

FOUNDATIONS FOR GENETIC ENGINEERING IN THE SHIKIMATE
PATHWAY OF *BACILLUS METHANOLICUS* MGA3

By

Megan Gruenberg

A THESIS

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

Chemistry – Master of Science

2020

ABSTRACT

FOUNDATIONS FOR GENETIC ENGINEERING IN THE SHIKIMATE PATHWAY OF *BACILLUS METHANOLICUS* MGA3

By

Megan Gruenberg

Numerous widely-used chemicals, such as adipic acid, are produced from hydrocarbons that come from petroleum. Consumption of petroleum contributes to global climate change and ocean acidification, which has created a great need for the development of petroleum independent chemical synthesis methods. Recently, considerable interest has turned toward using methanol as a feedstock for chemical synthesis because of its abundance, ease of purification, and lack of ties to the food industry. Adipic acid precursor *cis,cis*-muconic acid has been biosynthesized in a variety of production hosts, but it has not been biosynthesized using a methanol feedstock. In this work, groundwork was laid for genetically engineering the thermophilic, methylotrophic bacterium *Bacillus methanolicus* MGA3 for *cis,cis*-muconic acid production from methanol via the shikimate pathway. The first enzyme in the shikimate pathway is 3-deoxy-D-arabinoheptulosonate-7-phosphate (DAHP) synthase. This enzyme has been shown to play an important role in the regulation of carbon flow into the pathway. *B. methanolicus* MGA3 DAHP synthase isozymes AroG1 and AroG2 were characterized. Both isozymes were found to be Type I β DAHP synthases, with AroG1 being unregulated at the protein level and AroG2 being allosterically regulated by the aromatic amino acid precursors chorismate and prephenate. Additionally, methods for storing and screening *B. methanolicus* shikimate pathway mutants were developed. The established storage method allows *B. methanolicus* to be mutagenized in large batches and subsequently stored as spores. Agar plate growth medium recipes were developed for germinating *B. methanolicus* spores and phenotypic screening of mutants.

ACKNOWLEDGMENTS

I am blessed with incredibly loving, supportive (and large) family. They offered a great deal of support and encouragement, especially my parents, Jacob, Grace, and Tony. They believed in me even when I didn't believe in myself and helped me through my most difficult moments in graduate school. I would also like to thank all of my friends and colleagues that took the time to read drafts and offer advice; namely: Amaya, Katie, Yasheen, Hadi, Jorge, and Kio. I appreciate them for catching my mistakes and offering helpful suggestions. I thank them for their friendship and words of encouragement. They all made coming to the lab each day worth it. I also owe thanks to my advisor, Dr. Karen Draths, who accepted me into her group even with my initial lack of knowledge in this area of study. Additionally, Dr. Marta Irla provided insightful advice on experiments and checked in with me often, despite being halfway across the world. I have learned so much both about science and life working with this group of dedicated researchers. I am so thankful for these people, and I wish them all happiness and success in the years to come.

TABLE OF CONTENTS

LIST OF TABLES	v
LIST OF FIGURES	vi
KEY TO ABBREVIATIONS	viii
CHAPTER 1 – Introduction	1
BIOECONOMY	1
ADIPIIC ACID	3
THE SHIKIMATE PATHWAY	6
NATIVE METHYLOTROPHY	10
SYNTHETIC METHYLOTROPHY	13
PRODUCTION OF <i>CIS,CIS</i> -MUCONIC ACID BY <i>B. METHANOLICUS</i>	14
REFERENCES	16
CHAPTER 2 – The <i>Bacillus methanolicus</i> MGA3 DAHP Synthases	23
OVERVIEW	23
MULTIPLE SEQUENCE ALIGNMENT	28
PLASMID CONSTRUCTION	37
OVEREXPRESSION AND PURIFICATION	41
MOLECULAR WEIGHT	43
ENZYMATIC ASSAYS	46
DISCUSSION	50
REFERENCES	59
CHAPTER 3 – <i>Bacillus methanolicus</i> MGA3 Culturing	64
OVERVIEW	64
CULTURING	66
SPORULATION	70
PLATING AND SELECTIVE PLATING	72
CONCLUSION	75
REFERENCES	76
CHAPTER 4 – Experimental	80
MATERIALS	80
INSTRUMENTAL MEASUREMENTS	80
BACTERIAL STRAINS AND PLASMIDS	81
CULTURE MEDIA	83
GENERAL DNA MANIPULATION TECHNIQUES	84
PLASMID CONSTRUCTION	85
TRANSFORMATIONS	86
PROTEIN EXPRESSION AND PURIFICATION	87
ENZYME ASSAYS	90
E4P SYNTHESIS AND QUANTIFICATION	92
REFERENCES	93

LIST OF TABLES

Table 1. DAHP synthase classification systems.	23
Table 2. Representative DAHP synthases and their allosteric inhibitors.	28
Table 3. The MSA results of AroG1 and AroG2 with representative DAHP synthases.	30
Table 4. The predicted and calculated masses of AroG1 and AroG2 subunits.	43
Table 5. Kinetic values for AroG1 and AroG2.	48
Table 6. All known bacteria with multiple experimentally studied DAHP synthases and the quantity of each type the organism possesses.	52
Table 7. All known bacteria containing experimentally studied Type I β DAHP synthases and the quantity of each Type I β the organism possesses.	53
Table 8. The top 10 closest relatives to <i>B. methanolicus</i> MGA3 based on 16s rRNA sequence identity and their putative DAHP synthases.	54
Table 9. Michaelis-Menten constants and turnover numbers for DAHP synthases from thermophilic bacteria.	57
Table 10. Select <i>B. methanolicus</i> production strains.	65
Table 11. Growth rates of MGA3 cultured in varied conditions.	70
Table 12. Tested plate recipes for <i>B. methanolicus</i> shikimate pathway mutant screening.	73
Table 13. Bacterial strains and plasmids used in this work.	82
Table 14. Primers for amplification of <i>aroG1</i> , <i>aroG2</i> , and <i>tktA</i> for Gibson Assembly.	86

LIST OF FIGURES

Figure 1. Industrial synthesis of adipic acid.	3
Figure 2. Select examples of indirect biobased synthesis of adipic acid.	5
Figure 3. Direct biosynthesis of adipic acid.	6
Figure 4. The shikimate pathway.	8
Figure 5. Production of ccMA by manipulations to the shikimate pathway.	9
Figure 6. The ribulose monophosphate pathway for methanol assimilation.	11
Figure 7. The three RuMP pathway proteins necessary for introducing synthetic methylotrophy into native nonmethylotrophs.	14
Figure 8. DAHP and KDOP synthase classification tree.	24
Figure 9. Renderings of the Type I DAHP synthases.	26
Figure 10. The condensation reactions performed by DAHP synthase and KDOP synthase.	27
Figure 11. Multiple sequence alignment of MGA3 AroG1 with unregulated Type I β DAHP synthases from <i>P. furiosus</i> and <i>A. pernix</i> .	31
Figure 12. Multiple sequence alignment of MGA3 AroG1 with regulated Type I β DAHP synthases from <i>T. maritima</i> (regulated by phenylalanine and tyrosine) and <i>B. subtilis</i> (regulated by chorismate and prephenate).	32
Figure 13. Multiple sequence alignment of MGA3 AroG1 with Type I α DAHP synthases from <i>C. glutamicum</i> (regulated by phenylalanine and tyrosine) and <i>E. coli</i> (regulated by tyrosine).	33
Figure 14. Multiple sequence alignment of MGA3 AroG2 with unregulated Type I β DAHP synthases from <i>P. furiosus</i> and <i>A. pernix</i> .	34
Figure 15. Multiple sequence alignment of MGA3 AroG2 with regulated Type I β DAHP synthases from <i>T. maritima</i> (regulated by phenylalanine and tyrosine) and <i>B. subtilis</i> (regulated by chorismate and prephenate)	35
Figure 16. Multiple sequence alignment of MGA3 AroG2 with Type I α DAHP synthases from <i>C. glutamicum</i> (regulated by phenylalanine and tyrosine) and <i>E. coli</i> (regulated by tyrosine).	36
Figure 17. Comparison of the traditional and the modified TEV protease recognition sites.	38

Figure 18. Representation of the <i>E. coli</i> codon-optimized DAHP synthase gene (DAHPS) fused to a modified TEV recognition site (ENLYFQM) and a His ₆ tag under the control of a T7 promoter.	38
Figure 19. Construction of pMG2.109.	39
Figure 20. SDS-PAGE analysis of AroG1 and AroG2 purification.	42
Figure 21. Results from mass spectral analysis of AroG1 and AroG2.	44
Figure 22. Results from mass deconvolution of AroG1 and AroG2 mass spectra.	45
Figure 23. The effect of pH on AroG1 and AroG2 activity.	46
Figure 24. The effect of temperature on AroG1 and AroG2 activity.	47
Figure 25. The thermostability of AroG1 and AroG2.	47
Figure 26. The Michaelis-Menten plots for AroG1 and AroG2.	49
Figure 27. The effect of potential inhibitors on AroG1 and AroG2 activity.	50
Figure 28. Model structures of AroG1 and AroG2.	55
Figure 29. The most common alkylation products of thymine and guanine that result from MNNG or EMS treatment.	66
Figure 30. Low levels of Hps cause formaldehyde accumulation which interferes with native biological processes and hinders the growth of <i>B. methanolicus</i> .	67
Figure 31. <i>B. methanolicus</i> MGA3 growth curves in a variety of conditions.	69
Figure 32. Images of <i>B. methanolicus</i> MGA3 cultures after exposure to sporulation conditions.	71
Figure 33. Method for replica plating <i>B. methanolicus</i> MGA3.	74
Figure 34. The TktA assay.	90

KEY TO ABBREVIATIONS

A5P	arabinose-5-phosphate
Amp ^R	ampicillin resistance
aros	aromatic amino acids
ATP	adenosine triphosphate
BTP	bis(tris)propane
C ₁	one carbon
CA	chorismate
ccMA	<i>cis</i> , <i>cis</i> -muconic acid
CM	chorismate mutase
Cm ^R	chloramphenicol resistance
CRISPR	clustered regularly interspaced palindromic repeats
DAHP	3-deoxy-D-arabinoheptulosonate-7-phosphate
DHAP	dihydroxyacetone
DHQ	3-dehydroquininate
DHS	3-dehydroshikimate
E4P	D-erythrose-4-phosphate
EMS	ethyl methanesulfonate
EPA	Environmental Protection Agency
EPSP	5-enolpyruvylshikimate-3-phosphate
ESI-MS	electrospray ionization mass spectrometry
F1,6P	fructose-1,6-bisphosphate
F6P	fructose-6-phosphate
FA	formaldehyde
Fba	fructosebisphosphate aldolase

FPLC	fast protein liquid chromatography
G3P	glyceraldehyde-3-phosphate
G6P	glucose-6-phosphate
G6PD	glucose-6-phosphate dehydrogenase
GHG	greenhouse gas
Glp _x	sedoheptulose-1.7-bisphosphatase
H6P	hexulose-6-phosphate
His ₆	hexa-histidine tag
His ₇	hepta-histidine tag
Hps	hexulose phosphate synthase
HR	homologous region
IPTG	β-D-isopropylthiogalactopyranoside
k_{cat}	turnover number
KDOP	3-deoxy-D-mannooctulosonate-8-phosphate
K_m	Michaelis-Menten constant
lacI	lactose operator repressor
Mann	mannitol
Mdh	methanol dehydrogenase
MeOH	methanol
MNNG	N-methyl-N'-nitro-N-nitrosoguanidine
MSA	multiple sequence alignment
MSU	Michigan State University
MVcM	minimal vitamin containing media
MVcMY	minimal vitamin containing media with yeast extract
MW	molecular weight
NADH	nicotinamide adenine dinucleotide

NADPH	nicotinamide adenine dinucleotide phosphate
NEB	New England Biolabds
OD	optical density
ori	origin of replication
PA	prephenate
PCA	protocatechuate
PCR	polymerase chain reaction
PEP	phosphoenolpyruvate
Pfk	phosphofructo kinase
PGI	phosphoglucose isomerase
Phe	phenylalanine
Phi	phosphohexulose isomerase
psi	pounds per square inch
R5P	ribose-5-phosphate
RAD	reverse adipate degradation
Rpe	ribulosephosphate-3-epimerase
Rpi	ribose-5-phosphate isomerase
RTSF	Research Technology Support Facility
Ru5P	ribulose-5-phosphate
RuBP	ribulose biphosphate pathway
RuMP	ribulose monophosphate pathway
S1,7P	sedoheptulose-1,7-bisphosphate
S3P	shikimate-3-phosphate
S7P	sedoheptulose-7-phosphate
SA	shikimate
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis

SOB	super optimal broth
TEV	tobacco etch virus
Tkt	transketolase
TktA	<i>E. coli</i> transketolase A
TPP	tyrosine pyrophosphate
Trp	tryptophan
Tyr	tyrosine
USD	United States Dollar
V_f	final volume
V_i	initial volume
V_{max}	maximum velocity
X5P	xylose-5-phosphate
YE	yeast extract
YEMAA	yeast extract mimic amino acids
β -HPA	β -hydroxypyruvic acid
ϵ	extinction coefficient

CHAPTER 1 – Introduction

BIOECONOMY

A prominent issue affecting the well-being of the planet and the human population is the dependence on petroleum, also known as crude oil. Petroleum serves industry in transportation fuels, consumer products, and petrochemicals by providing the energy and carbon resources necessary to sustain many of society's activities.^{1,2} However, the wide use of petroleum comes with some serious concerns. Consumption of this non-renewable resource is in large part responsible for the greenhouse gas (GHG) emissions causing global climate change and ocean acidification.^{3,4} The refining, storage, and transportation of petroleum is the second largest stationary source of GHG emissions in the United States, accounting for 30% of total GHG pollution in 2015.⁵ In addition to atmospheric pollution, petroleum is also responsible for contamination of soil and water.² It is estimated that 25% of underground petroleum storage tanks leak their contents into the ground.⁶ This leaked petroleum can find its way to water sources if it is not cleaned quickly and effectively.²

For these reasons, using alternative carbon sources for manufacturing traditional petrochemicals and other value-added chemicals is of paramount importance. Currently, there is widespread support for creating a bioeconomy sector that involves replacing crude oil with renewable carbon sources and replacing harmful chemical processes with microbial catalysts.⁷ Developing such biobased alternatives could potentially reduce the dependence on petroleum.⁸ The production of ethanol through biological processes (i.e. bioethanol) as a gasoline additive is a prime example of steps being taken toward establishing a bioeconomy. First-generation bioethanol is produced from crops such as sugarcane, wheat, and corn.⁹ The concern regarding use of food sources for chemical synthesis lead to the development of second-generation bioethanol (cellulosic ethanol) which is produced from non-food sources such as the inedible portion of plants.⁹ In 2007, the 110th Congress of the United States passed the Energy

Independence and Security Act which mandated the production of 16 billion liters of cellulosic ethanol by 2016; however, the Environmental Protection Agency (EPA) reduced the mandate to just 0.87 billion liters citing an inadequate feedstock supply.¹⁰ Since the original Congressional mandate was not met, first-generation bioethanol was thus used to meet the demand. In 2016, first-generation bioethanol produced from corn accounted for 95% of total bioethanol production at 55 billion liters.¹¹ The large production of first-generation bioethanol raised the market prices of their food-derived feedstocks.¹² Unfortunately, while crops are renewable, their application for chemical production is unsustainable because crop production levels often do not match the feedstock volume required for large-scale chemical synthesis. Consequently, the association of the agriculture and chemical sectors drives unproductive competition among the industries' goals.

Another approach to a sustainable bioeconomy is to employ non-food carbon sources that are more abundant. One-carbon (C_1) molecules have been identified as promising candidates because of their natural abundance and low production costs.¹³ Methanol is an excellent C_1 feedstock due to its availability, affordability, ease of purification, and lack of ties to the food industry.^{14,15} Methanol prices steadily decreased during 2019, and are currently the lowest they have been in the last four years.¹⁶ With the construction of mega-methanol production plants, this price is expected to drop below the price of sugar feedstocks to about \$100 per ton (USD).¹⁷ Methanol can be produced by steam reformation from methane which is a potent GHG.¹⁸ Channeling methane into methanol production for use in the bioeconomy would help reduce its release into the atmosphere and provide an alternative carbon resource other than those derived from petroleum or agricultural crops.

In addition to replacing petroleum, a bioeconomy also aims to increase the use of biological methods in chemical synthesis. Microbes engineered to produce commodity chemicals offer a valuable alternative to traditional chemical synthesis because they reduce the need for organic solvents and dangerous starting materials.¹⁹ These microbial-based production methods also reduce the formation of unwanted byproducts and hazardous waste.¹⁹ Genetically

engineering methylotrophic microbes could provide a green process of producing value-added chemicals from methanol.

ADIPIIC ACID

Adipic acid is an important dicarboxylic acid in high demand.⁸ The synthesis of adipic acid relies on a petroleum-derived benzene and harsh chemical processes. Adipic acid synthesis (Figure 1) requires high temperatures, pressures, and generates the potent GHG nitrous oxide,²⁰ thus making it a process which could greatly benefit from bio-based alternatives. An estimated 3 million tons of adipic acid are synthesized each year from petroleum-derived benzene.²¹ Of this, about 80% is used in the production of nylon 6,6.⁸ In addition, demand for petroleum-derived benzene has been on the rise.²² Benzene production cannot meet the demand which inflates the price of both benzene and products derived from benzene, such as adipic acid.²³

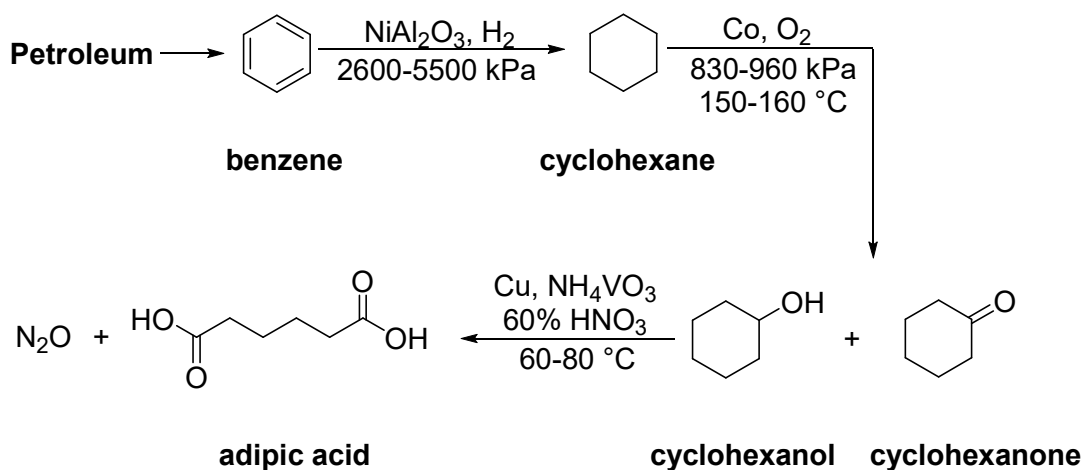


Figure 1. Industrial synthesis of adipic acid.²⁰

Biosynthetic strategies for producing adipic acid have gained much attention because they potentially decrease use of hazardous chemicals and dependence on petroleum.²¹ Adipic acid can be biosynthesized either directly or indirectly.⁸ The indirect routes are the most explored because, until recently, a direct route to adipic acid was unknown. An in-depth description and analysis of these pathways are beyond the scope of this discussion, and detailed reviews on this topic are described elsewhere.^{8,24} The pathways are reviewed briefly as follows.

The indirect production of bio-adipic acid from biosynthesized *cis,cis*-muconic acid (ccMA) or D-glucaric acid have been reported.^{8,24} Adipic acid can be produced from ccMA by hydrogenation and from D-glucaric acid by reduction (Figure 2).²⁴ The highest titer (36.8 g L⁻¹ ccMA) of an adipic acid precursor is reported using an engineered strain of *E. coli* that contains a modified shikimate pathway which redirects carbon flow from aromatic amino acid biosynthesis toward ccMA production.²⁵

There are three known direct pathways to adipic acid (Figure 3). One of which is the reverse adipate degradation (RAD) pathway.⁸ The RAD pathway was recently discovered in *Thermobifida fusca*, which has been shown to directly produce 2.23 g L⁻¹ of adipic acid.²⁹ By incorporating the *T. fusca* RAD genes into a strain of *E. coli* engineered to utilize glycerol as a sole carbon source, a titer of 68 g L⁻¹ of adipic acid has been achieved.³⁰ Additionally, a strain of yeast isolated from petroleum-contaminated soil has been found to produce a variety of dicarboxylic acids through a combination of the β - and ω -oxidation pathways.³¹ By engineering this strain to selectively funnel six-carbon fatty acids into the β -oxidation pathway, it has produced 50 g L⁻¹ of adipic acid using coconut oil as a substrate.

Incorporation of the enzyme enoate reductase has generated another direct biosynthetic route to adipic acid.³² This enzyme can catalyze the conversion of ccMA to adipic acid. Unfortunately, the enzyme turnover is slow and requires optimization to be implemented industrially as a viable source of adipic acid.³³ In proof of concept studies, enoate reductase has been introduced into engineered microorganisms that produce ccMA. In one case, a coculturing

system was created in which one strain engineered of *E. coli* produced ccMA, while another strain overexpressed the enoate reductase. This coculturing system was reported to generate 27.6 mg L⁻¹ of adipic acid.³³ Enoate reductase was also introduced into a ccMA producing strain of *Saccharomyces cerevisiae* which produced 2.59 mg L⁻¹ of adipic acid.³⁴ Notably, the *S. cerevisiae* strain also generated >284 mg L⁻¹ of ccMA. Although the strategy of introducing enoate reductase into ccMA-producing organisms shows potential toward developing a new direct pathway to adipic acid, their mg L⁻¹ scale production is not yet competitive with the g L⁻¹ scale of the RAD or β - and ω -oxidation pathways for bio-adipic acid production.

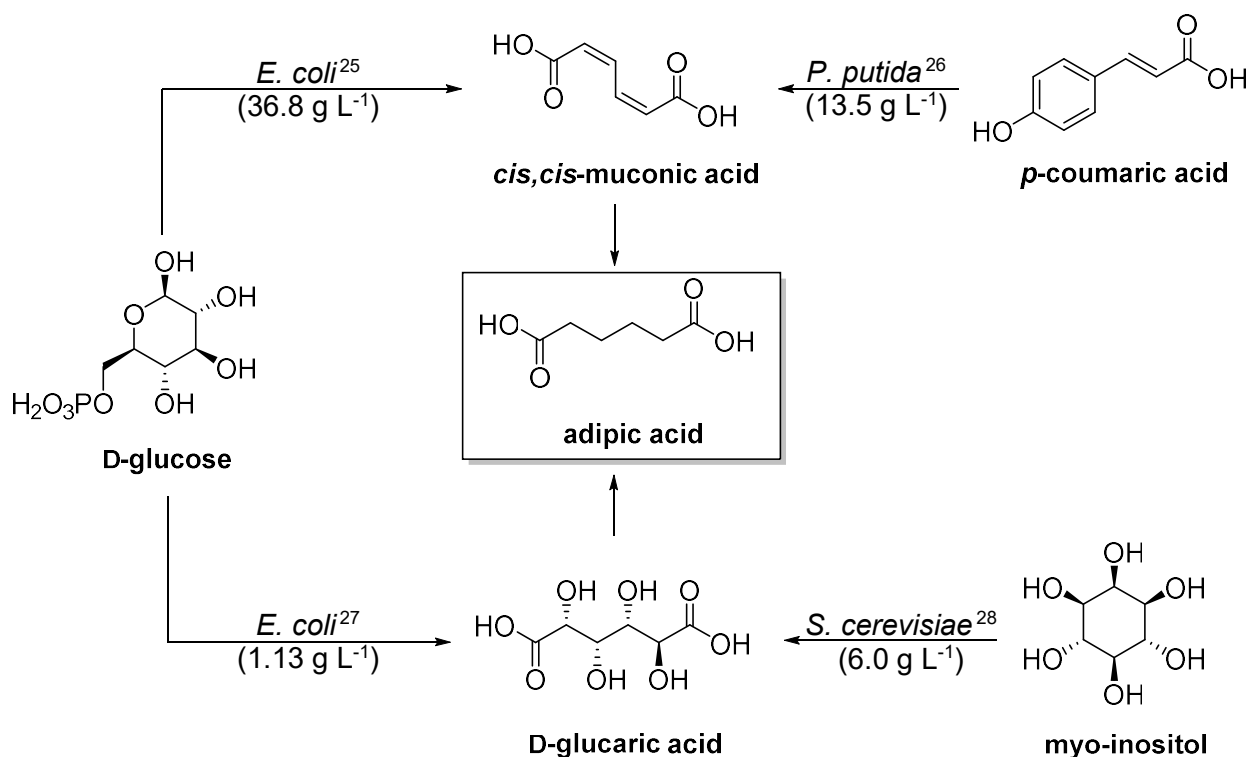


Figure 2. Select examples of indirect biobased synthesis of adipic acid. Titrers of ccMA or D-glucaric acid are shown in parentheses below the arrow.

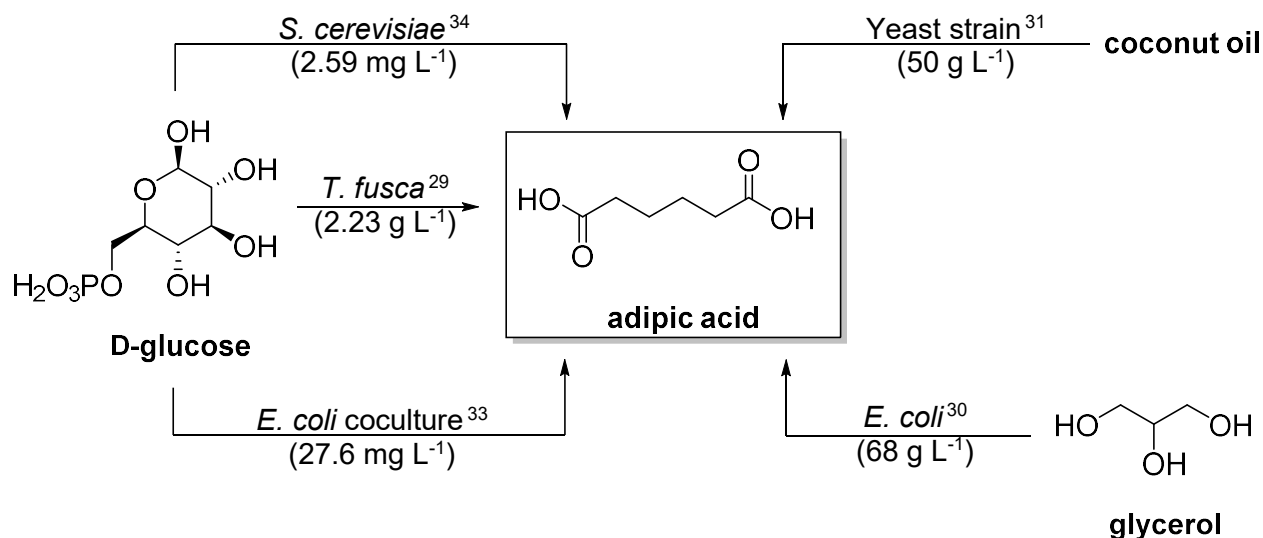


Figure 3. Direct biosynthesis of adipic acid. Titer of adipic acid is shown in parentheses below the arrow. The *T. fusca* and *E. coli* strains produce adipic acid via the RAD pathway. The yeast strain produces adipic acid via the β -oxidation pathway. The *S. cerevisiae* strain and the *E. coli* coculture system produce adipic acid by a modified shikimate pathway.

THE SHIKIMATE PATHWAY

In bacteria and plants, the shikimate pathway (Figure 4) is responsible for producing the metabolite chorismate, which serves as a branch point on the biosynthetic pathway to the aromatic amino acids (L-phenylalanine, L-tyrosine, and L-tryptophan) and aromatic vitamins (*p*-hydroxybenzoic acid, *p*-aminobenzoic acid, and 2,3-dihydroxybenzoic acid).³⁵ The first step in the pathway is the irreversible condensation of D-erythrose-4-phosphate (E4P) and phosphoenolpyruvate (PEP) to form 3-deoxy-D-arabinoheptulosonate-7-phosphate (DAHP) and inorganic phosphate. Six subsequent steps then convert DAHP to chorismate. The condensation of E4P and PEP is catalyzed by DAHP synthase. This enzyme plays an important role in regulating the influx of carbon into the shikimate pathway.³⁶ Being the first committed step of aromatic amino acid biosynthesis, DAHP synthases have evolved allosteric regulation to downstream metabolites.³⁷⁻⁴¹ Various regulatory features separate members of the DAHP synthase family into distinct classes.³⁶

Manipulation of the shikimate pathway is a productive method for engineering a ccMA-producing microorganism. This is accomplished by knocking out the gene encoding for shikimate dehydrogenase (*aroE*) which causes an accumulation of 3-dehydroshikimic acid (DHS). Expression of DHS dehydratase, protocatechuate decarboxylase, and catechol-1,2-dioxygenase in the *aroE* mutant then converts DHS to ccMA (Figure 5).

Although the ccMA-producing organisms described generate high titers, they depend either on glucose, a glucose derivative (*myo*-inositol), or *p*-coumaric acid as a feedstock. The use of glucose and glucose derivatives like *myo*-inositol is not a sustainable source of carbon for chemical synthesis as mentioned earlier. The compound *p*-coumaric acid is often used as a model monomer of lignin.²⁶ Lignocellulosic biomass is a potential source of renewable carbon which comes from the inedible structural support tissues of plants.⁴³ These support tissues are comprised mainly of three natural polymers: cellulose, hemicelluloses, and lignin.⁴² The glucose monomers of cellulose are sought-after as a non-food derived feedstock for renewable microbial chemical production, especially in regard to bioethanol production.⁴² One of the major problems with using lignocellulosic biomass for this purpose is the mixture of compounds released during depolymerization.⁴² The organisms employed for chemical production industrially (e.g. *E. coli* and *S. cerevisiae*) can utilize the glucose monomers of cellulose, but not the monomers obtained from lignin, thus the attempt at developing an *E. coli* strain which consumes *p*-coumaric acid.^{43,26} Lignin is an intricate polymer whose composition varies greatly, which poses many problems in its use as a carbon source.⁴⁴ While lignocellulose is not edible, it requires valuable farming land for its production, which would continue to perpetuate the competition between the agricultural and chemical industries. It is necessary to develop a production host that can biosynthesize ccMA from a sustainable, easily accessible source of carbon like methanol.

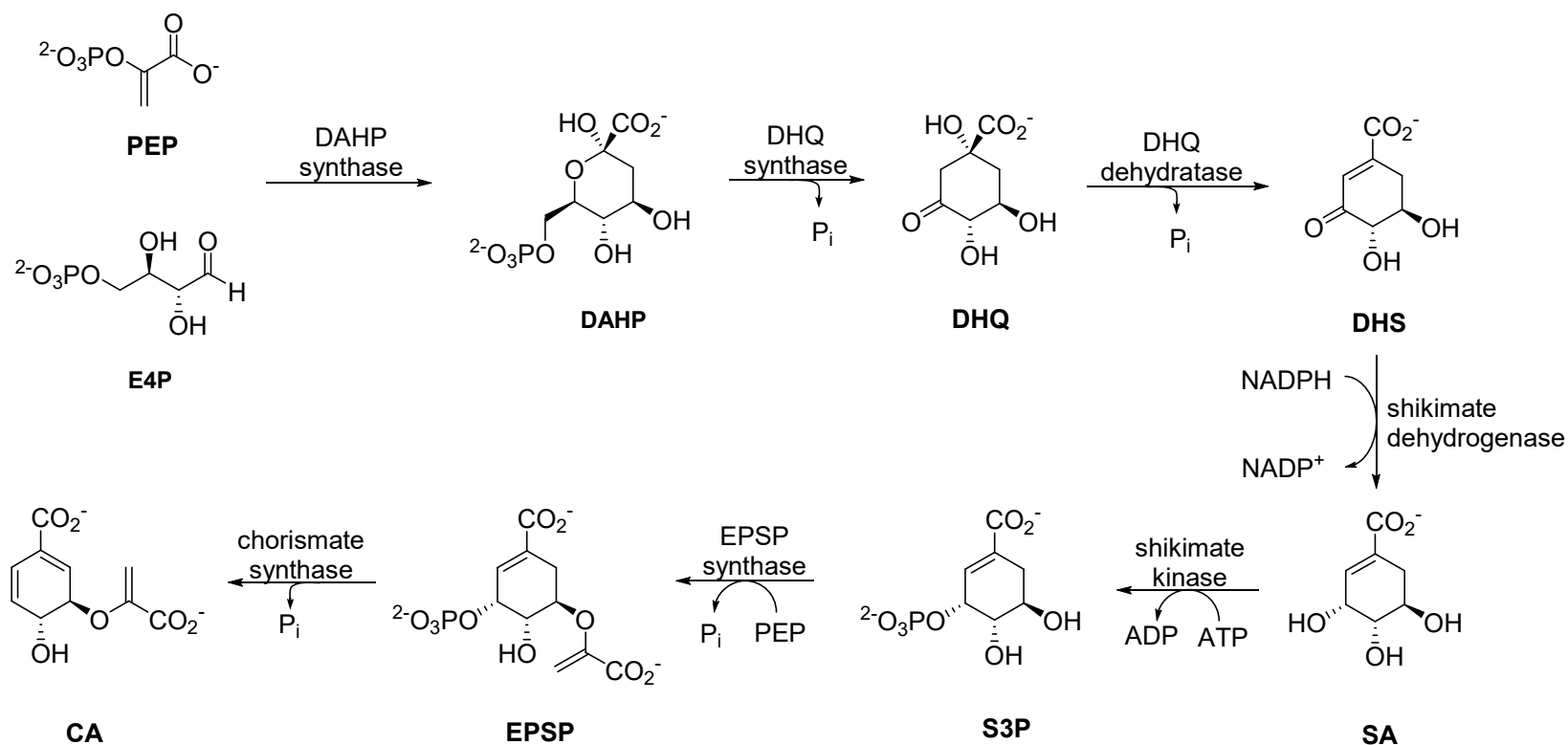


Figure 4. The shikimate pathway. Chorismic acid serves as a branching point for the production of the aromatic amino acids (L-phenylalanine, L-tyrosine, and L-tryptophan) and the aromatic vitamins (p-hydroxybenzoic acid, p-aminobenzoic acid, and 2,3-dihydroxybenzoic acid). Abbreviations: phosphoenolpyruvate (PEP), D-erythrose-4-phosphate (E4P), 3-deoxy-D-arabinoheptulosonate-7-phosphate (DAHP), 3-dehydroquinate (DHQ), 3-dehydroshikimate (DHS), shikimate (SA), shikimate-3-phosphate (S3P), 5-enopyruvylshikimate-3-phosphate (EPSP), chorismate (CA).

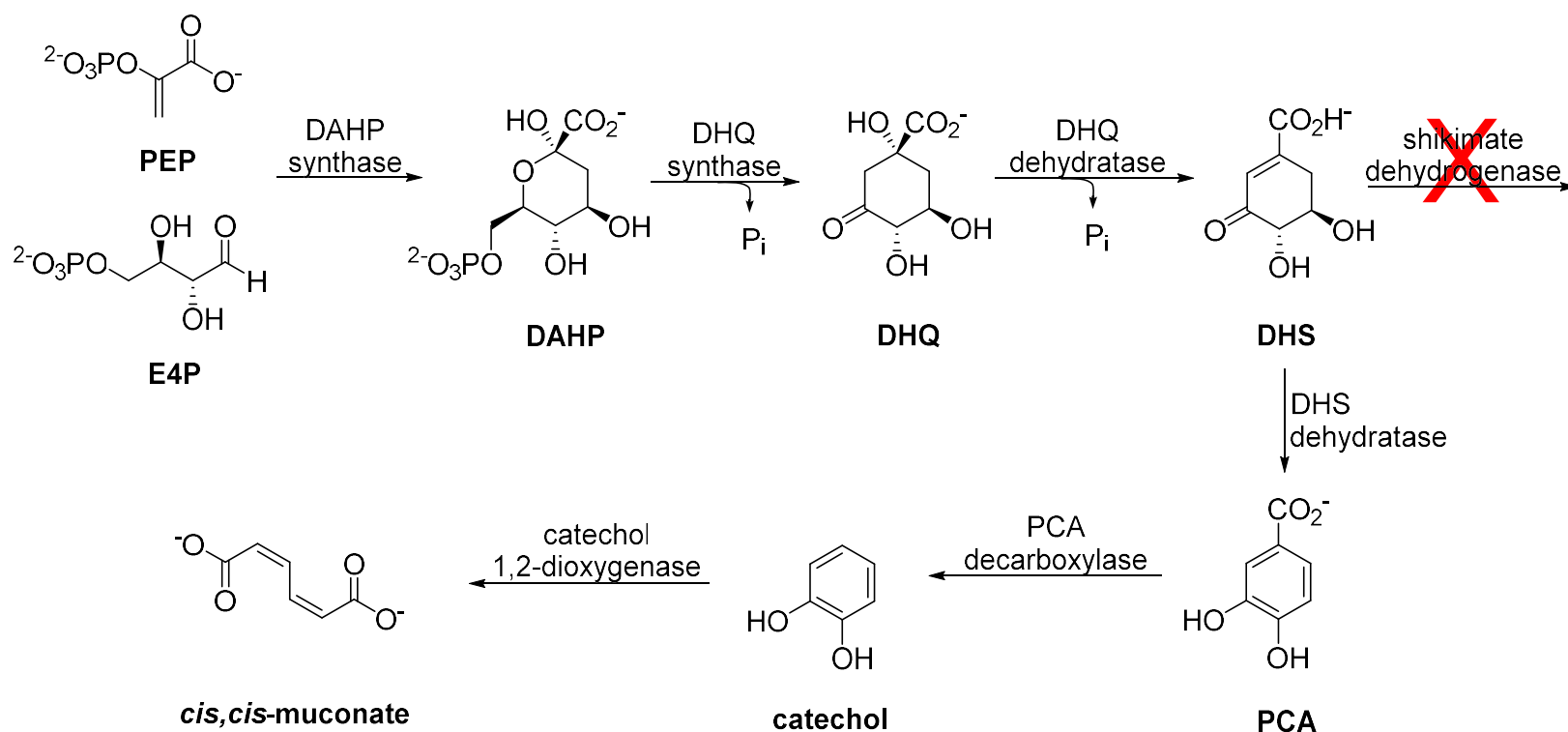


Figure 5. Production of ccMA by manipulations to the shikimate pathway. Knock-out of shikimate dehydrogenase results in the accumulation of DHS. Introduction of three exogenous genes (DHS dehydratase, PCA decarboxylase, and catechol 1,2-dioxygenase) convert DHS to ccMA. Abbreviations: phosphoenolpyruvate (PEP), D-erythrose-4-phosphate (E4P), 3-deoxy-D-arabinoheptulosonate-7-phosphate (DAHP), 3-dehydroquininate (DHQ), 3-dehydroshikimate (DHS), protocatechuate (PCA).

NATIVE METHYLOTROPHY

One potential method for transitioning to a methanol feedstock for chemical synthesis is to engineer native methylotrophs to produce value-added chemicals. There are three known biosynthetic pathways that exist for methanol assimilation: the serine pathway, the ribulose monophosphate (RuMP) pathway, or the ribulose biphosphate (RuBP) pathway.⁴⁵ The serine pathway requires an input of two ATP and two NADH equivalents to produce of one pyruvate molecule, and the RuBP pathway requires an input of seven ATP equivalents to produce of one pyruvate molecule.¹⁴ The RuMP pathway, however, requires an input of one ATP and one NAD⁺ equivalent for every molecule of pyruvate made, which makes it the most bioenergetically favorable (Figure 6).¹⁴

Many native methylotrophs are also thermophiles, which is cost beneficial as volumes increase. At 200 L, the cost of cooling organisms growing on glucose at 35 °C and those growing on methanol at 50 °C is the same. As the volume increases, the cost difference between the culturing conditions increases, and growth on methanol at 50 °C is, in fact, cheaper.⁴⁶ Additionally, the increased temperature and use of methanol as the only carbon source decreases the likelihood of contamination from other microorganisms during chemical production.¹⁷

Genetic engineering of native methylotrophs is challenging because these organisms are generally less studied than common production hosts such as *E. coli* or *C. glutamicum*.¹⁷ Consequently, the genetic engineering options are limited; however, efforts toward pursuing native methylotrophy are focused on expanding the genetic toolboxes for these organisms.

3 $\text{H}_3\text{C}-\text{OH}$

methanol

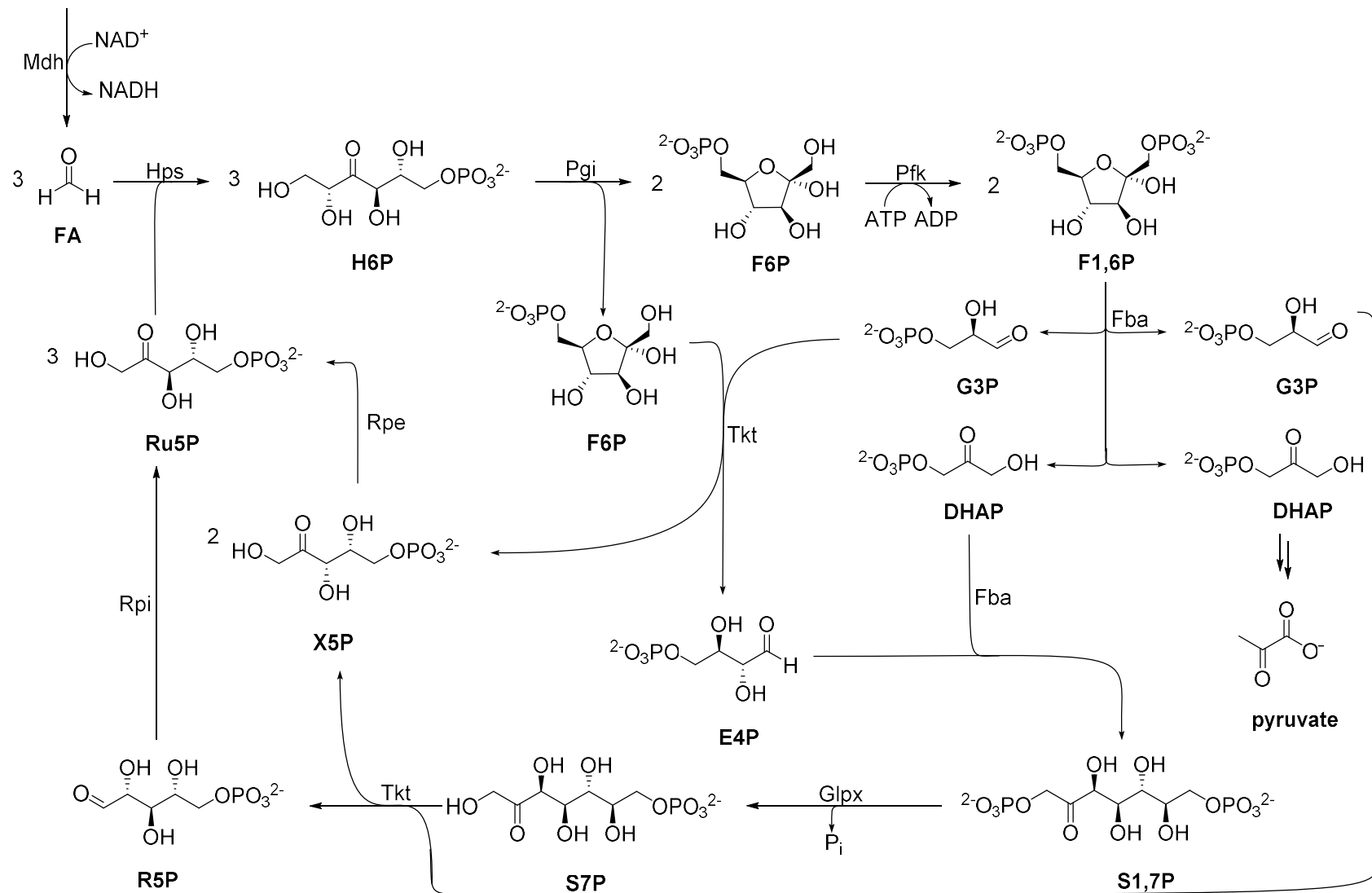


Figure 6. The ribulose monophosphate pathway for methanol assimilation. Three molecules of methanol produce one molecule

Figure 6 (cont'd) of pyruvate. Abbreviations: formaldehyde (FA), hexulose-6-phosphate (H6P), fructose-6-phosphate (F6P), fructose-1,6-biphosphate (F1,6P), glyceraldehyde-3-phosphate (G3P), dihydroxyacetonephosphate (DHAP), D-erythrose-4-phosphate (E4P), sedoheptulose-1,7,-biphosphate (S1,7P), sedoheptulose-7-phosphate (S7P), ribose-5-phosphate (R5P), xylose-5-phosphate (X5P), ribulose-5-phosphate (Ru5P), methanol dehydrogenase (Mdh), hexulose phosphate synthase (Hps), phosphohexulose isomerase (Phi), phosphofructokinase (Pfk), fructose-bisphosphate aldolase (Fba), sedoheptulose-1,7-bisphosphatase (Glp), transketolase (Tkt), ribose-5-phosphate isomerase (Rpi), ribulosephosphate-3-epimerase (Rpe).

SYNTHETIC METHYLOTROPHY

Another potential method for switching to a methanol feedstock is to genetically engineer a common production host such as *E. coli*, *S. cerevisiae*, or *C. glutamicum* that is currently used for industrial-scale chemical synthesis, to consume methanol. Synthetic methylotrophy has been a hot topic in current research; however, no organism has been successfully engineered or identified through selection methods to grow solely on methanol, despite many years of developing the technology.⁴⁷ Some challenges are slow cellular growth and inefficient methanol uptake.⁴⁸ Native nonmethylotrophs engineered to assimilate methanol require a sugar co-substrate, which makes them pseudo-synthetic methylotrophs. In addition, these pseudo-synthetic methylotrophs do not produce the same quantity of target molecule as native methylotrophs.⁴⁹ For example, a pseudo-synthetic methylotrophic strain of *C. glutamicum* can only produce 1.5 g L⁻¹ of cadaverine, while *Bacillus methanolicus*, a native methylotroph, can produce 11.3 g L⁻¹.^{49, 50}

To introduce methylotrophy into native nonmethylotrophs, only the genes encoding three proteins must be used to transform the host: methanol dehydrogenase (Mdh), hexulose phosphate synthase (Hps), and phosphohexulose isomerase (Phi) (Figure 7).⁵¹ These enzymes convert methanol into fructose-6-phosphate, which can then be funneled into glycolysis. In theory, this should provide all the carbon and energy the cell needs.⁵¹ The genes encoding the proteins Mdh, Hps, and Phi from *Bacillus methanolicus* are a gold standard for synthetic methylotrophy because they have high activity and do not require expression of additional co-factors.⁴⁷ While this trifecta of enzymes can assimilate carbon from methanol in native nonmethylotrophs, the cells cannot tolerate the concentration of methanol needed for sufficient carbon flux because formaldehyde accumulates to toxic levels.⁵¹ Expression of enzymes required for ribulose-5-phosphate (Ru5P) regeneration are tightly regulated by methanol concentration in native methylotrophs. This regulation is nonexistent in native nonmethylotrophs, which limits Ru5P availability. With an inadequate supply of Ru5P, formaldehyde accumulates to toxic levels. Since

methanol alone is not enough to sustain the cells, current pseudo-synthetic methylotrophs continue to rely on sugar substrates while being cultured in methanol.

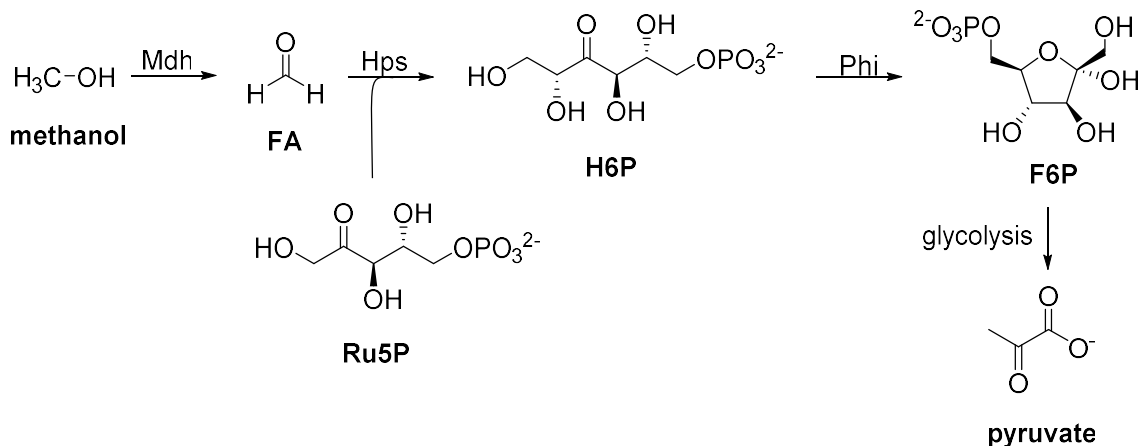


Figure 7. The three RuMP pathway proteins necessary for introducing synthetic methylotrophy into native nonmethylotrophs. Methanol dehydrogenase, Hexulosephosphate synthase, and phosphohexulose isomerase produce fructose-6-phosphate from methanol which can be converted to pyruvate by glycolysis. Abbreviations: formaldehyde (FA), hexulose-6-phosphate (H6P), fructose-6-phosphate (F6P), ribulose-5-phosphate (Ru5P), methanol dehydrogenase (Mdh), hexulosephosphate synthase (Hps), phosphohexulose isomerase (Phi).

PRODUCTION OF CIS,CIS-MUCONIC ACID BY *B. METHANOLICUS*

As discussed earlier, genes from *Bacillus methanolicus* are widely used for the implementation of synthetic methylotrophy.⁴⁷ For the same reasons, this organism is also a promising candidate for the bioconversion of methanol into value-added chemicals.⁵² *B. methanolicus* is a gram-positive, thermophilic, and facultative methylotrophic bacterium. It has the ability to form spores at lower temperatures, but not at its optimal culturing temperature of 50 °C.⁵³ Engineering this organism has historically been challenging because its metabolism was not well understood, and its genetic toolbox is been limited.⁴⁵ However, much progress has been made in recent years on these issues. The genome, transcriptome, and proteome of *B. methanolicus* have been elucidated and analyzed.⁵⁴⁻⁵⁶ A thermotolerant Cas9 was discovered in *Geobacillus*

thermodenitrificans T12 that may be used in developing a CRISPR-Cas system in this organism, although not yet demonstrated in *B. methanolicus*.⁵⁷ Most recently, a CRISPRi system for *B. methanolicus* has been described and tested.⁵⁸

Ultimately, our goal is to generate a *B. methanolicus* strain that can produce ccMA from methanol. *B. methanolicus* was chosen as the engineering host because it uses the RuMP for methanol assimilation. We will delete *aroE* in *B. methanolicus* and introduce three exogenous genes encoding for DHS dehydratase, protocatechuic acid decarboxylase, and catechol 1,2-dioxygenase (Figure 5).

Factors limiting carbon flow into the shikimate pathway include E4P and PEP availability.^{59,60} In *E. coli*, PEP is required for the transport of glucose into the cell, which decreases the PEP cellular concentration. Methanol uptake by the RuMP pathway does not have the same requirement and thus may increase the availability of PEP. E4P may also be more available in *B. methanolicus* because of its key role in the RuMP pathway. It has been shown in *E. coli* that higher concentrations of transketolase increases carbon flow into the shikimate pathway by producing more intracellular E4P.⁵⁹ *B. methanolicus* encodes two transketolase genes: one on the chromosome and one on its native plasmid (pBM19).⁶¹ The plasmid pBM19 carries genes that encode for a majority of the RuMP proteins.⁶² When grown on methanol, the mRNA level of the plasmid-encoded transketolase increases almost 15-fold while the mRNA level of the chromosome-encoded transketolase remains unchanged.⁶¹ This result suggests that the two transketolases serve different roles in the cell. The chromosome-encoded transketolase may function as an E4P producer for the shikimate pathway.

The shikimate pathway has not yet been studied in *B. methanolicus*. Here, characterization of the first enzyme of the shikimate pathway (DAHP synthase) and foundations for chemical mutagenesis on *B. methanolicus* are reported.

REFERENCES

REFERENCES

1. Ungerer, P. Petroleum. In *Encyclopedia of Geochemistry: A Comprehensive Reference Source on the Chemistry of the Earth*; White, W. M., Ed.; Springer International Publishing: Cham, 2018; pp 1207–1219.
2. Sillanpää, M.; Ncibi, C. Legacy of Petroleum-Based Economy. In *A Sustainable Bioeconomy: The Green Industrial Revolution*; Springer International Publishing: Cham, 2017. pp. 1-28
3. IPCC, 2013: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.
4. Correa, D. F.; Beyer, H. L.; Fargione, J. E.; Hill, J. D.; Possingham, H. P.; Thomas-Hall, S. R.; Schenk, P. M. Towards the Implementation of Sustainable Biofuel Production Systems. *Renewable Sustainable Energy Rev.* **2019**, *107*, 250–263.
5. EPA. *Greenhouse Gas Emissions from Large Facilities*. U.S. Environmental Protection Agency: Washington, DC, **2016**.
6. Purushothama Raj, P. *Soil Mechanics and Foundation Engineering*; Pearson Education India, 2008.
7. Sillanpää, M.; Ncibi, C. Bioeconomy: The Path to Sustainability. In *A Sustainable Bioeconomy: The Green Industrial Revolution*; Springer International Publishing: Cham, 2017. pp. 29-54
8. Skoog, E.; Shin, J. H.; Saez-Jimenez, V.; Mapelli, V.; Olsson, L. Biobased Adipic Acid - The Challenge of Developing the Production Host. *Biotechnol. Adv.* **2018**, *36* (8), 2248–2263.
9. Carriquiry, M. A.; Du, X.; Timilsina, G. R. Second Generation Biofuels: Economics and Policies. *Energy Policy* **2011**, *39* (7), 4222–4234.
10. Jones, J. P. H.; Wang, Z. M.; McCarl, B. A.; Wang, M. Policy Uncertainty and the US Ethanol Industry. *Sustain. Sci. Pract. Policy* **2017**, *9* (11), 2056.
11. Mohanty, S. K.; Swain, M. R. Chapter 3 - Bioethanol Production From Corn and Wheat: Food, Fuel, and Future. In *Bioethanol Production from Food Crops*; Ray, R. C., Ramachandran, S., Eds.; Academic Press, 2019; pp 45–59.

12. Frona, D.; Szenderak, J.; Harangi-Rakos, M. The Challenge of Feeding the World. *Sustain. Sci. Pract. Policy* **2019**, *11* (20). <https://doi.org/10.3390/su11205816>.
13. Cotton, C. A.; Claassens, N. J.; Benito-Vaquerizo, S.; Bar-Even, A. Renewable Methanol and Formate as Microbial Feedstocks. *Curr. Opin. Biotechnol.* **2019**, *62*, 168–180.
14. Whitaker, W. B.; Sandoval, N. R.; Bennett, R. K.; Fast, A. G.; Papoutsakis, E. T. Synthetic Methylophony: Engineering the Production of Biofuels and Chemicals Based on the Biology of Aerobic Methanol Utilization. *Curr. Opin. Biotechnol.* **2015**, *33*, 165–175.
15. Irla, M.; Nærdal, I.; Brautaset, T.; Wendisch, V. F. Methanol-Based γ -Aminobutyric Acid (GABA) Production by Genetically Engineered *Bacillus methanolicus* Strains. *Ind. Crops Prod.* **2017**, *106*, 12–20.
16. *Methanex Monthly Average Regional Posted Contract Price History*; Methanex. (<https://www.methanex.com/our-business/pricing>)
17. Pfeifenschneider, J.; Brautaset, T.; Wendisch, V. F. Methanol as Carbon Substrate in the Bio-Economy: Metabolic Engineering of Aerobic Methylophony Bacteria for Production of Value-Added Chemicals: Methanol as Carbon Substrate in the Bio-Economy. *Biofuels Bioprod. Biorefin.* **2017**, *11* (4), 719–731.
18. Su, Y.; Lü, L.; Shen, W.; Wei, S. An Efficient Technique for Improving Methanol Yield Using Dual CO₂ Feeds and Dry Methane Reforming. *Frontiers of Chemical Science and Engineering* **2019**. <https://doi.org/10.1007/s11705-019-1849-5>.
19. Du, Jing; Shao, Zengyi; Zhao, Huimin. Engineering Microbial Factories for Synthesis of Value-Added Products. *J. Ind. Microbiol. Biotechnol.* **2011**, *38*, 873–890.
20. Adipic Acid. In *Kirk-Othmer Encyclopedia of Chemical Technology*; John Wiley & Sons, Inc., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2014; Vol. 34, p 53.
21. Aryapratama, R.; Janssen, M. Prospective Life Cycle Assessment of Bio-Based Adipic Acid Production from Forest Residues. *J. Clean. Prod.* **2017**, *164*, 434–443.
22. Gentry, J. C. Benzene Production and Economics: A Review. *Asia-Pacific Jnl of Chem. Eng* **2007**, *2* (4), 272–277.
23. Qingdao Echemi Technology Co., Ltd. Adipic Acid. <https://www.echemi.com/productsInformation/pd20150901270-adipic-acid.html> (accessed Jan 25, 2020).
24. Deng, Y.; Ma, L.; Mao, Y. Biological Production of Adipic Acid from Renewable Substrates: Current and Future Methods. *Biochem. Eng. J.* **2016**, *105*, 16–26.

25. Niu, W.; Draths, K. M.; Frost, J. W. Benzene-Free Synthesis of Adipic Acid. *Biotechnol. Prog.* **2002**, *18* (2), 201–211.
26. Vardon, D. R.; Franden, M. A.; Johnson, C. W.; Karp, E. M.; Guarnieri, M. T.; Linger, J. G.; Salm, M. J.; Strathmann, T. J.; Beckham, G. T. Adipic Acid Production from Lignin. *Energy Environ. Sci.* **2015**, *8* (2), 617–628.
27. Moon, T. S.; Yoon, S.-H.; Lanza, A. M.; Roy-Mayhew, J. D.; Prather, K. L. J. Production of Glucaric Acid from a Synthetic Pathway in Recombinant *Escherichia coli*. *Appl. Environ. Microbiol.* **2009**, *75* (3), 589–595.
28. Chen, N.; Wang, J.; Zhao, Y.; Deng, Y. Metabolic Engineering of *Saccharomyces cerevisiae* for Efficient Production of Glucaric Acid at High Titer. *Microb. Cell Fact.* **2018**, *17* (1), 67.
29. Deng, Y.; Mao, Y. Production of Adipic Acid by the Native-Occurring Pathway in *Thermobifida fusca* B6. *J. Appl. Microbiol.* **2015**, *119* (4), 1057–1063.
30. Zhao, M.; Huang, D.; Zhang, X.; Koffas, M. A. G.; Zhou, J.; Deng, Y. Metabolic Engineering of *Escherichia coli* for Producing Adipic Acid through the Reverse Adipate-Degradation Pathway. *Metab. Eng.* **2018**, *47*, 254–262.
31. Beardslee, T.; Picataggio, S. Bio-Based Adipic Acid from Renewable Oils. *Lipid Technology* **2012**, *24* (10), 223–225.
32. Joo, J. C.; Khusnutdinova, A. N.; Flick, R.; Kim, T.; Bornscheuer, U. T.; Yakunin, A. F.; Mahadevan, R. Alkene Hydrogenation Activity of Enoate Reductases for an Environmentally Benign Biosynthesis of Adipic Acid. *Chem. Sci.* **2017**, *8* (2), 1406–1413.
33. Sun, J.; Raza, M.; Sun, X.; Yuan, Q. Biosynthesis of Adipic Acid via Microaerobic Hydrogenation of Cis,cis-Muconic Acid by Oxygen-Sensitive Enoate Reductase. *J. Biotechnol.* **2018**, *280*, 49–54.
34. Raj, K.; Partow, S.; Correia, K.; Khusnutdinova, A. N.; Yakunin, A. F.; Mahadevan, R. Biocatalytic Production of Adipic Acid from Glucose Using Engineered *Saccharomyces cerevisiae*. *Metab Eng Commun* **2018**, *6*, 28–32.
35. Bentley, R. The Shikimate Pathway--a Metabolic Tree with Many Branches. *Crit. Rev. Biochem. Mol. Biol.* **1990**, *25* (5), 307–384.
36. Light, S. H.; Anderson, W. F. The Diversity of Allosteric Controls at the Gateway to Aromatic Amino Acid Biosynthesis. *Protein Sci.* **2013**, *22* (4), 395–404.
37. Bai, Y.; Lang, E. J. M.; Nazmi, A. R.; Parker, E. J. Domain Cross-Talk within a Bifunctional Enzyme Provides Catalytic and Allosteric Functionality in the Biosynthesis of Aromatic Amino Acids. *J. Biol. Chem.* **2019**, *294* (13), 4828–4842

38. Fan, Y.; Cross, P. J.; Jameson, G. B.; Parker, E. J. Exploring Modular Allostery via Interchangeable Regulatory Domains. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (12), 3006–3011.
39. Nazmi, A. R.; Lang, E. J. M.; Bai, Y.; Allison, T. M.; Othman, M. H.; Panjikar, S.; Arcus, V. L.; Parker, E. J. Interdomain Conformational Changes Provide Allosteric Regulation En Route to Chorismate. *J. Biol. Chem.* **2016**, *291* (42), 21836–21847.
40. Cross, P. J.; Dobson, R. C. J.; Patchett, M. L.; Parker, E. J. Tyrosine Latching of a Regulatory Gate Affords Allosteric Control of Aromatic Amino Acid Biosynthesis. *J. Biol. Chem.* **2011**, *286* (12), 10216–10224.
41. Jiao, W.; Blackmore, N. J.; Nazmi, A. R.; Parker, E. J. Quaternary Structure Is an Essential Component That Contributes to the Sophisticated Allosteric Regulation Mechanism in a Key Enzyme from *Mycobacterium tuberculosis*. *PLoS One* **2017**, *12* (6), e0180052.
42. Liu, C.-G.; Xiao, Y.; Xia, X.-X.; Zhao, X.-Q.; Peng, L.; Srinophakun, P.; Bai, F.-W. Cellulosic Ethanol Production: Progress, Challenges and Strategies for Solutions. *Biotechnol. Adv.* **2019**, *37* (3), 491–504.
43. Marriott, P. E.; Gómez, L. D.; McQueen-Mason, S. J. Unlocking the Potential of Lignocellulosic Biomass through Plant Science. *New Phytol.* **2016**, *209* (4), 1366–1381.
44. Ponnusamy, V. K.; Nguyen, D. D.; Dharmaraja, J.; Shobana, S.; Banu, J. R.; Saratale, R. G.; Chang, S. W.; Kumar, G. A Review on Lignin Structure, Pretreatments, Fermentation Reactions and Biorefinery Potential. *Bioresour. Technol.* **2019**, *271*, 462–472.
45. Wang, X.; Wang, X.; Lu, X.; Ma, C.; Chen, K.; Ouyang, P. Methanol Fermentation Increases the Production of NAD(P)H-Dependent Chemicals in Synthetic Methylophilic *Escherichia coli*. *Biotechnol. Biofuels* **2019**, *12*, 17.
46. Komives, C. F.; Cheung, L. Y.-Y.; Pluschkell, S. B.; Flickinger, M. C. Growth of *Bacillus methanolicus* in Seawater-Based Media. *J. Ind. Microbiol. Biotechnol.* **2005**, *32* (2), 61–66.
47. Heux, S.; Brautaset, T.; Vorholt, J. A.; Wendisch, V. F.; Portais, J. C. Synthetic Methylophilic: Past, Present, and Future. In *Methane Biocatalysis: Paving the Way to Sustainability*; Kalyuzhnaya, M. G., Xing, X.-H., Eds.; Springer International Publishing: Cham, 2018; pp 133–151.
48. Chistoserdova, L.; Kalyuzhnaya, M. G. Current Trends in Methylophilic. *Trends Microbiol.* **2018**, *26* (8), 703–714.
49. Zhang, W.; Song, M.; Yang, Q.; Dai, Z.; Zhang, S.; Xin, F.; Dong, W.; Ma, J.; Jiang, M. Current Advance in Bioconversion of Methanol to Chemicals. *Biotechnol. Biofuels* **2018**, *11*, 260.

50. Naerdal, I.; Pfeifenschneider, J.; Brautaset, T.; Wendisch, V. F. Methanol-Based Cadaverine Production by Genetically Engineered *Bacillus methanolicus* Strains. *Microb. Biotechnol.* **2015**, *8* (2), 342–350.
51. Antoniewicz, M. R. Synthetic Methylophony: Strategies to Assimilate Methanol for Growth and Chemicals Production. *Curr. Opin. Biotechnol.* **2019**, *59*, 165–174.
52. Müller, J. E. N.; Heggeset, T. M. B.; Wendisch, V. F.; Vorholt, J. A.; Brautaset, T. Methylophony in the Thermophilic *Bacillus methanolicus*, Basic Insights and Application for Commodity Production from Methanol. *Appl. Microbiol. Biotechnol.* **2015**, *99* (2), 535–551.
53. Schendel, F. J.; Bremmon, C. E.; Flickinger, M. C.; Guettler, M.; Hanson, R. S. L-Lysine Production at 50 Degrees C by Mutants of a Newly Isolated and Characterized Methylophony *Bacillus* sp. *Appl. Environ. Microbiol.* **1990**, *56* (4), 963–970.
54. Irla, M.; Neshat, A.; Winkler, A.; Albersmeier, A.; Heggeset, T. M. B.; Brautaset, T.; Kalinowski, J.; Wendisch, V. F.; Rückert, C. Complete Genome Sequence of *Bacillus methanolicus* MGA3, a Thermotolerant Amino Acid Producing Methylophony. *J. Biotechnol.* **2014**, *188*, 110–111.
55. Irla, M.; Neshat, A.; Brautaset, T.; Rückert, C.; Kalinowski, J.; Wendisch, V. F. Transcriptome Analysis of Thermophilic Methylophony *Bacillus methanolicus* MGA3 Using RNA-Sequencing Provides Detailed Insights into Its Previously Uncharted Transcriptional Landscape. *BMC Genomics* **2015**, *16*, 73.
56. Müller, J. E. N.; Litsanov, B.; Bortfeld-Miller, M.; Trachsel, C.; Grossmann, J.; Brautaset, T.; Vorholt, J. A. Proteomic Analysis of the Thermophilic Methylophony *Bacillus methanolicus* MGA3. *Proteomics* **2014**, *14* (6), 725–737.
57. Mougiakos, I.; Mohanraju, P.; Bosma, E. F.; Vrouwe, V.; Finger Bou, M.; Naduthodi, M. I. S.; Gussak, A.; Brinkman, R. B. L.; van Kranenburg, R.; van der Oost, J. Characterizing a Thermostable Cas9 for Bacterial Genome Editing and Silencing. *Nat. Commun.* **2017**, *8* (1), 1647.
58. Schultenkämper, K.; Brito, L. F.; López, M. G.; Brautaset, T.; Wendisch, V. F. Establishment and Application of CRISPR Interference to Affect Sporulation, Hydrogen Peroxide Detoxification, and Mannitol Catabolism in the Methylophony Thermophile *Bacillus methanolicus*. *Appl. Microbiol. Biotechnol.* **2019**, *103* (14), 5879–5889.
59. Draths, K. M.; Pompliano, D. L.; Conley, D. L.; Frost, J. W.; Berry, A.; Disbrow, G. L.; Staversky, R. J.; Lieve, J. C. Biocatalytic Synthesis of Aromatics from D-Glucose: The Role of Transketolase. *J. Am. Chem. Soc.* **1992**, *114* (10), 3956–3962.
60. Chandran, S. S.; Yi, J.; Draths, K. M.; von Daeniken, R.; Weber, W.; Frost, J. W. Phosphoenolpyruvate Availability and the Biosynthesis of Shikimic Acid. *Biotechnol. Prog.* **2003**, *19* (3), 808–814.

61. Markert, B.; Stolzenberger, J.; Brautaset, T.; Wendisch, V. F. Characterization of Two Transketolases Encoded on the Chromosome and the Plasmid pBM19 of the Facultative Ribulose Monophosphate Cycle Methylothermophilic *Bacillus methanolicus*. *BMC Microbiol.* **2014**, *14*, 7.
62. Brautaset, T.; Jakobsen M, Ø. M.; Flickinger, M. C.; Valla, S.; Ellingsen, T. E. Plasmid-Dependent Methylothermophily in Thermotolerant *Bacillus methanolicus*. *J. Bacteriol.* **2004**, *186* (5), 1229–1238.

CHAPTER 2 – The *Bacillus methanolicus* MGA3 DAHP Synthases

OVERVIEW

The specific characteristic(s) which separated DAHP synthase enzymes into their computationally generated subclass was unknown for many years.¹ It is now understood that the classification of DAHP synthases is based on the structural features that have evolved to provide allosteric regulation.² Multiple research teams have established a DAHP synthase classification system, which occasionally leads to confusion when comparing studies on various DAHP synthases. Table 1 relates the classification systems to each other.³⁻⁶ As the body of knowledge regarding DAHP synthases grows, tiers of the classification system are added. The Walker-Parker classification system is referred to in this work as it is the most comprehensive thus far.

Table 1. DAHP synthase classification systems. Each subclass is represented by a row. NA (not applicable).

Walker-Parker ^{5,6}	Woodard ⁴	Jensen ³
Type I α	Class II	AroAl α
Type I β	Class I	AroAl β
Type II $_A$	NA	NA
Type II $_B$	NA	NA

A catalytic core comprised of a TIM (β/α)₈ barrel and a 2 β -2 α loop is universal to all DAHP synthases that have been experimentally examined to date.² The most general classification is bacteria or plant-like where Type I DAHP synthases are bacteria-like and Type II DAHP synthases are plant-like.¹ Structurally, Type II DAHP synthases contain an N-terminal extension which Type I does not. Regulatory elements then divide each type more finely (Figure 8).

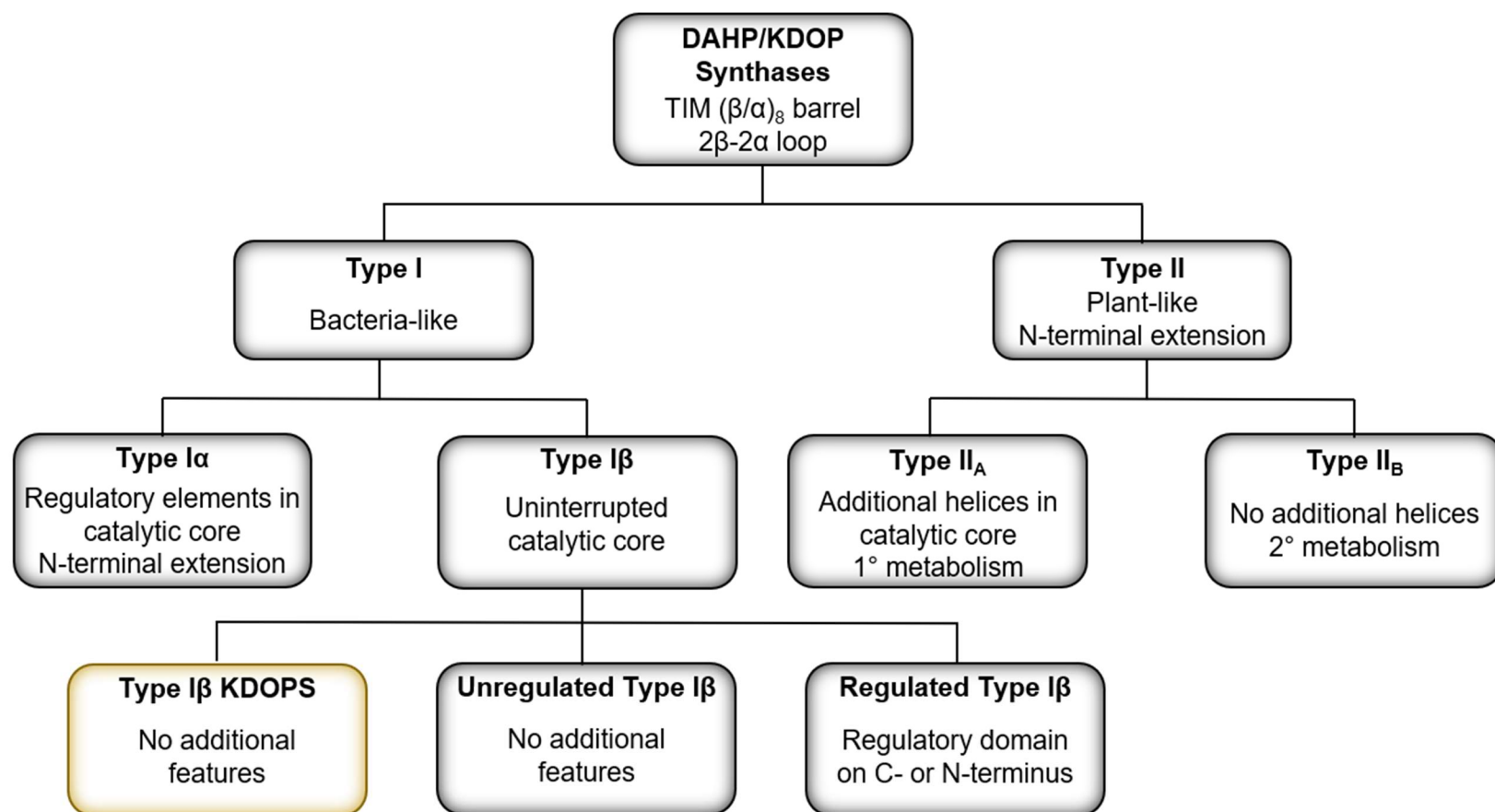


Figure 8. DAHP and KDOP synthase classification tree. The only KDOP synthase (KDOPS) classification is highlighted in yellow.

Type I β DAHP synthases are the simplest, containing an uninterrupted catalytic core which can either be regulated or unregulated. The unregulated Type I β DAHP synthases comprise only the catalytic core domain, while regulated Type I β DAHP synthases have a regulatory domain fused to either the C- or N-terminus. This regulatory domain can either be chorismate mutase (CM) domain, or an ACT-like domain.² Independently, CM is an enzyme that converts chorismate, the product of the shikimate pathway, to prephenate. When DAHP synthase is linked to a CM domain the enzyme is allosterically inhibited by chorismate and prephenate.² ACT-like domains are common in enzymes regulated by amino acids.⁸ This regulatory domain is found in a variety of enzymes and is named eponymously: aspartate kinase, chorismate mutase, and TyrA (prephenate dehydrogenase). When an ACT-like domain is present on Type I β DAHP synthases, the enzyme is regulated by one or more of the aromatic amino acids.²

All Type I α DAHP synthases are allosterically regulated. Their catalytic core is interrupted by regulatory elements and contain an N-terminal extension.² The general structural differences between the various Type I DAHP synthases are compared in Figure 9.

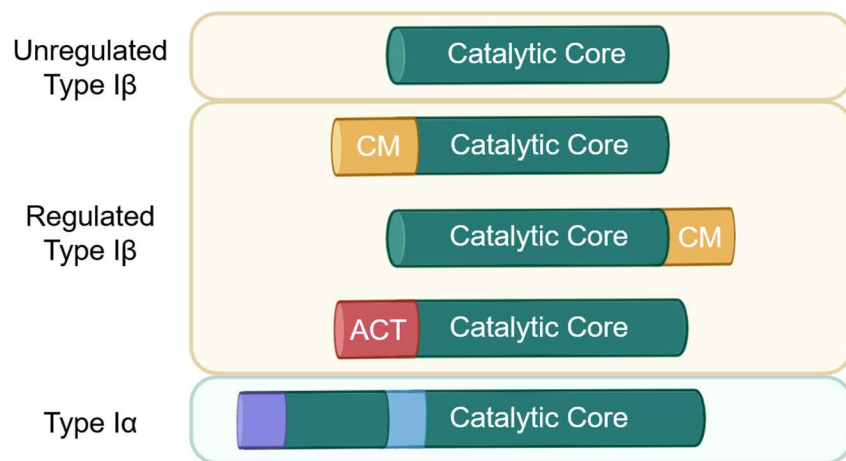


Figure 9. Renderings of the Type I DAHP synthases. The Type I β DAHP synthases are shown in yellow boxes. The Type I α DAHP synthase is shown in a green box. The DAHP synthase catalytic core DAHPS is depicted by a teal cylinder. The chorismate mutase (CM) domain is depicted by a yellow cylinder. The ACT domain is depicted by an orange cylinder. The regulatory domain in the Type I α DAHP synthase is depicted by a blue cylinder. The N-terminal extension is represented by a purple cylinder. All Type I β DAHP synthases contain an uninterrupted catalytic core. Unregulated Type I β DAHP synthases contain only the catalytic core, while regulated Type I β DAHP synthases contain either a CM domain fused to the C- or N-terminus or an ACT domain fused to the N-terminus. Fusion to a CM domain renders the Type I β DAHP synthase allosterically inhibited by chorismate and prephenate. Fusion to an ACT domain renders the Type I β DAHP synthase allosterically inhibited by one or more of the aromatic amino acids. The catalytic core of Type I α DAHP synthases are interrupted by regulatory elements, which render the DAHP synthase allosterically inhibited by one or more of the aromatic amino acids.

Finally, Type II DAHP synthase can be divided even further into subclasses based on their roles in either primary or secondary metabolism.⁵ Type II_A DAHP synthases are involved in primary metabolism and contain additional helices within the catalytic core. Type II_B DAHP synthases are involved in secondary metabolism and lack these additional helices.

A similar enzyme, 3-deoxy-D-mannooctulosonate-8-phosphate synthase (KDOPS) is classified in tandem with DAHP synthases.⁹ The DAHP synthase classification system is based upon structural similarities (Figure 8). KDOP synthases also comprise an uninterrupted TIM (β/α)₈ barrel catalytic core and a 2 β -2 α loop. These structural features classify KDOP synthases as Type I β . KDOP synthases catalyze the condensation of PEP and arabinose-5-phosphate (A5P), a similar reaction to those catalyzed by DAHP synthases (Figure 10).

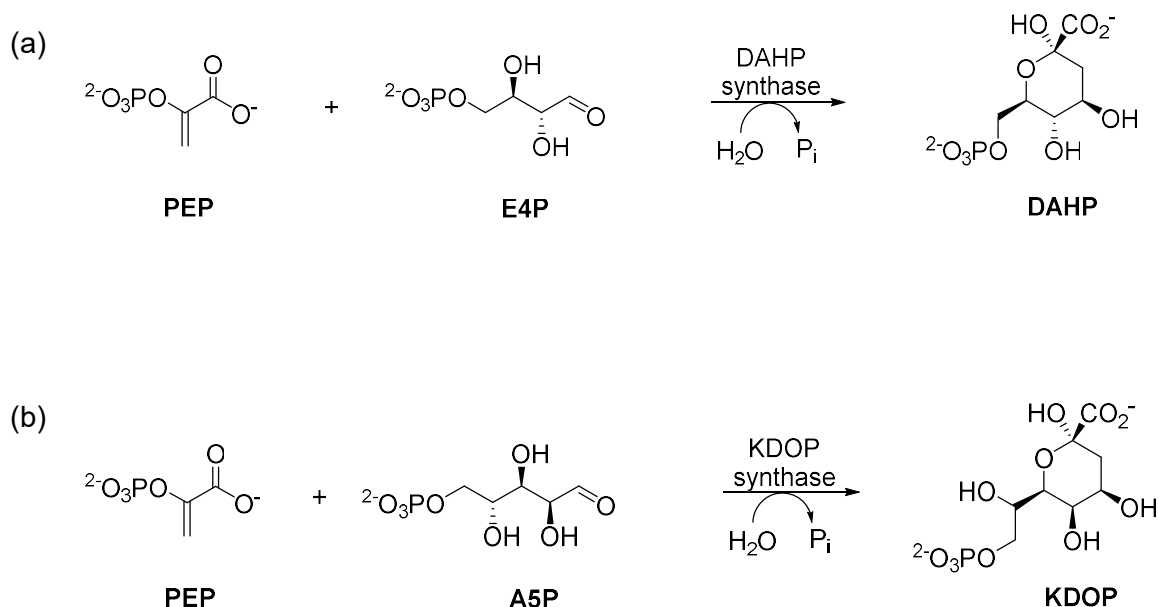


Figure 10. The condensation reactions performed by DAHP synthase and KDOP synthase. (a) the reaction catalyzed by DAHP synthase. (b) the reaction catalyzed by KDOP synthase. Abbreviations: phosphoenolpyruvate (PEP); D-erythrose-4-phosphate (E4P); 3-deoxy-D-arabinoheptulosonate-7-phosphate (DAHP); D-arabinose-5-phosphate (A5P); 3-deoxy-D-mannooctulosonate-8-phosphate (KDOP), inorganic phosphate (P_i).

The DAHP synthases are known to be allosterically inhibited by the aromatic amino acids, chorismate, or prephenate.² It was not known whether the *B. methanolicus* MGA3 DAHP synthases were inhibited by any of these molecules. As described in Chapter 1, a goal of this project is to manipulate the shikimate pathway in *B. methanolicus* MGA3 to produce ccMA from methanol. For maximum yields, the gene *aroE* must be knocked-out, abolishing aromatic amino acid production (Figure 5). An *aroE* mutant would be auxotrophic for the aromatic amino acids and aromatic vitamins, necessitating supplementation of the culturing media with these molecules. It is important to know what effect, if any, this supplementation will have on the DAHP synthases of MGA3. The shikimate pathway has not yet been studied in *B. methanolicus*, and it is crucial to understand the role of DAHP synthases in feedback inhibition and carbon flow regulation related to the shikimate pathway.

MULTIPLE SEQUENCE ALIGNMENT

As described above, the DAHP synthases are classified based on structural features. AroG1 and AroG2 were sorted into classes by aligning their amino acid sequences with those of representative DAHP synthases from each Type I class using the multiple sequence alignment tool, T-Coffee.¹⁰ Representative enzymes were chosen from DAHP synthases which have been experimentally verified to belong to their class. The unregulated Type I β representatives were from *Pyrococcus furiosus* (Pfu_unreg_I β) and *Aeropyrum pernix* (Ape_unreg_I β).^{11,12} The regulated Type I β representatives were from *Thermotoga maritima* and *Bacillus subtilis*. The *T. maritima* representative is inhibited by phenylalanine and tyrosine (Tma_phe/tyr_I β), while the *B. subtilis* representative is inhibited by chorismate and prephenate (Bsu_CA/PA_I β) (Table 2)^{13,14} The Type I α representatives were from *E. coli* and *Corynebacterium glutamicum*. The representative from *E. coli* is inhibited by tyrosine (Eco_tyr_I α), and the one from *C. glutamicum* is inhibited by phenylalanine and tyrosine (Cgl_phe/tyr_I α) (Table 2).^{15,16}

Table 2. Representative DAHP synthases and their allosteric inhibitors. Abbreviations: chorismate (CA), prephenate (PA), phenylalanine (Phe), tyrosine (Tyr), not applicable (NA).

	Unregulated Type I β		Regulated Type I β		Type I α	
Organism	<i>P. furiosus</i>	<i>A. pernix</i>	<i>T. maritima</i>	<i>B. subtilis</i>	<i>E. coli</i>	<i>C. glutamicum</i>
Inhibitor(s)	NA	NA	Phe, Tyr	CA, PA	Tyr	Phe, Tyr

The multiple sequence alignments (MSA) showed that AroG1 had highest percent identity with the unregulated Type I β models (Table 3). The MSA of AroG1 with representatives from both the unregulated Type I β (Figure 11) and the regulated Type I β (Figure 12) both showed high similarity (42-52% identity). As described in the overview, regulated Type I β DAHP synthases contain a regulatory domain fused to either terminus of the catalytic core. For the regulated Type I β representatives (from *T. maritima* and *B. subtilis*), the regulatory domain is present on the N-terminus.^{14,13} Notably, the MSA with the regulated Type I β representatives revealed an extended gap on the N-terminus of AroG1 which contains the regulatory domains of the regulated Type I β representatives. This MSA indicated that a fused regulatory domain is not present on AroG1. Further, the MSA of AroG1 with Type I α representatives showed very few conserved residues (Figure 13). Based on all these results, I hypothesize that AroG1 is an unregulated Type I β DAHP synthase.

The MSA of AroG2 with representatives from the unregulated Type I β (Figure 14) revealed a gap on the N-terminus of the representative sequences. The regulated Type I β (Figure 15) showed high similarity, especially with Bsu_CA/PA_I β . AroG2 had a 43% identity with Tma_phe/tyr_I β and a 78% identity with Bsu_CA/PA_I β . In contrast, the MSA of AroG2 with the Type I α representatives showed little similarity (Figure 16). Opposite to the AroG1 results, AroG2 appeared to have a regulatory domain on its N-terminus and showed the highest percent identity with the regulated Type I β models (Table 3). Based on these results, I hypothesize that AroG2 is a regulated Type I β DAHP synthase.

Table 3. The MSA results of AroG1 and AroG3 with representative DAHP synthases.

	Unregulated Type I β		Regulated Type I β		Type I α	
	<i>P. furiosus</i>	<i>A. pernix</i>	<i>T. maritima</i>	<i>B. subtilis</i>	<i>C. glutamicum</i>	<i>E. coli</i>
AroG1	51% (67%)	52% (69%)	42% (56%)	50% (59%)	12% (22%)	17% (32%)
AroG2	23% (42%)	29% (38%)	43% (57%)	78% (89%)	10% (22%)	17% (32%)

Percent identities between the *B. methanolicus* DAHP synthases and each representative DAHP synthase are shown. The percent similarities are included in parentheses.

Bme_AroG1	1	MSDKLQQEKSP	LLVSRHHKGE	DTIINIKQAV	I	GGFSHTL	I	AGPCSVESREQL	LVTVASVL	KKAGLTVLRGGAF	KPRTSPYDF	81						
Pfu_unreg_Iβ	1	MK	-	-	-	-	-	YSKEYKEK	TVVKINDV	KFGE	-FTI	I	AGPCSIESRDQ	IMKVAEFL	AEVGI	KVLRGGAF	KPRTSPYSF	68
Ape_unreg_Iβ	1	MAG	-	-	-	F	KGVKL	ALKSEERRE	TVVEVEGVR	I	GGGSKAV	I	AGPCSVESWEQ	VREAALAV	KEAGAHML	RGGAF	KPRTSPYSF	77
Bme_AroG1	82	QGLGEEGLK	I	MKDVADE	I	GLVTISE	I	MSPTQIDTAARY	I	D	I	FQIGARNMQNF	DLLKAVGQT	KKPILLKRG	LSATLEEF	MLA	162	
Pfu_unreg_Iβ	69	QGYGEKAL	RWMREAADE	YGLVT	VTEVMD	TRHVEL	VAKYS	D	I	LQIGARNSQNF	ELLKEVGK	VENPVLL	KRGMGNT	I	QELLYS	149		
Ape_unreg_Iβ	78	QGLGLEGLK	LLRRAGDE	AGLPVVTE	VLDPRH	VETVSRY	A	D	M	LQIGARNMQNF	P	LLREVGRSGK	VPVLLKRG	F	GNTVEELLA	158		
Bme_AroG1	163	AEYILH	QGNDRV	MLMERG	IRTFE	KATRNTLD	I	SAVPILKQETH	LPVLVDVTH	STGRKDILL	PCAKAALA	AAGADG	I	MVEVHP	243			
Pfu_unreg_Iβ	150	AEYIMA	QGNENV	ILCERG	IRTFE	TATRFTLD	I	SAVPVVKELSHLP	I	IVDP	SHPAGRRSLV	I	PLAKAAYA	I	GADGIMVEVHP	230		
Ape_unreg_Iβ	159	AEYIL	LEGWQV	VLVERG	IRTFE	PSTRFTLD	V	AAVAVLKEATH	LPVIVDP	SHPAGRRSLV	PALAKAGLA	AAGADGL	I	VEVHP	239			
Bme_AroG1	244	NPPAAL	SDAKQQVN	I	DEF	HEFWKGL	C	-	-	-	DSGLYH	275						
Pfu_unreg_Iβ	231	EPEKAL	SDSQQLT	FDD	FLQL	LKELEAL	GWKG	-	-	-	262							
Ape_unreg_Iβ	240	NPEEAL	SDAKQQLT	PGEF	ARLMGEL	R	-	-	-	WHRL	-	L	270					

Figure 11. Multiple sequence alignment of MGA3 AroG1 with unregulated Type I β DAHP synthases from *P. furiosus* and *A. pernix*. Dark purple highlights show 100% identity across all three sequences. Light purple highlights show 100% identity across two sequences. Unhighlighted residues are not conserved in these sequences. Green boxes encompass residues which are a part of the catalytic core.

Bme_AroG1	1	MSDK	4
Tma_phe/tyr_Iβ	1	MIVV - - - LKPGSTEEDIRKVVKLAESYNLKCHI SKGQERTVIGIIG - - DDYVVDKFE SLDC - - - - - V	59
Bsu_CA/PA_Iβ	1	MSNTELELLRQKADELNLQILKL INERGNVVKELGKAKEAAGVNRFPVRRER TMLNNI I ENNDGPFENSTIQHIFKEIFKA	81
Bme_AroG1	5	--- LQQE - - KSPLLVSRHHKGEDTI INIKQAVIGGFSHTL IAGPCSVESREQLVTVASVLKKAGLTVLRGGAFKPRTSPYD	80
Tma_phe/tyr_Iβ	60	ESVVRVL - - KPYKLVSREFHPEDTVIDLGDVKIGNGYFTI IAGPCSVEGREMLMETAHFLSELGVKVLRRGGAYKPRTSPYS	138
Bsu_CA/PA_Iβ	82	GLELQEEHDSKALLVSRKKKPEDTIVDIKGEKIGDQQRFI VGPCAVESYEQVAEVAAAANKQGIK I LRGGAFKPRTSPYD	162
Bme_AroG1	81	FQGLGEEGLKIMKDVADIEGLVTISEIMSPQTIDTAARYIDIFQIGARNMQNFDLLKAVGQTKKPI LLKRGLSATLEEFML	161
Tma_phe/tyr_Iβ	139	FQGLGEKGLLEYLREAADKYGMVVTEALGEDDLPKVAEYADI IQIGARNAQNFRLLSKAGSYNKPVLLKRGFMNTIEEFLL	219
Bsu_CA/PA_Iβ	163	FQGLGVEGLQILKRVADEFDLAVISEIVTPAHIEEALDYIDVIQIGARNMQNFELLKAAGAVKKPVLLKRGLAATISEFIN	243
Bme_AroG1	162	AAEYILHQGNDRVMLMERGIRTFEKATRNTLDISAVPILKQETHLPVLVDVTHSTGRKDI LLPCAKAALAAGADGIMVEVH	242
Tma_phe/tyr_Iβ	220	SAEYIANSNTKIILCERGIRTFEKATRNTLDISAVPIIRKESHLPILVDPSHSGRRDLVIPLSRAAIAVGAHGIIVEVH	300
Bsu_CA/PA_Iβ	244	AAEYIMSQNDQIILCERGIRTYETATRNTLDISAVPILKQETHLPVFDVTHSTGRRDLLPTAKAALAIGADGVMAEVH	324
Bme_AroG1	243	PNPPAALSDAKQQVNI DEFHEFWKGLCDSG - - - - - LY-H	275
Tma_phe/tyr_Iβ	301	PEPEKALSDGKQSLDFELFKELVQEMKKLADALGVKV-N	338
Bsu_CA/PA_Iβ	325	PDPSVALSDSAQQMAIPEFEKWLNELKPMV - - - - - KVNA	358

Figure 12. Multiple sequence alignment of MGA3 AroG1 with regulated Type Iβ DAHP synthases from *T. maritima* (regulated by phenylalanine and tyrosine) and *B. subtilis* (regulated by chorismate and prephenate). Dark purple highlights show 100% identity across all three sequences. Light purple highlights show 100% identity across two sequences. Unhighlighted residues are not conserved in these sequences. Green boxes encompass residues which are a part of the catalytic core. Yellow boxes encompass residues which are part of the CM domain. Red boxes encompass residues which are part of the ACT domain.

Bme_AroG1	1	MSDKL	-----	QQEKSP	LLVSRHHK	GEDT	I	-----	INIKQ	---AVIGG	34							
Cgl_phe/tyr_la	1	MNRGVSWTVD	IPKEVL	PDLPPL	PEGM	QQQF	EDT	I	SRDAK	QQPTW	DRAQAENVRK	I	LESVPP	IVVAPEV	LELKQKL	ADVANG	81	
Eco_tyr_la	1	MQKDALNNVH	ITDEQVL	---MTPEQL	KAAFPL	SLQQEAQ	IADSR	-----	KSIS	---DIIAG	50							
Bme_AroG1	35	F-SHT	-LIAGPCS	VESR	-----	EQLVTVAS	VLKK	-----	AGLTVLR	GGAF	KPRTSPY	DFQGLGE	-----	86				
Cgl_phe/tyr_la	82	K-AFL	-LQGGD	CAET	FESNTEPH	IRANVK	TLLQMAV	VLTYGAST	-PVI	KMAR	IAGQYAK	KPRSSDL	DGNGLPN	-----	150			
Eco_tyr_la	51	RDPRLL	VVCGPCS	IHDP	-----	ETALEY	ARRFKALAAE	VSDSLYL	VMRVYFE	KPRTTV	-GWKGL	INDPHMDGS	FDF	119				
Bme_AroG1	87	-E	-----													GL	89	
Cgl_phe/tyr_la	151	-	YRGDI	VNGVEAT	PEARRHDP	ARMIRAYAN	ASAAMNL	VRLTSSGTAD	LYRLSEWN	REFVANS	PAGARYEAL	AREIDS	GL	229				
Eco_tyr_la	120	VEA	-----													GL	124	
Bme_AroG1	90	KIM	-----	KDVADE	I	---GLVT	ISEIMSPTQ	IDTAARY	IDI	-FQ	I	GARNMQ	-NFDLLKAVG	QTKKPI	146			
Cgl_phe/tyr_la	230	RFME	ACGVSD	ESLRAADI	YCSHE	ALLVDY	ERSMLRLAT	---DEEGNEEL	YDL	SAHQLW	I	GERTRGMDD	FHVN	FASMI	SNPI	307		
Eco_tyr_la	125	QIA	-----	RKLLLE	LVNMGL	PLATEAL	DPNSPQY	LGD	LSW	-SA	I	GARTTE	-SQTHREMAS	GSLMPV	184			
Bme_AroG1	147	LLKRGL	SATLEEF	ML	-----	AAE	-----	YILHQ	GNDRV	MLMERG	I	RTFEKATRNT	LDISAVP	-ILKQETHL	-P	207		
Cgl_phe/tyr_la	308	GIKIGPGIT	PEEAVA	-----	YADKLD	PNFEPGR	LTI	VARM	SHDKVRS	VLPGVI	QA	-----	VEASGHKVI	366				
Eco_tyr_la	185	GFKNGTDG	SLATAINAMRAA	AQPHRF	VGINQAGQV	ALLQTQ	GNPDGHV	I	LRGGKAPN	---YSPADVAQCEKEME	QAGLRPS	262						
Bme_AroG1	208	VLVDVTH	STGRKD	-----	IL-L	-----	PCA-KAALAA	-----	GADGIM	VEVHP	---NPP	---AALSDAKQQV	NID	259				
Cgl_phe/tyr_la	367	WQSD	PMHGNTFTAS	NGYKTRH	-FDKV	I	DEVQGF	FEV	-H	RALGT	-----	HPGGI	HIEFTG	-EDVT	-ECLGGAED	ITD	VD	435
Eco_tyr_la	263	LMVDCSH	GNNSKD	-----	YRRQ	-----	PAVAESV	VQIKDGNRS	I	I	GLMIESNI	HEGNQSSE	QPRSE	MYGVSVT	327			
Bme_AroG1	260	EFHEFWKGL	CD	SG	---	LY	-----	-H									275	
Cgl_phe/tyr_la	436	LP	-GRYESACD	PR	---	LNTQQSLE	L	AF	LVAEMLRN								466	
Eco_tyr_la	328	DAC	I	SWE	-MTDALLRE	I	HQDLNGQL	TARVA	-----								356	

Figure 13. Multiple sequence alignment of MGA3 AroG1 with Type Ia DAHP synthases from *C. glutamicum* (regulated by phenylalanine and tyrosine) and *E. coli* (regulated by tyrosine). Dark purple highlights show 100% identity across all three sequences. Light purple highlights show 100% identity across two sequences. Unhighlighted residues are not conserved in these sequences. Green boxes encompass residues which are a part of the catalytic core. Blue boxes encompass residues which are a part of the Type Ia regulatory domains.

Bme_AroG2	1	MSNKELDQLRQRVDELNLELLSLINERARLVQELGRVKETQGVFRYPVRERKMLDLIKENNDGPFENSTIEHIFKEIFKA	81
Pfu_unreg_I β	1	MK-----YSKEY-----	7
Ape_unreg_I β	1	MAGFKGVKLALKSEE-----	15
Bme_AroG2	82	GLELQKDDHRKALLVSRKKKPDNTVIDLKGEKIGDGNPHFVFGPCAVESYEQVATVAESVKAKGLKLLRGGAYKPRTSPYD	162
Pfu_unreg_I β	8	-----KEKTVVKINDVKFGEF-TIIAGPCSIESRDQIMKVAEFLAEVGIKVLRRGGAFKPRTSPYS	67
Ape_unreg_I β	16	-----RRETVEVEGVRIGGSGKAVIAGPCSVESWEQVREAALAVKEAGAHMLRGGAFKPRTSPYS	76
Bme_AroG2	163	FQGLGLEGLKILKRIADEYDLAVISEITNPADIEAVQYIDVIQIGARNMQNFDLLKAAGAVNKPVLLKRGLAATIEEFIN	243
Pfu_unreg_I β	68	FQGYGEKALRWMREAADYGLVTVEVMDTRHVELVAKYSDILQIGARNSQNFELLKEVGKVENPVLLKRGMGNTIQELLY	148
Ape_unreg_I β	77	FQGLGLEGLKLLRRAGDEAGLPVVTEVLDPRHVETVSRYADMLQIGARNMQNFPLLREVGRSGKPVLLKRGFGNTVEELLA	157
Bme_AroG2	244	AAEYIMAHGNGQIILCERGIRTYERATRNTLDISAVPILKQETHLPVFVDVTHSTGRRDLL-----	305
Pfu_unreg_I β	149	SAEYIMAGNENVILCERGIRTFETATRTLDISAVPVVKELSHLP IIVDPSPHPAGRRSLVPLAKAAYAIGADGIMVEVH	229
Ape_unreg_I β	158	AAEYILLEGNWQVVVVERGIRTFEPSTRFTLDVAAVAVLKEATHLPVIVDPSPHPAGRRSLVPALAKAGLAAGADGLIVEVH	238
Bme_AroG2		-----	
Pfu_unreg_I β	230	PEPEKALSDSQQLTFDDFLQLLKELE-----ALGWKG-----	262
Ape_unreg_I β	239	PNPEEALSDAKQQLTPGEFARLMGELRWHRLPAAKAALAIGADGVMAEVHPDPAVALSDSAQQMDLKQFDEFYREIKKSV	319
Bme_AroG2		----	
Pfu_unreg_I β		----	
Ape_unreg_I β	320	LARV	323

Figure 14. Multiple sequence alignment of MGA3 AroG2 with unregulated Type I β DAHP synthases from *P. furiosus* and *A. pernix*. Dark purple highlights show 100% identity across all three sequences. Light purple highlights show 100% identity across two sequences. Unhighlighted residues are not conserved in these sequences. Green boxes encompass residues which are a part of the catalytic core. Yellow boxes encompass residues which are part of the CM domain.

Bme_AroG2	1	MSN	KELDQ	LQRVDE	LNLE	LS	INERAR	LVQE	IGRVKETQGV	FRYPV	RERK	MLDL	I	KENND	---	GP	FENS	-	T	I	E	H	I	F	K	E	I	F	K	A	81																																																						
Tma_phe/tyr_Iβ	1	M	I	---	---	V	V	L	KPGST	E	ED	IRKVV	KLA	E	SYN	L	KCH	I	SKGQ	---	---	---	---	---	---	---	---	---	---	---	---	63																																																					
Bsu_CA/PA_Iβ	1	MSN	TELEL	L	RQKADE	LN	LQ	I	L	KL	I	NERGN	VVKE	I	GKAKEA	QGVN	R	FDPV	RERT	MLNN	I	I	ENND	---	GP	FENS	-	T	I	Q	H	I	F	K	E	I	F	K	A	81																																													
Bme_AroG2	82	GLE	LQKDD	H	RKAL	L	VSR	KKK	P	DNT	V	I	D	L	KGEK	I	GDG	NPH	F	V	G	PCAVES	YEQ	VAT	V	A	E	S	V	K	A	K	G	L	K	L	R	G	G	A	Y	K	P	R	T	S	P	Y	D	F	Q	G	L	166																															
Tma_phe/tyr_Iβ	64	---	---	---	---	R	V	L	KPYK	L	VSR	EFH	P	E	D	T	V	I	D	L	G	D	V	K	I	G	N	G	Y	F	T	I	I	A	G	P	C	S	V	E	G	R	E	M	L	M	E	T	A	H	F	L	S	E	L	G	V	K	V	L	R	G	G	A	Y	K	P	R	T	S	P	Y	S	F	Q	G	L	142							
Bsu_CA/PA_Iβ	82	GLE	LQEED	H	S	KAL	L	VSR	KKK	P	E	D	T	I	V	D	I	KGEK	I	GDG	Q	Q	R	F	I	V	G	PCAVES	YEQ	V	A	E	V	A	A	A	A	K	Q	G	I	K	I	L	R	G	G	A	F	K	P	R	T	S	P	Y	D	F	Q	G	L	166																							
Bme_AroG2	167	GLE	GLK	I	L	K	R	I	A	D	E	Y	D	L	A	V	I	S	E	I	T	N	P	A	D	I	E	T	A	V	Q	Y	I	D	V	I	Q	I	G	A	R	N	M	Q	N	F	D	L	L	K	A	A	G	A	V	N	K	P	V	L	L	K	R	G	L	A	A	T	I	E	E	F	I	N	A	A	E	Y	I	M	A	H	251		
Tma_phe/tyr_Iβ	143	GEK	GLE	Y	L	R	E	A	A	D	K	Y	G	M	Y	V	V	T	E	A	L	G	E	D	L	P	K	V	A	E	Y	A	D	I	I	Q	I	G	A	R	N	A	Q	N	F	R	L	L	S	K	A	G	S	Y	N	K	P	V	L	L	K	R	G	F	M	N	T	I	E	E	F	L	L	S	A	E	Y	I	A	N	S	227			
Bsu_CA/PA_Iβ	167	GVE	GLQ	I	L	K	R	V	A	D	E	F	D	L	A	V	I	S	E	I	V	T	P	A	H	I	E	E	A	L	D	Y	I	D	V	I	Q	I	G	A	R	N	M	Q	N	F	E	L	L	K	A	A	G	A	V	N	K	P	V	L	L	K	R	G	L	A	A	T	I	S	E	F	I	N	A	A	E	Y	I	M	S	Q	251		
Bme_AroG2	252	GNG	Q	I	I	L	C	E	R	G	I	R	T	Y	E	R	A	T	R	N	T	L	D	I	S	A	V	P	I	L	K	Q	E	T	H	L	P	V	F	D	V	T	H	S	T	G	R	R	D	L	L	L	P	A	A	K	A	A	L	A	I	G	A	D	G	V	M	A	E	V	H	P	D	P	A	V	A	L	S	D	S	A	Q	336	
Tma_phe/tyr_Iβ	228	GNT	K	I	I	L	C	E	R	G	I	R	T	F	E	K	A	T	R	N	T	L	D	I	S	A	V	P	I	I	R	K	E	S	H	L	P	I	L	V	D	P	S	H	S	G	G	R	R	D	L	V	I	P	L	S	R	A	A	I	A	V	G	A	H	G	I	I	V	E	V	H	P	E	P	E	K	A	L	S	D	G	K	Q	312
Bsu_CA/PA_Iβ	252	GND	Q	I	I	L	C	E	R	G	I	R	T	Y	E	T	A	T	R	N	T	L	D	I	S	A	V	P	I	L	K	Q	E	T	H	L	P	V	F	D	V	T	H	S	T	G	R	R	D	L	L	L	P	T	A	K	A	A	L	A	I	G	A	D	G	V	M	A	E	V	H	P	D	P	S	V	A	L	S	D	S	A	Q	336	
Bme_AroG2	337	QMD	L	K	Q	F	D	E	F	Y	R	E	I	K	K	S	V	---	---	---	---	L	A	R	V	358																																																											
Tma_phe/tyr_Iβ	313	S	L	D	F	E	L	F	K	E	L	V	Q	E	M	K	K	L	A	D	A	L	G	V	K	V	N	-	338																																																								
Bsu_CA/PA_Iβ	337	QMA	I	P	E	F	E	K	W	L	N	E	L	K	P	M	V	---	---	---	---	K	V	N	A	358																																																											

Figure 15. Multiple sequence alignment of MGA3 AroG2 with regulated Type Iβ DAHP synthases from *T. maritima* (regulated by phenylalanine and tyrosine) and *B. subtilis* (regulated by chorismate and prephenate). Dark purple highlights show 100% identity across all three sequences. Light purple highlights show 100% identity across two sequences. Unhighlighted residues are not conserved in these sequences. Green boxes encompass residues which are a part of the catalytic core. Yellow boxes encompass residues which are part of the CM domain. Red boxes encompass residues which are part of the ACT domain.

Bme_AroG2	1	MSNKELDQLRQRVDELNLELLSLINERARLVQEI	GRVKETQGVFRYPVRRERKMLDLIKENNDGPF	FENSTIEHIFKEIFKA	81		
Cgl_phe/tyr_la	1	MNRGVSWTVD-IPKEVLPDLP		PLPEGMQQQFED	32		
Eco_tyr_la	1	MQKDALNNVH-ITDEQVLM		TPEQLKAAFPL	29		
Bme_AroG2	82	GLE--LQ-----KDDHRKALLVSRKKK	PDNTVIDLK--GEKIGDG-NPH-F-VFGPCAVESYE	EQ-----	133		
Cgl_phe/tyr_la	33	TISRDAKQQPTWDRAQAENVRKILESVPPIV	VAPEVLELKQKLADVANG-KAF-LLQGGDCA-ETFE	SNTEPHIRANVKTL	110		
Eco_tyr_la	30	SLQQEAQ-----IADSRKSI	-----SDIIAGRDPRLLVVCGPCS	IHDPE	69		
Bme_AroG2	134	-----VATVAESVK-----AKGLKLLR--GGAYKPRTSPYDFQGLGLE	-----GLKILK		175		
Cgl_phe/tyr_la	111	LQMAVVLTYGASTP-----VIKMARIAGQYAKPRSSDLGNGLPNY	-----RGDIVNGVEATPEARRHDP	A	171		
Eco_tyr_la	70	-----ALEYARRFKALAAEVSDSLYLVMR--VYFEKPRTTV-GWKGL	INDPHMDGSFDEAGLQIAR		128		
Bme_AroG2	172	RMIRAYANASAMNLRALTSSGTADLYRLSEWNREFVANSPAGARYEALARE	IDSGLRFMEACGVSD	ESLRAADIYCSHE	252		
Bme_AroG2	176	RIADEY--DLAVISEITNPADIEAVQYID-VIQIGARNMQN-FDLLKAA	GAVNKPVLLKRGLAATIEEF	INAAE--	246		
Cgl_phe/tyr_la	253	ALLVDY--ERSMLRLATDEEGNEELYDL	SAHQLWIGERTGMDDFHVNFASMI	SNPIGIGKIGPGITPEE	AVAYAD--	325	
Eco_tyr_la	129	KLLLELVNMGLPLATEALDPNSPQYL	GDLFS-WSAIGARTTES-QTHREMASGL	SMPVGFKN	GTGSLATAINAMRAAAQP	207	
Bme_AroG2	247	-YI-MAHGNQIILCERG	-----IRTYERATRNTLDISAVP-ILKQETH-LPVFVDVTHSTGRR--	DLLL-P	306		
Cgl_phe/tyr_la	326	-KLDPNFEPGRLTIVARM	SHDKVRSVLPG-----VIQAVEASGHKVIWQSDPMHGN	TFTASNGYKT--	385		
Eco_tyr_la	208	HRFVGINQAGQVALLQTO	GNDPGHVILRGKAPNYS	P-----ADVAQCEKEME	QAGLRPSLMVDCSHGNSNK--	DYRRQP	280
Bme_AroG2	307	A-AKAALAIGADG-----VMAEVHPDPA	-----VALSDSAQQ--MDLKQFDE	FYREIKKSVLARV--		358	
Cgl_phe/tyr_la	386	RHFDKVIDEVQG---FFE	VHRLGTHPGGIHIEFTGEDVTECLGGAEDITD	VDLPGRYESA---CDPR	LNTQQSLELAF	458	
Eco_tyr_la	281	AAESVVAQIKDGNRSIIGLMI	ESNIHEGNQSSE---QPRSE-MKYGVSVTD	ACISWEMTDALLREIHQD	LNGQLTARVA-	356	
Bme_AroG2	459	LVAEMLRN				466	

Figure 16. Multiple sequence alignment of MGA3 AroG2 with Type Ia DAHP synthases from *C. glutamicum* (regulated by phenylalanine and tyrosine) and *E. coli* (regulated by tyrosine). Dark purple highlights show 100% identity across all three sequences. Light purple highlights show 100% identity across two sequences. Unhighlighted residues are not conserved in these sequences. Teal boxes encompass residues which are a part of the catalytic core. Yellow boxes encompass residues which are part of the CM domain. Blue boxes encompass residues which are a part of the Type Ia regulatory domains.

PLASMID CONSTRUCTION

Affinity-tags fused to a protein of interest allows for easy purification via an affinity column, but this technique typically introduces non-native amino acids. Since we were interested in engineering the shikimate pathway *in vivo*, it was desirable to study the properties of the native enzymes without affinity-tags and linker regions. The tobacco etch virus (TEV) protease is a popular tool for removing affinity tags post purification. However, TEV protease leaves behind one non-native amino acid after site-specific cleavage. The TEV protease recognition site typically includes glycine or serine in the P1' position (Figure 17a), but 2002 study by David Waugh and colleagues shows that the TEV protease can cleave modified TEV protease recognition sites with a variety of substitute amino acids in the P1' position.¹⁷ As a consequence, if the first naturally occurring amino acid of a protein is substituted into the P1' position, then only the native protein will remain after TEV protease digestion. Methionine is the first amino acid in both AroG1 and AroG2. In the study mentioned above, TEV protease cleaves between 30 - 40% of the methionine variant in 3.5 h, only 10 - 20% less than when the wild-type glycine is used in the P1' position. In order to purify large amounts of native AroG1 and AroG2 with ease, a hexa-histidine (His₆) tag and a modified TEV protease recognition site (Figure 17b) were added to the N-terminus of both proteins.

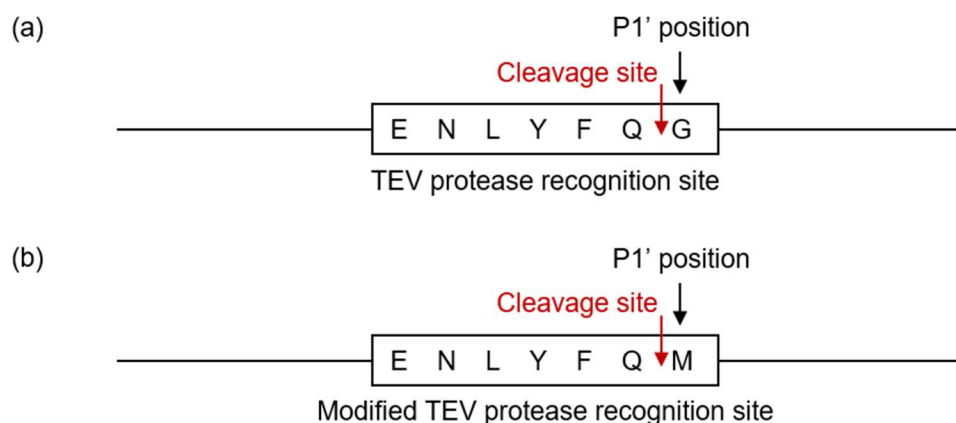


Figure 17. Comparison of the traditional and the modified TEV protease recognition sites. (a) The TEV protease recognition site. Traditionally, S or G are used in the P1' position. (b) The modified TEV protease recognition site used in this study.

The sequence analysis highlighted putative function of AroG1 and AroG2. Plasmids were constructed to express genes encoding AroG1 and AroG2 with a modified His₆-tag containing methionine in the P1' position. The *aroG1* and *aroG2* genes from *B. methanolicus* were codon optimized for expression in *E. coli*, synthesized, and cloned into the pUC-IDTamp vector by Integrated DNA Technologies, Inc. The codon-optimized genes were then subcloned into pET-15b downstream of the T7 promoter via Gibson Assembly (Figure 18). The resulting plasmids were named pMG2.109 and pMG2.128 and contain *aroG1* and *aroG2*, respectively (Figure 19).



Figure 18. Representation of the *E. coli* codon-optimized DAHP synthase gene fused to a modified TEV recognition site (ENLYFQM) and a His₆ tag under the control of a T7 promoter.

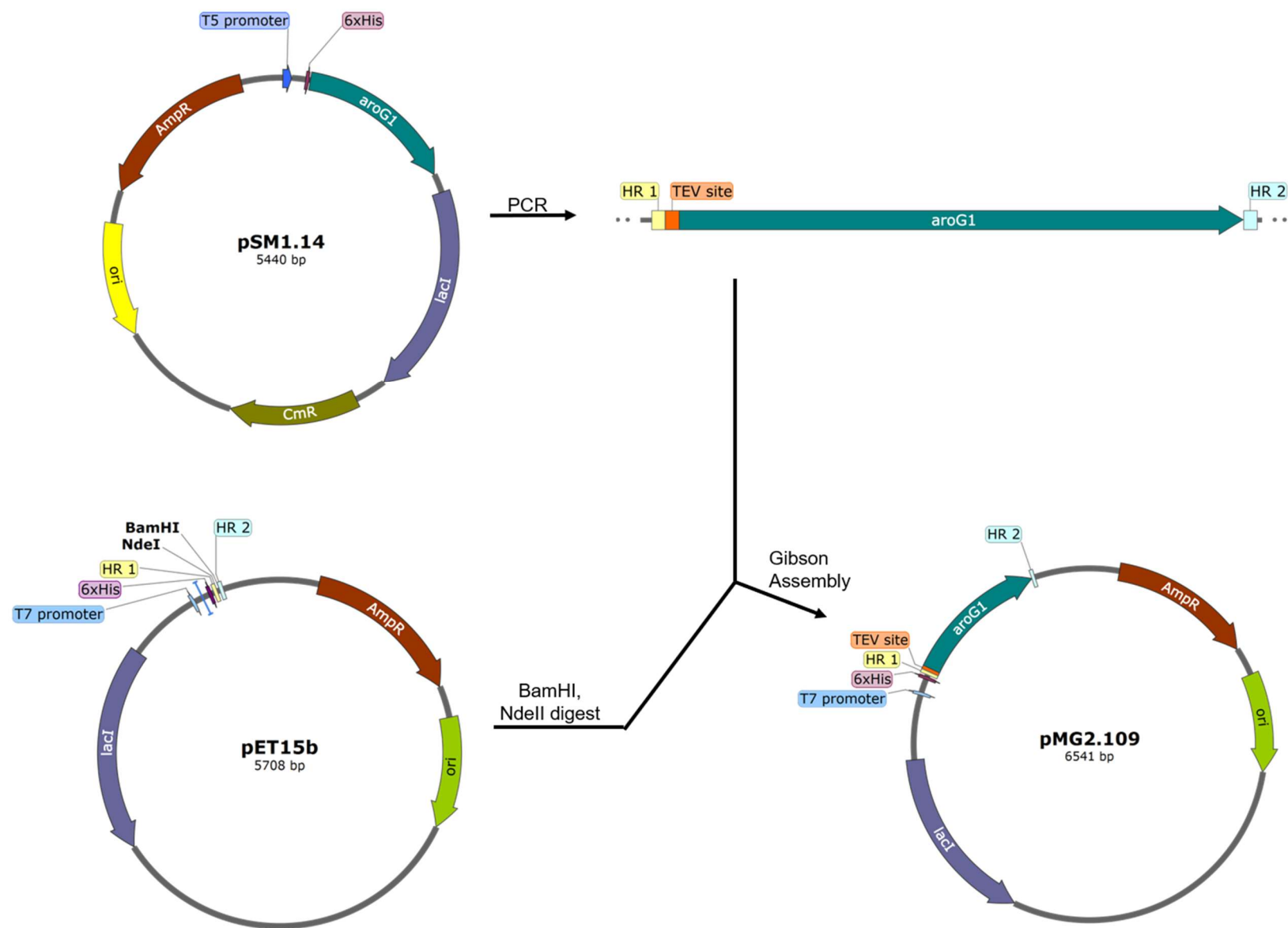


Figure 19. Construction of pMG2.109. The aroG1 gene was amplified from pSM1.14 by PCR. PCR primers added regions

Figure 19 (cont'd) (homologous to pET-15b on either end of the gene (HR1 and HR2). The vector pET-15b was linearized by digestion with restriction enzymes BamHI and NdeI. The linearized vector and the aroG1 fragment were ligated by Gibson assembly. pMG2.128 was constructed using the same method. Abbreviations: ampicillin resistance (AmpR), chloramphenicol resistance (CmR), lactose operator repressor (lacI), origin of replication (ori), hexa-histidine tag (6xHis), polymerase chain reaction (PCR).

OVEREXPRESSION AND PURIFICATION

Plasmids pMG2.109 and pMG2.128 were transformed into *E. coli* BL21(DE3). Expression was induced by addition of β -D-1-isopropylthiogalactopyranoside (IPTG) at a final concentration of 1.0 mM. Overexpression of the DAHP synthases was allowed to proceed overnight at 30 °C. Cells pellets were collected by centrifugation and subsequently lysed by French Press. Cell lysates were clarified by centrifugation. The His₆-DAHP synthases were isolated from the clarified cell lysates via a nickel affinity column. A desalting column was used to remove the imidazole and exchange the protein into a buffer A His₇-tagged TEV protease was used to remove the His₆-tag from AroG1 and AroG2 to facilitate easier separation of the protease from the native DAHP synthases. After five hours of incubation with TEV protease at 30 °C, the DAHP synthases were re-loaded onto the nickel affinity column. This time, the His₇-TEV protease and cleaved His₆ tag remained bound to the column while the native DAHP synthases eluted in the flow through.

A large portion of His₆-AroG1 and His₆-AroG2 were detected in the insoluble fraction of cell lysates which indicated the formation of inclusion bodies, aggregates of improperly folded proteins.¹⁸ Luckily, extended induction time allowed proper folding of sufficient amounts of DAHP synthases which could be recovered in the soluble lysate. An induction time of 5-6 hours resulted in little to no DAHP synthase recovery while overnight induction allowed for recovery of up to 52 mg of His₆-tagged DAHP synthase from a 1 L culture. Decreasing the temperature at which induced cells are cultured can help slow down protein folding and thus combat formation of inclusion bodies.¹⁹ To test whether this would improve His₆-AroG1 and His₆-AroG2 solubility, protein expression was induced either at 19 °C or 30 °C. No difference in protein yield was observed, however. Consequently, protein expression was induced at 30 °C. Since sufficient yields of His₆-AroG1 and His₆-AroG2 were obtained, purification was conducted without further overexpression modifications (Figure 20).

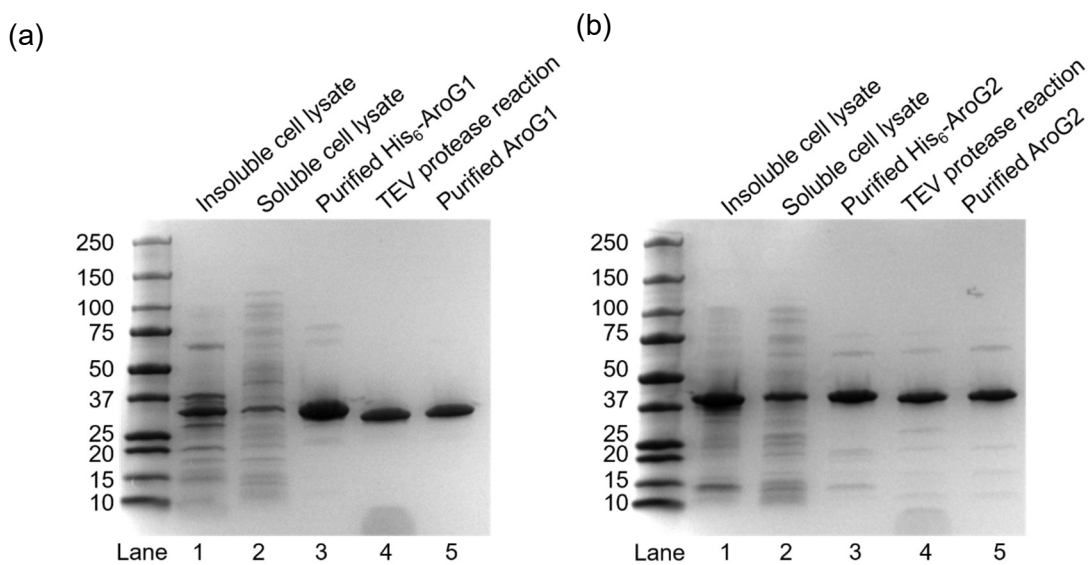


Figure 20. SDS-PAGE analysis of AroG1 and AroG2 purification. (a) SDS-PAGE of the purification of AroG1 (30 kDa). Lane 1: the insoluble lysate of BL21(DE3)/pMG2.109. Lane 2: the soluble lysate of BL21(DE3)/pMG2.109. Lane 3: the purified His₆-AroG1. Lane 4: the TEV protease reaction after 5 hours of incubation with His₆-AroG1. Lane 5: the purified AroG1. (b) SDS-PAGE of the purification of AroG2 (40 kDa). Lane 1: the insoluble lysate of BL21(DE3)/pMG2.128. Lane 2: the soluble lysate of BL21(DE3)/pMG2.128. Lane 3: the purified His₆-AroG2. Lane 4: the TEV protease reaction after 5 hours of incubation with His₆-AroG2. Lane 5: the purified AroG2.

MOLECULAR WEIGHT

The predicted masses of the DAHP synthase subunits were calculated using Protein Calculator v3.4.²⁰ AroG1 and AroG2 subunits were calculated to be 30,133 Da and 39,968 Da, respectively. Electron spray time-of-flight mass spectrometry was used to measure the subunits' masses (Figures 21).²¹ As each protein attains multiple charge states at ionization, the resulting mass spectrum is much more complex than those obtained for small molecules. MaxEnt mass deconvolution software collectively analyzed mass-to-charge ratios to determine the zero-charge spectra, and the measured masses of the AroG1 and AroG2 subunits (Figures 22) were calculated to be $30,165 \pm 10$ Da and $39,970 \pm 10$ Da, respectively (Table 4).²²

Table 4. The predicted and calculated masses of AroG1 and AroG2 subunits.

	AroG1	AroG2
Predicted Mass (Da)	30,133	39,968
Calculated Mass (Da)	$30,165 \pm 10$	$39,970 \pm 10$

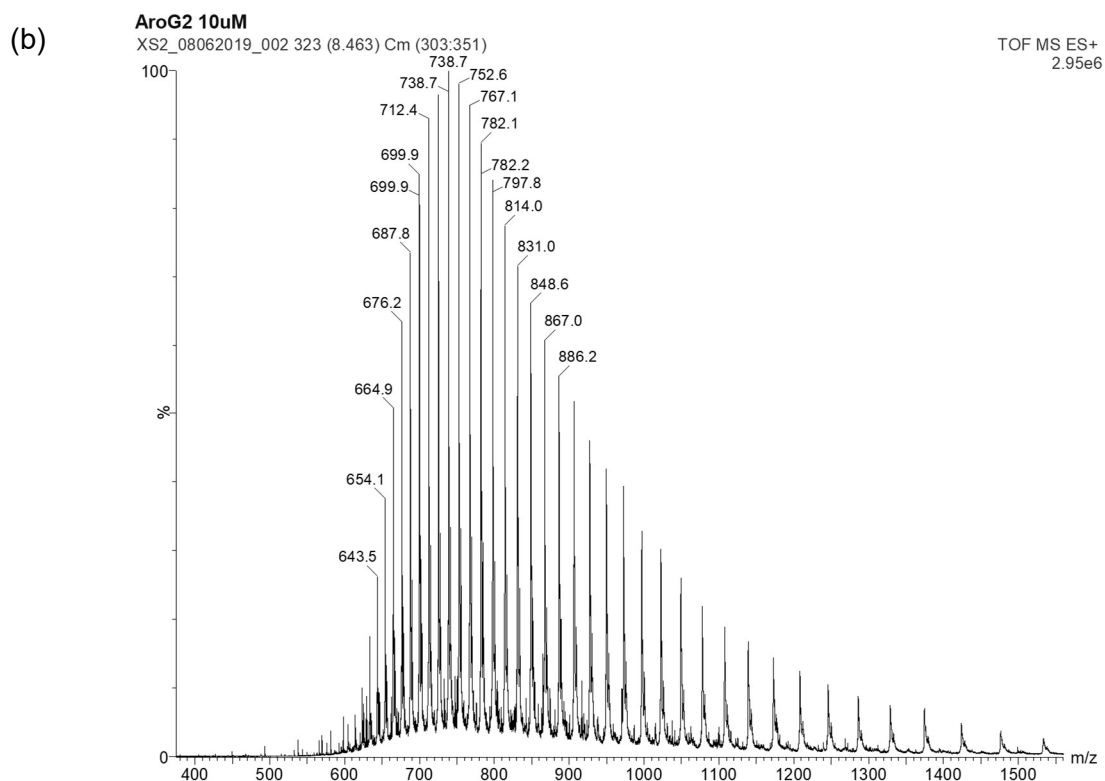
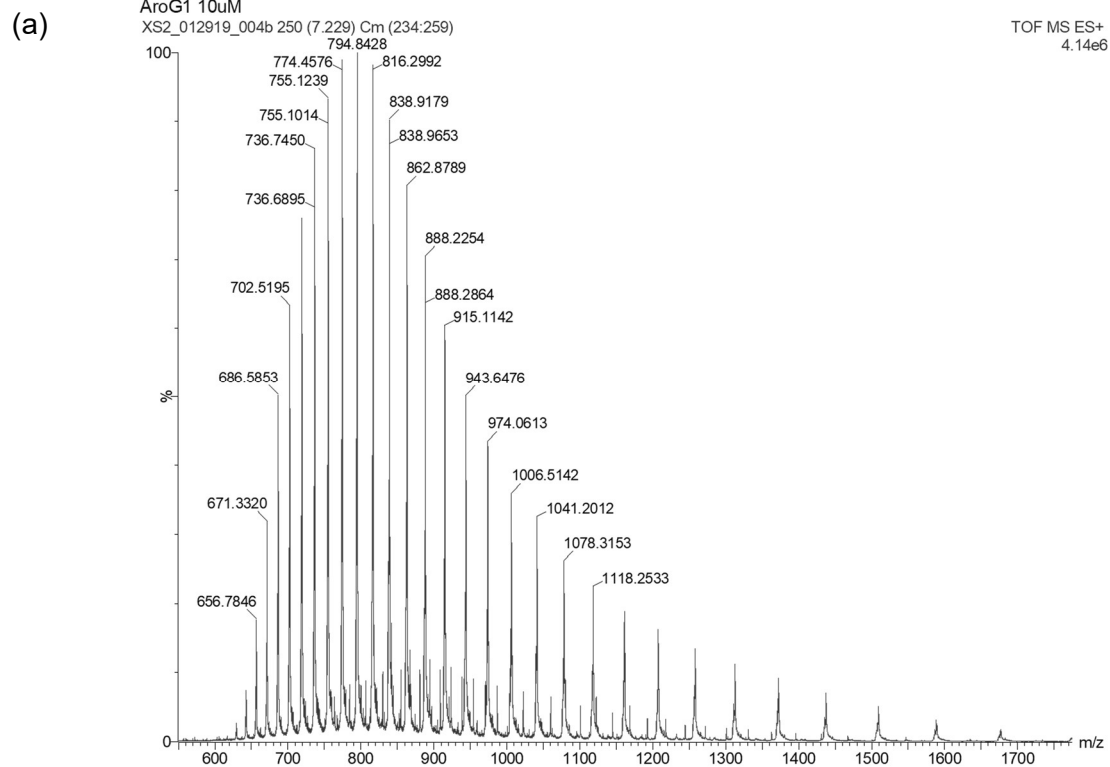


Figure 21. Results from mass spectral analysis of AroG1 and AroG2. (a) Mass spectrum of AroG1. (b) Mass spectrum of AroG2.

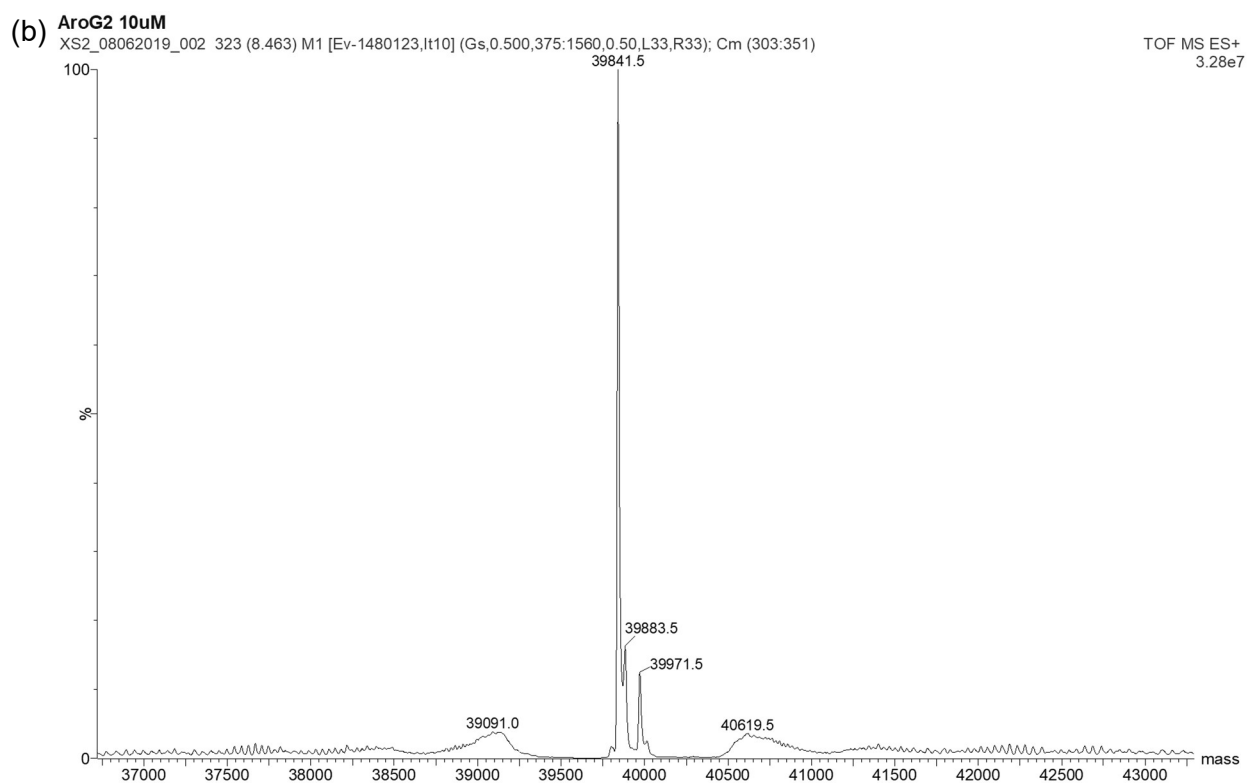
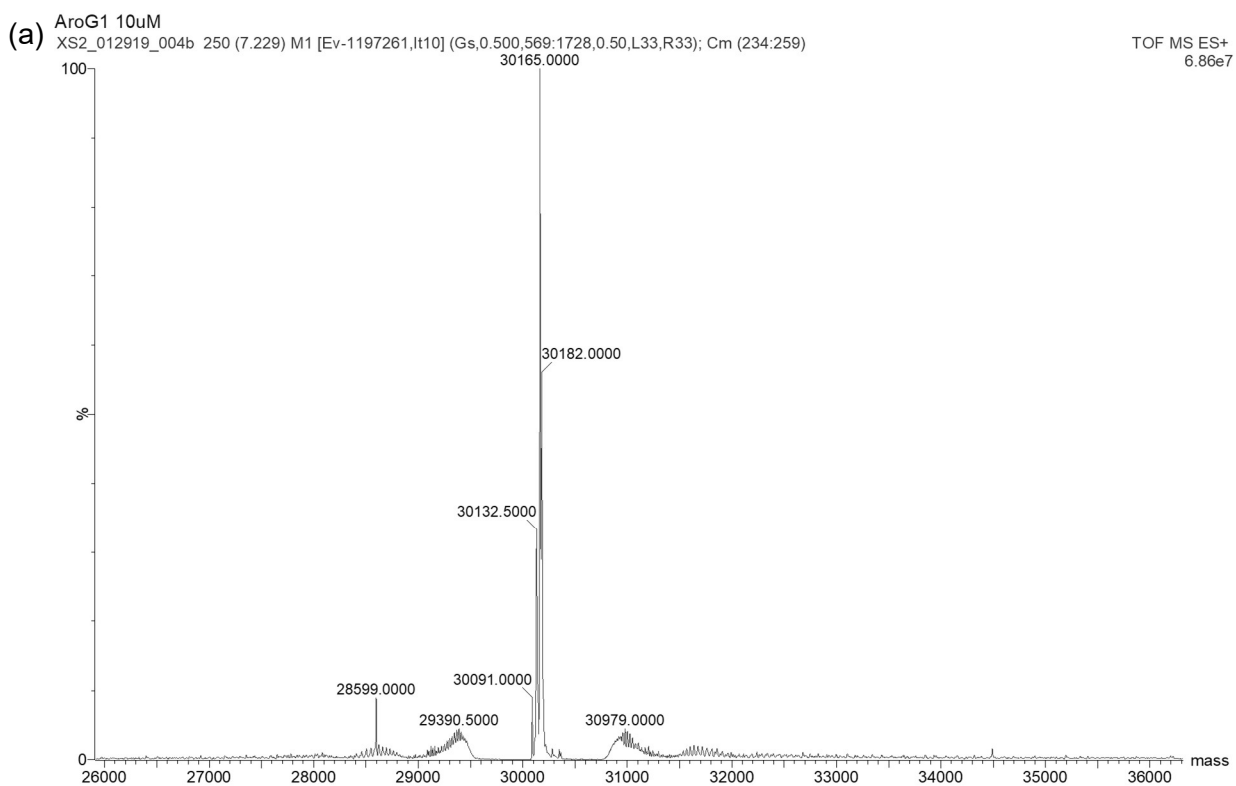


Figure 22. Results from mass deconvolution of AroG1 and AroG2 mass spectra. (a) Mass deconvolution of AroG1. (b) Mass deconvolution of AroG2.

ENZYMATIC ASSAYS

Effect of pH – AroG1 exhibited the highest specific activity ($3.0 \mu\text{mol}/\text{min}\cdot\text{mg}$) at pH 6.9 (Figure 23a). AroG2 exhibited the highest specific activity ($0.52 \mu\text{mol}/\text{min}\cdot\text{mg}$) at pH 7.3 (Figure 23b). Both AroG1 and AroG2 activity decreased in more basic conditions.

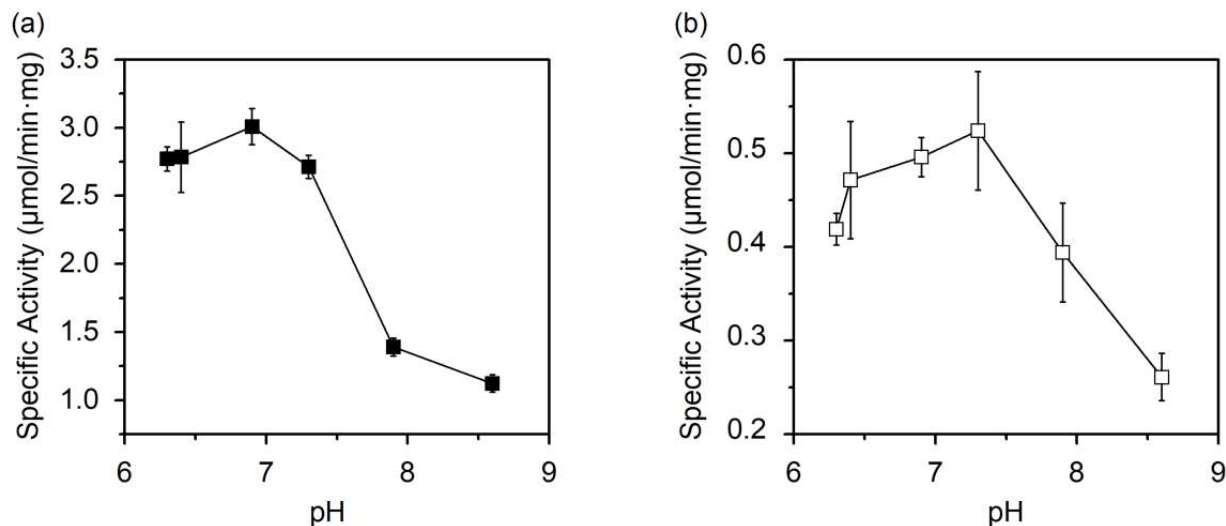


Figure 23. The effect of pH on AroG1 and AroG2 activity. (a) The effect of pH on AroG1 activity. AroG1 exhibits maximum activity at pH 6.9. (b) The effect of pH on AroG2 activity. AroG2 exhibits maximum activity at pH 7.3.

Effect of temperature – Both AroG1 and AroG2 activity increased at higher temperatures. AroG1 reached maximum activity ($8.1 \mu\text{mol}/\text{min}\cdot\text{mg}$) at 50°C . Of the temperatures tested, AroG2 displayed highest activity ($2.9 \mu\text{mol}/\text{min}\cdot\text{mg}$) at 60°C (Figure 24). Due to instrument constraints, higher temperatures could not be tested.

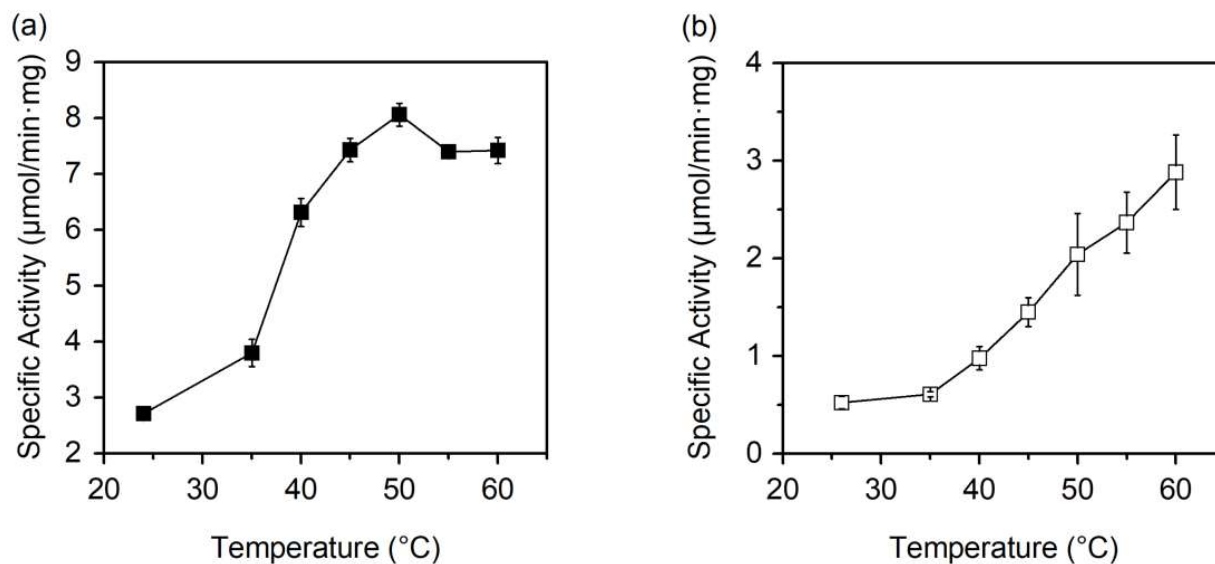


Figure 24. The effect of temperature on AroG1 and AroG2 activity. (a) The effect of temperature on AroG1 activity. AroG1 exhibits maximum activity at 50 $^{\circ}\text{C}$. (b) The effect of temperature on AroG2 activity. AroG2 exhibits maximum activity at 60 $^{\circ}\text{C}$.

Thermostability –AroG1 lost activity much quicker than AroG2. AroG1 lost $96 \pm 1.3\%$ activity after only 10 minutes at 50 $^{\circ}\text{C}$. AroG2 lost only $30 \pm 11\%$ activity after 1 hour. After 5 hours, AroG1 lost all activity while AroG2 still retained $17 \pm 8\%$ of its original activity (Figure 25).

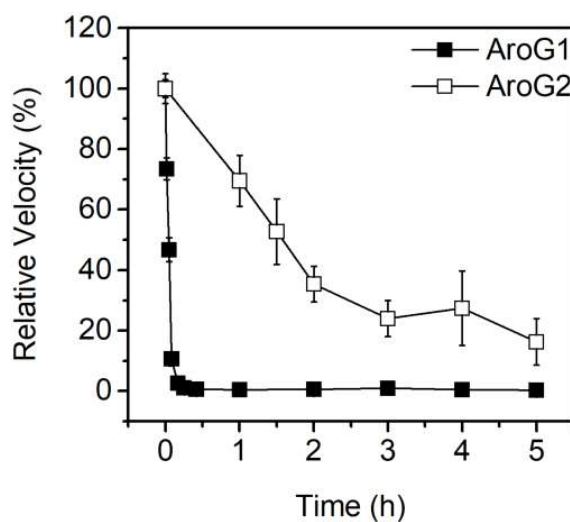


Figure 25. The thermostability of AroG1 and AroG2.

Kinetics – AroG1 displayed a K_m for PEP of $1124 \pm 132 \mu\text{M}$ and a K_m for E4P of $525 \pm 96 \mu\text{M}$. AroG2 displayed a K_m for PEP of $92 \pm 23 \mu\text{M}$ and a K_m for E4P of $127 \pm 37 \mu\text{M}$ (Table 5, Figure 26). AroG1 had a higher $V_{\max}$ than AroG2 ($10 \pm 0.73 \mu\text{M/s}$ and $1.6 \mu\text{M/s}$, respectively), but both enzymes bound to E4P stronger than PEP.

Table 5. Kinetic values for AroG1 and AroG2.

	AroG1	AroG2
$V_{\max} (\mu\text{M/s})$	10 ± 0.73	1.6 ± 0.12
$K_m^{E4P} (\mu\text{M})$	525 ± 96	127 ± 37
$K_m^{PEP} (\mu\text{M})$	1124 ± 132	92 ± 23
$k_{\text{cat}} (\text{s}^{-1})$	13 ± 1	2.1 ± 0.2

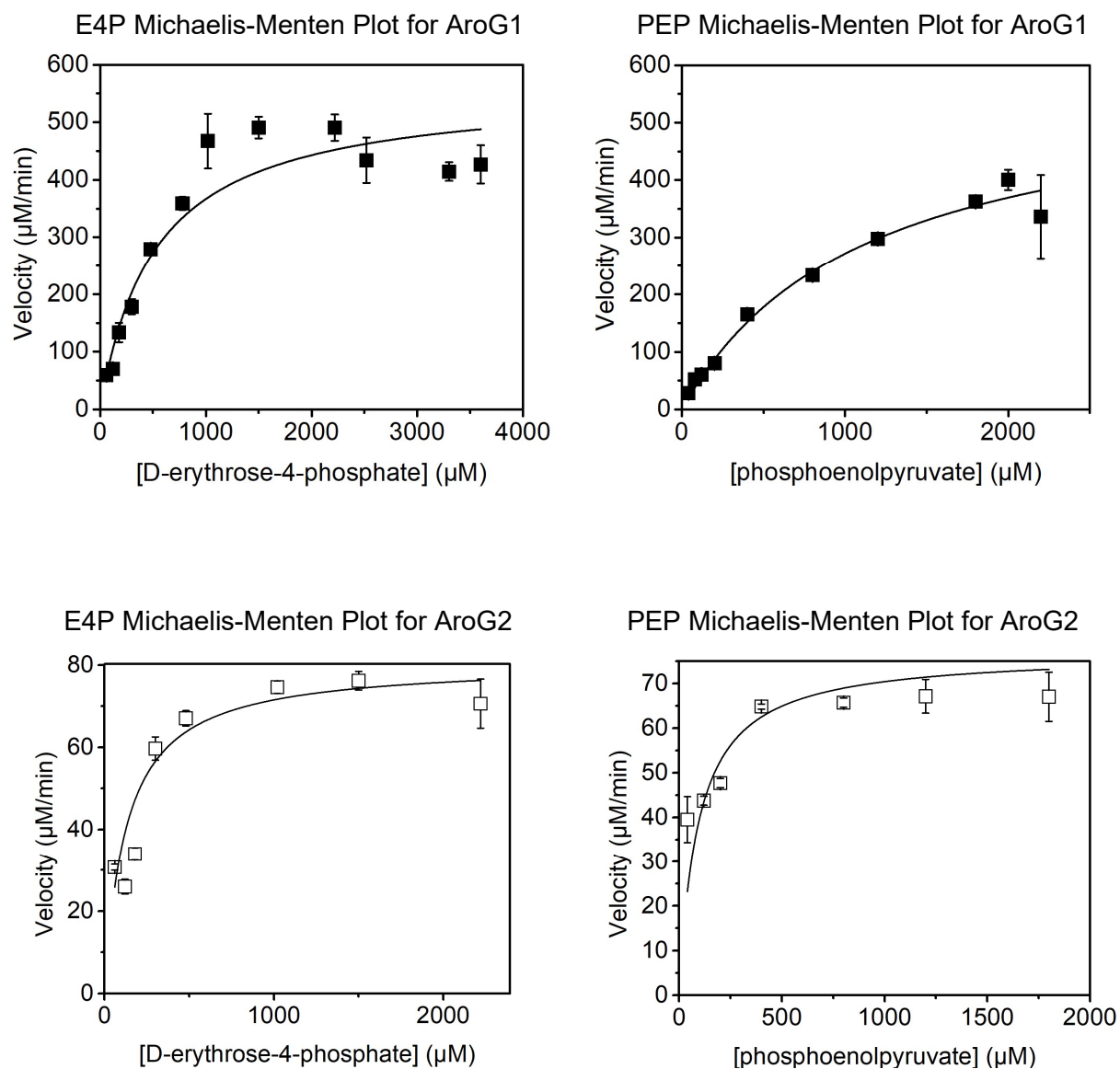


Figure 26. The Michaelis-Menten plots for AroG1 and AroG2.

Feedback inhibition – Based on the amino acid sequences of the enzymes, AroG1 was predicted to be unregulated, while AroG2 was predicted to be regulated by chorismate and prephenate. At the time of this writing, the only known inhibitors of DAHP synthases are phenylalanine, tyrosine, tryptophan, chorismate, and/or prephenate. The activity of AroG1 and AroG2 were compared to their activity in the absence of potential inhibitors (Figure 27). As predicted, AroG1 was not inhibited by any of the tested compounds, and AroG2 was inhibited by chorismate and prephenate. In the presence of 0.25 mM chorismate, AroG2 activity decreased by $81 \pm 3\%$. In the presence of 0.25 mM prephenate, AroG2 activity decreased by $72 \pm 8\%$.

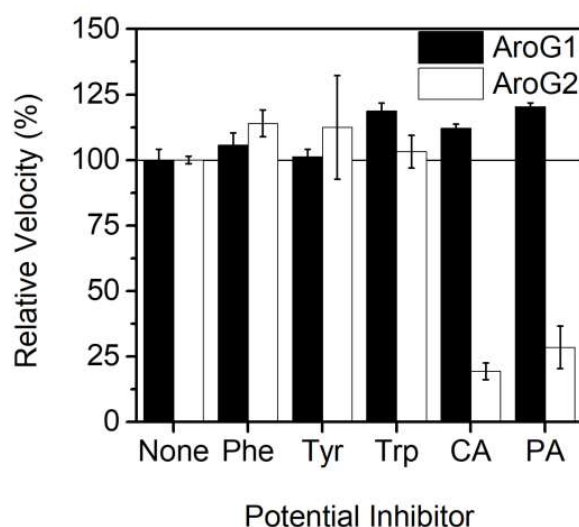


Figure 27. The effect of potential inhibitors on AroG1 and AroG2 activity.

DISCUSSION

Allosteric regulation of Type I β DAHP synthases is likely to have been adopted multiple times over the course of evolutionary history.¹ The adoption of fused regulatory domains is expected to occur through gene-fusion events.²³ This idea is supported by many experimental studies.^{1,24-26,35} Regulated Type I β DAHP synthases that are truncated so only their catalytic core

remains yield an active enzyme that is unaffected by its previous allosteric inhibitor.^{1,24,25} Allosteric regulation is also easily transferred from one Type I β DAHP synthase to another by swapping regulatory domains.³⁵ This domain swapping requires no modifications to the amino acid sequences to effectively transfer allosteric regulation. In addition, allosteric regulation can be provided to unregulated Type I β DAHP synthases by fusing a regulatory domain to its N-terminus.²⁶ The ease in which allostery can be transferred between Type I β DAHP synthases is evidence that these regulatory domains were obtained through gene-fusion events.

Both DAHP synthase isozymes present in *B. methanolicus* MGA3 are Type I β , with AroG1 being unregulated and AroG2 being regulated via a CM domain. While no less than twenty bacteria have been studied for their DAHP synthase activity, MGA3 presents the first that contains two Type I β DAHP synthases. Only six of these twenty bacteria contain multiple DAHP synthase isozymes. Usually when bacteria contain multiple DAHP synthases, they are either isozymes of Type I α or a mix of Type I α and Type II. All previously characterized bacteria containing a Type I β DAHP synthase do not have multiple isozymes. Bacteria with multiple DAHP synthase isozymes include *E. coli*^{27,28}, *P. aeruginosa*⁵, *Corynebacterium glutamicum*^{29,30}, *Ralstonia solanacearum*³¹, *Amycolatopsis methanolica*,³⁸ and *Stigmatella aurantiaca*³² (Table 6). Both *E. coli* and *R. solanacearum* contain isozymes of Type I α DAHP synthases. *C. glutamicum*, *P. aeruginosa*, *A. methanolica*, and *S. aurantiaca* all have both Type I α and Type II DAHP synthases. (The *R. solanacearum* and *A. methanolica* DAHP synthases have not been categorized into a type by the authors of the referenced papers; I took the liberty of performing a sequence alignment to determine their most likely classification). To the best of my knowledge, MGA3 is the first organism characterized that contains two Type I β DAHP synthases.

Table 6. All known bacteria with multiple experimentally studied DAHP synthases and the quantity of each type the organism possesses.

Organism	Type I α	Unregulated Type I β	Regulated Type I β	Type II
<i>Bacillus methanolicus</i>	----	1	1	----
<i>Escherichia coli</i>	3	----	----	----
<i>Ralstonia solanacearum</i>	2	----	----	----
<i>Corynebacterium glutamicum</i>	1	----	----	1
<i>Pseudomonas aeruginosa</i>	2	----	----	2
<i>Amycolatopsis methanolica</i>	1	----	----	1
<i>Stigmatella aurantiaca</i>	1	----	----	1

B. subtilis is the only other bacterium in the *Bacillus* genus whose DAHP synthase has been studied experimentally. This organism only possesses one Type I β DAHP with a CM domain on the N-terminus which puts it in the same class as AroG2 (regulated Type I β). It is most common for Type I β to be regulated via a CM domain, with only one bacterium, *T. maritima*, displaying a DAHP synthase regulated via an ACT-like domain (Table 7).

Table 7. All known bacteria containing experimentally studied Type I β DAHP synthases and the quantity of each Type I β the organism possesses.

Organism	Unregulated Type I β	Regulated Type I β (CM domain)	Regulated Type I β (ACT-like domain)
<i>Bacillus methanolicus</i>	1	1	-----
<i>Pyrococcus furiosus</i>	1	-----	-----
<i>Aeropyrum pernix</i>	1	-----	-----
<i>Bacillus subtilis</i>	-----	1	-----
<i>Geobacillus sp.</i>	-----	1	-----
<i>Listeria monocytogenes</i>	-----	1	-----
<i>Porphyromonas gingivalis</i>	-----	1	-----
<i>Prevotella nigrescens</i>	-----	1	-----
<i>Thermotoga maritima</i>	-----	-----	1

Since only one other *Bacillus* species has been examined for DAHP synthase activity, it would be advantageous to know what types of DAHP synthases other *Bacillus* species are likely to possess. A BLASTn search of the *B. methanolicus* MGA3 16s rRNA was used to find the closest relatives to MGA3. The top 10 bacteria whose genomes had been completely sequenced were subjected to a BLASTp search using the amino acid sequences of AroG1 and AroG2 as the query. The results showed that even the 10 closest known relatives of MGA3 contain only one putative Type I β DAHP synthase, each containing an amino acid sequence homologous to a CM domain on the N-terminus (Table 8). These findings highlight the uniqueness of *B. methanolicus* possessing two Type I β DAHP synthases.

Table 8. The top 10 closest relatives to *B. methanolicus* MGA3 based on 16s rRNA sequence identity and their putative DAHP synthases.

Organism	16s rRNA % Identity	Number of Putative DAHP Synthases	Type of Putative DAHP Synthase
<i>Bacillus ciccensis</i>	96.82	1	Type I β (CM domain)
<i>Bacillus</i> sp. FJAT-25496	96.76	1	Type I β (CM domain)
<i>Bacillus</i> sp. 1NLA3E	96.76	1	Type I β (CM domain)
<i>Bacillus foraminis</i>	96.44	1	Type I β (CM domain)
<i>Bacillus</i> sp.S3	96.44	1	Type I β (CM domain)
<i>Bacillus marisflavi</i>	96.30	1	Type I β (CM domain)
<i>Bacillus jeotgali</i>	96.31	1	Type I β (CM domain)
<i>Bacillus oceanisediminis</i> 2691	96.18	1	Type I β (CM domain)
<i>Bacillus mesonae</i>	96.12	1	Type I β (CM domain)
<i>Bacillus asahii</i>	95.92	1	Type I β (CM domain)

The greater thermostability of AroG2 relative to AroG1 may be explained by its CM domain which may promote dimerization and increase stability. AroG2-like DAHP synthases are homotetramers containing CM domains which have been shown to dimerize.^{25,33-35} This dimerization of CM domains may provide more stability to the protein.

AroG1 has a higher maximum velocity and a lower turnover number than AroG2. Unregulated Type I β DAHP synthases have been shown in the literature to have higher activity than their regulated counterparts. The difference in activity is likely related to the CM domain present on AroG2.^{25,35} A study by Emily J. Parker and colleagues from 2016 shows that removing the CM domain from a *Geobacillus* sp. DAHP synthase increases the turnover rate (k_{cat}) from $45 \pm 4 \text{ s}^{-1}$ to $63 \pm 5 \text{ s}^{-1}$.³⁵ Additionally, when the ACT-like regulatory domain on the *T. maritima* DAHP synthase was replaced with a CM domain, the k_{cat} value decreased from $14 \pm 0.3 \text{ s}^{-1}$ to 4.8 ± 0.1

s⁻¹.²⁵ Taken together these studies suggest that the physical barrier of the CM domain may be the reason for the lower activity of AroG2 when compared to AroG1.

The tertiary and quaternary structures of AroG1 and AroG2 were modeled using SWISS-MODEL.³⁶ The structure predictions of AroG1 and AroG2 exhibit how the position of the AroG2 CM domain may slow down its activity. Each subunit has one active site, two on each face of the assembled enzyme. The two active sites on the forward-presenting face of AroG1 are circled (Figure 27a). On the AroG2 model the CM domains hide the active sites of AroG2 (Figure 27b). This physical barrier may be the reason for the slower activity of AroG2 when compared to AroG1.

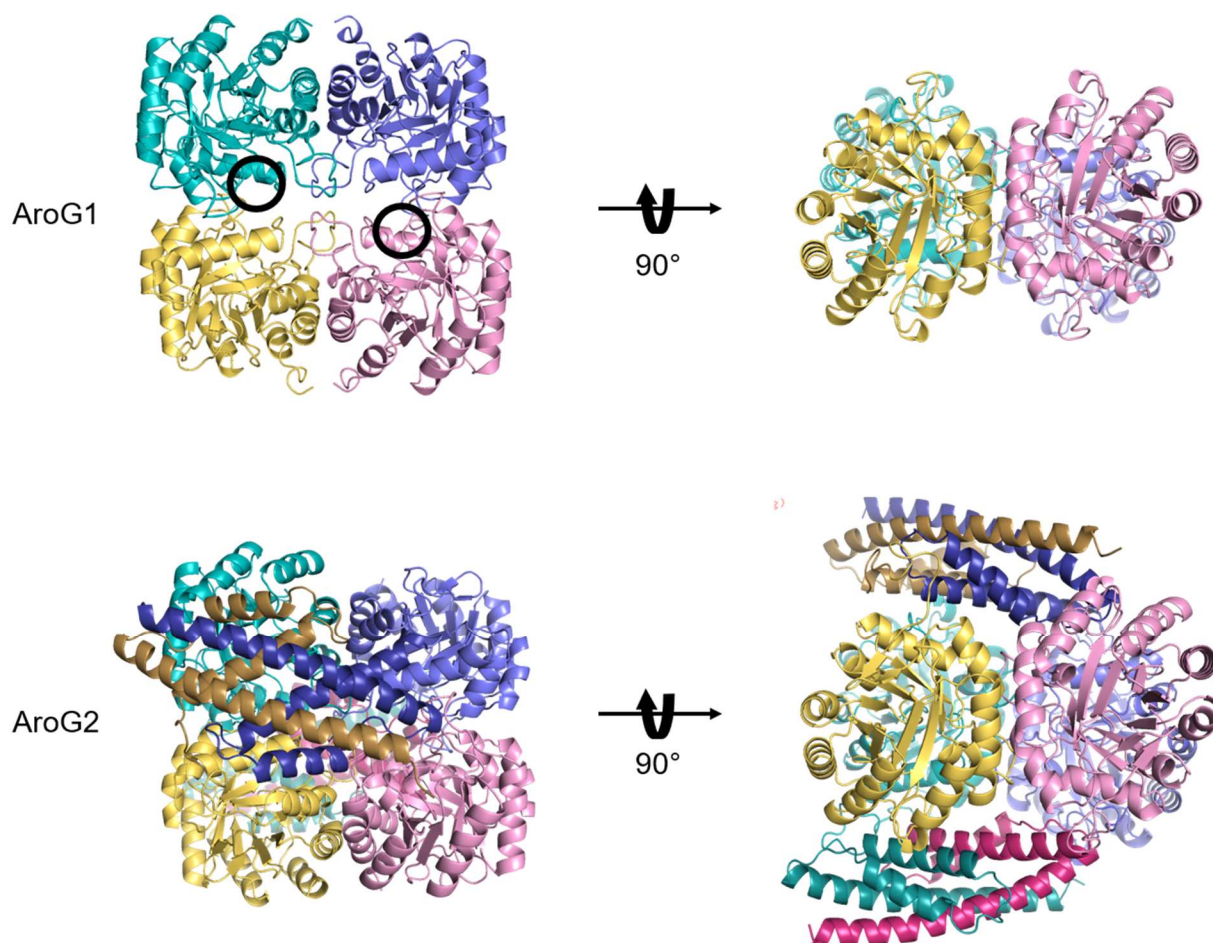


Figure 28. Model structures of AroG1 and AroG2. Each subunit is represented by a different color. The CM domain is a darker shade of the same color as the catalytic core it is fused to. The circle represent the two active sites on the front-facing plane of AroG1.

B. methanolicus MGA3 is a thermophile. Five DAHP synthases from thermophiles have been studied experimentally, including DAHP synthases from a *Geobacillus* sp.²⁵, *Aeropyrum pernix*¹², *Pyrococcus furiosus*¹¹, *Thermotoga maritima*¹⁴, and *Helicobacter pylori*.³⁷ Of these, there is no clear trend on which substrate the thermophilic DAHP synthases bind stronger. K_m^{PEP} values range from 3 to 890 μ M. The K_m^{PEP} value for AroG2 (92 ± 23 μ M) falls well within this range, however the K_m^{PEP} value for AroG1 is much higher (1124 ± 132 μ M). The thermophilic DAHP synthase K_m^{E4P} values range from 6 to 280 μ M. Again, the K_m^{E4P} value for AroG2 (264 ± 61 μ M) is within this range while the value for AroG1 is significantly higher (525 ± 96 μ M). The higher binding affinity of substrates by AroG2 may be explained again by the CM domain. Both PEP and E4P must be present in the active site for the enzyme to perform its function. The CM domain may assist in retention of one substrate while the other is still making its way to the active site. Since AroG1 does not have the CM domain, it may be more difficult for the enzyme to bind both substrates at once. The turnover number (k_{cat}) for thermophilic DAHP synthases ranges from 1 to 45 s⁻¹. Both AroG1 and AroG2 have k_{cat} values well within this range (Table 9).

Table 9. Michealis-Menten constants and turnover numbers for DAHP synthases from thermophilic bacteria.

Type	Organism	K_m^{PEP}	K_m^{E4P}	k_{cat}	Allosteric Inhibitor(s)	Reference
		μM	μM	s^{-1}		
Type I β	<i>B. methanolicus</i> (AroG1)	1124 $\pm$ 132	525 $\pm$ 96	13 $\pm$ 1.0	None	This work
	<i>B. methanolicus</i> (AroG2)	92 $\pm$ 23	264 $\pm$ 61	2.1 $\pm$ 0.2	CA; PA	This work
	<i>Geobacillus</i> sp. Y412MC61	87 $\pm$ 9	95 $\pm$ 8	45 $\pm$ 4.0	CA; PA	25
	<i>A. pernix</i>	890 $\pm$ 30	280 $\pm$ 20	1.0 $\pm$ 0.2	None	12
	<i>P. furiosus</i>	120 $\pm$ 20	28 $\pm$ 4	1.5 $\pm$ 0.1	None	11
Type I α	<i>T. maritima</i>	13 $\pm$ 2.6	57.3 $\pm$ 4.3	2.3 $\pm$ 0.2	Phe; Tyr	14
Type II	<i>H. pylori</i>	3.0 $\pm$ 1.0	6.0 $\pm$ 1.0	3.3 $\pm$ 0.3	None	37

The only two studied DAHP synthases from a methylotrophic bacterium were characterized in 1995 from the organism *Amycolatopsis methanolica*.³⁸ Based on a sequence alignment I preformed, *A. methanolica* appears to have one Type Ia and one Type II DAHP synthase, fitting with the notion that multiple Type Iβ isozymes in a single species is rare. Both *A. methanolica* DAHP synthases were found to be inhibited by one or more of the aromatic amino acids. The Michaelis-Menten constants were only measured for the Type II enzyme. The K_m values for PEP and E4P were $400 \pm 60 \mu\text{M}$ and $1100 \pm 200 \mu\text{M}$, respectively.

Neither AroG1 nor AroG2 are inhibited by the aromatic amino acids. When an *aroE* mutant of *B. methanolicus* MGA3 is created, it appears that the supplementation of aromatic amino acids in the culturing media will not disrupt the carbon flow into the shikimate pathway. This is good news in terms of genetically engineering the shikimate pathway, because no feedback resistant DAHP synthase must be introduced into the mutant. If the slow DAHP synthase activity creates a bottleneck, it may be necessary to introduce an additional copy of a DAHP synthase gene. Based on the results presented here, AroG2 would be the best candidate for duplication because of its greater thermostability. The activity of AroG2 may be improved by deletion of its CM domain which would remove the physical barrier covering the active site.

REFERENCES

REFERENCES

1. Wu, J.; Woodard, R. W. New Insights into the Evolutionary Links Relating to the 3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate Synthase Subfamilies. *J. Biol. Chem.* **2006**, *281* (7), 4042–4048.
2. Light, S. H.; Anderson, W. F. The Diversity of Allosteric Controls at the Gateway to Aromatic Amino Acid Biosynthesis. *Protein Sci.* **2013**, *22* (4), 395–404.
3. Subramaniam, P. S.; Xie, G.; Xia, T.; Jensen, R. A. Substrate Ambiguity of 3-Deoxy-D-Manno-Octulosonate 8-Phosphate Synthase from *Neisseria Gonorrhoeae* in the Context of Its Membership in a Protein Family Containing a Subset of 3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate Synthases. *J. Bacteriol.* **1998**, *180* (1), 119–127.
4. Birck, M. R.; Woodard, R. W. Aquifex Aeolicus 3-Deoxy-D-Manno-2-Octulosonic Acid 8-Phosphate Synthase: A New Class of KDO 8-P Synthase? *J. Mol. Evol.* **2001**, *52* (2), 205–214.
5. Sterritt, O. W.; Lang, E. J. M.; Kessans, S. A.; Ryan, T. M.; Demeler, B.; Jameson, G. B.; Parker, E. J. Structural and Functional Characterisation of the Entry Point to Pyocyanin Biosynthesis in *Pseudomonas Aeruginosa* Defines a New 3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate Synthase Subclass. *Biosci. Rep.* **2018**, *38* (5).
6. Walker, G. E.; Dunbar, B.; Hunter, I. S.; Nimmo, H. G.; Coggins, J. R. Evidence for a Novel Class of Microbial 3-Deoxy-D-Arabino-Heptulosonate-7-Phosphate Synthase in *Streptomyces coelicolor* A3(2), *Streptomyces Rimosus* and *Neurospora Crassa*. *Microbiology* **1996**, *142* (Pt 8), 1973–1982.
7. Schofield, L. R.; Anderson, B. F.; Patchett, M. L.; Norris, G. E.; Jameson, G. B.; Parker, E. J. Substrate Ambiguity and Crystal Structure of *Pyrococcus furiosus* 3-Deoxy-D-Arabino-Heptulosonate-7-Phosphate Synthase: An Ancestral 3-Deoxyald-2-Ulosonate-Phosphate Synthase? *Biochemistry* **2005**, *44* (36), 11950–11962.
8. Lang, E. J. M.; Cross, P. J.; Mittelstädt, G.; Jameson, G. B.; Parker, E. J. Allosteric ACTion: The Varied ACT Domains Regulating Enzymes of Amino-Acid Metabolism. *Curr. Opin. Struct. Biol.* **2014**, *29*, 102–111.
9. Jensen, R. A.; Xie, G.; Calhoun, D. H.; Bonner, C. A. The Correct Phylogenetic Relationship of KdsA (3-Deoxy-D-Manno-Octulosonate 8-Phosphate Synthase) with One of Two Independently Evolved Classes of AroA (3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate Synthase). *J. Mol. Evol.* **2002**, *54* (3), 416–423.

10. Notredame, C.; Higgins, D. G.; Heringa, J. T-Coffee: A Novel Method for Fast and Accurate Multiple Sequence Alignment. *J. Mol. Biol.* **2000**, *302* (1), 205–217.
11. Schofield, L. R.; Patchett, M. L.; Parker, E. J. Expression, Purification, and Characterization of 3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate Synthase from *Pyrococcus furiosus*. *Protein Expr. Purif.* **2004**, *34* (1), 17–27.
12. Zhou, L.; Wu, J.; Janakiraman, V.; Shumilin, I. A.; Bauerle, R.; Kretsinger, R. H.; Woodard, R. W. Structure and Characterization of the 3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate Synthase from *Aeropyrum pernix*. *Bioorg. Chem.* **2012**, *40* (1), 79–86.
13. Wu, J.; Sheflyan, G. Y.; Woodard, R. W. *Bacillus subtilis* 3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate Synthase Revisited: Resolution of Two Long-Standing Enigmas. *Biochem. J* **2005**, *390* (Pt 2), 583–590.
14. Wu, J.; Howe, D. L.; Woodard, R. W. *Thermotoga maritima* 3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate (DAHP) Synthase: The Ancestral Eubacterial DAHP Synthase? *J. Biol. Chem.* **2003**, *278* (30), 27525–27531.
15. Cui, D.; Deng, A.; Bai, H.; Yang, Z.; Liang, Y.; Liu, Z.; Qiu, Q.; Wang, L.; Liu, S.; Zhang, Y.; et al. Molecular Basis for Feedback Inhibition of Tyrosine-Regulated 3-Deoxy-D-Arabino-Heptulosonate-7-Phosphate Synthase from *Escherichia coli*. *J. Struct. Biol.* **2019**, *206* (3), 322–334.
16. Cross, P. J.; Pietersma, A. L.; Allison, T. M.; Wilson-Coutts, S. M.; Cochrane, F. C.; Parker, E. J. *Neisseria meningitidis* Expresses a Single 3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate Synthase That Is Inhibited Primarily by Phenylalanine. *Protein Sci.* **2013**, *22* (8), 1087–1099.
17. Kapust, R. B.; Tözsér, J.; Copeland, T. D.; Waugh, D. S. The P1' Specificity of Tobacco Etch Virus Protease. *Biochem. Biophys. Res. Commun.* **2002**, *294* (5), 949–955.
18. Carrió, M. M.; Villaverde, A. Construction and Deconstruction of Bacterial Inclusion Bodies. *J. Biotechnol.* **2002**, *96* (1), 3–12.
19. Schein, C. H. Production of Soluble Recombinant Proteins in Bacteria. *Biotechnology* **1989**, *7* (11), 1141–1149.
20. Putnam C. Protein calculator v 3.4. The Scripps Research Institute; USA. available at <http://protcalc.sourceforge.net>.
21. Michigan State University RTSF Mass Spectrometry Facility
22. Cottrell, J. C.; Green, B. N. *MaxEnt: An Essential Maximum Entropy Based Tool for Interpreting Multiply-Charged Electrospray Data*; 212; Micromass UK Limited, 1998.

23. Pasek, S.; Risler, J.-L.; Brézellec, P. Gene Fusion/fission Is a Major Contributor to Evolution of Multi-Domain Bacterial Proteins. *Bioinformatics* **2006**, *22* (12), 1418–1423.
24. Cross, P. J.; Dobson, R. C. J.; Patchett, M. L.; Parker, E. J. Tyrosine Latching of a Regulatory Gate Affords Allosteric Control of Aromatic Amino Acid Biosynthesis. *J. Biol. Chem.* **2011**, *286* (12), 10216–10224.
25. Nazmi, A. R.; Lang, E. J. M.; Bai, Y.; Allison, T. M.; Othman, M. H.; Panjikar, S.; Arcus, V. L.; Parker, E. J. Interdomain Conformational Changes Provide Allosteric Regulation En Route to Chorismate. *J. Biol. Chem.* **2016**, *291* (42), 21836–21847.
26. Cross, P. J.; Allison, T. M.; Dobson, R. C. J.; Jameson, G. B.; Parker, E. J. Engineering Allosteric Control to an Unregulated Enzyme by Transfer of a Regulatory Domain. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (6), 2111–2116.
27. Shumilin, I. A.; Kretsinger, R. H.; Bauerle, R. Purification, Crystallization, and Preliminary Crystallographic Analysis of 3-Deoxy-D-Arabino-Heptulosonate-7-Phosphate Synthase from *Escherichia coli*. *Proteins* **1996**, *24* (3), 404–406.
28. Ramilo, C. A.; Evans, J. N. Overexpression, Purification, and Characterization of Tyrosine-Sensitive 3-Deoxy-D-Arabino-Heptulosonic Acid 7-Phosphate Synthase from *Escherichia coli*. *Protein Expr. Purif.* **1997**, *9* (2), 253–261.
29. Chen, C. C.; Liao, C. C.; Hsu, W. H. The Cloning and Nucleotide Sequence of a *Corynebacterium glutamicum* 3-Deoxy-D-Arabinoheptulosonate-7-Phosphate Synthase Gene. *FEMS Microbiol. Lett.* **1993**, *107* (2-3), 223–229.
30. Liu, Y.-J.; Li, P.-P.; Zhao, K.-X.; Wang, B.-J.; Jiang, C.-Y.; Drake, H. L.; Liu, S.-J. *Corynebacterium glutamicum* Contains 3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate Synthases That Display Novel Biochemical Features. *Appl. Environ. Microbiol.* **2008**, *74* (17), 5497–5503.
31. Zhang, W.; Li, J.; Shi, X.; Hikichi, Y.; Zhang, Y.; Ohnishi, K. Functional Characterization of Two Putative DAHP Synthases of AroG1 and AroG2 and Their Links With Type III Secretion System in *Ralstonia solanacearum*. *Front. Microbiol.* **2019**, *10*, 183.
32. Silakowski, B.; Kunze, B.; Müller, R. Stigmatella Aurantiaca Sg a15 Carries Genes Encoding Type I and Type II 3-Deoxy-D-Arabino-Heptulosonate-7-Phosphate Synthases: Involvement of a Type II Synthase in Aurachin Biosynthesis. *Arch. Microbiol.* **2000**, *173* (5-6), 403–411.
33. Light, S. H.; Halavaty, A. S.; Minasov, G.; Shuvalova, L.; Anderson, W. F. Structural Analysis of a 3-Deoxy-D-Arabino-Heptulosonate 7-Phosphate Synthase with an N-Terminal Chorismate Mutase-like Regulatory Domain. *Protein Sci.* **2012**, *21* (6), 887–895.

34. Bai, Y.; Lang, E. J. M.; Nazmi, A. R.; Parker, E. J. Domain Cross-Talk within a Bifunctional Enzyme Provides Catalytic and Allosteric Functionality in the Biosynthesis of Aromatic Amino Acids. *J. Biol. Chem.* **2019**, *294* (13), 4828–4842.
35. Fan, Y.; Cross, P. J.; Jameson, G. B.; Parker, E. J. Exploring Modular Allostery via Interchangeable Regulatory Domains. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (12), 3006–3011.
36. Bertoni, M.; Kiefer, F.; Biasini, M.; Bordoli, L.; Schwede, T. Modeling Protein Quaternary Structure of Homo- and Hetero-Oligomers beyond Binary Interactions by Homology. *Sci. Rep.* **2017**, *7* (1), 10480.
37. Webby, C. J.; Patchett, M. L.; Parker, E. J. Characterization of a Recombinant Type II 3-Deoxy-D-Arabin-Heptulosonate-7-Phosphate Synthase from *Helicobacter pylori*. *Biochem. J* **2005**, *390* (Pt 1), 223–230.
38. Euverink, G. J.; Hessels, G. I.; Franke, C.; Dijkhuizen, L. Chorismate Mutase and 3-Deoxy-D-Arabin-Heptulosonate 7-Phosphate Synthase of the Methylophilic Actinomycete *Amicorhynchus methanolicus*. *Appl. Environ. Microbiol.* **1995**, *61* (11), 3796–3803.

CHAPTER 3 – *Bacillus methanolicus* MGA3 Culturing

OVERVIEW

Bacillus methanolicus is a thermophilic, endospore-forming, facultative methylotrophic bacterium that has been identified as a promising production host for the conversion of methanol into value-added chemicals.^{1,2} *B. methanolicus* was first isolated from freshwater marsh soil in Minnesota by R. S. Hanson.³ This bacterium uses the ribulose monophosphate pathway (RuMP) for assimilating methanol (see “Native Methylotrophy” in Chapter 1). Methylotrophy is associated with high oxygen demands.⁴ With an optimal growth temperature of 50 °C, *B. methanolicus* requires less external cooling input, reducing the cost of culturing.⁵ Methanol and the elevated culturing temperature can also act as deterrents against contaminating microbes.¹ Additional cost reduction in *B. methanolicus* cultivation can be achieved by culturing in seawater-based media.⁵ The *B. methanolicus* strain MGA3 is generally referred to as the model strain because of its faster growth rate.⁶ As such, MGA3 has been used for metabolic engineering.

While it is true that *B. methanolicus* MGA3 has many attractive qualities of a production host, it is also limited by its minimal genetic engineering tools. Implementation of synthetic methylotrophy in traditional production hosts such as *E. coli* or *C. glutamicum* has emerged as a potential solution to this problem. This too has its limitations as no organism has been successfully engineered in the laboratory to grow solely on methanol.⁷ Current synthetic methylotrophs require a sugar co-substrate and produce less product than native methylotrophs (see “Synthetic Methylotrophy” in Chapter 1).⁸

Recent advances provide insights into development of genetic engineering tools for *B. methanolicus*. The genome, transcriptome, and proteome have been elucidated and analyzed.⁹⁻
¹¹ Plasmids have been designed that can be shuttled between *E. coli* and *B. methanolicus*.¹² These plasmids allow for exogenous gene expression in *B. methanolicus* inducible by mannitol. Targeted gene knock-outs have not yet been described for *B. methanolicus*. The elevated growth

temperatures of *B. methanolicus* has made developing a CRISPR-Cas system for this organism challenging. However, a CRISPRi system utilizing the traditional *Streptococcus pyogenes* dCas9 has been established, which utilizes a catalytically inactive Cas9 to block transcription of the targeted gene.¹³ Additionally, a thermotolerant Cas9 has been discovered that may find use in *B. methanolicus* in the future.¹⁴

While the genetic toolbox for *B. methanolicus* is somewhat limited, mutants have been created through random mutagenesis or by the expression of exogenous genes. Wild-type MGA3 accumulates 55 g L⁻¹ of L-glutamate in a 14 L methanol fermentation. This titer was improved upon by random mutagenesis of MGA3 which afforded strain M168-20 that produces 69 g L⁻¹ of L-glutamate.¹⁵ All *B. methanolicus* mutants reported in the literature produce either lysine, glutamate or molecules derived from lysine by the introduction of an exogenous gene (Table 10).

Table 10. Select *Bacillus methanolicus* production strains.

Product	Strain	Titer (g L ⁻¹)	Reference
L-Glutamate	MGA3	58	16
L-Glutamate	M168-20	69	15
L-Lysine	NOA2#13A52-8A66	60	17
GABA	MGA3(pTH1mp-gad St)	9.0	18
Cadaverine	MGA3(pTH1mp-cadA)	11.3	19

Random mutagenesis requires culturing *B. methanolicus* in media containing either N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) or ethyl methansulfonate (EMS).³ Both MNNG and EMS alkylate the DNA bases, mainly guanine (typically on O⁶) and thymine (typically on O⁴) (Figure 29).²⁰⁻²² Bases are methylated by MNNG but ethylated by EMS. The alkylation of the bases causes disruptions to their hydrogen bonding patterns. Without the appropriate hydrogen

bonding patterns, errors during DNA replication occur. Specifically in the case with MNNG and EMS, T:A to C:G and G:C to A:T transversion mutations take place.^{21,22} One point mutation can be the difference between a functional and non-functional enzyme.

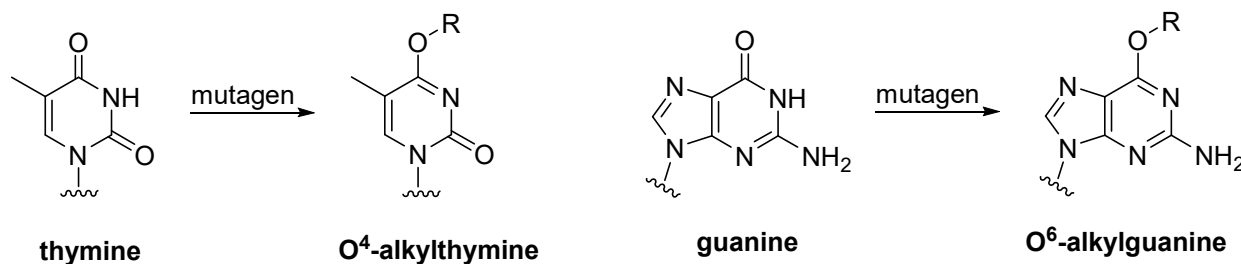


Figure 29. The most common alkylation products of thymine and guanine resulting from MNNG or EMS treatment.

CULTURING

We received our *B. methanolicus* strain and culturing recommendations from Dr. Trygve Brautaset at the Norwegian University of Science and Technology. *B. methanolicus* was cultured in a minimal salt media (MVcM) supplemented with trace vitamins, metals, and a carbon source. The media components are described in detail in Chapter 4. Yeast extract (0.25 g L⁻¹) was added to the seed cultures to shorten the initial lag phase after inoculation from a glycerol freeze. At this concentration, the yeast extract only supports growth up to an OD₆₀₀ of 0.2.²³ When yeast extract is present, the media is referred to as MVcMY. This organism can use glucose, mannitol, methanol, or arabitol as carbon sources.^{24,25} Generally, mannitol is used for culturing when methanol-free conditions are preferred. The carbon source and its concentration are indicated as a subscript with Mann standing for mannitol and MeOH standing for methanol. For example, MVcMY_{MeOH200} is the minimal salt media supplemented with 0.25 g L⁻¹ of yeast extract and 200 mM methanol.

The first *B. methanolicus* strain we received was faulty (F-MGA3) but this was not immediately evident. F-MGA3 exhibited a faster growth rate (0.53 h^{-1}) than that reported in the literature (0.32 h^{-1}).²³ F-MGA3 also appeared to grow on methanol only after first being grown in a seed culture containing a mixture of mannitol and methanol. This phenotype was not originally concerning to us or to our collaborators because formaldehyde accumulation can occur when MGA3 is shocked with high methanol concentrations.²⁶ Formaldehyde is reactive and is toxic to plants, animals, and bacteria. The reactivity of formaldehyde is a result of its nucleophilicity, which allows it to react with electrophiles on proteins and DNA such as thiols or amines, thereby disrupting native biological processes.²⁷ In order to assimilate methanol, the alcohol must first be converted to formaldehyde by methanol dehydrogenase (Mdh). Formaldehyde is then coupled with ribulose-5-phosphate by hexulose-6-phosphate synthase (Hps). Mdh is constitutively expressed, while Hps expression is induced by methanol (Figure 30).²³ Hps levels are low when MGA3 is growing on mannitol. When the cells are exposed to high levels of methanol, Mdh produces formaldehyde at a rate too quickly for the limited amount of Hps to process, which leads to formaldehyde accumulation.

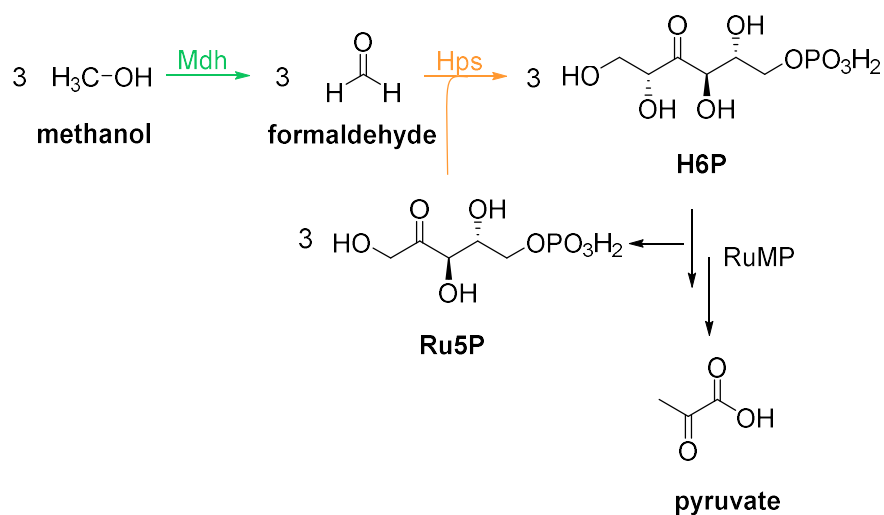


Figure 30. Low levels of Hps cause formaldehyde accumulation which interferes with native biological processes and hinders the growth of *B. methanolicus*.

The above information was how we mistakenly explained the need for a seed culture of mannitol and methanol to initiate methylotrophic growth. In reality, the strain we had did not grow on methanol at all. This was not immediately apparent, as residual mannitol transferred from seed cultures during re-inoculation sustained MGA3 growth to an acceptable OD₆₀₀. It was not until the summer of 2019 when Yasheen Jadidi and I visited Dr. Brautaset's lab in Trondheim, Norway that we were able to bring back the correct strain.

Table 11 shows the growth rates of MGA3 in various culturing conditions. For this experiment, MGA3 cultures were grown overnight in MVcMY_{mann50} or MVcMY_{MeOH200}. The OD₆₀₀ of the seed cultures were measured and the appropriate volumes were inoculated into a variety of media to an initial OD₆₀₀ of 0.2 (Figure 31). The OD₆₀₀ was measured periodically, and the growth rate was calculated from the exponential growth phase (Table 11). The conditions resulting in the fastest growth rate was a methanol seed culture inoculated into methanol media containing yeast extract. Close behind is the methanol seed culture inoculated into a methanol-only media. Notably, the mannitol seed culture could not be inoculated into a methanol-only media for successful growth. This is likely because of formaldehyde accumulation, as described above.

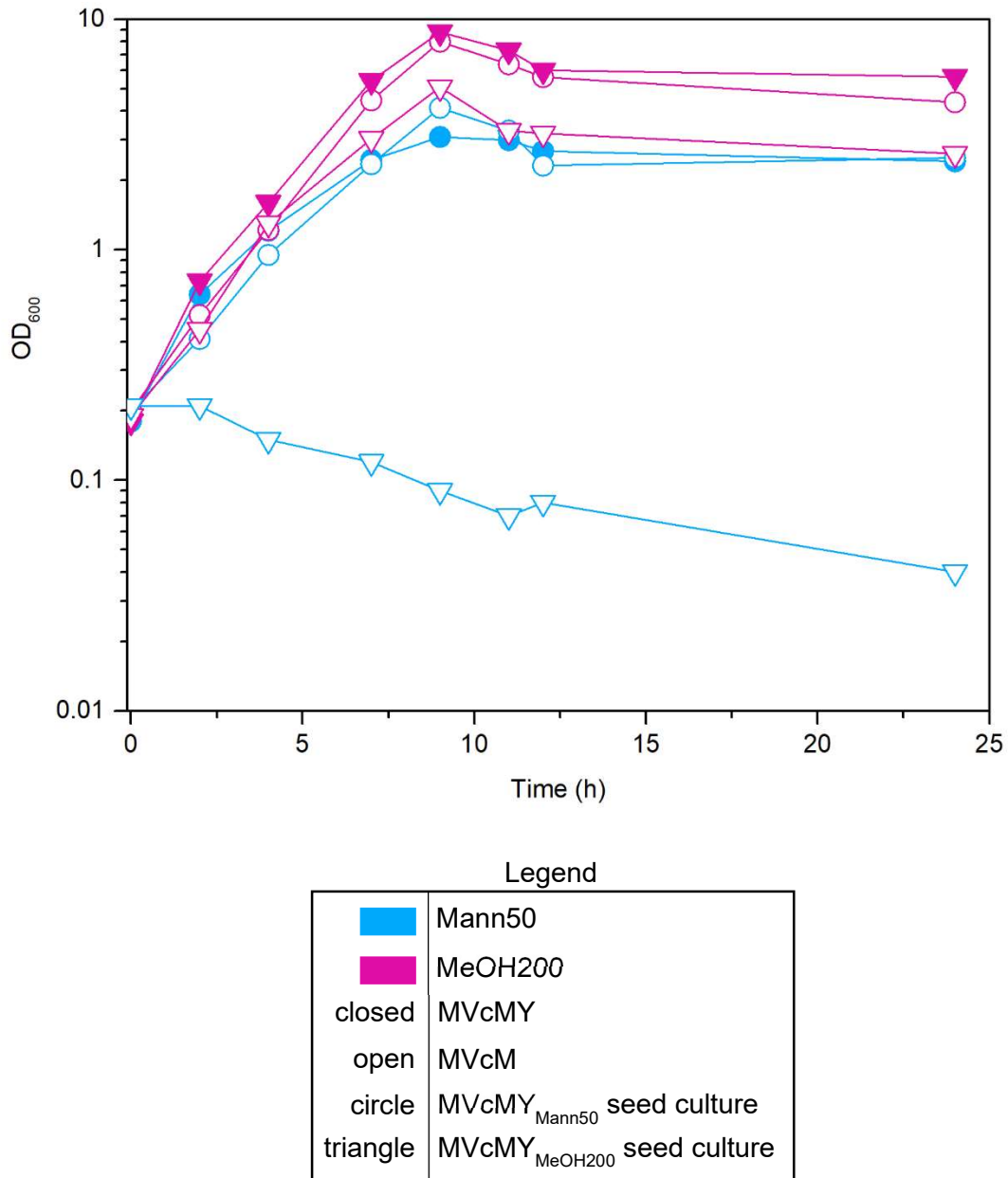


Figure 31. *B. methanolicus* MGA3 growth curves in a variety of conditions. The shape of the data point corresponds to which seed culture was used to inoculate the culture (circle = MVcMY_{Mann50}; triangle = MVcMY_{MeOH200}). The fill of the data point corresponds to the type of media used for the culture (closed = MVcMY; open = MVcM). The color of the data point corresponds to the carbon used for the culture (blue = Mann50; pink = MeOH200).

Table 11. Growth rates of MGA3 cultured in varied conditions.

Conditions	Seed Culture Media	Media	Growth Rate (h ⁻¹)
A	MVcMY _{mann50}	MVcMY _{mann50}	0.30
B	MVcMY _{mann50}	MVcM _{mann50}	0.34
C	MVcMY _{mann50}	MVcM _{MeOH200}	0.00
D	MVcMY _{meOH200}	MVcMY _{MeOH200}	0.42
E	MVcMY _{meOH200}	MVcM _{mann50}	0.37
F	MVcMY _{meOH200}	MVcM _{MeOH200}	0.41

SPORULATION

Before random chemical mutagenesis could be performed on MGA3, a reliable plating and storage method for the mutants had to be established. Sporulation has been used previously to aid in screening and storage of MGA3 mutants.³ Spores are a dormant form of bacteria that are comprised of its DNA protected by a hard protein coat.¹³ Favorable environmental conditions such as the presence of sufficient nutrients and optimal temperatures can germinate the spore which allows the cell to resume metabolic activity. Sporulation of the future *B. methanolicus* mutants was an attractive solution to circumvent possible complications caused by their unknown behavior during storage. This way, a large batch of cells could be mutagenized and stored in a dormant state until they could be screened.

Two sporulation methods were explored: the salts method and the starvation method, with the starvation method being the most successful. After 48 hours of sporulation, samples from both sporulation methods were heat-fixed to a microscope slide. The slides were stained with Malachite Green followed by the counterstain Safranin O.²⁸ The cells were visualized using a Nikon A1Rsi Confocal Laser Scanning Microscope.²⁹ Malachite green and Safranin O were excited at 615 nm and 543 nm, respectively. Emission of malachite green was measured at 650

nm and emission of Safranin O was measured at 580 nm. The resulting fluorescent images were then overlaid.

The salts method was shared by Marta Irla and was later published in the literature.^{13,30} To sporulate MGA3 with salts, the cells were grown to exponential phase at 50 °C, 200 rpm in MVcM_{MeOH200}. The sporulation salts (MnCl₂, MgCl₂, and CaCl₂) were then added to the media and the cells were cultured at 37 °C, 200 rpm for an additional 48 hours. The starvation method was based on a protocol published by Schendel et al. Using this method, MGA3 cells were grown to exponential phase at 50 °C, 200 rpm in MVcM_{MeOH200}. The media was then diluted three-fold with carbon-free media and the cells were cultured at 37 °C, 200 rpm for an additional 48 hours.³ Using the salts method, approximately 52% of MGA3 sporulated after 48 hours whereas only about 17% of MGA3 sporulated after 48 hours using the starvation method (Figure 32). Assistance with this experiment was provided by Yasheen Jadidi, a PhD student in our lab.

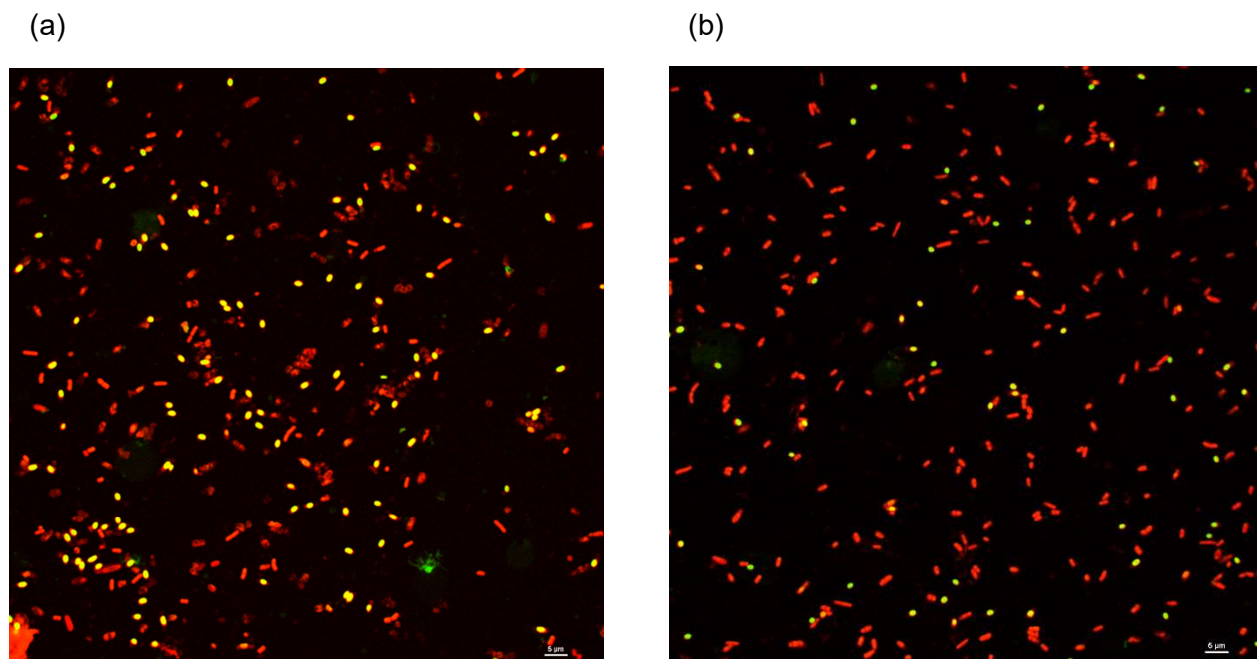


Figure 32. Images of *B. methanolicus* MGA3 cultures after exposure to sporulation conditions. Vegetative cells are shown in red and spores are shown in yellow. (a) cells collected after 48 hours of sporulation using the salts method. (b) cells collected after 48 hours of sporulation using the starvation method.

PLATING AND SELECTIVE PLATING

B. methanolicus shikimate pathway mutants will be screened by selective plating. Plating this organism directly from a glycerol freeze or a liquid culture is a challenge faced by both our lab and the lab of our collaborators, as MGA3 does not readily form single colonies on solid media. The plating of MGA3 spores does result in single colonies. Various plate recipes were tested for application in replica plating (Table 12). Cells were sporulated then plated onto SOB solid media. The plates were then incubated at 50 °C for 24 hours. Single colonies were then replica plated onto a defined media plate without aromatic amino acids (aros), then a defined media plate with aros, followed by SOB again (Figure 33). The spores were always replica plated onto plates without aros first in order to prevent false positives. These plates were then incubated at 50 °C until growth was observed, typically 8-12 hours. The final SOB plate served as a positive control, ensuring that enough biomass was being transferred to the plates. The percentage of colonies transferred was recorded to determine which recipe supported MGA3 the best. Some plates contain amino acids at concentrations that mimic yeast extract. Concentrations of the yeast extract mimic amino acids (YEMAA) can be found in chapter 4.

Table 12. Tested plate recipes for *B. methanolicus* shikimate pathway mutant screening.

Recipe	Mannitol (mM)	Vitamins	Metals	YE ^a (g L ⁻¹)	YEMAA ^a	aros ^a	Aromatic Vitamins	Results
SOB ^b	-----	-----	-----	-----	-----	-----	-----	100%
A ^c	50	1X	1X	0.25	0.0	0.0	0.0	0%
B ^c	50	2X	1X	0.25	0.0	0.0	0.0	0%
C ^c	50	3X	1X	0.25	0.0	0.0	0.0	0%
D ^c	50	4X	1X	0.25	0.0	0.0	0.0	0%
E ^d	50	4X	1X	0.0	1X	0.0	0.0	40%
F ^d	50	4X	1X	0.0	1X	1X	0.0	25%
G ^e	50	4X	1X	0.0	10X	0.0	0.0	100%
H ^e	50	4X	1X	0.0	10X	10X	0.0	0.0%
I ^f	100	4X	1X	0.0	1X	0.0	0.0	100%
J ^f	100	4X	1X	0.0	1X	1X	0.0	100%
K ^g	100	4X	1X	0.0	1X	0.0	0.0	100%
L ^g	100	4X	1X	0.0	1X	1X	1X	100%

^a Yeast extract (YE); yeast extract mimic amino acids (YEMAA); aromatic amino acids (aros)

^b Super optimal broth (SOB)

^c MVcMY_{Mann50} plates with varying concentrations of trace vitamins

^d MVcM_{Mann50} plates with 4X vitamins either with or without aros

^e MVcM_{Mann50} plates with 4X vitamins, 10X YEMAA, either with or without 10X aros

^f MVcM_{Mann100} plates with 4X vitamins either with or without aros

^g MVcM_{Mann100} plates with 4X vitamins either with or without aros and aromatic vitamins

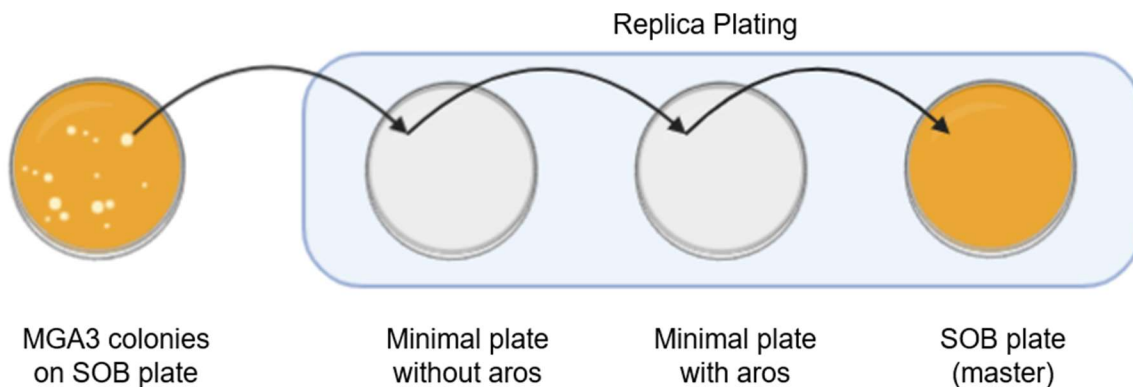


Figure 33. Method for replica plating *B. methanolicus* MGA3. MGA3 spores were plated onto a SOB plate then incubated at 50 °C for 24 hours. Single colonies were then picked up on a pipette tip and transferred to three plates: a minimal plate without aros, a minimal plate with aros, and a SOB master plate.

All plate recipes examined included mannitol as the carbon source, as we have not yet been successful in developing methanol solid media. The first plate component that was varied was the concentration of the trace vitamins (plates A-D). Yeast extract was included in these plates at a concentration of 0.25 g L⁻¹. No biomass transfer was observed using these plates. Since yeast extract contains aros, plates A-D would have been ineffective at identifying mutants of the shikimate pathway. To address this issue, yeast extract was replaced by adding amino acids at concentrations that mimicked that of yeast extract (plate set E & F). Only 25 – 40% of the colonies transferred grew on these plates, so the amino acid concentration was increased ten-fold (plate set G & H). Plate G supported 100% of the colonies transferred, but plate H which contained the aromatic amino acids did not. The aros in plate H were present in too high of a concentration, which caused them to precipitate from the media upon incubation at 50 °C.

Next, the mannitol concentration was doubled in plate set I & J, which also contained 4X trace vitamins and 1X YEMAA. This higher concentration of mannitol supported growth of all colonies transferred to these plates. Finally, *p*-aminobenzoic acid was removed from the trace vitamins in plate K and all the aromatic vitamins were present in plate L. Plate set K & L supported

growth of all colonies that were transferred to them and contained all the necessary supplementation that would be required for a shikimate pathway mutant of *B. methanolicus*.

Plate set K & L (renamed Sel and SelAros, respectively) supported growth of MGA3 both with and without aros and aromatic vitamins. These plate recipes are likely to be an effective screening method for shikimate pathway mutants of *B. methanolicus*.

CONCLUSION

B. methanolicus has been recognized as a potential production host for value-added chemical from methanol. A major limitation in generating *B. methanolicus* production strains is the lack of targeted gene knock-out technology. Being a thermophile, the traditional methods for disrupting genes are not viable in *B. methanolicus*.³¹ Successfully generating genomic mutations via random mutagenesis is possible but requires a meticulous screening method which must be tailored specifically for the desired mutant. The work presented here lays the foundation for engineering the shikimate pathway in *Bacillus methanolicus* MGA3 by random chemical mutagenesis. Culturing protocols were established and modified to accommodate mutagenesis techniques.

The starvation method for sporulating MGA3 was found to be a reliable way to store mutants. Germination of these spores can be achieved by culturing in either liquid or solid media. For selection of a mutation in the shikimate pathway, aros and aromatic vitamins must be supplemented in the growth media. Due to the hydrophobicity of the aros, their concentration must be kept to a minimum.

Solid media for selecting shikimate pathway mutants were defined as Sel and SelAros. *B. methanolicus* containing a mutation in the shikimate pathway should not grow on Sel plates but should grow on SelAros plates. Plates should be spotted in the following order with a pipette tip to avoid false positives: Sel, SelAros, SOB. SOB plates act as a master plate and a positive control, ensuring enough biomass was transferred to the minimal plates.

REFERENCES

REFERENCES

1. Brautaset, T.; Jakobsen, Ø. M.; Josefsen, K. D.; Flickinger, M. C.; Ellingsen, T. E. *Bacillus methanolicus*: A Candidate for Industrial Production of Amino Acids from Methanol at 50 Degrees C. *Appl. Microbiol. Biotechnol.* **2007**, *74* (1), 22–34.
2. Müller, J. E. N.; Heggeset, T. M. B.; Wendisch, V. F.; Vorholt, J. A.; Brautaset, T. Methylo-trophy in the Thermophilic *Bacillus methanolicus*, Basic Insights and Application for Commodity Production from Methanol. *Appl. Microbiol. Biotechnol.* **2015**, *99* (2), 535–551.
3. Schendel, F. J.; Bremmon, C. E.; Flickinger, M. C.; Guettler, M.; Hanson, R. S. L-Lysine Production at 50 Degrees C by Mutants of a Newly Isolated and Characterized Methylo-trophic *Bacillus* sp. *Appl. Environ. Microbiol.* **1990**, *56* (4), 963–970.
4. Arfman, N.; de Vries, K. J.; Moezelaar, H. R.; Attwood, M. M.; Robinson, G. K.; van Geel, M.; Dijkhuizen, L. Environmental Regulation of Alcohol Metabolism in Thermotolerant Methylo-trophic *Bacillus* Strains. *Arch. Microbiol.* **1992**, *157* (3), 272–278.
5. Komives, C. F.; Cheung, L. Y.-Y.; Pluschkell, S. B.; Flickinger, M. C. Growth of *Bacillus methanolicus* in Seawater-Based Media. *J. Ind. Microbiol. Biotechnol.* **2005**, *32* (2), 61–66.
6. Heggeset, T. M. B.; Krog, A.; Balzer, S.; Wentzel, A.; Ellingsen, T. E.; Brautaset, T. Genome Sequence of Thermotolerant *Bacillus methanolicus*: Features and Regulation Related to Methylo-trophy and Production of L-Lysine and L-Glutamate from Methanol. *Appl. Environ. Microbiol.* **2012**, *78* (15), 5170–5181.
7. Heux, S.; Brautaset, T.; Vorholt, J. A.; Wendisch, V. F.; Portais, J. C. Synthetic Methylo-trophy: Past, Present, and Future. In *Methane Biocatalysis: Paving the Way to Sustainability*; Kalyuzhnaya, M. G., Xing, X.-H., Eds.; Springer International Publishing: Cham, 2018; pp 133–151.
8. Zhang, W.; Song, M.; Yang, Q.; Dai, Z.; Zhang, S.; Xin, F.; Dong, W.; Ma, J.; Jiang, M. Current Advance in Bioconversion of Methanol to Chemicals. *Biotechnol. Biofuels* **2018**, *11*, 260.
9. Irla, M.; Neshat, A.; Winkler, A.; Albersmeier, A.; Heggeset, T. M. B.; Brautaset, T.; Kalinowski, J.; Wendisch, V. F.; Rückert, C. Complete Genome Sequence of *Bacillus methanolicus* MGA3, a Thermotolerant Amino Acid Producing Methylo-troph. *J. Biotechnol.* **2014**, *188*, 110–111.
10. Irla, M.; Neshat, A.; Brautaset, T.; Rückert, C.; Kalinowski, J.; Wendisch, V. F. Transcriptome Analysis of Thermophilic Methylo-trophic *Bacillus methanolicus* MGA3 Using RNA-Sequencing Provides Detailed Insights into Its Previously Uncharted Transcriptional Landscape. *BMC Genomics* **2015**, *16*, 73.

11. Müller, J. E. N.; Litsanov, B.; Bortfeld-Miller, M.; Trachsel, C.; Grossmann, J.; Brautaset, T.; Vorholt, J. A. Proteomic Analysis of the Thermophilic Methylophilic *Bacillus methanolicus* MGA3. *Proteomics* **2014**, *14* (6), 725–737.
12. Irla, M.; Heggeset, T. M. B.; Nærdal, I.; Paul, L.; Haugen, T.; Le, S. B.; Brautaset, T.; Wendisch, V. F. Genome-Based Genetic Tool Development for *Bacillus methanolicus*: Theta- and Rolling Circle-Replicating Plasmids for Inducible Gene Expression and Application to Methanol-Based Cadaverine Production. *Front. Microbiol.* **2016**, *7*, 1481.
13. Schultenkämper, K.; Brito, L. F.; López, M. G.; Brautaset, T.; Wendisch, V. F. Establishment and Application of CRISPR Interference to Affect Sporulation, Hydrogen Peroxide Detoxification, and Mannitol Catabolism in the Methylophilic Thermophile *Bacillus methanolicus*. *Appl. Microbiol. Biotechnol.* **2019**, *103* (14), 5879–5889.
14. Mougiakos, I.; Mohanraju, P.; Bosma, E. F.; Vrouwe, V.; Finger Bou, M.; Naduthodi, M. I. S.; Gussak, A.; Brinkman, R. B. L.; van Kranenburg, R.; van der Oost, J. Characterizing a Thermostable Cas9 for Bacterial Genome Editing and Silencing. *Nat. Commun.* **2017**, *8* (1), 1647.
15. Brautaset, T.; Jakobsen, Ø. M.; Degnes, K. F.; Netzer, R.; Naerdal, I.; Krog, A.; Dillingham, R.; Flickinger, M. C.; Ellingsen, T. E. *Bacillus methanolicus* Pyruvate Carboxylase and Homoserine Dehydrogenase I and II and Their Roles for L-Lysine Production from Methanol at 50 Degrees C. *Appl. Microbiol. Biotechnol.* **2010**, *87* (3), 951–964.
16. Brautaset, T.; Williams, M. D.; Dillingham, R. D.; Kaufmann, C.; Bennaars, A.; Crabbe, E.; Flickinger, M. C. Role of the *Bacillus methanolicus* Citrate Synthase II Gene, citY, in Regulating the Secretion of Glutamate in L-Lysine-Secreting Mutants. *Appl. Environ. Microbiol.* **2003**, *69* (7), 3986–3995.
17. Nærdal, I.; Netzer, R.; Irla, M.; Krog, A.; Heggeset, T. M. B.; Wendisch, V. F.; Brautaset, T. L-Lysine Production by *Bacillus methanolicus*: Genome-Based Mutational Analysis and L-Lysine Secretion Engineering. *J. Biotechnol.* **2017**, *244*, 25–33.
18. Irla, M.; Nærdal, I.; Brautaset, T.; Wendisch, V. F. Methanol-Based γ -Aminobutyric Acid (GABA) Production by Genetically Engineered *Bacillus methanolicus* Strains. *Ind. Crops Prod.* **2017**, *106*, 12–20.
19. Nærdal, I.; Netzer, R.; Irla, M.; Krog, A.; Heggeset, T. M. B.; Wendisch, V. F.; Brautaset, T. L-Lysine Production by *Bacillus methanolicus*: Genome-Based Mutational Analysis and L-Lysine Secretion Engineering. *J. Biotechnol.* **2017**, *244*, 25–33.
20. Haerlin, R.; Süßmuth, R.; Lingens, F. Mechanism of Mutagenesis by N-Methyl-N'-Nitro-N-Nitroso-Guanidine (MNNG) V. Methylation of DNA by N-Trideuteriomethyl-N'-Nitro-N-Nitroso-Guanidine (D(3)-MNNG). *FEBS Lett.* **1970**, *9* (3), 175–176.

21. Wang, X.; Kitamura, K.; Yamamoto, K. Mutagenic Specificity of N-Methyl-N'-Nitro-N-Nitrosoguanidine in the tonB Gene on the Chromosome of *Escherichia coli* recA⁺ and recA⁻ Cells. *Biochem. Biophys. Res. Commun.* **1996**, *227* (2), 334–339.
22. Sega, G. A. A Review of the Genetic Effects of Ethyl Methanesulfonate. *Mutat. Res.* **1984**, *134* (2-3), 113–142.
23. Jakobsen, Ø. M.; Benichou, A.; Flickinger, M. C.; Valla, S.; Ellingsen, T. E.; Brautaset, T. Upregulated Transcription of Plasmid and Chromosomal Ribulose Monophosphate Pathway Genes Is Critical for Methanol Assimilation Rate and Methanol Tolerance in the Methylophilic Bacterium *Bacillus methanolicus*. *J. Bacteriol.* **2006**, *188* (8), 3063–3072.
24. Arfman, N.; Dijkhuizen, L.; Kirchhof, G.; Ludwig, W.; Schleifer, K. H.; Bulygina, E. S.; Chumakov, K. M.; Govorukhina, N. I.; Trotsenko, Y. A.; White, D. *Bacillus methanolicus* Sp. Nov., a New Species of Thermotolerant, Methanol-Utilizing, Endospore-Forming Bacteria. *Int. J. Syst. Bacteriol.* **1992**, *42* (3), 439–445.
25. López, M. G.; Irla, M.; Brito, L. F.; Wendisch, V. F. Characterization of D-Arabitol as Newly Discovered Carbon Source of *Bacillus methanolicus*. *Front. Microbiol.* **2019**, *10*, 1725.
26. Bozdag, A.; Komives, C.; Flickinger, M. C. Growth of *Bacillus methanolicus* in 2 M Methanol at 50 °C: The Effect of High Methanol Concentration on Gene Regulation of Enzymes Involved in Formaldehyde Detoxification by the Ribulose Monophosphate Pathway. *J. Ind. Microbiol. Biotechnol.* **2015**, *42* (7), 1027–1038.
27. Chen, N. H.; Djoko, K. Y.; Veyrier, F. J.; McEwan, A. G. Formaldehyde Stress Responses in Bacterial Pathogens. *Front. Microbiol.* **2016**, *7*, 257.
28. Schaeffer, A. B.; Fulton, M. D. A Simplified Method of Staining Endospores. *Science* **1933**, *77* (1990), 194.
29. Center for Advanced Microscopy, Michigan State University
30. Irla, M. Norwegian University of Science and Technology, Trondheim, Norway. Personal communication, 2018-2020.
31. Drejer, E. B.; Hakvåg, S.; Irla, M.; Brautaset, T. Genetic Tools and Techniques for Recombinant Expression in Thermophilic Bacillaceae. *Microorganisms* **2018**, *6* (2).

CHAPTER 4 – Experimental

MATERIALS

All chemicals were purchased from MilliporeSigma unless otherwise stated. DNA primers, codon optimized genes, and *E. coli* BL21(DE3) were purchased from Integrated DNA Technologies. High efficiency *E. coli* NEB5 α competent cells, restriction enzymes, Monarch gel extraction kit, Monarch DNA clean-up kit, Q5 DNA polymerase, HiFi master mix, and 1 kb DNA ladder were purchased from New England Biolabs. The 1 kb+ DNA ladder was purchased from Promega. Pre-cast 4-20 % (w/v) acrylamide Mini-PROTEAN TGX gels and Precision Plus Protein All Blue Ladder were purchased from BioRad. Plasmid extraction kits were purchased from Qiagen. Nickel affinity columns for protein purification were HisTrap Fast Flow Sepharose from GE life sciences. Protein samples were concentrated using an Amicon Ultra-15 10K Centrifugal Filter with an Ultracel-50 membrane. Buffer exchanges and desalting of protein samples were performed using a GE Life Sciences PD-10 column. *B. methanolicus* MGA3 was received as a gift from Dr. Trygve Brautaset (Norwegian University of Science and Technology, Trondheim, Norway). *E. coli* BL21-CodonPlus(DE3)-RP and pMHT Δ 238 were received as gifts from Dr. Heedok Hong (Michigan State University, East Lansing, MI). All commercial kits were used according to the manufacturer's instructions.

INSTRUMENTAL MEASUREMENTS

UV-visible measurements were recorded on a SpectraMax iD3 microplate reader and visualized using Softmax Pro 7.1. DNA and protein concentrations were measured using a Thermo Scientific™ Nanodrop™ One^C. Fast protein liquid chromatography (FPLC) was carried out using a GE Life Sciences ÄKTA Start. Mass spectral analysis of AroG1 and AroG2 were performed using electrospray ionization mass spectrometry (ESI-MS) on a Xevo G2-XS time-of-flight mass spectrometer provided by the Michigan State University Research Technology

Support Facility (MSU RTSF) Mass Spectrometry Core. Images of agarose gels and SDS-PAGE gels were recorded using an Axygen GD-1000 Gel Documentation System.

BACTERIAL STRAINS AND PLASMIDS

All bacterial strains and plasmids are shown in Table 13. DH5 α was used for long-term plasmid storage while BL21(DE3) and BL21-CodonPlus(DE3)-RP were used for protein expression. High efficiency *E. coli* NEB5 α competent cells were used for transformation of NEB HiFi DNA Assembly products. The *aroG1* and *aroG2* gene sequences were codon-optimized for expression in *E. coli* and inserted into pUC vectors by Integrated DNA Technologies, resulting in pUC-IDTamp+aroG1 and pUC-IDTamp+aroG2. Plasmid pET-15b was used as the general cloning vector for preparation of plasmids pMG2.109 and pMG2.128 via HiFi DNA Assembly. Plasmid pET-21(a)+ was used as the cloning vector for preparation of pMG2.209 via HiFi DNA Assembly.

All plasmids were transformed into electrocompetent DH5 α and frozen at -80 °C for long-term storage. All bacterial strains were stored in 20% glycerol at -80 °C. *E. coli* glycerol freezes were prepared by combining 250 μ L of sterile 80% glycerol with 750 μ L of an overnight cell culture in the late log growth phase in a sterile cryo-vial. This mixture was incubated at room temperature for two hours and then flash frozen in liquid nitrogen prior to storage at -80 °C. *B. methanolicus* glycerol freezes were prepared in the same manner except the samples were immediately flash frozen without an incubation period.

Table 13. Bacterial strains and plasmids used in this work.

Strain or Plasmid	Description	Ref
<i>E. coli</i> strains		
DH5 α	Host for plasmid storage	Invitrogen
NEB5 α	General cloning host	NEB
BL21(DE3)	Expression host for AroG1, AroG2, and TktA	Invitrogen
BL21-CodonPlus(DE3)-RP	Expression host for TEV protease	Hong ^b
<i>B. methanolicus</i> strains ^a		
MGA3	Wild-type	1
F-MGA3	Faulty MGA3 strain	This work
Plasmids		
pUC-IDTamp+aroG1	aroG1 cloning vector, Amp ^R	IDT
pUC-IDTamp+aroG2	aroG2 cloning vector, Amp ^R	IDT
pET-15b	General cloning vector, Amp ^R	Lab plasmid
pET-21(a)+	General cloning vector, Amp ^R	Lab plasmid
pMG2.109	AroG1 in pET-15b, Amp ^R	This work
pMG1.128	AroG2 in pET-15b, Amp ^R	This work
pMG2.208	TktA in pET-21(a)+, Amp ^R	This work
pMHT Δ 238	TEV protease expression vector, Kan ^R	Hong ^b

^a The faulty *B. methanolicus* strain (F-MGA3) was received by mail from the lab of Dr. Trygve Brautaset in January, 2018. The correct *B. methanolicus* strain (MGA3) was transported personally from the lab of Dr. Trygve Brautaset in July, 2019.

^b BL21-CodonPlus(DE3)-RP and pMHT Δ 238 were gifts from Dr. Heedok Hong at Michigan State University.

CULTURE MEDIA

LB medium contained Bacto tryptone (10 g L^{-1}), Difco yeast extract (5 g L^{-1}), and NaCl (10 g L^{-1}).² SOC medium contained Bacto tryptone (20 g L^{-1}), Difco yeast extract (5 g L^{-1}), NaCl (0.5 g L^{-1}), KCl (0.186 g L^{-1}), Mg_2SO_4 (2.4 g L^{-1}), and glucose (10 mM). Commercial SOC medium (NEB) was used for recovery of NEB5 α transformants. 2xYT medium contained Bacto Tryptone (16 g L^{-1}), Difco yeast extract (10 g L^{-1}), and NaCl (5 g L^{-1}). Solid media was prepared by addition of 1.5% agar. When required for plasmid maintenance, ampicillin (Amp) or kanamycin (Kan) were added to a final concentration of $50 \text{ } \mu\text{g mL}^{-1}$. Solutions of Mg_2SO_4 and glucose were autoclaved separately from the rest of the media and were combined after sterilization.

B. methanolicus MGA3 was cultured in MVcM minimal medium, MVcMY minimal medium, or SOB media.³ MVcM and MVcMY were supplemented with 50 mM mannitol and/or 200 mM methanol. Mannitol solutions were autoclaved separately from other media components. Methanol was not sterilized. Methanol or sterilized mannitol was added to sterile media as needed. MVcM contained K_2HPO_4 (40.93 g L^{-1}), KH_2PO_4 (12.96 g L^{-1}), and $(\text{NH}_4)_2\text{SO}_4$ (21.14 g L^{-1}). MVcMY was MVcM media supplemented with Difco yeast extract (0.25 g L^{-1}). The minimal media described above was supplemented with 1X trace metals, 1X trace vitamins, and 0.1 mM Mg_2SO_4 . The trace metals and trace vitamins were prepared as 1000X concentration solutions, sterilized by passage through a $0.22 \text{ } \mu\text{m}$ filter, and added to sterile media. Trace metals (1000X) contained $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ (5.56 g L^{-1}), $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ (0.04 g L^{-1}), $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ (7.35 g L^{-1}), $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ (0.04 g L^{-1}), $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ (9.90 g L^{-1}), $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ (0.29 g L^{-1}), $(\text{NH}_4)_2\text{MoO}_4 \cdot 4\text{H}_2\text{O}$ (0.01 g L^{-1}). 1000X trace vitamins contained D-biotin (0.1 g/L), thiamine HCl (0.1 g/L), riboflavin (0.1 g/L), pyridoxine HCl (0.1 g/L), pantothenate (0.1 g L^{-1}), nicotinamide (0.1 g L^{-1}), *p*-aminobenzoic acid (0.1 g L^{-1}), folic acid (0.02 g L^{-1}), and alphamine (0.1 g L^{-1}). SOB media contained Bacto tryptone (20 g L^{-1}), Difco yeast extract (5 g L^{-1}), NaCl (0.5 g L^{-1}), KCl (0.186 g L^{-1}), and Mg_2SO_4 (2.4 g L^{-1}). When using the salts sporulation method MnCl_2 (15 mg L^{-1}), MgCl_2 (1.0 g L^{-1}), and CaCl_2 (0.8 g L^{-1}) were added to MGA3 cultures growing in $\text{MVcM}_{\text{MeOH}200}$.

The selection plates for *B. methanolicus* MGA3 shikimate pathway mutant selection are as follows. All solid media was prepared by addition of 1.5% agar. Sel medium contained K_2HPO_4 (40.93 g L^{-1}), KH_2PO_4 (12.96 g L^{-1}), $(\text{NH}_4)_2\text{SO}_4$ (21.14 g L^{-1}), 1X trace metals, 4X trace vitamins excluding *p*-aminobenzoic acid, 1X yeast extract mimic amino acids (YEMAA), 0.1 mM Mg_2SO_4 , and 100 mM mannitol. SelAros medium was Sel medium supplemented with 1X YEMAA and 1X YEMAAaros. YEMAA (1000X) contained L-alanine (10.25 mg mL^{-1}), L-asparagine (7.0 mg mL^{-1}), L-cystine (1.5 mg mL^{-1}), L-glutamate (27.5 mg mL^{-1}), L-glycine (6.25 mg mL^{-1}), L-histidine (2.75 mg mL^{-1}), L-isoleucine (6.5 mg mL^{-1}), L-leucine (9.25 mg mL^{-1}), L-lysine (10.5 mg mL^{-1}), L-methionine (2.0 mg mL^{-1}), L-proline (5.5 mg mL^{-1}), L-serine (6.75 mg mL^{-1}), L-threonine (6.5 mg mL^{-1}), L-valine (7.75 mg mL^{-1}). YEMAAaros (1000X) contained L-phenylalanine (6.25 mg mL^{-1}), L-tryptophan (2.0 mg mL^{-1}), and L-tyrosine (2.0 mg mL^{-1}). SOB medium contained Bacto tryptone (20 g L^{-1}), Difco yeast extract (5 g L^{-1}), NaCl (0.5 g L^{-1}), KCl (0.186 g L^{-1}), and Mg_2SO_4 (2.4 g L^{-1}). Solutions of Mg_2SO_4 and mannitol were autoclaved separately and added to sterile media as needed. Solutions of YEMAA and YEMAAaros were sterilized separately by passage through a 0.22 μm syringe filter and added to sterile media as needed.

GENERAL DNA MANIPULATION TECHNIQUES

For small-scale plasmid purification, the appropriate DH5 α strain was cultured for 12-15 h in a 5 mL culture of LB supplemented with the appropriate antibiotic. The cells were pelleted by centrifugation at 14,100 x *g* for 2 minutes at room temperature. The plasmids were extracted from the cell pellets. For large-scale plasmid purification, the appropriate DH5 α was cultured for 12-15 h in a 100 mL culture of LB supplemented with the appropriate antibiotic. The cells were pelleted by centrifugation at 7,200 x *g* for 30 minutes at 4 °C. The plasmids were extracted from the cell pellets.

Restriction digests were performed by incubating the DNA with the appropriate restriction enzyme(s) using the recommended buffer supplied by the manufacturer. The reactions were

incubated at 37 °C for one hour and terminated by addition of 0.1 volumes of 10X endostop (1.1 mM EDTA pH 8.0, 50% glycerol, 0.1% xylene cyanol, 1.0% sodium dodecyl sulfate, 0.1% bromophenol blue). The products were visualized by agarose gel electrophoresis.

Agarose gels were prepared by dissolving 0.7 % (w/v) agarose in TAE buffer (40 mM Tris, 20 mM acetic acid, 50 mM EDTA) by heating. Ethidium bromide was added to the gel at a concentration of 0.5 $\mu\text{g mL}^{-1}$. Gels were submerged in TAE buffer and loaded with the DNA samples and a DNA ladder. A 100 V current was applied to the gel for approximately one hour.

Polymerase chain reaction (PCR) was preformed using the NEB Q5 polymerase in the provided Q5 reaction buffer. The manufacturer's recommended reaction component concentration and thermocycler conditions were used.

PLASMID CONSTRUCTION

Plasmids pMG2.109, pMG2.128, and pMG2.208 were prepared from templates pUC-IDTamp+aroG1, pUC-IDTamp+aroG1, and pKD17.037, respectively by the Gibson Assembly method using NEB's HiFi DNA Assembly master mix. The plasmids pMG2.109 and pMG2.128 were prepared using the vector pET-15b linearized with BamHI and NdeI. The plasmid pMG2.208 was prepared using the vector pET-21(a)+ linearized with NdeI and XhoI. The linearized vector was isolated from the restriction digest reaction using a DNA clean-up kit. Gene fragments were amplified by PCR with primers that added a TEV protease recognition site and contained regions homologous to both the vector and the gene fragment (Table 14). The PCR-amplified gene fragments were purified by agarose gel electrophoresis prior to additional processing.

The linearized vector and PCR-amplified gene fragment were incubated at 50 °C with the HiFi DNA Assembly master mix for 15 minutes at a picomolar ratio of 1:2. High efficiency NEB5 α competent cells were transformed with 2 μL of the reaction product. Transformants were plated onto LB media supplemented with the appropriate antibiotic and incubated at 37 °C overnight. Plasmid DNA was isolated from 5 mL cultures of 5 to 10 single colonies, and DNA was analyzed

by restriction digest. Large-scale plasmid purifications were performed on two colonies containing insert-carrying vectors. The isolated plasmids were confirmed to carry the insert with a restriction digest and were submitted to the MSU RTSF Genomics Core for Sanger sequencing.

Table 14. Primers for amplification of *aroG1*, *aroG2*, and *tktA* for Gibson Assembly.

Amplification of	Primer	Sequence
<i>aroG1</i>	Forward	5'- <u>TGGTGCCGCGCGGCAGCCA</u> GAAAACCTGTATTTTC AGATGAGTGACAAATTGCAGCAAG -3'
	Reverse	5'- <u>TCGGGCTTTGTTAGCAGCCG</u> TTAATGGTACAAACCC GAATC-3'
<i>aroG2</i>	Forward	5'- <u>TGCCGCGCGGCAGCCA</u> GAAAACCTGTATTTTCAGA TGTCTAATA AAGAACTGGATCAAT-3'
	Reverse	5'- <u>TCGGGCTTTGTTAGCAGCCG</u> TTAGACACGTGCAAGG AC-3'
<i>tktA</i>	Forward	5'- <u>CTTTAAGAAGGAGATATACAAT</u> GTCTCACGTAAAGA G-3'
	Reverse	5'- <u>AGTGGTGGTGGTGGTGGTGCAC</u> GCCCTGAAAATAC AGGTTTTCCAGCAGTTCTTTTGCTT -3'

*Underlined bases are regions homologous to the vector. Bolded bases encode for the modified TEV protease recognition site⁴ (in the case of *aroG1* and *aroG2*) or the unmodified TEV protease recognition site (in the case of *tktA*). Remaining bases are regions homologous to the gene fragment.*

TRANSFORMATIONS

Electrocompetent *E. coli* cells were prepared using a modified procedure.⁵ Bacterial cells were cultured initially in 5 mL of LB medium at 37 °C, 300 rpm for 12-15 h. A 2 mL aliquot of the culture was then inoculated into 200 mL of 2xYT medium and cultured at 37 °C, 300 rpm to an OD₆₀₀ between 0.5 and 0.7. The culture was chilled on ice for 30 minutes. All following solutions were kept on ice. Chilled cells were collected by centrifugation (2,700 x g, 5 min, 4 °C). The supernatant was removed, and the cell pellet was resuspended in 200 mL of cold, sterile water. The cells were harvested by centrifugation (2,700 x g, 10 min, 4 °C) and the supernatant was

removed. The wash step was repeated with 100 mL of cold, sterile water. Cells were harvested by centrifugation (2,700 x g, 10 min, 4 °C) and the supernatant was removed. Finally, the cell pellet was resuspended in 100 mL of cold, sterile 10 % (v/v) glycerol. The cells were harvested by centrifugation (2,700 x g, 10 min, 4 °C) and the supernatant was removed. The cell pellet was resuspended in 500 µL of cold, sterile 10 % (v/v) glycerol and were aliquoted (50 µL) into microfuge tubes, flash frozen in liquid nitrogen, and stored at -80 °C.

Transformation of electrocompetent *E. coli* proceeded as follows: 0.5 µL of plasmid (10 ng µL⁻¹) was added to 50 µL of thawed, chilled electrocompetent cells. The sample was mixed with gentle tapping and was transferred to an ice-cold Bio-Rad Gene Pulse cuvette. A single pulse (2.5 kV, 25 µF, 200 Ω) was applied using a Bio-Rad Gene Pulser II Electroporator. The cells were then resuspended in 1.0 mL of SOC media and transferred to a sterile culture tube. Following a recovery period of 1 h at 37 °C with shaking (300 rpm), the transformants were selected by plating onto LB media supplemented with the appropriate antibiotic.

PROTEIN EXPRESSION AND PURIFICATION

The buffers used for protein purifications are: lysis buffer (20 mM sodium phosphate, pH 7.4, 500 mM sodium chloride, and 20 mM imidazole), buffer A (20 mM sodium phosphate, pH 7.4, and 500 mM sodium chloride), buffer B (20 mM sodium phosphate, pH 7.4, 500 mM sodium chloride, and 500 mM imidazole), TEV reaction buffer (50 mM sodium phosphate, pH 8.0, 0.5 mM EDTA, and 1.0 mM DTT), TktA storage buffer (100 mM potassium phosphate, pH 7.4 and 1.0 mM MgCl₂), and DAHPS storage buffer (10 mM BTP, pH 7.5, 0.05 mM CoCl₂, and 10 % (v/v) glycerol). Buffers A and B were used simultaneously by the FPLC. Concentrations of imidazole used during an FPLC run are denoted as a percentage of buffer B.

TktA-His₆ overexpression utilized BL21(DE3)/pMG2.208. The strain was cultured in 1 L of LB containing Amp at 37 °C, 300 rpm to an OD₆₀₀ of 0.5 to 0.7. Expression was induced by addition of 1.0 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and culturing was resumed at 30

°C, 300 rpm for an additional 5-6 hours. Cells were collected by centrifugation (7,200 x g, 15 min, 4 °C) and stored at -20 °C (<1 month). Cells were resuspended in 2 mL of lysis buffer per gram of wet cell weight. Cells were lysed with two passages through a SLM Aminco French Pressure Cell Press (18,000 psi) and cell lysate was clarified by centrifugation (38,500 x g, 30 min, 4 °C). Clarified lysate was filtered through a 0.45 µm syringe filter. A nickel affinity column was pre-equilibrated with 4% buffer B and then the clarified lysate was loaded onto the column. The unbound proteins were eluted from the nickel affinity column by 14 column volumes of 4% buffer B. TktA-His₆ was eluted from the nickel affinity column by 5 column volumes of 30% buffer B. Eluate fractions that absorbed at 280 nm were pooled and concentrated. The concentrated protein sample was desalted and exchanged into TktA storage buffer. A 100 µL sample of TktA-His₆ was diluted 10-fold with 6 M guanidine HCl and the absorbance at 280 nm was measured to calculate the protein concentration ($MW_{\text{TktA-His6}} = 73.9 \text{ kDa}$, $\epsilon_{\text{TktA-His6}} = 93.3 \text{ mM}^{-1}\text{cm}^{-1}$).

The expression and purification of His₇-TEV protease was performed in the same manner as TktA-His₆ described above with the following modifications. The strain BL21-CodonPlus(DE3)-RP/pMHTΔ238 was used to overexpress His₇-TEV protease. His₇-TEV protease was eluted from the nickel affinity column using 50% buffer B. The concentration of His₇-TEV protease was measured using the following constants: $MW_{\text{His7-TEV}} = 32.5 \text{ kDa}$, $\epsilon_{\text{His7-TEV}} = 32.3 \text{ mM}^{-1}\text{cm}^{-1}$.

His₆-AroG1 overexpression utilized BL21(DE3)/pMG2.109. This strain was cultured in 1 L of LB containing Amp at 37 °C, 300 rpm to an OD₆₀₀ of 0.5 to 0.7. Expression was induced by addition of 1.0 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and culturing was resumed at 30 °C, 300 rpm for an additional 10-12 hours. Cell lysis and nickel affinity column equilibration was carried out as described previously. His₆-AroG1 was eluted from the nickel affinity column using 10 column volumes of 70% buffer B. Eluate fractions that absorbed at 280 nm were pooled and concentrated as described previously. The concentrated protein sample was desalted and exchanged into TEV reaction buffer. A 100 µL sample of His₆-AroG1 was diluted 10-fold with 6 M guanidine HCl and the absorbance at 280 nm was measured to calculate the protein

concentration ($MW_{\text{His6-AroG1}} = 33.1 \text{ kDa}$, $\epsilon_{\text{His6-AroG1}} = 12.10 \text{ mM}^{-1}\text{cm}^{-1}$). TEV protease was added to the sample at a ratio of 1:50 (TEV protease: His₆-AroG1). The reaction mixture was then incubated at 30 °C for five hours. The reaction mixture was subsequently passed through a 0.45 μm syringe filter and reloaded onto a nickel affinity column pre-equilibrated with 4% buffer B. AroG1 eluted in the flow through. Flow through fractions that absorbed at 280 nm were pooled and concentrated as described above. AroG1 was exchanged into the DAHPS storage buffer and the concentration was measured as described above ($MW_{\text{AroG1}} = 30.1 \text{ kDa}$, $\epsilon_{\text{AroG1}} = 10.81 \text{ mM}^{-1}\text{cm}^{-1}$).

The expression and purification of AroG2 was performed in the same manner as AroG1 described above with the following modifications. The strain BL21(DE3)/pMG2.128 was used to overexpress His₆-AroG2. His₆-AroG2 was eluted from the nickel affinity column using 10 column volumes of 30% buffer B. The concentration of His₆-AroG2 and AroG2 were measured using the following constants: $MW_{\text{His6-AroG2}} = 42.9 \text{ kDa}$, $\epsilon_{\text{His6-AroG2}} = 12.80 \text{ mM}^{-1}\text{cm}^{-1}$, $MW_{\text{AroG2}} = 40.0 \text{ kDa}$, $\epsilon_{\text{AroG2}} = 11.52 \text{ mM}^{-1}\text{cm}^{-1}$.

Protein samples for SDS-PAGE analysis were quantified using the Bradford Assay.⁶ Protein samples (100 μL) were diluted with 5 mL of Bradford reagent (Bio-Rad). The solution was incubated at room temperature for five minutes and then the absorbance at 595 nm was measured. The concentration of protein was calculated using a calibration curve prepared from bovine serum albumin. SDS-PAGE samples were prepared by diluting 25 μL of a 300 ng/ μL protein sample with 25 μL of 2X Laemmli sample buffer (0.125 M Tris-HCl pH 6.8, 4.0% (w/v) sodium dodecyl sulfate, 20% (v/v) glycerol, 0.005% (w/v) bromophenol blue). Samples were vortexed and then incubated at 95 °C for five minutes. Pre-cast gels were loaded with 25 μL of each sample. The gel was submerged in SDS-PAGE running buffer containing 1.0% sodium dodecyl sulfate, 192 mM L-glycine, and 25 mM Tris. A constant voltage of 200 V was applied to the gel for 25 minutes. The gel was then rinsed with water, removed from the casing, and stained with a solution containing 50% (v/v) methanol, 10% (v/v) acetic acid, and 0.025% (w/v) Coomassie

blue for two hours. The gels were rinsed with water and then destained with a solution of 5.0% (v/v) methanol and 7.5% (v/v) acetic acid overnight.

ENZYME ASSAYS

The transketolase (TktA) assay was used to quantify E4P. This coupled enzyme assay followed the production of NADPH ($\epsilon = 6.22 \text{ mM}^{-1}\text{cm}^{-1}$) at 340 nm. The assay was carried out in 150 mM triethanolamine, pH 7.6 and contained 5 mM MgCl_2 , 0.4 mM NADP^+ , 0.2 mM thiamine pyrophosphate, 0.4 mM β -hydroxypyruvic acid, 10 U phosphoglucose isomerase, 3 U glucose-6-phosphate dehydrogenase, varying amounts of E4P, and TktA- His_6 . Figure 33 shows the reactions occurring during this assay.

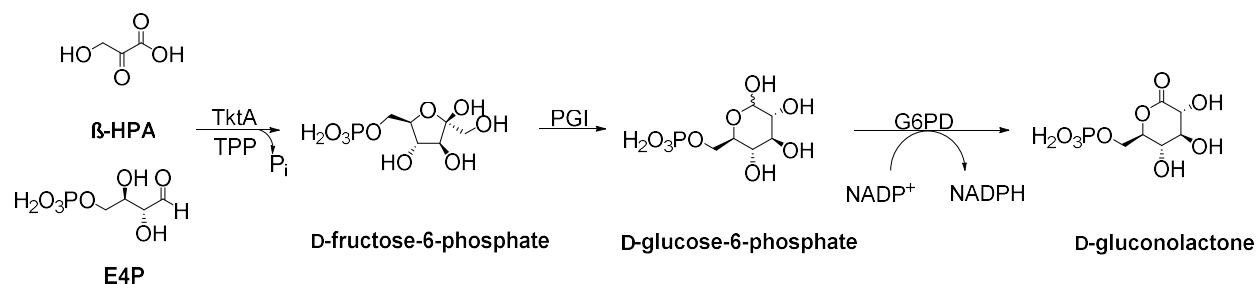


Figure 34. The TktA assay. Production of NADPH is monitored at 340 nm.

AroG1 and AroG2 activity was measured using a continuous assay that followed the loss of PEP ($\epsilon = 2.84 \text{ mM}^{-1} \text{ cm}^{-1}$) at 232 nm.⁷ Assays were monitored for 30-60 seconds. Initial rates were calculated from linear portions of the absorbance slope (Equation 1). Specific activities and velocities were calculated from the initial rates (Equation 2 and Equation 3, respectively).

Equation 1: $\text{rate (AU/min)} = |\text{Slope (mAU/min)}| \cdot 1000$

Equation 2: $\text{specific activity } (\mu\text{M/min}\cdot\text{mg}) = \text{rate} \cdot \left(\frac{1}{[\text{enzyme}] \text{ (mg/mL)}} \right) \cdot \left(\frac{1}{\epsilon} \right) \cdot \left(\frac{V_f}{V_i} \right) \cdot 1000$

Equation 3: $\text{Velocity } (\mu\text{M/min}) = \text{rate} / \left(\frac{\epsilon}{1,000,000} \right)$

The effect of pH on DAHP synthase activity was evaluated between pH 6.3 and 8.6 at room temperature with 1.6 mM PEP, 3.0 mM E4P, 0.05 mg of AroG1 or AroG2 in 50 mM bistrispropane (BTP). The effect of temperature on DAHP synthase activity was evaluated between 35 °C and 60 °C with 1.6 mM PEP, 3.0 mM E4P, 0.05 mg of AroG1 or AroG2 in 50 mM BTP pH 7.2 ± 0.1. The assay mixture (except E4P) was incubated at the assay temperature for five minutes, and the reaction was initiated with the addition of E4P.

Thermostability was evaluated by incubating AroG1 and AroG2 at 50 °C and periodically measuring the activity. Assay reactions contained 0.05 mg of AroG1 or AroG2, 3.0 mM E4P, 1.6 mM PEP in 50 mM BTP pH 7.2 ± 0.1.

Kinetic assays were conducted with 0.05 mg of AroG1 or AroG2 in 50 mM BTP buffer pH 7.2 ± 0.1 at 50 °C. For PEP measurements, the E4P concentration was held steady at 3.0 mM and the PEP concentration was varied from 0.04 mM to 1.8 mM. For E4P measurements, the PEP concentration was held steady at 1.6 mM and the E4P concentration was varied from 0.06 mM to 3.6 mM. The assay mixture (except E4P) was incubated at 50 °C for five minutes, and the reaction was initiated with the addition of E4P. The Michaelis-Menten constant (K_m) was calculated using the Michaelis-Menten function in Origin Pro 9.0.

Feedback inhibition assays were conducted with 0.05 mg of AroG1 or AroG2, 3.0 mM E4P, 1.6 mM PEP, and 0.25 mM potential inhibitor in 50 mM BTP buffer pH 7.2 ± 0.1 at 50 °C. The potential inhibitors were L-phenylalanine, L-tyrosine, L-tryptophan, chorismate, and prephenate. The assay mixture (except E4P) was incubated at 50 °C for five minutes, and the reaction was initiated with the addition of E4P.

E4P SYNTHESIS AND QUANTIFICATION

D-Erythrose-4-phosphate (E4P) was synthesized from glucose-6-phosphate (G6P).⁸ An addition funnel was attached to a 3-neck round bottom flask and flushed with nitrogen gas. Glacial acetic acid (500 mL) and 2.0 mmol (1 equivalent) of G6P were added to the flask. To the addition funnel, 80 mL of glacial acetic acid, 192 μ L (1.8 equivalents) of concentrated sulfuric acid, and 1.5 g (1.7 equivalents) of lead tetraacetate were added. The contents of the addition funnel were added to the round bottom over the course of an hour with stirring. The resulting solution contained lead sulfate which precipitated. The lead sulfate was removed by two passages through a celite pad. The filtrate was then concentrated to 80 mL using a Buchi rotary evaporator. The solution was diluted 2-fold with water and concentrated to 80 mL. This step was repeated an additional four times. The final 80 mL solution was passed through a 50 mL Dowex 50WX column that had been equilibrated with 200 mL of 2 N HCl and 500 mL of water. 300 mL of water was used to elute E4P from the column. Eluate was collected from the addition of the E4P solution and stored at room temperature in an amber bottle. E4P was quantified using the transketolase assay described above. The concentration of E4P was calculated using Equation 4.

Equation 4:
$$[\text{E4P}] \text{ (mM)} = \left(\frac{\Delta\text{OD}_{340}}{\epsilon} \right) \cdot \left(\frac{V_f}{V_i} \right)$$

REFERENCES

REFERENCES

1. Schendel, F. J.; Bremmon, C. E.; Flickinger, M. C.; Guettler, M.; Hanson, R. S. L-Lysine Production at 50 Degrees C by Mutants of a Newly Isolated and Characterized Methylophilic *Bacillus* sp. *Appl. Environ. Microbiol.* **1990**, *56* (4), 963–970.
2. Miller, J. H. Experiments in Molecular Genetics; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1972.
3. Jakobsen, Ø. M.; Benichou, A.; Flickinger, M. C.; Valla, S.; Ellingsen, T. E.; Brautaset, T. Upregulated Transcription of Plasmid and Chromosomal Ribulose Monophosphate Pathway Genes Is Critical for Methanol Assimilation Rate and Methanol Tolerance in the Methylophilic Bacterium *Bacillus methanolicus*. *J. Bacteriol.* **2006**, *188* (8), 3063–3072.
4. Kapust, R. B.; Tözsér, J.; Copeland, T. D.; Waugh, D. S. The P1' Specificity of Tobacco Etch Virus Protease. *Biochem. Biophys. Res. Commun.* **2002**, *294* (5), 949–955.
5. Sambrook, J.; Russell, D. W. Molecular Cloning: A Laboratory Manual 3rd ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 2001.
6. Bradford, M. M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **1976**, *72* (1), 248–254.
7. Wu, J.; Sheflyan, G. Y.; Woodard, R. W. Bacillus Subtilis 3-Deoxy-D-Arabinose-Heptulosonate 7-Phosphate Synthase Revisited: Resolution of Two Long-Standing Enigmas. *Biochem. J* **2005**, *390* (Pt 2), 583–590.
8. Sieben, A. S.; Perlin, A. S.; Simpson, F. J. An Improved Preparative Method for D-Erythrose-4-phosphate. *Can. J. Biochem.* **1966**, *44* (6), 663–669.