Engineering *Saccharomyces cerevisiae* for conversion of methanol into biomass

A Thesis

in

The Department

of

Biology

Presented in Partial Fulfillment of the Requirements

For the Degree of Master of Arts and Sciences (Biology)

Concordia University

Montreal, Quebec, Canada

December 2018

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Alex Vallerand-Legault

Entitled: Engineering *Saccharomyces cerevisiae* for conversion of methanol into biomass

And submitted in partial fulfillment of the requirements for the degree of

Master of Arts and Science (Biology)

Complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____Chair and Examiner

Dr. David Walsh

_____Examiner

Dr. Malcolm Whiteway

_____Examiner

Dr. David Kwan

_____Supervisor

Dr. Vincent Martin

Approved by _____

Chair of Department or Graduate Program Director

_____ 2018 _____

Date Dean of Faculty

Abstract

Engineering *Saccharomyces cerevisiae* for conversion of methanol into biomass

Alex Vallerand-Legault

The effect of greenhouse gases on the environment has been well documented. In effort to mitigate the addition of greenhouse gases to the environment, there has been a collective effort to reduce our dependency on fossil fuel. One of the many ways to reduce this dependency is to find alternative methods for producing high-valued chemicals that are currently derived from fossil fuels. Due to the increasing bioengineering tools made available, this branch of science has greatly evolved and has facilitated the production of different high-valued chemical commodities in model organisms such as *Escherichia coli* or *Saccharomyces cerevisiae*.

However, a great barrier to commercialization of these high-valued bioproducts is their high production cost. Thus, in this thesis we focus on reducing the cost of the carbon feedstock by engineering yeast *Saccharomyces cerevisiae* to utilize methanol instead of glucose, the preferred but expensive carbon source. A three gene pathway found in *Bacillus methanolicus* was introduced in a strain of *S. cerevisiae*, enabling it to assimilate methanol. Overall, we were able to engineer *S. cerevisiae* to assimilate methanol and observed a 13.54% increase in biomass synthesis in the presence of 5mM methanol. We show the ability of our engineered strains to convert methanol to amino acids using ^{13}C-labeled methanol. As well, we show a 2.2 to 3.7-fold increase in ^{13}C-labeled amino acids when the media was supplemented with yeast extract.

Acknowledgements

First and foremost, I would like to acknowledge my supervisor, Dr. Vincent Martin, for giving me this opportunity. He provided me with the necessary tools to get me where I am today, but also, he went above and beyond to provide me with the learning experience I was looking for and by being an amazing mentor. I'm truly thankful for this experience and the guidance you provided me.

A big thank you to my committee members Dr. Malcom Whiteway and Dr. David Walsh for their time, guidance and for helping me along the way.

I would also like to thank all the members of the Martin lab for their constant support and the time they have given me over the years. A special thanks to Michael Pyne and Leanne Bourgeois for their amazing help and support throughout my journey in the lab.

Lastly, I would like to thank my family for their support over the years and specifically my mother Lynn Legault for pushing me and helping me write this amazing thesis! Last but not least, a huge thank you to Kimberly Emond for being by my side the entire way and for being so understanding when I had to spend late nights and weekends in the lab.

Table of contents

List of Tables

List of Figures

List of Equations

List of Abbreviations

AMP:	Adenosine Monophosphate
ADP:	Adenosine Diphosphate
ATP:	Adenosine Triphosphate
Pi:	Phosphate
RNA	Ribonucleic acid
DNA:	Deoxyribonucleic acid
KanMX:	Kanamycine
MDH.M:	Methanol dehydrogenase from *Bacillus methanolicus* MGA3
MDH.P:	Methanol dehydrogenase from *Bacillus methanolicus* PB1
HPS.M:	3-Hexulose-6-phosphate synthase from *Bacillus methanolicus* MGA3
HPS.P:	3-Hexulose-6-phosphate synthase from *Bacillus methanolicus* PB1
PHI.M:	3-Hexulose-6-phosphate isomerase from *Bacillus methanolicus* MGA3
PHI.P:	3-Hexulose-6-phosphate isomerase from *Bacillus methanolicus* PB1
ACT.M:	Activator protein from *Bacillus methanolicus* MGA3
ACT.P:	Activator protein from *Bacillus methanolicus* PB1
AOX:	Alcohol oxidase
MO:	Methanol Oxidase
ADH:	Alcohol dehydrogenase
DAS:	Dihydroxyacetone synthase

CAT: Catalase

FrmA: S-(hydroxymethyl)glutathione dehydrogenase or formaldehyde dehydrogenase

XPK: Phosphoketolase

FPK: Fructose phosphate synthase

H6P: 3-hexulose-6-phosphate

F6P: Fructose-6-phosphate

Xu5P: Xylulose-5-phosphate

Ru5P: Ribulose-5-phosphate

FALD: Formaldehyde

MEOH: Methanol

CO: Carbon monoxide

CO_2: Carbon Dioxide

ddH_2O: Double distilled water

PQQ: Pyrroloquinoline quinone

NAD(H): Nicotinamide adenine dinucleotide (NAD) + hydrogen

NADP(H): *Nicotinamide adenine dinucleotide phosphate*

FAD: *Flavin adenine dinucleotide*

RuMP: Ribulose MonoPhosphate

XuMP: Xylulose MonoPhosphate

MCC: Methanol condensation cycle

NOG: Non-oxidative glycolysis

YPD: Yeast Peptone Dextrose

ETS: Electron Transport System

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

Cas9: CRISPR associated protein 9

K_M: Mechaelis-Menten constant

K_{CAT}: Catalytic activity of the enzyme

V_{MAX}: Maximum velocity of the enzyme

1 THESIS STATEMENT

The present study aims to engineer *Saccharomyces cerevisiae* to assimilate methanol as a carbon and energy source.

An important issue of today is climate change and the many ways the global population aims to halt its accelerated development. The reduction of greenhouse gases (GHG) is a fundamental strategy to reduce global warming. Gasification is a cost-effective process that efficiently transforms any biomass, such as landfill waste or municipal waste into methanol [1].

Reduction of fossil fuel dependency via large-scale biotechnological production of valuable chemicals from *S. cerevisiae*, is one way of reducing GHGs. However, these biotechnological processes depend on expensive sugars such as glucose to serve as the carbon and energy source. Around 60% of the total cost of production of a target bioproduct comes from glucose [2]. Thus, there has been a recent shift towards identifying alternative carbon sources to minimize production costs.

A large diversity of microorganisms have evolved the necessary functions to utilize methanol as a carbon and energy source [3]. Since, methanol can be economically produced from biomass in large quantities, substituting glucose for methanol as the carbon source can reduce production cost and help bring these bioproducts to market. Thus, we set out to engineer a tractable industrial host, *S. cerevisiae*, to assimilate methanol into biomass.

We have identified one promising pathway, found in the gram-positive methylotrophic bacterium, such as *Bacillus methanolicus*, that could be introduced into *S. cerevisiae*, thus providing this budding yeast the capacity to utilize methanol.

2 INTRODUCTION

The reduction of GHGs is an important issue due to the accelerated climate change and its adverse effects on the world's ecosystems [4]. Additionally, concerns over fossil fuel supply and prices have led the research community to find alternative routes to commodity chemical production and energy generation [5][6]. Many countries have mandated bioenergy processes to support the production of fuels from renewable sources. A shift toward a global bioeconomy would decrease fossil fuel dependency and reduce GHG emissions [7].

The first step in bioprocess development is choosing the feedstock that will provide carbon and energy to drive synthesis of a target bioproduct. Renewable feedstocks can be acquired from cellulosic biomass, animal manure, and other organic sources [5]. Cellulosic biomass is regarded as a superior feedstock due to its abundance and renewability. Although cellulosic biomass is projected to supply much of the future's energy need, it is currently the fourth largest energy source after coal, oil, and natural gas [8][9].

Currently, the application of renewable feedstocks is focused on electricity and biofuel production. However, producing electricity from coal and fuel from fossil fuels remains more economical [5]. Hence, rather than producing electrical energy from renewable feedstocks, effort should be focused on the production of higher-value chemicals, such as those currently produced from petrochemical derivatives [10][11]. In this regard, renewable feedstocks must first be converted to a source of carbon that microorganisms can use. One way this can be achieved is by the conversion of renewable feedstocks to biosyngas via a gasification process [5]. Finally, biosyngas can be economically converted to methanol, a carbon and energy source that microorganisms can harness to produce higher-valued chemicals.

2.1 Biosyngas Production

Depending on the method used, biogas or biosyngas can be produced from renewable feedstocks (Figure 2-1). Anaerobic digestion (AD) of a limited number of renewable feedstocks, such as animal manure, organic waste, and energy crops, although lignin cannot be digested, produces biogas, a composition of methane (CH_4) and carbon dioxide (CO_2) [12][16]. On the other hand,

gasification of any renewable feedstock produces biosyngas, a mixture of carbon monoxide (CO), CO_2, and hydrogen (H_2) [16]. However [12]. Additionally, coal refinery residues, municipal waste, and biogas can be converted to biosyngas [9]. Regardless of the feedstock used during the gasification process, the same biosyngas is produced, thus making this process attractive and universal.

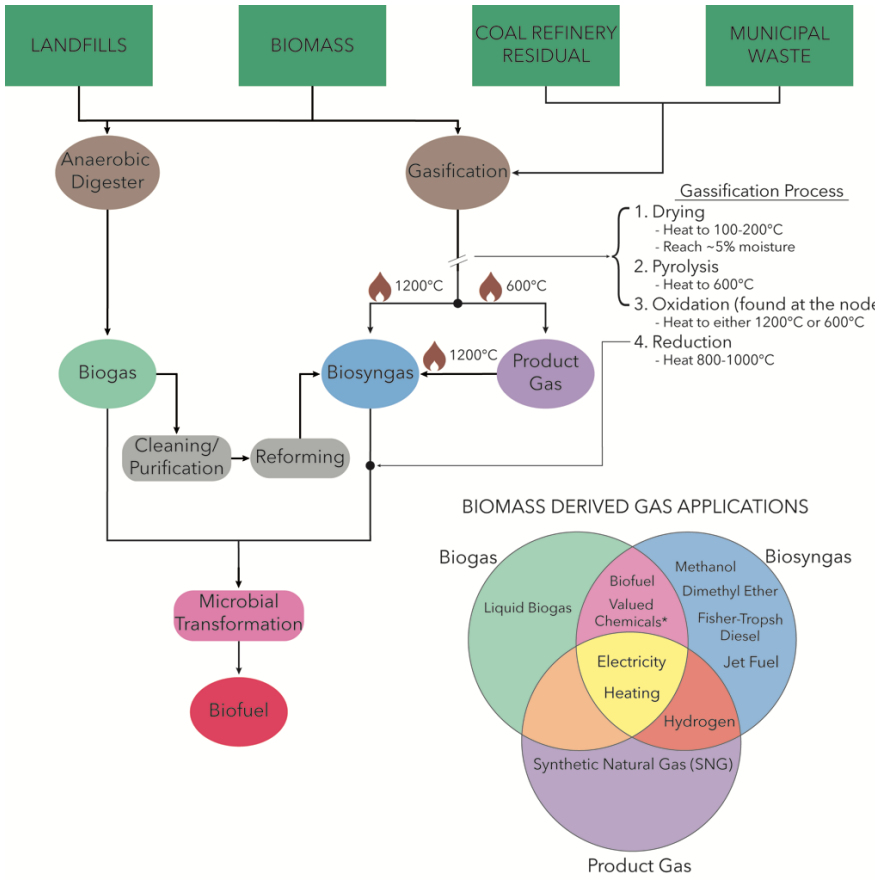


Figure 2-1: Renewable feedstock conversion to biofuel. (A) Conversion route of renewable feedstocks to methanol. (B) Current use of biomass derived gas. Figure adapted from L. Yang et al. [5].

Gasification of biomass is a large-scale process that converts biomass to biosyngas via a four-stage process: drying, pyrolysis, oxidation, and reduction [13]. In the first step, the renewable feedstock is heated between 100-200°C to achieve a ~5% moisture content [14]. Then, the feedstock is gasified through a two-stage endothermic reaction process that includes pyrolysis and oxidation [14]. During pyrolysis, the feedstock is heated above ~600°C in the absence of air, releasing carbon monoxide, carbon dioxide, hydrogen, tar and hydrocarbons [15]. This process also produces the byproducts char (fixed carbon) and ash [15]. After pyrolysis, char is heated to high temperatures, above 1200°C, releasing carbon monoxide and hydrogen (Equation 2-1 and Equation 2-2)

[15][10]. Finally, during the reduction step, products of gasification are heated near 800-1000°C (Equation 2-3 and Equation 2-4) [15][10].

Table 2-1: Oxidation and reduction reactions during gasification process

Equation 2-1: Partial oxidation	$C + 0.5O_2 \leftrightarrow CO$	$\Delta H_{vap} = -268 \frac{MJ}{kg\ mol}, \Delta G° = -\frac{151kJ}{mol}$	(1)
Equation 2-2: Water gas reaction	$C + H_2O \leftrightarrow CO + H_2$	$\Delta H_{vap} = +118 \frac{MJ}{kg\ mol}, \Delta G° = -\frac{100kJ}{mol}$	(2)
Equation 2-3: Water gas-shift reaction	$CO + H_2O \leftrightarrow CO_2 + H_2$	$\Delta H_{vap} = -42 \frac{MJ}{kg\ mol}, \Delta G° = -\frac{20kJ}{mol}$	(3)
Equation 2-4: Methane formation	$CO + 3H_2 \leftrightarrow CH_4 + H_2O$	$\Delta H_{vap} = -88 \frac{MJ}{kg\ mol}, \Delta G° = -\frac{151kJ}{mol}$	(4)

Overall, this process converts biomass into biosyngas however, biosyngas is an unattractive carbon source because of its low aqueous solubility [16]. A more favoured carbon source among microorganisms is methanol, a liquid at room temperature and a biosyngas derivative. Methanol production was first produced from wood distillation, which is an inefficient process. Presently, hydrogenation of CO and/or CO_2 efficiently produces methanol [16]. Thus, via reversal of the water-gas shift reactions, biosyngas is converted to methanol (Equation 2-5 and Equation 2-6) [16]. The reverse water-gas shift reactions are exothermic, necessitating milder processing conditions leading to lower production costs [16]. As a carbon source, methanol is assimilated by many methylotrophic microorganisms such as *Pichia pastoris* and *B. methanolicus*. Thus, methanol represents an attractive feedstock for upgrading biosyngas to value-added bioproducts.

Table 2-2: Hydrogenation reactions for the conversion of biosyngas to methanol

Equation 2-5: CO Hydrogenation	$CO + 2H_2 \leftrightarrow CH_3OH$	$\Delta H_{298} = -\frac{91kJ}{mol}$	(5)
Equation 2-6: CO_2 Hydrogenation	$CO_2 + 3H_2 \leftrightarrow CH_3OH + H_2O$	$\Delta H_{298} = -\frac{49kJ}{mol}$	(6)

2.2 Methylotrophic fermentation

The microbial conversion of methanol depends on specific enzymes only found in methylotrophic species, microorganisms capable of utilizing C_1 carbon sources, such as methanol, for biosynthesis. In the last 50 years, research on methanol-utilizing microorganisms has focused on the elucidation of enzymes and pathways required for microbial methanol assimilation [17]. More recently, research on methylotrophic microbes has focused on the mechanisms involved in converting methanol into value-added products, such as amino acids, biofuels, and biopolymers [18].

2.2.1 Methanol assimilation

It has been established that a minimum of three enzymes are necessary for microorganisms to assimilate methanol [17]. The first step is the conversion of methanol to formaldehyde, catalyzed by either alcohol oxidases (AOX) (Table 2-3), or alcohol dehydrogenases (ADH) (Table 2-4 and Table 2-5).

Table 2-3: AOX characteristic comparison.

Characteristics	Short-chain alcohol oxidase	Long-chain alcohol oxidase	Secondary alcohol oxidase	Aromatic alcohol oxidase
Source organism	Yeast, filamentous fungi	Yeast, filamentous fungi, plant	Bacteria	Fungi, insect, mollusk
Localization	Peroxisome	Microsome, glyoxisome	Extracellular	Extracellular
Protein structure	Homooctomer (65-80kDa/subunit)	Dimeric (70-94kDa/subunit)	Monomeric (40-85kDa)	Monomeric (66-80kDa)
Primary substrate	Methanol, ethanol	Long-chain alcohols	Polyvinyl alcohol, 2º alcohols: chain length C4-C12	Benzyl-, 3- and 4-methoxy benzyl-, 2,4-dimethoxybenzyl alcohol, veratryl alcohol, cinnamyl alcohol, naphthyl methanol
Optimum pH range	~6.0-8.5	~6.0-9.5	~6.5-10	5-8.5
Optimum temperature range	25-30ºC, 37 ºC	20-30 ºC	40-50 ºC	25-45 ºC

Methylotrophic yeasts convert methanol to formaldehyde using a subclass of AOX, known as methanol oxidase (MO). The AOXs have an FAD, a flavoprotein, as a cofactor, which is associated with the reaction center and is involved in accepting the hydride ion from methanol and

transferring it to molecular oxygen, producing hydrogen peroxide as a by product [19]. The enzymes involved in assimilating methanol are found in the peroxisome [20]. Post-methanol induction leads to the proliferation of peroxisomes, which can account for 80% of the cytoplasmic space [20].

Table 2-4: XoxF-MDH and MxaFI-MDH comparison.

Properties	**MxaFI-MDH**			**XoxF-MDH**		
Catalytic properties	V_{max} (%)	K_m(mM)	k_{eff} ($s^{-1}mM^{-1}$)	V_{max} (%)	K_m(mM)	k_{eff} ($s^{-1}mM^{-1}$)
Methanol	100	0.014-0.45	13-800	100	0.0008-0.29	3-11,600
Ethanol	76-114	0.016-3.58	2-430	90-160	0.003-0.014	4-3,100
1-propanol	43-106	0.036-3.69	1-120	90-100	0.007	1,330
1-butanol	42-89	0.040-6.38	2-100	70-100	0.006	1,550
Formaldehyde	78-81	0.1-3.300	1-34	93-100	0.007-0.065	0.5-1,330
Subunit	$\alpha_2\beta_2$ or α_2			α_2		
Molar mass (kDa)	125-148.5			126-127		

*Table was adapted from Krog et al. [18] [22].

Methylotrophic bacteria have evolved with subclasses of ADHs, known as methanol dehydrogenase (MDH) [21]. There exist two major different subclasses of MDHs: PQQ-dependent MDHs, and NAD(P)-dependent MDHs [21]. Of the PQQ-dependent MDHs, two different sub-types have been characterized, XoxF-type and MxaFI-type [21]. Both isozymes are found in the periplasm of gram-negative bacteria, and are linked to the electron transport system (ETS) [21]. The PQQ (PyrroloQuinoline Quinone) accepts the electron from methanol and funnels it down the ETS to generate ATP [21]. XoxF-type MDH requires the addition of lanthanide ion, while MxaFI-MDH requires calcium ion for their activities [21][22]. PQQ-dependent methanol dehydrogenases require the assistance of a minimum of 11 other proteins and cytochrome c [15][21].

NAD(P)-dependent MDHs are cytoplasmic enzymes found in gram-positive bacteria and make use of NAD(P) as the cofactor. The NAD(P) accepts the hydride ion during the catalysis of the alcohol and generates NAD(P)H, which can donate its electron pair to the ETS [19]. As a result, a microorganism that assimilates methanol via the NAD(P)-dependent MDH and/or the PQQ-dependent MDH has an energetic advantage [19]. Since the PQQ-dependent MDHs are membrane bound enzymes, requires several accessory proteins, and requires a PQQ molecule, which is not natively synthesized by many model organisms, the NAD(P)-dependent MDHs have been regarded as better candidates for engineering methanol-utilizing microorganisms.

2.2.2 NAD(P)-dependent MDH from *Bacillus methanolicus*

The presence of cytosolic NAD(P)-dependent MDHs have been identified in *B. methanolicus* and other thermotolerant methanol-utilizing *Bacillus spp.* [18]. Gram-positive methylotrophic bacteria, of the family *B. methanolicus* C1, encodes an NAD(P)-dependent MDH located in the cytoplasm [18]. Two sub-species, *B. methanolicus* PB1 and *B. methanolicus* MGA3, each encode three NAD(P)-dependent MDHs (Table 2-5) [15][16]. Of the three different NAD(P)-dependent MDHs in the strain *B. methanolicus* MGA3, two are found on the chromosome ($MDH2^M$ and $MDH3^M$), while the third MDH (MDH^M) is found on an endogenous multi-copy plasmid, pBM19 [24]. Similarly, *B. methanolicus* PB1 also encodes two different NAD(P)-dependent MDHs found on the chromosome ($MDH1^P$ and $MDH2^P$), and a third (MDH^P) found on a naturally occurring multi-copy plasmid, pBM20 [23]. The plasmid-dependent methylotrophy is a widespread trait among the *B. methanolicus* strains [24]. An NAD(P)-dependent MDH has also been reported in *Bacillus stearothermophilus*. This enzyme shares about 20-23% sequence similarity with the MDHs of *B. methanolicus,* and has a Mechaelis-Menten constant (K_m) of 20mM and a V_{max} of 2.1U/mg [23][24].

Table 2-5: *B. methanolicus* NAD(P)-dependent MDH enzymatic activity.

Strains	**MDH^M**		**$MDH2^M$**		**$MDH3^M$**		**MDH^P**		**$MDH1^P$**		**$MDH2^P$**	
Kinetics	K_M	V_{Max}	K_M	V_{Max}	K_M	V_{Max}	K_M	V_{Max}	K_M	V_{Max}	K_M	V_{Max}
Methanol	170 ±20	0.06 ±0.002	360 ±30	0.09 ±0.003	200 ±70	0.07 ±0.005	220 ±30	0.03 ±0.001	170 ±60	0.015 ±0.001	330 ±0.05	0.08 ±0.004
Methanol + ACT	26 ±7.0	0.4 ±0.02	200 ±20	0.2 ±0.008	150 ±10	0.4 ±0.008	10 ±1.0	0.2 ±0.003	5 ±1.0	0.05 ±0.002	110 ±50	0.38 ±0.04
Formaldehyde	1.1 ±0.2	0.6 ±0.030	4.5 ±0.4	1.8 ±0.06	7.1 ±0.9	4.6 ±0.2	3.0 ±0.2	0.5 ±0.007	7 ±1.0	0.6 ±0.01	1.0 ±0.1	1.1 ±0.03
Methanol*		1.7		30		23		0.7		0		3
Methanol + ACT*		0.7		21.3		n.a		n.a		n.a		n.a

*Table was adapted from Krog et al. [18].
-FALD stands for formaldehyde.
-MeOH, MeOH + ACT and FALD are *in vitro* experiments. Their Km units are in mM and the Vmax units are in U/mg.
-MeOH* and MeOH + ACT* are *in vivo* experiments. Their Vmax units are in mU/mg.

The crystal structure of the *B. methanolicus* MDHs reveals that it is made up of ten identical subunits arranged in a sandwich of two pentagonal rings, displaying a Rossman-fold protein structure [27]. Each of the subunits have a tightly but non-covalently bound NAD(P), one Zn^{2+}-ion found at the active site and 1-2 Mg^{2+}-ions [28]. Even though the MDHs of *B. methanolicus* contains a Zn^{2+}-ion and can convert C1-C4 alcohols, it does not belong to the family of zinc-

containing NAD(P)-dependent alcohol dehydrogenases [29]. In fact, multiple sequence alignment studies have shown that the NAD(P)-dependent MDHs belong to the type III alcohol dehydrogenases family [29]. The identification of such decameric protein containing magnesium ions have been found in gram-positive methanol utilizing bacteria, the actinomycetes: *Acidomonas methanolica* and *Mycobacterium gastri* [29]. Additionally, other decameric alcohol dehydrogenases have been found in non-methylotrophic gram-positive bacteria [29]. Thus, evidence suggests a widespread distribution of the enzymes in the microbial world.

Kinetic studies suggest a ping-pong reaction mechanism for the NAD(P)-dependent MDHs. This is a two-step mechanism where first the electrons from methanol are transiently transferred to the $NAD(P)^+$ cofactor, which is then re-oxidized by a free $NAD(P)^+$ to form NAD(P)H[28]. *In vitro* substrate specificity experiment of all six MDHs found across the *B. methanolicus* species suggest lower methanol specificity in contrasts to higher alcohols: ethanol, butanol, propanol, isopropanol and 1,2-propanediol [18].

In vitro enzyme-kinetic analyses have shown that the MDHs of *B. methanolicus* can be stimulated up to 40-fold by an activator protein (ACT) [18]. ACT protein belongs to the Nudix hydrolase ribose pyrophosphatases (EC. 3.6.1.13) family and can hydrolyze ADP-ribose and NAD(P)H degradation products [18]. An *in vitro* activity study done by Dijkhuizen et al., replicated the stimulation of the ACT protein on MDH^{C1} with an S97G mutation for methanol and ethanol substrates [18]. Later, Vorholt et al. replicated the S97G mutant on the MDH^{M} but reported no *in vitro* and *in vivo* increase activity [17].

2.2.3 Phylogenetic relationship of the *Bacillus methanolicus* NAD(P)-dependent methanol dehydrogenases

The thermotolerant gram-positive bacteria, *B. methanolicus* MGA3 and PB1, each encode three MDH homologs (Figure 2-2)[18]. Using BLAST, the pairwise alignment of $MDH3^{M}$ and $MDH2^{M}$ revealed 96% sequence similarity, and 61% and 62% similarity to MDH^{M} respectively [30]. Pairwise alignment of MDH^{P} with $MDH1^{P}$ and with $MDH2^{P}$ revealed 93% and 60% sequence similarity respectively, while the pairwise alignment of $MDH1^{P}$ and $MDH1^{P}$ reveal 60% sequence similarity [30].

The phylogenetic relationship of the NAD(P)-dependent methanol dehydrogenases from *B. methanolicus* can be separated into two groups sharing high sequence similarity. Group 1 includes $MDH2^{M}$, $MDH3^{M}$, and $MDH3^{P}$, while group 2 includes MDH^{P}, $MDH1^{P}$, and MDH^{M} (Figure 2-2).

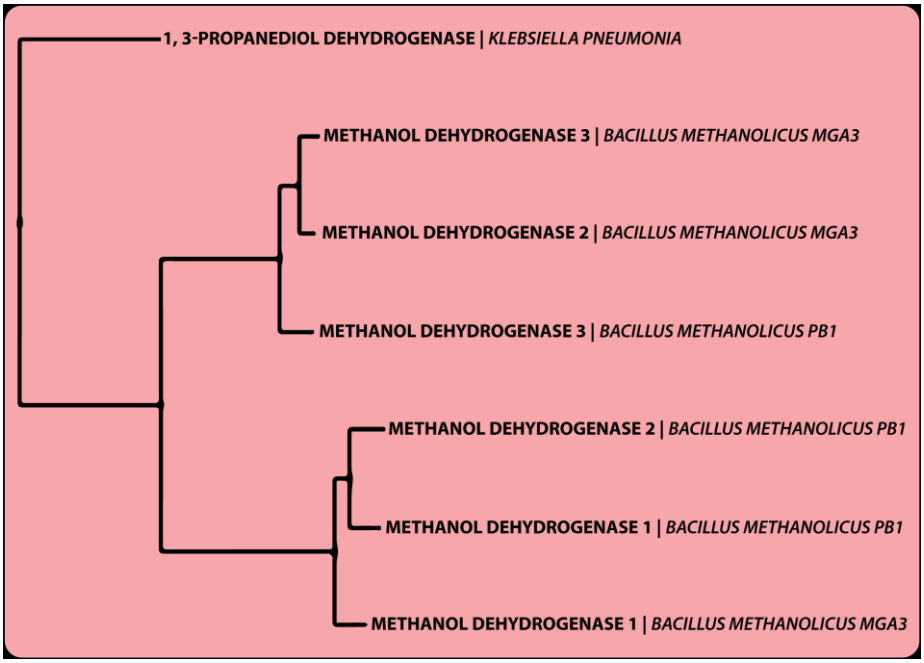


Figure 2-2: Phylogenetic relationship of NAD(P)-Dependent methanol dehydrogenase from *B. methanolicus*. Phylogenetic representation of NAD(P)-dependent MDHs from *B. methanolicus* PB1 and MGA3 in comparison to the most similar enzyme from the BLAST analysis against the PDB database. Figure was adapted from Krog et al. [16].

2.2.4 Formaldehyde assimilation pathway

Formaldehyde is almost a hundred times more toxic to yeast than methanol, therefore when formaldehyde is produced in the cell it must be quickly removed [31]. Formaldehyde can either be assimilated for biomass or dissimilated into CO_2 for energy production [32]. To assimilate formaldehyde, methylotrophic bacteria do so via the Serine cycle or the Ribulose MonoPhosphate cycle (RuMP), while methylotrophic yeast employ the Xylulose MonoPhosphate cycle (XuMP) [32]. The Serine cycle works by producing two molecules of 2-phosphoglycerate (2PG) from three molecules of formaldehyde and three molecules of glyoxylate [33]. One molecule of 2PG is recycled closing the cycle, while the second molecule is converted to 3-phosphoglycerate (3PG) (Figure 2-3) [33]. Studies by Goldberg et al. showed greater carbon assimilation efficiency in respect to ATP and NADH requirements for the RuMP cycle in comparison to the Serine cycle

[34]. Precedent work using four bacterial strains – two bacterial strains using the RuMP cycle and two bacterial strains using the Serine cycle – showed a maximum yield of 19.1g cell dry wt/mol and 13.5g cell dry wt/mol for the strains using the RuMP and Serine cycles respectively [34].

Two additional enzymes – 3-hexulose-6-phosphate synthase (HPS), and 3-hexulose-6-phosphate isomerase (PHI) – are associated with the RuMP cycle [33]. The first enzyme, HPS, catalyzes a bond between formaldehyde and ribulose-5-phosphate (Ru5P) to produce 3-hexulose-6-phosohate (H6P) [33]. The second enzyme, PHI, isomerizes H6P into fructose-6-phosphate (F6P) [33]. In a functional RuMP cycle, three assimilated methanol molecules produces one F6P molecule for biosynthesis and three Ru5P molecules are recycled [33]. The remaining two F6P molecules are used to replenish the Ru5P pool, keeping the cycle running (Figure 2-4). Methylotrophic yeast assimilates formaldehyde in the peroxisomes through the XuMP cycle [20]. In comparison to the

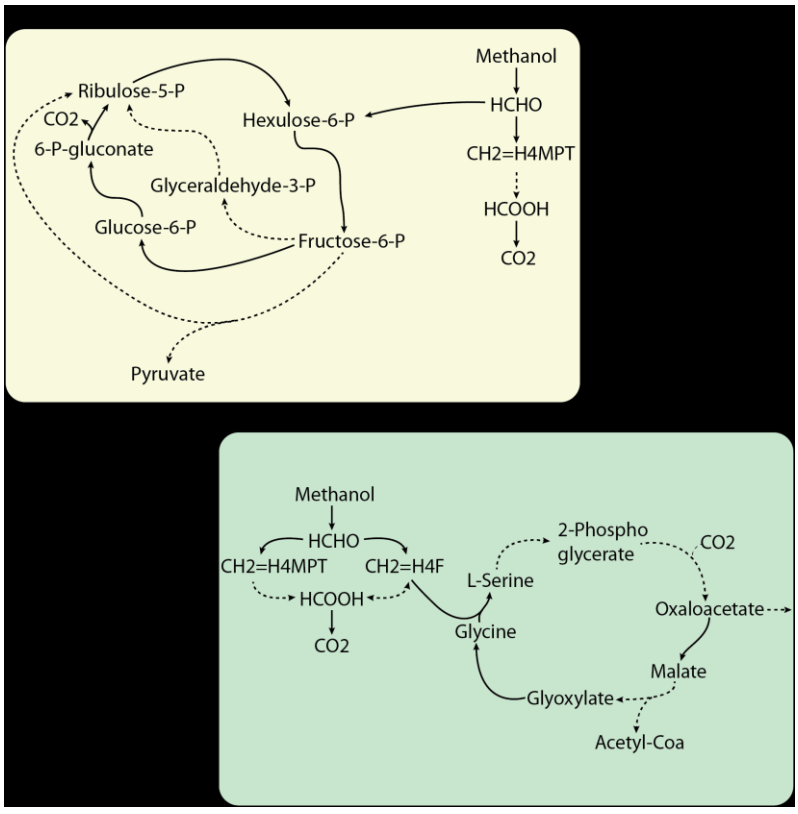


Figure 2-3: Methylotrophic bacterial methanol assimilation pathway. The RuMP cycle shown at the top, displays ribulose-5-phosphate as the recycling molecule of the cycle. The Serine cycle shown at the bottom. Figure adapted from Fei et al. [33].

RuMP cycle, the recycled molecule of the XuMP cycle is xylulose-5-phosphate (Xu5P) [20]. The molecule Ru5P and Xu5P are isomerases of each other, however depending on which molecule is

catalyzed with formaldehyde different products are generated [20] [33]. Dihydroxyacetone synthase (DHAS) catalyzes the hydroxyl transfer of xylulose-5-phosphate to formaldehyde producing one molecule of dihydroxyacetone and one molecule of glyceraldehyde-3-phosphate [35]. Overall, three formaldehyde molecules are needed to produce one molecule of glyceraldehyde for biosynthesis (Figure 2-4) [35].

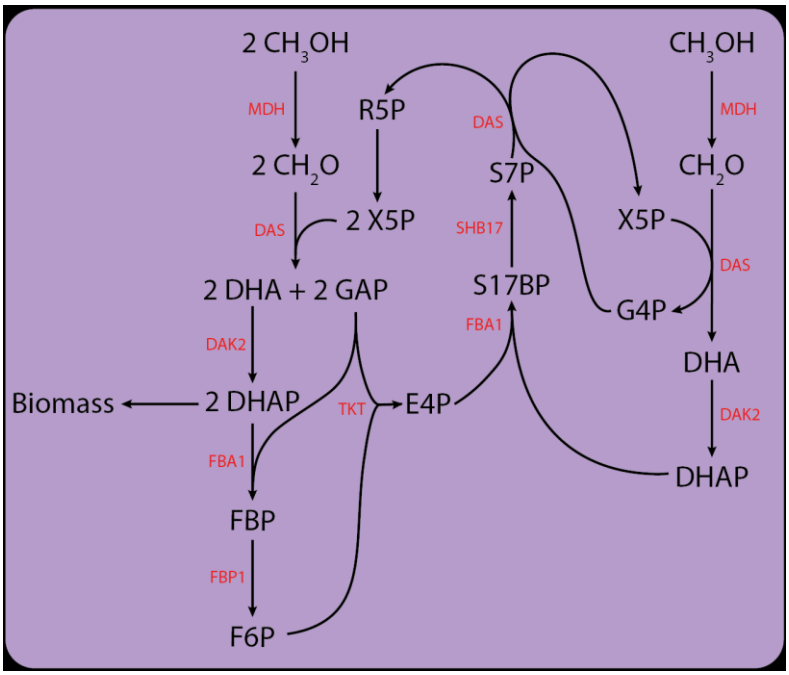


Figure 2-4: Methylotrophic yeast methanol assimilation pathway. Oxidation of methanol by Alcohol Oxidase (AOX) and the recycling of Xylulose-5-phosphate. Figure adapted from Geier et al. [35].

2.2.5 Methanol Condensation Cycle (MCC) engineered pathway

The different methanol assimilation pathways – RuMP cycle, Serine cycle, and the XuMP cycle – all depend to some degree on ATP (Table 2-6). It can be noted that the Serine cycle demands the most ATP/NADH. The RuMP cycle ultimately assimilates three methanol molecules to generate one pyruvate molecule, which is then decarboxylated to acetyl-CoA and CO_2 [36]. Therefore, the

Table 2-6: ATP and NADH demand on different methanol assimilation pathway.

Methanol Assimilation Cycle	ATP demand / mole of methanol	NADH demand / mole of methanol
Ribulose MonoPhosphate Cycle	2 ATP / 3 FALD	-3 NADH / 3 FALD
Serine Cycle	3 ATP / 2 FALD	3 NADH / 2 FALD
XuMP Cycle	3 ATP/ 3 FALD	-
Methanol Condensation Cycle	-	-

maximum carbon efficiency this pathway can achieve is 67% [36]. As of today, there has been no identified pathway that theoretically could reach 100% carbon efficiency [36]. However, the group of Bogorad et al. have conceptualized a Methanol Condensation Cycle (MCC), which theoretically could reach 100% carbon efficiency (Figure 2-5) [36]. Although, this pathway has not been engineered *in vivo*, using ^{13}C-labeled formaldehyde they show the fully functioning MCC in a cell free assay [36]. One drawback of the MCC is the sugar phosphate requirement. Previous studies suggest sugar phosphate regeneration as the rate-limiting step of the methanol assimilation pathway [17]. Secondly, the pathway performed well for about 5 hours but then the pathway activity started to decline [36]. Even though this pathway was performed in an *in vitro* setting, a decline in activity suggests a possible instability of intermediates [36].

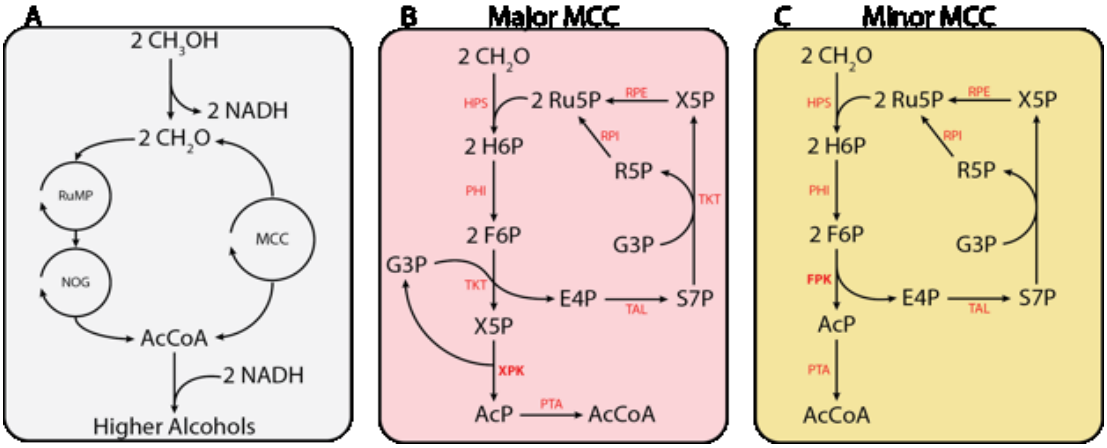


Figure 2-5: Methanol Condensation Cycle. (A) The MCC is the combination of RuMP cycle and NOG that bypasses ATP dependency. (B) The major MCC mode uses the more active X5P-phosphoketolase (XPK). (C) The minor MCC can achieve the same result with the less active F6P-Phosphotetolase (FPK). Figure adapted from Bogorad et al. [36].

2.3 Engineering methanol-assimilation pathway in model organisms

Engineering non-methylotrophs for methanol assimilation only began in the last few years [17]. J.E.N. Müller et al. were among the pioneers, and evaluated the MDH activities of both *B. methanolicus* PB1 and MGA3 as well as different HPS and PHI genes in *E. coli* [17]. Each gene was cloned into a high-copy number plasmid and controlled by an IPTG-inducible promoter [17]. Through flux analysis using ^{13}C-labeled methanol they determined the gene combination that maximized methanol assimilation. MDH activity was only reported for MDH2^{M} and MDH3^{P} and of the two MDH3^{P}, provided better methanol assimilation [17].

The prevalence and toxicity of formaldehyde has pressured organisms to evolved formaldehyde detoxification pathways. One of the major routes is the formaldehyde oxidation II pathway, where glutathione is needed as the recycled molecule and CO_2 is the final product; this is a potential route for carbon loss. In *E. coli*, the first enzyme of this pathway is formaldehyde dehydrogenase (*frmA*). J.E.N. Müller et al. deleted *frmA* in hopes to funnel more carbon down their engineered pathway [17]. However, *frmA* is also needed for other pathways involved in particular amino acid degradation [37]. As such, *in vitro* and *in vivo* data suggested that deleting *frmA* allowed for increased methanol assimilation [17]. Those results were contradicted by the ^{13}C-labeled methanol experiment, where strains lacking *frmA* had less ^{13}C-labeled incorporation [17]. In this study, *E. coli* did not assimilate methanol into biomass. Instead, they showed that with the addition of MDH, HPS, and PHI genes, 39.4% of the ^{13}C- labeled methanol could be assimilated as glycolytic and PPP intermediates in *E. coli* [17]. This served as a proof of concept that engineering a methanol-assimilation pathway could be achieved in *E. coli*.

Later, this proof of concept was taken a step further by Whitaker et al. [26]. This group used a different methanol dehydrogenase, from the family of *B. stearothermophilus*, and reported a 20-fold increase activity in comparison to MDH2^{P} [26]. The group cloned their MDH, HPS and PHI enzymes into high copy number plasmids, and transformed these into *E. coli* [26]. By changing the first enzyme of the pathway, Whitaker et al. reported a 53% ^{13}C incorporation in glycolytic intermediates, PPP intermediates, TCA intermediates and in hydrolyzed biomass components [26]. Methanol assimilation in *E. coli* was maximized when grown in a mixture of methanol and another carbon source, such as yeast extract peptone dextrose (YPD) [26].

Recently a group led by Min Jiang, were the first to report effective methanol assimilation in *S. cerevisiae*, and also reported growth using methanol as the sole carbon source [38]. They incorporated MDH3^{P}, HPS and PHI genes from *B. subtilis* into the genome in hopes to attain an active RuMP cycle [38]. However, using this approach they were unable to detect methanol incorporation into *S. cerevisiae*, as noted in the paper, this could be in part because codon-optimization for *S. cerevisiae* was not performed [38]. Instead, they decided to focus on incorporating the *P. pastoris* pathway, XuMP cycle, by introducing AOX, catalase (CAT), DAS, and dihydroxyacetone kinase (DAK) into *S. cerevisiae*'s genome [38]. Maximum growth of the engineered strains was observed when 1g/L yeast extract was added to the media [38]. In

comparison to the wild-type strain, their engineered strain harboring DAS2 showed an 11.7% increase in biomass when monitoring the cell densities of the cultures [38].

E. coli and *S. cerevisiae* are not the only microorganisms that have been engineered to assimilate methanol. The group of Witthoff et al., engineered *Corynebacterium glutamicum* a non-methylotrophic bacteria for methanol assimilation [39]. *C. glutamicum* contains an alcohol dehydrogenase A (ADHA) capable of converting methanol to formaldehyde [39]. An Increase in methanol oxidation compared to the wild-type was observed when either $MDH1^{M}$ and ACT.M or HPS and PHI from *B. subtilis* were expressed [39]. Deleting aldolase (ALD) and alcohol dehydrogenase E (ADHE), two enzymes involved in detoxifying formaldehyde, caused a decrease in fitness in the presence of methanol [39]. J.E.N. Müller et al. also reported better methanol assimilation when the formaldehyde oxidation II pathway was present, reinforcing the fact that formaldehyde concentration must be well balanced in the cell for proper methanol assimilation [17]. Methanol assimilation was maximized in the engineered strain harboring $MDH1^{M}$, ACT.M and HPS and PHI from *B. subtilis* [39]. Flux analysis using ^{13}C-labeled methanol, reported a biomass titer increase of 33% growth on a mixture of glucose minimal media and methanol in comparison to growth on glucose alone [39]. However, measurements of ^{13}C-labeled CO_2 suggested that 78% of formaldehyde was being lost to the detoxification of formaldehyde [39].

Leßmeier et al. also engineered *C. glutamicum*, however they introduced the MDH, PHI and HPS genes from *B. methanolicus* [40]. In contrast to J.E.N. Müller et al. who saw a 39.4% carbon labeled in the PPP intermediates, they only saw a maximum of 25% carbon labeling in the glycolytic and PPP intermediates [17][36]. This could be explained by the concentration of ^{13}C carbon added to the culture, almost a 20-fold difference [40]. Although Leßmeier et al. were not able to grow *C. glutamicum* solely on methanol, they showed coupling between methanol metabolism and the production of cadaverine, an exogenous molecule to *C. glutamicum* [40].

Research on the engineering of model organisms for the metabolism of methanol is starting to trend. However, we are still in the neophyte stages of this research area. Nevertheless, to reach a state of industrialization, current and further research is much needed.

2.4 Thesis Objective

Previous studies, as summarized in the introduction section, have been able to engineer both bacterial – *E. coli* and *C. glutamicum* – and yeast, *S. cerevisiae*, for the metabolism of methanol. Despite the achievements made thus far, only one lab has been able to show growth using methanol as the sole carbon source [38]. This was achieved in a *S. cerevisiae* strain; it was engineered using genes from its close relative *P. pastoris* and the pathway was targeted to the lysosome. This imparts a difficulty when pushing for growth under methanol as a sole carbon source; as stated earlier, methanol induces peroxisome proliferation in methylotrophic yeast, for which it occupies roughly 80% of the total cytoplasmic space [20]. Additionally, peroxisome proliferation is dependent on many accessory proteins [20]. Engineering such a drastic change may lead to unforeseen complications. As such, one may target this exact pathway to the cytoplasm, however the enzyme involved in converting methanol to formaldehyde creates reactive oxygen species as by-products, most likely why this pathway has evolved in an organelle. Alternatively, the RuMP cycle and the Serine cycle could be engineered for methanol assimilation, which the former has shown to be functional in both *E. coli* and *C. glutamicum*.

In this thesis, we sought to engineer *S. cerevisiae* for the metabolism of methanol toward biomass synthesis. We exploit the energetic advantage of the RuMP pathway as we introduce MDH, HPS, and PHI in *S. cerevisiae.* Rather engineering a native methylotrophic yeast for production of a high-valued biochemical, we opted for *S. cerevisiae* as our model organism because the of the abundant library of information and the wealth of available tools developed for engineering and studying it.

We tested six NAD(P)-dependent MDHs variants, two HPS variants, and two PHI variants from two bacterial species – *B. methanolicus* MGA3 and *B. methanolicus* PB1 – as well as the two possible MDH accessory proteins. We used different genetic manipulation tools such as Clustered Regularly Interspaced Short Palindromic Repeats with Caspase 9 (CRISPR-Cas9) to introduce and delete genes in the genome of *S. cerevisiae*. As a result, we introduced up to 3 copies of each enzyme, as well as deleting the formaldehyde dehydrogenase (SFA1) gene.

3 MATERIALS & METHODS

3.1 Reconstruction of *S. cerevisiae* metabolic model iAV3000

Prior to engineering *S. cerevisiae*, a metabolic model of the model organism, named iAV3000, was built. The *S. cerevisiae* model was used to better understand and predict the outcomes of adding the RuMP pathway. Specifically, it was used to predict any cofactor imbalance, to predict ATP production, and to predict the overall growth when using methanol and the RuMP pathway. In this chapter, the methodology behind the reconstruction of the iAV3000 model is described.

A published article by Palsson et al., studied and outlined necessary steps towards the reconstruction of a high-quality genome scale metabolic model [41]. This article was used as a guide for the reconstruction of the iAV3000 model, however, some steps were omitted while others were altered. The following will describe each step that was accomplished during the reconstruction of the iAV3000 model.

A basic *S. cerevisiae* model from Jol et al. was used as the foundation of the iAV3000 model [42]. The model includes 240 reactions involved in central carbon metabolism and amino acid synthesis pathway, encompassing the cytosolic and mitochondrial compartment as well as other parallel pathways [42]. To assess the quality of the model constructed by Jol et al., different Cobra Toolbox functions were used in conjunction with MATLAB [43]. The primary solver used for analysis was the Gurobi5, however it was used in comparison with the GNU Linear Programming Kit (GLPK) solver [44][45][46]. This approach identified weak areas in the model, such as the electron transport system (ETS), the biomass reaction, and some other important missing reactions.

3.1.1 Electron Transport System

The theoretical yeast aerobic mole ATP/mole glucose (Y_{ATP}^{Theo}) production is 28.1, however, proton leakage across the mitochondrial membrane is present in all organisms and so the Y_{ATP}^{Theo} is rarely if ever observed [47][48]. Additionally, proton leakage can significantly impact the cell's phosphate/oxygen (P/O) ratio, amount of ATP produced per atom of oxygen consumed. A previous study found that increasing the growth temperature of yeast *S. cerevisiae* CEN.PK 113D from 30-37°C, the P/O ratio changed from 0.7-1.4 [49]. An acceptable P/O ratio has been established at

0.95 for growth at 30°C and at a pH of 5 [47]. Previous models have set the experimental ATP/glucose (Y_{ATP}^{Exp}) to be ~12.5, and thus this was also our target [48].

The reactions of the oxidative phosphorylation system found in the Jol et al. model produced 890 mole ATP/mole glucose, which is over 30-fold the theoretical value. One of the main reasons this value was much larger was due to the many perpetuate cycles in the model, thus, these had to be removed. This was done by assessing the directionality of each reaction. At times, the flux was forced toward a certain direction, the one observed under the current state of the model. Once these were removed, the *in silico* production of mole ATP/mole glucose was 2.

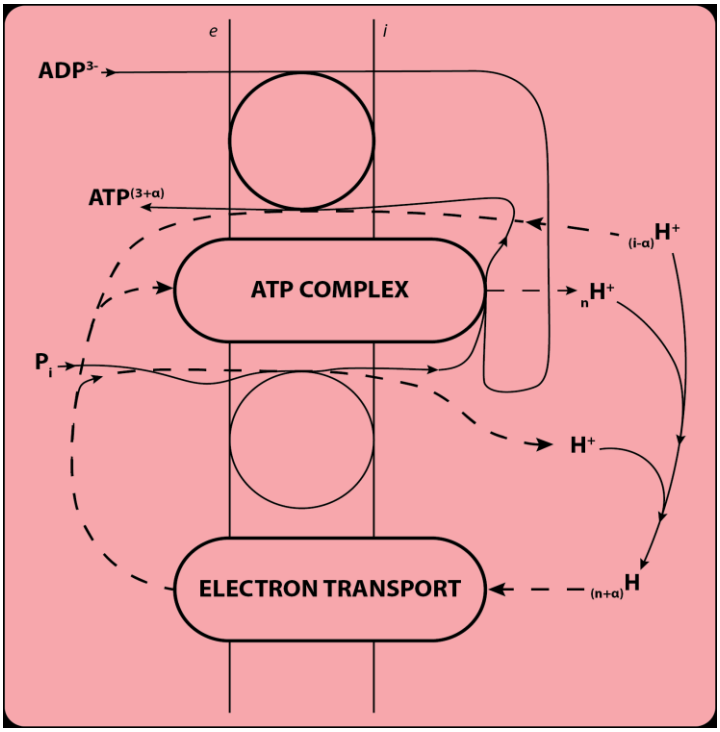


Figure 3-1: The H+ movement during the activity of ATP synthesis and the ATP/ADP balance across the mitochondrial inner membrane space. Figure was adapted from Wulf et al. [50].

The yeast ATPase synthase has been studied extensively and it has been reported that 4 protons are needed to generate 1 ATP molecule, and subsequently 1 of these 4 protons is consumed for ATP extrusion (Figure 3-1) [50]. Thus, the ATP synthesis reaction was changed accordingly. The change in this reaction was sufficient to fix the ETS and to reflect experimental results. With those changes, the *in silico* P/O ratio of the iAV3000 model is 0.98 and 11.5 moles of ATP are produced per mole of glucose. Once the ETS was able to generate an appropriate amount of mole ATP/ mole glucose with a proper P/O ratio, the model was ready for expansion. Since the aim of the model is

to predict whether methanol is a sufficient carbon source, it was necessary to include additional reactions involved in: the metabolism of alcohols, the balance of cofactors, NAD(H) and NADP(H), and missing secondary reactions.

3.1.2 The Expansion of Available Reactions

The basic model contained reactions found in primary pathways and in amino acid synthesis pathways and some found in secondary pathways. The pathways found in the model were compared to the Saccharomyces Genome Database (SGD), Braunschweig Enzyme Database (BRENDA), and Kyoto Encyclopedia of Genes and Genomes (KEGG) and to previously published models: iMM904, iIN800, iKC, and iSce926 model [51][52][53][54][55][56].

The directionality, activity, and the inclusion of all reactions in the model produced by Jol et al. were verified using the different databases. To verify the subcellular localization of each reaction, many different databases were used in a comparative manner and with any available published articles about the specific enzyme. No high-throughput subcellular protein localization data was used. Databases used for identification of subcellular protein localization for the reconstruction of the iAV3000 model were: the Yeast Resource Center Informatics Platform (YRC) database, the Centre for Cellular and Biochemical Research (CCBR) database, the Yeast Protein Localization (YPL) Database, the Yeast GFP Fusion Localization Database, and A Database Of Organelle And Protein Complexes [57][58][59][60][61].

Using the previously published models, new reactions were added. Each newly added reaction was compared across all models and to the SGD and BRENDA databases. Reactions that included the hydrolysis of ATP were constrained in the forward direction only, unless otherwise proven based on experimental data. If no experimental results could back up the directionality of the reaction, the lower and upper bounds were left unconstrained.

The charge and mass balance of all reactions in the metabolic model were determined using the Cobra Toolbox in MATLAB and then balanced appropriately [43].

Once all the necessary reactions were added to the model the biomass equation was revised. To be able to predict with confidence the growth rate of *in vivo* experiments, it is empirical that the biomass reaction be as precise as possible.

3.1.3 Development of the biomass equation

The composition of the biomass is growth-condition dependent and therefore can affect the results of the simulation if some nutrients are missing. It is crucial to build a proper biomass equation when assessing the *in silico* growth rate. The *in silico* growth rate of the model is represented by a biomass equation that includes: proteins, carbohydrates, lipids, amino acids, RNA, DNA, and ATP. An Aerobic glucose-limited condition was evaluated using this model. The availability of glucose allowed for glucose derepressed reactions to be turned on [59]. To make sure the right reactions were turned on or off in this condition, two independent studies tested different protein expression in different chemostat conditions, and thus were used as evidence to include or exclude certain reactions [59][60].

3.1.3.1 Protein composition of the biomass

The protein composition of the cell was shown to have the greatest effect on the bioenergetics and growth rates [64][62]. Additionally, depending on the growth condition, many transcription factors may activate or repress the transcription of specific genes, thus changing the composition of available proteins in the cell and ultimately changing the concentration of total protein [58][59].

As previously discussed, many genes are glucose repressed while others are glucose derepressed. Previous experiments done by Schulze have shown that the amino acid ratios under different conditions are relatively unchanged and that only the total protein percent per dry cell weight(DCW) is affected (Table 3-1)[62].

Table 3-1: Amino acid composition at D = 0.1h-1 under glucose-limited chemostat.

Amino Acid	mmol/gDCW	Amino Acid	mmol/gDCW
ALA	0.35734	LEU	0.25014
ARG	0.13579	LYS	0.23942
ASN	0.17152	MET	0.050027
ASP	0.17152	PHE	0.11435
CYS	0.04288	PRO	0.12864
GLN	0.268	SER	0.25371
GLU	0.268	THR	0.19653
GLY	0.32518	TRP	0.028
HIS	0.075041	TYR	0.096481
ILE	0.17152	VAL	0.25728

*Table adapted from Oura (1972), Bruinenberg (1983), and Förster (2003), [63][88][48].

Additionally, Schulze shows that protein and RNA content increases with increasing dilution rates at the expense of carbohydrates, while the other cellular components are virtually independent of the dilution rate [62]. The percent composition of the macromolecules composing the biomass was retrieved from Schulze Ph.D. dissertation (Table 3-2)[62].

Table 3-2: Cellular composition under aerobic glucose-limited chemostat.

D=0.1 h^{-1}	Cellular content %(w/w) Condition: Aerobic Glucose-limited chemostat
Protein	45.0
Glycogen	8.4
Trehalose	0.8
Mannan	13.1
Other carbohydrates (13BDglcn)	18.4
RNA	6.3
DNA	0.4
Lipid	2.9
Ash	5.0
Sum	100.3

*Table adapted from Schulze (1995) [62].

3.1.3.2 Carbohydrates, lipids, and nucleic acids

Experimental data show that carbohydrates components are inversely correlated to protein content; as the concentration of protein rises, carbohydrate content decreases [62]. This is because when glucose is abundant the cells main objective is to grow leading to an increase in carbohydrates production, however during starvation, *S. cerevisiae*'s focus is on survival, thus carbohydrate

synthesis is decreased and protein production is increased. Only the major carbohydrates were included in the iAV3000 model (Table 3-3).

Table 3-3: Carbohydrate composition in *S. cerevisiae.*

Carbohydrates	Cellular content %(w/w)	mmol/gDCW	Reference
Glycogen	8.4	0.519	[62]
Trehalose	0.8	0.023	
Mannan	13.1	0.821	
1,3-beta-D-Glucan	18.4	1.136	

*Table adapted from Schulze (1995) and Förster (2003), [62] and [48].

The DNA calculation was based on a G/C content of 40.3% [65]. The RNA and DNA content accounts for 6.3% and 0.4% respectively of the total cell content (Table 3-4)[62].

Table 3-4: RNA and DNA compositions in *S. cerevisiae.*

RNA	Cellular content %(w/w)	mmol/gDCW	Reference
AMP		0.051	[63]
GMP		0.051	
CMP		0.050	
UMP		0.067	
DNA	**Cellular content %(w/w)**	**mmol/gDCW**	**Reference**
dAMP	0.313	0.0036	[65]
dGMP	0.211	0.0024	
dCMP	0.287	0.0024	
dTMP	0.289	0.0036	

*Table adapted from Schulze (1995) and Förster (2003), [62] and [48].

The sterol and sterol esters were assumed to be composed solely of ergosterol and zymosterol respectively (Table 3-5) [59][63]. Monoacylglycerols were not included in the model since Ratledge and Evans (1989) believed that components of this measurement was most likely an artifact of extractions [66].

Table 3-5: Sterols and triglycerides content in *S. cerevisiae.*

Sterols and Triglycerides	Cellular content %(w/w) lipid	mmol/gDCW	Reference
Sterols (Ergosterol)	1	0.001	[62]
Sterol esters (Zymosterol)	2	0.002	
Triglyceride	20	0.007	
Monoacylglycerol (artifact)	8	0.0066	

*Table adapted from Schulze (1995) and Förster (2003), [62] and [48].

Aerobic glucose-limited chemostat experiments suggest that Lipids and nucleic acids only accounts for 12-15% of the total biomass, thus any fluctuation of the lipid content will marginally

changes the *in silico* growth rate [62][63]. Only the major phospholipids were included in this model (Table 3-6).

Table 3-6: Phospholipid composition in *S. cerevisiae.*

Phospholipids	% Phospholipids	mmol/gDCW	Reference
Phosphatidate	2.5	0.0006	[89]
Phosphatidylcholine	29.2	0.006	
Phosphatidylserine	8.0	0.0017	
Phosphatidylethanolamine	20.5	0.0048	
Phosphatidylinositol	28.7	0.0053	

*Table adapted from Henry (1982) and Förster (2003), [89] and [48].

3.1.4 Growth Associated ATP Maintenance (GAM) & Non-Growth Associated ATP Maintenance (Non-GAM)

The microbial energy requirement is dependent on the composition of its biomass and the available carbon source. The required ATP amount for one amino acid assembly into a protein is 4 moles, thus it is evident that an increase in protein content will require additional ATP [47]. Additionally, depending on the carbon source available, different amounts of ATP will be needed for the assimilation and dissimilation reactions [47]. ATP requirements are highly dependable on the protein content, and amino acid polymerization [47]. A positive correlation has been observed between protein content and ammonium transport in the cell, both of which demand energy [47]. The required ATP amount for growth formation under glucose-limited chemostat was calculated based on the requirements to meet a biomass yield of 0.51g DCW/g glucose [47].

Previous research studies on glucose-limited chemostat obtained an Y_{ATP}^{Exp} in the range of 12 – 16 while the Y_{Theo}^{Max} equals 28.1 [47]. There are two possible answers that could explain the discrepancy between the experimental and theoretical Y_{ATP}. The first is synthesis of certain cell components requires more ATP than used in theoretical calculations, or that ATP produced during dissimilation is expended elsewhere than for biomass formation [47]. Little research has provided solid explanation for this phenomenon. Therefore, Y_{ATP} can be divided into two categories, ATP required for growth (GAM), and ATP not required for growth (Non-GAM) [47]. The Non-GAM is predicted to be very small and almost negligible, and since calculation of this parameter is very difficult in yeast it has been omitted by previous studies and is also omitted in this current study

[47]. The GAM value for *S. cerevisiae* utilizing glucose as a carbon source is 59.276 ATP/gDCW*h (Table 3-8 and Table 3-7).

Table 3-8: Calculation of the theoretical energy requirements $\left(Y_{theo}^{Max}\right)$ for biomass formation of *S. cerevisiae* in aerobic glucose-limited chemostat growth at a dilution rate of 0.10h^{-1}.

	Total (mmol ATP/g biomass)
Amino acid synthesis	1.62
Polymerization	15.87
Carbohydrate synthesis	4.82
Lipid synthesis	1.79
RNA synthesis	1.35
Polymerization	0.47
Turnover of mRNA	0.71
NADPH generation	0.77
Transport	
Ammonium	5.80
Potassium and Phosphorus	2.40
Sum	35.60

Table 3-7: Energy requirement for polymerization of macromolecules.

	Cellular content %(w/w)	Polymerization Energy	
		Per Molecule (mmol ATP/g polymer)	Total (mmol ATP/g biomass)
Protein	45	37.7	16.965
Carbohydrate	40.7	12.8	5.2096
RNA	6.3	26.0	1.638
DNA	0.4	26.0	0.104
Sum	93.4%		23.911

*Table adapted from Verdyun (1991) and Förster (2003), [64] and [48].

The overall cell composition, which makes up the biomass equation is as follows:

Table 3-9: Overall biomass composition.

Molecule	Content (mM/gDCW/h)	Molecule	Content (mM/gDCW/h)	Molecule	Content (mM/gDCW/h)	Molecule	Content (mM/gDCW/h)
ALA	0.35734	**LYS**	0.23942	**CMP**	0.050	**PE**	0.0048
ARG	0.13579	**MET**	0.050027	**UMP**	0.067	**PC**	0.006
ASN	0.17152	**PHE**	0.11435	**dAMP**	0.13579	**PI**	0.0053
ASP	0.17152	**PRO**	0.12864	**dGMP**	0.17152	**PS**	0.0017
CYS	0.04288	**SER**	0.25371	**dCMP**	0.17152	**Glycogen**	0.519
GLN	0.268	**THR**	0.19653	**dTMP**	0.04288	**Tre**	0.023
GLU	0.268	**TRP**	0.028	**UMP**	0.067	**Mannan**	0.821
GLY	0.32518	**TYR**	0.096481	**Triglyceride**	0.007	**13bdglc**	1.136
HIS	0.075041	**VAL**	0.25728	**Zymosterol**	0.0020	**SLF**	0.096481
ILE	0.17152	**AMP**	0.051	**Ergosterol**	0.001	**ATP**	59.276
LEU	0.25014	**GMP**	0.051	**PA**	0.0006		

*Table adapted from Förster (2003) [48].

3.1.5 Steps Used for the Generation of the iAV3000 Model

Only selective steps in the reconstruction of the model proposed by Palsson et al, have been retained [41]. The proposed protocol is composed of 31 steps, but only 24 of these were used. Additionally, some steps were modified. In fact, the present study differs from Palsson et al. since the model was build off previously published models. The following summarizes the steps used in the reconstruction of the model (Table 3-10).

Table 3-10: Steps used in the Reconstruction and Analysis of the Model iAV3000.

Stage	Step	Description
1	**Creating a draft reconstruction** The core of the iAV3000 model used previously published models by Jol et al. and Zampar et al. [42][68]. Different models were used to expand the iAV3000 model: the model iMM904 by O Palsson et al. [54], the model iKC by Correia et al. [69] and the model S7 by Chowdhury et al. [56], were used to expand the model. This was believed to be a better and faster way to begin reconstructing a model since these models have already been manually curated and should be theoretically errorless, therefore making the manual curation a faster process.	
	Step 1: Genome annotation	Previously published model by Jol et al and Zampar et al. were used as the core genome.
	Step 2: Candidate metabolic functions	Three models, iMM904, iKC, and S7 model were used to add missing reactions to the core model.
2	**Manual reconstruction refinement.** Metabolic functions and reactions collected during the draft reconstructions were individually evaluated against yeast specific literature as well as expert opinion [41]. Therefore, each enzyme was evaluated on the following criteria: their presence/expression or absence/repression (condition dependent), the compartment they are found in (condition dependent), and the metabolites they interact with. Subsequently, each reaction was mass and charge balanced.	
	Step 3: Reconstruction assembly	The absence and presence of each enzyme was evaluated under several conditions.

	Step 4: Verification of metabolic functions – cofactor usage	The metabolites and cofactors involved in each reaction and therefore associates with individual enzymes were evaluated.
	Step 5 and 6: Charged formula for each metabolite	The Cobra Toolbox function *CheckMassChargeBalance* was used to analyze each reaction. Any unbalanced reaction was corrected for [43].
	Step 7: Reaction stoichiometry	The Cobra Toolbox function *CheckMassChargeBalance* was used to analyze each reaction. Any unbalanced reaction was corrected [43].
	Step 8: Reaction directionality	Using the BRENDA database, the directionality of each reaction was determined. Reactions involving the hydrolysis of ATP were made irreversible unless proven otherwise. To remove these perpetuate cycles, some reactions were forced into one direction. Some reaction's directionality was based off the reference models.
	Step 9: Information for gene and reaction localization	Using different databases: The Yeast Protein Localization Images database (YPL), Yeast Resource Center database (YRC), Collection of Yeast Cells and Localization Patterns database (Cyclops), and yeast GFP fusion localization database and A Database of Organelle and Protein Complexes, were used to identify the localization of each enzyme. For each of the determine location, at least 3 databases were used and compared. If each database agreed the search was terminated. However, if they did not, additional databases were used. From each database, the reference data was checked, and the experiment done was validated. Data from specific enzymatic experiments were given a higher confidence over large scale studies. If no localization could be found for a reaction, the reference models were used.
	Step 10: Spontaneous reactions	These were identified using the table made available by Palsson et. al. and added to the model [41].
	Step 11: Intracellular transport reactions	Reactions were search for in the literature. If no data was found for a reaction the reference models were use.
	Construction of Biomass Identity	The biomass equation includes, protein, carbohydrates, lipids, fatty acids, and ATP for maintenance and non-maintenance. A different biomass equation was used for different conditions.
	Step 13: Growth-associated ATP maintenance reaction (GAM)	The energy needed to grow and to maintain cell viability is known as the growth-associated ATP maintenance. This value, depending on the condition the host finds itself in, will vary. Therefore, the

		ATP must be adapted accordingly. The GAM has been previously calculated by Verduyn et al. [64] and thus used in the model.
	Step 14: Non-GAM reactions (NGAM)	Omitted
	Step 15: Sink reactions	Sink reactions allow one to identify the maximum flux through a reaction based on a specific amount of carbon source. This is useful when analyzing the electron transport system in the model, or to evaluate the amount of NADH and NADPH produced. In this step, an ATP, NADH, and NADPH sink reactions were added to analyze the model. Because ATP is hydrolyzed in the yeast as a normal function, the sink reaction was kept.
	Step 16: Growth medium requirements	As mentioned earlier, the biomass is condition depended and thus specific reactions were added to supply the cells with the appropriate nutrients.
	Conversion from reconstruction to mathematical model.	Using the Cobra Toolbox function *xls2model,* the model created in an xlsm format could be converted to a mathematical model and further analyzed [43].
	Step 17: Simulation constraints	Simulation constraints were done using the Cobra Toolbox functions *FluxBalanceAnalysis* and *FluxVariability* [43].
3	**Network evaluation = 'Debugging mode'**	
	Step 18: Blocked reactions and Metabolic dead end	Using the *FluxVariability* function, reactions within the model that had no allowable flux were found and corrected for. A reaction where no flux is permitted may be due to a metabolic dead end, where downstream metabolite is formed but not consumed or secreted from the cell. Therefore, this may be due to the fact that the model is missing a metabolic function. During the construction of the model, several of these came up. All culprits were found and added to the model.
	Step 19: Candidate reactions for gap filling	Based on step 22 and other analyses, some metabolic functions may not be present in the model when in fact should. Using the reference models, many of these gaps were found and fixed.
	Step 20: Biomass precursor production	To make sure the model is functioning properly, it is important to produce all major precursors of the biomass reactions. This is done by changing the objective reaction to the efflux reaction of the precursor.

4	Step 21: By-product secretion	In this step, all by products are established and added a secretion reaction. This provides a way to see if some by products are being overproduced by the model.
	Step 22: Known incapability	Each reaction was deleted, and the growth rate was determined. The values were then compared to the literature.
5	Step 23: Comparison of predicted physiological properties with known properties	The amount of ATP made per oxygen atom is the P/O ratio. This is important to calculate in order to have an *in silico* P/O ratio that reflects the *in vivo* P/O ratio.
	Step 24: Quantitative evaluation of growth rate	The biomass during glucose-limited aerobic chemostat was lower than previously reported. Therefore, to identify the limiting factor, each precursor was added to the media and the biomass was recalculated. The limiting factor was determined this way and fixed appropriately.
	Prospective use	The addition of three reactions allowed the model to have a complete RuMP cycle and allowed for the metabolism of methanol. The model then predicted growth and ATP expenditure using methanol as a carbon source.

*Table adapted from Palsson (2010), [41].

3.2 Strain engineering protocols

The following includes a list of parts for the assembly of our engineered strains.

3.2.1 Candidate selection

All genes used in this study are found in (Table 3-11). All genes were chosen based on previous research done by J.E.N. Müller et al.. A total of six NAD(P)-dependent methanol dehydrogenase variants were chosen, three from the strain *B. methanolicus* MGA3 and three from *B. methanolicus* PB1. Additionally, two 3-hexulose-6-phosphate synthase (HPS) and 3-hexulose-6-phosphate isomerase (PHI) variants were chosen. One HPS and PHI gene came from *B. methanolicus* MGA3 and the other from *B. methanolicus* PB1. Later, two NAD(P)-dependent MDH activator proteins (ACT) were added, one from *B. methanolicus* MGA3 and the other came from *B. methanolicus* PB1. All genes were synthesized by Gen9 (Boston, MA). All genes were cloned into our custom pBOT vectors, derived from pGREG *E. coli-S. cerevisiae* shuttle vectors (Table 3-12) [70].

Table 3-11: Gene candidate used in this study.

Gene name (original)	Gene name (used in this study)	Description	Host strain	GenBank accession number
MDH	MM1	NAD(P)-dependent methanol dehydrogenase	*Bacillus methanolicus* MGA3	AIE61787.1
MDH2	MM2			AIE59127.1
MDH3	MM3			AIE60275.1
MDH	MP1		*Bacillus methanolicus* PB1	EIJ77618.1
MDH1	MP2			EIJ78790.1
MDH2	MP3			EIJ78397.1
HPS	HPS.M	3-hexulose-6-Phosphate synthase	*Bacillus methanolicus* MGA3	AIE59794.1
HPS	HPS.P		*Bacillus methanolicus* PB1	EIJ81375.1
PHI	PHI.M	3-hexulose-6-Phosphate isomerase	*Bacillus methanolicus* MGA3	AIE59793.1
PHI	PHI.P		*Bacillus methanolicus* PB1	EIJ81376.1
ACT	ACT.M	NAD(P)-dependent methanol dehydrogenase activator protein	*Bacillus methanolicus* MGA3	AIE60888.1
ACT	ACT.P		*Bacillus methanolicus* PB1	EIJ78755.1

3.2.2 Strains and plasmids construction

All strains and plasmids used in this study (Table 3-12 and Table 3-14) were assembled using overlapping DNA parts and transformation-assisted recombination in yeast. Plasmid propagation was done using *E. coli* DH5-α. All primers for PCR amplifications of the DNA parts are listed in (Table 3-14). Sites of genome integrations are described in (Table 3-15). All chemicals were obtained from Sigma-Aldrich. Phusion polymerase (New England Biolabs) was used for cloning purposes as well as for producing all DNA parts according to the manufacture's recommended protocols. Plant Phire polymerase (Fermentas) was used for colony PCR, and for generating DNA fragments sent for sequencing (Eurofins Genomics). Plasmids were isolated using the GeneJET Plasmid Miniprep Kit (Thermo Scientific) and DNA isolated from agarose gels and PCR reactions were purified using the QIAquick Gel and QIAquick PCR purification kit (Qiagen), respectively.

Table 3-12: List of plasmids used in this study.

Plasmid	Resistance	Gene
pBOT-HIS	HIS3 amp KanMX	MM11, MM2, MM3, MP1, MP2, MP3
pBOT-LEU	LEU2 amp	HPS.M, HPS.P

	KanMX	
pBOT-URA	URA3 amp KanMX	PHI.M, PHI.P
pYES2-URA	URA3 amp	ACT.M, ACT.P
pCAS-TYR	G418/Hygromycin	

Genes found on pBOT-HIS plasmids have a TEF1 promoter and PGI1 terminator. Genes found on pBOT-LEU plasmids have a FBA1 promoter and ADH1terminator. Genes found on pBOT-URA have PMA1 promoter and TPI terminator. Finally, genes found on pYES2-URA have a TEF1 promoter and PGI1 terminator.

Table 3-13: *Saccharomyces cerevisiae* CEN.PK 1D strains tested in this study.

Strain name	Phenotype of host	Genes added on plasmid
WT*	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$	
MM1-LP1	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MM1^+$	
MM2-LP1	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MM2^+$	
MM3-LP1	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MM3^+$	
MP1-LP1	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MP1^+$	
MP2-LP1	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MP2^+$	
MP3-LP1	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MP3^+$	
MM1-LP2	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MM1^+$ (2 copies)	
MM2-LP2	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MM2^+$ (2 copies)	
MM3-LP2	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MM3^+$ (2 copies)	
MP1-LP2	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MP1^+$ (2 copies)	
MP2-LP2	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MP2^+$ (2 copies)	
MP3-LP2	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MP3^+$ (2 copies)	
MM1-LP1+LP2	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MM1^+$ (3 copies)	
MM2-LP1+LP2	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MM2^+$ (3 copies)	
MM3-LP1+LP2	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MM3^+$ (3 copies)	
MP1-LP1+LP2	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MP1^+$ (3 copies)	
MP2-LP1+LP2	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MP2^+$ (3 copies)	
MP3-LP1+LP2	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $MP3^+$ (3 copies)	
WT ΔSFA1	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $Sfa1^-$	
MM1 ΔSFA1	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $Sfa1^-$ $MM1^+$(3 copies)	
WT HPS.M PHI.M	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $HPS.M^+$(3 copies) $PHI.M^+$(3 copies)	
WT HPS.M PHI.P	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $HPS.M^+$(3 copies) $PHI.P^+$(3 copies)	
WT HPS.P PHI.M	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $HPS.P^+$ (3 copies) $PHI.M^+$(3 copies)	
WT HPS.P PHI.P	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $HPS.P^+$(3 copies) $PHI.P^+$(3 copies)	
WT ΔSFA1 HPS.M PHI.M	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $Sfa1^-$ $HPS.M^+$(3 copies) $PHI.M^+$(3 copies)	
WT ΔSFA1 HPS.M PHI.P	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $Sfa1^-$ $HPS.M^+$(3 copies) $PHI.P^+$(3 copies)	
WT ΔSFA1 HPS.P PHI.M	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $Sfa1^-$ $HPS.P^+$ (3 copies) $PHI.M^+$(3 copies)	
WT ΔSFA1-HPS.P PHI.P	$His3^-$ $Ura3^-$ $Leu2^-$ $Trp1^-$ $Sfa1^-$ $HPS.P^+$(3 copies) $PHI.P^+$(3 copies)	

MM1 HPS.M PHI.M	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ MM1.M$^+$(3 copies) HPS.M$^+$(3 copies) PHI.M$^+$(3 copies)	
MM1 HPS.M PHI.P	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ MM1.M$^+$(3 copies) HPS.M$^+$(3 copies) PHI.P$^+$(3 copies)	
MM1 HPS.P PHI.M	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ MM1.M$^+$(3 copies) HPS.P$^+$ (3 copies) PHI.M$^+$(3 copies)	
MM1 HPS.P PHI.P	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ MM1.M$^+$(3 copies) HPS.P$^+$(3 copies) PHI.P$^+$(3 copies)	
MM1 ΔSFA1 HPS.M PHI.M	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ Sfa1$^-$ MM1.M$^+$(3 copies) HPS.M$^+$(3 copies) PHI.M$^+$(3 copies)	
MM1 ΔSFA1 HPS.M PHI.P	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ Sfa1$^-$ MM1.M$^+$(3 copies) HPS.M$^+$(3 copies) PHI.P$^+$(3 copies)	
MM1 ΔSFA1 HPS.P PHI.M	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ Sfa1$^-$ MM1.M$^+$(3 copies) HPS.P$^+$ (3 copies) PHI.M$^+$(3 copies)	
MM1 ΔSFA1 HPS.P PHI.P	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ Sfa1$^-$ MM1.M$^+$(3 copies) HPS.P$^+$(3 copies) PHI.P$^+$(3 copies)	
MM1 ACT.M	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ MM1.M$^+$(3 copies)	ACT.M
MM2 ACT.M	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ MM2.M$^+$(3 copies)	ACT.M
MM3 ACT.M	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ MM3.M$^+$(3 copies)	ACT.M
MP1 ACT.P	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ MM1.P$^+$(3 copies)	ACT.P
MP2 ACT.P	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ MM2.P$^+$(3 copies)	ACT.P
MP3 ACT.P	His3$^-$ Ura3$^-$ Leu2$^-$ Trp1$^-$ MM3.P$^+$(3 copies)	ACT.P

*System B (SysB) refers to a proprietary strain created by the Martin Lab, Concordia University. This strain contains these "Landing pads" allows for quick integration of 1-4 (LP1-LP4) gene copies using a single gRNA [73].

Table 3-14: Primer List.

Primer Name	Description	Primer Sequence 5'- 3'
LP1_TEF1_F	Used to amplified NAD-MDH variants for integration at FGF20 site.	TTGGCGCGTGACTGTCAGCGCAATCCGAGGAATACTCTGAATAAAACAACTTATATAATAAAAATGCGGACTTTTAATTTTCGAGGACCG
LP1_TEF1_R	Used to amplified NAD-MDH variants for integration at FGF20 site.	GAGATGTAGGCCGCCTTGTTAGCGTCTGTTATAATTATTTTCTTATTTTGATGTAATATAAAGAACGGGGTATACTGGAGGCTTCATGAG
LP2_TEF1_F	Used to amplified NAD-MDH variants for integration at FGF18 and FGF24 sites.	TACACCTTAATCTCCGGTTCATGCTAGGGATGTGGCTGCATGCTACGTTGACACACCTACACTGCTCGGACTTTTAATTTTCGAGGACCG
LP2_TEF1_R	Used to amplified NAD-MDH variants for integration at FGF18 and FGF24 sites.	AACAGTAAACGAACACGTGACGATGCGGAACGGCTCCGGCCAGGTCGTACACTTCATCTCGCTACTTCGGTATACTGGAGGCTTCATGAG
LP3.A_FBA1p_F	Used to amplified NAD-MDH variants for integration at FGF18 and FGF24 sites.	GAACGCTGTCTGAAGGATGAGTGTCAGCGAGTGTAACTCGATGAGCTACCCAGTAGTCGTACTGGTCGAGACAACATCCAACTGGCACCG

LP3.Z_TPIt_R	Used to amplified HPS+PHI variants for integration at FGF7, FGF19, USERXii-1 sites.	GTCAACACCCCCGCCCTAGTCTAAGCTG ATATGTCGCGTCAGGCCTCGGGATTGAG TGCATTAGAGCTGGGGTATACTGGAGGC TTCATG
USERXII-2_FBA1p_F	Used to amplified HPS+PHI variants for integration at USERxii-2.	AACGAAAAAGAAAAGAAAGACCATGTC ATGTACGGGCAATCAGAATCTGTAACAA GCGCCATCCAACTGGCACCGCTGGC
USERXII-2_TPIt_R	Used to amplified HPS+PHI variants for integration at USERxii-2.	ATTTCCTATAATAGAAATCCAAGTGGCA AAAGCGTTAGACGCAGTACAAGGACGC GTTAAGGTGATATCAGATCCACTAGTGG CC
AHD1t-PMAp_F	Used to amplified HPS+PHI variants for integration.	CAGAGTGATGTATGGGTGATCCCAGCAA TGCCTGTGCATGCCGGTAGAGGTGT
PMA1p-ADH1t_R	Used to amplified HPS+PHI variants for integration.	GCTCTTATTGACCACACCTCTACCGGCA TGCACAGGCATTGCTGGGATCA
PAYGE 127. EB100. LV5_(PGI1t)_R	For generating ACT donor for cloning into pYES-URA vector	TGTTATAATTATTTTCTTATTTTGATGTA ATATAAAGAGGGGTATACTGGAGGCTTC AT
PAYGE 121. EB257. LV3_(TEF1p)_F	For generating ACT donor for cloning into pYES-URA vector	AGGAATACTCTGAATAAAACAACTTATA TAATAAAAATGCATAGCTTCAAAATGTT TCTACTC
PAYGE 148. EB160. LV5_pYES(70)_F	For generating pYES2-URA backbone for ACT cloning	CCTCTTTATATTACATCAAAATAAGAAA ATAATTATAACACCTGCATTAATGAATC GGCCAAC
PAYGE 149. EB161. LV3_pYES(70)_R	For generating pYES2-URA backbone for ACT cloning	GCATTTTTATTATATAAGTTGTTTTATTC AGAGTATTCCTACTAGTGGATCATCCCC ACGC
GC1776	For cPCR of proper cloning of ACT in pYES-URA	CCACTACGTGAACCATCACCC
193.seqTEF1p_F	For cPCR of proper cloning of ACT in pYES-URA	CGATGACCTCCCATTGATATTTAAG
FgF20_cPCR_F	For cPCR at FGF20	CGGAGTTATTGGATATACTGTGTAAACT
FgF20_cPCR_R	For cPCR at FGF20	CGTTAGTATCGTCAAAACACTCG
FGF18_cPCR_F	For cPCR at FGF18	CTGTTTTCAGTAGATTTGGTAACTGTGC AACC
FGF18_cPCR_R	For cPCR at FGF18	GAGCATTTCGTTCACTTACCAAACAATT AAGG
FGF24_cPCR_F	For cPCR at FGF24	CGTAGTGGCGATCTTGTGATTTTCGTAC
FGF24_cPCR_R	For cPCR at FGF24	GCACTGTGGATTCTATGTACTTGGCAAT AG
FGF7_cPCR_F	For cPCR at FGF7	GCGACTTTTGGTGGAATATTATGATATG TGTTG
FGF7_cPCR_R	For cPCR at FGF7	CAAATTATGATAAGAACCAATCATCATC CATCG
FGF19_cPCR_F	For cPCR at FGF19	GATTCCGCGCTTCCATCATTTAGTATAA TCC
FG19_cPCR_R	For cPCR at FGF19	GGTAATGTCAGTAATTAGCGGATGATAG TTGG
USERxii-1_cPCR_F	For cPCR at USERxii-1	CTTGGACTTCTTACCACCAGCAAG
USERxii-1_cPCR_R	For cPCR at USERxii-1	ATCAATCCTCGCATTTCAGCTTCC
USERxii-2_cPCR_F	For cPCR at USERxii-2	TCGAGAGAGTCGCCGATAGTGTTTTAGA GCTAGAAATAGCAAGT
USERxii-2_cPCR_R	For cPCR at USERxii-2	ACTATCGGCGACTCTCTCGAAAAGTCCC ATTCGCC
pCAS_F1	For amplifying pCAS-Tyr-Hyg or pCAS-TYR-G418 backbone	TAGGTCTAGAGATCTGTTTAGCTTG

pCAS_R1	For amplifying pCAS-Tyr-Hyg or pCAS-TYR-G418 backbone	GCATTTAAGCATAAACACGC
gRNA-FGF20_F	For amplifying guide RNA for FGF20	GTTAGAGCTGTTACAAGTTAGTTTTAGA GCTAGAAATAGCAAGT
gRNA-FGF20_R	For amplifying guide RNA for FGF20	TAACTTGTAACAGCTCTAAC AAAGTCCCATTCGCC
gRNA-FGF18_F	For amplifying guide RNA for FGF18	ATAGAATTACTATTGAAGAG GTTTTAGAGCTAGAAATAGCAAGT
gRNA-FGF18_R	For amplifying guide RNA for FGF18	CTCTTCAATAGTAATTCTAT AAAGTCCCATTCGCC
gRNA-FGF24_F	For amplifying guide RNA for FGF24	CCTATTGGACAAGATTTACG GTTTTAGAGCTAGAAATAGCAAGT
gRNA-FGF24_R	For amplifying guide RNA for FGF24	CGTAAATCTTGTCCAATAGG AAAGTCCCATTCGCC
gRNA-FGF7_F	For amplifying guide RNA for FGF7	ACTCCTGGGAGAGAACATTCGTTTTAGA GCTAGAAATAGCAAGT
gRNA-FGF7_R	For amplifying guide RNA for FGF7	GAATGTTCTCTCCCAGGAGTAAAGTCCC ATTCGCC
gRNA-FGF19_F	For amplifying guide RNA for FGF19	ATTCACTCTGCTAAGATTATGTTTTAGA GCTAGAAATAGCAAGT
gRNA-FGF19_R	For amplifying guide RNA for FGF19	ATAATCTTAGCAGAGTGAATAAAGTCCC ATTCGCC
gRNA-Userxii-1_F	For amplifying guide RNA for USERXii-1	AACAAACTTGTGTGCTTCATGTTTTAGA GCTAGAAATAGCAAGT
gRNA-Userxii-1_R	For amplifying guide RNA for USERXii-1	ATGAAGCACACAAGTTTGTTAAAGTCCC ATTCGCC
gRNA-Userxii-2_F	For amplifying guide RNA for USERXii-2	TCGAGAGAGTCGCCGATAGTGTTTTAGA GCTAGAAATAGCAAGT
gRNA-Userxii-2_R	For amplifying guide RNA for USERXii-2	ACTATCGGCGACTCTCTCGAAAAGTCCC ATTCGCC

3.2.3 Cas9 amplification and purification

CRISPR associated protein 9 (Cas9) vector was cloned into a pBOT plasmid and transformed into *E. coli* using heat shock protocol before being plated onto Luria Broth (LB) agar plate with the respective drug selection [70][71]. One colony was picked and grown overnight in 5mL of LB with resistance. Using ThermoFisher Plasmid DNA Miniprep Kit the plasmid was isolated.

3.2.4 CRISPR-Cas9 assisted homologous recombination

Yeast cells from one colony were picked from a fresh plate and inoculated into a 50mL glass tube with 5mL of 2x Yeast Extract Peptone Dextrose (YPD) and left for overnight growth at 30ºC and 200rpm. The cell density (OD600) of the overnight culture was determined using Tecan Infinite M200 and diluted to an initial concentration of $0.5x10^7$cells/mL into an Erlenmeyer flask containing 50mL of 2×YPD. The initial culture was grown at 30ºC and 200rpm until two doublings was reached. The cells were then harvested by centrifuging at 1,975g for 5 minutes. The cells were

washed once with 50mL of double distilled water (ddH_2O) and a second time with 5mL of 100mM lithium acetate. To each transformation well, 300ng of purified pCas9 plasmid, 600ng of purified gRNA plasmid, 1,000-2,000ng of DNA repair template, and ddH_2O was added for a final volume of 40µL. The final cell pellet was then suspended in 110µL of ice-cold transformation mix per reaction (per reaction: 100µl of 50% w/v PEG 3350, 5.6µL of 3M LiAc, 4.4µL of 10mg/mL salmon sperm DNA). A culture of 50mL gives 25 transformations. A 110µL volume of cell suspension was aliquoted into each well containing DNA and ddH_2O. The cells were mixed by pipetting. The

Table 3-15: Gene Integration Sites. The transcriptional activities of these integration sites, except for the site USERXII, were investigated by Bai Flagfeldt et. al [74]. The transcriptional activity of the USERXII site was investigated by Mikkelsen et al. [75].

Integration Site Name	Content	Integration Site	Chromosome
FGF20	MDH	YPCR∆9	XVI
FGF18	MDH	YORW∆17	XV
FGF24	MDH	PDC6	VII
FGF7	HPS-PHI	YERC∆8	V
FGF19	HPS-PHI	YORW∆22	XV
USERXII	HPS-PHI	795787..796720	XII

suspension was then incubated at 30ºC for 30 minutes followed by a 30-minute incubation in a 42ºC water bath. The suspension was then transferred to a 2mL deep-well plate and pelleted down at 1,975g for 5 minutes. The supernatant was then removed and the pellet was suspended in 600µL of YPD. The cell suspension was incubated for 16h at 30ºC with 450rpm shaking. For each well roughly $1/100^{th}$ (6µL) was plated onto YPD and either G418 or Hygromycin depending on the pCas9 variant.

3.2.5 Growth media

For routine propagation, *E. coli* was grown in LB and *S. cerevisiae* was cultured in YPD. When plates were needed agar was added at 1.5%w/v. Appropriate antibiotics (100µg/mL ampicillin, and/or 200mg/mL geneticin, and/or 200mg/mL hygromycin) were added to cultures and plates for plasmid maintenance. When required leucine was added at 380mg/L while all other amino acids were added at 76mg/L.

3.2.6 NAD(P)-dependent methanol dehydrogenase *in vitro* activity assay

To prepare the starter culture, cells from one fresh colony was picked and inoculated into a 50mL glass tube containing 5mL of Yeast Nitrogen Base (YNB) with added 1% glucose and supplemented with uracil, histidine, tryptophan, and leucine amino acid. The starter culture was

grown overnight at 30°C and 200rpm. The cell density (OD600) of the overnight culture was determined using Tecan Infinite M200 and used to inoculate an initial concentration of $5.0x10^6$cells/mL into a 250mL Erlenmeyer flask containing 50mL of YNB with added 1% glucose and supplemented with uracil, histidine, tryptophan, and leucine amino acids. The cultures were incubated for about 20 hours at 30°C and 200rpm before harvesting the cells by centrifugation at 1,975g for 5 minutes. The cell pellet was washed twice with ddH_2O and once with 50mM K_2HPO_4 pH 7.5 buffer. Finally, the cells were suspended in 700µL of protease inhibitor cocktail (Roche).

To a 2mL round bottom tube, 100µL of acid-washed glass beads (400µm) were added with 500µL of suspended cells. Using a bead beater (Precellys) kept at 4°C, the cells were lysed using two times 3 cycles of 30sec at 6,000rpm with 30 second rest between each cycle and 5 minutes rest before repeating. The cells were verified for proper lysing under the microscope. The lysate was then spun down at 16,873g and 4°C three times, once for 5 minutes the second time for 15 minutes and the third time for 30-45 minutes. The protein concentration of each lysate was determined and normalized using the Thermo Fisher Coomassie Kit with bovine serum albumin (BSA) as a standard. For enzyme assays, 30µL of lysate (~16-26mg of proteins) was added to 7.5µL of 5mM $MgSO_4$, 30µL of 500µM NAD, and buffer for a total volume of 270µL. NADH production was recorded using a Tecan Infinite M200 set at wavelength 340nm. First, the reaction was monitored without the addition of methanol until it plateaued (10-15 minutes). At that time, 30µL of 1M methanol was added to each well and the reaction was monitored for 30 minutes. Each reading was repeated three times and averaged.

Using the law of Lambert–Beer ($\varepsilon_{NADH} = 6220M^{-1}cm^{-1}$) the enzyme activity was determined. One unit (U) was defined as the amount of enzyme that is required to process 1mmol of substrate per minute. Also, each dataset was normalized to mg proteins. The variance between the triplicates were determined by calculating the standard error, $SE = {}^{\sigma}/_{\sqrt{N}}$. The initial velocities were used to compare the different NAD(P)-dependent MDHs' efficiencies toward methanol. The initial velocities were determined by averaging the linear region of the activity curve while keeping a correlation higher than 0.9 for most enzymatic activities. The interval [0, 170.5] was determined as the initial velocity segment. For each NAD(P)-dependent MDH variant, we performed a one-

way ANOVA test with an $\alpha = 0.05$ to determine the significance between the different copy numbers. The same procedure was followed for the experiment with the activator protein.

3.2.7 Effects of methanol on growth of *S. cerevisiae*

Cells from one colony were picked from a YPD plate and inoculated into a 50mL glass tube containing 5mL of YNB with added 5% glucose and supplemented with uracil, histidine, tryptophan, and leucine amino acid. The starter culture was grown overnight at 30°C and 200rpm. The cell density (OD600) of the overnight culture was determined, and a dilution was made so that the inoculum would contain $5.0x10^7$cells/mL. To each well of a 96-well plate the following was added: 140µL of YNB with added 1% glucose and supplemented with uracil, histidine, tryptophan, and leucine amino acid, 20µL of methanol (at different concentrations), and 20µL of cell culture. Each strain was tested for growth under four different methanol concentrations – 1M, 100mM, 5mM, and 0mM – and for each concentration, growth from triplicate cultures were recorded. The cell density (OD600) was then measured on a Tecan Sunrise every 20 minutes for a 5-day period. The temperature was kept at 30ºC for the duration of experiment and the plate was constantly being shaken to keep cells in suspension. The area under the exponential phase of each growth curve was used to determine to the total growth. To determine the percent growth difference, each of the total growths were subtracted from their respective growth value in the absence of methanol.

3.2.8 ^{13}C-labeled methanol and flux analysis

To prepare the starter culture, cells from one colony were transferred from a YPD plate into a 50mL glass tube containing 5mL of YNB with added 1% glucose and supplemented with uracil, histidine, tryptophan, and leucine amino acid. The cultures were grown for 24 hours at 30°C and 200rpm. The cultures were then harvested by centrifugation at 1,975g for 5 minutes and washed twice with 5mL ddH_2O. The density of the culture was measured at OD660 and used to inoculate an initial confluency of 0.7 $x10^7$cells/mL in 5mL of YNB or YNB with 1g/L yeast extract with added 200mM methanol and supplemented with uracil, histidine, tryptophan, and leucine amino acid. After 24 hours, cells from 1mL of culture were transferred to a 1.7mL Eppendorf tube and harvested at 16,873g for 5 minutes. The supernatant was discarded and 0.7mL of pure methanol -40ºC was added to the pellet. The cells were mixed by vortexing the mixture for 30 seconds followed by a 30 minutes incubation in a dry ice bath made with 75% ethanol. The cells were then

centrifuged at 16,873g and 4ºC for 20 minutes. The supernatant was then transferred to a 3kDa 96-well AcroPrep filter plate and spun for 15 minutes. The flow through was collected and transferred to a 1.7mL Eppendorf tube. The supernatant was evaporated to dryness in a speedvac at room temperature. The crystals were then suspended in 100uL of pure methanol with 0.1% formic acid. The mixture was sonicated for 30 minutes to allow for good mixing prior to being analyzed. Samples were separated using the Agilent LC 1100 system with the Phenomenex Synergi 4 micron Hydro-RP 150 x 2 mm column. An injection of 10µL was used per run. Each run lasted 25 minutes with a constant flow rate of 300µL/min. Each run made use of two solutions: solution A containing 0.1 formic acid in ddH_2O and solution B containing 0.1% formic acid in acetonitrile. The run was started with 100% solution A for the first 5 minutes followed by 90% solution B for 20 minutes. The LC was hooked up to a mass spectrometer (MS) 7 tesla LTQ-FTMS from ThermoFisher. The FTMS scan was set from 50 to 210 m/z at 50000 resolution and 200 m/z with automatic gain control set for 500 msec maximum trap fill time or 1x10e6 ions.Using the Agilent software, the peaks associated with amino acids were identified and the relative abundances of non-labeled (m) and labeled (m+1) were determined.

4 RESULTS

4.1 ATP and growth prediction using the iAV3000 model

In comparison to other available models, the iAV3000 model, the model used in this study, focuses on reactions found in the cytoplasm, mitochondrial, and extracellular (Table 4-1).

Table 4-1: Comparison of total reactions and compartments of different *S. cerevisiae* models

Model	Total Count Rxns	Mitochondrial Rxns	Cytosolic Rxns	Cytosolic Exchange Flux	Mitochondrial Exchange fluxes
iFF708	1175	104 (8.85%)	723 (61.53%)	286	62
iLN800	1446	161 (11.13%)	906 (62.66%)	304	75
iAV3000	798	72 (9.02%)	336 (42.11%)	245	83

Constraint-based metabolic models like the iAV3000 model do not reflect everything that is going on inside a living cell. For example, there are no regulation on the transcription and translation of genes, which would affect K_{cat} values, rate of turnover of a metabolite. Therefore, bottlenecks are

not created by the concentration of available enzymes and increasing copy number of an enzyme will not change the outcome of the results. Additionally, toxicity of metabolites, such as methanol or formaldehyde, is not quantifiable in a metabolic model. There is currently no method used to model toxicity: providing a high concentration of methanol to the metabolic model, which in our case is used as carbon source, would result in higher biomass production when in reality it could have the opposite effect. It is with understanding of the limitation of the model that we begin to predict certain outcomes.

4.1.1 *In silico* ATP production under glucose and methanol carbon source

Generating chemical energy in the form of ATP is essential for the survival of *S. cerevisiae*. The model needs to generate the right amount of ATP molecules per mole of carbon source [48].

Table 4-2: ATP production predicted *In silico* under glucose and methanol carbon source. The Pichia pastoris column is used as a reference to compare the engineered strain to a native methylotrophic yeast.

Theoretical ATP Production	*P. pastoris*	*S. cerevisiae* (Engineered Strain)
Per Glucose (ATP/mole carbon)	1.92	1.92
Per Methanol (ATP/mole carbon)	2.75	4.41

Theoretical mole ATP production per mole of glucose has been established at 28.1 [48]. However, the P/O ratio, which is dependent on the microenvironment of the organism such as temperature and salinity and thus reflects a more appropriate estimate of the mole ATP production per mole of glucose [72]. In *S. cerevisiae* the P/O value is close to one, which experimental results puts it at around 12.5 mole ATP per mole glucose [72]. The *in silico* model was modified in order to satisfy both ATP production and P/O ratio. The iAV3000 *in silico* P/O ratio is 0.98 and produces 11.5 moles of ATP per mole of glucose. Once the model was able to output the correct amount of ATP per mole of glucose we introduced the foreign reactions that makes up the RuMP cycle and ran the model again, however this time changing the carbon source to methanol. Since, glucose consists of six carbon atoms we decided to normalize the results as per carbon atoms (Table 4-2). The *in silico* model predicts higher ATP production under methanol. Once the model predicted positive ATP production under methanol carbon source, we then looked at the predicted growth.

4.1.2 *In silico* growth prediction under glucose and methanol carbon source

ATP production is one essential part of the survival of *S. cerevisiae*. But, for growth to occur there must also be a good balance between all molecules of the system. This is dictated by the different reactions inside the system, as well the growth constraints placed on the system. Using the growth equation established in (Section 3.1.3), we modeled the growth output under different levels of glucose-limited chemostat (Table 4-3).

Table 4-3: Comparison of the predicted aerobic glucose-limited chemostat growth rate results to published model.

q_{GLC} mmol/gDCW/h	1.15		1.17		1.69		2.26		3.29	
	Model	**%**	**Model**	**%**	**Model**	**%**	**Model**	**%**	**Model**	**%**
iFF708	0.1103	10.3	0.1123	12.3	0.1621	8.0666	0.2168	8.4	0.3157	12.75
iND750	0.0263	73.7	0.0267	73.3	0.0386	74.266	0.0516	74.2	0.0752	73.142
iIN800	0.1071	7.1	0.109	9	0.1601	6.7333	0.2161	8.05	0.3174	13.357
iMM904	0.0221	77.9	0.0225	77.5	0.033	78	0.0445	77.75	0.0652	76.714
Yeast 4	0.1883	88.3	0.1916	91.6	0.2767	84.466	0.3701	85.05	0.5387	92.392
iAZ900	0.0373	62.7	0.038	62	0.0548	63.466	0.0733	63.35	0.1067	61.892
iMM904bs	0.0221	77.9	0.0225	77.5	0.033	78	0.0445	77.75	0.0652	76.714
Yeast 5	0.0909	9.1	0.0924	7.6	0.1335	11	0.1786	10.7	0.2599	7.1785
iTO977	0.1065	6.5	0.1085	8.5	0.1593	6.2	0.215	7.5	0.3157	12.75
Yeast 6	0.1096	9.6	0.1116	1(1.6	0.1611	7.4	0.2155	7.75	0.3137	12.035
Yeast 7	0.1104	10.4	0.1123	12.3	0.1622	8.1333	0.2169	8.45	0.3158	12.785
iAV3000	0.0899	10.05	0.1321	32.19	0.1767	17.85	0.2252	12.63	0.2573	8.0931
Dilution rate h^{-1}	0.1		0.1		0.15		0.2		0.28	

The percent sign (%) represents percent error between the respective model and the experimental growth rate.
The table was adapted from Heavner (2015) [92].

Using experimental results found in the literature we can compare our predicted results (Table 4-3). The values outputted from this experiment have between 8 and 32 percent error in reference to experimental values (Table 4-3). Since these percent error were in the range of previously published models, we decided that these growth values were sufficient to predict growth under methanol.

When we tested for the growth rate under methanol certain reactions were removed. This is because transcription of certain genes are only available while others are repressed in the presence of glucose [78]. However, the biomass reaction was left unchanged. Our model prediction under methanol carbon source in comparison to expected values from a previously published *P. pastoris* model is similar at low dilution rates (**Error! Reference source not found.**). At higher dilution r ates, our model predicts much lower growth rates than that of *P. pastoris* (**Error! Reference source not found.**).

Table 4-4: Methanol growth comparison between *P. pastoris* and *S. cerevisiae*.

Dilution Rate h^{-1}	q_{MET} mmol/gDCW/h		q_{MET}* mmol/gDCW/h		Expected mmol/gDCW/h	Ref
	Model	%	Model	%		
0.31	0.7916	19.94	0.99	50.30	0.66	[93]
0.54	0.7958	30.80	1.73	50.70	1.15	
0.66	0.7979	28.12	1.67	50.31	1.11	
1.20	0.8072	64.12	3.39	50.71	2.25	
1.62	0.8145	78.16	5.64	51.10	3.73	

* The P. pastoris model iLC915 was taken from Caspeta (2012) [79] and the table was adapted from Morales et al. [94].
The percent sign (%) represents percent error between the respective model and the experimental growth rate.

This could be caused by the differences in the biomass reaction. While taking a closer look at the biomass reaction that the iLC915 model uses, we found a reduction in every macromolecule used in comparison to our model (**Error! Reference source not found.**). Most notably, we can see a l arge difference in the protein and ATP requirements, which is most likely the cause behind the discrepancy of the two models [79]. However, previous studies on the toxicity of methanol in *S. cerevisiae* reveals an induction of 375 ORFs and the repression of 314 ORFs, suggesting that protein synthesis is increased in the presence of methanol, and thus requires additional ATP [80].

Table 4-5: Comparison of the biomass composition.

iAV3000 mmol/gDCW/h	iLC915 mmol/gDCW/h	Macromolecule
0.5	0.37	Protein
0.31	0.369	Carbohydrates
0.1	0.0583	RNA
0.01	0.0013	DNA
0.03	0.0756	Lipids
59.276	35.7	ATP

Comparison of the predicted growth under methanol and glucose carbon source from the iAV3000 model shows higher biomass production than under glucose carbon source (Figure 4-1). Possibly, this is because of the increase of available NADH in the cell.

Therefore, we introduced the foreign reactions and changed the carbon source to methanol. The model predicts higher growth under methanol carbon source (Figure 4-1). We believe this is because the extra NADH pool generated from methanol assimilation is then used to drive ATP production, which is then used towards biomass synthesis.

The model predicted both ATP production and growth under methanol and glucose carbon source. With this information, we decided to add the necessary genes to complete the RuMP cycle in *S. cerevisiae*.

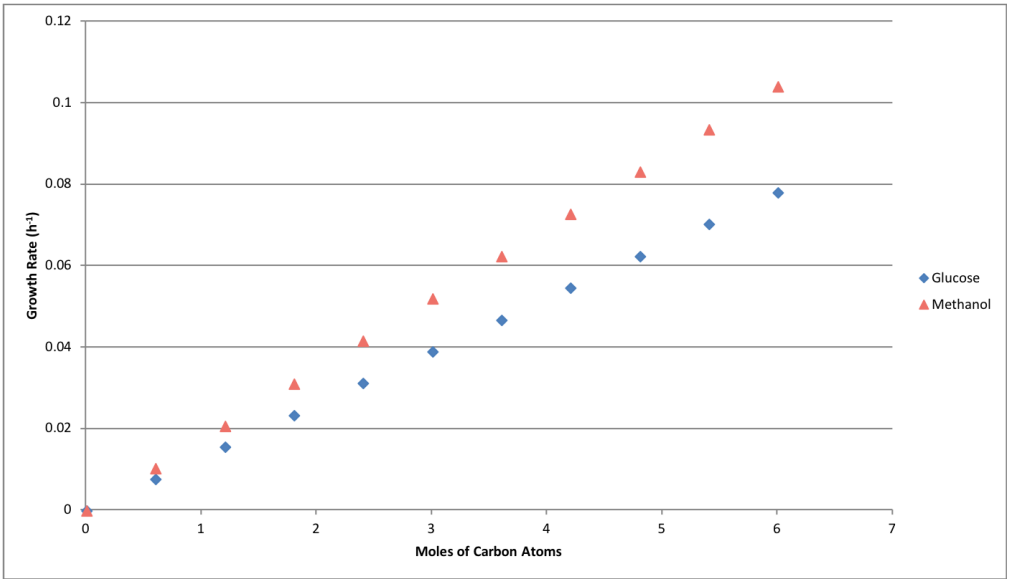


Figure 4-1: Growth values under glucose and methanol carbon source. Blue line show growth levels under glucose, normalized to moles of carbon atoms. Orange line show growth levels under methanol, normalized to moles of carbon atoms.

4.2 NAD(P)-dependent methanol dehydrogenase *in vitro* activity

The first enzymatic step in the methanol assimilation pathway is the conversion of methanol to formaldehyde [17]. A variety of different enzymes can accomplish this conversion. In this study we decided to test six NAD(P)-dependent MDH variants (Table 4-6).

Table 4-6: Characteristics of the different engineered strains.

Strain Name	Enzyme Added	Origin of Enzyme
MM1	Methanol Dehydrogenase found in the pMB19 plasmid Original name: MDH^M	*B. methanolicus* MGA3
MM2	Methanol dehydrogenase found in the genome Original name: $MDH2^M$	
MM3	Methanol dehydrogenase found in the genome Original name: $MDH3^M$	
MP1	Methanol Dehydrogenase found in the pMB20 plasmid Original name: MDH^P	*B. methanolicus* PB1
MP2	Methanol dehydrogenase found in the genome Original name: $MDH1^P$	
MP3	Methanol dehydrogenase found in the genome Original name: $MDH2^P$	

Using CRISPR methodology we introduced each variant in different copy numbers – one, two and three copies – into the genome of a *S. cerevisiae* CEN P.K 113D strain [73]. Three different sites – YPRCτ3 (FGF20), YORWΔ17 (FGF18), PDC6 (FGF24) – were chosen based on a previous study that reported high transcription activity at these sites [74]. Each strain was lysed and the total protein lysate was determined (Table 4-7).

Table 4-5: Protein concentration of the different strain lysate.

Strain	Protein Concentration [mg/mL]
MM1	1.3327
MM2	1.0937
MM3	1.1259
MP1	2.1902
MP2	2.1628
MP3	1.9022
WT	1.1255

We then performed *in vitro* activity assays on the NAD(P)-dependent MDHs. We first added all the necessary reagents except methanol and allowed all initial reactions to proceed. Once the conversion of NAD to NADH reached a plateau we added 1M of methanol and continued to measure the process. This allowed us to remove all background reactions and focus on the NAD(P)-dependent MDHs. Based on the p-values we can conclude that the addition of a gene copy – 1 vs. 2 vs. 3 – significantly increases the initial velocity of the assay (Table 4-).

Table 4-7: Slope significance between different copy numbers of the same MDH family.

Strain	p-value
WT	P = 0.0536
MM1	P < 0.0001
MM2	P < 0.0001
MM3	P = 0.0421
MP1	P < 0.0001
MP2	P < 0.0001
MP3	P < 0.0001

The raw data and the graphs of each NAD(P)-dependent MDHs can be found in the appendix (Section 7.1). The linear portion of the graph was used to determine the initial velocity of the enzyme. Results of the activity assays provide evidence of low wild-type methanol oxidation activity (Figure 4-2). Possibly, this is because of promiscuous alcohol dehydrogenases activity. The different *B. methanolicus* PB1 NAD(P)-dependent MDHs show significantly lower activity than those from *B. methanolicus* MGA3 (Figure 4-2). The increase in enzyme copy significantly increases the overall turnover rate of methanol (Figure 4-2).

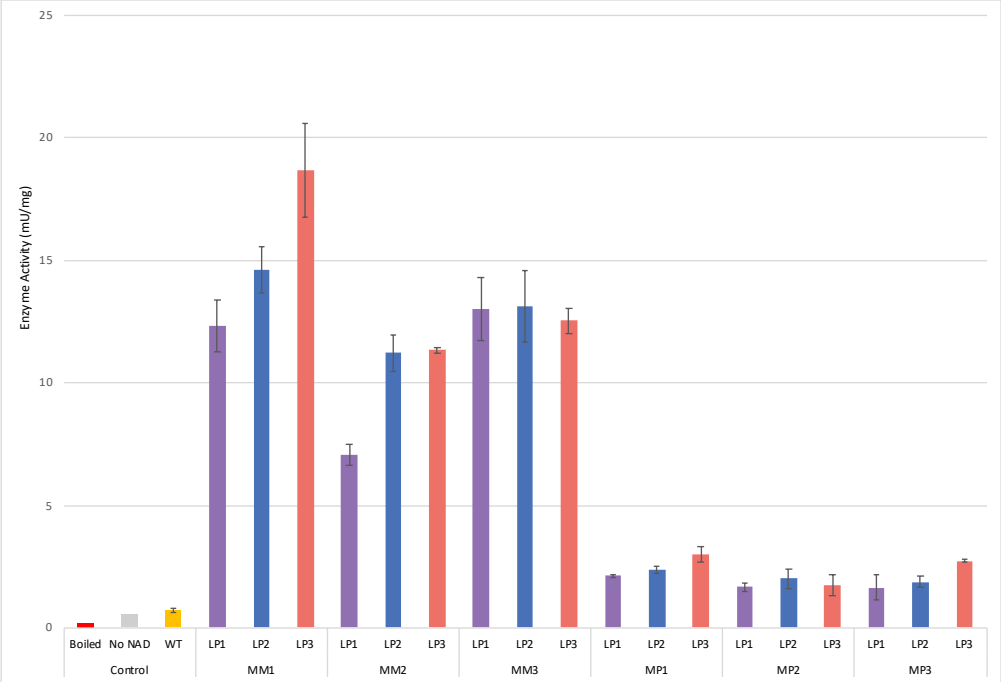


Figure 4-2: NAD(P)-dependent MDH activity comparison. In the first segment we have boiled lysate, then we have WT with no NAD added and finally we have WT. For each methanol dehydrogenases there are three bars title LP1, LP2 and LP3 to represent 1 copy, 2 copies and 3 copies of the genes respectively.

We conclude that MM1 has the highest activity towards methanol (Figure 4-2). For each strain, as a control, we also tested the enzyme activity in the absence of methanol, which can be found in

the appendix. The data reveals that enzyme activity observed is in fact due to methanol oxidation. The presence of an activator protein has previously been shown to increase the activity of the NAD(P)-dependent MDH [28]. The following section will investigate the effects of this activator protein on the six NAD(P)-dependent MDH variants.

4.3 In vitro effects of activator protein on methanol dehydrogenase activity

The activator protein has been observed to increase the catalytic activity of the NAD(P)-dependent MDHs by enabling the enzyme to perform methanol dehydrogenation in one step rather than via a ping-pong reaction mechanism [28]. Although the activator protein is from a family of protein ubiquitously found in nature, we wanted to test the effect of this activator protein on the activity of the six NAD(P)-dependent MDH variants.

The activator proteins were first cloned into a high copy number plasmid, pYES2-URA; this plasmid harbors the URA3 gene. Then, the plasmid was transformed into each strain containing three NAD(P)-dependent MDH copies. The same protocol was employed as in Section 4.2. The lysate was used to determine the concentration of proteins and to then normalize each data point (Table 4-).

Table 4-7: Protein concentration of the different strain lysate.

Strain	Protein Concentration [mg/mL]		
	No Plasmid	With Activator Protein	Empty Plasmid
WT	1.8456	1.5952	2.0119
MM1	1.9129	1.5765	1.5882
MM2	1.1724	2.2356	1.7413
MM3	1.1367	1.2341	1.1907
MP1	1.0407	1.0002	1.2158
MP2	1.5695	1.1846	1.5259
MP3	1.9485	1.4276	1.6672

As mentioned earlier, the linear portion of the graph was used to determine the initial velocity of the enzyme. The same interval was used as earlier. To determine the ACT protein significance on the NAD(P)-dependent MDH activities, the slopes of the initial velocities were compared using one-way ANOVA test. Within each group, no significant differences between the slopes were observed and thus the ACT protein does not appear to have any significant effects on the different NAD(P)-dependent MDH activities (Table 4-).

Table 4-8: The slope significance of the different MDH strains. The p-value represents the significance between the initial velocities of the strains with the Activator protein, with an empty plasmid, and with no plasmid.

Strain	p-value (vs. ACT)
WT	P = 0.2886
MM1	P = 0.3061
MM2	P = 0.3944
MM3	P = 0.5968
MP1	P = 0.4787
MP2	P = 0.9648
MP3	P = 0.3253

Even though we could not see an increase in NAD(P)-dependent MDH activity in the presence of the ACT protein, we could still see positive NAD(P)-dependent MDH activity (Figure 4-4). Therefore, the activator protein does not hinder the NAD(P)-dependent MDH activity.

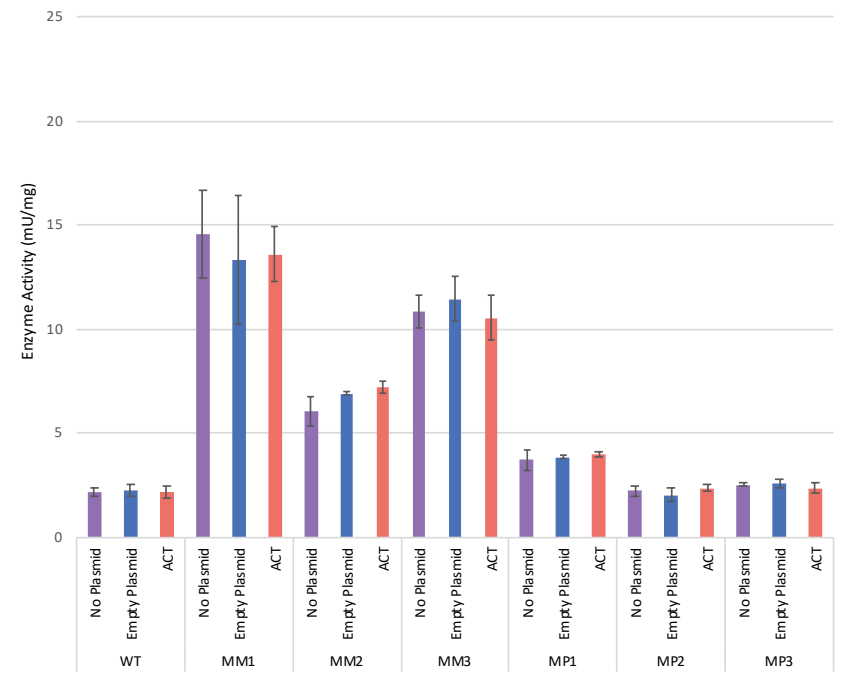


Figure 4-3: Effect of activator protein on various NAD(P)-dependent MDH. Purple bars are strains which were not transformed with any plasmid. The blue bars are strains that have been transformed with a PYES2-URA vector carrying no gene cassette. The orange bars are strains that have been transformed with a PYES2-URA vector carrying the activator protein gene cassette.

Sections 4.2 and 4.3 provide evidence of the NAD(P)-dependent MDH activities and therefore indirectly provides evidence of methanol oxidation. The next step was to introduce the next enzymes of the pathway, HPS and PHI, and to then test for the pathway activity.

4.4 Introduction of 3-hexulose-6-phosphate synthase and 3-hexulose-6-phosphate isomerase and its effect on methanol tolerance

To complete the engineered pathway, two additional genes, HPS and PHI, need to be added to the MM1 strain. As previously mentioned, two 3-hexulose-6-phosphate synthase genes and two 3-hexulose-6-phosphate isomerase genes were chosen from two separate bacterial species (Table 4-8). We introduced the two gene cassettes into the genome using CRISPR. Both genes where introduced in tandem at the sites: YERC∆8, YORW∆22, and USERXii, providing three copies of each genes. [74][75]. To test both HPS and PHI activities, the growths of our mutant strains were monitored in the presence and absence of methanol. The following graphs will represent the differences in growths in the absence of methanol and the respective methanol concentrations: 5mM, 100mM and 1M. Thus, each bar will be normalized to its own growth in the absence of methanol.

Table 4-8: Characteristics of the different engineered strains containing HPS and PHI genes.

Strain Name	Enzyme Added	Origin of Enzyme
HPS.M	3-hexulose-6-phosphate synthase	*B. methanolicus* MGA3
PHI.M	3-hexulose-6-phosphate isomerase	
HPS.M	3-hexulose-6-phosphate synthase	*B. methanolicus* MGA3
PHI.P	3-hexulose-6-phosphate isomerase	*B. methanolicus* PB1
HPS.P	3-hexulose-6-phosphate synthase	*B. methanolicus* PB1
PHI.P	3-hexulose-6-phosphate isomerase	*B. methanolicus* MGA3
HPS.P	3-hexulose-6-phosphate synthase	*B. methanolicus* PB1
PHI.P	3-hexulose-6-phosphate isomerase	

Strain names may include HPS and PHI, the name of the strain will be added to the name of the background strain. As an example, wild-type strains with both HPS and PHI from *B. methanolicus* MGA3 will be named WT HPS.M PHI.M, and MM1 background strains with both HPS and PHI from *B. methanolicus* PB1 will be name MM1 HPS.P PHI.P.

4.4.1 WT parental strain

As a control, the HPS and PHI genes were introduced into the wild-type strain. It appears that at 5mM methanol concentration, the cells can harvest the carbon toward biomass synthesis, shown by the positive differences (Figure 4-4).

This could be explained by the promiscuous activity of alcohol dehydrogenase on methanol, which would produce formaldehyde. However, yeast native formaldehyde detoxification pathway, the formaldehyde oxidation II pathway, can rid the cell of formaldehyde. By doing so, for every molecule of formaldehyde being detoxified, one molecule of NADH is produce, and this cofactor

Figure 4-4: WT growth in the presence of methanol. The WT strain represents the unmodified wild-type strain. All strains have been normalized to their own growth under no methanol, therefore the graphs represent the effects of methanol on the cells. For name reference to Table 4-8.

can be used to generate energy used to drive biomass synthesis, which would explain the positive growth difference in WT observed at 5mM methanol concentration. For all other strains both the formaldehyde oxidation II pathway and the engineered pathway is at work. Above 5mM methanol concentration, we observe a negative growth difference, which can be attributed to the methanol toxicity (Figure 4-4). The activity of HPS and PHI are best seen at 1M methanol concentration, where enough methanol is converted to formaldehyde allowing for the observation of the activity of those two enzymes. The combination of HPS.P and PHI.M and of HPS.P and PHI.P appears to be best at rescuing the cell from methanol toxicity.

4.5 MM1 parental strains

The different HPS and PHI combinations were also introduced in the strain harboring three MM1 copies. Differences between wild-type and MM1 growths at 5mM methanol concentration provides additional evidence of MM1 activity since we are observing a negative growth difference at 5mM methanol concentration for the MM1 strain (Figure 4-5). This is because enough methanol is converted to formaldehyde, which only a portion of it can be handled by the formaldehyde oxidation II pathway. Thus, some formaldehyde never gets detoxified and leads to cell toxicity. The positive difference at 5mM methanol concentration for the strains with HPS and PHI suggest methanol utilization by the recombinant strain (Figure 4-5). The decrease in growths above 5mM methanol concentration, suggests possible methanol and/or formaldehyde toxicity (Figure 4-5).

The combination of HPS.P and PHI.M and HPS.P and PHI.P have the largest positive growth difference which suggest to be better at metabolizing methanol in comparison to the other strains (Figure 4-5).

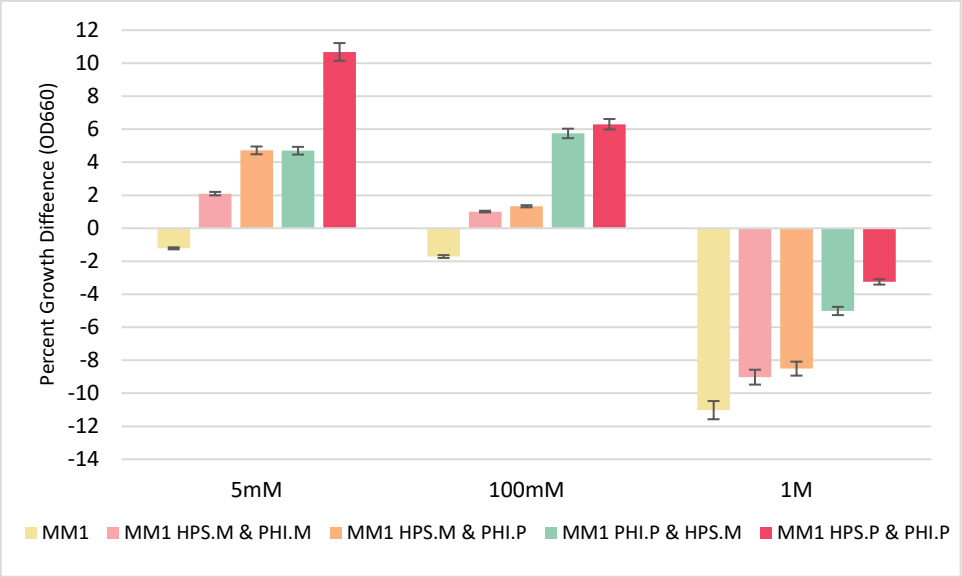


Figure 4-5: MM1 growth in the presence of methanol. The MM1 strain represents the strain harboring three copies of the MDH1.M gene. All strains have been normalized to their own growth under no methanol, therefore the graphs represent the effects of methanol on the cells.

4.5.1 Deletion of SFA1 and its effects on methanol tolerance

Due to the high toxicity of formaldehyde, many microorganisms have evolved a pathway to rid the cell of it [31]. In yeast, the formaldehyde oxidation II pathway is the major route of formaldehyde detoxification [76]. Since we want to funnel formaldehyde in a different direction, we have decided to delete SFA1 in the formaldehyde producing strains. The SFA1 enzyme is a bifunctional alcohol dehydrogenase and formaldehyde dehydrogenase, and it is the first enzyme of the formaldehyde oxidation II pathway [71][72]. We hoped that by removing SFA1, there would be an increase in carbon flux through our engineered pathway, and ultimately increased biomass synthesis.

4.5.1.1 WT SFA1 parental strains

We first looked at the effects of SFA1 in the wild-type strain. At 5mM methanol concentration, the absence of SFA1 appears to have little effect on wild-type growths (Figure 4-6). However, we

begin to see more of the effects of SFA1 at higher methanol concentrations (Figure 4-6). This is also apparent if one compares the chart with that of the wild-type with SFA1 (Figure 4-4).

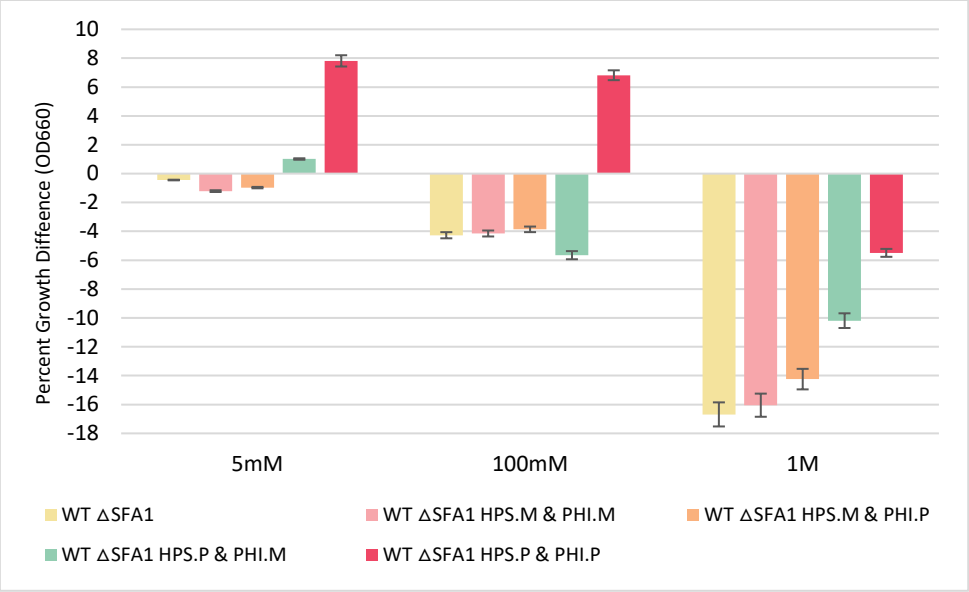


Figure 4-6: WT ΔSFA1 growth in the presence of methanol. The WT ΔSFA1 strain represents the wild-type strain with SFA1 deletion. All strains have been normalized to their own growth under no methanol, therefore the graphs represent the effects of methanol on the cells.

Additionally, the comparison to the wild-type suggest endogenous methanol to formaldehyde conversion. The data presented suggests that SFA1 is helping in the detoxification of formaldehyde since we are observing greater negative growth difference in the absence of SFA1 (Figure 4-6). It must be noted that the strain harboring HPS.P and PHI.P appears to be active at both 5mM and 100mM methanol concentrations (Figure 4-6). This could be explained by the fact that the promiscuous activity of the alcohol dehydrogenase on methanol generates a pool of formaldehyde, and because SFA1 has been deleted from this strain, formaldehyde has no where to go other than down the engineered pathway. As such we are seeing positive growth difference in the wild-type strain having both HPS and PHI from *B. methanolicus* PB1. The reason we do not observe positive difference in the other strains may be due to lower K_M values of those HPS and PHI, and thus cannot funnel this low pool of formaldehyde down the engineered pathway. In accordance with the previous figures, the combinations of HPS.P and PHI.M and HPS.P and PHI.P, appears to rescue the cell better than other combinations of HPS and PHI genes at rescuing the cell from methanol toxicity.

4.5.1.2 MM1 SFA1 parental strains

Finally, SFA1 was removed from the different MM1 strains. In comparison to the MM1 (Figure 4-5), we can see a greater overall positive trend suggesting that the absence of SFA1 helps to funnel methanol towards the glycolysis pathway, and thus assimilate more methanol for biomass

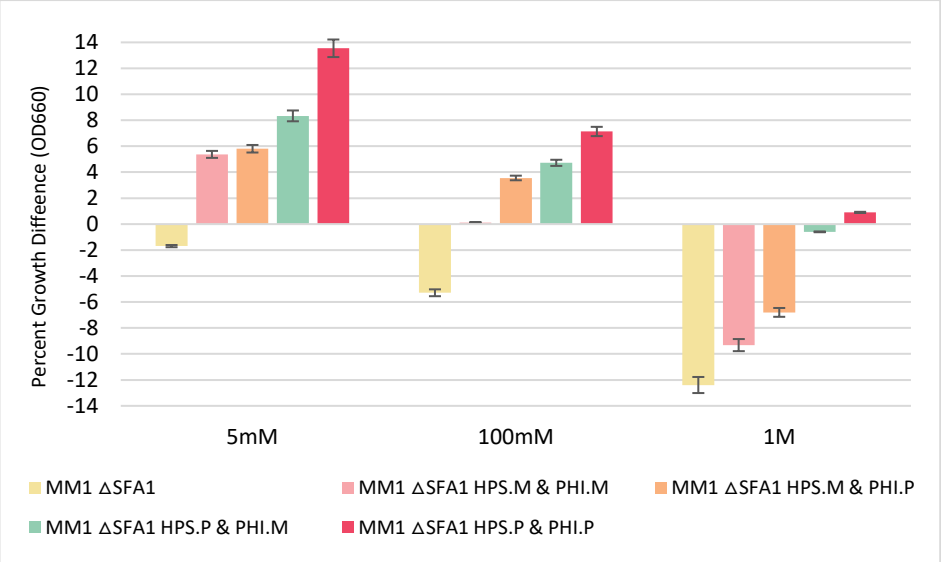


Figure 4-7: MM1 ΔSFA1 growth in the presence of methanol. The MM1 ΔSFA1 strain represents the strain harboring three copies of the MDH1.M with the deletion of SFA1 gene. All strains have been normalized to their own growth under no methanol, therefore the graphs represent the effects of methanol on the cells.

synthesis (Figure 4-7). The MM1 strain that harbors HPS.P and PHI.P (MM1 ΔSFA1 HPS.P PHI.P)has a positive difference in growth at 1M methanol, suggesting that the SFA1 pathway is highly active and robs our engineered pathway from additional carbon flow (Figure 4-7). From these results, during the absence of all three genes we observe negative biomass growth difference. This suggests that all three enzymes are needed to complete the engineered pathway and funnel methanol for biomass synthesis. Of the four different combination of HPS and PHI genes, higher growth is observed for HPS.P and PHI.P, suggesting higher activity of these enzymes.

In conclusion, we have compared the activity of twenty-four different strains: six wild-type background, six wild-type ΔSFA1 background strains, six MM1 background strains, and six MM1 ΔSFA1 background strains. We tested four different growth environments: no methanol, 5mM methanol, 100mM methanol, and 1M methanol.

Overall, as the concentration of methanol increases the fitness of the cells decreases. However, for most strains that contain all three genes necessary to complete the engineered pathway, the addition of 5mM and 100mM methanol lead to higher growth. Even higher growth was observed when SFA1 was deleted from the host strain, this is because more methanol was being funnelled through the engineered pathway. This suggests that methanol is being either used for biomass synthesis, or that methanol is being detoxified allowing for greater glucose assimilation. At 1M methanol concentration, we observe a negative difference in growths, suggesting that this methanol concentration hinders the fitness of the cells.

Since the combination of MM1 with HPS and PHI from *B. methanolicus* PB1 outperformed the other strains we decided to test whether we could observe assimilation of ^{13}C-labeled methanol in biomass components.

4.6 ^{13}C-labeling of amino acids

Direct testing for the ability of the engineered strains to assimilate methanol for biomass production was accomplished by feeding ^{13}C-labeled methanol and quantifying the fraction of labeled to unlabeled amino acids. To obtain a steady state, each strain was pre-grown in 1% glucose for 24 hours before being washed and transferred into YNB or YNB with yeast extract. The experiment was started when the addition of 200mM ^{13}C-labeled methanol (99%) was added to the media. After twenty-four hours the strains were quenched and lysed, and the intracellular metabolites were concentrated and analyzed via LC-MS. The difference between labeled and unlabeled amino acids were calculated (Figure 4-8).

Not all amino acids were detected via this method and from the ones that were detected only three amino acids – glutamate, aspartate and alanine – showed incorporation of ^{13}C-labeled methanol. Of the three amino acids that were found to be labeled, alanine showed the greatest incorporation of labeled carbon. As well, the strain MM1 ΔSFA1 HPS.P PHI.P was able to incorporate the most methanol, which is shown by the amount of labeled amino (Figure 4-8). The effects of SFA1 towards formaldehyde detoxification and ultimately on methanol assimilation is shown by the increase of labeled amino acids in its absence (Figure 4-8).

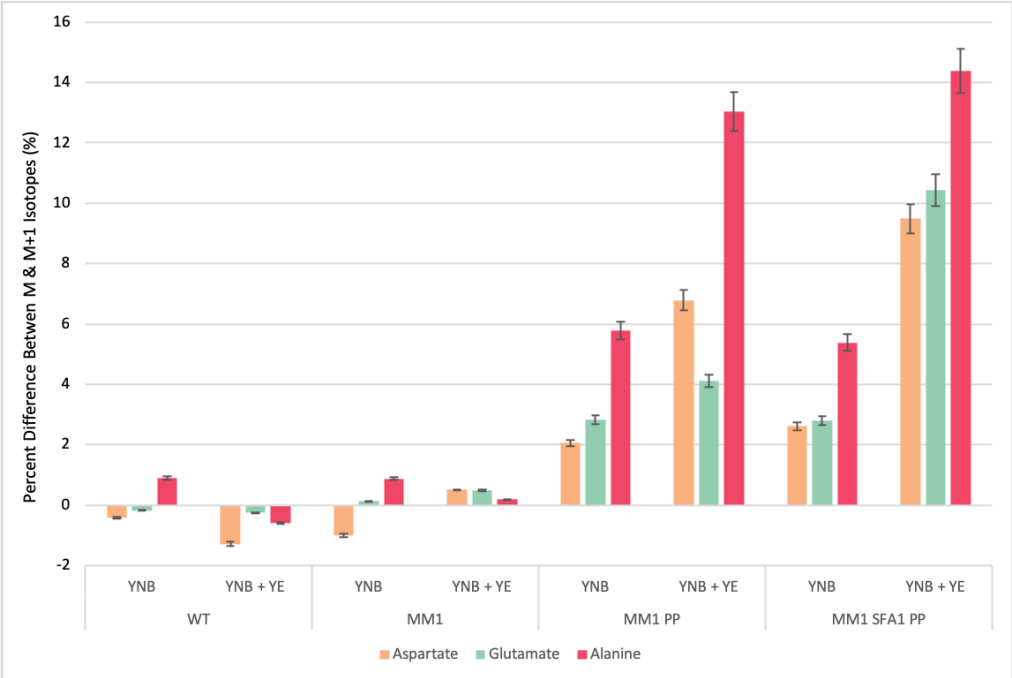


Figure 4-8: ^{13}C-labeled methanol incorporation into amino acids. Four different strains were tested for incorporation of ^{13}C-labeled methanol – WT, MM1, MM1 HPS.P PHI.P and MM1 ΔSFA1 HPS.P PHI.P – in two different media, YNB and YNB with 1g/L yeast extract. Each strain was grown in duplicates, in both labeled and unlabeled methanol. The ratio between labeled and unlabeled methanol was found and shown in this figure.

In summary, the results provide evidence of the necessity of all three genes – MM1, HPS, and PHI – for a complete methanol assimilation pathway and for the observation of labeled amino acids. Different NAD(P)-dependent MDH enzymes were tested for their ability to act on methanol (Section 4.2) and in conclusion we have shown that MM1 variant has greatest activity. Finally, to provide evidence that methanol is indeed being used for biomass synthesis we supplied our engineered strains with ^{13}C-labeled methanol and looked at amino acid labeling (Section 4.6). Evidence suggests that both MM1 HPS.P PHI.P and MM1 ΔSFA1 HPS.P PHI.P strains can assimilate methanol for biomass production (Figure 4-8). Also, we provide extensive evidence on the activity of SFA1 and conclude that its absence benefits the growth of our engineered strains in the presence of methanol (Figure 4-6 and Figure 4-7).

5 DISCUSSION

Saccharomyces cerevisiae is a widely used model organism that is used to produce various bioproducts among other things, amino acids, polymers, and pharmaceutical bioproducts [18]. To industrialize these bioprocesses it must be economically feasible. Current method of producing

these bioproduct come from plant extraction or from oil and gas derivatives which is cheaper to produce than via engineered microbial strains. There are many ways we can try and reduce the cost of production; we decided to focus on the carbon source used during fermentation. As of current, the preferred feedstock for *S. cerevisiae* is glucose, however it is an expensive commodity due to its high demand [2]. As such, it has been reported that nearly 60% of production cost comes from the use of glucose feedstock [2]. Thus, being able to use a more cost-effective feedstock, could help bring these bioproducts to market. Close relatives to *S. cerevisiae*, such as *P. pastoris* or *Hansenula polymorpha*, can ferment on methanol, a cheaper carbon source than glucose. As well, methanol can be produced through the process of gasification [12]. In this study, we sought to engineer a strain of *S. cerevisiae* that would be capable of converting methanol to biomass. We used *S. cerevisiae* as the model organism over *P. pastoris* because, historically *S. cerevisiae* has been more engineered to produce high-valued bioproducts. Therefore, to help bring these bioproducts to market, we engineered a strain of *S. cerevisiae* for assimilation of methanol, a cheaper carbon source, through the addition of three genes: NAD(P)-dependent MDH, HPS, and PHI.

Before moving into the genetic manipulation, we first developed a *S. cerevisiae* metabolic model in order to better understand the flux of this carbon source and in order to elucidate any cofactor imbalance caused by the addition of these genes.

5.1 *In silico* prediction of growth rate and ATP production

Even the simplest model organisms tend to be complex, filled with thousands of reactions found in different cellular compartments. We therefore delegate these convoluted reaction networks to computational models, which can account for different variables at once and output a single number. Before designing our yeast strain, we decided to test our design using a *S. cerevisiae* metabolic model. The ATP production as well as the specific growth were compared *in silico* between the wild-type and the engineered strain under glucose and methanol carbon sources.

5.1.1 Growth rate comparison

Looking at the iAV3000 predicted growth under glucose-limited chemostat, we can conclude that the percent error at different dilution rates differ between 8 and 32 percent. Although a 32 percent error seems high, we compared our values to other published metabolic models (Table 4-3) and it

suggests that the growth rate predicted by our model lies on the lower portion of all percent errors (Table 4-3).

Prediction under methanol carbon source shows higher biomass production than under glucose carbon source (Figure 4-1). Possibly, this is because of the increase of available NADH in the cell. The model utilizes this cofactor to drive ATP production and this available energy is then used to increase biomass. Looking at the flux balance analysis of the iAV3000 model under methanol carbon source, the NDE1 and NDE2 reactions have increased fluxes. These two enzymes are part of the ETS and oxidize NADH by reducing one mole of quinone [81]. Reoxidation of NADH is primarily done through the NDE1/NDE2 reactions, however in times of starvation, such as providing only methanol as a carbon source to *S. cerevisiae*, the main route for NADH reoxidation is the glycerol-3-phosphate shuttle [82][83]. From the model we can observe some increase flux through the glycerol pathway, however this flux is probably underestimated by the model since the NDE1/NDE2 reaction is highly active.

Nevertheless, we have a good estimate of the reliability of the engineered pathway. Since, the engineered pathway creates an excess NADH supply, we wanted to look at the prediction of ATP production.

5.1.2 ATP production comparison

Using the metabolic model, we looked at the ability to generate ATP under glucose-limited chemostat and methanol-limited chemostat. Previously published experimental results of ATP production in *S. cerevisiae* suggests 12.5 mole ATP per mole of glucose [48]. Our model predicts the production of 11.5 moles of ATP per mole of glucose, which results in a 4.17% error. Under methanol-limited chemostat the model predicts 26.5 moles of ATP per 6 moles of methanol.

Looking at the flux results, we see that the increase in available NADH allows the model to generate more ATP. In reality, as aforementioned, NAD and NADH equilibrium is regulated by the glycerol pathway, and for this reason we believe it is overestimating the ATP production. The lack of published research on methylotrophic yeast ATP production under methanol carbon source inhibits us from quantifying the output of our model.

5.2 NAD(P)-dependent methanol dehydrogenase activity

The NAD(P)-dependent MDHs were introduced at one to three different locations of the genome (Table 3-15), chosen from previously studied integration sites [74]. This allowed us to test the effects of increase copy numbers – 1 copy, 2 copies, and 3 copies. Results of this experiment provides evidence of the activity of the six NAD(P)-dependent MDH variants (Figure 4-2). Additionally, we observed an increase in activity as the enzyme copy number increases (Figure 4-2). We conclude that activity of MM1 is superior to the other NAD(P)-dependent MDH variants (Figure 4-2).

When comparing the kinetics of the six NAD(P)-dependent MDHs (Table 2-5), we see that the K_M constant, which represents the concentration of the substrate at which 50% enzyme activity is achieved, is lowest for both MM1 and MP2 (MDH^M and $MDH1^P$ respectively). Therefore, previously publish data is in agreement with our results. However, it does not explain the low activity of the MDHs that came from *B. methanolicus* PB1, in comparison to their kinetics (Table 2-5). We can suspect that perhaps improper folding of the proteins resulted in their poor activities, fluorescently tagging these proteins could help explain these observations. Our reported *B. methanolicus* MGA3 MDH activities lie between the values obtained by J.E.N Müller et al. for his *in vitro* activity assay with and without the ACT protein (Table 5-1). While the activities of the enzymes from *B. methanolicus* PB1 lie below the values obtained by J.E.N Müller et al. (Table 5-1). It must be stated that the values J.E.N Müller et al. obtained from ^{13}C-labeling experiments do not correlate with his *in vitro* activity assays. This makes it difficult to benchmark the activity of the NAD(P)-dependent MDHs when experimental results from the same paper do not align.

The ^{13}C-labeling experiments done by J.E.N Müller et al., showed that in the absence of ACT protein, MP3 lead to the most assimilated methanol [17]. Our study suggests that MP3 has amongst the lowest activity in comparison to the six different MDHs. However, *in vitro* activity studies by J.E.N Müller et al. also suggested low MP3 activity in comparison to the other MDHs and is similar to the reported activity in this study (Table 5-1) [17]. When performing *in vitro* experiments, only a very specific reaction is being monitored, while when performing an *in vivo* study the entire metabolism is taken into account.

Table 5-1: *In vitro* NAD(P)-dependent MDH activities on methanol substrate

Enzyme	1 copy (mU/mg)	2 copies (mU/mg)	3 copies (mU/mg)	Ref
MM1	12.31	14.61	18.67	This study
	1.4 and 15.5*	-	-	[17]
MM2	7.06	11.22	11.32	This study
	2.3 and 45	-	-	[17]
MM3	13.01	13.14	12.53	This study
	1.5 and 27.3	-	-	[17]
MP1	2.12	2.36	3.01	This study
	2.6 and n.a	-	-	[17]
MP2	1.66	2.01	1.73	This study
	10 and n.a	-	-	[17]
MP3	1.65	1.88	2.72	This study
	3.3 and n.a	-	-	[17]
	7.2	-	-	[26]
MDH^S	19.8	-	-	[26]

*The second value is the MDH activity in the presence of ACT protein.

Overall lower *in vitro* enzyme activities were obtained in this study in comparison to those obtained by J.E.N Müller et al (Table 5-1). Temperature variation of the enzymatic assay has been shown to affect the overall activities of the enzymes[18]. The *in vitro* activity assay done by J.E.N Müller et al. was performed at 37ºC while the temperature of the *in vitro* activity assay performed in this study was at 30ºC. The differences in temperature could also explain the enzyme activity differences between the two studies but for this study, testing the enzymes' activities at 30ºC is a better prediction of their activities since *S. cerevisiae* grows at this temperature.

The activity of the native *Methylomonas* MDH has been reported to be 28±5 mU/mg [84]. Also, the *B. stearothermophilus* MDH, which was used by Whitaker et al. for engineering *E. coli*, was reported to have an activity of 19.8 mU/mg [26]. With this activity Whitaker et al. were able to provide evidence of methanol metabolism towards biomass, which J.E.N Müller et al were unable to show. The chosen MDH in this study is MM1 and we report an activity of 18.67 mU/mg when three copies are integrated in the strain. This enzyme activity was later shown to incorporate methanol and increase biomass titers in our engineered strains.

The addition of the ACT proteins did not result in higher NAD(P)-dependent MDH activity, (Figure 4-3). This questions the true purpose of the ACT protein since the original paper by Krog et al. suggests higher *in vitro* activity of the *B. methanolicus* MDHs in the presence of the ACT protein [18]. However, due to the lack of genetic engineering tools the ACT protein has yet to be deleted from *B. methanolicus*, which could provide evidence of its function. Additionally, subsequent studies have shown increase *in vitro* MDH activity in the presence of the ACT protein in *E. coli,* but show reduction of the *in vivo* MDH in the presence of ACT protein [17].

As stated earlier, the ACT protein is of the Nudix hydrolase family, which is ubiquitously found across all kingdoms [18]. Perhaps orthologs of the ACT protein in *E. coli* and *S. cerevisiae* can function as the ACT protein. However, this does not explain the contradicting effects of the ACT protein in an *in vitro* and *in vivo* setting. *In vivo* studies do represent a more natural environment versus an *in vitro* study and for this reason, it is more acceptable to think that ACT protein has no positive effect on the activity of the MDHs. As such, we did not pursue with the addition of the ACT protein in our engineered strains.

5.3 Effects of HPS and PHI variants on biomass production

To complete the RuMP cycle, both HPS and PHI genes must be introduced into our engineered strain containing three copies of MM1. Both genes, HPS and PHI, were introduced in tandem at three different locations of the genome (Table 3-15). The functional activity of the two enzymes were indirectly tested by monitoring growth under mixed carbon sources, glucose and methanol. Previous studies evaluated the toxicity of methanol on *S. cerevisiae*, and reported that cells were inviable above 1.23M methanol concentration [80]. For this reason, we decided to test growth under four different concentrations of methanol – 0mM, 5mM, 100mM, and 1M – while keeping the glucose concentration at 10g/L.

At 5mM methanol concentration, the wild-type background strains provided evidence of increased biomass (Figure 4-4). However, since there is no way for methanol carbon to be used for biomass synthesis the observation could in fact be attributed to the promiscuous activity of the alcohol dehydrogenases on methanol. This would convert methanol to formaldehyde, which would then enter the formaldehyde oxidation II pathway and generate NADH molecules. The NADH cofactor

can then be used to produce ATPs used to drive biomass synthesis (Figure 4-4) [76]. Since we have shown promiscuous methanol oxidation in wild-type (Section 4.2), perhaps enough formaldehyde is pushed through the formaldehyde oxidation II pathway, which creates an additional energy source that can used for biomass synthesis [76]. We also observed a major cell fitness reduction at 1M methanol concentration, which is in accordance with previously published results [80].

In comparison to the wild-type strain, MM1 strain has higher methanol toxicity at both 5mM and 100mM methanol concentration (Figure 4-4 and Figure 4-5). The overall observation of this study suggest activity of all three enzymes at 5mM and 100mM methanol concentrations. Therefore, I believe that at lower methanol concentrations MM1 is properly functioning, that is methanol conversion to formaldehyde is happening. However, at higher methanol concentrations the engineered RuMP cycle may have funneled some methanol down glycolysis before eventually being hindered by methanol toxicity. Nevertheless, methanol concentration would have been reduced, which could explain why we observe greater reduction in wild-type cell fitness.

The reduce fitness of the MM1 background strain is rescued through the addition of HPS and PHI genes (Figure 4-5). This suggests a two-part conclusion, first, that both HPS and PHI enzymes are active in the cell and secondly that the cooperation of all three enzymes – MM1, HPS, and PHI – results in an active RuMP cycle. The synergy of the four different combinations of HPS and PHI genes were tested. From our results, we observe high activity for the HPS.P and PHI.P combination (Figure 4-4 and Figure 4-5).

The presence of formaldehyde in *S. cerevisiae* will induce the transcription of SFA1 gene by 8.3-folds after two hours [80]. Previous studies where they engineered the RuMP cycle in *E. coli* showed a reduction of fitness in strains where *frmA* was deleted [17]. Formaldehyde dehydrogenase, *frmA* in *E. coli* and SFA1 in *S. cerevisiae*, is involved in other pathways such as amino acid degradation [37]. Regardless, the presence of SFA1 robs the engineered RuMP cycle from utilizing carbon that could be funnel towards biomass production. Therefore, we tested the activity of our pathway in strains lacking SFA1. Comparison of the wild-type with and without SFA1 show greater cell fitness reduction from methanol in the absence of SFA1 (Figure 4-4 and Figure 4-6). As well, comparison of MM1 with and without SFA1 also show a greater cell fitness

reduction in the absence of SFA1 (Figure 4-5 and Figure 4-7). In contrast, we observe greater growth in the absence of SFA1 when all three enzymes of the engineered RuMP cycle are active. This suggests that SFA1 does indeed rob the cell of potential methanol carbon and that SFA1 deleted strains as a result end with a higher cell density (Appendix 7.1.7).

The strain harbouring three copies of MM1 with three copies of HPS.P and PHI.P with the SFA1 gene deleted was able to assimilate the most methanol for biomass synthesis. A final increase of 13.54% was observed after 33 hours (Figure 4-7 and Section 7.1.7). The engineered *E. coli* built by Whitaker et al. was able to gain 30% in biomass after 72 hours only in the presence of added yeast extract. Additionally, Dai et al. whom engineered a *S. cerevisiae* strain for methanol assimilation had a maximum increase in biomass of 11.70% only in the presence of added yeast extract after 72 hours. We did not add yeast extract to our media, therefore media optimization by the addition of yeast extract may increase the assimilation of methanol.

Whitaker et al. also provides evidence of methanol assimilation toward biomass synthesis starting after 8 hours. They conclude that their engineered *E. coli* harvested the available yeast extract nutrients in the first 8 hours before assimilating methanol. However, our growth experiment shows faster entry into the exponential phase and a higher final cell density for our strain harbouring all three genes (Section 7.1.7). The growth experiments were performed in the presence of 10g/L glucose, which we observed was being simultaneously metabolized with methanol. As for Dai et al., they only observe methanol assimilation after the 6-hour mark. Additionally, Whitaker et al. report methanol assimilation toward biomass to occur between the 8 and 72 hours, suggesting that *E. coli* used the available nutrients in the yeast extract for the first 8 hours before utilizing methanol.

5.4 Tracking methanol assimilation for biomass synthesis

In order to obtain direct evidence of an active RuMP cycle in our engineered strains we decided to feed ^{13}C-labeled methanol and track its incorporation in amino acids. Previous research by J.E.N Müller et al. showed a reduction in assimilated methanol in the absence of *frmA*, the equivalence of SFA1 [17]. They concluded that *frmA* is not only needed for methanol detoxification but also for other cellular processes [17]. In our engineered yeast strains, we show that in the absence of

SFA1 an increase in labeled amino acids is observed. This suggests that in the presence of SFA1, some methanol is funneled away from our engineered pathway which results in lower assimilated methanol and ultimately lower biomass production. Perhaps the 1M methanol concentration used by J.E.N Müller et al. during their ^{13}C-labeling experiment reached a critical point where the toxicity of formaldehyde outweighed the benefits of assimilating additional methanol. The same experiment should have been performed using different concentrations of ^{13}C-labeled methanol in order to fully understand the results of this experiment.

We also tested the effects of two different medias, YNB and YNB with 1g/L yeast extract, on methanol assimilation. Previous published research by Whitaker et al. showed increase methanol assimilation in the presence of yeast extract. In our strain lacking SFA1 we observed a 2.2 to 3.7-fold increase of labeled amino acids in the presence of 1g/L yeast extract. Therefore, our observation aligns with that of Whitaker et al [26]. Previous studies suggest that yeast extract may help replenish the pool of threonine amino acid, which has been shown to help with methanol assimilation [85]

6 CONCLUSION AND FUTURE DIRECTIONS

Overall, we showed the individual activities of the different NAD(P)-dependent methanol dehydrogenases. We then showed the inability to increase their activities in the presence of ACT protein. We showed that with the addition of HPS and PHI, higher growth in the presence of methanol could be achieved. Finally, we showed that methanol is being assimilated and used to increase biomass synthesis.

The central hypothesis of this thesis is that by engineering the genome of *Saccharomyces cerevisiae* through the addition of three genes – NAD(P)-dependent MDH, HPS, and PHI – it would create a new functional pathway, the RuMP cycle, and would allow *S. cerevisiae* to metabolize methanol and use this carbon for biomass production.

In this study, I integrated six variants of NAD(P)-dependent MDHs in different copy numbers and showed their individual *in vitro* activities. From this experiment I found that MM1 outperformed the other MDHs. These results agree with the kinetics that had been previously published.

However, it did not align with J.E.N Müller et al. experimental results. I explained this discrepancy by acknowledging the differences in the hosts, one being *E. coli* and the other *S. cerevisiae*, and the temperature difference of each *in vitro* assay. Later work by Whitaker et al., tested the activity of an NAD(P)-dependent MDH from *B. stearothermophilus* strain, and was able to show higher activity than with $MDH2^P$ [26]. One of the major drawbacks of these NAD(P)-dependent MDHs are their low affinities for methanol. There are two ways we can try and alleviate this negative effect, one is by increasing the copy number of these enzymes, which was done in this study, and secondly, is to find a better MDH. As such, it would be interesting to test the MDH from *B. stearothermophilus* in our system and see if methanol conversion increases or use this newly identified NAD(P)-dependent MDH to find new homologs with similar activity.

We also tested the *in vitro* effects of two ACT proteins on the various MDHs and found no positive effects. This area of study has been debated, as aforementioned, previous published data showed positive *in vitro* effect of this cofactor but failed to show any *in vivo* activity. For this reason, it is hard to conclude the necessity of this cofactor. Due to the lack of supporting evidence from this study I decided to leave out the ACT protein from our engineered strain.

The last two enzymes, HPS and PHI, were integrated into the genome of *S. cerevisiae* in three copies. This completed the RuMP cycle and theoretically allowed for assimilation of methanol. Activity of both RuMP cycle and of the different HPS and PHI combinations were tested by monitoring the growths of our engineered strains. Different methanol concentrations were tested and we observed methanol toxicity at 1M methanol concentration; suggesting that the RuMP cycle is not active enough to handle such high methanol concentrations. However, at 5mM and 100mM methanol concentrations we observed an increase in cell density and relate this to an increase in biomass synthesis. We must keep in mind the limitation of this observation as this is an indirect assumption. To know if methanol is being utilized for biomass synthesis we had to trace conversion of methanol to a biomass component such as amino acids. This was accomplished using ^{13}C-labled methanol. In this experiment we show that methanol was indeed being used to produce amino acids and ultimately biomass (Figure 4-8). We also observed differences in the growth of our engineered strains harboring different combinations of HPS and PHI genes. There have been no published studies on the kinetics of these two enzymes, and previously published results from both J.E.N Müller et al. and Whitaker et al. did not assessed the combination that would result in higher

RuMP cycle activity. It would be interesting to identify their kinetics and to determine how they compare in a synergetic pathway.

Finally, we show that deleting SFA1 leads to higher methanol assimilation. This is also an area of debate because ^{13}C-labeling experiment done by J.E.N Müller et al. showed lower methanol incorporation in the absence of *frmA*, SFA1 equivalence in *E. coli*. However, Whitaker et al., which based their research on that of J.E.N Müller et al., did not assess the effects of *frmA* in *E. coli* and only tested for methanol assimilation in an *frmA* deleted background strain. Both SFA1 and *frmA* are also found in amino acid degradation pathways [37]. Whitaker et al. observed highest methanol incorporation when the growth media was supplemented with yeast extract [26]. Thus, perhaps the addition of yeast extract, which contains free amino acids, in the media alleviated the negative effects of a *frmA* deletion. In comparison to the aforementioned research, we tested both MM1 HPS.P PHI.P and MM1 ΔSFA1 HPS.P PHI.P for the ability to convert methanol into amino acids. Through this experiment we were able to show incorporation of labeled carbon into glutamate, aspartate and alanine, suggesting that methanol is used for biomass synthesis. Additionally, we were able to show the effects of SFA1 deletion. We show that in the absence of SFA1, our engineered strain was able to assimilate more methanol (Figure 4-8). This aligns with our previous experiments where we show higher growth in the absence of SFA1 (Section 0). We also show that addition of yeast extract to the media increases the total count of labeled amino acids and therefore suggests higher methanol assimilation rate (Figure 4-8).

Recent study by Dai et al. showed by the incorporation of the XuMP cycle, the methylotrophic yeast methanol assimilation pathway, that they were able to growth their engineered strain solely on methanol, and observed a 11.70% increase in growth in the presence of yeast extract [38]. This was done by expressing – AOX, CAT, DAS1 or DAS2, and DAK – proteins that were targeted to the peroxisome. The reason I did not proceed to engineer *S. cerevisiae* with this cycle was because of studies on methylotrophic yeast, which shows that methanol induces peroxisome proliferation to such an extent that it can take up to 80% of the total cytoplasmic space. This suggests that for *S. cerevisiae* to solely grow on methanol the XuMP cycle must be highly active and compartmentalized. I believe that compartmentalization needs to occur because the first step of the pathway is achieved by a methanol oxidase, which generates one molecule of reactive oxygen species (ROS) for every molecule of methanol conversion to formaldehyde: compartmentalizing

these ROS mitigates their negative effects. To achieve such peroxisome proliferation many accessory proteins need to be activated, which would be overly complicated to orchestrate as well as being energy intensive. Thus, I do not see this pathway being able to scale up to the industrial scale like the engineering of the RuMP cycle can. However, both RuMP and XuMP cycles are highly similar.

This brings forth a new engineered pathway that could outperform both pathways; the integration of an NAD(P)-dependent MDH with a DAS2 enzyme would allow methanol to be assimilated in the cytosol and would remove the dependency on Ru5P metabolites. Additionally, DAS2 is from a methylotrophic yeast, a closer relative to *S. cerevisiae* than *B. methanolicus*, which is a thermophilic bacterium. Using enzymes from distant thermophiles and incorporating them into a host which thrives at room temperature reduces the activity of certain enzymes [18]. The addition of DAS2 instead of HPS and PHI would eliminate the need of a third step and as a result could lead to a higher pathway activity. It is also worth noting that a lot of research around engineering *S. cerevisiae* for the assimilation of xyloses as a carbon source has come out in recent years [86]. Assimilation of xylose results in increase Xu5P concentration, which is the recycled molecule of the XuMP cycle [86]. Thus, a synergetic advantage might arise from combining these two carbon sources.

A major XuMP cycle advantage is the compartmentalization of formaldehyde. Due to formaldehyde toxicity, keeping this molecule in an organelle confines its cross-linking effects and allows for a healthier host. Previously published research on building synthetic organelles provided an opportunity to confine this pathway [87]. Although this area of study is still in its neophyte stages, it is something to keep in mind for future improvements of the engineered methanol assimilation pathway.

Lastly, a new pathway known as the Methanol Condensation Cycle (MCC), suggest 100% carbon efficiency (Figure 2-5) [36]. The MCC pathway has yet to be engineered in a host, but preliminarily studies provide evidence of a successful methanol assimilation pathway [36]. Thus, attempting to engineer the MCC in yeast could result in higher methanol assimilation.

Overall, we were able to engineer *S. cerevisiae* to assimilate methanol and observed a 13.54% increase in biomass synthesis in the presence of 5mM methanol and 1% glucose. We show the

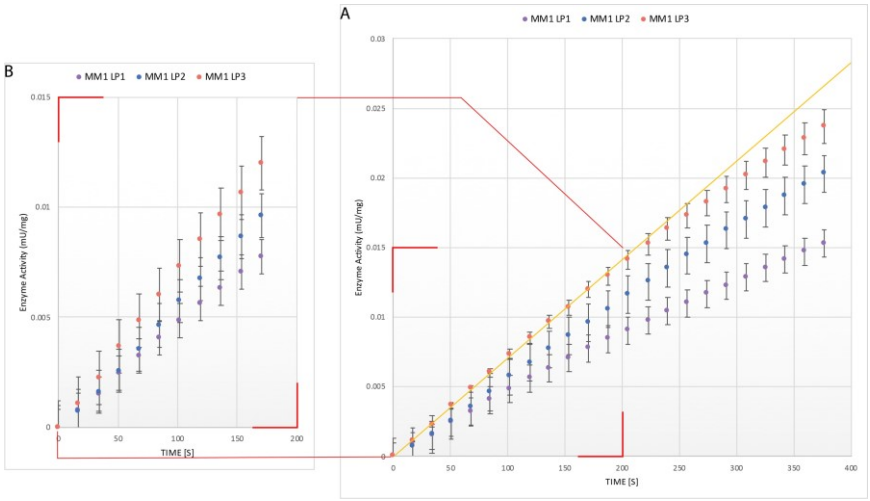
ability of our engineered strains to convert methanol to amino acids using ^{13}C-labeled methanol. We also show a 2.2 to 3.7-fold increase in ^{13}C-labeled amino acids when the media was supplemented with yeast extract. Perhaps even better media optimization is possible by the addition of yeast extract and threonine as it has been shown to help with methanol assimilation for biomass synthesis [85].

7 APPENDIX

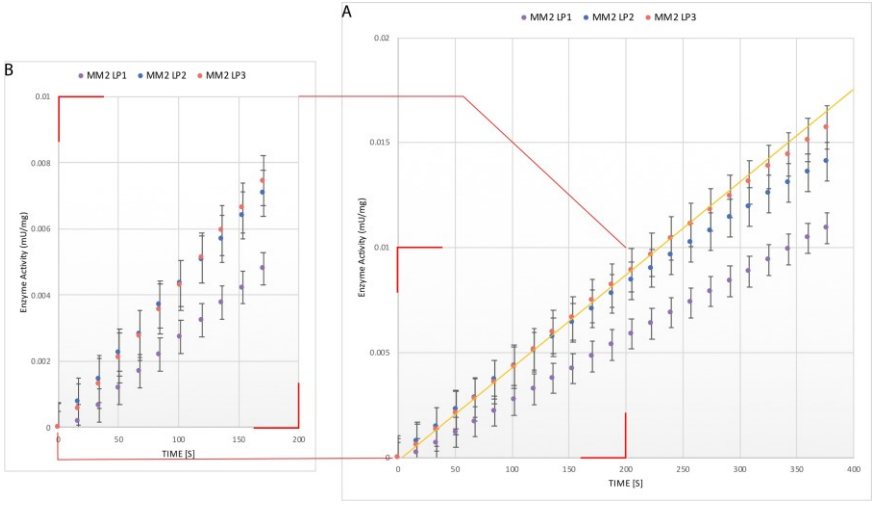
The appendix contains the raw data of some of the experiments as well as some different representations of the data.

7.1 *In vitro* NAD(P)-dependent methanol dehydrogenase activity data

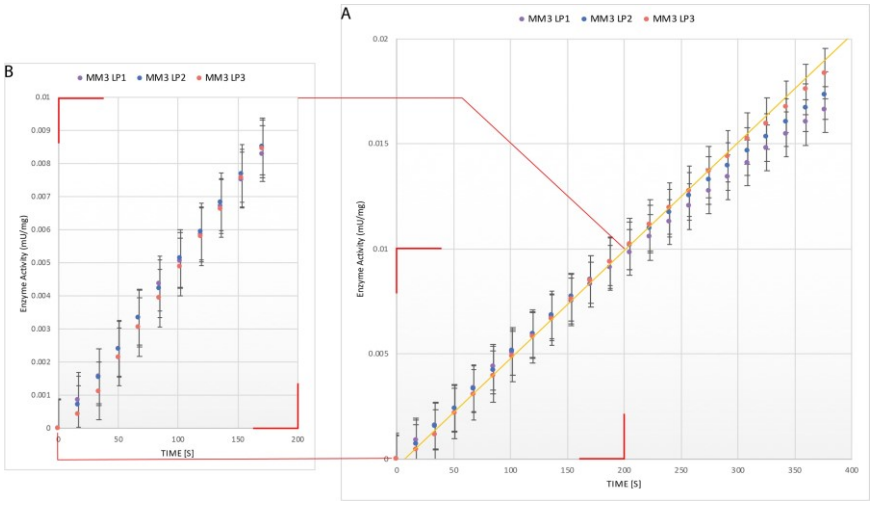
7.1.1 MM1



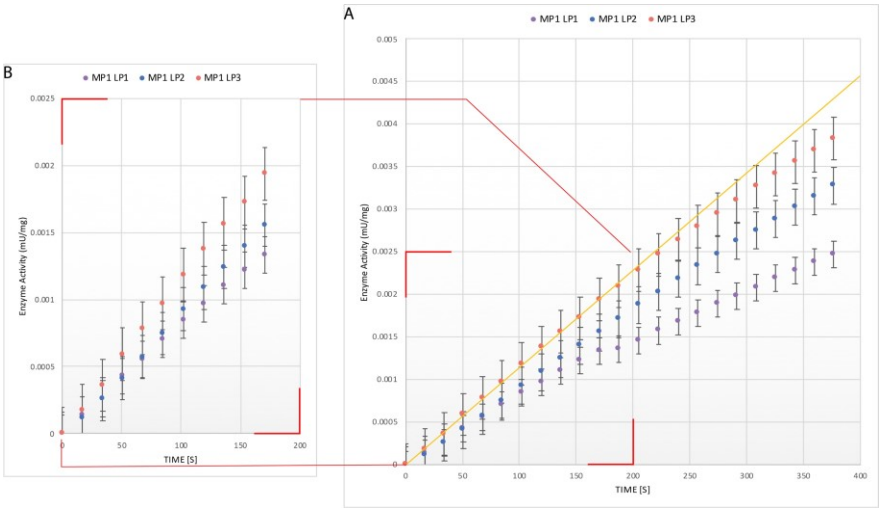
7.1.2 MM2



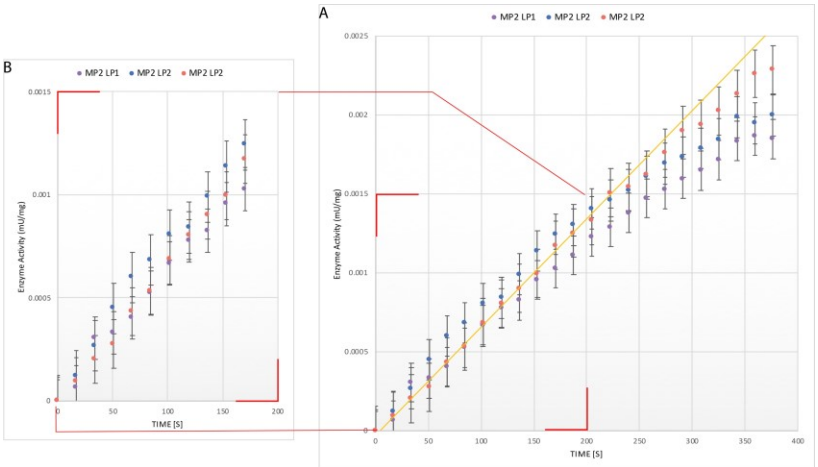
7.1.3 MM3



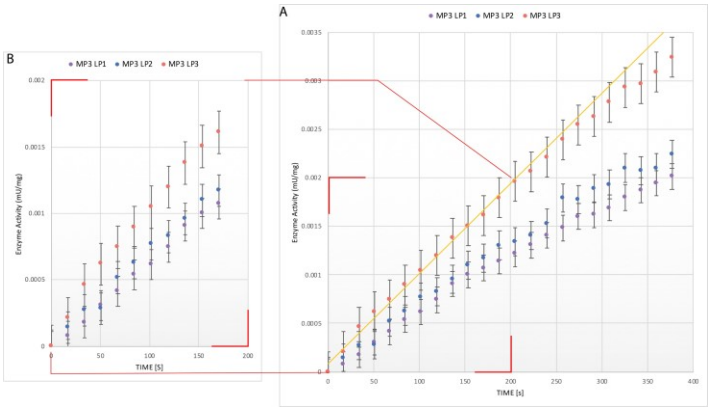
7.1.4 MP1



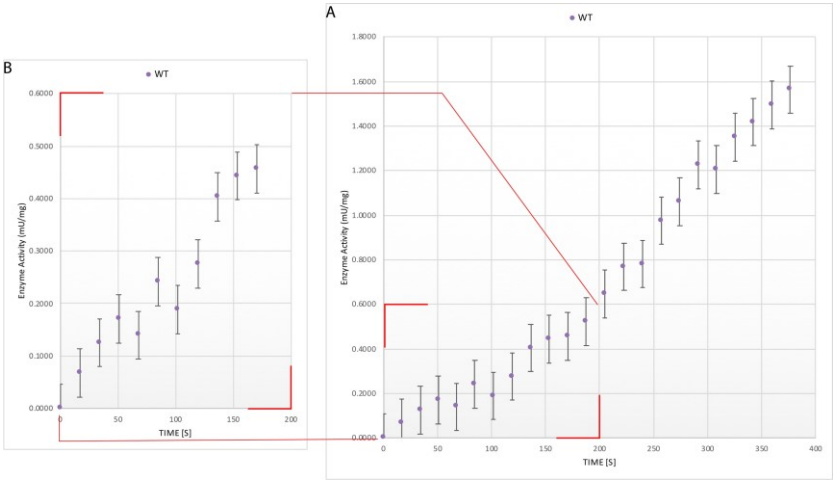
7.1.5 MP2



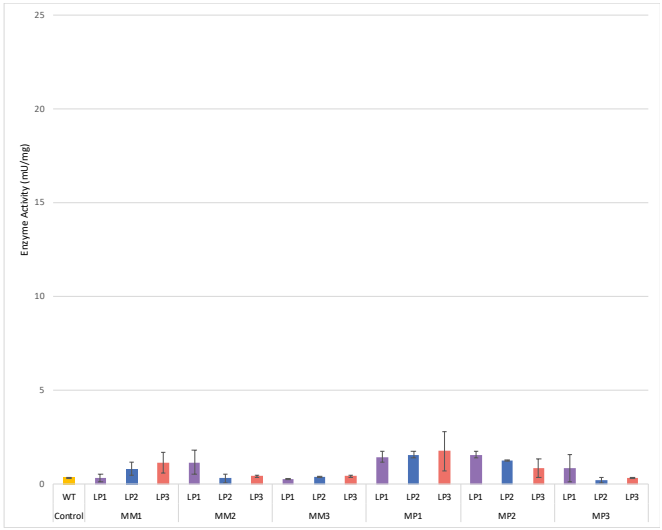
7.1.6 MP3



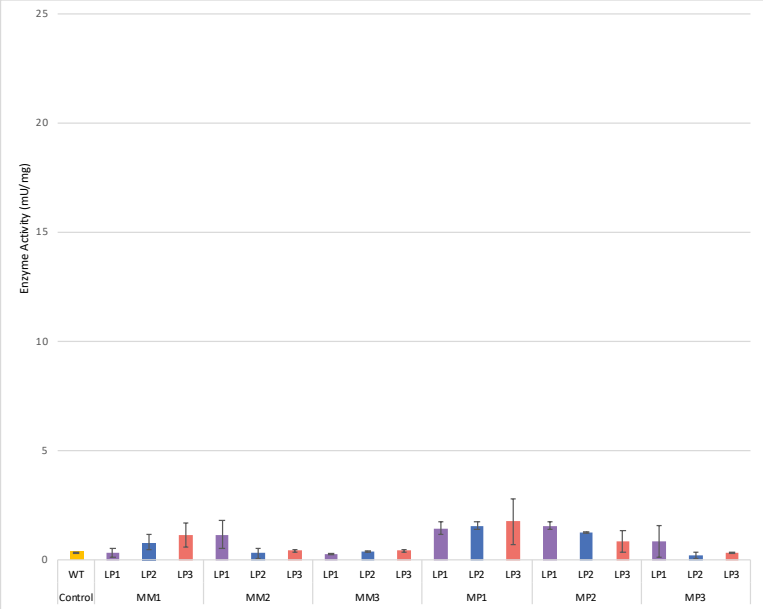
7.1.7 WT



7.1.8 Controls without ACT

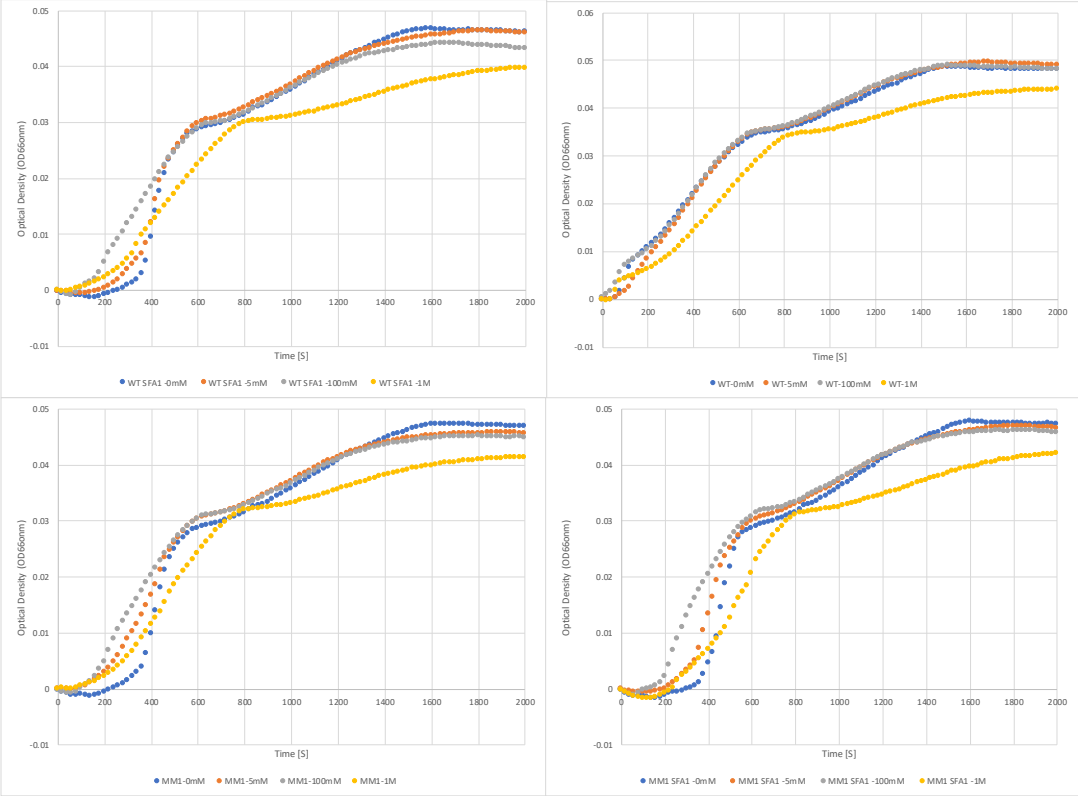


7.1.9 Controls with ACT

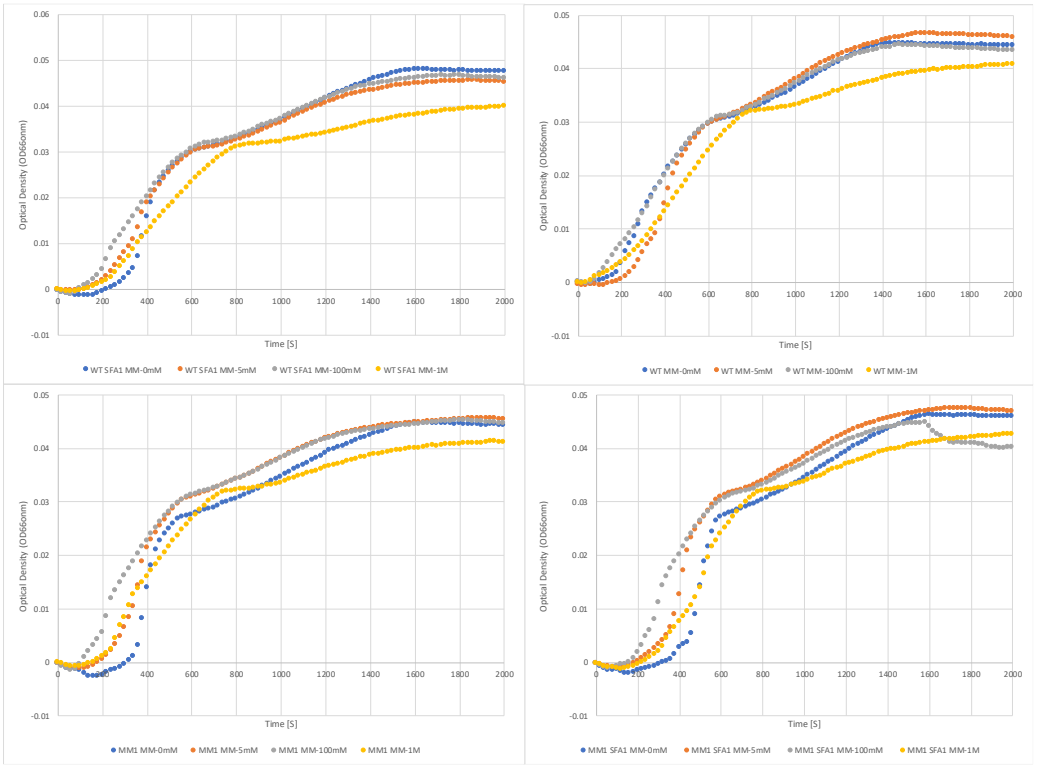


7.2 *In vivo* growth curves

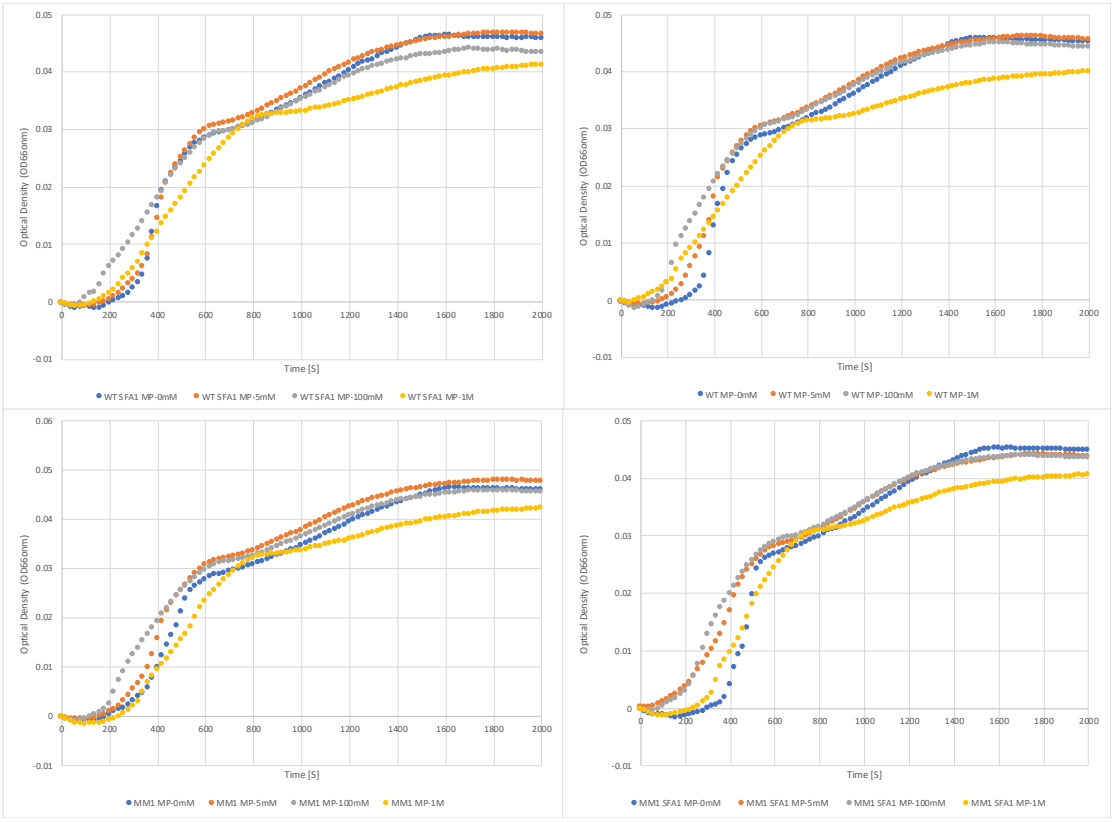
7.2.1 Controls data – no HPS or PHI genes



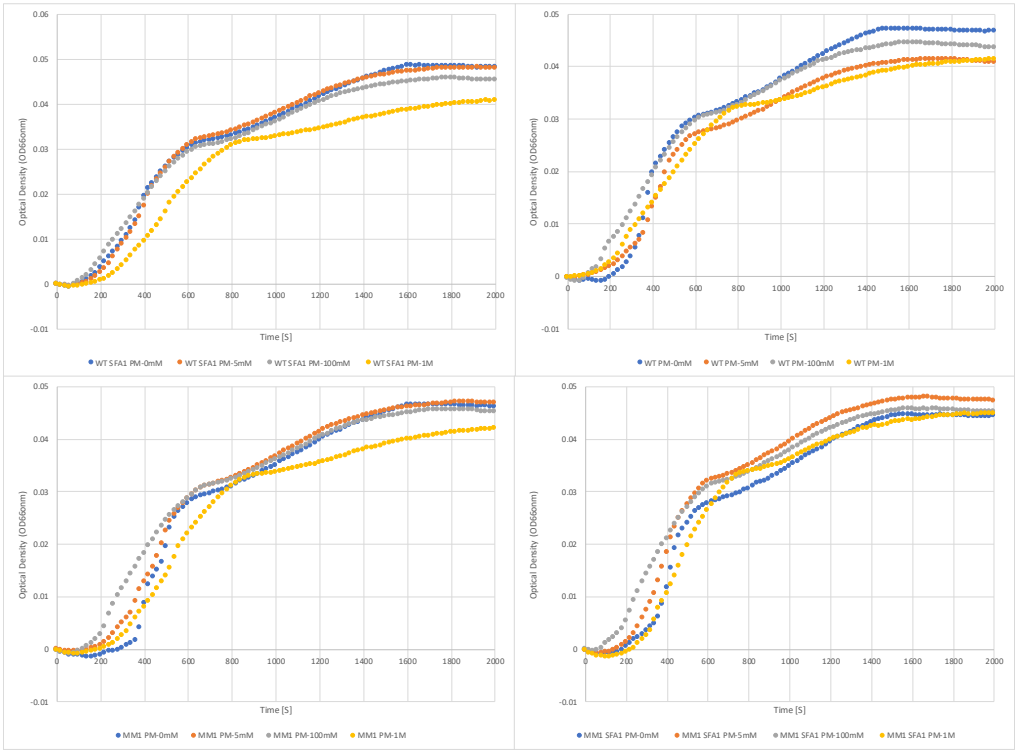
7.2.2 HPS.M and PHI.M



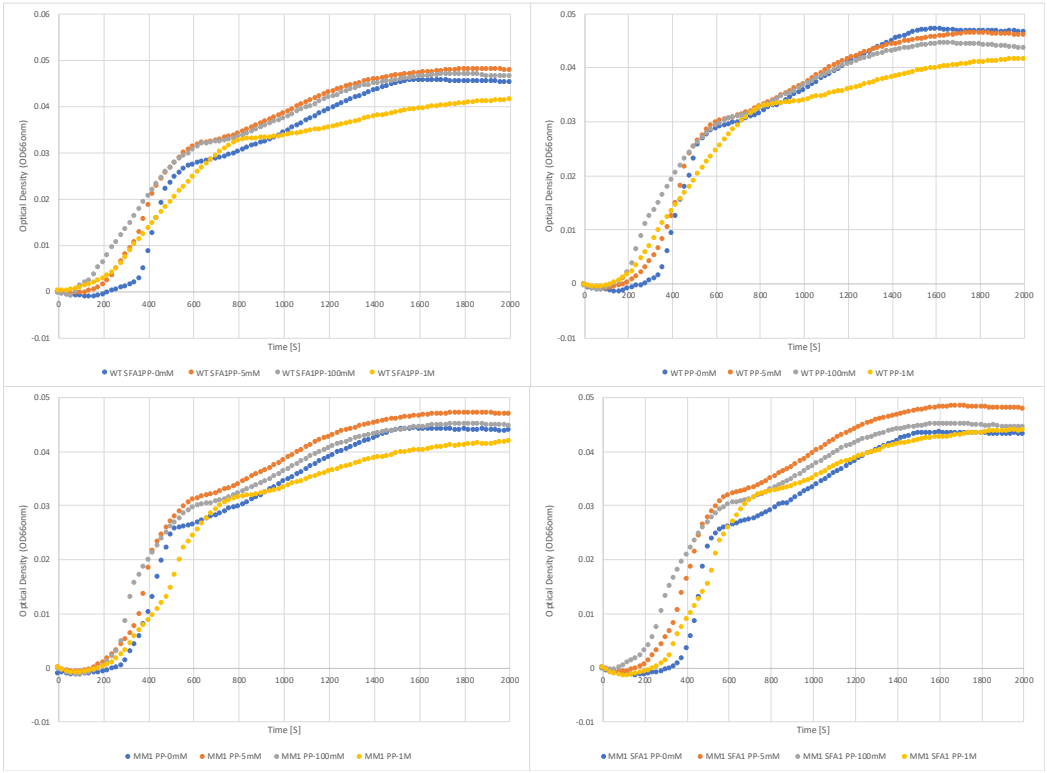
7.2.3 HPS.M and PHI.P



7.2.4 HPS.P and PHI.M



7.2.5 HPS.P and PHI.P



8 BIBLIOGRAPHY

[1] K. Ghasemzadeh, S. M. Sadati Tilebon, M. Nasirinezhad, and A. Basile, "Economic Assessment of Methanol Production," in *Methanol*, Elsevier, 2018, pp. 613–632.

[2] Q. Fei, H. N. Chang, L. Shang, and J. D. R. Choi, "Exploring low-cost carbon sources for microbial lipids production by fed-batch cultivation of Cryptococcus albidus," *Biotechnol. Bioprocess Eng.*, vol. 16, no. 3, pp. 482–487, 2011.

[3] J. Daniell, M. Köpke, and S. Simpson, "Commercial Biomass Syngas Fermentation," *Energies*, vol. 5, no. 12, pp. 5372–5417, Dec. 2012.

[4] J. L. D. C. P. Brandão, "Summary for Policymakers," in *Climate Change 2013 - The Physical Science Basis*, Intergovernmental Panel on Climate Change, Ed. Cambridge: Cambridge University Press, 2011, pp. 1–30.

[5] L. Yang, X. Ge, C. Wan, F. Yu, and Y. Li, "Progress and perspectives in converting biogas to transportation fuels," *Renew. Sustain. Energy Rev.*, vol. 40, pp. 1133–1152, 2014.

[6] T. Amon, B. Amon, V. Kryvoruchko, A. Machmüller, K. Hopfner-Sixt, V. Bodiroza, R. Hrbek, J. Friedel, E. Pötsch, H. Wagentristl, M. Schreiner, and W. Zollitsch, "Methane production through anaerobic digestion of various energy crops grown in sustainable crop rotations," *Bioresour. Technol.*, vol. 98, no. 17, pp. 3204–3212, 2007.

[7] P. Havlík, U. A. Schneider, E. Schmid, H. Böttcher, S. Fritz, R. Skalský, K. Aoki, S. De Cara, G. Kindermann, F. Kraxner, S. Leduc, I. McCallum, A. Mosnier, T. Sauer, and M. Obersteiner, "Global land-use implications of first and second generation biofuel targets," *Energy Policy*, vol. 39, no. 10, pp. 5690–5702, 2011.

[8] H. Boerrigter and a Van Der Drift, "'Biosyngas' Key Intermediate in Production of Renewable Transportation Fuels, Chemicals, and Electricity: Optimum Scale and Economic Prospects of Fischer-Tropsch Plants.," *Proc. 14th Eur. Biomass*, no. October, pp. 17–21, 2005.

[9] N. A. Samiran, M. N. M. Jaafar, J. H. Ng, S. S. Lam, and C. T. Chong, "Progress in biomass gasification technique - With focus on Malaysian palm biomass for syngas production," *Renew. Sustain. Energy Rev.*, vol. 62, pp. 1047–1062, 2016.

[10] C. J. Paddon and J. D. Keasling, "Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development," *Nat. Rev. Microbiol.*, vol. 12, no. 5, pp. 355–367, 2014.

[11] K. A. Curran, J. M. Leavitt, A. S. Karim, and H. S. Alper, "Metabolic engineering of muconic acid production in Saccharomyces cerevisiae," *Metab. Eng.*, vol. 15, pp. 55–66, Jan. 2013.

[12] M. Guo, W. Song, and J. Buhain, "Bioenergy and biofuels: History, status, and perspective," *Renew. Sustain. Energy Rev.*, vol. 42, pp. 712–725, 2015.

[13] R. De Maria, I. Diaz, M. Rodriguez, and A. Saiz, "Industrial methanol from syngas: Kinetic study and process simulation," *Int. J. Chem. React. Eng.*, vol. 11, no. 1, pp. 469–477, 2013.

[14] P. McKendry, "Energy production from biomass (part 3): Gasification technologies," *Bioresour. Technol.*, vol. 83, no. 1, pp. 55–63, 2002.

[15] M. Puig-Arnavat, J. C. Bruno, and A. Coronas, "Review and analysis of biomass gasification models," *Renew. Sustain. Energy Rev.*, vol. 14, no. 9, pp. 2841–2851, 2010.

[16] H. Boerrigter and R. Rauch, "Review of applications of gases from biomass gasification," in *ECN Biomass, Coal and Environmental …*, no. June, 2006, p. 33.

[17] J. E. N. Müller, F. Meyer, B. Litsanov, P. Kiefer, E. Potthoff, S. Heux, W. J. Quax, V. F. Wendisch, and T. Brautaset, "Engineering Escherichia coli for methanol conversion," *Metab. Eng.*, vol. 28, pp. 190–201, 2015.
[18] A. Krog, T. M. B. Heggeset, J. E. N. Müller, C. E. Kupper, O. Schneider, J. a. Vorholt, T. E. Ellingsen, and T. Brautaset, "Methylotrophic Bacillus methanolicus Encodes Two Chromosomal and One Plasmid Born NAD+ Dependent Methanol Dehydrogenase Paralogs with Different Catalytic and Biochemical Properties," *PLoS One*, vol. 8, no. 3, 2013.
[19] P. Goswami, S. S. R. Chinnadayyala, M. Chakraborty, A. K. Kumar, and A. Kakoti, "An overview on alcohol oxidases and their potential applications," *Appl. Microbiol. Biotechnol.*, vol. 97, no. 10, pp. 4259–4275, 2013.
[20] M. A. Gleeson and P. E. Sudbery, "The methylotrophic yeasts," *Yeast*, vol. 4, no. 1, pp. 1–15, Mar. 1988.
[21] F. Chu and M. E. Lidstrom, "XoxF acts as the predominant methanol dehydrogenase in the type I methanotroph Methylomicrobium buryatense," *J. Bacteriol.*, vol. 198, no. 8, pp. 1317–1325, 2016.
[22] J. T. Keltjens, A. Pol, J. Reimann, and H. J. M. Op Den Camp, "PQQ-dependent methanol dehydrogenases: Rare-earth elements make a difference," *Appl. Microbiol. Biotechnol.*, vol. 98, no. 14, pp. 6163–6183, 2014.
[23] A. M. Ochsner, J. E. N. Müller, C. A. Mora, and J. A. Vorholt, "In vitro activation of NAD-dependent alcohol dehydrogenases by Nudix hydrolases is more widespread than assumed," *FEBS Lett.*, vol. 588, no. 17, pp. 2993–2999, 2014.
[24] T. Brautaset, ??yvind M. Jakobsen, M. C. Flickinger, S. Valla, and T. E. Ellingsen, "Plasmid-Dependent Methylotrophy in Thermotolerant Bacillus methanolicus," *J. Bacteriol.*, vol. 186, no. 5, pp. 1229–1238, 2004.
[25] M. C. Sheehan, C. J. Bailey, B. C. Dowds, and D. J. McConnell, "A new alcohol dehydrogenase, reactive towards methanol, from Bacillus stearothermophilus.," *Biochem. J.*, vol. 252, no. 3, pp. 661–666, 1988.
[26] W. B. Whitaker, J. A. Jones, R. K. Bennett, J. E. Gonzalez, V. R. Vernacchio, S. M. Collins, M. A. Palmer, S. Schmidt, M. R. Antoniewicz, M. A. Koffas, and E. T. Papoutsakis, "Engineering the biological conversion of methanol to specialty chemicals in Escherichia coli," *Metab. Eng.*, vol. 39, no. November 2016, pp. 49–59, 2017.
[27] S. N. Ruzheinikov, J. Burke, S. Sedelnikova, P. J. Baker, R. Taylor, P. a. Bullough, N. M. Muir, M. G. Gore, and D. W. Rice, "Glycerol Dehydrogenase," *Structure*, vol. 9, no. 9, pp. 789–802, 2001.
[28] N. Arfman, H. J. Hektor, L. V Bystrykh, N. I. Govorukhina, L. Dijkhuizen, and J. Frank, "Properties of an NAD(H)-containing methanol dehydrogenase and its activator protein from Bacillus methanolicus.," *Eur. J. Biochem.*, vol. 244, no. 2, pp. 426–433, 1997.
[29] N. Arfman, J. Van Beeumen, G. E. De Vries, W. Harder, and L. Dijkhuizen, "Purification and characterization of an activator protein for methanol dehydrogenase from thermotolerant Bacillus spp," *J. Biol. Chem.*, vol. 266, no. 6, pp. 3955–3960, 1991.
[30] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman, "Gapped BLAST and PS I-BLAST: a new generation of protein database search programs," *Nucleic Acids Res*, vol. 25, no. 17, pp. 3389–3402, 1997.
[31] R. K. Bajpai, M. Reub, and W. Held, "Regulation Phenomena in Methanol Consuming Yeasts : An Experiment with Model Discrimination," vol. XXIII, pp. 499–521, 1981.
[32] T. J. Lawton and A. C. Rosenzweig, "Methane-Oxidizing Enzymes: An Upstream

Problem in Biological Gas-to-Liquids Conversion," *J. Am. Chem. Soc.*, vol. 138, no. 30, pp. 9327–9340, 2016.
[33] Q. Fei, M. T. Guarnieri, L. Tao, L. M. L. Laurens, N. Dowe, and P. T. Pienkos, "Bioconversion of natural gas to liquid fuel: Opportunities and challenges," *Biotechnol. Adv.*, vol. 32, no. 3, pp. 596–614, 2014.
[34] F. Baneyx, *Protein Expression Technologies: Current Status and Future Trends*. Wymondham, Norfolk, U.K: Horizon Bioscience, 2003.
[35] M. Geier, C. Brandner, G. A. Strohmeier, M. Hall, F. S. Hartner, and A. Glieder, "Engineering Pichia pastoris for improved NADH regeneration: A novel chassis strain for whole-cell catalysis," *Beilstein J. Org. Chem.*, vol. 11, pp. 1741–1748, 2015.
[36] I. W. Bogorad, C.-T. Chen, M. K. Theisen, T.-Y. Wu, A. R. Schlenz, A. T. Lam, and J. C. Liao, "Building carbon–carbon bonds using a biocatalytic methanol condensation cycle," *Proc. Natl. Acad. Sci.*, vol. 111, no. 45, pp. 15928–15933, 2014.
[37] J. R. Dickinson, L. E. J. Salgado, and M. J. E. Hewlins, "The catabolism of amino acids to long chain and complex alcohols in Saccharomyces cerevisiae," *J. Biol. Chem.*, vol. 278, no. 10, pp. 8028–8034, 2003.
[38] Z. Dai, H. Gu, S. Zhang, F. Xin, W. Zhang, W. Dong, J. Ma, H. Jia, and M. Jiang, "Metabolic construction strategies for direct methanol utilization in Saccharomyces cerevisiae," *Bioresour. Technol.*, vol. 245, pp. 1407–1412, 2017.
[39] S. Witthoff, K. Schmitz, S. Niedenführ, K. Nöh, S. Noack, M. Bott, and J. Marienhagen, "Metabolic engineering of Corynebacterium glutamicum for methanol metabolism," *Appl. Environ. Microbiol.*, vol. 81, no. 6, pp. 2215–2225, 2015.
[40] L. Leßmeier, J. Pfeifenschneider, M. Carnicer, S. Heux, J. C. Portais, and V. F. Wendisch, "Production of carbon-13-labeled cadaverine by engineered Corynebacterium glutamicum using carbon-13-labeled methanol as co-substrate," *Appl. Microbiol. Biotechnol.*, vol. 99, no. 23, pp. 10163–10176, 2015.
[41] I. Thiele and B. Ø. Palsson, "A protocol for generating a high-quality genome-scale metabolic reconstruction," *Nat. Protoc.*, vol. 5, no. 1, pp. 93–121, 2010.
[42] S. J. Jol, A. Kümmel, M. Terzer, J. Stelling, and M. Heinemann, "System-level insights into yeast metabolism by thermodynamic analysis of elementary flux modes," *PLoS Comput. Biol.*, vol. 8, no. 3, pp. 1–9, 2012.
[43] S. A. Becker, A. M. Feist, M. L. Mo, G. Hannum, B. Ø. Palsson, and M. J. Herrgard, "Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox.," *Nat. Protoc.*, vol. 2, no. 3, pp. 727–38, 2007.
[44] A. Makhorin, "GNU Linear Programming Kit." Free Software Foundation, Inc, Moscoq, p. 229, 2010.
[45] Gurobi Optimization Inc., "Gurobi Optimizer Reference Manual - Version 6.5." p. 592, 2016.
[46] "{GLPK} ({GNU} Linear Programming Kit)." 2006.
[47] C. Verduyn, a H. Stouthamer, W. a Scheffers, and J. P. van Dijken, "A theoretical evaluation of growth yields of yeasts.," *Antonie Van Leeuwenhoek*, vol. 59, no. 1, pp. 49–63, 1991.
[48] I. Famili, J. Forster, J. Nielsen, and B. O. Palsson, "Saccharomyces cerevisiae phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 23, pp. 13134–9, 2003.
[49] J. Postmus, I. Tuzun, M. Bekker, W. H. Muller, M. J. Teixeira de Mattos, S. Brul, and G. J. Smits, "Dynamic regulation of mitochondrial respiratory chain efficiency in

Saccharomyces cerevisiae," *Microbiology*, vol. 157, no. 12, pp. 3500–3511, 2011.
[50] R. Wulf, A. Kaltstein, and M. Klingenberg, "H+ and Cation Movements Associated With Adp, Atp Transport in Mitochondria," *Eur. J. Biochem.*, vol. 82, no. 2, pp. 585–592, 1978.
[51] J. M. Cherry, E. L. Hong, C. Amundsen, R. Balakrishnan, G. Binkley, E. T. Chan, K. R. Christie, M. C. Costanzo, S. S. Dwight, S. R. Engel, D. G. Fisk, J. E. Hirschman, B. C. Hitz, K. Karra, C. J. Krieger, S. R. Miyasato, R. S. Nash, J. Park, M. S. Skrzypek, M. Simison, S. Weng, and E. D. Wong, "Saccharomyces Genome Database: The genomics resource of budding yeast," *Nucleic Acids Res.*, vol. 40, no. D1, pp. 700–705, 2012.
[52] I. Schomburg, A. Chang, C. Ebeling, M. Gremse, C. Heldt, G. Huhn, and D. Schomburg, "BRENDA, the enzyme database: updates and major new developments.," *Nucleic Acids Res.*, vol. 32, no. Database issue, pp. D431-3, 2004.
[53] M. Kanehisa and S. Goto, "Yeast Biochemical Pathways. KEGG: Kyoto encyclopedia of genes and genomes," *Nucleic Acids Res*, vol. 28, no. 1, pp. 27–30, 2000.
[54] M. L. Mo, B. Ø. Palsson, and M. J. Herrgard, "Connecting extracellular metabolomic measurements to intracellular flux states in yeast.," *BMC Syst. Biol.*, vol. 3, p. 37, 2009.
[55] I. Nookaew, M. C. Jewett, A. Meechai, C. Thammarongtham, K. Laoteng, S. Cheevadhanarak, J. Nielsen, and S. Bhumiratana, "The genome-scale metabolic model iIN800 of Saccharomyces cerevisiae and its validation: a scaffold to query lipid metabolism.," *BMC Syst. Biol.*, vol. 2, no. 1, p. 71, Jan. 2008.
[56] R. Chowdhury, A. Chowdhury, and C. D. Maranas, "Using Gene Essentiality and Synthetic Lethality Information to Correct Yeast and CHO Cell Genome-Scale Models.," *Metabolites*, vol. 5, no. 4, pp. 536–70, Jan. 2015.
[57] M. Riffle, L. Malmstr??m, and T. N. Davis, "The Yeast Resource Center Public Data Repository," *Nucleic Acids Res.*, vol. 33, no. DATABASE ISS., pp. 378–382, 2005.
[58] J. L. Y. Koh, Y. T. Chong, H. Friesen, A. Moses, C. Boone, B. J. Andrews, and J. Moffat, "CYCLoPs: A Comprehensive Database Constructed from Automated Analysis of Protein Abundance and Subcellular Localization Patterns in Saccharomyces cerevisiae.," *G3 (Bethesda).*, vol. 5, no. 6, pp. 1223–32, Apr. 2015.
[59] G. Habeler, K. Natter, G. G. Thallinger, M. E. Crawford, S. D. Kohlwein, and Z. Trajanoski, "YPL.db: the Yeast Protein Localization database.," *Nucleic Acids Res.*, vol. 30, no. 1, pp. 80–83, 2002.
[60] W.-K. Huh, J. Falvo, W.-K. Huh, J. V Falvo, L. Gerke, A. Carroll, L. C. Gerke, A. S. Carroll, R. Howson, J. Weissman, R. W. Howson, E. O'Shea, J. S. Weissman, and E. K. O'Shea, "Global analysis of protein localization in budding yeast.," *Nature*, vol. 425, no. 6959, pp. 686–91, 2003.
[61] N. Wiwatwattana, C. M. Landau, G. J. Cope, G. A. Harp, and A. Kumar, "Organelle DB: An updated resource of eukaryotic protein localization and function," *Nucleic Acids Res.*, vol. 35, no. SUPPL. 1, pp. 810–814, 2007.
[62] U. Schulze, "Anaerobic Physiology of Saccharomyces cerevisiae," Technical University of Denmark, 1995.
[63] E. Oura, "The effect of aeration intensity on the biochemical composition of Baker's yeast: Activities of enzymes of the glycolytic and pentose phosphate pathways," *Biotechnol. Bioeng.*, vol. 18, no. 3, pp. 415–420, Mar. 1976.
[64] C. Verduyn, "Physiology of yeasts in relation to biomass yields.," *Antonie Van Leeuwenhoek*, vol. 60, no. 3–4, pp. 325–353, 1991.
[65] A. N. N. Vaughan-martini and A. Martini, "A Taxonomic Key for the Genus Saccharomyces," *Syst. Appl. Microbiol.*, vol. 16, no. 1, pp. 113–119, 1993.

[66] Colin Ratledge and Christopher T. Evans, “Lipids and their Metabolism,” *The Yeasts*, vol. 3, pp. 367–444, 1989.
[67] G. G. Zampar, A. Kümmel, J. Ewald, S. Jol, B. Niebel, P. Picotti, R. Aebersold, U. Sauer, N. Zamboni, M. Heinemann, a. Kummel, J. Ewald, S. Jol, B. Niebel, P. Picotti, R. Aebersold, U. Sauer, N. Zamboni, M. Heinemann, A. Kü Mmel, J. Ewald, S. Jol, B. Niebel, P. Picotti, R. Aebersold, U. Sauer, N. Zamboni, and M. Heinemann, “Temporal system-level organization of the switch from glycolytic to gluconeogenic operation in yeast,” *Mol. Syst. Biol.*, vol. 9, no. 651, pp. 651–651, 2014.
[68] K. Correia and R. Mahadevan, “Pan-genome-scale network reconstruction: a framework to increase the quantity and quality of metabolic network reconstructions throughout the tree of life,” *bioRxiv*, p. 412593, 2018.
[69] G. Jansen, C. Wu, B. Schade, D. Y. Thomas, and M. Whiteway, “Drag&Drop cloning in yeast,” *Gene*, vol. 344, pp. 43–51, 2005.
[70] O. W. Ryan, J. M. Skerker, M. J. Maurer, X. Li, J. C. Tsai, S. Poddar, M. E. Lee, W. DeLoache, J. E. Dueber, A. P. Arkin, and J. H. D. Cate, “Selection of chromosomal DNA libraries using a multiplex CRISPR system,” *Elife*, vol. 3, no. August2014, pp. 1–15, Aug. 2014.
[71] B. Bakker, “Stoichiometry and compartmentation of NADH metabolism in Saccharomyces cerevisiae,” *FEMS Microbiol. Rev.*, vol. 25, no. 1, pp. 15–37, Jan. 2001.
[72] L. Bourgeois, M. E. Pyne, and V. J. J. Martin, “A Highly Characterized Synthetic Landing Pad System for Precise Multicopy Gene Integration in Yeast,” *ACS Synth. Biol.*, vol. 7, no. 11, pp. 2675–2685, Nov. 2018.
[73] D. B. Flagfeldt, V. Siewers, L. Huang, and J. Nielsen, “Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae.,” *Yeast*, vol. 26, no. 10, pp. 545–551, 2009.
[74] M. D. Mikkelsen, L. D. Buron, B. Salomonsen, C. E. Olsen, B. G. Hansen, U. H. Mortensen, and B. A. Halkier, “Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform,” *Metab. Eng.*, vol. 14, no. 2, pp. 104–111, 2012.
[75] M. Grey, M. Schmidt, and M. Brendel, “Overexpression of ADH1 confers hyper-resistance to formaldehyde in Saccharomyces cerevisiae.,” *Curr. Genet.*, vol. 29, no. 5, pp. 437–440, 1996.
[76] E. P. Wehner, E. Rao, and M. Brendel, “Molecular structure and genetic regulation of SFA, a gene responsible for resistance to formaldehyde in Saccharomyces cerevisiae, and characterization of its protein product.,” *Mol. Gen. Genet.*, vol. 237, no. 3, pp. 351–8, Mar. 1993.
[77] K. Weinhandl, M. Winkler, A. Glieder, and A. Camattari, “Carbon source dependent promoters in yeasts,” *Microb. Cell Fact.*, vol. 13, no. 1, p. 5, 2014.
[78] L. Caspeta, S. Shoaie, R. Agren, I. Nookaew, and J. Nielsen, “Genome-scale metabolic reconstructions of Pichia stipitis and Pichia pastoris and in-silico evaluation of their potentials,” *BMC Syst. Biol.*, vol. 6, no. 1, p. 24, 2012.
[79] D. Yasokawa, S. Murata, Y. Iwahashi, E. Kitagawa, R. Nakagawa, T. Hashido, and H. Iwahashi, “Toxicity of methanol and formaldehyde towards Saccharomyces cerevisiae as assessed by DNA microarray analysis,” *Appl. Biochem. Biotechnol.*, vol. 160, no. 6, pp. 1685–1698, 2010.
[80] M. A. H. Luttik, K. M. Overkamp, P. Kötter, S. De Vries, J. P. Van Dijken, and J. T. Pronk, “The Saccharomyces cerevisiae NDE1 and NDE2 genes encode separate

mitochondrial NADH dehydrogenases catalyzing the oxidation of cytosolic NADH," *J. Biol. Chem.*, vol. 273, no. 38, pp. 24529–24534, 1998.
[81] F. Rodrigues, P. Ludovico, and C. Leão, "Sugar Metabolism in Yeasts: an Overview of Aerobic and Anaerobic Glucose Catabolism," in *Biodiversity and Ecophysiology of Yeasts*, Berlin/Heidelberg: Springer-Verlag, 2006, pp. 101–121.
[82] K. Larsson, R. Ansell, P. Eriksson, and L. Adler, "A gene encoding sn-glycerol 3-phosphate dehydrogenase (NAD +) complements an osmosensitive mutant of Saccharomyces cerevisiae," *Mol. Microbiol.*, vol. 10, no. 5, pp. 1101–1111, Dec. 1993.
[83] C. M. Bussineau and E. T. Papoutsakis, "Intracellular reaction rates, enzyme activities and biomass yields in Methylomonas L3: growth rate and substrate composition effects," *Appl. Microbiol. Biotechnol.*, vol. 24, no. 6, pp. 435–442, 1986.
[84] J. E. Gonzalez, R. K. Bennett, E. T. Papoutsakis, and M. R. Antoniewicz, "Methanol assimilation in Escherichia coli is improved by co-utilization of threonine and deletion of leucine-responsive regulatory protein," *Metab. Eng.*, vol. 45, no. August 2017, pp. 67–74, 2018.
[85] T. Hasunuma, T. Sanda, R. Yamada, K. Yoshimura, J. Ishii, and A. Kondo, "Metabolic pathway engineering based on metabolomics confers acetic and formic acid tolerance to a recombinant xylose-fermenting strain of Saccharomyces cerevisiae.," *Microb. Cell Fact.*, vol. 10, no. 1, p. 2, 2011.
[86] Y. H. Lau, T. W. Giessen, W. J. Altenburg, and P. A. Silver, "Prokaryotic nanocompartments form synthetic organelles in a eukaryote," *Nat. Commun.*, vol. 9, no. 1, 2018.
[87] P. M. Bruinenberg, J. P. Van Dijken, and W. a. Scheffers, "A Theoretical Analysis of NADPH Production and Consumption in Yeasts," *Microbiology*, vol. 129, no. 4, pp. 953–964, 1983.
[88] S. A. Henry, *Membrane Lipids of Yeast: Biochemical and Genetic Studies*, vol. 11B. 1982.
[89] B. D. Heavner and N. D. Price, "Comparative Analysis of Yeast Metabolic Network Models Highlights Progress, Opportunities for Metabolic Reconstruction," *PLoS Comput. Biol.*, vol. 11, no. 11, pp. 1–26, 2015.
[90] W. Zhang, C. P. Liu, M. Inan, and M. M. Meagher, "Optimization of cell density and dilution rate in Pichia pastoris continuous fermentations for production of recombinant proteins," *J. Ind. Microbiol. Biotechnol.*, vol. 31, no. 7, pp. 330–334, 2004.
[91] Y. Morales, M. Tortajada, J. Picó, J. Vehí, and F. Llaneras, "Validation of an FBA model for Pichia pastoris in chemostat cultures," *BMC Syst. Biol.*, vol. 8, p. 142, 2014.