Carbon Dioxide. *Eur. J. Org. Chem.* **2013**, *2013* (10), 1867–1871.

(18) Mathes Hewage, A. N. S. Synthesis of 3-Hydroxypropionic Acid from Acetylenecarboxylic Acid, Michigan State University, Ann Arbor, United States, 2022.

(19) Kumar, A.; Singh, J.; Baskar, C. Lactic Acid Production and Its Application in Pharmaceuticals. **2020**. https://doi.org/10.1007/978-981-15-1394-7_16.

(20) Alexandri, M.; Schneider, R.; Mehlmann, K.; Venus, J. Recent Advances in D-Lactic Acid Production from Renewable Resources: Case Studies on Agro-Industrial Waste Streams. *Food Technol. Biotechnol.* **2019**, *57* (3), 293–304.

(21) Eş, I.; Mousavi Khaneghah, A.; Barba, F. J.; Saraiva, J. A.; Sant'Ana, A. S.; Hashemi, S. M. B. Recent Advancements in Lactic Acid Production - a Review. *Food Res. Int.* **2018**, *107*, 763–770.

(22) Hafid, H. S.; Nor 'Aini, A. R.; Mokhtar, M. N.; Talib, A. T.; Baharuddin, A. S.; Umi Kalsom, M. S. Over Production of Fermentable Sugar for Bioethanol Production from Carbohydrate-Rich Malaysian Food Waste via Sequential Acid-Enzymatic Hydrolysis Pretreatment. *Waste Manag.* **2017**, *67*, 95–105.

(23) Wang, P.; Feng, J.; Guo, L.; Fasina, O.; Wang, Y. Engineering *Clostridium Saccharoperbutylacetonicum* for High Level Isopropanol-Butanol-Ethanol (IBE) Production from Acetic Acid Pretreated Switchgrass Using the CRISPR-Cas9 System. *ACS Sustain. Chem. Eng.* **2019**, *7* (21), 18153–18164.

(24) Huo, Z.; Fang, Y.; Yao, G.; Zeng, X.; Ren, D.; Jin, F. Improved Two-Step Hydrothermal Process for Acetic Acid Production from Carbohydrate Biomass. *J. Energy Chem.* **2015**, *24* (2), 207–212.

(25) Silva, A. P. M. da; Sica, P.; Pires, L. de A. N.; Spironello, L.; Mota, L. A.; Peixoto, G. T.; Calegari, R. P.; Basso, T. O.; Tonso, A.; Gomes, M. P.; Somessari, S. L.; Duarte, H. G.; Somessari, E. S. R.; Carvalho, R. de S.; Baptista, A. S. Integration of Corn and Cane for

Ethanol Production: Effects of Lactobacilli Contamination on Fermentative Parameters and Use of Ionizing Radiation Treatment for Disinfection. *Fermentation* **2023**, *9* (2), 89.

(26) McFarlane, J.; Robinson, S. M. *Survey of Alternative Feedstocks for Commodity Chemical Manufacturing*; Office of Scientific and Technical Information (OSTI), 2008. https://doi.org/10.2172/931080.

(27) Francois, J. M.; Alkim, C.; Morin, N. Engineering Microbial Pathways for Production of Bio-Based Chemicals from Lignocellulosic Sugars: Current Status and Perspectives. *Biotechnol. Biofuels* **2020**, *13*, 118.

(28) Morrow, G. W. The Shikimate Pathway: Biosynthesis of Phenolic Products from Shikimic Acid. **2016**. https://doi.org/10.1093/oso/9780199860531.003.0009.

(29) Bochkov, D. V.; Sysolyatin, S. V.; Kalashnikov, A. I.; Surmacheva, I. A. Shikimic Acid: Review of Its Analytical, Isolation, and Purification Techniques from Plant and Microbial Sources. *J. Chem. Biol.* **2012**, *5* (1), 5–17.

(30) Fujiwara, R.; Noda, S.; Tanaka, T.; Kondo, A. Metabolic Engineering of *Escherichia coli* for Shikimate Pathway Derivative Production from Glucose-Xylose Co-Substrate. *Nat. Commun.* **2020**, *11* (1), 279.

(31) Singh, P.; Gupta, E.; Mishra, N.; Mishra, P. Shikimic Acid as Intermediary Model for the Production of Drugs Effective against Influenza Virus. *Phytochemicals as Lead Compounds for New Drug Discovery* **2020**, 245.

(32) Seca, A. M. L.; Pinto, D. C. G. A. Plant Secondary Metabolites as Anticancer Agents: Successes in Clinical Trials and Therapeutic Application. *Int. J. Mol. Sci.* **2018**, *19* (1). https://doi.org/10.3390/ijms19010263.

(33) Li, X.; Mo, K.; Tian, G.; Zhou, J.; Gong, J.; Li, L.; Huang, X. Shikimic Acid Regulates the NF-KB/MAPK Signaling Pathway and Gut Microbiota to Ameliorate DSS-Induced Ulcerative Colitis. *J. Agric. Food Chem.* **2023**, *71* (23), 8906–8914.

(34) Draths, K. M.; Knop, D. R.; Frost, J. W. Shikimic Acid and Quinic Acid: Replacing Isolation from Plant Sources with Recombinant Microbial Biocatalysis. *J. Am. Chem. Soc.* **1999**, *121* (7), 1603–1604.

(35) Sakhuja, D.; Ghai, H.; Rathour, R. K.; Kumar, P.; Bhatt, A. K.; Bhatia, R. K. Cost-Effective Production of Biocatalysts Using Inexpensive Plant Biomass: A Review. *3 Biotech* **2021**, *11* (6). https://doi.org/10.1007/s13205-021-02847-z.

(36) Sóti, V.; Lenaerts, S.; Cornet, I. Of Enzyme Use in Cost-Effective High Solid Simultaneous Saccharification and Fermentation Processes. *J. Biotechnol.* **2018**, *270*, 70–76.

(37) van Vugt-Lussenburg, B. M. A.; van der Weel, L.; Hagen, W. R.; Hagedoorn, P.-L. Biochemical Similarities and Differences between the Catalytic [4Fe-4S] Cluster

Containing Fumarases FumA and FumB from *Escherichia coli*. *PLoS One* **2013**, *8* (2), e55549.

(38) Flint, D. H. Escherichia Coli Fumarase A Catalyzes the Isomerization of Enol and Keto Oxalacetic Acid. *Biochemistry* **1993**, *32* (3), 799–805.

(39) Flint, D. H.; McKay, R. G. Escherichia Coli Fumarase A Catalyzed Transfer of 18O from C-2 and 2H from C-3 of Malate to Acetylene Dicarboxylate To Form 18O and 2H Labeled Oxalacetate. *J. Am. Chem. Soc.* **1994**, *116* (13), 5534–5539.

(40) Schubert, C.; Unden, G. C4-Dicarboxylates as Growth Substrates and Signaling Molecules for Commensal and Pathogenic Enteric Bacteria in Mammalian Intestine. *J. Bacteriol.* **2022**, *204* (4), e0054521.

(41) Janausch, I. G.; Zientz, E.; Tran, Q. H.; Kröger, A.; Unden, G. C4-Dicarboxylate Carriers and Sensors in Bacteria. *Biochim. Biophys. Acta* **2002**, *1553* (1–2), 39–56.

(42) Lázár, V.; Nagy, I.; Spohn, R.; Csörgő, B.; Györkei, Á.; Nyerges, Á.; Horváth, B.; Vörös, A.; Busa-Fekete, R.; Hrtyan, M.; Bogos, B.; Méhi, O.; Fekete, G.; Szappanos, B.; Kégl, B.; Papp, B.; Pál, C. Genome-Wide Analysis Captures the Determinants of the Antibiotic Cross-Resistance Interaction Network. *Nat. Commun.* **2014**, *5*, 4352.

(43) Kohanski, M. A.; DePristo, M. A.; Collins, J. J. Sublethal Antibiotic Treatment Leads to Multidrug Resistance via Radical-Induced Mutagenesis. *Mol. Cell* **2010**, *37* (3), 311–320.

(44) Driscoll, A. J.; Bhat, N.; Karron, R. A.; O'Brien, K. L.; Murdoch, D. R. Disk Diffusion Bioassays for the Detection of Antibiotic Activity in Body Fluids: Applications for the Pneumonia Etiology Research for Child Health Project. *Clin. Infect. Dis.* **2012**, *54 Suppl 2*, S159-64.

CHAPTER FIVE: Experimental

5.1 Materials and equipment

Culture media components and buffer salts were obtained from Becton, Dickinson and Company, Millipore Sigma USA, and Thermo Fisher Scientific. Organic acids including acetylenedicarboxylic acid (ADCA), fumaric acid, malic acid, oxaloacetic acid, pyruvic acid, and D- or L-lactic acid were obtained from Millipore Sigma. Isopropyl-β-D-thiogalactopyranoside (IPTG) was obtained from Gold Biotechnology. Bio-Rad Protein Assay Dye Reagent Concentrate, Precision Plus Protein Electrophoresis Standards, and Mini-PROTEAN® TGX™ Precast 4-20% Polyacrylamide Gels were sourced from Bio-Rad Laboratories. Molecular cloning and DNA modification enzymes including restriction enzymes, T4 DNA Ligase, Gibson Assembly® Cloning Kit, NEBuilder® HiFi DNA Assembly Cloning Kit, and Q5® Site-Directed Mutagenesis Kit were purchased from New England Biolabs. QIAprep Spin Miniprep and Maxiprep Kits were purchased from Qiagen. Genomic DNA was purified with the Wizard® Genomic DNA Purification Kit from Promega. Optical density measurements at 600 nm (OD_{600}) of cell cultures were taken using a NanoDrop™ One Microvolume UV-Vis Spectrophotometer.

Protein purification was conducted using pre-packed HisTrap™ FF (1 mL and 5 mL) columns on ÄKTA Start FPLC from Cytiva, and molecular weight cut-off filtration was performed with Amicon Ultra-15 30K Centrifugal Filter Units. Genomic DNA purification kits were obtained from Promega and oligonucleotides were purchased from Integrated DNA Technologies (IDT). Enzymatic assays and specific reactions utilized aldehyde dehydrogenase (CAS Number 9028-88-0) from *Saccharomyces cerevisiae* and L-lactate dehydrogenase (CAS Number 9001-60-9) from rabbit muscle, obtained from Millipore Sigma. Enzyme assays and cell growth monitoring were carried out using Molecular Devices SpectraMax® iD3 multi-mode microplate reader. Standard assays were conducted using a Grenier UV-Star® clear 96-well plates (catalog number M3812) and Corning costar black clear-bottomed 96-well plates (catalog number CLS3603) for NADH, NAD^+, and ATP quantification. ATP quantification kit (catalog number MAK433) and NAD^+/NADH quantification kit (catalog number MAK037) were obtained from Millipore Sigma. NMR spectra analysis was performed using a 500 MHz Varian NMR spectrophotometer. Metabolomic analysis was performed using a GC-Time-of Flight MS (Leco GC-HRT+). All HPLC analysis was carried out using the Agilent 1100 HPLC system with an Aminex HPX-87H column. PCR amplifications utilized a Bio-Rad DNA Engine® Peltier Thermal Cycler.

5.2 Media preparation and stock solutions

Culture media were prepared using double-deionized water. All media were sterilized by autoclaving solutions at 121 °C for 25 min. LB media (1 L) included 10 g of tryptone, 5 g of yeast extract, and 10 g of NaCl. All 2x Yeast-tryptone (2xYT, 1L) media contained 16 g of tryptone, 10 g of yeast extract, and 5 g of NaCl. SOB medium (1L) contained 20 g of tryptone, 5 g of yeast extract, 0.5 g of NaCl, and 10 mL of 250 mM KCl were mixed, and the pH was adjusted to 7.0 using 10 N NaOH. Sterile 2 M $MgCl_2$ (5 mL) was added immediately before use. SOC media was prepared by adding 20 mL of 1 M D-glucose to SOB medium. M9 salts (1L) contained 6.0 g of Na_2HPO_4, 3.0 g of KH_2PO_4, 1.0 g of NH_4Cl, and 0.5 g of NaCl. The complete M9 medium was prepared by addition of 20% (w/v) glucose (20 mL), 1 M $MgSO_4$ (2 mL), and thiamine hydrochloride (1 mg mL^{-1}, 1 mL) to 1 L of M9 salts When supplementation with L-serine was necessary, L-serine (40 mg mL^{-1}, 1 mL) was added per L of medium.

Stock solutions of 2 M $MgCl_2$, 1 M $MgSO_4$, 1 M glucose, and 20% (w/v) glucose, 0.1 M fumaric acid, 1 M ADCA, 0.1 M malic acid, 0.1 M pyruvic acid, 0.1 M oxaloacetic acid, and 0.1 M lactic acid were separately prepared in deionized water and sterilized by filtering with a 0.22

μm filter. The media used in this study were prepared as follows: M9G, M9F, and M9A refer to mineral salt minimal media containing glucose, fumarate, or acetylenedicarboxylate (ADCA), respectively. Unless otherwise stated, the substrate concentration in these media was 20 mM for glucose and 40 mM for fumarate and ADCA. PowerTrack™ SYBR™ Green Master Mix (Cat. No. A46110) was obtained from Thermo Fisher. Solutions for RNA isolation and analysis were prepared as follows: 1 M NaOH (4 g of pellets stirred into 100 mL ddH$_2$O). Stock solutions of antibiotics, thiamine hydrochloride (1 mg mL^{-1}), L-serine (40 mg mL^{-1}), and isopropyl β-D-1-thiogalactopyranoside (IPTG, 1M) were prepared in deionized water and sterilized using 0.22 μm syringe filters. Antibiotics were added to media where neccessary, achieving final concentrations of 50 μg mL^{-1} for ampicillin (Ap) and kanamycin (Kan). Solid media contained either 1.5% (w/v) Agar (rich media) or 1.5% (w/v) low EEO agarose (minimal media).

5.3 Preparation and transformation of competent cells

Electrocompetent cells were prepared following standard operating procedures adapted from Sambrook and Russell.[1] A single colony of the desired bacterial strain was obtained from a freshly streaked plate (< 3 days old), and was used to inoculate 5 mL of LB medium. This culture was incubated overnight at 37 °C with agitation at 200 rpm. Overnight culture (2 mL) was used to inoculate 100 mL 2xYT medium in a 500 mL baffled shake flask. The culture was incubated under the same conditions until reaching an OD$_{600}$ of 0.5 - 0.7. Cells were harvested by centrifugation (6,000 x *g*, 10 min, 4 °C) and gently resuspended in 100 mL of cold, sterile deionized water to wash away residual salts. After a repeat centrifugation under identical conditions, the supernatant was decanted, and the washing step was performed again with another 100 mL of cold, sterile deionized water. The cell pellet was then resuspended in 100 mL of cold, sterile 10% (v/v) glycerol, centrifuged (6,000 x *g*, 5 min, 4 °C), and the supernatant was discarded. The cell pellet was resuspended in 0.5 mL of 10% (v/v) glycerol and aliquoted (50 µL) volumes into pre-chilled sterile microcentrifuge tubes and flash-frozen in liquid nitrogen, and stored at -80 °C.

For transformation, 2 µL of plasmid DNA (1-5 ng/µL in sterile deionized water) or similarly prepared PCR or ligation products were combined with 50 µL of thawed electrocompetent *E. coli* cells. The mixture was gently pipetted to ensure homogeneity and transferred to a pre-chilled, sterile Gene Pulser® electroporation cuvette (0.2 cm electrode gap). Electroporation was performed using a Bio-Rad Gene Pulser II system set to 2.5 kV, 25 µF, and 200 Ω, achieving a time constant of 5.16 – 5.25 ms. Immediately following electroporation, the cuvette was cooled on ice, and 1 mL of SOC medium (RT) was added. The recovery culture was incubated at 37 °C with shaking (200 rpm) for 1 h. Cells were collected by centrifugation (14,000 rpm, 2 min, RT) in a microcentrifuge, and 800 µL of the supernatant was discarded. The cell pellet was resuspended in the residual medium, and aliquots of 20 µL and 80 µL were spread onto LB agar plates containing the appropriate antibiotic. Plates were incubated at 37 °C for 12 – 16 h.

5.4 Genetic manipulation

5.4.1 Plasmid DNA isolation

Plasmid DNA was purified from *E. coli* cultures using small-scale (mini) and large-scale (midi) protocols tailored to experimental requirements. To begin, a single colony of freshly transformed *E. coli* was selected from an agar plate containing the appropriate antibiotic. For small-scale purification, this colony was inoculated into 5 mL of LB medium containing the appropriate antibiotic and incubated overnight at 37 °C with shaking at 200 rpm. Cells from this culture were harvested by centrifugation (13,000 rpm, 5 min, RT). For large-scale plasmid purification, a single colony was used to inoculate 100 mL of LB medium containing the appropriate antibiotic, followed by overnight incubation under identical conditions. The cells were then collected by centrifugation (6,000 x *g*, 15 min, RT).

Plasmid isolation was performed using the QIAGEN Plasmid Miniprep and Midi prep Kits following protocols provided by the manufacturer. For the small-scale procedure, plasmid DNA was eluted from the spin column using 50 μL of sterile, deionized water. For the large-scale purification, the air-dried plasmid pellet was resuspended in 100 – 200 μL of sterile, deionized water. Purified plasmids were stored at 4 °C.

5.4.2 Isolation of genomic DNA

Genomic DNA was isolated from *Escherichia coli* MG1655 or W3110 using the Wizard® Genomic DNA Purification Kit and following manufacturer's protocol. A single bacterial colony was cultured in 5 mL of LB medium and incubated at 37 °C with shaking for 12 h. Cells were pelleted by centrifugation (14,000 x *g*, 2 min, 4 °C). Culture supernatant removed and the resulting cell pellet was subjected to lysis according to the manufacturer's instruction. Briefly, the lysate was then processed through subsequent steps as outlined in the kit's protocol to purify the genomic DNA, effectively removing proteins, RNA, and other impurities. The purified genomic DNA was quantified and assessed for purity using spectrophotometric methods, ensuring that the DNA was of high quality and suitable for downstream molecular biology applications.

5.4.3 Restriction digest of DNA

A standard restriction digest protocol was employed to confirm the identity of purified plasmid DNA, each restriction digest (20 μL) contained 500 – 1000 ng of plasmid DNA, 1x specific buffer, and 10 U of restriction enzyme(s). For double digests, 1 μL of the two required restriction enzymes was added to the mixture. When a larger quantity of digested DNA was necessary for cloning or DNA assembly reactions, the reaction volume was increased to 50 μL.

Restriction digests were incubated at 37 °C for 1 h to ensure complete cleavage of the DNA. Reactions were terminated by incubation at 60 °C for 20 min followed by addition of 6X loading dye (4 µL) which included SDS. Digested DNA fragments were visualized using electrophoresis on a 0.7 – 1.0 % agarose gel, stained with ethidium bromide (0.5 µg/mL) to allow for DNA detection under UV light. The gel was submerged in 1x TAE (Tris-Acetate EDTA) buffer, which also served as the running buffer. Electrophoresis was conducted at a constant voltage of 95 V for 1 - 2 h(s), for optimal separation of DNA fragments.

5.4.4 Cloning of DNA

Insertion of a specific DNA fragment into a plasmid vector was performed using T4 ligase following the manufacturer's protocol. Ligation reactions typically use a vector: insert ratio 1:5 or 1:3. Electrocompetent cells were transformed with 5 µL of ligation reaction. Transformed cells were plated on LB agar plates containing the appropriate antibiotic. Colonies were cultured in liquid LB medium containing the appropriate antibiotic. Plasmid DNA was extracted and purified using a plasmid miniprep kit, following the manufacturer's instructions. Isolated plasmid was sequenced at the Michigan State University Research Technology Support Facility (MSU RTSF) Genomics Core using custom primers via Sanger sequencing methodology.

5.4.5 PCR amplification

PCR amplifications (20 - 50 µL) were performed in a Bio-Rad DNA Engine® Peltier Thermal Cycler. Reactions utilized Q5® High-Fidelity DNA Polymerase following PCR program described by manufacturer. Following amplification, loading dye (1x) containing SDS (from New England Biolabs) was added to each PCR sample to prepare for gel analysis. The samples were then visualized on a 0.7% agarose gel to assess the PCR products.

5.4.6 Sanger sequencing of plasmids

DNA plasmids generated in this study were verified via Sanger sequencing. The DNA was initially confirmed by appropriate restriction digestion gel electrophoresis and candidate plasmids were sequenced at the MSU RTSF Genomic Core. This sequencing employed custom primers specifically designed to span 600 base pairs of targets, enabling complete sequencing of both strands of each plasmid. The sequencing outputs from MSU RTSF were then analyzed using the "Align imported sequence to reference sequence" function in the latest version of SnapGene software.

5.5 Protein overexpression and purification

Multiple proteins were overexpressed with an N-terminal His_6-tag, which facilitated their purification from cellular lysate. Protein production was executed in *E. coli* BL21(DE3) employing a T7 promoter system using the pET-15b expression vectors. Transformation of BL21(DE3) with plasmids encoding His_6-tagged proteins was followed by the selection of a single colony. This colony was used to inoculate 50 mL of LB medium containing ampicillin, and the culture was grown at 37 °C with agitation at 200 rpm for 12 - 16 h. For small-scale protein expression, this overnight culture was used to inoculate 250 mL of LB medium containing ampicillin, starting at an OD_{600} of 0.05. For large-scale expression, the overnight culture was used to inoculate 1 L of LB medium containing ampicillin to a start OD_{600} of 0.05. These cultures were incubated under at 37 °C, 200 rpm, until an OD_{600} of 0.5 - 0.7 was reached, at which point IPTG was added to a final concentration of 0.1 mM. Cultures were incubated at 30 °C for 8 - 10 h with continuous shaking at 200 rpm. Following the induction period, cells were harvested by centrifugation (12,000 x *g*, 30 min, 4 °C). Cell pellets were stored at -20 °C pending protein purification.

Cell pellets were thawed at room temperature and resuspended in lysis buffer (20 mM sodium phosphate, pH 7.4, containing 20 mM imidazole) (2 mL /g of cell paste). Lysis was achieved by two passages through a SLM Aminco French® Pressure cell (Thermo Scientific, Waltham, MA) at gauge pressure of 1,100. The lysate was then clarified by centrifugation (47,500 x *g*, 4 °C, 30 min), and the supernatant was filtered through a 0.45 μm sterile syringe filter.

The protein purification protocol consisted of five stages: column equilibration, sample application, washing, elution and fractionation, and post-elution equilibration carried out with two buffers. Binding buffer contains 20 mM sodium phosphate, pH 7.4, 500 mM NaCl and elution buffer contains 20 mM sodium phosphate, pH 7.4, 500 mM NaCl, 500 mM imidazole. Except during elution and fractionation, a mix of 96% binding buffer and 4% elution buffer was used to maintain an imidazole concentration consistent with the lysis buffer (20 mM). The HisTrap FF column was equilibrated with five column volumes at a flow rate of 4.0 mL/min. Sample application proceeded at 1.0 mL/min, followed by a wash with 10 column volumes at 4.0 mL/min to remove unbound proteins. His-tagged proteins were eluted using an imidazole gradient from 20 mM to 500 mM over 10 - 20 column volumes. Fractions containing the target protein were pooled (as determined by the UV signal at 280 nm), concentrated, and desalted using Amicon Ultra-15 30K filters, then stored in 100 mM sodium phosphate, pH 8.0, with 20% glycerol, unless specified otherwise. Typical yields ranged from 50 to 200 mg of enzyme per liter of cell culture. The effectiveness of each purification step was monitored by SDS-PAGE, which analyzes cell lysate and the final purified enzyme.

5.6 Protein quantification

Protein concentrations in cell lysates were determined using the Bradford Assay, with the reagent supplied by Bio-Rad. Cell lysate (10 μL) was diluted with 90 μL of deionized water, then mixed with 5 mL of Bradford reagent at room temperature. A control was similarly prepared, consisting of 100 µL deionized water and 5 mL Bradford reagent. Following incubation at room temperature for 5 min, the optical density at 595 nm (OD_{595}) was measured. The concentration of the enzyme in the samples was determined using a standard curve generated with bovine serum albumin (BSA) (Figure 5.1).

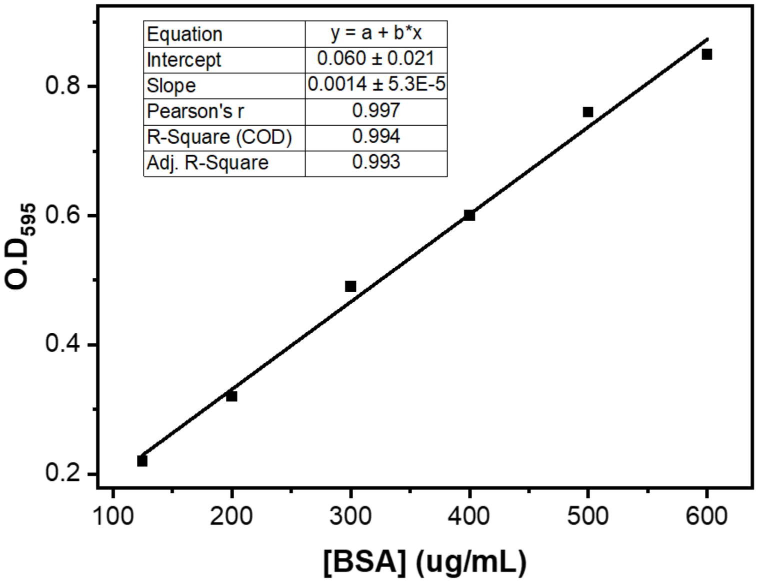


Figure 5.1. Calibration curve for Bradford assay.

Purified enzymes were quantified using guanidinium hydrochloride. Purified enzyme (10 µL) was diluted into 990 µL of 6.0 M guanidinium-hydrochloride. A control sample was prepared by diluting 10 µL of the storage buffer for the enzyme into 990 µL of the same solution. These samples were incubated at room temperature for 5 min before the absorbance at 280 nm (OD_{280}) was measured. Enzyme concentrations were calculated using their respective molar extinction coefficients at 280 nm and molecular weights, factoring in the His-tag, with calculations performed using Protein Calculator v3.46.

5.7 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Purified proteins were subjected to SDS-PAGE analysis using 4 – 20% (w/v) acrylamide Mini-PROTEAN® TGX™ precast gels (Bio-Rad). Prior to electrophoresis, each protein sample was quantified as described above and adjusted to stock solutions of 300 ng/μL in their respective storage buffers. For sample preparation, 2X Laemmli Buffer containing 4% (w/v) SDS, 20% (w/v) glycerol, 0.004% (w/v) bromophenol blue, and 125 mM Tris-HCl (pH 6.8) was mixed with 200 μL of 2 M DTT to achieve a final DTT concentration of 400 mM. This mixture (50 μL) was then combined with an equal volume of the protein sample and heated at 98 °C for 5 min to ensure complete denaturation. The samples were centrifuged at 17,000 x g for 2 min to pellet any insoluble material. For the electrophoresis, 30 μL of each prepared protein sample and 10 μL of Precision Plus Protein™ All Blue Ladder (Bio-Rad) were loaded onto the gel. The SDS-PAGE was performed in a running buffer (pH 8.3) containing 0.1% (w/v) SDS, 192 mM L-glycine, and 25 mM Tris base, applying a constant voltage of 200 V for 35 - 40 min. Following electrophoresis, gels were removed from their cassettes and stained with a solution of 0.1% (w/v) Coomassie Brilliant Blue R, 45% (v/v) methanol, and 10% (v/v) acetic acid in water for 1.5 – 2 h to visualize the protein bands. Destaining was then performed over 6 – 12 h in a solution containing 45% (v/v) methanol and 10% (v/v) acetic acid in water, until clear bands were observed against a transparent background.

5.8 Chapter Two: STALE *E. coli* on acetylenedicarboxylic acid

5.8.1 Materials and solutions

To make 0.1% DEPC, 100 μL of diethylpyrocarbonate (DEPC) in 1 L ddH$_2$O and stirred vigorously. The solution was incubated for 12 h at 37 °C. It was then autoclaved for 21 min to remove any trace of DEPC. DEPC reacts with primary amines. EDTA solution (0.5 M, pH 8) was prepared by adding 18.6 g of disodium EDTA.2H_2O to 80 mL of ddH_2O. The solution was stirred vigorously, and the pH was adjusted to 8.0 by addition of NaOH. The volume was adjusted to 100 mL. Tris-HCl (1 M, pH 8) was prepared by adding 12.11 g Tris base to 80 mL of ddH_2O followed by addition of concentrated HCl to adjust the pH to 8.0. The volume was adjusted to 100 mL by addition of ddH_2O. Tris-EDTA (TE) buffer (1 M, pH 8) was prepared by adding 2 mL of EDTA (0.5 mM, pH 8) and 10 mL of Tris-HCl (1 M, pH 8) to ddH_2O to a 1 L final volume. The solution was autoclaved to sterilize on a liquid cycle. Ethanol (70%) was made by adding 70 mL of ethanol (100%) to 30 mL ddH_2O. MOPS (3-(N-morpholino)propanesulfonic acid) (1 M, pH 7) was prepared by adding 41.86 g of MOPS free acid, 4.1 g of sodium acetate, 3.72 g of Na_2EDTA to 800 mL of ddH_2O. The solution was stirred vigorously, and the pH was adjusted using NaOH to 7. The volume was adjusted to 1 L with ddH_2O. Sodium acetate (1 M, pH 5.2) was made by dissolving 4.1 g sodium acetate in 40 mL RNase-free ddH_2O. The pH of the solution was adjusted to 5.2 with glacial acetic acid. The volume was adjusted to 50 mL with RNase-free ddH_2O. Formaldehyde (FA) gel buffer 10x (pH 7) was made by combining MOPS (200 mM), sodium acetate (50 mM), EDTA (10 mM), and ddH_2O to a final 1 L volume. The pH was adjusted to 7 with NaOH. The working concentration of FA gel running buffer (1X) was made by combining 100 mL 10X FA gel buffer, 20 mL 37% (12.3 M) formaldehyde, and 880 mL RNase-free ddH_2O. RNA loading buffer 5X was prepared by combining 16 µL saturated aqueous bromophenol blue solution, 80 µLof 500 mM EDTA (pH 8), 720 µL of 37% (12.3 M) formaldehyde, 2 mL of 100% glycerol, 3.084 mL of formamide, 4 mL of 10X FA gel buffer and RNase-free water to 10 mL. Lysozyme (15 mg mL^{-1}) was made by dissolving 60 mg lysozyme in 4 Ml TE buffer.

5.8.2 Bacterial strains and plasmids

Escherichia coli strains BL21(DE3), MG1655, and DH5α were obtained from Invitrogen. These cells were cultured at 37 °C in LB medium or specified mineral salts media, supplemented with the relevant antibiotic as required. Plasmid pET-15b was acquired from Invitrogen. Strains and plasmids used in Chapter Two are listed in Tables 5.1. Primers are listed in Table 5.2.

5.8.3 Growth of *E. coli* MG1655 and RB791 on M9G, M9F, and M9A

Growth of *E. coli* MG1655 and RB791 on minimal agar plates containing glucose, fumarate, or ADCA was assessed as follows. Five microliters of an overnight culture of each strain grown in M9G was diluted into M9 salts without any carbon source to achieve desired dilutions (1:10^{-5} and 1:10^{-6}). A new pipet tip is used for each transfer. Approximately 100 µL of culture from each tube is plated onto separate labeled dry agar plates and incubated at 37°C overnight or until colonies appear. Plates are stored for future use by wrapping in parafilm and storing at 4°C. Single colonies can also be obtained by streaking. Inoculum is streaked down one side of a plate using a sterile biochemical technique. The inoculation loop is sterilized or replaced and used to streak a sample from the first streak across a fresh portion of the plate. This is repeated until the plate is covered. The plate is then incubated at 37°C until colonies appear.

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

Strains	**Genotype/ Description**	**Reference/ Source**
MG1655	F^{-} λ^{-} *ilvG*$^{-}$ *rfb*-50 *rph*-1	ATCC
RB791	*lacIp-4000(lacI*^{Q}*),* λ^{-}*, IN(rrnD-rrnE)1*	ATCC
ACG1	MG1655 after round 1 of evolution in ADCA (ADCA Consuming Generation 1)	This study
ACG2	MG1655 after round 2 of evolution in ADCA (ADCA Consuming Generation 2)	This study
ACG3	MG1655 after round 3 of evolution in ADCA (ADCA Consuming Generation 3)	This study
ACG1_1	ACG1 of lineage 1	This study
ACG1_2	ACG1 of lineage 2	This study
ACG1_3	ACG1 of lineage 3	This study
ACG2_1	ACG2 of lineage 1	This study
ACG2_2	ACG2 of lineage 2	This study
ACG2_3	ACG2 of lineage 3	This study
ACG3_1	ACG3 of lineage 1	This study
ACG3_2	ACG3 of lineage 2	This study
ACG3_3	ACG3 of lineage 3	This study
ACG3_4_1	ACG3 of lineage 4 clone 1	This study
ACG3_4_3	ACG3 of lineage 4 clone 3	This study
ACG3_4_4	ACG3 of lineage 4 clone 4	This study
ACG3_5_1	ACG3 of lineage 5 clone 1	This study
ACG3_5_3	ACG3 of lineage 5 clone 3	This study
ACG3_5_10	ACG3 of lineage 5 clone 10	This study
ACG3_6_2	ACG3 of lineage 6 clone 2	This study
ACG3_6_3	ACG3 of lineage 6 clone 3	This study
ACG3_6_4	ACG3 of lineage 6 clone 4	This study
ACG3_X_Y	ACG3 of lineage X clone Y; where X and Y are numbers between 1 and 10	This study
Δ*fumA*	MG1655 Δ*fumA*	This study
Δ*fumB*	MG1655 Δ*fumB*	This study
Δ*fumC*	MG1655 Δ*fumC*	This study
Δ*fumAfumB*	MG1655 Δ*fumAfumB*	This study
Δ*dcuA*	MG1655 Δ*dcuA*	This study
Δ*dcuB*	MG1655 Δ*dcuB*	This study
Δ*dcuC*	MG1655 Δ*dcuC*	This study
Δ*dcuD*	MG1655 Δ*dcuD*	This study
Δ*dcuR*	MG1655 Δ*dcuR*	This study
Δ*dcuS*	MG1655 Δ*dcuS*	This study

Table 5.1. CONTD. Bacterial strains and plasmids used in this study.

Strains	Genotype/ Description	Reference/ Source
Δ*dctA*	MG1655 Δ*dctA*	This study
Δ*pckA*	MG1655 Δ*pckA*	This study
Δ*gltP*	MG1655 Δ*gltP*	This study
Δ*frdA*	MG1655 Δ*frdA*	This study
B. circulans	wild-type	ATCC
B. subtilis	wild-type	ATCC
C. braakii	wild-type	ATCC
K. oxytoca	wild-type	ATCC
K. pneumoniae	wild-type	ATCC
P. denitrificans	wild-type	ATCC
P. putida	wild-type	ATCC
MG1655/pFumA	MG1655/pBKA1.418	This study
Δ*dcuC*/pFumA	MG1655 Δ*dcuC*/pBKA1.418	This study
JW	F-, *Δ(araD-araB)567*, *ΔlacZ4787*(::rrnB-3), λ^-, *rph-1*, *Δ(rhaD-rhaB)568*, *hsdR514*	CSGC
JW1604	JW *ΔfumA::kan*	CSGC
JW4083-1	JW *ΔfumB::kan*	CSGC
JW1603-1	JW *ΔfumC::kan*	CSGC
JW5735-1	JW *ΔdcuA::kan*	CSGC
JW4084-1	JW *ΔdcuB::kan*	CSGC
JW0616-2	JW *ΔdcuC::kan*	CSGC
JW3196-7	JW *ΔdcuD::kan*	CSGC
JW4085-3	JW *ΔdcuR::kan*	CSGC
JW4086-1	JW *ΔdcuS::kan*	CSGC
JW3496-1	JW *ΔdctA::kan*	CSGC
JW3366-1	JW *ΔpckA::kan*	CSGC
JW4038-1	JW *ΔgltP::kan*	CSGC
JW4115-1	JW *ΔfrdA::kan*	CSGC
plasmids		
pET-15b	Ap^R, *lacI*, P_{T7}, ori (3882) replicon	Invitrogen
pLBA110	*fumA* cloned into pET-15b	This study
pBKA1.56	*fumC* cloned into pET-15b	This study
pBKA1.58	*fumB* cloned into pET-15b	This study
pBKA1.412	*mdh* cloned into pET-15b	This study
pBbA1a-RFP	Ap^R, *lacI*, P_{Trc}, p15A ori replicon, *mRFP*	This study
pBKA1.418	*fumA* cloned into pBbA1a-RFP to replace *mRFP*	This study
pBKA1.420	*fumC* cloned into pBbA1a-RFP to replace *mRFP*	This study

Table 5.2. Primers used in this study.

Primer ID	SEQUENCE	Purpose
BKA118_dcuC_R	CCAGCCGCCATCTGATTATTGC	Primers used to confirm gene knockout
BKA117_dcuC_F	TTATTTACTTTGGGGCCTGGAGACAG	
BKA145_ompF_ R	ATTATTCTCGAGTATTTATTACCCTCATGGTTTTTTTATGACAC	
BKA146 _ompR_F	ATTATTCATATGTTTAGCGTCTTCAAGAGCCAGCG	
BKA058_fumA_R_qPCR	TCCGTTGGCAGTTCATCATAG	
BKA057_fumA_F_qPCR	CCTGTCCTCCGTATCATATTG	
BKA148_ompC_R	ATTTATCTCGAGGTTATTAACCCTCTGTTATATGCCT	
BKA147_ompC_F	ATTATTCATATGGTTTATGACCTGTTAAAACTTCGCG	
BKA143_maeB_R	TTATAACAGATCTCCCTGAATATTCAGGGTAAGCGTGAGAG	
BKA142_maeB_F	TATTACGAATTCGGGGCTGTTGCCCACACACTTTATTTGTG	
BKA120_dcuD_R	CAATCTGCCGTTTATGGGATTGACC	
BKA119_dcuD_F	TCGACCGGCATTACGCGTTG	
BKA127_pck_R	CGTTTCGTGACAGGAATCACGGAG	
BKA126_pck_F	AAATAAACGGGAGGCGAAGGTGC	
BKA125_dctA_R	AAAAAACTGCCCGTTGTCGGTC	
BKA124_dctA_F	GTGATGTTGTCTGGGTGTTATTTTAAGGCC	
BKA132_gltP_R	CCAACGCAACACAATTCATGCCC	
BKA131_gltP_F	AGGCATTCACACATTGCCGG	
BKA116_dcuB_R	TTTCCCCATCGGGAAAGGTGCC	
BKA115_dcuB_F	GTTCATACAAAACAGAACGTGACTGTGA	
BKA123_dcuR_R	ACACCGTTCACCTGAAGAGATATTAATTT	
BKA122_dcuR_F	GGAACCCGGGATTTTCACACAATTTTTTG	
BKA121_dcuS_R	TTTGGCTTTCTCCAGCGTCGAG	
BKA_dcuS_F	GCCCCGGAATCTTCATACAGCAC	
BKA114_dcuA_R	CATAGCACCACGACAGACGC	
BKA113_dcuA_F	CGGCTTACAAAGCAAAACGCT	
BKA129_frdA_R	CAGGTTGTCTTTGATGTAGCCCAGCG	
BKA128_frdA_F	GTGGAATAGCGTTCGCAGACC	
BKA_qPCR_R_fumA	TCCGTTGGCAGTTCATCATAG	Primers used for gene expression analysis
BKA_qPCR_R_fumB	TTGCTGGAAGAAGTCGTTACC	
BKA_qPCR_R_fumC	GACGGAAGACGTTCAGTTCAA	
BKA_qPCR_R_pck	CAACGCATCACGACGGATAG	
BKA_qPCR_R_dcuB	CGATACGCTCACACCAATCA	
BKA_qPCR_R_dctA	TACCAGCACCACAAACAGGA	
BKA_qPCR_R_gltP	TGACCGTATCCTCGACATGG	
BKA_qPCR_R_dcuA	GGTCGGCAGAATGAACAGAC	
BKA_qPCR_R_dcuC	CAGCAGCATCACCACGTTAG	
BKA_qPCR_R_dcuD	TGTTTCGCGCAAGTTATGCT	
BKA_qPCR_R_mdh	TGCAACTGTGGTGTTAACCG	
BKA_qPCR_R_gltA	TAACATCCAGCGAGTCGTGA	

Table 5.2. CONTD. Primers used in this study.

Primer ID	SEQUENCE	Purpose
BKA_qPCR_R_dcuR	GTAACGGTAAACTGGACGCC	Primers used for gene expression analysis
BKA_qPCR_R_dcuS	CGAAGTGCGTCAGCATAGTT	
BKA_qPCR_R_frdA	GACATCCTGCTCACACAACC	
BKA_qPCR_R_phoE	GATAACGCCGAAGAAGTCGG	
BKA_qPCR_R_ompC	CCTGGAATTTCAGACCTGCG	
BKA_qPCR_R_ompF	GTTCGCGTCGTACTTCAGAC	
BKA_qPCR_R_eda	TAGCCGGAGAAATACCACCC	
BKA_qPCR_R_maeA	GTTGGTTGTTCGTTCCGACA	
BKA_qPCR_R_maeB	TTCATACAGGCGATTGCTGC	
BKA_qPCR_R_gyrA	TGAACTCGGTGAGGACAGTT	
BKA_qPCR_F_gyrA	AGCAGGACGAACGTATCACT	
BKA_qPCR_F_maeB	TTCCGGTATTCCACGACGAT	
BKA_qPCR_F_maeA	CGATCGGTAAACTGTCGCTC	
BKA_qPCR_F_eda	ATGCTGGGTATGGACTACGG	
BKA_qPCR_F_ompF	ATCGTTGGTGCTTATGGTGC	
BKA_qPCR_F_ompC	CCTACATGCGTCTTGGCTTC	
BKA_qPCR_F_phoE	TGGAAGCCTGGACCGATATG	
BKA_qPCR_F_frdA	AGTATACCCGATGCGTAGCC	
BKA_qPCR_F_dcuS	TGCGCAGTAATGGCGTTATC	
BKA_qPCR_F_dcuR	CGAGCTAGCCAACGAGGTTA	
BKA_qPCR_F_gltA	TTACCCGTCTGTTCCATGCT	
BKA_qPCR_F_mdh	GGTATGGATCGTTCCGACCT	
BKA_qPCR_F_dcuD	CGTTAATCCTGATGCTCGGC	
BKA_qPCR_F_dcuC	TCCATTCTGGAGTTCCTCCG	
BKA_qPCR_F_dcuA	CAACCGCAAAAGCACTGATG	
BKA_qPCR_F_gltP	CAGCGCTTTCTTACGGTCAA	
BKA_qPCR_F_dctA	ATCATGCGTCTGGCACCTAT	
BKA_qPCR_F_dcuB	ATCGGTCGCTTCGTCATCAA	
BKA_qPCR_F_pck	TGATTGGCGATGACGAACAC	
BKA_qPCR_F_fumC	CCAACACAGTGTGAGGCATTA	
BKA_qPCR_F_fumB	TCTGGAGTGCGTCGCTTCTC	
BKA_qPCR_F _fumA	CCTGTCCTCCGTATCATATTGC	
LBA_FumA_fwd	TGCCGCGCGGCAGCCAAGAAAACCTGTATTTTCAGGGCATGTCAAACAAACCCTTTCATT	For cloning *fumA* into pET-15b by Gibson assembly
LBA_FumA_rev	TCGGGCTTTGTTAGCAGCCGTTATTTCACACAGCGGGTG	
BKA001_FumB	ATTATACATATGATGTCAAACAAACCCTTTATCTACC	For cloning *fumB* into pET-15b
BKA002_FumB	ATTATAGGATCCTTACTTAGTGCAGTTCGCGCAC	
BKA003_FumC	AATTAATCAGGTGAGGAGCAGCATATGAATACAGTACGCAGCG	For cloning *fumC* into pET-15b
BKA004_FumC	GGCTGCACCTGTATGTGGATCCTTAACGCCCGGCTTTCATACTG	
BKA005_Mdh	ATTAGGATCCTTACTTATTAACGAACTCTTCGCC	For cloning *mdh* into pET-15b
BKA006_Mdh	ATATACATATGAAAGTCGCAGTCCTCG	
BKA007_FumA	TTATACTTACATATGTCAAACAAACCCTTTCATTATCAGGC	For cloning *fumA* into pBbA1a-RFP
BKA008_FumA	ATATACAGGATCCTTATTTCACACAGCGGGTGCATTG	

5.8.4 Culturing of *E. coli* in liquid media

For liquid cultures, 5 mL of liquid medium (with antibiotics if necessary) is transferred to a sterile 18 mm culture tube. The medium was then inoculated with a single bacterial colony. The tube was capped and incubated overnight or until the stationary phase (12 - 16 h) at 37 °C with shaking (200 rpm). To grow large cultures, the overnight culture was diluted (1:100 or to an initial OD_{600} of 0.05) in baffled flask that is ≥ 5 times the volume of the culture. The flasks were incubated at 37 °C with shaking (200 rpm). The OD_{600} of the cultures was measured manually using the Nanodrop instrument by diluting the stationary phase culture (100:900 culture: fresh medium).

5.8.4.1 Using the plate reader

For plate reader culturing, a single colony from an overnight plate is inoculated into 1 mL of media in a sterile deep well plate and incubated at 37°C to saturation. A typical plate reader start culture is 2:200, 1:200, or 1:250 overnight culture: fresh media. The first and last columns and rows are typically filled with blank media and used as controls. This is to reduce the effect of bias due to evaporation, temperature difference, or other parameters that are more pronounced at the edges of the plate. The plates are sealed with a breathable seal that allows air exchange. The appropriate plate type is selected in the instrument setup panel, and the culture is incubated with orbital shaking at 30 °C or 37 °C, as required. The OD_{600} of the cultures is automatically measured every 30 min for the duration of the culture. Orbital shaking is performed during the reading intervals. The plates must shake most of the time when measurements are not taken. All strains grown have at least three replicates. The plate chamber is preheated to the desired temperature before culturing begins. Once the culturing is complete, the data is exported as an Excel spreadsheet and formatted to create a tab-delimited file for further processing and determination of growth parameters using the R Studio program.

5.8.5 Determination of growth characterization parameters

All strains cultured using shake flasks were characterized by calculating fitness values for each culture relative to *E. coli* MG1655 in glucose minimal media (M9G). The fitness was determined as the slope/maximum growth rate of the growth curves obtained from OD_{600} measurements determined from the log-linear regions. The relative fitness of each population was calculated using the equation:

$$Relative\ fitness = \frac{\mu_{max,c}}{\mu_{max,M9G}}$$

where $\mu_{max,c}$ = maximum growth rate of strain in specific media

$\mu_{max,M9G}$ = maximum growth rate of *E. coli* MG1655 in M9G media

The lag-phase times were extrapolated from growth profiles for both shake flask and plate reader experiments. All plots were obtained using Origin 2022b software.

For plate reader experiments, the Growthcurver package in R studio is used to determine growth parameters by analyzing and fitting the optical density data to a logistic function from which the growth rate (r), doubling time (t), and carrying capacity (K) can be determined. This is achieved by using codes provided by the developers.[2] Growthcurver uses the non-linear least-squares Levenberg-Marquardt algorithm to find the best values of N_0, K, and r. The logistic equation:

$$N_t = \frac{K}{1 + (\frac{K - N_0}{N_0})e^{-rt}}$$

Where N_t = final population size

N_0 = population size at the beginning of the growth curve

K = carrying capacity

r = growth rate of the population

t = doubling time

5.8.6 Culturing *E. coli* in varying concentrations of substrates

To investigate the effect of varying ADCA concentrations on the specific growth rate of *E. coli* MG1655 in M9A, the strain was cultivated in M9A media containing 5, 20, 50, and 100 mM ADCA using the plate reader method. Similar experiments were conducted in M9G and M9F with identical substrate concentrations. Three additional media conditions, M9FA (50 mM fumarate and 50 mM ADCA), M9GA (50 mM glucose and 50 mM ADCA), and M9GF (50 mM glucose and 50 mM fumarate), were also studied. MG1655 was cultivated using the plate reader method for 72 h. The maximum growth rates and lag-phase times were determined for each condition.

5.8.7 Adaptive Laboratory Evolution (ALE)

ALE experiments were performed in shake flasks using *E. coli* MG1655 as the ancestor strain. A single colony was used to cultivate a seed culture in culture tubes using 5 mL M9G media. The seed culture was washed twice with M9 salts and then used to inoculate fresh 50 mL M9G (20 mM glucose), M9F (40 mM fumarate), and M9A (40 mM ADCA) media in a 250 mL flask to an initial OD_{600} of 0.02 and incubated at 37 ^{o}C with 200 rpm shaking. The OD_{600} of all cultures was measured at 4 – 6 h intervals. At late log phase, an aliquot from the M9A culture (ADCA-Consuming Generations 1 - ACG1) was inoculated into fresh M9G, M9F, and M9A media as described above and cultivated to obtain ADCA-Consuming Generations 2 (ACG2). This process was repeated to obtain ACG3. The ALE experiment was performed in triplicates to generate three distinct lineages. The nomenclature for each sample is as shown in Figure 2.7. The corresponding number of generations for ACG1, ACG2, and ACG3 were determined using the equation:

$$No. of\ generations = log_2 \frac{initial\ OD_{600}}{final\ OD_{600}}$$

When the cultures reached $OD_{600} = 1$, an aliquot of each was removed and cells were collected by centrifugation (12, 000 x *g*, 10 min, 4 °C) for subsequent transcriptional analysis and whole genome sequencing analysis. Culture supernatants obtained were also analyzed by ^{1}H-NMR, HPLC-RID/UV, LCMS/MS, and GCMS/MS for characterization and quantification of pyruvate, acetate, succinate, and lactate secreted into the media (See Analytical methods).

5.8.8 Whole genome sequencing of ACG populations and clones

5.8.8.1 Isolation of gDNA

Genomic DNA from ACG1, ACG2, and ACG3 triplicates and *E. coli* MG1655 in M9G (G1) was extracted using a Promega Wizard® genomic DNA purification kit for Illumina sequencing and separately using Promega Wizard® HMW DNA Extraction Kit for Oxford Nanopore Technologies sequencing, using the manufacturer's protocol. Briefly, cell pellets were lysed using a custom nucleic lysis solution followed by incubation at 80 ºC for 5 min, RNase was added to degrade RNA molecules. Protein is precipitated and removed in subsequent steps. The supernatant from this step containing genomic DNA is transferred to a new tube containing isopropanol. Precipitated DNA was washed with ethanol and air-dried. Extracted DNA was rehydrated and quantified using Nanodrop. For Oxford Nanopore Technologies sequencing, the genomic DNA was sheared using a g-TUBE (Covaris®, PerkinElmer Company, Part number 520079) following manufacturer's protocol to obtain more fragments between 3 – 12 kbp (Figure 5.2.1).

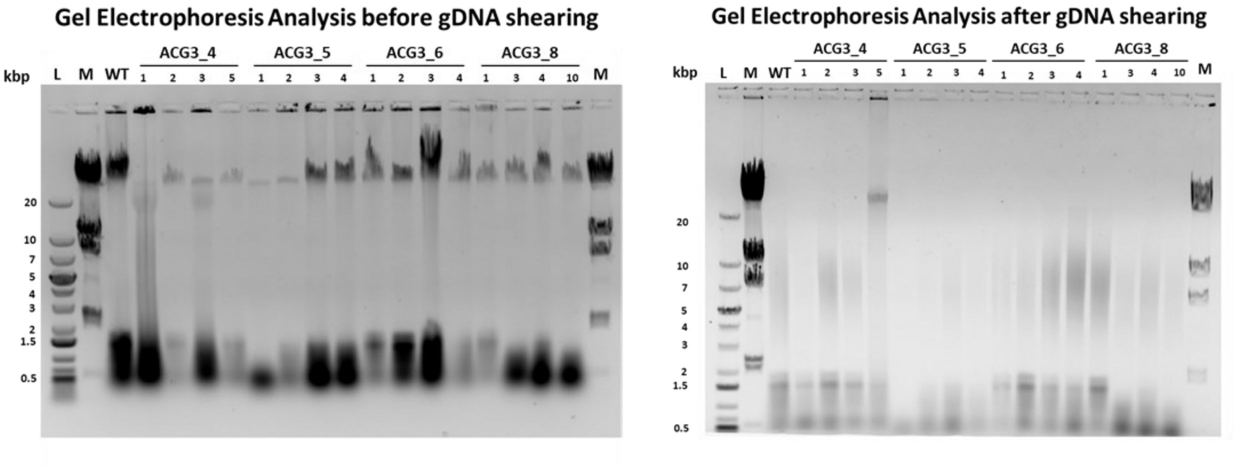


Figure 5.2. Gel electrophoresis analysis of genomic DNA before and after shearing for ONT sequencing.

5.8.8.2 Sequencing of population samples and clonal samples

Samples were submitted to the MSU RTSF Genomic Core for NGS library preparation and sequencing. Population samples were sequenced using Illumina NovaSeq 6000. Libraries were prepared using the Illumina TruSeq Nano DNA Library Preparation kit with IDT for Illumina Unique Dual Index adapters, following the manufacturer's recommendations. Completed libraries were QC'd and quantified using a combination of Qubit dsDNA HS and Agilent 4200 TapeStation HS DNA1000 assays. The libraries were pooled in equimolar quantities, and the pool was quantified using the Invitrogen Collibri Quantification qPCR kit. The pool was loaded onto one lane of an Illumina NovaSeq 6000 SP flow cell using the Xp loading workflow. Sequencing was performed in a 1x100bp single-read format using a NovaSeq v1.5 100-cycle reagent cartridge. Base calling was done by Illumina Real Time Analysis (RTA) v3.4.4, and the output of RTA was demultiplexed and converted to FastQ format with Illumina Bcl2fastq v2.20.0.

Clonal samples were sequenced using Oxford Nanopore Technology using an R10.4.1 PromethION flow cell. Variants calling was performed using breseq pipeline to identify SNP and structural variants.[3,4] See Figure in Chapter Two for variant calling using breseq, Medaka, and downstream analysis strategies.

5.8.8.2 Predicting effect of mutations using mutfunc database

The precomputed mechanistic effects of observed mutations were investigated using the mutfunc database (www.mutfunc.com).[5] Mutfunc is a database of mutations occurring in functionally essential regions or that are predicted to disrupt protein structure stability, protein interaction interfaces, post-translation motifs (PTMs), protein translation, conserved regions, and regulatory regions. This allows for a quick annotation of variants displaying mutants that are likely deleterious to function and predicted consequences. A list of the mutations is submitted to the database as a query and once completed, a results page with various sections is displayed. A detailed discussion on how to interpret the data can be found here (www.mutfunc.com/help).

5.8.9 Transcriptomic analysis of ACG series

5.8.9.1 Isolation of total RNA

Total RNAs were extracted from about 2.9 x 10^9 cells from G1, F1, ACG1, ACG2, and ACG3 populations using the RNeasy mini kit from QIAGEN following the manufacturer's protocol. About 500 µL of bacterial culture was added to 1 mL of RNAprotect^{TM} Bacterial Reagent (QIAGEN), vortexed, and incubated for 5 min at room temperature. Cells were harvested using a microcentrifuge (14 000 x g, 10 min). The supernatant was removed by inversion of the tubes. RNA was extracted from cell pellets following exact instructions in the QIAGEN protocol with RNA cleanup procedure to remove genomic DNA contaminations. The integrity of isolated RNAs was evaluated by agarose gel electrophoresis.

For the RNA gel electrophoresis, an agarose gel was prepared by heating 1 g of agarose in 100 ml of water and heated with a microwave until it fully dissolved. The solution was allowed to cool to 60 °C before adding 10 ml of 10X FA running buffer and 18 mL of 37% formaldehyde (12.3 M) solution. After vigorous mixing, the gel was poured using a comb to create wells that could accommodate at least 25 µl of sample. The gel was then placed in an electrophoresis tank and covered with enough 1X FA running buffer to submerge the gel by a few millimeters. The comb was removed once the setup was complete. RNA samples ranging (100 ng) were mixed with 5X RNA loading buffer. Ethidium bromide was added to a final concentration of 10 µg mL^{-1} to stain the RNA. Samples were then heat-denatured at 65 - 70°C for 10 min. The prepared RNA samples were loaded into the gel and electrophoresed at 5 V/cm. Electrophoresis was continued until the bromophenol blue dye had migrated at least up to 2/3 of the gel length.

Post-electrophoresis, the gel was visualized on a. The presence of sharp 28S and 18S rRNA bands indicated intact RNA, with the 28S band being approximately twice as intense as the 18S band (Figure 5.3). This 2:1 ratio served as an indicator of RNA integrity. Partially degraded RNA appeared smeared and lacked sharp rRNA bands or the 2:1 ratio. Completely degraded RNA was evident as a low molecular weight smear.

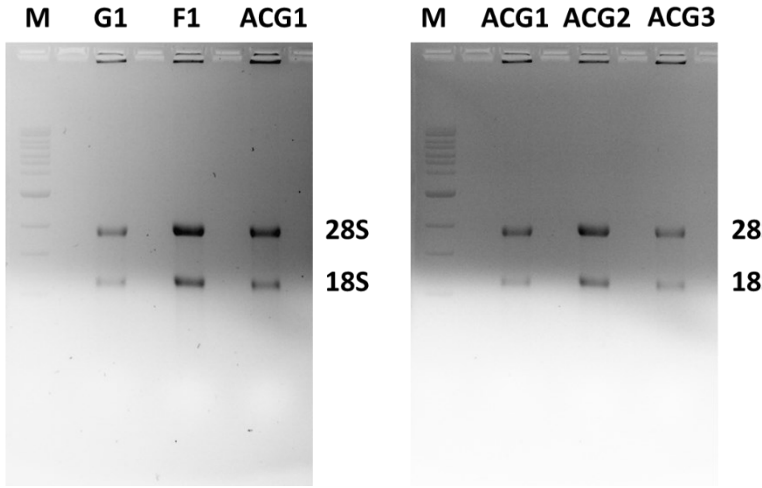


Figure 5.3. RNA gel electrophoresis for samples extracted from G1, F1, ACG123. Marker lane is indicated as M. Gel was run with 1 µg of each sample on 1% agarose. Two bands (28S and 18S) are seen for all samples showing about 2:1 ratio. Less smear is observed indicating high intergrity of isolated RNA.

5.8.9.2 RNA-Seq analysis

RNA samples for G1 (MG1655 cultivated in glucose medium as control) and ACG3 (triplicates) were shipped to Seq Coast Genomics for RNA-Seq analysis. Samples were prepared for whole genome sequencing using an Illumina Stranded Total RNA Prep Ligation with Ribo-Zero Plus Microbiome and IDT For Illumina Unique Dual Indexes. Sequencing was performed on the Illumina NextSeq2000 platform using a 300-cycle flow cell kit to produce 2X 150 bp paired reads. Read demultiplexing, read trimming, and run analytics were performed using DRAGEN v3.10.12, an on-board analysis software on NextSeq2000.

5.8.9.3 Synthesis of cDNA

The cDNA for RT-qPCR analysis was synthesized using the Applied Biosystems^{TM} High Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (Thermo Fisher Cat. No. 4374966). The cDNA synthesis was performed following the manufacturer’s protocol. In brief, a x2 reverse transcription master mix was prepared as shown in Table 5.3. About 10 µL of the master mix was added into wells of a 96-well reaction plate. RNA sample (10 µL of 100 ng µL^{-1}) was then added into each well, pipetting up and down to mix. The plate was sealed and briefly centrifuged to spin down the contents and eliminate air bubbles. The plate was placed on ice and then loaded onto the thermal cycler. The program used on the thermal cycler is as shown in Table 5.4. Good practices should be maintained for PCR and RT-qPCR procedures. For instance, a separate area was maintained, and dedicated equipment and supplies were used for sample preparation, reaction setup, amplification, and analysis of products.

Table 5.3. Components for 2X RT master mix.

Component	Volume (µL)
10X RT buffer	2.0
25X dNTP mix (100 mM)	0.8
10X RT random primers	2.0
MultiScribe™ reverse transcriptase	1.0
Rnase inhibitor	1.0
Nuclease-free H_2O	3.2
Total per reaction	10.0

Table 5.4. Thermal cycler program for reverse transcription reaction.

Settings	Step 1	Step 2	Step 3	Step 4
Temperature	25 °C	37 °C	85 °C	4 °C
Time	10 minutes	120 minutes	5 minutes	Hold

5.8.9.4 RT-qPCR for targeted gene expression analysis

RT-qPCR analysis was performed using PowerTrack™ SYBR™ Green Master Mix (Thermo Fisher Cat. No. A46110) with a QuantStudio™ 7 Pro Real-Time PCR system. About 5 ng of cDNA was used in all reactions. Four replicates of three biological samples were used for G1, F1, and ACG123. The master mix was thawed and swirled to mix thoroughly. The RT-qPCR was done following the manufacturer's protocol. The master mix, primers, and nuclease-free water with the

DNA and Yellow Sample Buffer were combined as shown in Table 5.5. The components were mixed thoroughly and then centrifuged briefly to collect the contents at the bottom of the tube. The 20 µL volume for each reaction was transferred to each well of a MicroAmp® Optical 96-Well Reaction Plate (Applied Biosystems by Life Technologies, Ref. N8010560).

Table 5.5. Components for RT-qPCR reactions.

Components	Stock concentration	Final concentration	Volume (1X)	Volume (4X) with 10% overage
Yellow Sample Buffer and DNA				
DNA	5 ng/µL	0.5 ng/µL	2 µL	8.8 µL
Yellow Sample Buffer	40X	1X	0.5 µL	2.2 µL
Master mix, primers, and nuclease-free water				
PowerTrack™ SYBR™ Green Master Mix	2X	1X	10 µL	44.0 µL
Forward and reverse primers	8,000 Nm	400 Nm	1 µL	4.4 µL
Nuclease-free water	~	~	6.5 µL	28.6 µL
Total PCR volume	~	~	20 µL	88.0 µL

The plates were sealed with an optical adhesive cover and then centrifuged briefly to collect the contents at the bottom of each well and eliminate any air bubbles. The PCR was performed on the reaction plate immediately after completing the setup. The thermal protocol used for the setup is as shown in Table 5.6. A dissociation step was performed immediately after the RT-qPCR run. The efficiencies of amplification of each set of primers were determined in a separate experiment. This was done by preparing a tenfold serial dilution of four points for a given cDNA sample (i.e 10^{1}, 10^{-1}, 10^{-2}, and 10^{-4}). The RT-qPCR setup was prepared as shown in Table 5.5 and 5.6. All plates included a no template control (NTC) for each primer set.

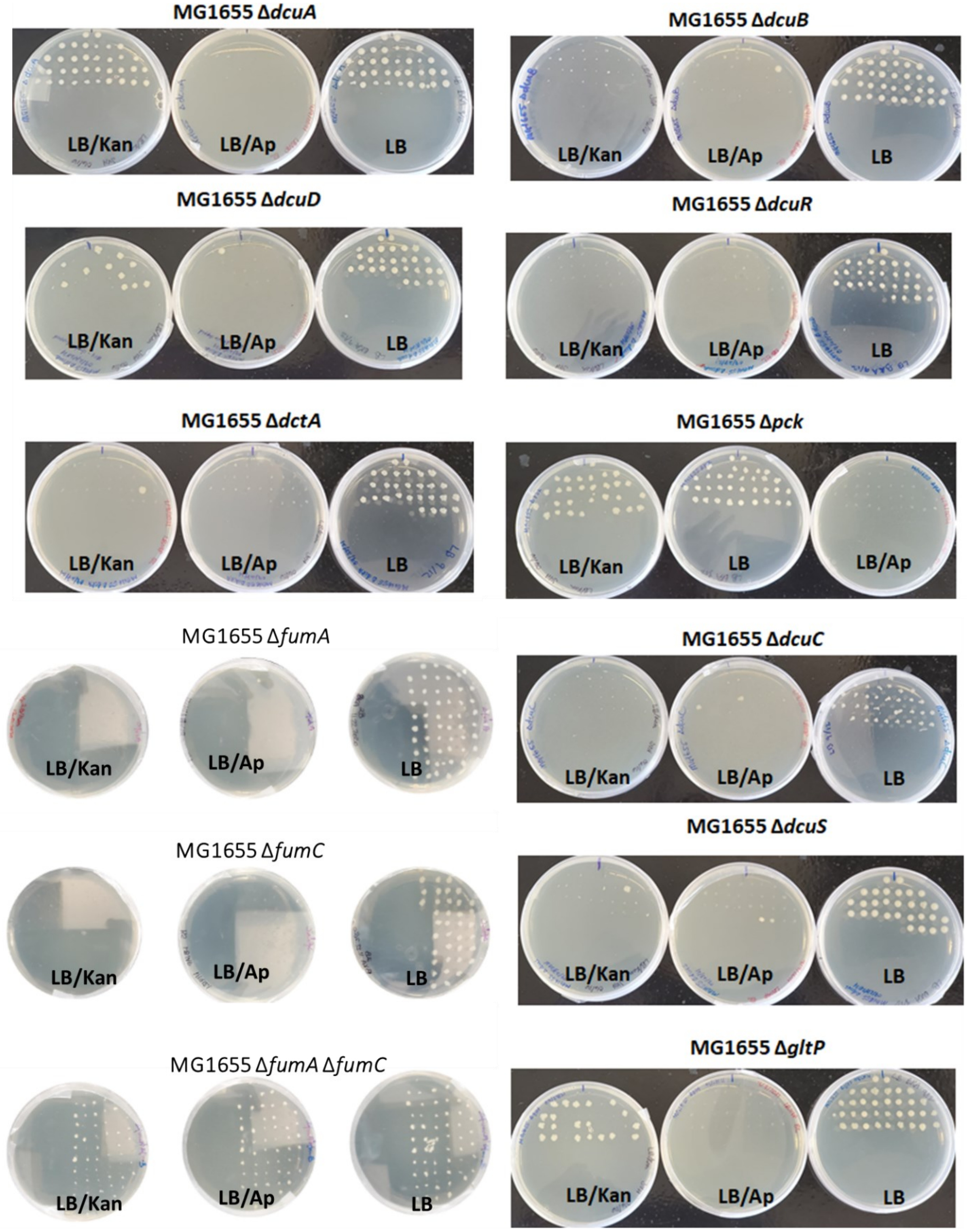
Table 5.6. Thermal protocol for RT-qPCR analysis.

Standard cycling mode			
Step	Temperature (°C)	Duration (seconds)	Cycles
Enzyme activation	95	120	1
Denature	95	15	40
Anneal/extend	60	60	
Dissociation step	Ramp rate	Temperature (°C)	Time (seconds)
1	1.99 °C/second	95	15
2	1.77 °C/second	60	60
3	0.075 °C/second	95	15
Options			
Experimental type :	Standard curve		
Reagent :	SYBR™ Green reagents		
Reporter :	SYBR™ Green		
Quencher :	None		
Passive reference dye :	ROX™ dye		
Ramp speed :	Standard or fast		
Melt curve ramp increment :	Continuous		
Reaction volume :	20 µL		

Analysis was performed using QuantStudio$^{\text{TM}}$ Design & Analysis Software v1.5.1 (Applied Biosystems by Thermo Fisher Scientific). Relative quantification was performed using the $\Delta\Delta C_T$ method, normalized against *gyrA* or *rssA* reference genes, and both had stable expression across all samples. The data were statistically analyzed to determine the expression levels of each target gene, using ANOVA followed by Tukey's multiple comparisons test, with a significance threshold set at $p < 0.05$.

5.8.10 Generating gene knockout mutants by P1 transduction

E. coli MG1655 was used as the recipient and wild-type strain for the gene knockout study. Donor strains with gene knockout cassettes were obtained from the Keio collection.[6] P1 transduction was used to introduce gene deletions in the recipient strain.[7] Table 5.8 lists all mutants generated. The approach utilized an excisable kanamycin resistance cassette as described by Saragliadis et al. (2018).[8] All donor strains were obtained for the Coli Genetic Stock Center (*E. coli* Genetic Resources at Yale University). These strains had kanamycin cassettes inserted into the target gene loci to facilitate the selection process. A P1vir bacteriophage lysate was prepared from the donor strains. This involved infecting the donor strains with P1vir and subsequently collecting phage from the lysed cultures. The recipient MG1655 cells were grown to an OD_{600} of approximately 1.0. These cells were then mixed with the phage lysate at a multiplicity of infection (MOI) of 0.5 and incubated in LB with calcium chloride to promote phage adsorption. After infection, the cells were treated with sodium citrate to halt the infection process, and the mixture was plated on LB agar containing kanamycin to select for transductants. Kanamycin-resistant colonies that were obtained were transformed with the pCP20 plasmid (obtained from the Draths lab plasmid collection) to provide the FLP recombinase in trans.



MG1655 Δ*dcuA*

MG1655 Δ*dcuB*

MG1655 Δ*dcuD*

MG1655 Δ*dcuR*

MG1655 Δ*dctA*

MG1655 Δ*pck*

MG1655 Δ*fumA*

MG1655 Δ*dcuC*

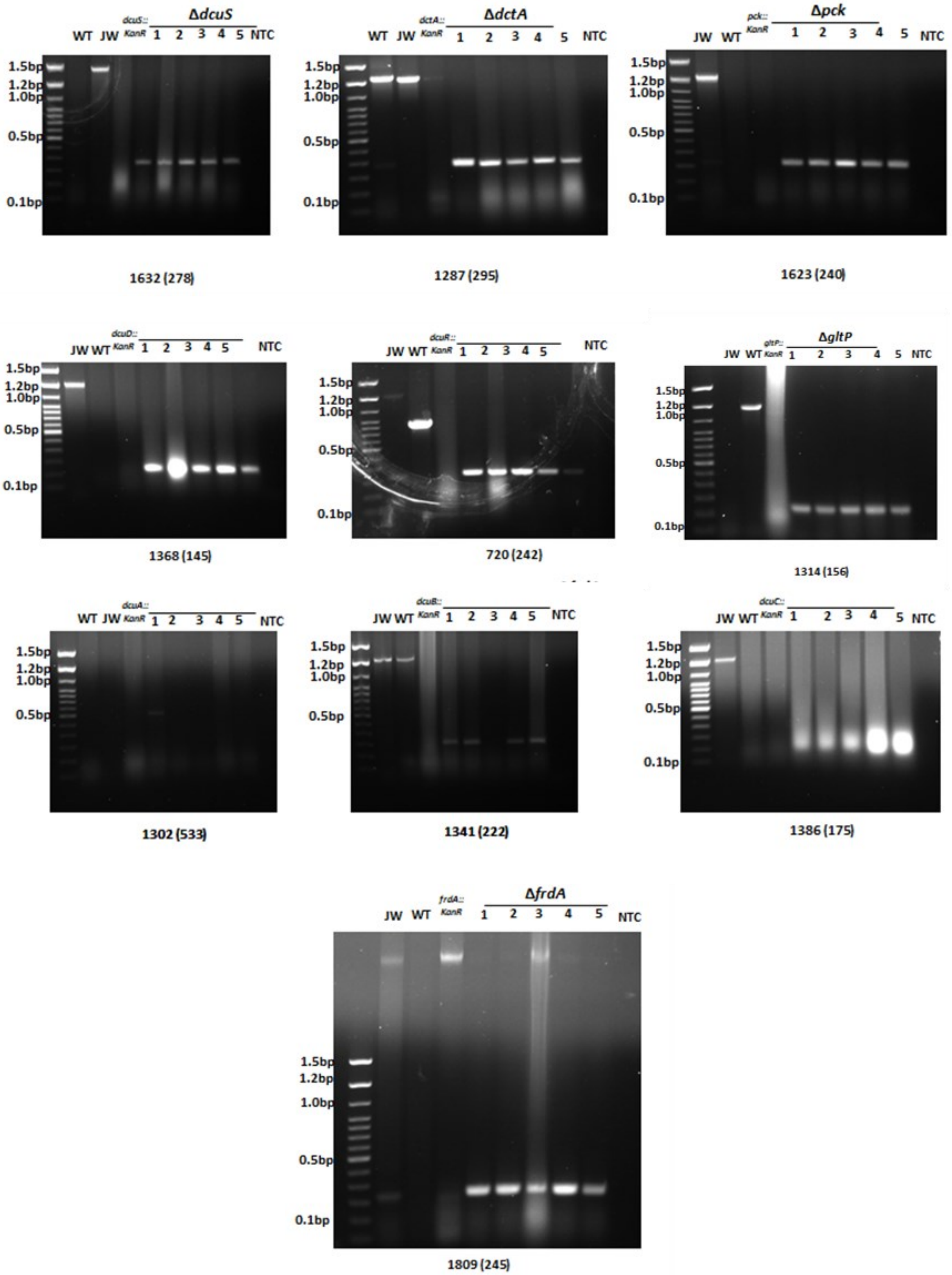
MG1655 Δ*fumC*

MG1655 Δ*dcuS*

MG1655 Δ*fumA* Δ*fumC*

MG1655 Δ*gltP*

Figure 5.4. Screening for kanamycin and ampicillin sensitivity to identify clones with excised kanamycin cassette and cured pCP20 plasmid.



1632 (278)

1287 (295)

1623 (240)

1368 (145)

720 (242)

1314 (156)

1302 (533)

1341 (222)

1386 (175)

1809 (245)

Figure 5.5. Colony PCR verification of gene knockouts.

The colonies were then incubated at 43 °C to cure the plasmid, and clones were screened for sensitivity to both kanamycin and ampicillin to confirm the excision of the cassette (Figure 5.4). Colony PCR was performed on selected clones to confirm the deletion of the target genes. Primers flanking the target locus (see Table 5.2) were used to amplify the genomic DNA. The expected size reduction in PCR products confirmed the successful excision of the gene (Figure 5.5). The wild-type strain and mutants were cultivated in both shake flasks and using plate reader as described above.

5.8.11 Determination of kinetic parameters for fumarases

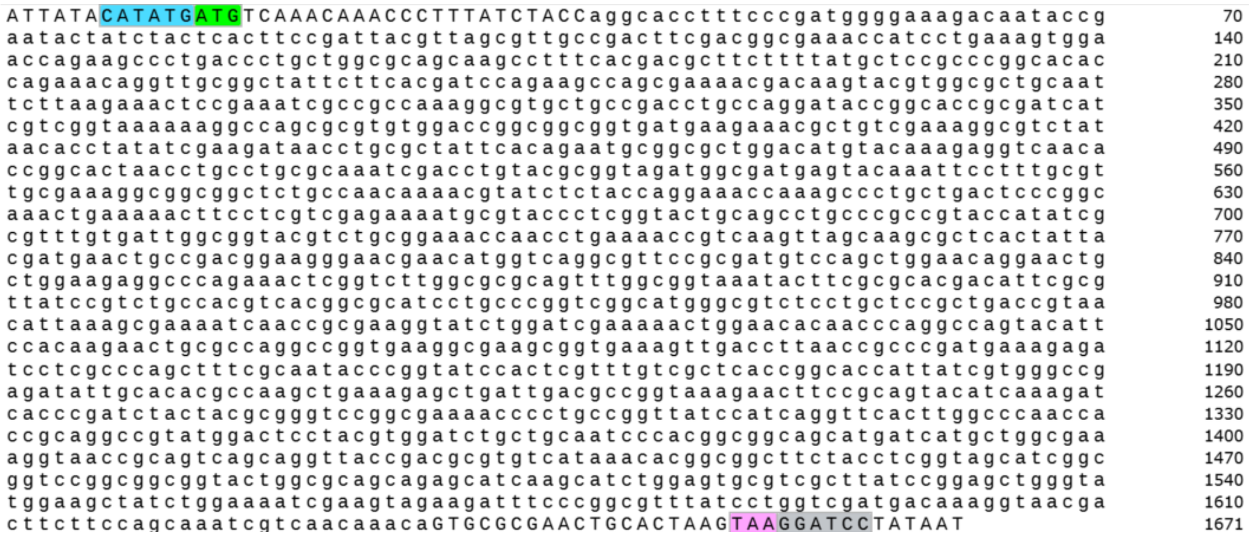
Enzyme kinetics were carried out at 25 °C using a Molecular Devices SpectraMax® iD3 multimode microplate reader and Shimadzu UV2600 spectrophotometer. All assays were performed in triplicates. Purified FumA and FumB were stored in 0.5 M ammonium sulfate buffer and used within 10 h of purification. FumA and FumB are both class I type fumarate hydratase and contain an iron-sulfur cluster making them unstable. FumC, a class II type fumurate hydratase lacks an iron-sulfur cluster, hence more stable. FumC was stored in 100 mM sodium phosphate buffer containing 20% ethylene glycol at -80 °C until use. All enzymes used for the kinetics overexpressed from cloneed plasmids and purified as discussed earlier.

5.8.11.1 Genes and plasmids

The fumarase genes were PCR amplified with primer sets LBA001 and LBA002 (*fumA*), BKA001 and BKA002 (*fumC*), BKA005 and BKA006 (*fumC*). Each gene fragment was cloned into pET-15b vector between NdeI and XhoI sites to afford pLBA110 (FumA), pBKA2.260 (FumB), and pBKA1.56 (FumC). This allowed for the overexpression of His_6-tagged fumarases downstream of T7 promoters inducible with IPTG. The gene sequences for the fumarases are shown in Figure 5.6 – 5.8. The cloning of pLBA110, pBKA1.58, and pBKA1.56 is shown in Figure 5.9 – 5.11.

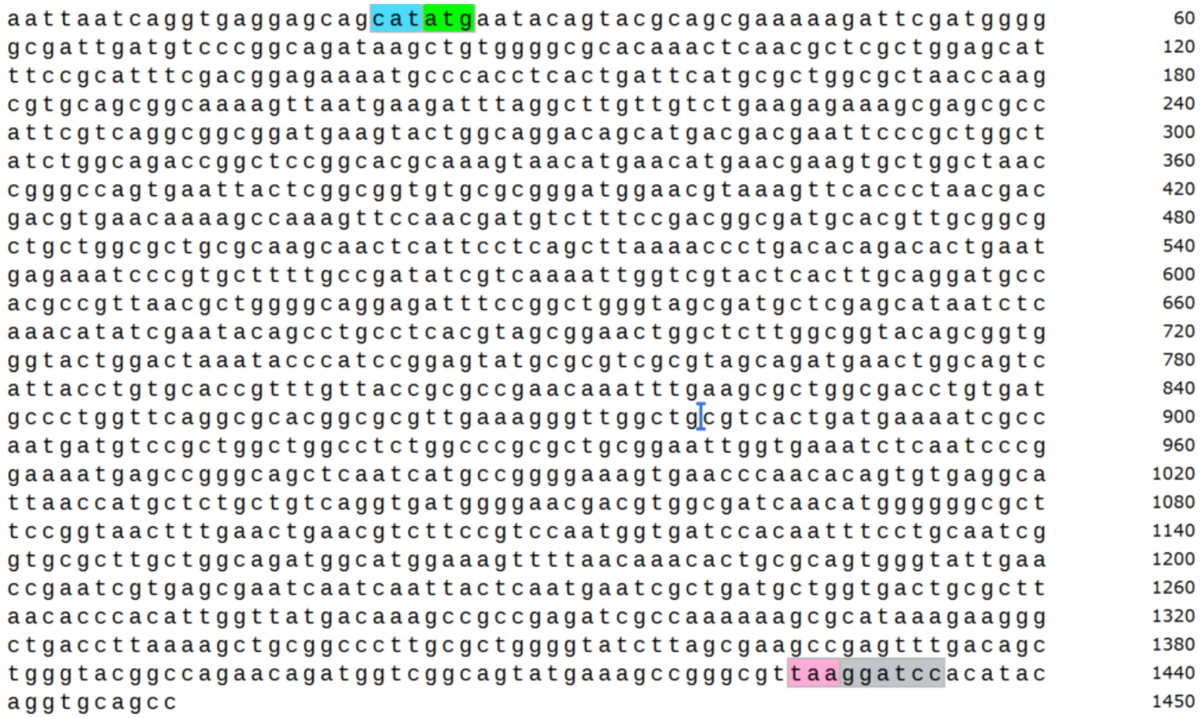
```
TGCCGCGCGGCAGCCAAGAAAACCTGTATTTTCAGGGCATGTCAAACAAACCCTTTCATTatcaggctcc     70
ttttccactcaaaaaagatgatactgagtattacctgctaaccagcgaacacgttagcgtatctgaattt    140
gaagggcaggagattttgaaagtcgcacccgaagcgttaactctgttggcgcgccaggcgtttcatgatg    210
cgtcgttcatgctgcgtccggcgcaccaacaacaggtggccgacattctgcgtgacccggaggccagcga    280
aaatgataaatatgtggcgctgcaattcctgcgtaactccgacatcgcggcgaaaggcgttctgccaacc    350
tgtcaggataccggcaccgcgattattgttggtaaaaagggcagcgtgtatggaccggtggtggtgatg    420
aagcggcgctggcgcgcggtgtctataacacttatatcgaagataatctgcgctactcgcaaaacgcgcc    490
gctggatatgtataaagaagtgaataccggcaccaatctgccagcgcagatcgatctttatgccgttgat    560
ggcgacgaatacaaattcctctgtatcgccaaaggtggtggttcggcaaacaagacgtatctctatcagg    630
aaaccaaagcgttactgacgccggggaaactgaaaaattacctggttgagaagatgcgcacgctgggtac    700
ggcggcctgtcctccgtatcatattgcgttcgttattggtggaacttctgcagaaacgaaccttaaaacg    770
gtgaaactggcttccgcgaaatactatgatgaactgccaacggaagggaatgagcacggtcaggcgttcc    840
gcgatgtggaactggaaaaagaattgctgatcgaagcgcaaaatcttggtctgggtgcgcagtttggtgg    910
taaatacttcgctcacgacatccgcgtgattcgcctgccacgtcacggcgcatcctgcccggtcggtatg    980
ggcgtctcctgctctgctgaccgtaatatcaaagcgaagatcaaccgtcaggggatctggatcgaaaaac   1050
tggaacataatccaggcaaatatatcccggaagagctgcgcaaagcgggagaaggcgaagcggtgcgcgt   1120
tgaccttaaccgtccgatgaaagagatcctcgcacagttgtcgcagtatcccgtttctacacgcttatcg   1190
cttaacggcacgattatcgtcggtcgtgatattgctcacgccaaactgaaagagcggatggataacggtg   1260
aagggctgccgcagtacatcaaagatcatccgatttactacgcgggtccggccaaaacgccggaaggtta   1330
tgcctccggttctcttggcccaacgaccgccggacggatggattcttatgtcgatcaactgcaagcgcag   1400
ggcggaagtatgatcatgctggcgaaaggcaaccgcagccagcaggtgacggatgcctgtaaaaaacacg   1470
gcggcttctaccttggcagtatcggtggtccggccgctgtattggcgcagggaagtattaagagcctgga   1540
atgtgttgaatatccggaactgggaatggaagccatctggaaaattgaagtggaagatttcccggcgttt   1610
atccttgtggatgataaaggaaatgacttcttccagcagatacaactcacacaatgCACCCGCTGTGTGA   1680
AATAACGGCTGCTAACAAAGCCCGA                                                1705
```

Figure 5.6. Sequence of *fumA* showing start and stop codons in green and red, respectively.



```
ATTATACATATGATGTCAAACAAACCCTTTATCTACCaggcacctttcccgatggggaaagacaataccg     70
aatactatctactcacttccgattacgttagcgttgccgacttcgacggcgaaaccatcctgaaagtgga    140
accagaagccctgaccctgctggcgcagcaagcctttcacgacgcttcttttatgctccgcccggcacac    210
cagaaacaggttgcggctattcttcacgatccagaagccagcgaaaacgacaagtacgtggcgctgcaat    280
tcttaagaaactccgaaatcgccgccaaaggcgtgctgccgacctgccaggataccggcaccgcgatcat    350
cgtcggtaaaaaaggccagcgcgtgtggaccggcggcggtgatgaagaaacgctgtcgaaaggcgtctat    420
aacacctatatcgaagataacctgcgctattcacagaatgcggcgctggacatgtacaaagaggtcaaca    490
ccggcactaacctgcctgcgcaaatcgacctgtacgcggtagatggcgatgagtacaaattcctttgcgt    560
tgcgaaaggcggcggctctgccaacaaaacgtatctctaccaggaaaccaaagccctgctgactcccggc    630
aaactgaaaaacttcctcgtcgagaaaatgcgtaccctcggtactgcagcctgcccgccgtaccatatcg    700
cgtttgtgattggcggtacgtctgcggaaaccaacctgaaaaccgtcaagttagcaagcgctcactatta    770
cgatgaactgccgacggaagggaacgaacatggtcaggcgttccgcgatgtccagctggaacaggaactg    840
ctggaagaggcccagaaactcggtcttggcgcgcagtttggcggtaaatacttcgcgcacgacattcgcg    910
ttatccgtctgccacgtcacggcgcatcctgcccggtcggcatgggcgtctcctgctccgctgaccgtaa    980
cattaaagcgaaaatcaaccgcgaaggtatctggatcgaaaaactggaacacaacccaggccagtacatt   1050
ccacaagaactgcgccaggccggtgaaggcgaagcggtgaaagttgaccttaaccgcccgatgaaagaga   1120
tcctcgcccagctttcgcaatacccggtatccactcgtttgtcgctcaccggcaccattatcgtgggccg   1190
agatattgcacacgccaagctgaaagagctgattgacgccggtaaagaacttccgcagtacatcaaagat   1260
cacccgatctactacgcgggtccggcgaaaacccctgccggttatccatcaggttcacttggcccaacca   1330
ccgcaggccgtatggactcctacgtggatctgctgcaatcccacggcggcagcatgatcatgctggcgaa   1400
aggtaaccgcagtcagcaggttaccgacgcgtgtcataaacacggcggcttctacctcggtagcatcggc   1470
ggtccggcggcggtactggcgcagcagagcatcaagcatctggagtgcgtcgcttatccggagctgggta   1540
tggaagctatctggaaaatcgaagtagaagatttcccggcgtttatcctggtcgatgacaaaggtaacga   1610
cttcttccagcaaatcgtcaacaaacaGTGCGCGAACTGCACTAAGTAAGGATCCTATAAT            1671
```

Figure 5.7. Sequence of *fumB* showing NdeI (blue) and BamHI (gray) restrictive sites. The start and stop codons are shown in green and red, respectively.



```
aattaatcaggtgaggagcagcatatgaatacagtacgcagcgaaaaagattcgatgggg      60
gcgattgatgtcccggcagataagctgtggggcgcacaaactcaacgctcgctggagcat     120
ttccgcatttcgacggagaaaatgcccacctcactgattcatgcgctggcgctaaccaag     180
cgtgcagcggcaaaagttaatgaagatttaggcttgttgtctgaagagaaagcgagcgcc     240
attcgtcaggcggcggatgaagtactggcaggacagcatgacgacgaattcccgctggct     300
atctggcagaccggctccggcacgcaaagtaacatgaacatgaacgaagtgctggctaac     360
cgggccagtgaattactcggcggtgtgcgcgggatggaacgtaaagttcaccctaacgac     420
gacgtgaacaaaagccaaagttccaacgatgtctttccgacggcgatgcacgttgcggcg     480
ctgctggcgctgcgcaagcaactcattcctcagcttaaaaccctgacacagacactgaat     540
gagaaatcccgtgcttttgccgatatcgtcaaaattggtcgtactcacttgcaggatgcc     600
acgccgttaacgctggggcaggagatttccggctgggtagcgatgctcgagcataatctc     660
aaacatatcgaatacagcctgcctcacgtagcggaactggctcttggcggtacagcggtg     720
ggtactggactaaatacccatccggagtatgcgcgtcgcgtagcagatgaactggcagtc     780
attacctgtgcaccgtttgttaccgcgccgaacaaatttgaagcgctggcgacctgtgat     840
gccctggttcaggcgcacggcgcgttgaaagggttggctgcgtcactgatgaaaatcgcc     900
aatgatgtccgctggctggcctctggcccgcgctgcggaattggtgaaatctcaatcccg     960
gaaaatgagccgggcagctcaatcatgccggggaaagtgaacccaacacagtgtgaggca    1020
ttaaccatgctctgctgtcaggtgatggggaacgacgtggcgatcaacatggggggcgct    1080
tccggtaactttgaactgaacgtcttccgtccaatggtgatccacaatttcctgcaatcg    1140
gtgcgcttgctggcagatggcatggaaagttttaacaaacactgcgcagtgggtattgaa    1200
ccgaatcgtgagcgaatcaatcaattactcaatgaatcgctgatgctggtgactgcgctt    1260
aacacccacattggttatgacaaagccgccgagatcgccaaaaaagcgcataaagaaggg    1320
ctgaccttaaaagctgcggcccttgcgctggggtatcttagcgaagccgagtttgacagc    1380
tgggtacggccagaacagatggtcggcagtatgaaagccgggcgttaaggatccacatac    1440
aggtgcagcc                                                      1450
```

Figure 5.8. Sequence of *fumC* showing NdeI (blue) and BamHI (gray) restrictive sites. The start and stop codons are shown in green and red, respectively.

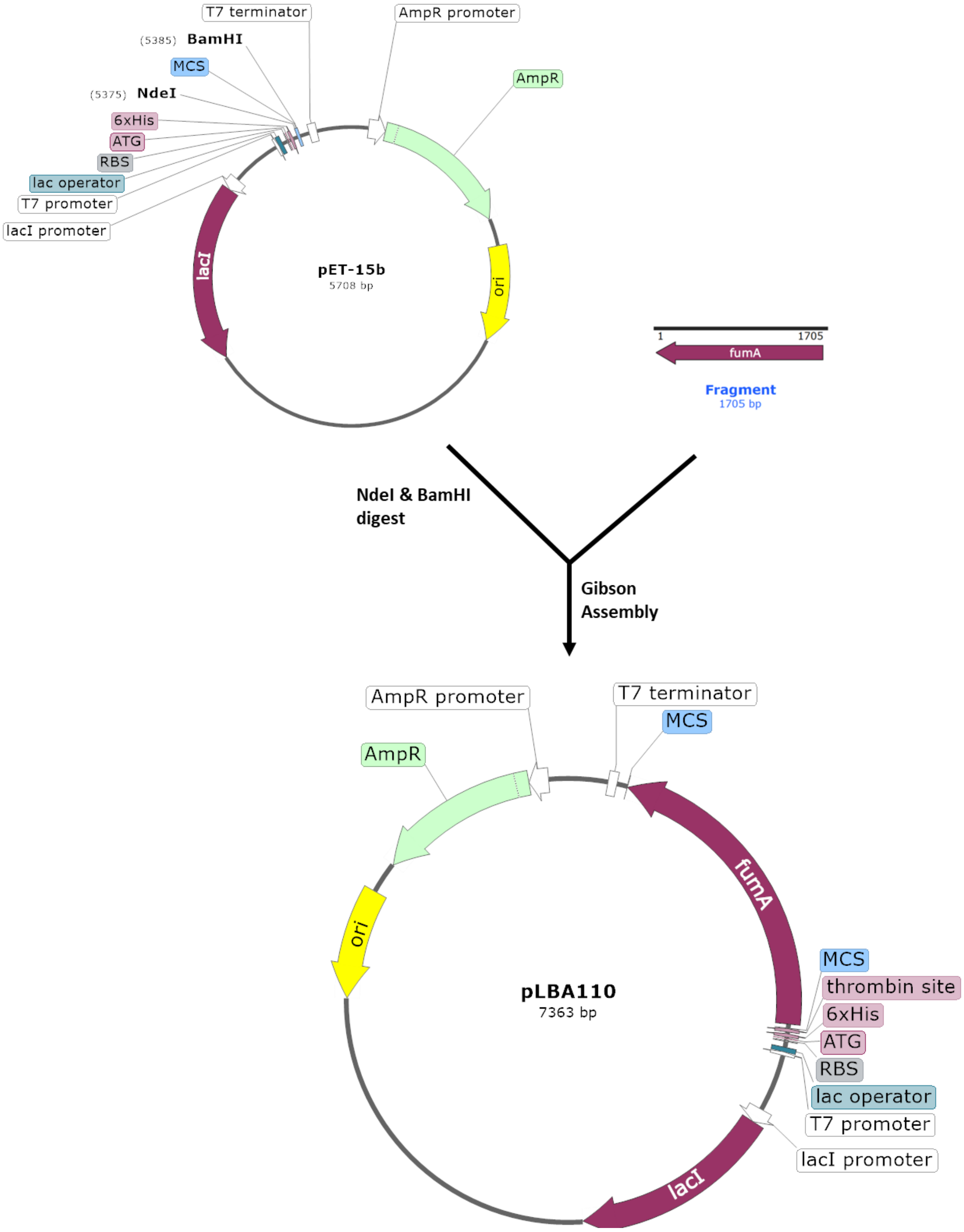


Figure 5.9. Construction of plasmid pLBA110.

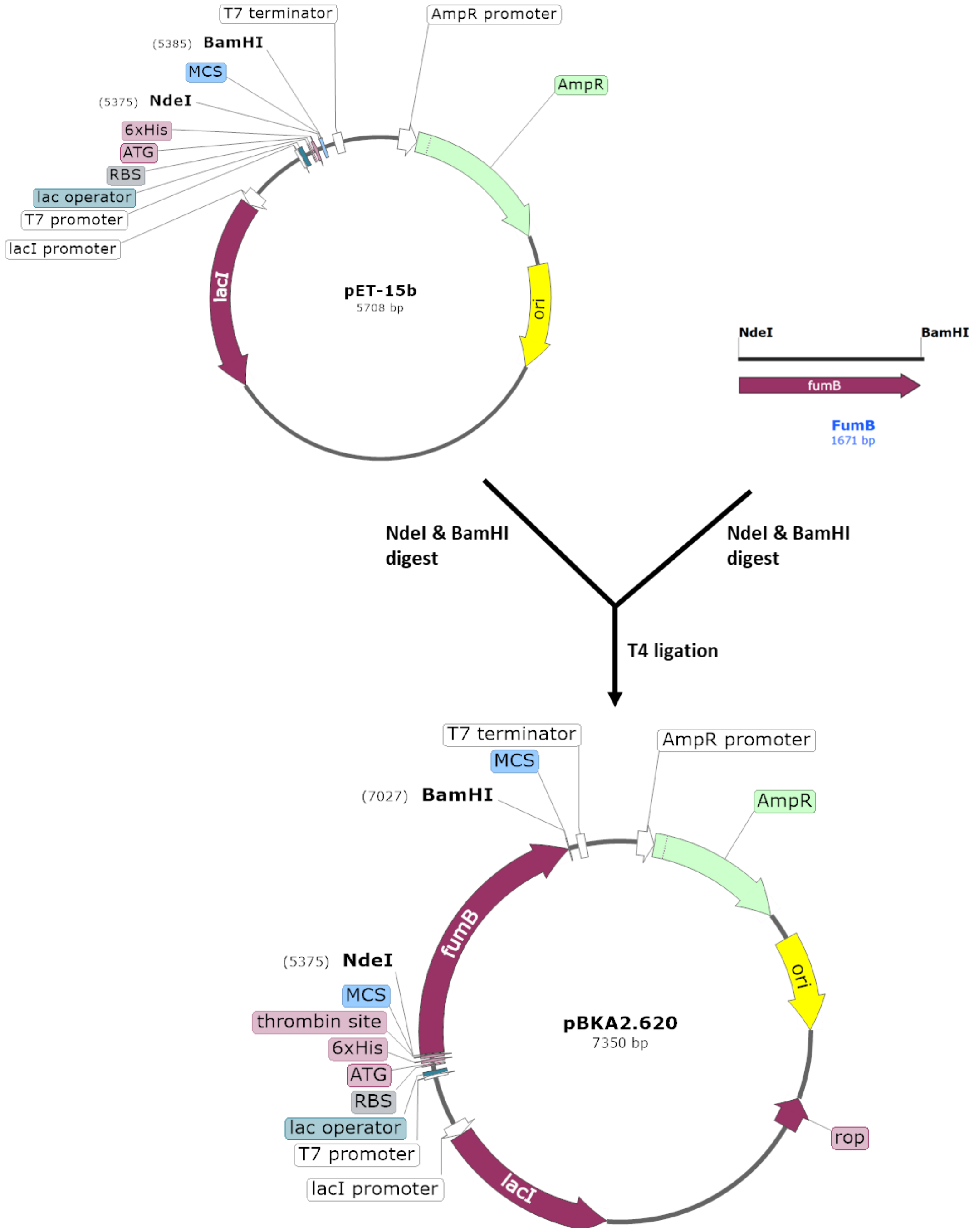


Figure 5.10. Construction of plasmid pBKA1.58.

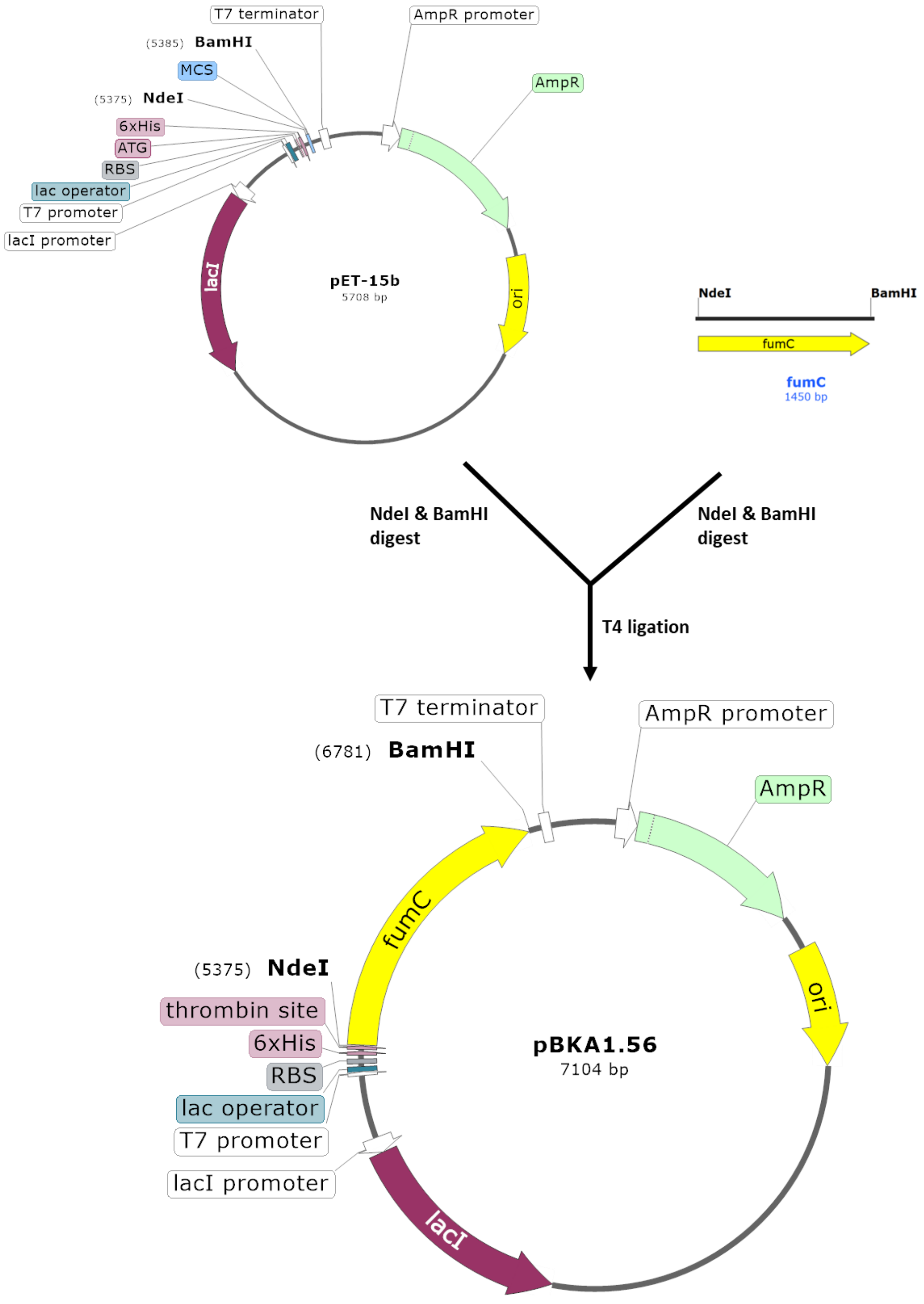


Figure 5.11. Construction of plasmid pBKA1.56.

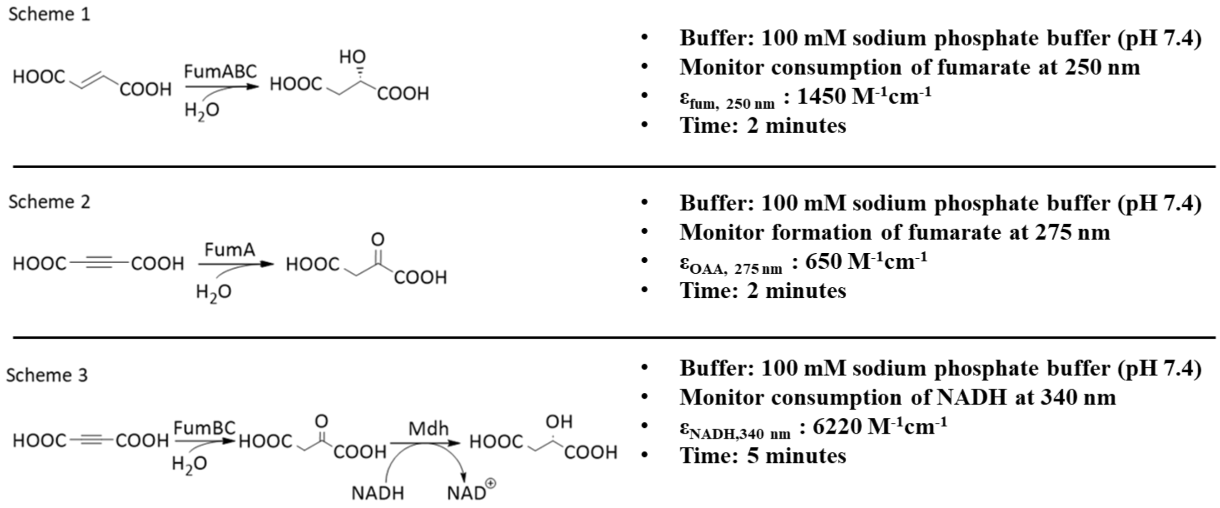
5.8.11.2 Assay for fumarate

Fumarase enzymes were purified using FPLC. The purification process used a Ni^{2+} His trap column, following the established protocol described previously. This method efficiently isolates His-tagged proteins, leveraging the affinity of histidine residues for nickel ions. The enzymatic activity of the purified fumarases was assessed in a direction-specific assay, measuring the conversion of fumarate to malate. The assay was performed at 25 °C using a 100 mM sodium phosphate buffer at pH 7.4. The progress of the reaction was monitored by measuring the decrease in absorbance at 250 nm. This method is based on the characteristic absorption of fumarate, and the specific extinction coefficient ($\varepsilon = 1450\ M^{-1}cm^{-1}$) was used to quantify the reaction progress. Each assay was conducted over a two-min period, allowing for the precise measurement of enzyme activity under the defined conditions. The assays were initiated by adding varying concentrations of fumarate to wells containing 5 ng μL^{-1} of a fumarase enzyme. Typical components of the assay are shown in Table 5.7.

5.8.11.3 Assay for ADCA

The enzymatic activity of FumA in the hydration of acetylenedicarboxylate to oxaloacetate was conducted at 25 °C in 100 mM sodium phosphate buffer, pH 7.4. The formation of oxaloacetate during the reaction was monitored by measuring the increase in absorbance at 275 nm due to oxaloacetate. The specific extinction coefficient ($\varepsilon = 650\ M^{-1}cm^{-1}$) for oxaloacetate was used to quantify the enzymatic activity. Each enzymatic reaction was monitored for 2 min to ensure consistent and comparable results. This time frame allows for the observation of the initial reaction rates necessary for kinetic analysis.

The hydration of ADCA by FumB and FumC was conducted using a coupled enzyme assay. This was necessary because high enzyme concentrations were needed to obtain measurable activity. Using the coupled enzyme assay enabled monitoring the assay at a different wavelength with protein inteference. The formation of oxaloacetate was coupled with it's reduction to malate by malate dehydrogenase (mdh) using NADH. The couple enzyme is monitored by following the consumption of NADH at 340 nm. The extinction coefficient of NADH is 6220 $M^{-1}cm^{-1}$. The wells were filled with all components for the assay except ADCA, which was added to initiate the reaction. A summary of all assays for the fumarases is shown in Figure 5.12



- **Buffer: 100 mM sodium phosphate buffer (pH 7.4)**
- **Monitor consumption of fumarate at 250 nm**
- **$\varepsilon_{fum,\ 250\ nm}$: 1450 $M^{-1}cm^{-1}$**
- **Time: 2 minutes**

- **Buffer: 100 mM sodium phosphate buffer (pH 7.4)**
- **Monitor formation of fumarate at 275 nm**
- **$\varepsilon_{OAA,\ 275\ nm}$: 650 $M^{-1}cm^{-1}$**
- **Time: 2 minutes**

- **Buffer: 100 mM sodium phosphate buffer (pH 7.4)**
- **Monitor consumption of NADH at 340 nm**
- **$\varepsilon_{NADH,340\ nm}$: 6220 $M^{-1}cm^{-1}$**
- **Time: 5 minutes**

Figure 5.12. Assays used for characterizing fumarases with ADCA and fumarate substrates.

Table 5.7. Components for various fumarases assays.

Assay for fumarate hydratase activity of FumA, FumB, and FumC

Component	Stock concentration	Volume	Final concentration
100 mM sodium phosphate buffer (pH 8)		140 µL	
FumA/FumB/FumC	100 ng/µL	10 µL	5 ng/µL
Fumarate	20 mM	50 µl	5 mM
Final		200 µL	

Assay for ADCA hydratase activity of FumA

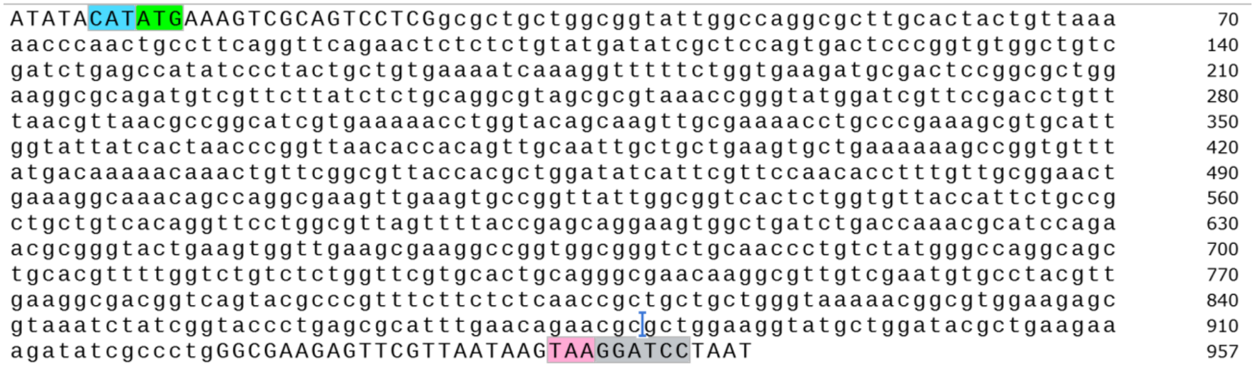
Component	Stock concentration	Volume	Final concentration
100 mM sodium phosphate buffer (pH 8)		140 µL	
FumA	100 ng/µL	10 µL	5 ng/µL
ADCA	20 mM	50 µl	5 mM
Final		200 µL	

Coupled enzyme assay for ADCA hydratase activity of FumB and FumC

Component	Stock concentration	Volume	Final concentration
100 mM sodium phosphate buffer (pH 8)		90 µL	
FumB/FumC	1 mg/mL	10 µL	50 ng/µL
ADCA	20 mM	50 µl	5 mM
Mdh	10 ng/µL	10 µl	0.5 ng/µL
NADH	2 mg/mL	40 µl	0.4 mg/mL
Final		200 µL	

5.8.11.4 Characterization of *E. coli* malate dehydrogenase (Mdh)

The malate dehydrogenase used for the coupled enzyme assay for FumB and FumC was purified and characterized. The gene encoding *E. coli* malate dehydrogenase was amplified from *E. coli* W3110 with primers sets BKA007 and BKA008 (See Primers Table). The amplified *mdh* fragment (Figure 5.13) was cloned into pET-15b vector to obtain pBKA1.412 (Figure 5.14). The protein was overexpression by induction with IPTG and purified on FPLC with Ni^{2+} His trap column. Purified enzyme was stored in 100 mM sodium phosphate buffer containing 20% ethylene glycol. The samples were divided into aliquots which were then flash frozen and stored at -80 °C until use. For assays, aliquots were diluted to 1 ng μL^{-1}. Mdh was characterized using assay with oxaloacetate as substrate. The reduction of oxaloacetate was measured by the decrease in absorbance at 340 nm due to oxidation of NADH. The assay was conducted at 25 °C in 100 mM sodium phosphate buffer (pH 7.4) and was initiated by addition of oxaloacetate to wells containing NADH and 0.05 ng μL^{-1} Mdh. Chromatograms for purification of enzymes and SDS-PAGE analysis are shown in Figure 5.15 – Figure 5.18



```
ATATACATATGAAAGTCGCAGTCCTCGgcgctgctggcggtattggccaggcgcttgcactactgttaaa     70
aacccaactgccttcaggttcagaactctctctgtatgatatcgctccagtgactcccggtgtggctgtc    140
gatctgagccatatccctactgctgtgaaaatcaaaggtttttctggtgaagatgcgactccggcgctgg    210
aaggcgcagatgtcgttcttatctctgcaggcgtagcgcgtaaaccgggtatggatcgttccgacctgtt    280
taacgttaacgccggcatcgtgaaaaacctggtacagcaagttgcgaaaacctgcccgaaagcgtgcatt    350
ggtattatcactaacccggttaacaccacagttgcaattgctgctgaagtgctgaaaaaagccggtgttt    420
atgacaaaaacaaactgttcggcgttaccacgctggatatcattcgttccaacacctttgttgcggaact    490
gaaaggcaaacagccaggcgaagttgaagtgccggttattggcggtcactctggtgttaccattctgccg    560
ctgctgtcacaggttcctggcgttagttttaccgagcaggaagtggctgatctgaccaaacgcatccaga    630
acgcgggtactgaagtggttgaagcgaaggccggtggcgggtctgcaaccctgtctatgggccaggcagc    700
tgcacgttttggtctgtctctggttcgtgcactgcagggcgaacaaggcgttgtcgaatgtgcctacgtt    770
gaaggcgacggtcagtacgcccgtttcttctctcaaccgctgctgctgggtaaaaacggcgtggaagagc    840
gtaaatctatcggtaccctgagcgcatttgaacagaacgcgctggaaggtatgctggatacgctgaagaa    910
agatatcgccctgGGCGAAGAGTTCGTTAATAAGTAAGGATCCTAAT                           957
```

Figure 5.13. Sequence of *mdh* showing NdeI (blue) and BamHI (gray) restrictive sites. The start and stop codons are shown in green and red, respectively.

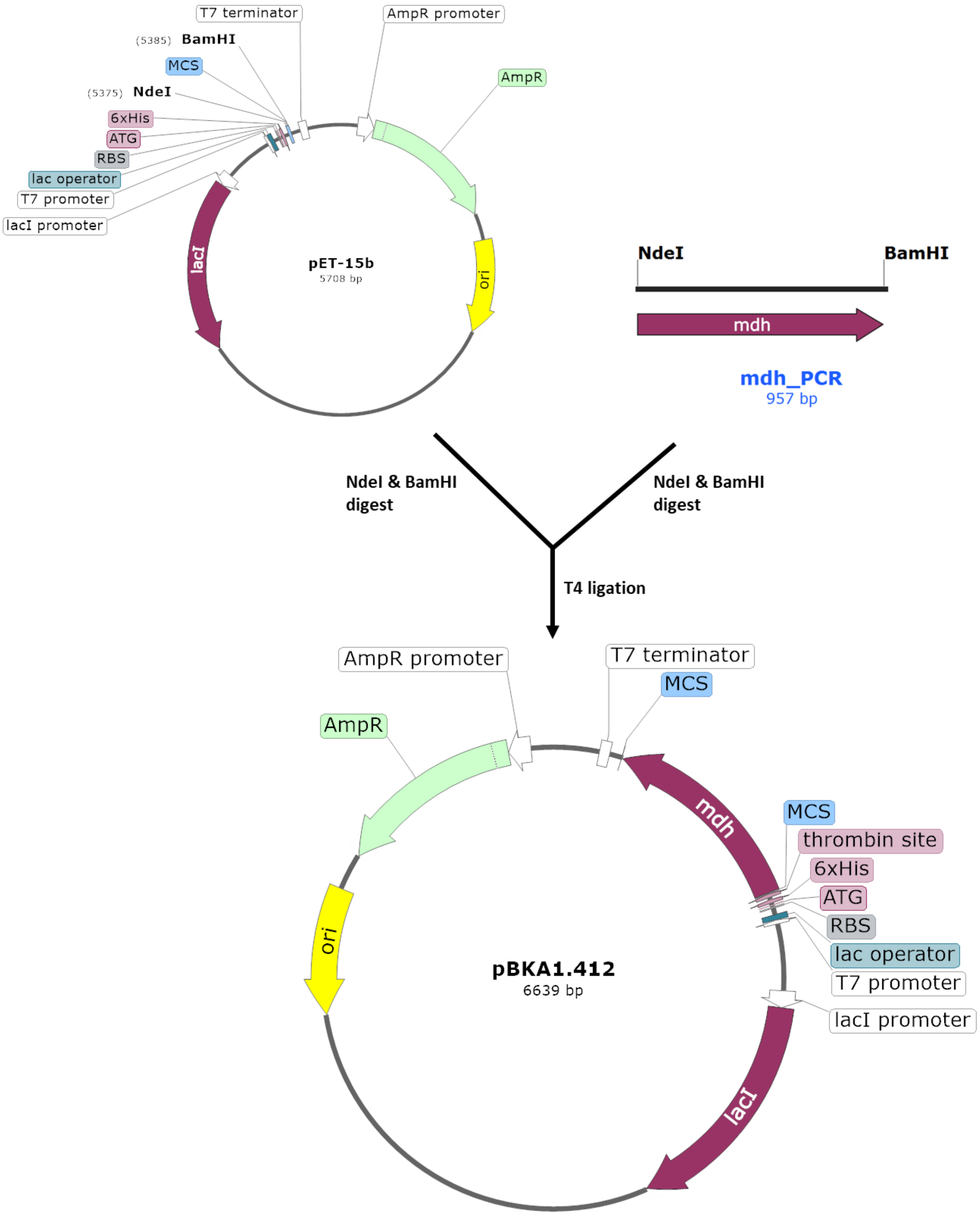


Figure 5.14. Construction of plasmid pBKA1.412.

5.8.11.5 Kinetic parameters of fumarases

To determine kinetic parameters for all fumarases, the Michaelis constant (K_m) and maximum reaction rate (V_{max}) were obtained by fitting initial rates of the reactions to the Michaelis-Menten model. Substrate concentrations were varied systematically around the anticipated K_m value to capture a comprehensive range of enzyme activities, while maintaining a constant enzyme concentration throughout the experiments. Initial reaction velocities (V_0) were quantitatively assessed by monitoring the rate for the first 10 – 15 seconds of the reaction.

The obtained data were first plotted as initial velocity (V_0) versus substrate concentration ([ADCA] or [Fumarate]). For kinetic analysis, data were subjected to two primary methods: Lineweaver-Burk linearization and non-linear regression fitting. The Lineweaver-Burk plot, a double reciprocal transformation of the Michaelis-Menten equation, was utilized to determine kinetic constants, plotting $1/V_0$ against $1/[S]$. The intercepts of the resulting line on the axes provided the inverse of V_{max} (y-intercept) and the negative inverse of K_m (x-intercept). This method also facilitated the evaluation of the goodness of fit, ensuring that the Michaelis-Menten model adequately described the hydratase activity under study. Results were validated through experimental replication, and comparisons of kinetic parameters were conducted to confirm consistency.

UNICORN start 1.1
1/30/2020 5:00:45 PM -05:00
Result: BKA HisTrap FF 5ml gradient FumA 003
Method: BKA HisTrap FF 5ml gradient FumA

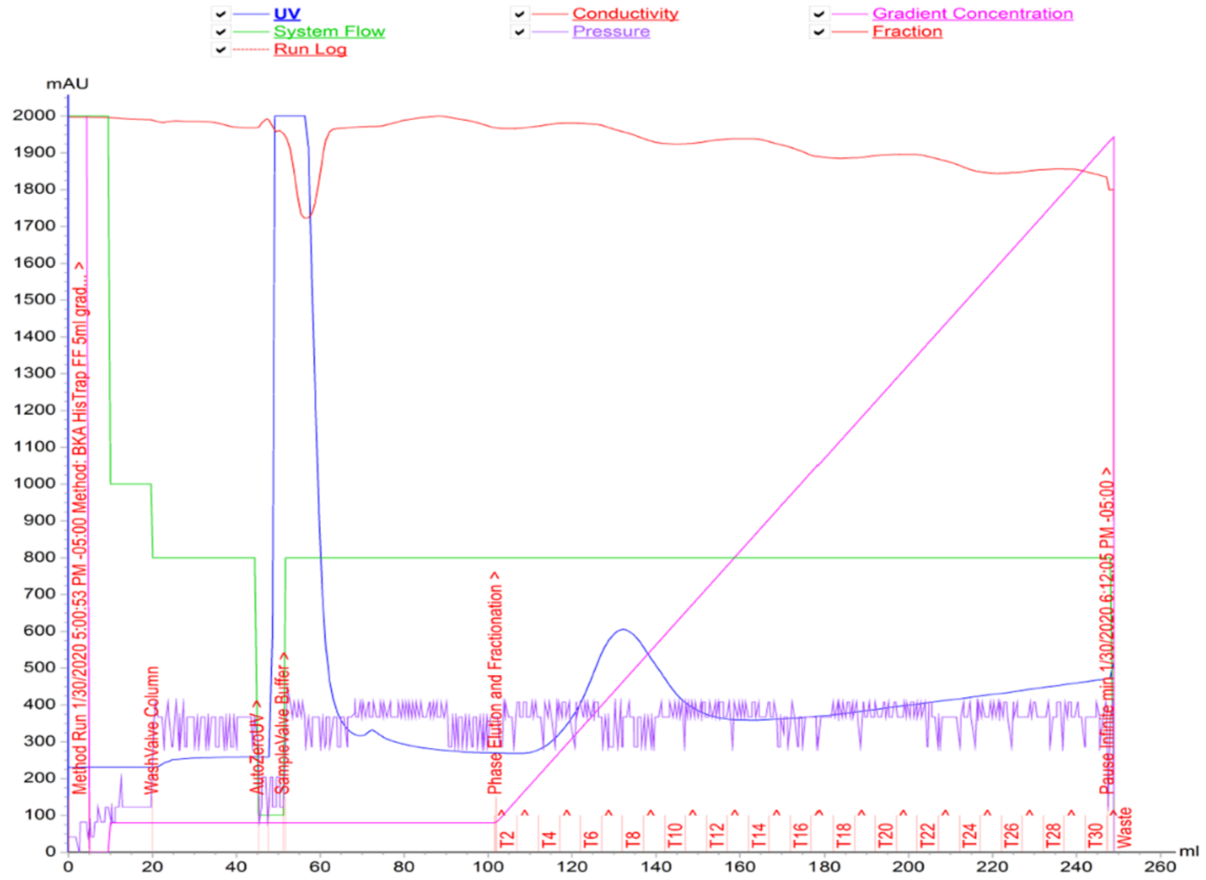
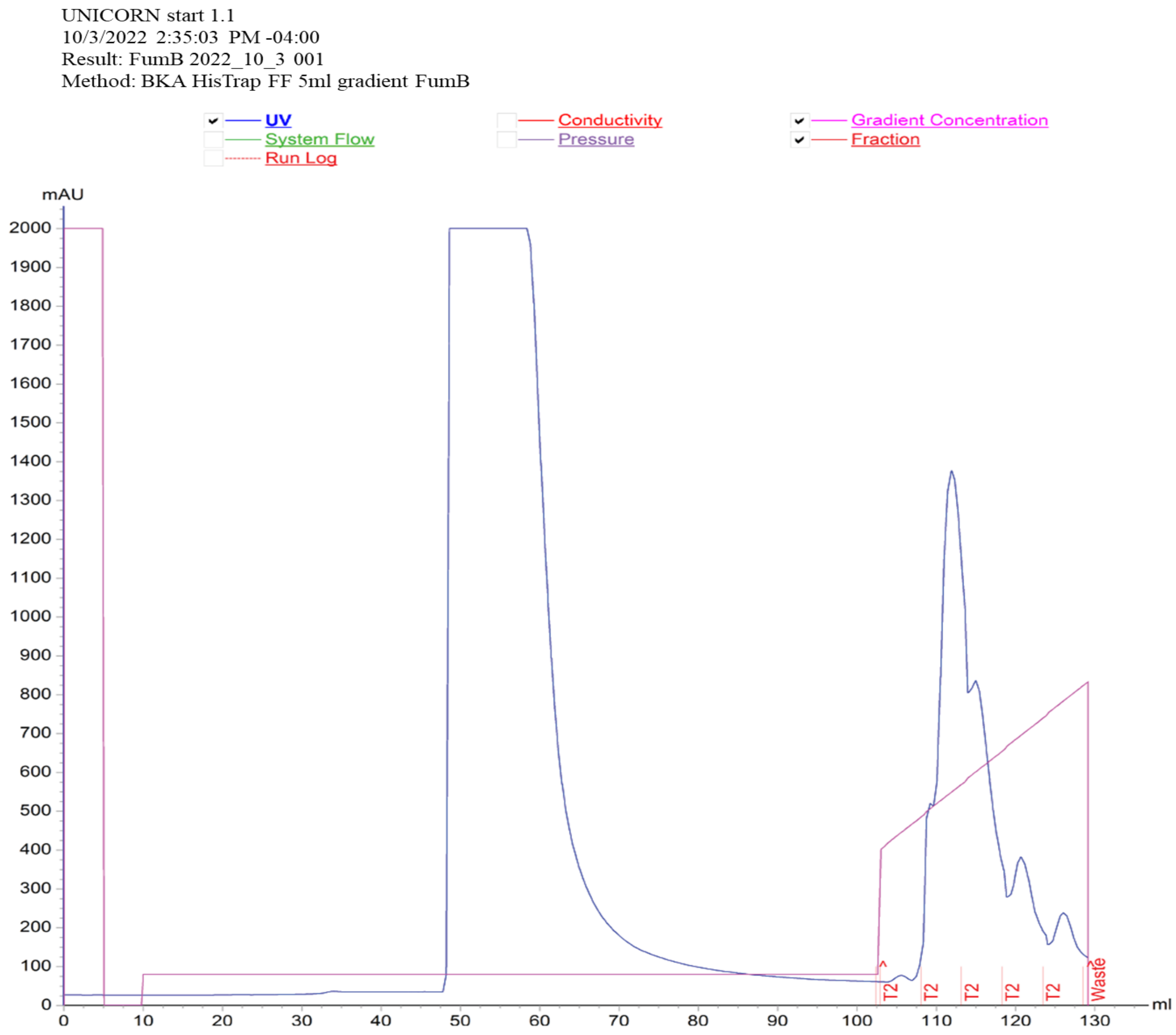
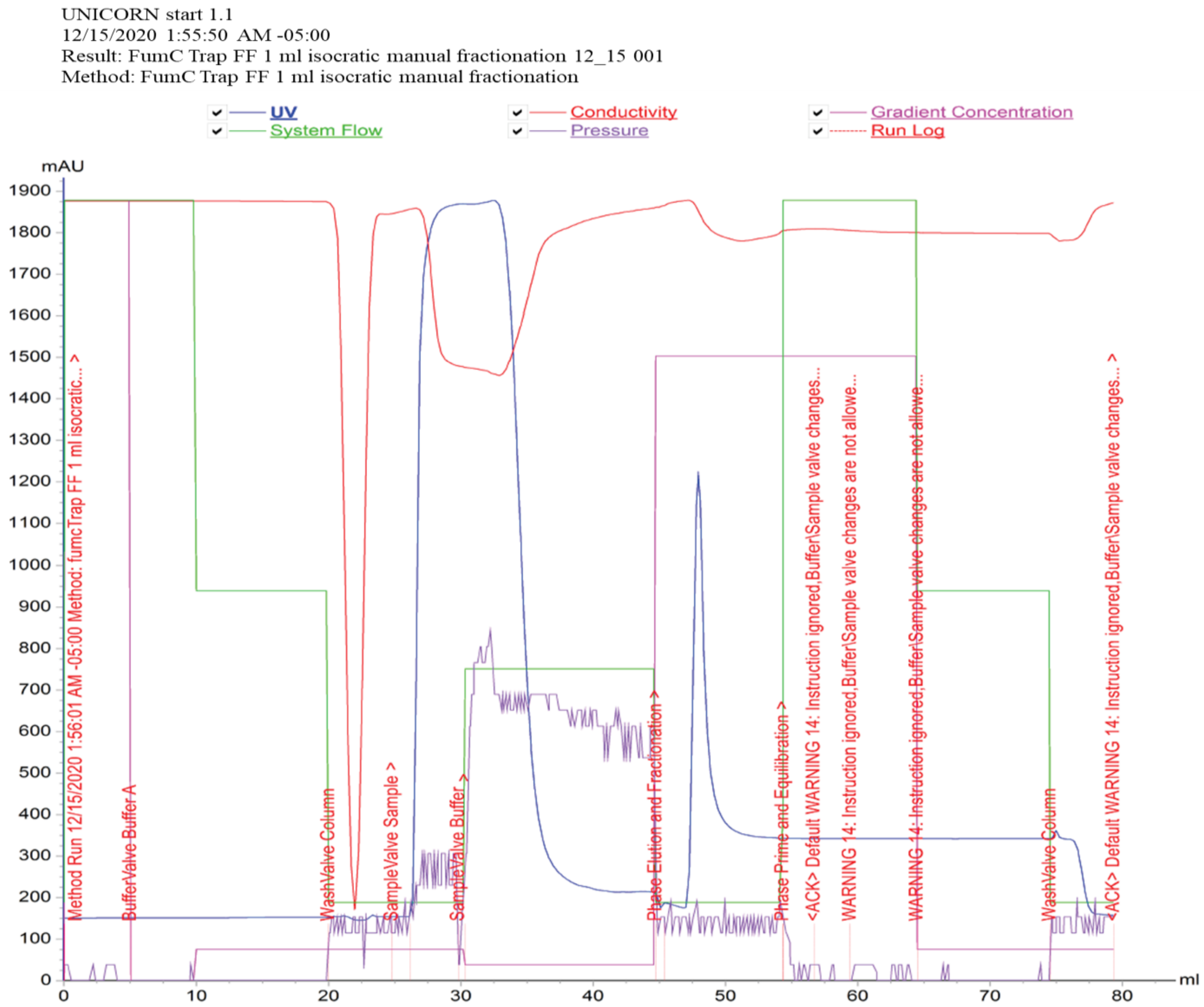


Figure 5.15. Chromatogram for FumA purification



UNICORN start 1.1
10/3/2022 2:35:03 PM -04:00
Result: FumB 2022_10_3 001
Method: BKA HisTrap FF 5ml gradient FumB

Figure 5.16. Chromatogram for FumB purification.



UNICORN start 1.1
12/15/2020 1:55:50 AM -05:00
Result: FumC Trap FF 1 ml isocratic manual fractionation 12_15 001
Method: FumC Trap FF 1 ml isocratic manual fractionation

Figure 5.17. Chromatogram for FumC purification

UNICORN start 1.1
5/24/2021 3:14:52 PM -04:00
Result: 2021_05_23_ MDH 002 001
Method: BKA HisTrap FF 5ml gradient MDH

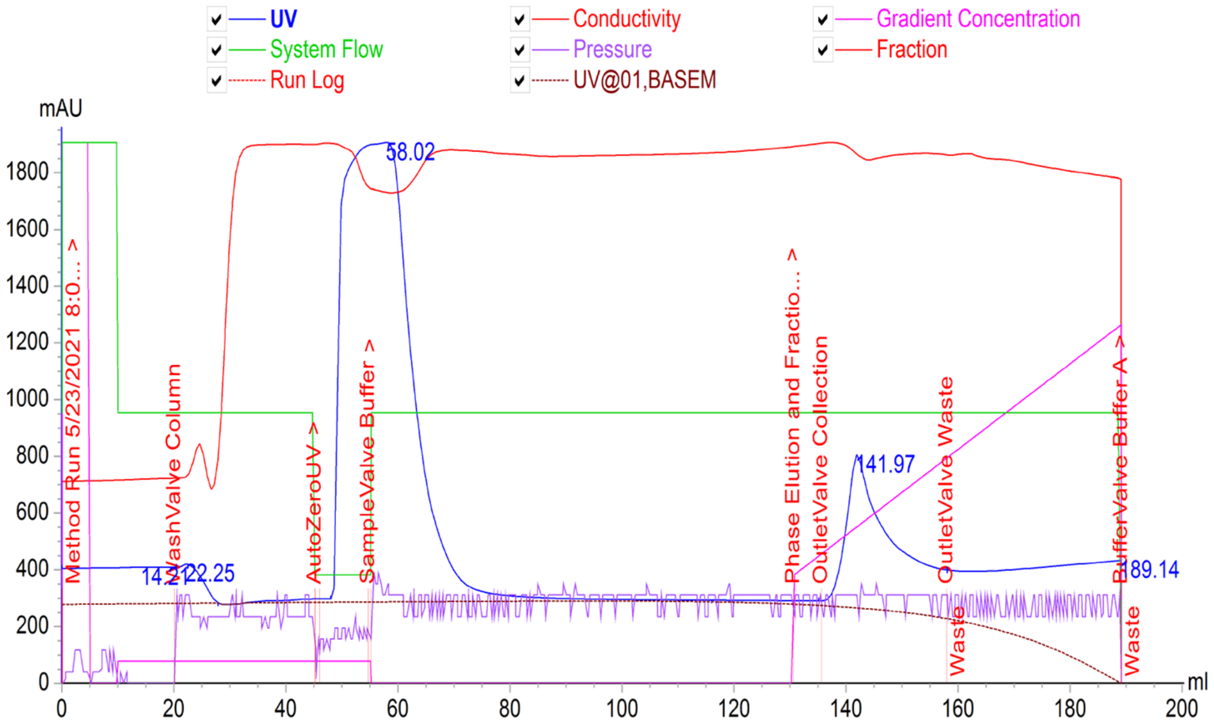


Figure 5.18. Chromatogram for MDH purification.

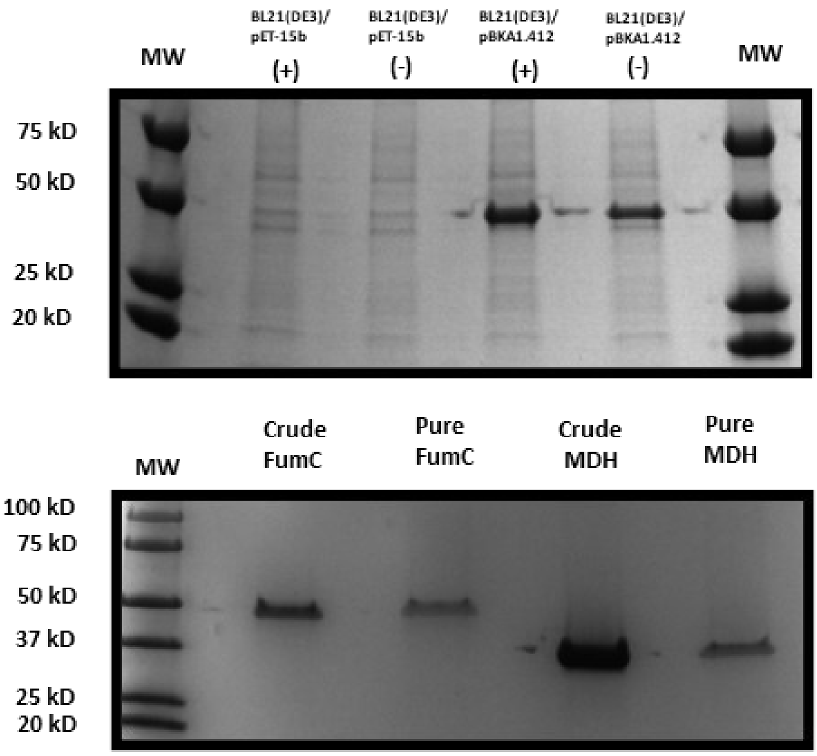
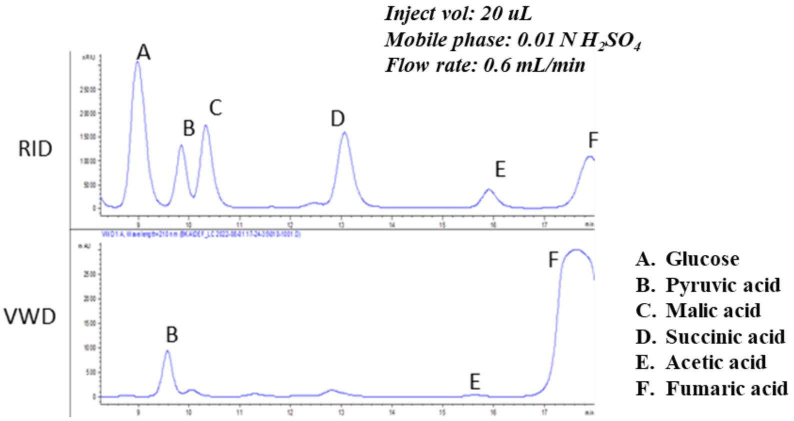


Figure 5.19. SDS-PAGE analysis of crude extract and purified FumC and MDH.

5.8.12 Differential metabolomics of ACG populations and engineered strains

Metabolic analysis of ACG populations (replicates of ACG1, ACG2, ACG3) was performed by GC-MS and HPLC. GC-MS was used for characterizing metabolites and HPLC analysis was used for quantification. Targeted excretion of carboxylic acids were acetic acid, formic acid, pyruvic acid, lactic acid, propionic acid, and succinic acid. For GC-MS analysis, a method that simultaneously extracts and derivatizes short chain fatty acids (SCFA) and subsequent detection by negative chemical ionization was used. For internal standards 7.6 µL of each 100 mM sodium formate ($^{13}C_1$), sodium acetate ($^{13}C_2$), sodium propionate ($^{13}C_3$), and sodium butyrate (^{13}C4) stock was dispensed into a final volume of 10 mL 1:1 methanol:water. About 100 µL of this mixture contains 0.35 µg of each internal standard, which will be to 50 µM in the final hexane extract. Since acetone and acetonitrile often contain trace amounts of acetic acid that could contribute to the background, these solvents were avoided. The derivatization reagent, PFBBr has also been known to contain trace amounts of formate, propionate, and butyrate, and high levels of acetate. In order to remove these background contaminants, a solution of 172 mM PFBBr dissolved in hexane (26 µL in 1 mL hexane) was made and washed three times with equal volumes of milliQ H_2O. The upper hexane layer was retained. The hexane is then evaporated with N_2 without heating. A small volume of PFFBr remaining after hexane has been evaporated is resuspended into an equal volume of methanol. For sample preparation, 100 µL of supernatant is added to 400 µL methanol, 100 µL internal standard solution, 100 µL of 172 mM PFBBr in methanol in a 1.6 microfuge tube, 10 µL diisopropylethylamine and vortexed to mix for 3 min and spin down briefly. This changes the pH to approximately 8. The mixture was incubated at 60 °C for 30 min. The tubes were allowed to cool down on ice and spun down briefly before opening the tubes. About 150 µL hexane and 150 µL 0.9% (w/v) sodium chloride in water is added to the sample vortexed and spun down briefly. The upper hexane layer is pipetted into autosampler vials. The vials are stored in the freezer at -20 °C until the extracts are analyzed.

For HPLC analysis, 1 mL of culture samples were harvested by centrifugation (12, 000 x *g*,10 min, 4 °C). The supernatant was filtered through a 0.22 µm filter and 20 µL were injected onto the HPLC column with flow rate 0.6 mL/min for 18 min using 0.01 N H_2SO_4 as mobile phase. Standards of compounds were used to generate calibration curves (Figure 5.20 and 5.21). The spectrum of ACG3 and identified compounds are shown in Figures 5.22 and 5.23.



Inject vol: 20 uL
Mobile phase: 0.01 N H_2SO_4
Flow rate: 0.6 mL/min

A. Glucose
B. Pyruvic acid
C. Malic acid
D. Succinic acid
E. Acetic acid
F. Fumaric acid

Figure 5.20. HPLC chromatograms of standards.

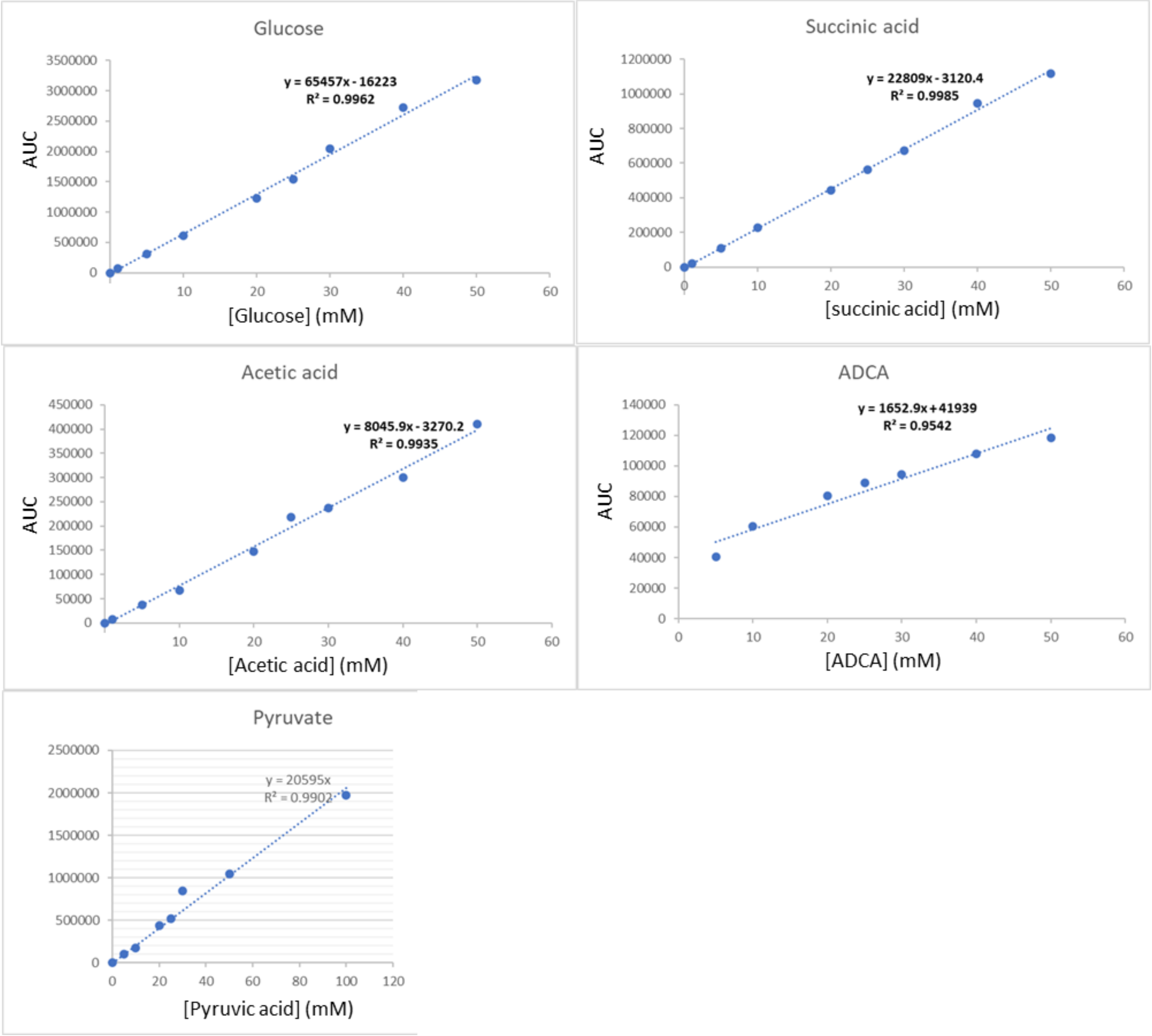


Figure 5.21. Calibration curves used for metabolic analysis.

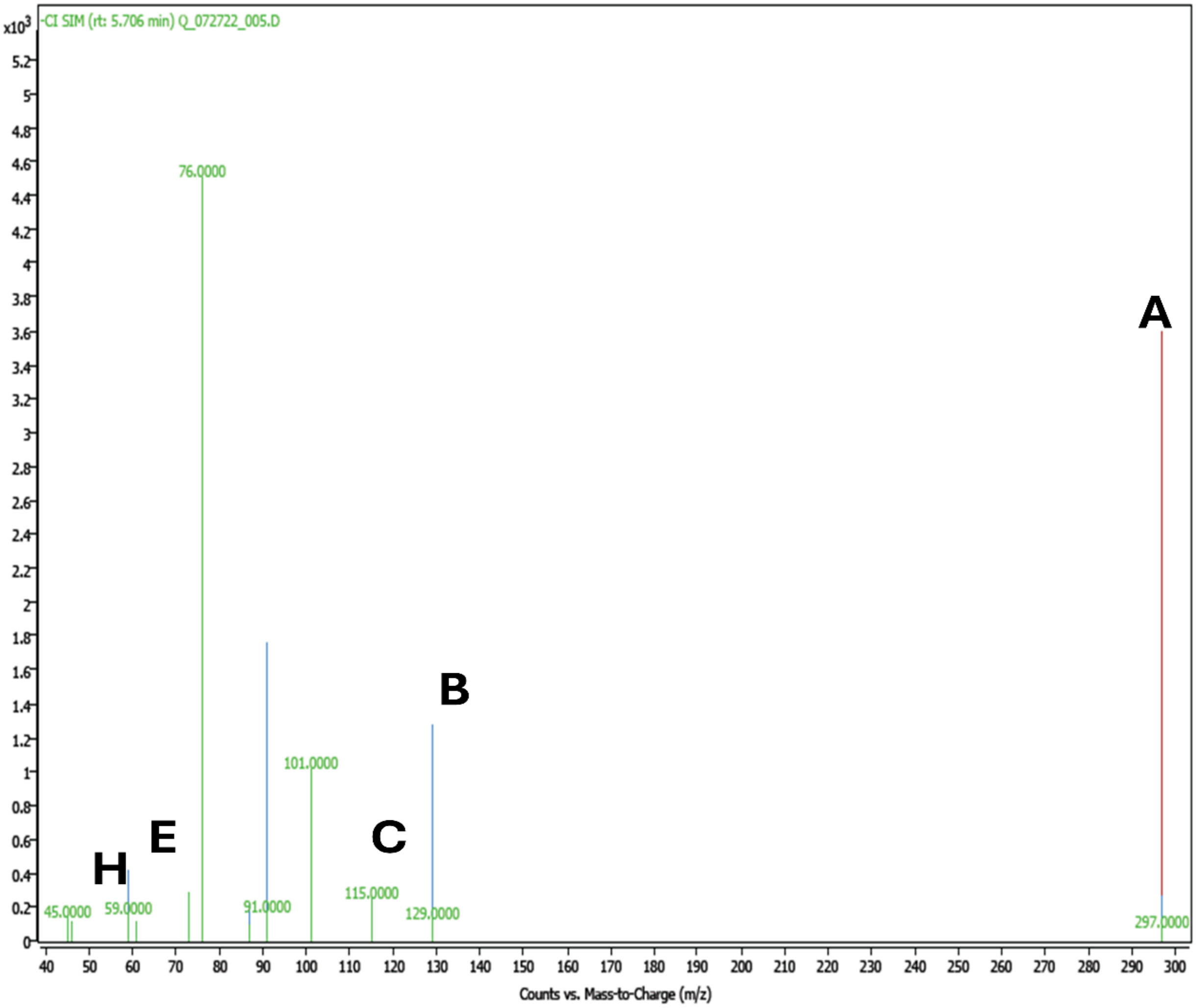


Figure 5.22. GC-MS spectrum of ACG3 sample.

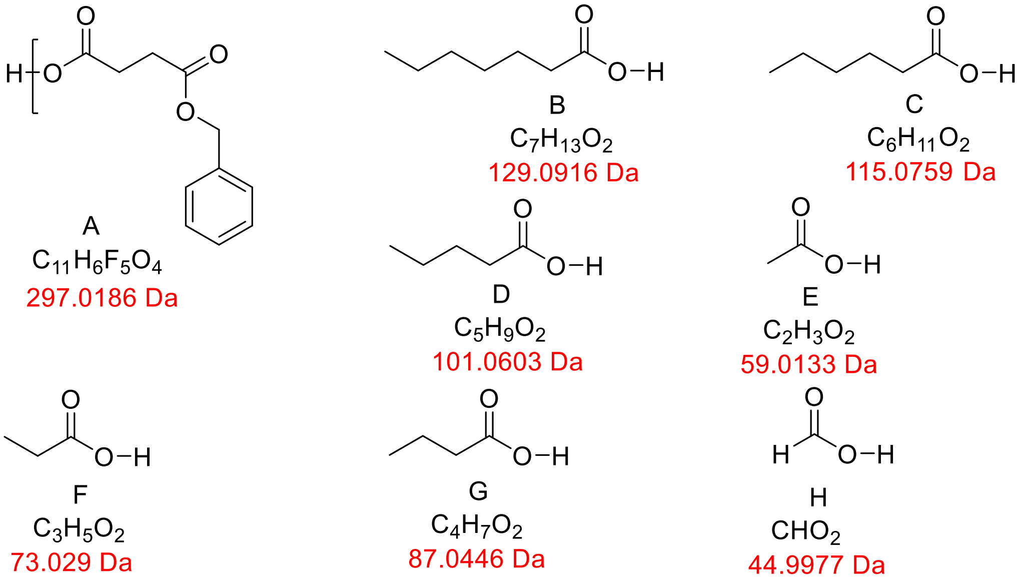


Figure 5.23. Proposed fragmentation patterns of identified compounds based on molecular ion peak.

5.8.13 SEM imaging of ACG populations

For scanning electron microscopy, ACG populations (ACG1_1, ACG2_1, ACG3_1, and ACG3_3) were generated as described before. For each sample, a small quantity of the liquid suspension was mixed with an equal amount of 4% glutaraldehyde in 100 mM sodium phosphate buffer (pH 7.4). Fixation was allowed to proceed for about an h at 4 °C. Two drops of 1% poly-L-Lysine (Sigma-Aldrich P1399) were placed on a plastic petri dish and a 12 mm round glass coverslip was placed on top of the drops and allowed to stand for 10 min. The coverslip was removed and gently washed with several drops of water. The coverslip was drained, but it was not allowed to dry. Two drops of the cells fixed in suspension were placed on the side of the coverslip that previously faced down. The suspension was allowed to settle for 10 min. Subsequently, the coverslip was gently washed with several drops of water and placed in a graded ethanol series (25%, 50%, 75%, 95%) for ten min at each step, with three 10-min changes in 100% ethanol. The samples were subjected to critical point drying in a Leica Microsystems model EM CPD300 critical point dryer (Leica Microsystems, Vienna, Austria) using carbon dioxide as the transitional fluid. The samples were mounted on aluminum stubs using epoxy glue (System Three Quick Cure 5, System Three Resins, Inc., Aubur, WA). The samples were coated with osmium (approximately 10 nm thickness) in a Tennant20 osmium chemical vapor deposition (CVD) coater (Meiwafosis Co., Ltd., Osaka, Japan). The samples were then examined in a JEOL 7500F field emission scanning electron microscope (JEOL Ltd., Tokyo, Japan).

5.9 Chapter Three: Characterization of ALE-inspired malic enzyme B variants

5.9.1 Materials and solutions

All chemicals, including malate, oxaloacetate, NADH, NAD^+, $MgCl_2$, Mn_2SO_4, Tris-Base were purchased from Sigma-Aldrich (St. Louis, MO, USA). Buffer solutions were prepared using ddH_2O.

5.9.2 Bacterial strains, plasmids, and growth condition

E. coli DH5α was used in all cloning steps. *E. coli* BL21(DE3) was used for protein overexpression and purification. Strains were cultivated in LB containing ampicillin (50 µg mL^{-1}) where appropriate. The gene encoding *E. coli* malic enzyme B was amplified by PCR from W3110 genomic DNA using primer sets BKA200 and BKA201 obtained from Integrated DNA Technologies (IDT^{TM}). The primers introduced unique restrictive sites XhoI and BamHI at the 5′ and 3′ ends of the PCR fragment (Figure 5.24 – 5.25), respectively. The fragment was cloned into pET-15b to generate pBKA2.202 (Figure 5.24). The construct was designed to place an N-terminal His_6-tag on MaeB for purification on Ni^{2+} His trap column. All strains and plasmids used in this study are shown in Table 5.8. All primers used in this study are shown in Table 5.9.

Table 5.8. Bacterial strains and plasmids used in this study.

Strains	**Genotype/ Description**	**Reference/ Source**
E. coli DH5α	F-, *Δ(argF-lac)169*, φ80d*lacZ58*(M15), *ΔphoA8*, *glnX44*(AS), λ^-, *deoR481*, *rfbC1?*, *gyrA96*(NalR), *recA1*, *endA1*, *thiE1*, *hsdR17*	Invitrogen
E. coli BL21(DE3)	F-, *lon-11*, *Δ(ompT-nfi:A)885*, *Δ(galM-ybhJ)884*, *λDE3 [lacI, lacUV5-T7 gene 1, ind1, sam7, nin5]*, *Δ46*, *[mal*$^+$*]*$_{K\text{-}12}$(λ^S), *hsdS10*	Invitrogen
E. coli W3110	F-, λ^-, *IN(rrnD-rrnE)1*, *rph-1*	Invitrogen

plasmids		
pET-15b	Ap^R, *lacI*, P$_{T7}$, ori (3882) replicon	Invitrogen
pBKA2.202	*maeB* cloned into pET-15b	This study
pBKA2.203	MaeB Y145E plasmid (Q5 mutagenesis on pBKA2.202)	This study
pBKA2.204	MaeB K149A plasmid (Q5 mutagenesis on pBKA2.202)	This study
pBKA2.205	MaeB R151I plasmid (Q5 mutagenesis on pBKA2.202)	This study
pBKA2.206	MaeB N155D plasmid (Q5 mutagenesis on pBKA2.202)	This study
pBKA2.207	MaeB A338V plasmid (Q5 mutagenesis on pBKA2.202)	This study

Table 5.9. Primers used in this study.

Primer ID	**SEQUENCE**	**Purpose**
BKA009_MaeB	ATTATCTCGAGATGGATGACCAGTTAAAACAAAG	For cloning *maeB* in pET-15b
BKA0010_MaeB	ATATTGGATCCAAATTACAGCGGTTGGGTTTG	
BKA011_MaeB_Y145E	AGAATGTTTCGAAATTGAACAGAAACTG	Q5 site directed mutagenesis to generate MaeB variants
BKA012_MaeB_Y145E	GGCGCTTTAATGTCTTC	
BKA013_MaeB_K149A	TAGAAACATTCTGGCG	
BKA014_MaeB_K149A	TATTGAACAGGCGCTGCGCGAGC	
BKA015_MaeB_R151C	TCAATATAGAAACATTCTGGC	
BKA016_MaeB_R151C	ACAGAAACTGATTGAGCGGATGAATATTC	
BKA017_MaeB_N155D	CGAGCGGATGGATATTCCGGTAT	
BKA018_MaeB_N155D	CGCAGTTTCTGTTCAATATAGAAAC	
BKA019_MaeB_A338V	TGGCGCAACCGTGATCAACGAAGAGATG	
BKA020_MaeB_A338V	ACGTCCAGCGCGCCA	

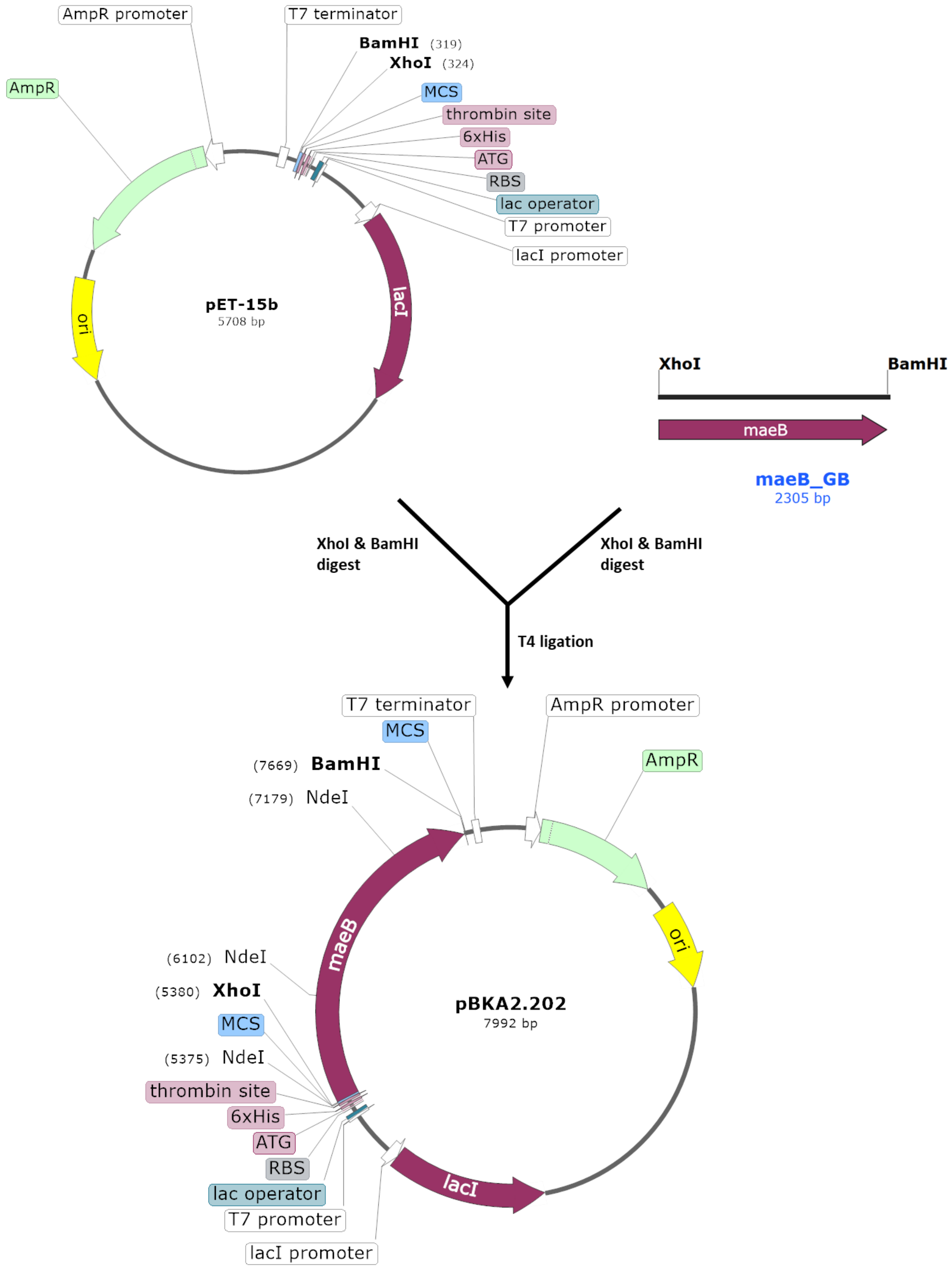


Figure 5.24. Cloning of pBKA2.202.

```
ATTATCTCGAGatggatgaccagttaaaacaaagtgcacttgatttccatgaatttccagttccagggaaaatccaggtttctccaacca      90
agcctctggcaacacagcgcgatctggcgctggcctactcaccaggcgttgccgcaccttgtcttgaaatcgaaaaagacccgttaaaag     180
cctacaaatataccgcccgaggtaacctggtggcggtgatctctaacggtacggcggtgctggggttaggcaacattggcgcgctggcag     270
gcaaaccggtgatggaaggcaagggcgttctgtttaagaaattcgccgggattgatgtatttgacattgaagttgacgaactcgacccgg     360
acaaatttattgaagttgtcgccgcgctcgaaccaaccttcggcggcatcaacctcgaagacattaaagcgccagaatgtttctatattg     450
aacagaaactgcgcgagcggatgaatattccggtattccacgacgatcagcacggcacggcaattatcagcactgccgccatcctcaacg     540
gcttgcgcgtggtggagaaaaacatctccgacgtgcggatggtggtttccggcgcgggtgccgcagcaatcgcctgtatgaacctgctgg     630
tagcgctgggtctgcaaaaacataacatcgtggtttgcgattcaaaaggcgttatctatcagggccgtgagccaaacatggcggaaacca     720
aagccgcatatgcggtggtggatgacggcaaacgtaccctcgatgatgtgattgaaggcgcggatattttcctgggctgttccggcccga     810
aagtgctgacccaggaaatggtgaagaaaatggctcgtgcgccaatgatcctggcgctggcgaacccggaaccggaaattctgccgccgc     900
tggcgaaagaagtgcgtccggatgccatcatttgcaccggtcgttctgactatccgaaccaggtgaacaacgtcctgtgcttcccgttca     990
tcttccgtggcgcgctggacgttggcgcaaccgccatcaacgaagagatgaaactggcggcggtacgtgcgattgcagaactcgcccatg    1080
cggaacagagcgaagtggtggcttcagcgtatggcgatcaggatctgagctttggtccggaatacatcattccaaaaccgtttgatccgc    1170
gcttgatcgttaagatcgctcctgcggtcgctaaagccgcgatggagtcgggcgtggcgactcgtccgattgctgatttcgacgtctaca    1260
tcgacaagctgactgagttcgtttacaaaaccaacctgtttatgaagccgatttctcccaggctcgcaaagcgccgaagcgcgttgttc    1350
tgccggaaggggaagaggcgcgcgttctgcatgccactcaggaactggtaacgctgggactggcgaaaccgatccttatcggtcgtccga    1440
acgtgatcgaaatgcgcattcagaaactgggcttgcagatcaaagcgggcgttgattttgagatcgtcaataacgaatccgatccgcgct    1530
ttaaagagtactggaccgaatacttccagatcatgaagcgtcgcggcgtcactcaggaacaggcgcagcgggcgctgatcagtaacccga    1620
cagtgatcggcgcgatcatggttcagcgtggggaagccgatgcaatgatttgcggtacggtgggtgattatcatgaacattttagcgtgg    1710
tgaaaaatgtctttggttatcgcgatggcgttcacaccgcaggtgccatgaacgcgctgctgctgccgagtggtaacacctttattgccg    1800
atacatatgttaatgatgaaccggatgcagaagagctggcggagatcaccttgatggcggcagaaactgtccgtcgttttggtattgagc    1890
cgcgcgttgctttgttgtcgcactccaactttggttcttctgactgcccgtcgtcgagcaaaatgcgtcaggcgctggaactggtcaggg    1980
aacgtgcaccagaactgatgattgatggtgaaatgcacggcgatgcagcgctggtggaagcgattcgcaacgaccgtatgccggacagct    2070
ctttgaaaggttccgccaatattctggtgatgccgaacatggaagctgcccgcattagttacaacttactgcgtgtttccagctcggaag    2160
gtgtgactgtcggcccggtgctgatgggtgtggcgaaaccggttcacgtgttaacgccgatcgcatcggtgcgtcgtatcgtcaacatgg    2250
tggcgctggccgtggtagaagcgcaaacccaaccgctgtaatttGGATCCAATAT                                       2305
```

Figure 5.25. Sequence of *maeB* showing start and stop codons in green and red respectively. XhoI and BamHI cut sites are shown in blue and gray , respectively.

5.9.3 Q5 site-directed mutagenesis

Specific primers designed for Q5 site directed mutagenesis targeting specific residues were purchased from IDT (Table). The desired mutations were encoded on either forward or reverse primers. Q5 site directed mutagenesis was performed using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs, Ipswich, MA, USA) following the manufacturer's protocol. The reaction mixture components and the PCR program are shown in Tables 5.9.2.2. Post-amplification, the methylated template DNA was digested by adding 1 µL of DpnI directly to the reaction mixture and incubating at 37 °C for 1 h. The DpnI-treated mixture was transformed into chemically competent *E. coli* DH5α cells by heat shock at 42 °C for 30 seconds, followed by rapid cooling on ice. Transformed cells were then recovered in SOC/LB medium at 37 °C for 1 h without shaking. The cells were plated on LB agar plates containing 50 µg mL^{-1} ampicillin to ensure selection of transformants carrying plasmids. Colonies were picked and grown in LB medium with ampicillin. Plasmid DNA was extracted using a plasmid miniprep kit (QIAGEN) and shipped for Sanger sequencing to confirm the presence of the desired mutation in *maeB* at the MSU Genomics Core. Sequencing results were analyzed using the current version of SnapGene software.

5.9.4 Protein overexpression and purification

The confirmed plasmids were transformed into BL21(DE3) for overexpression. Colonies of transformants were used to produce seed culture in 20 mL of LB with ampicillin (50 μg mL^{-1}) in 250 mL shake flasks. The overnight seed cultures were used to inoculate 1 L LB media (with ampicillin) to an initial OD_{600} of 0.05. The inoculum was incubated at 37 °C with 200 rpm agitation for about 2 – 3 h when the OD_{600} reached between 0.5 and 0.7. At this point the cultures were induced by adding 0.1 mM isopropyl β-D-thiogalactopyranoside (IPTG) and incubated at 30 °C as before for 12 h. Cells were then harvested by centrifugation (10,000 x *g* , 10 min, RT) then resuspended in lysis buffer. The cells were lyzed with a French press at 1100 psi, centrifuged at 20, 000 x*g* for 30 min, then filtered through a 0.45 μm filter to obtain the crude extract. MaeB was purified from crude extract using HisTrap™ FF pre-packed (1 mL and 5 mL) columns on ÄKTA Start FPLC from Cytiva. The chromatogram for MaeB (wt) is shown in Figure 5.24. The crude extract was quantified using a Bradford assay, while the purified MaeB was purified using the guanidine-HCl method on the Nanodrop instrument.

5.9.5 Enzyme kinetic assay

All enzyme assays were conducted using a Molecular Devices SpectraMax® iD3 multi-mode microplate reader. Assays volumes were 200 μL and performed at 25 °C using 100 Mm Tris-HCl (pH 8). All reactions were initiated by addition of substrates containing all other reaction components. All reactions were run in triplicates. The obtained data were first plotted as initial velocity (V_0) versus substrate concentration. For kinetic analysis, data were subjected to two primary methods: Lineweaver-Burk linearization and non-linear regression fitting (Hill equation).

The Lineweaver-Burk plot, a double reciprocal transformation of the Michaelis-Menten equation, was utilized to determine kinetic constants, plotting $1/V_0$ against $1/[S]$. The intercepts of the resulting line on the axes provided the inverse of V_{max} (y-intercept) and the negative inverse of K_m (x-intercept). This method also facilitated the evaluation of the goodness of fit, ensuring that the Michaelis-Menten model adequately describes the enzymatic activities. Results were validated through experimental replication, and comparisons of kinetic parameters were conducted to confirm consistency.

5.9.6.1 Malate oxidative decarboxylase activity

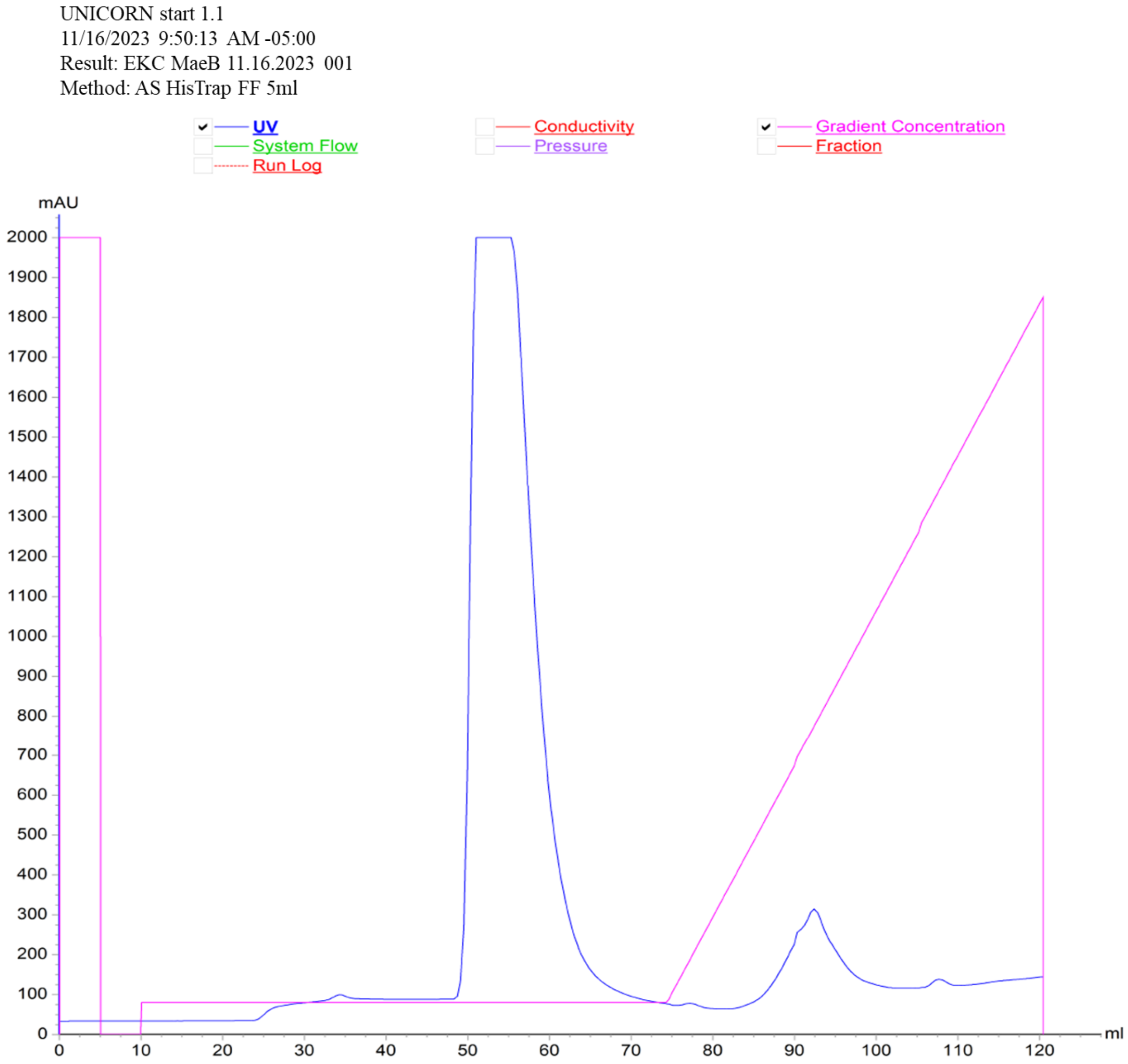
The oxidative decarboxylase activity of MaeB and variants was monitored for the conversion of malate to pyruvate. The enzyme assay was followed by measuring the change in absorbance associated with the reduction of NAD^+ to NADH. Control reactions without substrate or enzyme were included to correct for any non-enzymatic changes in absorbance. The change in absorbance was measured at 7 seconds interval for a 5 min run. The rates of NADH production was determined from the linear portion of the curve, using molar extinction coefficient of NADH (6220 $M^{-1}cm^{-1}$).

5.9.6.2 Oxaloacetate decarboxylase activity

The decarboxylation of oxaloacetate to pyruvate by MaeB (wt) and variants was monitored by measuring the decrease in absorbance at 275 nm, corresponding to the consumption of oxaloacetate. Assays without enzyme were performed to assess non-enzymatic consumption of oxaloacetate. The absorbance at 275 nm was recorded at 30 seconds intervals for 15 min. The initial rate of oxaloacetate consumption was calculated from the slope of the absorbance decrease during the first 2 min of the reaction, using molar extinction coefficient of oxaloacetate (650 $M^{-1}cm^{-1}$).

5.9.6.3 Oxaloacetate reductase activity

The reduction of oxaloacetate to malate eas monitored by measuring the decrease in absorbance at 340 nm due to oxidation of NADH or NADPH. Control reactions without enzymes were conducted to measure any non-enzymatic reduction of oxaloacetate. The decrease in absorbance at 340 nm was measured every 30 seconds for 10 min. The initial rate of NADH or NADPH consumption was calculated from the linear regions of the curve, using molar extinction coefficient of N6220 $M^{-1}cm^{-1}$.



UNICORN start 1.1
11/16/2023 9:50:13 AM -05:00
Result: EKC MaeB 11.16.2023 001
Method: AS HisTrap FF 5ml

Figure 5.26. Chromatogram for MaeB (wt) purification.

5.10 Chapter Four: Bioconversion of acetylenedicarboxylic acid to chemicals

5.10.2 Bacterial strains and plasmids

All bacterial strains and plasmids used in Chapter Four are shown in Table 5.10. For *in vivo* lactate production experiments, plasmids used possessed a P_{trc} promoter upstream the set of genes. The FumA gene was PCR amplified from *E. coli* W3110 and cloned into pBbA1a-RFP, replacing the mRFP gene using NdeI and BamHI restricitive sites to afford pBKA1.418. PCR fragments of gene encoding MaeB and LdhA were separately cloned into pBKA1.418 using EcoRI and BglII restrictive sites to afford pBKA1.1201 and pBKA1.1202, respectively. All cloning strategies are shown in Figure 5.27 – 5.29. All strains, plasmids and primers are listed in Table 5.10 and 5.11.

The *maeB* gene was then cloned into pBKA1.1201 between BamHI and XhoI sites to obtain pBKA1.1203. These plasmids were then transformed into MG1655 wild-type (W1) to obtain W2 (pBKA1.418), W3 (pBKA1.1201), W4 (pBKA1.1202), and W5 (pBKA1.1203). A similar transformation of plasmids into Δ*dcuC* afforded C2, C3, C4, and C5, respectively. In order to use expression of FumA as selective maker Δ*fumA* strain was transformed with plasmid pBKA1.1203 harboring all three genes. Three colonies of the transformants, Δ*fumA*/pBKA1.1203_123 were investigated for lactate production. Clones of ACG10 were also transformed with pBKA1.1203 to afford ACG10_X/pBKA1.1203, where X are different clones. For *in vitro* lactate production assays, FumA was overexpressed using pLBA110 plasmid and MaeBN155D using pBKA2.205 described earlier. For shikimic acid production, the RB791 *serA*::*aroB* Δ*aroL* Δ*aroK* shikimic acid producing (SP) strain was obtained from the Draths strain collection. This strain was evolved for ten rounds to obtain SPACG1 through SPACG10. Clones of the SPACG10 population (SPACG10_X) were transformed with plasmid pKD12.138a to afford SPACG10_Xa.

Table 5.10. Strains and plasmids used in this study.

Strains	Genotype/ Description	Reference/ Source
MG1655	F$^-$ λ^- *ilvG*$^-$ *rfb*-50 *rph*-1	ATCC
RB791	*lacIp-4000(lacI*^{Q}*)*, λ^-, *IN(rrnD-rrnE)1*	ATCC
MG1655/pFumA (W2)	MG1655/pBKA1.418	This study
MG1655/pFumA+MaeB (W3)	MG1655/pBKA1.1201	This study
MG1655/pFumA+LdhA (W4)	MG1655/pBKA1.1202	This study
MG1655/pFumA+MaeB+LdhA (W5)	MG1655/pBKA1.1203	This study
Δ*dcuC*/pFumA (C2)	MG1655 Δ*dcuC*/pBKA1.418	This study
Δ*dcuC*/pFumA+MaeB (C3)	MG1655 Δ*dcuC*/pBKA1.201	This study
Δ*dcuC*/pFumA+LdhA (C4)	MG1655 Δ*dcuC*/pBKA1.202	This study
Δ*dcuC*/pFumA+MaeB+LdhA (C5)	MG1655 Δ*dcuC*/pBKA1.203	This study
Δ*fumA*	MG1655 Δ*fumA*	This study
Δ*fumA*/pFumA+MaeB+LdhA _123(LAP123)	MG1655 Δ*fumA*/pBKA1.203	This study
ACG10_X	Clone X of ACG10	This study
ACG10_X/pFumA+MaeB+LdhA (LAPX)	ACG10_X/pBKA1.203	This study
SP1	RB791 *serA*::*aroB* Δ*aroL* Δ*aroK*	Draths Lab
SP2	RB791 *serA*::*aroB* Δ*aroL* Δ*aroK*/pKD12.138a	Draths Lab
SPACG1 - 10	SP1 after round 1 - 10 of evolution in ADCA	This study
SPACG10_X	Clone X of SPACG10	This study
SAPACG10_Xa	SPACG10_X/pKD12.138a	This study

Plasmids		
pET-15b	Ap^R, *lacI*, P$_{T7}$, ori (3882) replicon	Invitrogen
pLBA110	*fumA* cloned into pET-15b	This study
pBKA1.56	*fumC* cloned into pET-15b	This study
pBKA1.58	*fumB* cloned into pET-15b	This study
pBKA1.412	*mdh* cloned into pET-15b	This study
pBbA1a-RFP	Ap^R, *lacI*, P$_{Trc}$, p15A ori replicon, *mRFP*	This study
pBKA1.418	*fumA* cloned into pBbA1a-RFP to replace *mRFP*	This study
pBKA1.420	*fumC* cloned into pBbA1a-RFP to replace *mRFP*	This study
pKD12.138a	Ap^R, aroF^{FBR}, tktA, P$_{tac}$ aroE serA , p15 ori	Draths Lab
pBKA1.1201	*maeB* cloned into pBKA1.418	This study
pBKA1.1202	*ldhA* cloned into pBKA1.418	This study
pBKA1.1203	*ldhA* cloned into pBKA1.1201	This study

Table 5.11. Primers used in this study.

Primer ID	SEQUENCE	Purpose
BKA021_MaeB	TATTACGAATTCGGGGCTGTTGCCCACACACTTTATTTGTG	For cloning *maeB* into pBKA1.418
BKA022_MaeB	TTATAACAGATCTCCCTGAATATTCAGGGTAAGCGTGAGAG	
BKA023_Ldh	ATTCTTGGATCCGTAGCTTAAATGTGATTCAACATCAC	For cloning *ldhA* into pBKA1.418 and pBKA1.1201
BKA024_Ldh	TAATATCTCGAGCAGGGGAGCGGCAAGATTAAA	

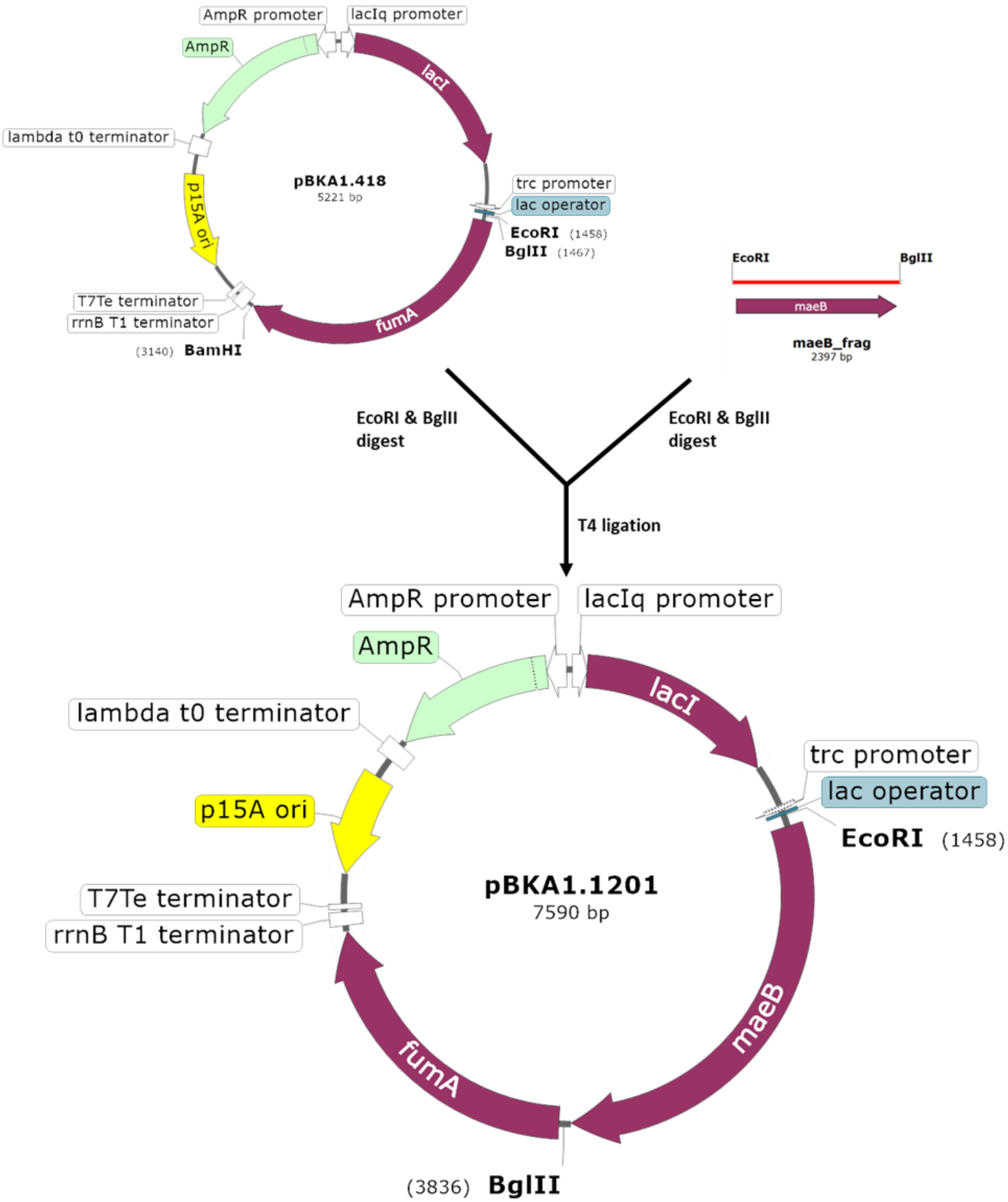


Figure 5.27. Cloning of pBKA1.1201.

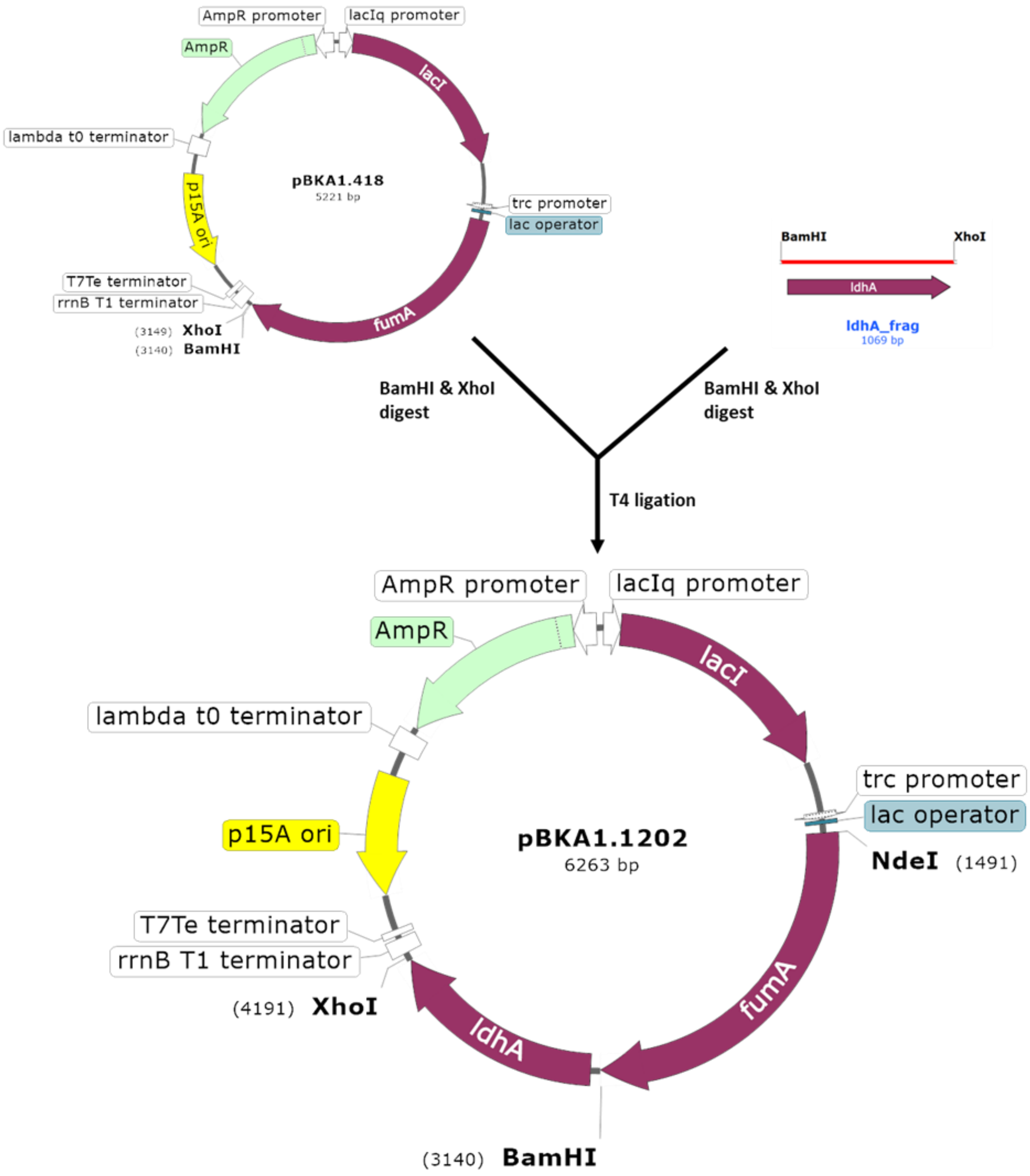


Figure 5.28. Cloning of pBKA1.1202.

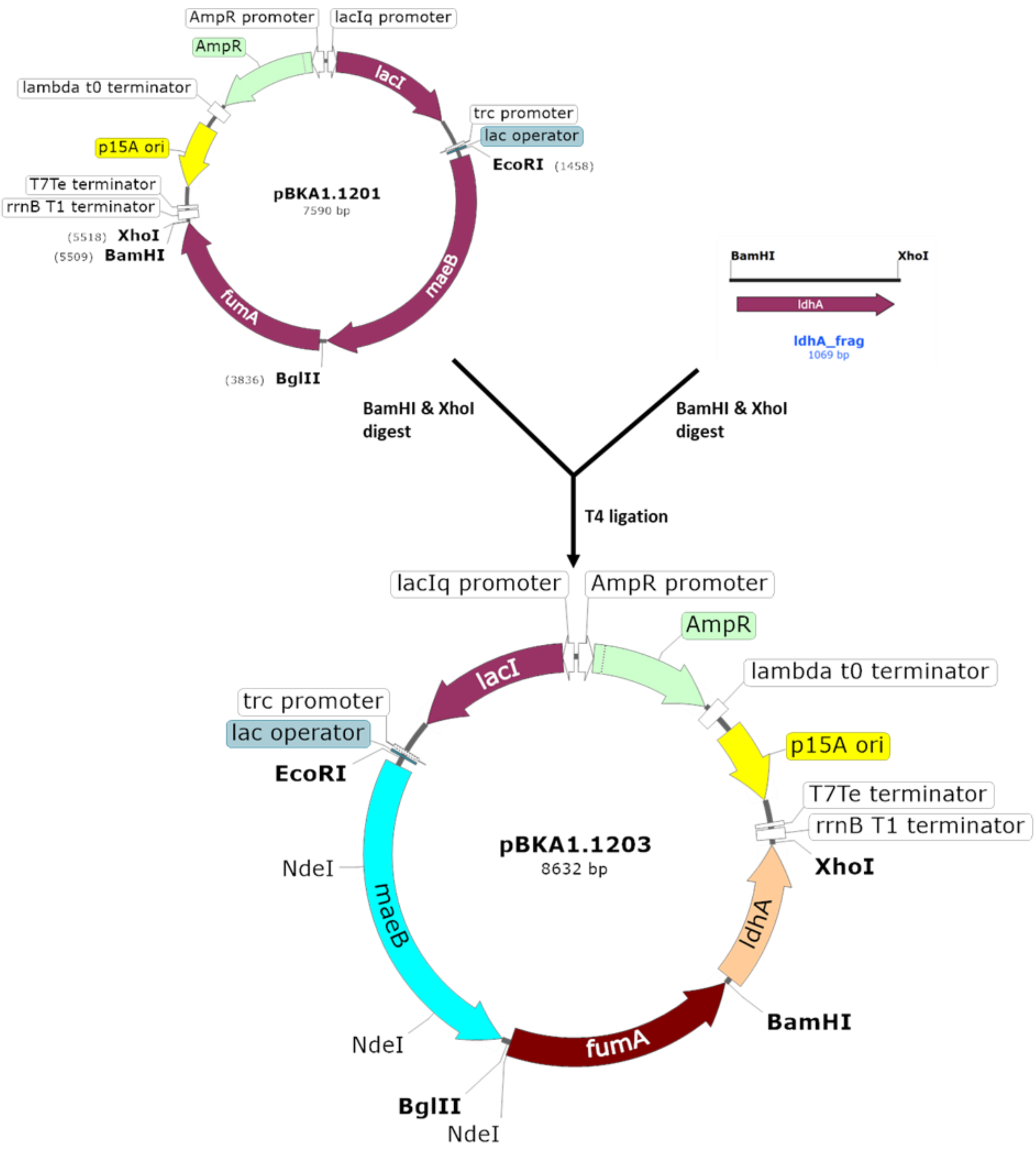


Figure 5.29. Cloning of pBKA1.1203.

5.10.3 Culture conditions and ADCA fermentation

For all fermentations, unless otherwise specified, cells were cultured in M9A containing 4.5 g L^{-1} of ADCA supplemented with 50 µg mL^{-1} (where necessary to maintain plasmids). Cultures were grown in 250 mL baffled shake flasks containing 50 mL of medium at 37 °C with shaking at 200 rpm. Protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM at the start of the culture or after the OD_{600} reached 0.1. Post-induction, cultures were incubated for an additional 24 h. The concentration of products and intermediates were quantified using HPLC equipped with refractive index detection (RID) or variable wavelength detection (VWD). Aliquots of samples were analyzed at 6 – 12 h intervals. About 1 mL of sample was centrifuged and the supernatant filtered through a 0.22 µm filter. The filtrate was analyzed on HPLC using 0.01 N H_2SO_4 mobile phase at flow rate 0.6 mL/min for 18 min. Calibration curves of standard using the same method were used for quantification.

5.10.3 Determination of kinetic parameters

5.10.3.1 Lactate dehydrogenase (LDH)

Lactate dehydrogenase (EC 1.1.1.27) used for *in vitro* production of lactic acid was purchased from Sigma-Aldrich® (CAS Number 9001-60-9). The purchased enzyme was characterized by monitoring conversion of pyruvic acid to lactic acid. The assay was conducted in 100 mM Tris-HCl (pH 8) using a plate reader measuring the consumption of NADH at 340 nm. The reaction was monitored for 5 min and was initiated by the addition of pyruvate to wells containing NADH and 0.05 ng μL^{-1} of LDH. Initial velocities for various substrate concentrations were fitted to the Hill equation to determine kinetic parameters as described earlier. Kinetic parameters were determined for both NADH and pyruvate substrates.

5.10.3.2. Aldehyde dehydrogenase (ALDH)

Aldehyde dehydrogenase (EC 1.2.1.5) was purchased from Sigma-Aldrich® (CAS Number 9028-88-0). The enzyme catalyzes the oxidation of aldehydes to carboxylates and in the process reducing NAD^+ to NADH. We exploited this activity by using ALDH as a cofactor regeneration system. In the vitro lactate production route, ADCA is converted to pyruvate by FumA and MaeBN155D in the presence of Mn^{2+} cofactor. The pyruvate is then converted to lactate by LDH, consuming a molecule of NADH per molecule of lactate produced. To ensure cost effectiveness of the process, a recycling system is coupled to the activity of LDH to regenerate NADH. For this reaction acetaldehyde was used as a sacrificial substrate. The kinetics for the conversion of acetaldehyde to acetic acid was investigated to identify appropriate concentrations and conditions for ALDH. An enzyme assay was developed for this conversion in 100 mM Tris-HCl buffer (pH 8) at 25 °C using the plate reader for a 200 µL reaction. Reaction rates were determined for various concentrations of acetaldehyde and NAD^+. The reactions were initiated by addition of various concentrations of substrates and monitored for 5 min. The rates were measured by monitoring production of NADH at 340 nm. The initial velocities for the various substrate concentrations were fitted to the Hill equation to determine kinetic parameters as described before.

5.10.4 Bioproduction of lactic acid from ADCA

5.10.4.1 *In vitro* approach

The enzymatic production of lactate was carried out in a 2- or 10-mL reaction. An initial reaction was performed varying NADH cofactor equivalence. In this reaction 25 mM ADCA was combined with 100 mM Tris-HCl buffer (pH 8) containing reaction components as shown in Table 5.12 in a 10 mL reaction. The reaction was incubated at 26 °C with gentle agitation at 80 rpm. At

30 min interval, 50 µL aliquots were taken from the reaction mixture for 3 h. The reaction was quenched by the addition of 50 µL of 5 M HCl.

Table 5.12. Components for *in vitro* production of lactate varying NADH cofactor molar equivalence.

Reaction ID	Cofactor (equiv.)	[ADCA] mM	[Acetaldehyde] mM	[NADH] mM	[Mn^{2+}] mM	[FumA] ng/µL	[MaeB] ng/µL	[LDH] ng/µL	[ALDH] ng/µL
1	1	25	0	25	5	0.3	125	0.009	125
2	0.1	25	25	2.5	5	0.3	125	0.009	125
3	0.01	25	25	0.25	5	0.3	125	0.009	125
4	0.001	25	25	0.025	5	0.3	125	0.009	125

The mixture was then diluted to either 1 mL with 100 µL of 10 mM 3-(Trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt (dissolved in D_2O) and 800 µL ddH_2O for ^1H NMR analysis or with 1.90 mL ddH_2O for HPLC analysis. About 700 µL of the resulting solution after spinning down was used for wet1D ^1H NMR.

For HPLC analysis, the samples were centrifuged to remove any precipitates and filtered through a 0.22 µm filter. For both samples and calibration standards, 20 µL were injected into the HPLC with 0.01 N H_2SO_4 mobile phase at 0.6 mL/min for 18 min. ADCA was quantified using the variable wavelength detection (VWD) occurring at about $R_T = 6$ min. Oxaloacetate, pyruvate, and lactate were quantified using refractive index detection (RID) occurring at $R_T = 8$, 10 and 13 min, respectively. A calibration curve was obtained for each compound using same run conditions. The 0.001 NADH equivalence was identified to be the appropriate cofactor concentration. The reaction was then scaled to 100 mM and 500 mM ADCA concentrations while maintaining same enzyme concentrations but using only 0.001 NADH molar equivalence (Table 5.13).

The 500 mM ADCA reaction was performed with two approaches: continuous and stepwise. For the continuous reaction, all components were combined in one pot at once and the reaction was allowed to run for 24 h, monitoring every 30 min for the first 3 h. The stepwise approach involved multiple steps where FumA was added to reaction buffer containing only ADCA, and the reaction allowed to run for an h. Then MaeBN155D and Mn^{2+} were added and then allowed to incubate with gentle shaking as before for another h. All remaining reaction components are then added, and the reaction run for another h. Both reactions were completely halted after 24 h.

Table 5.13. Components for *in vitro* production of lactate varying NADH cofactor molar equivalence.

Reaction ID	Cofactor (equiv.)	[ADCA] mM	[Acetaldehyde] mM	[NADH] mM	[Mn^{2+}] mM	[FumA] ng/µL	[MaeB] ng/µL	[LDH] ng/µL	[ALDH] ng/µL
A	0.001	25	25	0.025	5	0.3	125	0.009	125
B	0.001	100	100	0.100	5	0.3	125	0.009	125
C	0.001	500	500	0.500	5	0.3	125	0.009	125
D	0.001	500	500	0.500	5	0.3	125	0.009	125

Reaction A, B, and C: Continuous
Reaction D: stepwise/batch

5.10.4.2 *In vivo* approach

All fermentations were performed in 250 mL baffled shake flasks in 50 mL media. Seed cultures were produced by inoculating culture tubes containing 5 mL of M9G or LB (with antibiotics where necessary) with single colonies of strains from an overnight agar plate. The seed cultures were used to inoculate fermentation media to an initial OD_{600} of 0.05. The cultures were incubated at 37 °C with agitation (200 rpm). After every 12 h, 500 – 1000 µL aliquots of the cultures were centrifuged and filtered with a 0.22 µm filter to remove cells. The filtrate was analyzed immediately on HPLC as described before or flash frozen with liquid N_2 and stored at -80 °C for later analysis.

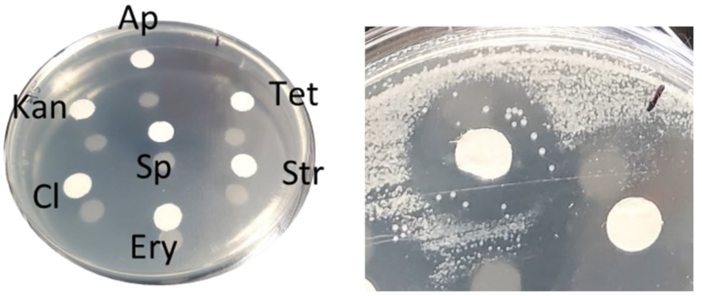
The strains C2, C3, C4, and C5 were cultivated under two conditions: continuous (condition A) or batch (condition B). For condition A, seed cultures were inoculated into 50 mL media containing 11.4 g L^{-1} ADCA and incubated with shaking (200 rpm) for 100 h at 37 °C. For condition B, seed cultures were used to inoculate 50 mL M9A media containing 6 g L^{-1} ADCA and cultivated as in condition A for 60 – 78 h (depending on when cells reached late log phase), then adjusted with 6 mL of 34 g L^{-1} ADCA. Aliquots of samples were taken at 12 h interval for HPLC analysis as described above.

5.10.4.3 Antibiotic susceptibility testing by disk diffusion assay

The Kirby-Bauer disk diffusion susceptibility testing was used to assess the sensitivity of evolved ACG populations (ACG1 – 10 and SPACG1 - 10) to ampicillin, chloramphenicol, erythromycin, kanamycin, spectinomycin, streptomycin, and tetracycline. LB agar plates where used instead of the commonly used Mueller-Hinton agar plate. Using a sterile inoculating loop, four isolated colonies from overnight (12 h incubation) agar plates of each population being tested were used to inoculate a 2 mL of sterile saline. The saline tube was vortexed briefly to create a uniform suspension. The OD_{600} was adjusted to 0.1 by addition of more bacterial cells or dilution with saline. A sterile swab was dipped and rotated against the side of the inoculum tube (above the liquid level) using firm pressure, to remove excess fluid. The dried surface of the LB agar plates were inoculated by streaking the swab three times over the entire agar surface, rotating the plates approximately 60 degrees each time to ensure an even distribution of the inoculum. The plates were rimmed with the swab to pick up any excess liquid. The inoculated plates were allowed to sit at room temperature for at least 5 min for the surface of the agar plate to dry. The appropriate antimicrobial-impregnated disks (see Table 5.14 for amount loaded) were placed on the surface of the agar plate as shown in Figure 5.30, using a forcep. Once all disks were placed, the plates were covered, inverted, and incubated at 37 °C for 18 h (Figure 5.31). After incubation, the zone sizes were measured to the nearest mm using a ruler.

Table 5.14 Amounts of antibiotics loaded for disk diffusion assay and measurements for susceptibility and resistance.

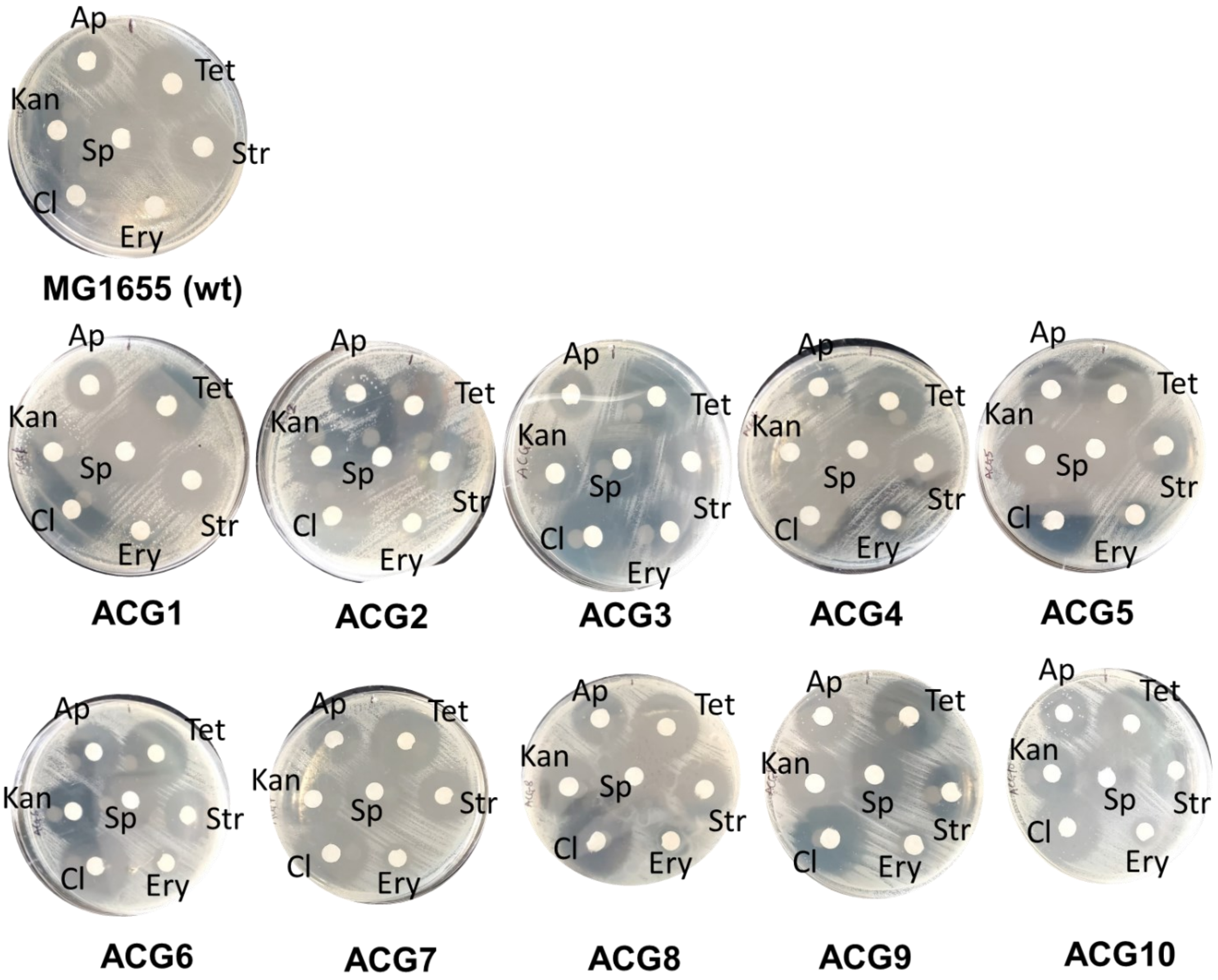
Target		Antibiotic	Amount loaded (μg)	Sensitive ≥ (mm)	Resistance < (mm)
Cell wall synthesis		Ampicillin	10	14	11
Protein synthesis	50S	Chloramphenicol	30	18	12
		Erythromycin	30	23	13
	30S	Streptomycin	10	15	11
		Tetracycline	30	15	11
		Spectinomycin	100	15	11
		Kanamycin	30	18	13



Before incubation

Showing colonies resistant to antibiotics

Figure 5.30. Placement of antibiotic-impregnated disks (left) showing ampicillin (Ap), tetracycline (Tet), streptomycin (Str), erythromycin (Ery), chloramphenicol (Cl), kanamycin (Kan), and spectinomycin (Sp). (Right) Plate after 18 h of incubation showing distinct colonies of resistant bacterial appearing in zones of inhibition. If there is no clear zone, the plate is repeated.



MG1655 (wt)

ACG1

ACG2

ACG3

ACG4

ACG5

ACG6

ACG7

ACG8

ACG9

ACG10

Figure 5.31. Plates after 12 h of incubation showing zones of inhibition for ampicillin (Ap), tetracycline (Tet), streptomycin (Str), erythromycin (Ery), chloramphenicol (Cl), kanamycin (Kan), and spectinomycin (Sp) across ACG populations.

5.10.4.4 Growth fitness of clones from ACG populations

The fitness of single colonies from ACG populations (ACG1 – 10) were tested. After ten rounds of ALE to afford ACG10, dilutions of the cultures were spread on M9A and LB agar plates. Five colonies from each plate were randomly selected and used to inoculate 1 mL M9A and LB media in sterile deep well plate. The plate was incubated at 37 °C overnight with agitation (200 rpm) to produce seed cultures. The seed cultures were used to inoculate 200 μL M9A media and cultivated using the plate reader method as described before.

5.10.4.5 Bioproduction of lactic acid using evolved ACG populations

Three clones from ACG10 populations were transformed with plasmid pBKA1.1203 to obtain ACG10_1/pBKA1.1203 (ACG_LAP1), ACG10_2/pBKA1.1203 (ACG_LAP2), and ACG10_3/pBKA1.1203 (ACG_LAP3). Three colonies from the transformants of each plate were inoculated in 5 mL M9A media and incubated at 37 °C overnight with agitation (200 rpm) to produce seed cultures. The seed cultures were used to inoculate 50 mL M9A media containing 4.5 g L^{-1} of ADCA in a 250 mL baffled shake flask to an initial OD_{600} of 0.05. The inoculum was incubated at 37 °C for 24 h with agitation (200 rpm). Aliquots of cultures were analyzed by HPLC after every 12 h as described before.

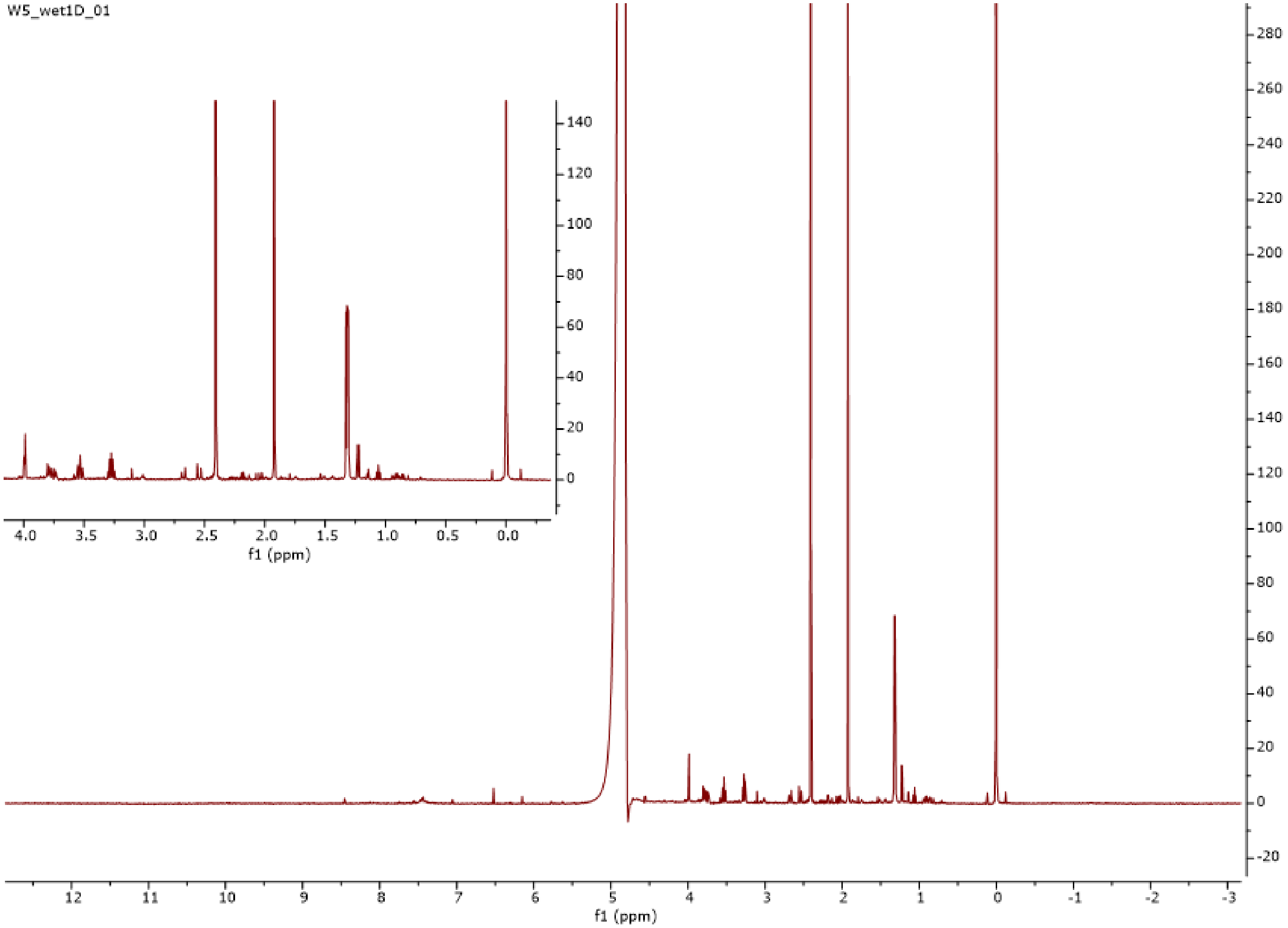


Figure 5.32. ^{1}H NMR for *in vivo* production of lactic acid using strain W5.

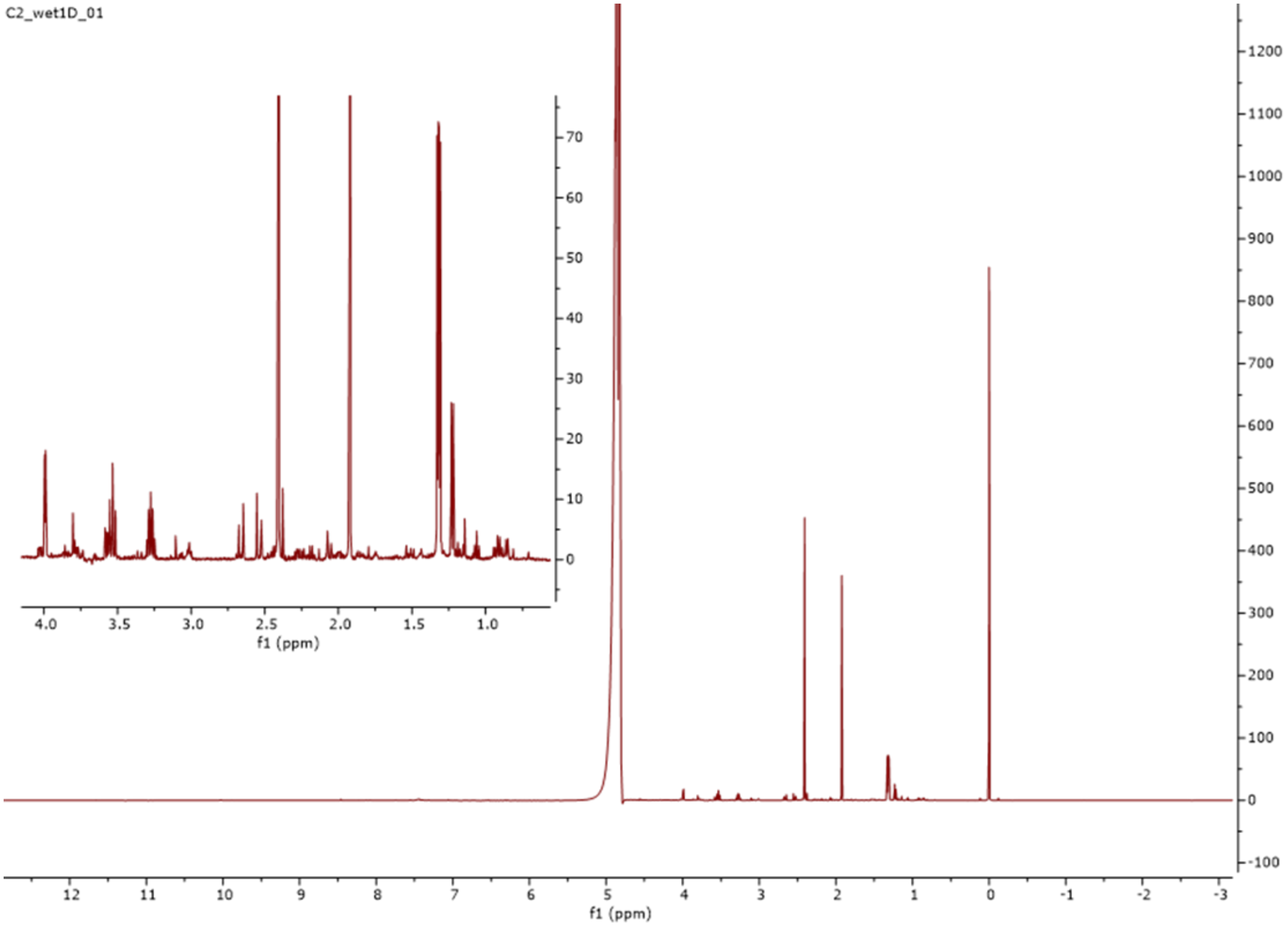


Figure 5.33. ^{1}H NMR for *in vivo* production of lactic acid using strain C5.

5.10.5 Bioconversion of ADCA to shikimic acid

The *E. coli* RB791 *ser*::*aroB* Δ*aroL* Δ*aroK* (SP) strain was obtained from the Draths lab. This strain evolved for ten rounds to obtain SPACG1 – 10. Two colonies of SPACG10 (SPACG10_4 and SPACG10_5) were transformed with plasmid pKD12.138a to obtain SPACG10_4a and SPACG10_5a. The pKD12.138a plasmid encodes genes for AroF^{FBR}, AroE, TktA, and SerA. For *in vivo* production of shikimic acid, all strains were cultivated in 50 mL M9A media containing 4.5 g L^{-1} of ADCA in a 250 mL baffled shake flask to an initial OD$_{600}$ of 0.05. The inoculum was incubated at 37 °C for 24 h with agitation (200 rpm). Aliquots of cultures were analyzed by HPLC after every 12 h as described before. HPLC traces of standards are shown in Figure 5.34

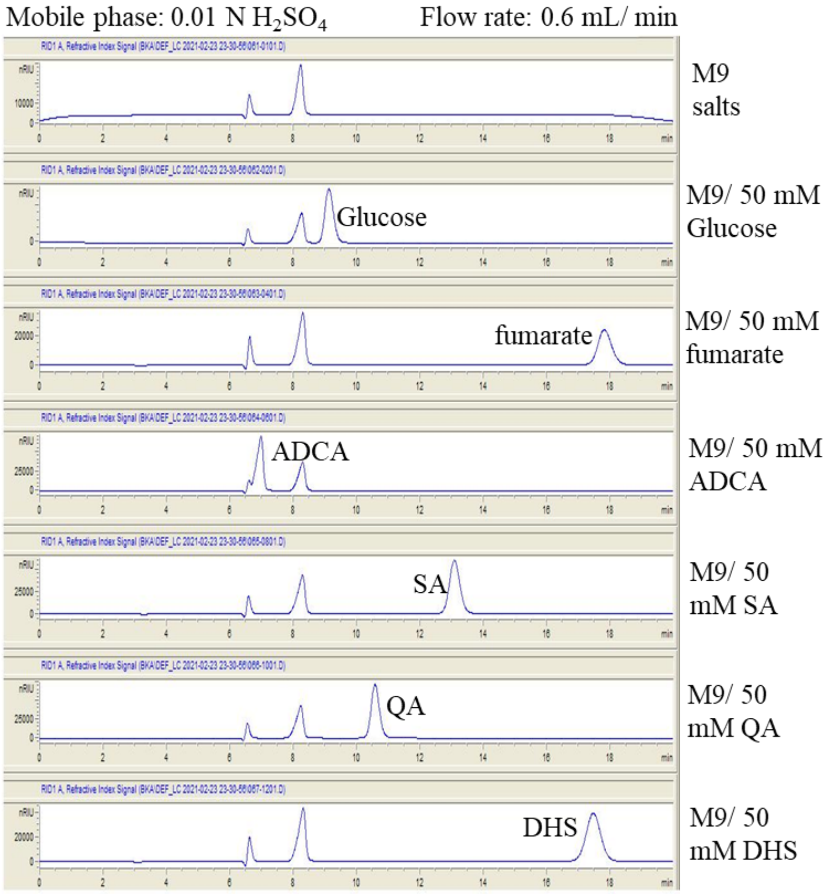


Figure 5.34. HPLC chromatogram of standards for bioconversion of ADCA to shikimate analysis.

REFERENCES

(1) Sambrook, J.; Fritsch, E. F.; Maniatis, T. Molecular Cloning: A Laboratory Manual. In *Molecular cloning: a laboratory manual.*; Cold Spring Harbor Laboratory Press, 1989.

(2) Sprouffske, K.; Wagner, A. Growthcurver: An R Package for Obtaining Interpretable Metrics from Microbial Growth Curves. *BMC Bioinformatics* **2016**, *17*, 172.

(3) Barrick, J. E.; Colburn, G.; Deatherage, D. E.; Traverse, C. C.; Strand, M. D.; Borges, J. J.; Knoester, D. B.; Reba, A.; Meyer, A. G. Identifying Structural Variation in Haploid Microbial Genomes from Short-Read Resequencing Data Using Breseq. *BMC Genomics* **2014**, *15* (1), 1039.

(4) Deatherage, D. E.; Traverse, C. C.; Wolf, L. N.; Barrick, J. E. Detecting Rare Structural Variation in Evolving Microbial Populations from New Sequence Junctions Using Breseq. *Front. Genet.* **2014**, *5*, 468.

(5) Wagih, O.; Galardini, M.; Busby, B. P.; Memon, D.; Typas, A.; Beltrao, P. A Resource of Variant Effect Predictions of Single Nucleotide Variants in Model Organisms. *Mol. Syst. Biol.* **2018**, *14* (12), e8430.

(6) Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H. Construction of Escherichia Coli K-12 in-Frame, Single-Gene Knockout Mutants: The Keio Collection. *Mol. Syst. Biol.* **2006**, *2*, 2006.0008.

(7) Baba, T.; Huan, H.-C.; Datsenko, K.; Wanner, B. L.; Mori, H. The Applications of Systematic In-Frame, Single-Gene Knockout Mutant Collection of Escherichia Coli K-12. *Methods Mol. Biol.* **2008**, *416*, 183–194.

(8) Saragliadis, A.; Trunk, T.; Leo, J. C. Producing Gene Deletions in Escherichia Coli by P1 Transduction with Excisable Antibiotic Resistance Cassettes. *J. Vis. Exp.* **2018**, No. 139. https://doi.org/10.3791/58267.