METABOLIC ENGINEERING FOR LIGNIN VALORIZATION AND MARTIAN *IN SITU* RESOURCE UTILIZATION

A Dissertation
Presented to
The Academic Faculty

By

Nicholas S. Kruyer

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemical and Biomolecular Engineering

Georgia Institute of Technology

May 2021

METABOLIC ENGINEERING FOR LIGNIN VALORIZATION AND MARTIAN *IN SITU* RESOURCE UTILIZATION

Approved by:

Dr. Pamela Peralta-Yahya, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Mark Styczynski
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Andreas Bommarius, Advisor
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Corey Wilson
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Matthew Realff
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Date Approved April 6, 2021

ACKNOWLEDGEMENTS

I'd first like to thank my support system over the last five years. Many family and friends have offered moral support keeping me motivated, inspired, and providing joy, even when results in the lab did not. My parents, Pete and Lauree Kruyer, and brother and sister, Sam and Sarah Kruyer, have supported me through my entire academic career, cheering me on from close and afar, giving me wonderful footsteps to follow and being my role models. Also, my partner Valeria. Building our life together has given me motivation to work hard, but also chances to enjoy life and explore Atlanta and beyond. You have been my constant support throughout my entire PhD, I would not have made it without your love, support, and the joy you bring me every day. I can't wait to have you by my side through our next steps in life.

Prior to coming to Georgia Tech, my scientific career was shaped by Dr. Michael Betenbaugh, Dr. An Goffin, and Jordan Baker at Johns Hopkins. Thanks to Jordan for taking me under his wing as my first research mentor in the Betenbaugh Lab, and to Dr. Betenbaugh for giving me the opportunity to learn the scientific process as well as my first taste of metabolic engineering in a class that sparked my interest in the topic. Dr. Betenbaugh, along with Dr. Goffin as my academic advisor, encouraged me to pursue a PhD, charting a path that I was reluctant to set myself on, but am forever grateful that I did.

At Georgia Tech, my thesis would not have been possible without the guidance of my advisors, Dr. Pamela Peralta-Yahya and Dr. Andy Bommarius. You've both provided invaluable advice and mentorship, directing both my thesis project as well as my future career. Thanks for accepting me into your labs and for teaching me to become the scientist

I am today. I'd also like to thank members of the Peralta-Yahya and Bommarius labs for their training, advice and friendship. Amy Ehrenworth and Stephen Sarria were excellent role models to have in my first few years of grad school, coaching me through my first mini-prep and transformation and providing advice on how to survive the long journey to a PhD. It was a pleasure to sit next to Emily Yasi throughout my time in the Peralta-Yahya lab. She helped me troubleshoot many research problems and taught me much of what I know about yeast. In the Bommarius Lab, older graduate students, especially Thomas Kwok, taught me to think critically as a chemical engineer, and how to think through many research struggles. Finally, members of the Georgia Tech Lignin Group and the Renewable Bioproducts Institute have become important scientific companions as well as friends. I am grateful for the opportunities provided by RBI, which funded much of my PhD work.

Lastly, the work presented in this thesis was not done alone. Undergraduate students Natalia Wauldron, who helped immensely in cloning and testing many strains, and Lily Torp, who helped on many side projects and generally kept our lab operating smoothly, were invaluable. I've also enjoyed the chance to work with a number of fantastic research technicians including Paul Branham, Sara Eisen, and Weldon Lane. Our many conversations about science and beyond made the lab an enjoyable place to work.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii

LIST OF TABLES ix

LIST OF FIGURES x

LIST OF SYMBOLS AND ABBREVIATIONS xvi

SUMMARY xx

CHAPTER 1. INTRODUCTION 1
1.1 Introduction 1
1.2 References 5

CHAPTER 2. METABOLIC ENGINEERING STRATEGIES TO BIO-ADIPIC ACID 7
2.1 Abstract 7
2.2 Introduction 8
2.3 Metabolic pathways to muconic, glucaric, and adipic acid 11
2.4 Muconic acid from cellulose- and hemicellulose-derived feedstocks 13
2.5 Muconic acid from lignin-derived feedstock 16
2.6 Key enzymes in microbial synthesis of muconic acid 17
2.7 Glucaric acid from cellulose-derived feedstock 18
2.8 Adipic acid production from lipids 19
2.9 Adipic acid production from tricarboxylic acid cycle intermediates 20
2.10 Alternative chassis and co-cultures to produce adipic acid and its precursors 21
2.11 Conclusions 22
2.12 References 24

CHAPTER 3. FULLY BIOLOGICAL PRODUCTION OF ADIPIC ACID ANALOGS FROM BRANCHED CATECHOLS 29
3.1 Abstract 29
3.2 Introduction 30
3.3 Materials and Methods 34
3.3.1 Reagents 34
3.3.2 Strains and plasmids 34
3.3.3 Screening CatA for muconic acid production from catechol 35
3.3.4 Screening MAR for adipic acid production from muconic acid 36
3.3.5 Cell lysate MAR activity on muconic acid isomers 36
3.3.6 Production of adipic acid analogs from lignin-derived aromatic monomers 37
3.3.7 Chemical analysis 38

3.3.8 SDS-PAGE analysis of CatA expression 39
3.3.9 Western blot analysis of CatA expression 40
3.4 Results and Discussion 40
3.4.1 Catechol to muconic acid conversion 40
3.4.2 Muconic acid to adipic acid conversion 44
3.4.3 Adipic acid production from catechol 46
3.4.4 MAR muconic acid isomer preference 46
3.4.5 MAR muconic acid analog preference 48
3.5 Conclusions 51
3.6 References 51

CHAPTER 4. DESIGNING THE PRODUCTION OF A MARTIAN ROCKET PROPELLANT USING IN SITU RESOURCE UTILIZATION 55
4.1 Abstract 55
4.2 Introduction 56
4.3 Results 60
4.3.1 Challenges in producing rocket propellant on Mars 60
4.3.2 Martian rocket propellant design 62
4.3.3 2,3-butanediol as a Martian rocket propellant 63
4.3.4 Process design overview 65
4.3.5 Cyanobacteria growth 67
4.3.6 Cyanobacteria growth modes 68
4.3.7 Cyanobacteria biomass concentration 72
4.3.8 Cyanobacteria biomass digestion 72
4.3.9 Microbial production of 2,3-butanediol on Mars 74
4.3.10 Rocket propellant extraction and separation 74
4.3.11 Water recycling 75
4.3.12 Full process analysis 76
4.4 Discussion 78
4.4.1 Biological process improvement strategies 78
4.4.2 Materials process improvement strategies 82
4.5 Conclusions 84
4.6 References 85

CHAPTER 5. MEMBRANE AUGMENTED CELL-FREE SYSTEMS: A NEW FRONTIER IN BIOTECHNOLOGY 90
5.1 Abstract 90
5.2 Introduction 91
5.3 Production of Membrane Proteins in CFSs 96
5.3.1 Membrane augmented CFSs 96
5.3.2 Chemical composition of membrane augmented CFSs 98
5.3.3 Physical properties of membrane augmented CFSs 99
5.3.4 Membrane protein insertion into membrane augmented CFS 100
5.4 Applications of Membrane Augmented CFSs 102
5.4.1 Enzyme Assays 108
5.4.2 Energy Production 108

5.4.3 Chemical Biosynthesis 109
5.4.4 Synthetic cells 110
5.5 Control Strategies for the Generation of Membrane Augmented CFSs 114
5.5.1 Genetic control of membrane properties 114
5.5.2 CRISPRa/i – a control system for the expression of membrane proteins in membrane augmented CFS 115
5.6 Future Directions 117
5.7 References 118

CHAPTER 6. FUTURE DIRECTIONS 126
6.1 Conclusions and Future Directions 126
6.1.1 Expanding use of lignin as a microbial feedstock 126
6.1.2 Engineering improved cyanobacterial growth 128
6.1.3 Applications of membrane associated proteins in cell-free systems 129
6.2 References 130

APPENDIX A. TABLE OF PLASMIDS 131
A.1 Plasmids used in this work 131
A.2 References 134

APPENDIX B. TABLE OF STRAINS 135
B.1 Strains used in this work 135
B.2 References 135

APPENDIX C. PRIMER TABLES 136
C.1 Primers used in this work 136

APPENDIX D. DNA SEQUENCES 138
D.1 Chapter 3 genes 138
D.2.1 Pseudomons putida CatA (WP_010954549) 138
D.2.2 Acinetobacter baylyi ADP1 CatA (CAG68305) 139
D.2.3 Candida albicans CatA (KCQ97177) 140
D.2.4 Rhodococcus sp. AN22 CatA (BAH56722) 140
D.2.5 Rhodococcus opacus CatA (3HGI_A) 141
D.2.6 Bacillus coagulans MAR (AEO99944) 142
D.2.7 Clostridium acetobutylicum MAR (AEI32805) 144
D.2 Ribosomal binding site (RBS) sequences 146
D.3 References 146

APPENDIX E. FLUX BALANCE ANALYSIS OF METABOLIC PATHWAYS 147
E.1 Flux Balance Analysis of Adipic Acid Producing Pathways 147
E.1.1 Muconic acid production from dehydroxyshikimic acid 147
E.1.2 Muconic acid production from anthranilate 148
E.1.3 Muconic acid production from p-hydroxybenzoate 148
E.1.4 Muconic acid production from isochorismate 149
E.1.5 Muconic acid production from 2,3-dihyroxybenzoate 149

E.1.6 Glucaric acid production from glucose 150
E.1.7 Adipic acid production from TCA cycle intermediates – route 1 150
E.1.8 Adipic acid production from TCA cycle intermediates – route 2 151
E.1.9 Muconic acid production from lignin-derived p-coumaric acid 152
E.1.10 Muconic acid production from lignin-derived ferulic acid 152
E.1.11 Muconic acid production from lignin-derived benzoate 153
E.2 Flux Balance Analysis of Short Chain Diols 153
E.2.1 1,2-Propanediol production from glucose 154
E.2.2 1,3-Butanediol production from glucose 154
E.2.3 2,3-Butanediol production from glucose 155
E.3 References 155

APPENDIX F. UNIT OPERATION CALUCLATIONS FOR BIOTECHNOLOGY-ENABLED IN SITU RESOURCE UTILIZATION 157
F.1 Unit Operation Specifications 157
F.2 Propellant Property Calculation 159
F.2.1 Theoretical specific impulse 159
F.2.2 Lower heating value 159
F.2.3 Required oxygen to fuel ratio 160
F.3 Process Water Use Calculation 160
F.4 Process Power Use Calculation 160
F.4.1 Suspended cyanobacteria growth 160
F.4.2 Biofilm cyanobacteria growth 161
F.4.3 Cyanobacteria concentrator 161
F.4.4 Digester, fermenter, and liquid-liquid extraction units 161
F.4.5 Membrane separation unit 161
F.5 Process Mass Calculation 162
F.5.1 Cyanobacteria growth 162
F.5.2 Cyanobacteria concentrator 163
F.5.3 Enzyme digester 163
F.5.4 Fermenter 163
F.5.5 Liquid-liquid extractor 163
F.5.6 Membrane separator 164
F.5.7 Storage tanks 164
F.5.8 Water treatment 164
F.6 Photosynthetic Oxygen Production 164
F.7 References 164

APPENDIX G. ADVANCING THE POTENTIAL FOR THE PRODUCTION OF CHEMICALS FROM CARBON DIOXIDE IN ESCHERICHIA COLI 166
G.1 Perspective 166
G.2 References 170

LIST OF TABLES

Table 2.1. Yields for muconic acid, glucaric acid and adipic acid from glucose **10**

Table 3.1. Rationale for CatA screening choice. **42**

Table 4.1. Design criteria for proposed Martian rocket propellants **63**

Table 4.2. Bioreachability of short chain diols **64**

Table 4.3. Cyanobacteria growth model parameters **70**

Table 4.4. Mass of cyanobacteria growth system components. **72**

Table 4.5. Effect of enzyme digester size on water usage and payload mass. **73**

Table 4.6. Water, power, and payload mass requirement for bio-ISRU **77**

Table 4.7. Comparison of bio-ISRU with NASA DRA 5.0 ISRU **78**

Table 5.1. Applications of membrane proteins in cell-free systems **103**

Table A.1. Plasmids used in Chapter 3 **131**

Table B.1. Strains used in Chapter 3 **135**

Table C.1. Primers used in Chapter 3 **136**

Table D.1. RBS sequences **146**

LIST OF FIGURES

Figure 2.1 Bio-adipic acid production. In the last two years an increasing number of feedstocks and microbial chassis have been used for the production of adipic acid and its immediate precursors, muconic acid and glucaric acid. Adipic acid is one of the most industrially important dicarboxylic acids being used in the synthesis of nylon and polyesters. 8

Figure 2.2 Metabolic engineering strategies to bio-adipic acid production. Shown, metabolic pathways highlighted in this review. Green box: lignin-derived substrates. Yellow box: cellulose- and hemicellulose-derived substrates. Grey box: lipid-derived substrates. Adipic acid, muconic acid and glucaric acid in blue. DHS: 3-dehydroshikimic acid; PHB: *p*-hydroxybenzoic acid; DHB: 2,3-dihydroxybenzoic acid**.** 11

Figure 3.1 Muconic acid production from catechol. (a) Schematic for bioproduction of adipic acid analogs CatA: catechol-1,2-dioxygenase, MAR: muconic acid reductase. (b) Catechol to muconic acid conversion using CatA. At low pH *cis, cis*-muconic acid (cc-muconic acid) can isomerize to *cis, trans*-muconic acid (ct-muconic acid), and *trans, trans*-muconic acid (tt-muconic acid). (c) Liquid chromatograph/mass spectrometry (LC/MS) chromatogram of *E. coli* expressing *Rhodococcus* sp. AN22 CatA fed 1 g/L of catechol after 23 h incubation resulted in ccMA and ctMA. No ttMA was observed. Also shown, standards for catechol (6.8 min), ccMA (9.4 min), ctMA (13.6 min), and ttMA (9.6 min). Sample data for LC/UV absorbance spectra found in Figure 3.2 (d) Optimization of muconic acid production by screening CatAs from five different sources using three different promoter strengths and in low and medium copy plasmids. Experiments were run in triplicate and error bars represent the standard deviation from the mean. 32

Figure 3.2 LC/UV spectra for standards of catechol, ccMA, ctMA, and ttMA, and the supernatant of *E. coli* expressing CatA-AN22 from the pTrc promoter in a medium copy plasmid (pNK56) when fed catechol. Corresponds to data in Figure 3.1c. 33

Figure 3.3 Expression of CatA homologs. (a) Coomassie stained SDS-PAGE gel for His_6-CatA homologs expressed under pTrc promoter, in low copy vector. (b) Western Blot His_6-CatA homologs. CatA Molecular 41

Weight: *P. putida* – 34.3 kDa, *A. baylyi* ADP1 – 34.3 kDa, *C. albicans* – 33.8 kDa, *R. opacus* – 30.7 kDa, *R. sp.* AN22 – 31.6 kDa.

Figure 3.4 Active site alignment of *Pseudomonas putida* CatA and *Acinetobacter baylyi* ADP1 CatA. Active site residues from Han *et al*. [27]. Colors represent different classes of amino acid. Key difference at residue 76 may account for conversion differences when using *P. putida* CatA and *A. baylyi* ADP1 CatA. 42

Figure 3.5 Isomerization of *cis, cis*-muconic acid to *cis, trans*-muconic acid as a function of pH. Ratio of concentration of ctMA to concentration of ccMA after 300 µM solution of ccMA was incubated for 24 h at 37°C and 250 RPM in M9 media of varying pH value. 42

Figure 3.6 Adipic acid production from muconic acid and catechol. (a) Reduction of *cis, cis*-muconic acid (cc-muconic acid) and *cis, trans*-muconic acid (ct-muconic acid) to adipic acid by muconic acid reductase (MAR). (b) LC/MS chromatograms of *E. coli* expressing *Bacillus coagulans* MAR fed 500 µM ccMA after 24 h incubation results in adipic acid, hexenedioic acid as well as unreacted *cis, cis*-muconic acid and *cis, trans*-muconic acid. Also shown, standards for adipic acid (m/z 146; 10.8 min), *cis, trans*-muconic acid (m/z 142; 16.1 min), and *cis, cis*-muconic acid (m/z 142; 10.0 min). (c) Adipic acid yields (mol%) obtained by screening MARs from two different sources using two different promoter strengths and in low, medium and high copy plasmids. (d) Adipic acid yields obtained from feeding 500 mg/L catechol to *E. coli* co-expressing *Rhodococcus sp.* AN22 CatA and *Bacillus coagulans* MAR in a two and one plasmid system. For c and d, experiments were run in triplicate and error bars represent the standard deviation from the mean. 44

Figure 3.7 Muconic acid reductase substrate preference. (a) Adipic acid production of *Escherichia coli* expressing *Bacillus coagulans* muconic acid reductase (MAR) or *Clostridium acetobutylilcum* MAR when feeding 500 µM of *cis, cis*-muconic acid (cc-muconic acid), *cis, trans*-muconic acid (ct-muconic acid), or *trans, trans*-muconic acid (tt-muconic acid) isomers. (b) Adipic acid yield of lysed *E. coli* expressing *B. coagulans* MAR and *C. acetobutylicum* MAR fed 583 µM of muconic acid isomers. All experiments were performed in triplicate and error bars represent the standard deviation from the mean. 46

Figure 3.8 Production of muconic acid and adipic acid analogs. Reaction schematic and LC/MS chromatograms of *Escherichia coli* co-expressing *Rhodococcus sp*. AN22 CatA and either *B. coagulans* MAR or *C. acetobutylicum* MAR and fed 1 mM (a) 3-methyl- 48

catechol results in 2-methyl hexenedioic acid and 2-methyl-adipic acid; (b) 4-methyl-catechol results in 3-methy hexenedioic acid, and (c) 4-ethyl-catechol results in 3-ethyl hexenedioic acid. *: Peaks visible in MS but not in UV channel, thus not muconic acid analog peaks (Figure 3.9).

Figure 3.9 Comparison of MS and UV active peaks for Figure 3.8. (a) 2-methyl muconic acid (m/z 155) and UV trace. (b) 3-methyl-muconic acid (m/z 155) and UV trace. (c) 3-ethyl-muconic acid (m/z 169) and UV trace. All traces come from *E. coli* ΔiscR expressing CatA from *Rhodococcus sp*. AN22 and MAR from *Bacillus coagulans* fed 1 mM of substituted catechol after 2-hr aerobic and 22-hr anaerobic incubation. Peaks appearing in both SIM and UV channels are muconic acid analogs. *: Peaks only present in SIM (likely media components). Difference in peak times between UV trace and SIM due to volume (tubing) distance between MWD and MS detectors. 49

Figure 4.1 Chemical enabled ISRU fuel production (A) Martian ISRU for rocket propellant production can greatly reduce mission costs and take advantage of abundant Martian CO_2 (B) Current ISRU fuel strategies usually target CH_4 as the fuel of interest. CH_4 can be sent from Earth or produced using the Sabatier reaction using H_2 shipped from Earth or produced from harvested Martian H_2O using water electrolysis. The required LOX can come from O_2 produced during water electrolysis (resulting in excess CH_4) or supplemented with solid oxide CO_2 electrolysis (SOCE). (C) Adding oxygen atoms to the fuel compound reduces the ratio of mass of required LOX per mass of required fuel. An unoxygenated compound, ex. CH_4, requires 4 times as much O_2 as fuel, whereas oxygenated compounds, ex. diols, have a ratio around 2. (D) Adding oxygen atoms to the fuel compound reduces the lower heating value (LHV) from 50.1 kJ/kg for CH_4 to 25-30 kJ/kg for alcohols and 20-25 kJ/kg for diols. 56

Figure 4.2 Metabolic pathways for short chain diols. Modeled pathways for conversion of glucose to 1,2-propanediol, 2,3-butanediol, and 1,3-butanediol in *E. coli*. 64

Figure 4.3 Process overview for continuous production of 2,3-BDO from cyanobacterial glucose. (A) The 2,3-BDO production process requires four modules. First, cyanobacterial cultivation for biomass production in photobioreactors or biofilm growth. Second, biomass preprocessing consisting of biomass concentration through membrane filtration and enzymatic breakdown in a stirred tank. Third, microbial fermentation of the cyanobacterial glucose to produce 2,3-BDO. Fourth, 2,3-BDO separation and purification using sequential liquid-liquid extraction and membrane separation. 65

(B) A rendering of what the process might look like on Mars, with a Martian rover included for scale. Importantly, the cyanobacterial cultivation modules makes up the bulk of the material and land footprints.

Figure 4.4 Comparison of biofilm and suspended cyanobacterial growth methods. (A) Representation of biofilm growth, where the cyanobacteria grow in a thin film on a porous substrate, suspended photobioreactor (PBR) growth where the cyanobacteria are grown in PBRs filled with water, and open pond growth where the cyanobacteria are grown in large horizontal, ponds open to the atmosphere. (B) Cyanobacterial growth model developed by Karemore *et al*. [38] to describe growth of *Arthrospira platensis* based on a light limited Monod growth model. The equation gives daily, areal productivity, taking into account both sunlight driven photosynthetic carbon fixation as well as light-independent respiration. (C) Total payload mass of the cyanobacterial cultivation unit for biofilm and suspended growth, taking into account the growth nutrients, support frame and growth substrate (biofilm) or PBR material (suspended). 68

Figure 4.5 Optimization of digester residence time. (A) The size of the enzyme digester depends on the residence time, which also determines glucose yield from cyanobacterial biomass. (B) Increasing digester reactor time decreases process water used and (C) payload mass for both biofilm and suspended growth modes. 72

Figure 4.6 Process metrics for 2,3-BDO production from cyanobacterial glucose. Comparison of (A) water, (B) power, and (C) mass requirements of biofilm and suspended growth modes, accounting for all unit operations in the proposed process. Results are based on 15 tons of 2,3-BDO produced over 500 sols. 75

Figure 4.7 Impact of biological process improvements on process specifications. Payload mass, power use, and water use can all be reduced by biological improvements in (A) cyanobacterial growth productivity, (B) the yield of glucose from biomass in the enzyme digester, and (C) the yield of 2,3-BDO from glucose in the fermenter. Solid vertical lines represent productivity and yields used in the unoptimized model, dashed vertical lines represent engineering targets used in the optimized model. 78

Figure 4.8 Process optimization to reduce payload mass. Biological and material optimization can reduce payload mass of the biotechnology-enabled ISRU process. Unoptimized is the original modelled scenario, material optimized uses HDPE and LDPE for reactor components instead of steel and cotton. Biological optimized 81

improves cyanobacterial biomass productivity 2-fold to 13.28 g/m^2/day, improves digester yield to 60 wt%, and improves fermenter yield to 95% of theoretical max. Full optimized implements both material and biological optimizations. The DRA 5.0 ISRU is the mass of shipping CH_4 as fuel from Earth along with process equipment to produce O_2 via SOCE in situ. This payload mass comes from NASA DRA 5.0 [2].

Figure 5.1 Biological roles of membrane proteins and cell membranes. (A) Cellular processes that occur at the cell membrane include cell signaling, lipid biosynthesis, methane utilization, metal remediation, and energy generation. (B) Membrane compositions in different organisms. PG: Phosphatidylglycerol, PA: Phosphatidic acid, PE: Phosphatidylethanolamine, CL: Cardolipin, PC: Phosphatidylcholine, PI: Phosphatidylinositol, PS: Phosphatidylserine, DAG: Diacylglycerol, MGDG: Monogalactosyldiacylglycerol, DGDG: Digalactosyldiacylglycerol. (C) Sample phospholipid structures. Phospholipid heads in blue. 92

Figure 5.2 Membrane augmented cell-free systems. Mimetic membranes, in the form of liposomes and nanodiscs, are added directly to CFSs to aid in membrane protein production. 97

Figure 5.3 Production of natural products using cell-free systems. Compounds produced to date do not require membrane proteins or membrane augmented CFSs. In the dashed box, natural products that depend on the production of transmembrane proteins that could be synthesized by using a membrane augmented CFS. 109

Figure 5.4 Applications of membrane augmented cell-free systems. (A) Potential bio-industrial applications of membrane proteins in membrane augmented cell-free systems. (B) Dynamic CRISPR-Cas control system can be implemented to stagger phospholipid and natural product biosynthesis, allowing for *in situ* liposome self-assembly prior to membrane protein production and chemical synthesis. (C) Sample two-stage control strategy for staggered phospholipid and natural product biosynthesis. First, constitutive (const) or induced expression of a modified guide RNA that enables the recruitment of a transcriptional activator (scaffold RNA, or scRNA) activates expression of the phospholipid pathway. Second, a user-supplied inducer or sufficient phospholipid concentration triggers expression of the natural product biosynthesis pathway (through targeted scRNA) as well as repression of the phospholipid pathway using a gRNA coupled with catalytically inactive Cas9 (dCas9). Suppression of lipid biosynthesis during stage two helps conserve precious CFS resources. 111

Figure G.1 Advancing the potential for the production of chemicals from carbon dioxide in *E. coli*. 167

LIST OF SYMBOLS AND ABBREVIATIONS

3MC	3-Methylcatechol
4EC	4-Ethylcatechol
4MC	4-Methylcatechol
AcCoA	Acetyl-CoA
ATP	Adenosine triphosphate
BDO	Butanediol
BIA	Benzylisoquinoline alkaloid
CatA	Catechol-1,2-dioxygenase
ccMA	*cis, cis*-Muconic acid
CFS	Cell-free system
CL	Cardiolipin
COBRA	Constraint-Based Reconstruction and Analysis
ctMA	*cis, trans*-Muconic acid
CYP	Cytochrome P450
DAG	Diacylglycerol
DAHP	3-Deoxy-D-arabino-heptulosonic acid 7-phosphate
DGDG	Digalactosyldiacylglycerol
DHB	2,3-Dihydroxybenzoic acid
DHS	3-Dehydroshikimic acid
DRA	Design Reference Architecture
E4P	Erythrose-4-phosphate

FDA	Food and Drug Administration
GFP	Green fluorescent protein
GPCR	G protein-coupled receptor
HDPE	High-density polyethylene
IPTG	Isopropyl β-d-1-thiogalactopyranoside
I_{sp}	Specific impulse
ISRU	*In situ* resource utilization
LC	Liquid chromatography
LDPE	Low-density polyethylene
LHV	Lower heating value
LLE	Liquid-liquid extraction
LOX	Liquid oxygen
MAR	Muconic acid reductase
MAR-BC	Muconic acid reductase from *Bacillus coagulans*
MAR-CA	Muconic acid reductase from *Clostridium acetobutylicum*
MAV	Mars ascent vehicle
MGDG	Monogalactosyldiacylglycerol
MIA	Monoterpene indole alkaloid
MIOX	Myo-inositol oxygenase
MMO	Methane monooxygenase
MS	Mass spectrometry
MTY	Maximum theoretical yield
MWD	Multiple wavelength detector
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate

NASA	National Aeronautics and Space Administration
PA	Phosphatidic acid
PAR	Photosynthetically active radiation
PBR	Photobioreactor
PBS	Phosphate buffered saline
PC	Phosphatidylcholine
PCA	Protocatechuic acid
PCR	Polymerase chain reaction
PDO	Propanediol
PE	Phosphatidylethanolamine
PEG	Polyethylene glycol
PEP	Phosphoenolpyruvate
PG	Phosphatidylglycerol
PHB	*p*-Hydroxybenzoic acid
PI	Phosphatidylinositol
PMME	Phosphatidylmonomethylethanolamine
PS	Phosphatidylserine
PURE	Purified cell machinery CFS
RBS	Ribosomal binding site
RPM	Rotations per minute
SIM	Selective ion monitoring
SOCE	Solid oxide CO_2 electrolysis
SuCoA	Succinyl-CoA
TCA	Tricarboxylic acid
TEA	Techno-economic analysis

ttMA *trans, trans*-Muconic acid

UV Ultraviolet

SUMMARY

Global climate change and depleting fossil fuel resources demand new strategies to replace petroleum-derived commodity chemicals. One strategy towards renewable chemical production is to take advantage of microbes that can naturally metabolize renewable carbon sources, such as simple sugars, to produce a variety of value-added compounds. Use of microbial fermentation takes advantage of natural pathways towards compounds of interest, often towards compounds that are difficult to produce chemically. Indeed, microbial production of commodity chemicals holds process advantages in terms of renewability, specificity, and use of mild process conditions such as low temperatures and atmospheric pressures. In this work, I explore how we can expand upon native metabolic pathways to engineer production of commodity chemicals through metabolic engineering, synthetic biology, and process design.

Chapter 1 introduces metabolic engineering, specifically metabolic engineering towards utilization of uncommon, non-plant derived sugar, feedstocks.

Chapter 2 provides a review of metabolic engineering strategies for microbial production of adipic acid, a commodity chemical used to make nylon-6,6. The most common adipic acid production pathway begins with petroleum derived cyclohenxanol and cyclohexanone and is responsible for ~10% of global nitrous oxide (N_2O) emissions. The non-renewable starting material and substantial production of greenhouse gases provide ample motivation for finding alternative pathways to adipic acid. While a native metabolic pathway to adipic acid has yet to be discovered, in the review we highlight engineered microbial pathways to produce adipic acid, as well as native pathways to produce the direct precursor, *cis, cis*-muconic acid. Our literature review revealed a distinct advantage to

using lignin-derived monomers, a non-sugar feedstock, instead of glucose as the starting material. The review also highlights the process simplification that comes with a fully biological pathway to adipic acid, rather than a hybrid biological-chemical process. However, distinct improvements in pathway yields are required before the fully biological pathways can compete with current industrial chemical routes.

In **Chapter 3**, we develop insight into fully biological production of adipic acid, exploring the enzyme cascade of catechol-1,2 dioxygenase (CatA) and muconic acid reductase (MAR) which converts catechol to adipic acid. This conversion makes up the final steps of biological routes to adipic acid that go through a *cis, cis*-muconic acid intermediate. Such pathways have shown the highest theoretical and experimental yields, regardless of starting feedstock. In particular, we are interested in further characterizing the substrate specificity of MAR, as the common substrate, *cis, cis*-muconic acid, is not the only muconic acid isomer found in fermentation broths. We show that MAR can metabolize all muconic acid isomers, with membrane transport playing an important role in *in vivo* conversions. Finally, towards lignin valorization, we tested the enzyme cascade against other lignin-derived monomers, namely alkyl-branched catechols. We showed that the position of the alkyl side chain in the resulting muconic acid analog impacts MAR activity with 2-methyl-muconic acid being converted to 2-methyl-adipic acid but 3-methyl- and 3-ethyl-muconic acids only producing 3-methyl- and 3-ethyl-hexendioic acid. This new adipic acid analog could be used to synthesize novel polyamides with altered glass transition properties.

In **Chapter 4**, we explore another target of metabolic engineering, focusing on process design for biotechnology-enabled production of rocket propellant on the surface of

Mars. As we start to plan a human mission to Mars, one of the questions that remains to be solved is how we will fuel the return trip from Mars to Earth. Rather than sending the required rocket propellant from Earth, we propose a process for *in situ* resource utilization (ISRU) to produce the rocket propellant on the surface of Mars using Martian resources. This reduces payload mass of mission launches, reducing mission cost by billions of dollars. Our proposed process uses cyanobacteria to fix abundant Martian CO_2 to produce simple sugars and other nutrients. These sugars can be extracted from the cyanobacteria biomass and upgraded into the proposed rocket propellant, 2,3-butanediol, by an engineered *Escherichia coli*. Our process model also includes the separations unit operations required to purify and concentrate the final product. Our process analysis indicated achievable process improvements in cyanobacterial growth, yield of glucose extraction from cyanobacteria biomass, yield of 2,3-butanediol, and lightweight process materials that are required to make our process competitive in terms of power requirement and total mass to other proposed, chemical-enabled ISRU strategies such as the one proposed by NASA. Here we offer both synthetic biology and process design approaches for overall process improvement. Importantly, the bio-ISRU produces 30 tons of excess O_2, while chemical processes produce none, a vital resource in the O_2-poor Martian environment, making continued process development desirable.

Chapter 5 provides a review of the use of membrane augmented cell-free systems (CFSs) towards chemical production and other applications. CFSs offer alternatives to microbes for enzymatic expression and chemical production with advantages in downstream processing and elimination of toxicity concerns. A key challenge for CFSs is the production of membrane associated proteins, given the absence of stabilizing structures

in cell-free reaction mixes. Despite this, CFSs offer a promising platform for membrane protein production as mimetic membranes, such as liposomes or nanodiscs, can be added to the reaction mix in a user-defined manner. This gives the researcher control over crucial membrane properties such as composition, elasticity, and size. Successful production of membrane proteins in CFSs has potential applications in chemical production, energy production, synthetic cells, and therapeutics.

Finally, looking forward, **Chapter 6** provides future directions towards microbial production of adipic acid from lignin-derived monomers, biotechnology-enabled production of rocket propellants and other chemicals on Mars, as well as improvements in production and applications of membrane proteins in CFSs. **Appendix G** offers a perspective on engineering CO_2 fixation in *E. coli*.

CHAPTER 1.

INTRODUCTION

1.1 Introduction

One piece of the puzzle in the fight against global warming is the reduction of dependence on fossil fuels by replacing their spot in the pipeline with more sustainably sourced compounds. While we are acutely aware of the fossil fuels used to power our cars and generate our electricity, we are often less aware of the role petroleum plays in the production of everyday compounds, materials, and pharmaceuticals. Alternative energy sources such as wind and solar power are helping to curb fossil fuel use towards energy production, yet other solutions are needed to help replace petroleum in production of our everyday goods. Towards this, we can look to again harness the power of nature, though unlike generating electricity with the wind or sunlight, we take advantage of biochemical processes instead of physical ones. Many of the compounds used to create polymers or pharmaceuticals are naturally found in nature, produced in a variety of micro- (i.e. unicellular bacteria and eukaryotes) and macro-organisms (i.e. multicellular plants and fungi). The fields of metabolic engineering and synthetic biology seek to discover and utilize these natural products to produce the compounds we need, renewably.

Renewable chemical production through metabolic engineering takes advantage of native metabolism of renewable substrates such as simple sugars. Focusing on micro-organisms, most prokaryotes and eukaryotes can uptake simple sugars, and through central carbon metabolism convert them to the building blocks of a wide variety of compounds from hydrocarbons [1] to biofuels [2] to active pharmaceutical intermediates [3]. Another

key component of metabolic engineering is the design of novel metabolic pathways to produce a compound of interest. Key aspects of this engineering include expression of pathway enzymes in non-native hosts, balancing metabolic flux through an engineered pathway, and combining enzymes from different pathways and organisms to create novel routes for a given metabolite. Regardless of the pathway, the source of simple sugar feedstock can determine sustainability and must not compete with food sources. Such feedstocks reduce fossil fuel dependence, but at the cost of exacerbation of food scarcity conflicts, especially for commodity chemicals which are produced on the scale of 1000s of tons per year.

Here, nature has again provided a solution to this problem in the form of micro-organisms that are able to metabolize non-sugar feedstocks for production of value added compounds [4]. One promising approach is to use lignin-derived compounds as the carbon source to produce a compound of interest [5]. Lignin, along with cellulose and hemi-cellulose, is one of the three main components of biomass. Whereas cellulose and hemi-cellulose are somewhat regular polymers of carbohydrates, lignin is a highly heterogeneous polymer of oxygenated aromatic units. While the heterogeneity of lignin provides challenges in valorization, the high aromatic content makes it a promising feedstock, especially as a direct replacement for aromatic compounds normally sourced from petroleum (ex. benzene, toluene, and xylene). Towards lignin valorization, extended efforts on depolymerization [6] and monomer upgrading [7] are ongoing.

Importantly, lignin sourcing does not create conflict with food sources or other processes. Indeed, in biorefining operations, such as pulp and paper manufacturing, lignin is extracted from biomass and burned for process heat and steam. While an integral part of

biorefining process integration, the heating value of the extracted lignin exceeds the process demand, signaling an opportunity for valorization of the excess lignin [5]. Furthermore, as the biorefining industry grows, so will the lignin supply, driving down feedstock costs to help drive down valorization process costs. Finally, a valorization strategy that makes lignin worth more than its heating value may drive a shift from burning lignin, a greenhouse gas producing process, towards upgrading lignin and utilizing renewable energy sources to make up any energy gap.

Another renewable feedstock with metabolic engineering potential is carbon dioxide (CO_2). Naturally fixed by photosynthetic organisms, such as cyanobacteria and algae, CO_2 can be directly converted into compounds of interest, or used to produce sugars destined for microbial upgrading [8]. Towards reduction of greenhouse gas emissions, these processes can be implemented in connection with CO_2 producing processes, taking advantage of increased CO_2 concentration, which promotes higher rates of microbial photosynthesis.

Implementing metabolic engineering strategies using uncommon feedstocks allows for use of biochemical processes in extreme environments to solve unique problems. One such application of metabolic engineering goes towards aiding in human exploration of space [9]. As living organisms, microbes serve as a renewable catalyst, where a small amount of starting material can be multiplied to produce compounds at scale. This characteristic is advantageous in the context of space exploration, where every kilogram and liter of mass and volume comes at a cost. Another advantage here is the modularity of biochemical processes, where a single reactor setup can be used to produce a wide range of compounds from the same inputs simply by switching the inoculated organism.

Researchers are pursuing applications of biotechnology towards pharmaceutical production, astronaut nutrition, and materials production [10].

One significant challenge with microbial utilization of non-sugar feedstocks is potential issues with feedstock toxicity. In the case of lignin-derived monomers, many aromatic species can be toxic to organisms that do not natively metabolize these compounds [11]. Along these same lines, concentrations of fed gasses, such as CO_2 and methane (CH_4), are limited to regimes in which organism growth is uninhibited. Crucially, limits on feedstock concentration could reduce titers and yields, driving up reactor sizes, batch times, and ultimately final product cost. One way to circumvent this problem is to use cell-free systems (CFSs). CFSs are mixture of the enzymes, cofactors, and building blocks of a cell without a living organism [12]. This offers advantages when using toxic feedstocks as concentrations can be kept high without worrying about cell death. While much of the work around chemical production in CFSs has focused on using sugar feedstocks [13], as these systems become more widely used and well-characterized, they can be expanded to use with lignin-derived compounds, CO_2, and CH_4.

In conclusion, metabolic engineering is an excellent strategy for renewable chemical production. The use of non-sugar feedstocks can help lower cost and reduce process risk by decoupling chemical production and crop growth. Towards this, this thesis helps push this field forward in applications of metabolic engineering for lignin valorization, *in situ* resource utilization for exploration of CO_2-rich Mars, and developments in the cell-free systems that will enable greater use of non-sugar feedstocks for production of chemicals.

1.2 References

1. Sarria, S., Kruyer, N.S. & Peralta-Yahya, P. Microbial synthesis of medium-chain chemicals from renewables. *Nat. Biotechnol.* **35,** 1158-1166 (2017).
2. Peralta-Yahya, P., Zhang, F., del Cardayre, S.B. & Keasling, J.D. Microbial engineering for the production of advanced biofuels. *Nature* **488,** 320-328 (2012).
3. Ehrenworth, A.M. & Peralta-Yahya, P. Accelerating the semisynthesis of alkaloid-based drugs through metabolic engineering. *Nat. Chem. Bio.* **13,** 249-258 (2017).
4. Yaguchi, A., Spagnuolo, M. & Blenner, M. Engineering yeast for utilization of alternative feedstocks. *Curr. Opin. Biotechnol.* **53,** 122-129 (2018).
5. Ragauskas, A.J. et al. Lignin valorization: improving lignin processing in the biorefinery. *Science* **344,** 1246843 (2014).
6. Schutyser, W. et al. Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chem. Soc. Rev.* **47,** 852-908 (2018).
7. Becker, J. & Wittmann, C. A field of dreams: lignin valorization into chemicals, materials, fuels, and health-care products. *Biotechnol. Adv.* **37,** 107360 (2019).
8. Angermayr, S.A., Gorchs Rovira, A. & Hellingwerf, K.J. Metabolic engineering of cyanobacteria for the synthesis of commodity products. *Trends Biotechnol.* **33,** 352-361 (2015).
9. Verseux, C. et al. Sustainable life support on Mars – the potential roles of cyanobacteria. *Int. J. Astrobiol.* **15,** 65-92 (2015).
10. Snyder, J.E., Walsh, D., Carr, P.A. & Rothschild, L.J. A makerspace for life support systems in space. *Trends Biotechnol.* **37,** 1164-1174 (2019).

11. Davis, K. & Moon, T.S. Tailoring microbes to upgrade lignin. *Curr. Opin. Chem. Biol.* **59,** 23-29 (2020).

12. Silverman, A.D., Karim, A.S. & Jewett, M.C. Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* **21,** 151-170 (2020).

13. Bogart, J.W. et al. Cell-free exploration of the natural product chemical space. *Chembiochem* **22,** 84-91 (2021).

CHAPTER 2.

METABOLIC ENGINEERING STRATEGIES TO BIO-ADIPIC ACID

Reproduced from:

Kruyer, N.S. and Peralta-Yahya, P. Metabolic engineering strategies to bio-adipic acid. *Curr. Opin. Biotechnol.* **45,** 136-143 (2017).

2.1 Abstract

Adipic acid is the most industrially important dicarboxylic acid as it is a key monomer in the synthesis of nylon. Today, adipic acid is obtained via a chemical process that relies on petrochemical precursors and releases large quantities of greenhouse gases. In the last two years, significant progress has been made in engineering microbes for the production of adipic acid and its immediate precursors, muconic acid and glucaric acid. Not only have the microbial substrates expanded beyond glucose and glycerol to include lignin monomers and hemicellulose components, but the number of microbial chassis now goes further than *Escherichia coli* and *Saccharomyces cerevisiae* to include microbes proficient in aromatic degradation, cellulose secretion and degradation of multiple carbon sources. Here, we review the metabolic engineering and nascent protein engineering strategies undertaken in each of these chassis to convert different feedstocks to adipic, muconic and glucaric acid. We also highlight near term prospects and challenges for each of the metabolic routes discussed.

2.2 Introduction

In 2010, global production of adipic acid was estimated at 2.6 million tons, with 65% going to the production of nylon-6,6 fibers [1,2]. Commercially, adipic acid is synthesized from petroleum-derived benzene, which is reduced to cyclohexane followed by oxidation to a mixture of cyclohexanol and cyclohexanone. This mixture is further oxidized using a vanadium or copper catalyst with nitric acid to produce adipic acid. The commercial synthesis of adipic acid not only generates 10% of the world's man-made nitrous oxide [3], a greenhouse gas 300-times more potent than carbon dioxide, but it also uses benzene, a known carcinogen, as a starting material. Production of bio-adipic acid, wherein the carbons are derived from a renewable feedstock, has the potential to reduce greenhouse gas emissions and eliminate the need for fossil fuel precursors. Based on the 2.6 million tons of adipic acid produced annually from petroleum, renewable production of adipic acid could eliminate the use of benzene found in 6 billion barrels of crude oil.

Advances in metabolic engineering and synthetic biology now allow the engineering of microbes for the production of advanced biofuels [4], commodity chemicals and pharmaceuticals [5]. In the last two years, a number of different metabolic engineering strategies have been applied to the microbial synthesis of adipic acid and its immediate precursors muconic acid and glucaric acid (Figure 2.1, Table 2.1). In particular, there has been an expansion in the renewable feedstocks used as starting materials, as well as an increase in the number of metabolic pathways engineered for the microbial synthesis of these compounds. The number of chassis has also risen and now goes beyond the workhorse chassis of *Escherichia coli* and *Saccharomyces cerevisiae*, to include the natural aromatic compound metabolizer *Pseudomonas putida* [6], the cellulolytic bacteria

Thermobifida fusca [7], and *Klebsiella pneumonia* [8], which can metabolize a number of different carbon sources. In this review, we focus on these latest advances and direct the reader to recent comprehensive reviews on adipic acid and muconic acid production [2,9,10].

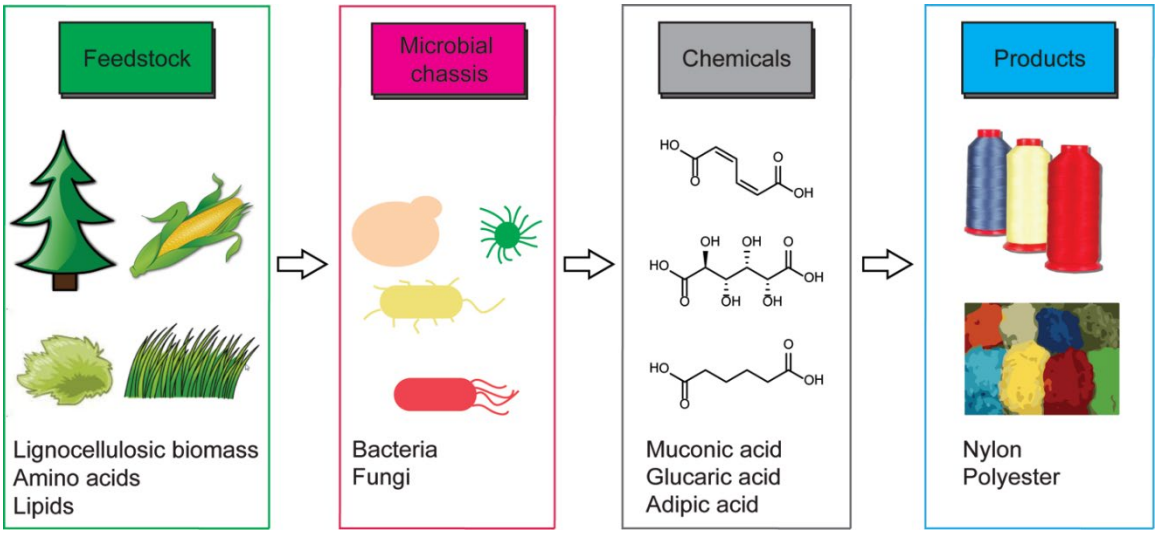

Figure 2.1. Bio-adipic acid production. In the last two years an increasing number of feedstocks and microbial chassis have been used for the production of adipic acid and its immediate precursors, muconic acid and glucaric acid. Adipic acid is one of the most industrially important dicarboxylic acids being used in the synthesis of nylon and polyesters.

Table 2.1. Yields for muconic acid, glucaric acid and adipic acid from glucose

Chassis	Pathway	Maximum Theoretical Yields [49]	Experimental percent of maximum theoretical yield[h, i]
E. coli	**DHS** → PCA → catechol → muconic acid	83%[a]	27% Niu *et al.* [25][j] 7.81% Zhang *et al.* [48][k]
E. coli	**Anthranilate** → catechol → muconic acid	50%	2.24% Sun *et al.* [29][k]
E. coli	**PHB** → PCA → catechol → muconic acid	52%	4.48% Sengupta *et al.* [17][k]
E. coli	**Isochorismate** → salicylate → catechol → muconic acid	53%	16.3% Lin *et al.* [15][k] 37.0% Noda *et al.* [16][k]
E. coli	**2,3-Dihydroxybenzoate** → catechol → muconic acid	67%	Wang and Zheng [13][l]
E. coli	**Glucose** → glucaric acid	99%[b]	9.79% Moon *et al.* [19][k]
E. coli	**Fatty acids** → adipic acid	NC[c]	Clomburg *et al.* [22]
E. coli	**AcCoA + SuccCoA** → adipyl-CoA → adipic acid from glycerol	53%[d]	6.6% Cheong *et al.* [23][k]
E. coli	**AcCoA + SuccCoA** → adipyl-CoA → adipyl-phosphate → adipic acid	92%[e]	<1% Yu *et al.* [44][k]
P. putida	Muconic acid from *p*-coumaric acid	76%	88% Vardon *et al.* [18][j]
P. putida	Muconic acid from benzoate	40%[f]	93% Vardon *et al.* [18][m]
S. cerevisiae	**DHS** → PCA → catechol → muconic acid	61%[g]	<1% Curran *et al.* [27][k] <1% Weber *et al.* [26][k] 2.9% Suastegui *et al.* [28][k]

For all *E. coli* and *S. cerevisiae* calculations, no reactions were removed from the model. For all *P. putida* calculations, PCA dioxygenase was replaced with PCA decarboxylase to allow for catechol synthesis; the muconic acid to β-ketoadipate conversion was not removed. Maximum theoretical yields from glucose unless specified. Genome scale models used for *E. coli, P. putida*, *and S. cerevisiae* were iAF1260b [50], iJN746 [51], iMM904 [52], respectively. Bold: last endogenous chemical in the model.

a) Zhang *et al*. reported a theoretical maximum molar yield of 86% based on stoichiometry [48].
b) The maximum theoretical yield (MTY) calculations did not account for cell growth. Practically, MTY will be lower as glucose will also need to be routed through glycolysis for cell growth.
c) NC: Not calculated. Without removing reactions from the *E.coli* model, hexanoic acid is produced via fatty acid biosynthesis.
d) All enzymes in this pathway are used in the reverse direction. The model does not take into account the kinetics of this process.

e) One enzyme in this pathway is used in the reverse direction and the rest have specificity for C4 rather than C6 substrates. The model does not take these facts into account.
f) Muconic acid to β-ketoadipate reaction could not be removed from the model and significant carbon is used for growth.
g) Predicted MTY from glucose matches that from Curran *et al.* [27].
h) Experimenters use a variety of glucose and glycerol concentrations in both complex and minimal medias
i) Percent of theoretical yield calculated as $\frac{raw\ percent\ yield}{calculated\ maximum\ theoretical\ yield}$
j) Raw percent yield reported by the author, no details for carrying out independent calculations provided
k) Raw percent yield calculated as $\frac{titer\ of\ product\ (muconic\ or\ adipic\ acid)\ in\ moles}{total\ moles\ of\ fed\ carbon\ source\ (glucose\ and\ glycerol)}$
l) No details for carrying out independent calculations provided
m) Due to limitations in the modeling of adipic acid production from benzoate (see f), reported yield is the raw percent yield reported by the author

2.3 Metabolic pathways to muconic, glucaric, and adipic acid

A total of ten biosynthetic pathways from lignocellulosic biomass, lipids and amino acids have been leveraged for the production of adipic acid, muconic acid and glucaric acid, with pathways utilizing lignin, lipids and amino acids as precursors demonstrated since 2014 (Figure 2.2). In 1994, Frost and Draths developed the first production of bio-adipic acid by exploiting *E. coli* primary metabolism to convert glucose to 3-dehydroshikimic acid (DHS), and using three heterologous enzymes to produce protocatechuic acid (PCA), catechol and ultimately *cis, cis*-muconic acid, which was in turn chemically hydrogenated to adipic acid [11]. More recent pathways to muconic acid include re-routing chorismate to anthranilic acid [12], 2,3-dihydroxybenzoic acid [13,14], salicylic acid [15,16], or *p*-hydroxybenzoic acid [17] to produce catechol, and a β-ketoadipate pathway modified to stop at PCA followed by a re-route to catechol [18]. Glucaric acid has also been microbially produced from glucose via a synthetic five-step

pathway [19]. Finally, adipic acid has been microbially synthesized directly from renewables by taking advantage of (1) the reversed adipate degradation pathway [20,21], (2) the reversed β-oxidation coupled to alkane degradation [22], and (3) the reversed phenylacetate degradation pathway coupled to reversed adipic acid degradation [23].

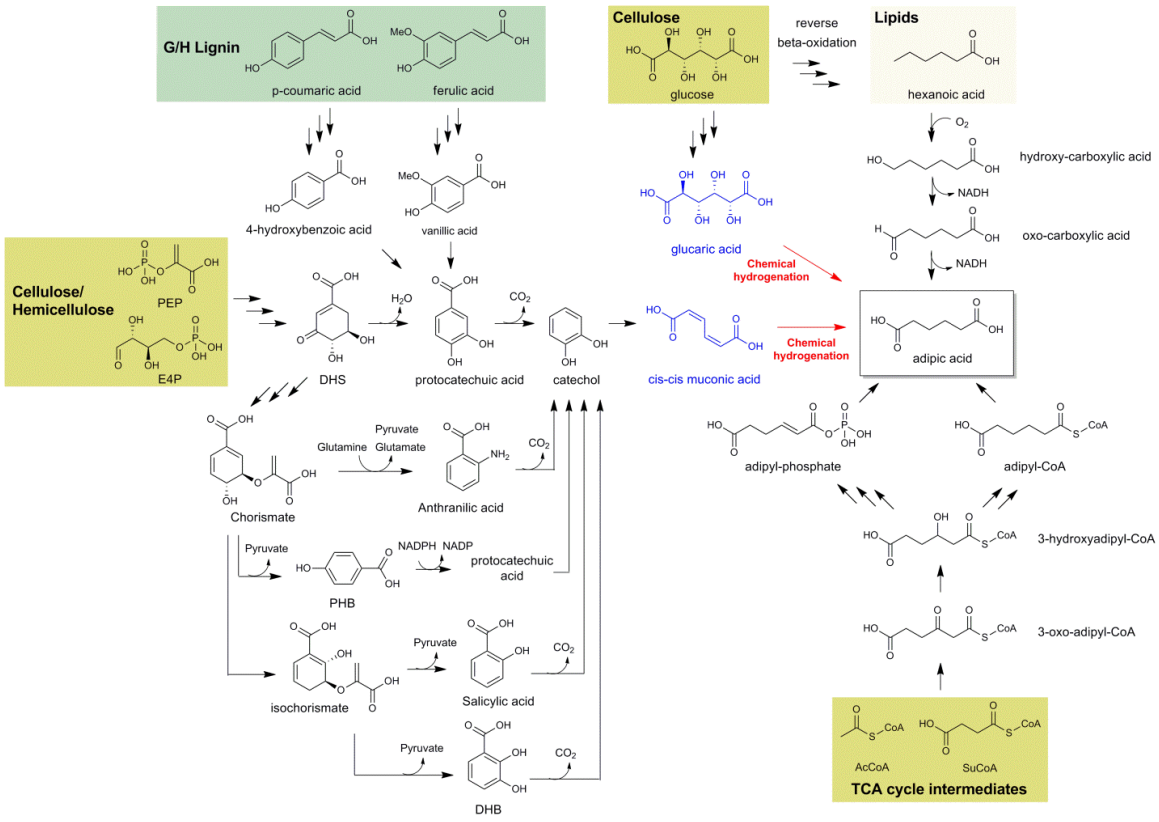

Figure 2.2. Metabolic engineering strategies to bio-adipic acid production. Shown, metabolic pathways highlighted in this review. Green box: lignin-derived substrates. Yellow box: cellulose- and hemicellulose-derived substrates. Grey box: lipid-derived substrates. Adipic acid, muconic acid and glucaric acid in blue. DHS: 3-dehydroshikimic acid; PHB: *p*-hydroxybenzoic acid; DHB: 2,3-dihydroxybenzoic acid.

2.4 Muconic acid from cellulose- and hemicellulose-derived feedstocks

Microbes can convert cellulose-derived glucose or hemicellulose-derived xylose to phosphoenolpyruvate (PEP) via glycolysis and erythrose-4-phosphate (E4P) via the pentose phosphate pathway. Aromatic amino acid biosynthesis takes PEP and E4P and eventually converts them to DHS, a key node in the biosynthesis of muconic acid. The concept of leveraging aromatic amino acid biosynthesis for muconic acid production is rooted in the successful engineering of *E. coli* for the production of aromatic amino acids at g/L levels [24].

In *E. coli*, DHS has been converted to muconic acid via PCA and catechol resulting in 36.8 g/L of muconic acid with a 22% yield from glucose [25]. Key to achieving this yield was the introduction of a feedback resistant 3-deoxy-D-arabino-heptulosonic acid 7-phosphate (DAHP) synthase, which condenses PEP and E4P in the first committed step of muconic acid biosynthesis. Of note, in this strain, the conversion of DHS to aromatic amino acids was blocked and cell growth required addition of aromatic amino acids and vitamins to the media. The same DHS to muconic acid pathway has been introduced in *S. cerevisiae* to produce 1.56 mg/L of muconic acid [26], and, in 2013, 141 mg/L of muconic acid with a 0.8% yield from glucose [27]. Interestingly, the low mg/L titers were obtained despite successful implementation of similar metabolic engineering strategies as in *E. coli*, including relieving feedback inhibition of DAHP synthase, and deletion of glucose-6-phosphate dehydrogenase to force the carbon flux from glycolysis to the pentose phosphate pathway via the transketolase reaction, which is kinetically hindered *in vivo*, rather than glucose-6-phosphate dehydrogenase. More recently, the DHS to muconic acid pathway was introduced into a diploid *S. cerevisiae* strain carrying similar up- and down-

regulations to increase the carbon flux to the shikimic acid pathway. Additionally, the diploid carried a 3-dehydroquinate dehydratase mutant (Aro:D1409A) that stopped conversion at DHS, reducing synthesis of the byproduct shikimate. When this strain was grown under anaerobic conditions to improve the activity of the oxygen sensitive protocatechuate decarboxylase, it produced 560 mg/L of muconic acid [28]. The *S. cerevisia*e results underscore the extensive regulation of aromatic amino acid biosynthesis in yeast, where limited pathway modifications do not result in the g/L titers seen in *E. coli*. Improving muconic acid yields in *S. cerevisiae* will likely require global regulation of the transcription machinery, which will require a muconic acid high-throughput assay for identification.

Aromatic amino acid biosynthesis can take DHS to chorismate, which *E. coli* can convert to anthranilic acid or *p*-hydroxybenzoic acid (PHB), intermediates that can enter muconic acid biosynthesis at the level of catechol or PCA, respectively. *E. coli* production of muconic acid via anthranilic acid results in 390 mg/L of muconic acid from a glycerol/glucose mixture [29]. Crucial to this strategy was screening for an efficient heterologous anthranilate-1,2-dioxygenase to convert anthranilate to catechol, increasing the levels of chorismate, deleting the tryptophan biosynthetic pathway after anthranilate, and introducing a glutamine regeneration system as glutamine is depleted upon conversion of chorismate to anthranilic acid. Due to deletion of the tryptophan biosynthetic pathway, supplementation of the media with tryptophan was needed for cell growth. Production of muconic acid in *E. coli* via PHB resulted in 170 mg/L of muconic acid from glucose [10]. Keys to this approach were, screening for an efficient heterologous PHB hydrolase to convert PHB to PCA, increasing PHB availability, knocking out genes predicted to

increase carbon flux *in silico* by Robust-Knock [30], and overexpressing a transketolase to route carbon flux from glycolysis to pentose phosphate pathway. A drawback of this strategy is the use of NADPH by PHB hydrolase, potentially resulting in a redox imbalanced strain.

E. coli converts chorismate to isochorismate and subsequently salicylic acid via the siderophore biosynthetic pathway, or converts chorismate to isochorismate and subsequently 2,3-dihydrobenzoic acid (DHB) via the enterobactin biosynthetic pathway. Both salicylic acid and DHB enter muconic acid biosynthesis at the level of catechol. Production of muconic acid in *E. coli* via salicylic acid results in 1.5 g/L of muconic acid from a glycerol/glucose mixture [15]. This work relied on previously identified efficient isochorismate synthase and isochrorismate pyruvate lyase [31] to produce salicylic acid, and a phenylalanine overproducing strain which could not synthesize phenylalanine or tyrosine due to deletions after chorismate. Further screening for an efficient heterologous salicylate-1-monooxygenase, and reducing the plasmid copy number for the expression of salicylate-1-monooxygease and catechol-1,2-dioxygenase resulted in the aforementioned yield. Due to deletions in the phenylalanine and tyrosine pathway, the production media needed to be supplemented with yeast extract for cell growth. *E. coli* production of muconic acid via DHB resulted in 480 mg/L of muconic acid from a glycerol/ glucose mixture [14], and more recently in 605 mg/L of muconic acid (carbon source not specified) [13]. The earlier work identified a novel prokaryotic *Klebsiella pneumoniae* DHB decarboxylase to convert DHB into catechol, and upregulated DHB synthesis to achieve the mentioned yield. The latter work took advantage of the *K. pneumoniae* DHB decarboxylase and increased the extent of DHB synthesis upregulation.

2.5 Muconic acid from lignin-derived feedstock

Lignin-derived monomers such as *p*-coumaric acid, coniferyl alcohol and caffeic acid can serve as substrates to produce muconic acid via the PCA branch of a modified β-ketoadipate pathway. In the β -ketoadipate pathway PCA is ultimately converted to acetyl-CoA and succinyl-CoA. Stopping the β-ketoadipate pathway at PCA and re-routing PCA to catechol leads to the production of muconic acid. Deleting the transformation of muconic acid to β-ketoadipate results in muconic acid accumulation. The β-ketoadipate pathway is not present in *E. coli* or *S. cerevisiae*, but it is endogenous of aromatic degrading organisms, such as *P. putida*. Since 2015, *P. putida* has been extensively engineered to produce muconic acid not just from lignin-derived monomers, but also directly from pretreated biomass.

In 2015, *P. putida* was genomically engineered to route guaiacyl (G) units, derived from coniferyl alcohol, and *p*-hydroxyphenyl (H) units, derived from *p*-coumaric acid monomers, to convert G/H lignin into muconic acid [18]. Specifically, PCA dioxygenase was replaced with PCA decarboxylase to convert PCA into catechol, and catB/catC were deleted to avoid degradation of muconic acid. This strain achieved 67% conversion of *p*-coumaric acid, 14% conversion of coniferyl alcohol, and 28% conversion of caffeic acid into muconic acid. It is worth highlighting that muconic acid is not the final product, but the immediate intermediate to adipic acid. Thus, any muconic acid production must be followed by a separation of muconic acid from the culture broth, purification and hydrogenation to produce the final adipic acid. The 2015 work in *P. putida* also tackled the downstream process of converting muconic acid into adipic acid. Specifically, by taking advantage of the different physical properties of the aromatic intermediates and the linear

muconic acid, aromatics were removed from the culture media using activated carbon. Muconic acid was then crystalized from the media by lowering the pH and temperature resulting in the recovery of 74% of muconic acid with >97% purity. Crystallization of muconic acid from the media reduced the number of inhibitors from the biological process that could inactivate the catalyst, and hydrogenation of muconic acid over palladium and carbon resulted >97% conversion of muconic acid with 1% Pd/C loading after 35 min. Finally, alkaline pretreated biomass was converted to muconic acid with 67% molar yield. The most recent titers for muconic acid using benzoate as the substrate and glucose to support cell growth are 34.5 g/L [32]. Follow up work has focused on improving the chemical hydrogenation process, making nylon-6,6 using bio-adipic acid [33], and increasing the activity of PCA decarboxylase [34], a bottleneck in the conversion of PCA to muconic acid. Remaining challenges in the conversion of lignin to muconic acid include improving the enzymatic activity or removing the endogenous regulation of vanillate-O-methylase, which converts vanillate from G-lignin substrates into PCA. Though this challenge can be circumvented by using a harsher biomass pretreatment processes that brings G/H-lignin substrates to phenol and guaiacol, which can be converted to catechol in a single step.

2.6 Key enzymes in microbial synthesis of muconic acid

Independent of the carbon source, the microbial synthesis of muconic acid depends on the activity of catechol-1,2-dioxygenase and, for pathways routing via PCA, PCA decarboxylase. Recently, a catechol-1,2-dioxygenase double mutant (Cat:L73F/P76A) was engineered using rational design to achieve a five-fold increase in turnover rate with an

equivalent increase in K_M resulting in a comparable catalytic efficiency as the wild type enzyme [35]. Cat:L73F/P76A also showed improved expression when compared to wild type. When Cat:L73F/P76A was used in the microbial synthesis of muconic acid from glucose, it increased titers 22%. In separate work, catechol-1,2-dioxygenase was purposefully engineered for improved expression via codon optimization using a stationary phase codon optimization matrix compiled from the fifty most highly expressed proteins in *S. cerevisiae* after 3 days of growth [36]. The codon optimized catechol-1,2-dioxygenase had a two-fold higher conversion of catechol to muconic acid. To increase the *in vivo* PCA decarboxylase activity 14-fold over wild type in *E. coli*, *K. pneumonia* PCA decarboxylase was co-expressed with an upstream gene found in the same operon (KpdB) [37]. This strategy to increase PCA decarboxylase is chassis agnostic as it has also been successfully applied in *P. putida* [34].

2.7 Glucaric acid from cellulose-derived feedstock

Glucaric acid is one of the top value-added chemicals from biomass identified by the US Department of Energy in 2004 [38], and can be reduced to adipic acid. Mammals synthesize glucaric acid from glucose in 10-steps; however, a 5 step synthetic pathway has been engineered in *E. coli* to route glucose via myo-inositol to produce 1 g/L of glucaric acid [19]. A bottleneck in the synthetic pathway is the conversion of myo-inositol to D-glucuronic acid by myo-inositol oxygenase (MIOX), the activity of which drops significantly after 24 hours. Using protein scaffolds to increase the myo-inositol concentration around MIOX increased MIOX specific activity and brought glucaric acid titers to 2.5 g/L [39]. Improved conversion of myo-inositol to glucaric acid has been

achieved by increasing the solubility of MIOX using solubility tags, yielding an enzyme that retains almost 40% of its activity after 72 hours [40]. More recently, the use of a dynamic pathway regulator based on quorum sensing circuitry allowed for control of the timing of up- and down-regulation of the glycolysis flux control enzyme phosphofructokinase to route carbon from glycolysis to glucaric acid production. Using this strategy, a glucaric acid production of 0.7 g/L was achieved without the need for inducers [41].

2.8 Adipic acid production from lipids

Lipids, in the form of fatty acids, can be converted directly to adipic acid. Bacteria and some yeast degrade long chain fatty acids and alkanes via oxidation of the ω-carbon. Although biological systems produce long-chain fatty acids, microbes can be modified to produce medium-chain fatty acids, most successfully via reversal of the β-oxidation cycle. By bringing together reversed β-oxidation and the ω-oxidation pathway, *E. coli* has been engineered to produce 170 mg/L of adipic acid from glycerol [22]. Key to this strategy has been the introduction of a thiolase that preferentially condenses acetyl-CoA and C4/C6 acyl-CoAs and a thioesterase that favors hydrolysis of medium-chain fatty acyl-CoAs. With C6-C10 fatty acids in hand, introduction of *P. putida* ω-hydroxylase, which enables *P. putida* growth on C6-C16 alkanes, allowed the production of ω-hydroxy-carboxylic acids. Despite evidence of *P. putida* ω-hydroxylase multi-oxidizing fatty acid methyl esters to dicarboxylic acids [42], adipic acid production required expression of NAD^+-dependent alcohol and aldehyde dehydrogenases [22]. Insight for the choice of these dehydrogenases came from the degradation of cyclohexanol in *Acinetobacter*, which pointed at their

activity on medium-chain ω-hydroxyacids [43]. Among the challenges to increasing adipic acid production via this strategy is the strain redox balance as two NADHs are produced per adipic acid generated. Further, as fatty acid hydroxylation requires molecular oxygen, using microaerobic conditions for cell growth to take advantage of oxidative phosphorylation to convert NADH into ATP is not an efficient strategy. Likely, an NAD^+ recycling system or a hydroxylase that requires a metal rather than NAD^+ or $NADP^+$ to carry out the oxidation will be needed to further increase adipic acid titers.

2.9 Adipic acid production from tricarboxylic acid cycle intermediates

Glucose, xylose, lipids, and amino acids can be metabolized to the tricarboxylic acid (TCA) cycle intermediates acetyl-CoA (AcCoA) and succinyl-CoA (SuCoA). Inspired by the *Penicillium chrysogenum* adipic acid degradation pathway and the broad substrate specificity of *Clostridium acetobutylicum* butanol synthesis enzymes, a six step synthetic pathway to adipic acid from AcCoA and SuCoA has been engineered [44]. For this pathway, a thiolase which hydrolyzes 3-oxo-adipyl-CoA into AcCoA and SuCoA was used in reverse to generate 3-oxo-adipyl-CoA. Next followed reduction, dehydration, and hydrogenation of 3-oxo-adipyl-CoA by enzymes previously shown to convert 3-ketohexanoyl-CoA into hexanoyl-CoA [45] to generate adipyl-CoA. Finally, adipyl-CoA was converted into adipyl-phosphate and ultimately into adipic acid using *C. acetobutylicum* enzymes with broad substrate specificity known to convert butyryl-CoA into butyrate. Introduction of the synthetic pathway in a succinic acid overproducing *E. coli* strain resulted in 639 mg/L of adipic acid from glucose. More recently, a second synthetic pathway to produce adipic acid from AcCoA and SuCoA in *E. coli* was

engineered [23]. This second synthetic pathway combines reversal of both the phenylacetate degradation pathway and the adipic acid degradation pathway in a mixed-acid fermentation-deficient *E. coli* that retained the reductive branch of the TCA cycle and overexpressed a CoA transferase for increased SuCoA availability to produce 2.5 g/L of adipic acid from glycerol in a bioreactor. Both adipic acid synthetic pathways rely on at least one enzyme used in reversed direction. For the butanol/butyrate inspired work, improving the thiolase reaction in the reversed direction, and optimizing the substrate specificity of the butanol/butyrate pathway enzymes will aid in increasing adipic acid titers. An eventual major concern will be the redox balance of the strain as 2 NADHs are needed per adipic acid produced, a similar problem to that faced in the production of butanol via the *C. acetobutylicum* pathway [46]. For the phenylacetate degradation inspired pathway, all the enzymes are used in the reversed direction and concerns about the thermodynamic and kinetic feasibility of this process need to be addressed.

2.10 Alternative chassis and co-cultures to produce adipic acid and its precursors

To date, production of adipic acid and its precursors has predominantly taken place in *E. coli*, *S. cerevisiae* and *P. putida*. Recently, alternative microbial chassis have been explored, including *T. fusca*, which has a reverse adipate degradation pathway, and *K. pneumoniae*, which has a catechol synthesis pathway. The pivotal work in *T. fusca* has been the identification of the reductase that converts 2,3-dehydroadipyl-CoA to adipyl-CoA as the flux control point in adipic acid biosynthesis [17]. Using *T. fusca* B6, in which the reductase is highly upregulated, resulted in 2.2 g/L of adipic acid from glucose. In *K. pneumoniae*, disruption of the aromatic amino acid biosynthetic pathway, introduction of

P. putida catechol-1,2-dioxygenase, and deletion of endogenous enzymes downstream of muconic acid resulted in 2.1 g/L of muconic acid from glucose [19.]. Co-cultures from the same species have also been explored for the production of muconic acid. Specifically, production of muconic acid from simple sugars via DHS, PCA and catechol has been split into two different *E. coli* strains, the first specialized in DHS production, the second specialized in DHS to muconic acid conversion [47,48]. Key to this strategy was DHS diffusion into the media from the first strain and using a DHS transporter to increase the DHS concentration in the second strain. This co-culture resulted in 2 g/L of muconic acid from glycerol. Introduction of all the enzymes in the same *E. coli* strain resulted in poor cell growth and significantly reduced muconic acid production. Further, by modifying the *E. coli* engineered to produce DHS to only utilize xylose and the *E. coli* engineered to convert DHS into muconic acid to only utilize glucose, the co-culture strategy circumvented the carbon catabolite repression and after media and reactor optimization, the co-culture system produced 4.7 g/L of muconic acid from a glucose/xylose mixture.

2.11 Conclusions

Production of bio-adipic acid has the potential to eliminate the reliance on petroleum precursors and reduce greenhouse emission from the chemical synthesis of adipic acid. Among the metabolic routes to produce muconic acid from cellulosic and hemicellulosic feedstock, the original pathway that routes DHS directly to PCA, catechol and ultimately muconic acid shows the highest theoretical yield (83%) and still has the highest muconic acid titers shown to date (36.8 g/L). However, if pre-treated lignin or lignin monomers are the desired substrate, *P. putida* with its endogenous β-ketoadipate

pathway is the most promising chassis with 67% molar yield from *p*-coumarate and ferulate present in pretreated lignin. There is currently a single pathway engineered for the microbial synthesis of glucaric acid, and while this pathway has potentially a 99% theoretical yield from glucose, the great pull on glucose from glycolysis has resulted in a maximum titer of 2.5 g/L of glucaric acid or 9.8% of theoretical yield. Both muconic and glucaric acid require purification from the microbial broth and chemical hydrogenation to be converted to adipic acid and an exciting solution has been the development of synthetic pathways and the discovery of natural pathways for the direct microbial synthesis of adipic acid. Although the yields for the adipic acid pathways are currently relatively low, with the best adipic acid titer being 2.5 g/L, the removal of the chemical hydrogenation step has the potential to eliminate the need for precursor (muconic acid, glucaric acid) purification and a second reactor for chemical hydrogenation. Thus far, there has been limited engineering of enzymes in the muconic acid and adipic acid pathways. Increasing the enzymatic activity and/or disrupting the regulation of enzymes in the PCA branch of the β-ketoadipate pathway, in particular vanillate-O-methylase, PCA decarboxylase and catechol-1,2-dioxygenase has the potential to increase muconic acid yields from lignin and cellulose/hemicellulose-derived substrates. With respect to adipic acid synthesis, engineering the substrate specificity of the ω-oxidation enzymes and reducing the dependence of the ω-oxidation enzymes on NAD^+ and molecular oxygen to carry out oxidations, will be necessary to improve yields. Additionally, increasing the activity of phenylacetate degradation enzymes in the reverse direction has the potential to increase adipic acid yields.

2.12 References

1. Merchant Reserach & Consulting, L. Adipic acid 2011 World Market Outlook and Forecast (2011).

2. Polen T, Spelberg M, Bott M: **Toward biotechnological production of adipic acid and precursors from biorenewables**. *J Biotechnol* 2013, **167**:75-84.

3. Alini S, Basile F, Blasioli S, Rinaldi C, Vaccari A: **Development of new catalysts for N_2O-decomposition from adipic acid plant**. *Appl Catal B-Environ* 2007, **70**:323-329.

4. Peralta-Yahya P, Zhang FZ, del Cardayre SB, Keasling J. D. **Microbial engineering for the production of advanced biofuels.** *Nature* 2012, **488**: 320-328.

5. Nielsen, J. & Keasling, J. D. **Engineering Cellular Metabolism.** *Cell* 2016, **164**: 1185-1197.

6. Thompson B, Machas M, Nielsen DR: **Creating pathways towards aromatic building blocks and fine chemicals**. *Curr Opin Biotechnol* 2015, **36**:1-7.

7. Bugg TDH, Rahmanpour R: **Enzymatic conversion of lignin into renewable chemicals**. *Curr Opin Chem Biol* 2015, **29**:10-17.

8. Jung HM, Jung MY, Oh MK: **Metabolic engineering of Klebsiella pneumoniae for the production of muconic acid**. *Appl Microbiol Biotechnol* 2015, **99**:5217-5225.

9. Deng Y, Ma L, Mao Y: **Biological production of adipic acid from renewable substrates: Current and future methods**. *Biochem Eng J* 2016, **105**:16-26.

10. Xie NZ, Liang H, Huang RB, Xu P: **Biotechnological production of muconic acid: current status and future prospects**. *Biotechnol Adv* 2014, **32**:615-622.

11. Draths, K. M. & Frost, J. W: **Environmentally Compatible Synthesis of Adipic Acid from D-Glucose.** *J Am Chem Soc* 1994, **116**: 399-400.

12. Sprenger GA: **From scratch to value: engineering Escherichia coli wild type cells to the production of L-phenylalanine and other fine chemicals derived from chorismate**. *Appl Microbiol Biotechnol* 2007, **75**:739-749.

13. Wang J, Zheng P: **Muconic acid production from glucose using enterobactin precursors in *Escherichia coli***. *J Ind Microbiol Biotechnol* 2015, **42**:701-709.

14. Sun X, Lin Y, Yuan Q, Yan Y: **Biological Production of Muconic Acid via a Prokaryotic 2,3-Dihydroxybenzoic Acid Decarboxylase**. *Chemsuschem* 2014, **7**:2478-2481.

15. Lin Y, Sun X, Yuan Q, Yan Y: **Extending shikimate pathway for the production of muconic acid and its precursor salicylic acid in Escherichia coli**. *Metab Eng* 2014, **23**:62-69.

16. Noda S, Shirai T, Oyama S, Kondo A: **Metabolic design of a platform Escherichia coli strain producing various chorismate derivatives**. *Metab Eng* 2016, **33**:119-129.

17. Sengupta S, Jonnalagadda S, Goonewardena L, Juturu V: **Metabolic engineering of a novel muconic acid biosynthesis pathway via 4-hydroxybenzoic acid in Escherichia coli**. *Appl Environ Microbiol* 2015, **81**:8037-8043.

18. Vardon DR, Franden MA, Johnson CW, Karp EM, Guarnieri MT, Linger JG, Salm MJ, Strathmann TJ, Beckham GT: **Adipic acid production from lignin**. *Energy Environ Sci* 2015, **8**:617-628.

19. Moon TS, Yoon SH, Lanza AM, Roy-Mayhew,JD, Prather KLJ: **Production of Glucaric Acid from a Synthetic Pathway in Recombinant Escherichia coli**. *Appl Environ Microbiol* 2009, **75**:4660-4660.

20. Yu JL, Xia XX, Zhong JJ, Qian ZG: **Direct biosynthesis of adipic acid from a synthetic pathway in recombinant Escherichia coli**. *Biotechnol Bioeng* 2014, **111**:2580-2586.

21. Deng Y, Mao Y: **Production of adipic acid by the native-occurring pathway in Thermobifida fusca B6**. *J Appl Microbiol* 2015, **119**:1057-1063.

22. Clomburg JM, Blankschien MD, Vick JE, Chou A, Kim S, Gonzalez R: **Integrated engineering of beta-oxidation reversal and omega-oxidation pathways for the synthesis of medium chain omega-functionalized carboxylic acids**. *Metab Eng* 2015, **28**:202-212.

23. Cheong S, Clomburg J, Gonzalez R: **Energy- and carbon-efficient synthesis of functionalized small molecules in bacteria using non-decarboxylative Claisen condensation reactions**. *Nat Biotechnol* 2016, **34**: 556-561.

24. Rodriguez A, Martinez JA, Flores N, Escalante A, Gosset G, Bolivar F: **Engineering Escherichia coli to overproduce aromatic amino acids and derived compounds**. *Microb Cell Fact* 2014, **13**:126.

25. Niu W, Draths KM, Frost JW: **Benzene-free synthesis of adipic acid**. *Biotechnol Progr* 2002, **18**:201-211.

26. Weber C, Bruckner C, Weinreb S, Lehr C, Essl C, Boles E: **Biosynthesis of cis,cis-muconic acid and its aromatic precursors, catechol and protocatechuic acid, from renewable feedstocks by Saccharomyces cerevisiae**. *Appl Environ Microbiol* 2012, **78**:8421-8430.

27. Curran, K. A., Leavitt, J. M., Karim, A. S., Alper, H. S: **Metabolic engineering of muconic acid production in Saccharomyces cerevisiae.** *Metab Eng* 2013, **15**: 55-66.

28. Suastegui, M., Matthiesen, J. E., Carraher, J. M., Hernandez, N., Quiroz, N. R., Okerlund, A., Cochran, E. W., Shao, Z. Y. & Tessonnier, J. P. **Combining Metabolic Engineering and Electrocatalysis: Application to the Production of Polyamides from Sugar**. *Angew Chem Int Edit* 2016, **55**: 2368-2373.

29. Sun X, Lin Y, Qin H, Qipeng Y, Yan Y: **A Novel Muconic Acid Biosynthesis Approach by Shunting Tryptophan Biosynthesis via Anthranilate**. *Appl Environ Microbiol* 2013, **79**:4024-4030.

30. Tepper N, Shlomi T: **Predicting metabolic engineering knockout strategies for chemical production: accounting for competing pathways**. *Bioinformatics* 2010, **26**:536-543.

31. Lin YH, Shen XL, Yuan QP, Yan YJ: **Microbial biosynthesis of the anticoagulant precursor 4-hydroxycoumarin**. *Nat Commun* 2013, **4**:2603.

32. Linger JG, Vardon DR, Guarnieri MT, Karp EM, Hunsinger GB, Franden MA, Johnson CW, Chupka G, Strathmann TJ, Pienkos PT et al: **Lignin valorization through integrated biological funneling and chemical catalysis**. *Proc Natl Acad Sci USA* 2014, **111**:12013-12018.

33. van Duuren JBJH, Wijte D, Karge B, Martins dos Santos VAP, Yang Y, Mars AE, Eggink G: **pH-stat fed-batch process to enhance the production of cis, cis-muconate from benzoate by Pseudomonas putida KT2440-JD1**. *Biotechnol Progr* 2012, **28**:85-92.

34. Johnson, C. W., Salvachúa, D., Khanna, P., Smith, H., Peterson, D. J. & Beckham, G. T: E**nhancing muconic acid production from glucose and lignin-derived aromatic compounds via increased protocatechuate decarboxylase activity.** *Metab Eng Commun* 2016 **3:**111-119.

35. Han, L., Liu, P., Sun, J. X., Wu, Y. Q., Zhang, Y. Y., Chen, W. J., Lin, J. P., Wang, Q. H. & Ma, Y. H: **Engineering catechol 1,2-dioxygenase by design for improving the performance of the cis, cis-muconic acid synthetic pathway in Escherichia coli.** *Sci Rep* 2015, **5**: 11.

36. Lanza, A. M., Curran, K. A., Rey, L. G. & Alper, H. S: **A condition-specific codon optimization approach for improved heterologous gene expression in Saccharomyces cerevisiae.** *Bmc Syst Biol* 2014, 8:Artn 33.

37. Sonoki T, Morooka M, Sakamoto K, Otsuka Y, Nakamura M, Jellison J, Goodell B: **Enhancement of protocatechuate decarboxylase activity for the effective production of muconate from lignin-related aromatic compounds**. *J Biotechnol* 2014, **192**:71-77.

38. Werpy T, Peterse G. In: N.R.E.L. and P.N.N.L. **Top Value added chemicals from biomass**, vol I: Results of screenign for potential cadidates from sugars and syntehsis gas, 2004.

39. Moon TS, Dueber JE, Shiue E, Prather KLJ: **Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered E. coli**. *Metab Eng* 2010, **12**:298-305.

40. Shiue E, Prather KLJ: **Improving D-glucaric acid production from myo-inositol in E. coli by increasing MIOX stability and myo-inositol transport**. *Metab Eng* 2014, **22**:22-31.

41. Gupta, A., Reizman, I. M., Reisch, C. R. & Prather, K. L. **Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit.** Nature Biotechnol 2017, 35:273-279.

42. Schrewe, M., Julsing, M. K., Lange, K., Czarnotta, E., Schmid, A. & Buhler, B. **Reaction and Catalyst Engineering to Exploit Kinetically Controlled Whole-Cell Multistep Biocatalysis for Terminal FAME Oxyfunctionalization.** *Biotechnol Bioeng* 2014, **111**: 1820-1830.

43. Cheng Q, Thomas SM, Kostichka K, Valentine JR, Nagarajan V: **Genetic analysis of a gene cluster for cyclohexanol oxidation in Acinetobacter sp. Strain SE19 by in vitro transposition**. *J Bacteriol* 2000, **182**:4744-4751.

44. Yu JL, Xia XX, Zhong JJ, Qian ZG: **Direct Biosynthesis of Adipic Acid From a Synthetic Pathway in Recombinant Escherichia Coli**. *Biotechnol Bioeng* 2014, **111**:2580-2586.

45. Dekishima Y, Lan EI, Shen CR, Cho KM, Liao JC: **Extending Carbon Chain Length of 1-Butanol Pathway for 1-Hexanol Synthesis from Glucose by Engineered Escherichia coli**. *J Am Chem Soc* 2011, **133**:11399-11401.

46. Shen, C. R., Lan, E. I., Dekishima, Y., Baez, A., Cho, K. M. & Liao, J. C. **Driving Forces Enable High-Titer Anaerobic 1-Butanol Synthesis in Escherichia coli.** *Appl Environ Microb* 2011, **77**: 2905-2915.

47. Zhang H, Li Z, Pereira B, Stephanopoulos G: **Engineering E. coli-E. coli cocultures for production of muconic acid from glycerol**. *Microb Cell Fact* 2015, **14**:134.

48. Zhang HR, Pereira B, Li ZJ, Stephanopoulos G: **Engineering Escherichia coli coculture systems for the production of biochemical products**. *Proc Natl Acad Sci USA* 2015, **112**:8266-8271

49. Schellenberger, J., Que, R., Fleming, R. M. T., Thiele, I., Orth, J. D., Feist, A. M., Zielinski, D. C., Bordbar, A., Lewis, N. E., Rahmanian, S., Kang, J., Hyduke, D. R. & Palsson, B. O: **Quantitative prediction of cellular metabolism with**

constraint-based models: the COBRA Toolbox v2.0. *Nat Protoc* 2011, **6**: 1290-1307.

50. Feist, A. M., Zielinski, D. C., Orth, J. D., Schellenberger, J., Herrgard, M. J. & Palsson, B. O: **Model-driven evaluation of the production potential for growth-coupled products of Escherichia coli.** *Metab. Eng* 2010. **12**:173-186.

51. Nogales, J., Palsson, B. O. & Thiele, I: **A genome-scale metabolic reconstruction of Pseudomonas putida KT2440: iJN746 as a cell factory.** *Bmc Syst Biol* 2008, 2:Artn 79

52. Mo, M. L., Palsson, B. O. & Herrgard, M. J: **Connecting extracellular metabolomic measurements to intracellular flux states in yeast.** *Bmc Syst Biol* 2009, 3:Artn 37.

CHAPTER 3.

FULLY BIOLOGICAL PRODUCTION OF ADIPIC ACID ANALOGS FROM BRANCHED CATECHOLS

Reproduced from:

Kruyer, N.S., Wauldron, N., Bommarius, A.S. and Peralta-Yahya, P. Fully biological production of adipic acid analogs from branched catechols. *Sci. Reports* **10,** 13367 (2020).

3.1 Abstract

Microbial production of adipic acid from lignin-derived monomers, such as catechol, is a greener alternative to the petrochemical-based process. Here, we produced adipic acid from catechol using catechol 1,2-dioxygenase (CatA) and a muconic acid reductase (MAR) in *Escherichia coli*. As the reaction progressed, the pH of the media dropped from 7 to 4-5 and the muconic acid isomerized from the *cis,cis* (ccMA) to the *cis,trans* (ctMA) isomer. Feeding experiments suggested that cells preferentially uptook ctMA and that MAR efficiently reduced all muconic isomers to adipic acid. Intrigued by the substrate promiscuity of MAR, we probed its utility to produce branched chiral diacids. Using branched catechols likely found in pretreated lignin, we found that while MAR fully reduced 2-methyl-muconic acid to 2-methyl-adipic acid, MAR reduced only one double bond in 3-substituted muconic acids. In the future, MAR's substrate promiscuity could be leveraged to produce chiral-branched adipic acid analogs to generate branched, nylon-like polymers with reduced crystallinity.

3.2 Introduction

Adipic acid is used in the production of nylon-6,6, a polyamide present in carpets, textiles and molded plastics. In 2016, the global production of adipic acid was ~ 3.3 million tons per year, with all adipic acid being produced from petroleum [1]. This process generated nearly 10% of global nitrogen oxide emissions [2].

Renewable production of adipic acid provides a greener alternative, and can be initiated from a variety of feedstocks, including simple sugars and lignin-derived aromatics [1]. Independent of the feedstock used, today, the renewable production of adipic acid is a semi-biological process, combining bioproduction of muconic acid followed by its chemical hydrogenation to adipic acid [3,4]. For example, *Pseudomonas putida* has been engineered to convert pretreated lignin to *cis,cis*-muconic acid (ccMA) at 100% yield from detectable monomers [5]. *Escherichia coli* has a maximum theoretical yield of 83% from glucose through the shikimate pathway [1], but experimentally a top yield of 22% has been achieved [6]. Purification of ccMA from the fermentation broth requires four unit operations to achieve the required purity (99.8%) at 81.4% yield [7]. Reduction of ccMA to adipic acid is performed using platinum, rhodium or palladium catalysts, with catalyst cost around $0.30 per kg of adipic acid [7–9], ~19% of the current market value of adipic acid from petroleum [10]. Direct production of adipic acid from glucose through a reverse adipate degradation pathway has been achieved, but has a lower maximum theoretical yield (67%) [11].

The fully biological production of adipic acid would eliminate (1) the need for a chemical reactor, (2) the catalyst cost, and (3) the cost of purifying muconic acid prior to chemical hydrogenation. Indeed, a recent techno-economic analysis (TEA) accounting for

both fixed and variable costs concluded that a fully biological route to adipic acid from glucose would result in an adipic acid price point of $1.36/kg, while fully chemical and hybrid biological and chemical routes would result in price points of $1.56/kg and $1.48/kg [10]. A separate TEA showed that switching feedstock from glucose to lignin monomers reduced adipic acid minimum selling price by 50% due to increased productivity and decreased feedstock cost [12]. Taken together, lignin is a more desirable feedstock than sugars for adipic acid production.

Lignin depolymerisation results in a number of aromatic compounds, including catechol, a key intermediate in the production of adipic acid from both glucose and lignin-derived aromatics such as ferulic acid and *p*-coumaric acid [1]. Niu *et al*. achieved the fully biological production of adipic acid from lignin-derived aromatics in *P. putida*. Specifically, 4-hydroxybenzoic acid was converted to catechol and subsequently 3-ketoadipoyl-CoA via the β-ketodipate pathway. Using three heterologous enzymes and one endogenous thioesterase, 3-ketoadipoyl-CoA was converted to adipic acid with 17.4% molar yield [13]. Sun *et al*. developed a shorter pathway in *E. coli*, where glucose is metabolized to catechol via the shikimate pathway and catechol is converted to adipic acid using catechol-1,2-dioxygenase (CatA) and an enoate reductase previously shown to reduce muconic acid to adipic acid (i.e. muconic acid reductase, MAR) [14]. This pathway achieved 80.6 μg/L/hr of adipic acid [15]. Therefore, a better understanding of the CatA-MAR enzyme cascade can be directly applied to adipic acid production pathways that route through catechol.

Here, we optimized the CatA-MAR enzyme cascade in *E. coli* and discovered that the cascade is capable of converting branched catechols to adipic acid analogs (Figure

3.1a). First, we optimized muconic acid production from catechol by screening and optimizing the expression of CatAs from five different organisms. Next, with the optimal CatA in hand, we maximized adipic acid production by optimizing the co-expression of two known MARs. Interestingly, during muconic acid production, the media acidified leading to the isomerization of ccMA to *cis,trans*-muconic acid (ctMA). To determine the extent to which MAR reduced ctMA, we fed ccMA and ctMA to cells expressing MAR and observed preference to reduce ctMA over ccMA. However, when we fed ccMA and ctMA to the cell lysate, MAR showed no preference between ccMA and ctMA. This data suggests that ctMA is preferentially transported into the cell over ccMA. Given the promiscuity of MAR, we explored its utility by feeding the CatA-MAR cascade branched catechols likely found in lignin [16] to produce chiral-branched diacids. We found that MAR reduced one double bond in muconic acid analogs with short alkyl chains at the 3-position and both double bonds of 2-methyl-muconic acid to produce 2-methyl-adipic acid. The promiscuity of MAR highlights the potential utility of the CatA-MAR enzyme cascade in lignin valorization, as lignin depolymerisation results in a heterogeneous mixture of aromatics.

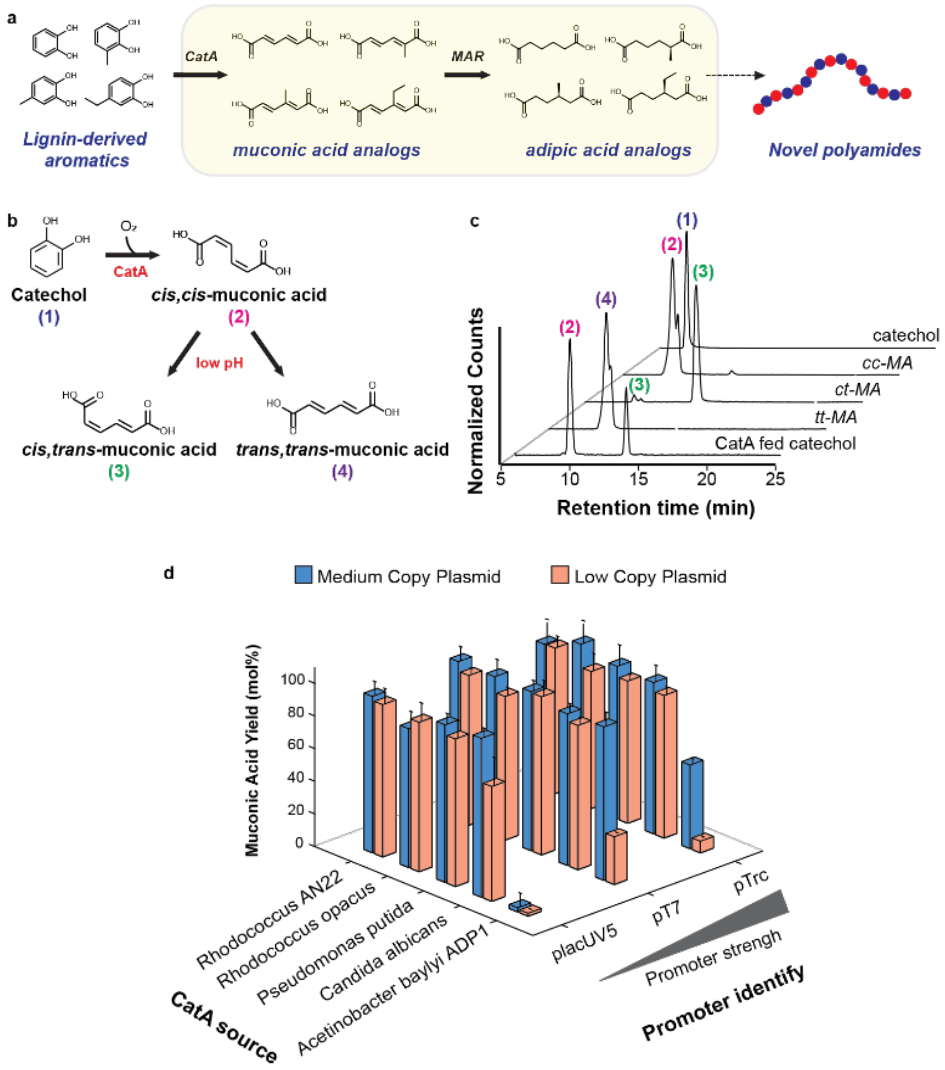

Figure 3.1. Muconic acid production from catechol. (a) Schematic for bioproduction of adipic acid analogs CatA: catechol-1,2-dioxygenase, MAR: muconic acid reductase. (b) Catechol to muconic acid conversion using CatA. At low pH *cis, cis*-muconic acid (cc-muconic acid) can isomerize to *cis, trans*-muconic acid (ct-muconic acid), and *trans, trans*-muconic acid (tt-muconic acid). (c) Liquid chromatograph/mass spectrometry (LC/MS) chromatogram of *E. coli* expressing *Rhodococcus sp.* AN22 CatA fed 1 g/L of catechol after 23 h incubation resulted in ccMA and ctMA. No ttMA was observed. Also shown, standards for catechol (6.8 min), ccMA (9.4 min), ctMA (13.6 min), and ttMA (9.6 min). Sample data for LC/UV absorbance spectra found in Figure 3.2 (d) Optimization of muconic acid production by screening CatAs from five different sources using three different promoter strengths and in low and medium copy plasmids. Experiments were run in triplicate and error bars represent the standard deviation from the mean.

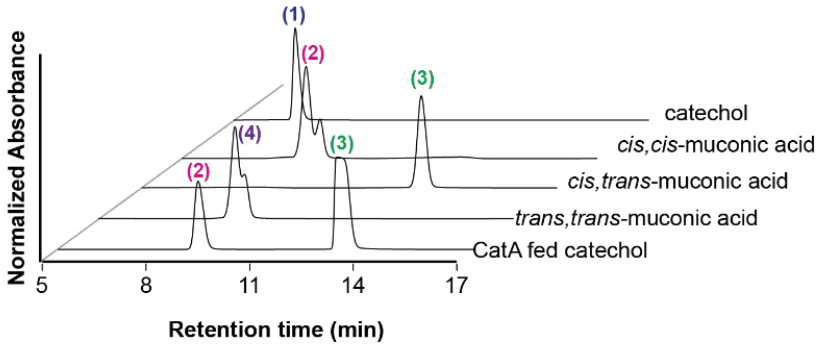

Figure 3.2. LC/UV spectra for standards of catechol, ccMA, ctMA, and ttMA, and the supernatant of *E. coli* expressing CatA-AN22 from the pTrc promoter in a medium copy plasmid (pNK56) when fed catechol. Corresponds to data in Figure 3.1c.

3.3 Materials and Methods

3.3.1 Reagents

All compounds were purchased from Sigma-Aldrich unless noted. Catechol (135011), ccMA (15992, ≥ 97% purity), *trans, trans*-muconic acid (ttMA) (M90003, 98% purity), 3-methyl-catechol (Alfa Aesar, A12324), 4-methyl-catechol (TCI, M0413) and 4-ethyl-catechol (Alfa Aesar, A12048). To synthesize ctMA, 5 mM ccMA was acidified to pH 4–4.5 and incubated for 1 h at 70 °C to achieve complete isomerization [17]. Solutions of ccMA were adjusted to pH 7 to avoid isomerization. All muconic acid solutions were prepared fresh prior to use.

3.3.2 Strains and plasmids

Gene sequences for the five CatA homologs (*Pseudomonas putida*—WP_010954549, *Acinetobacter baylyi* ADP1—CAG68305, *Candida albicans*—KGQ97177, *Rhodococcus sp*. AN22—BAH56722, and *Rhodococcus opacus*—3HGI_A)

and the two MAR homologs (*Bacillus coagulans*—AEO99944 and *Clostridium acetobutylicum*—AEI32805) were codon optimized for *E. coli* and commercially synthesized (Table D.1). High translation rate ribosome binding sites (RBSs) were designed using the RBS Calculator [18,19] and added to the synthesized genes via PCR (Table D.2). The genes carrying the high translation RBS were cloned into pBbA1a, pBbA1c, pBbS1a, pBbB1c, pBbA5c, pBbS5c, pBbB5a, pBbA7a and pBbS7a20 at BglII/ BamHI via Gibson assembly to generate CatA and MAR expression plasmids (Table C.1). To clone CatA and MAR as an operon, RBS2-CatA was amplified from pBbS1a_CatA-AN22 (pNK45), RBS1-MAR-BC was amplified from pNK102, and RBS1-MAR-CA was amplified from pBbA1a_MAR-CA (pNK146). The PCR products were cloned into pBbA1a at BglII/BamHI via Gibson assembly to generate pBbA1a_MAR-BC_CatA-AN22 (pNK111) and pBbA1a_MAR-CA_CatA-AN22 (pNW1). N-terminal 6x His Tags were added to CatA genes (pNK58-60 and pNK62) via PCR. The vector pBbS1a was amplified using primers NK153 and NK154. The His-tagged CatAs were cloned into pBbS1a backbone using Gibson Assembly to generate pNK167-171. Clones were confirmed via sequencing using primers NK18 (pTrc), NK19 (placUV5), NK20 (pT7) and NK22.

3.3.3 Screening CatA for muconic acid production from catechol

Overnight cultures of *E. coli* DH10B transformed with pTrc and placUV5 CatA expression plasmids (pNK53-72) or *E. coli* BL21 (DE3) transformed with pT7 CatA expression plasmids (pNK73-82) were diluted to OD600 = 0.1 in 5 mL of M9 minimal media (Difco 248,510: 33.9 g/L Na2HPO4, 15.0 g/L KH2PO4, 2.5 g/L NaCl, 5.0 g/L

NH4Cl, 2 mM MgSO4, 100 μM CaCl2, supplemented with 0.5% glucose) and 100 mg/L ampicillin or 50 mg/L chloramphenicol. Cultures were incubated (37 °C, 250 RPM) until reaching OD600 = 0.4–0.6. CatA expression was induced using 500 μM IPTG and 1 g/L (9.1 mM) catechol was fed to the cultures. Cultures were incubated for 23 h at 37 °C, 250 RPM. After incubation, cultures were centrifuged and the supernatant analysed via liquid chromatography, mass spectrometry (LC/MS).

3.3.4 Screening MAR for adipic acid production from muconic acid

Overnight cultures of *E. coli* Δ*iscR* transformed with MAR expression plasmids (pNK102, pNK146-149, pNK160-163, pNW5, pNW8 and pNW10) were diluted to OD600 = 0.1 in 5 mL of M9 minimal media with 100 mg/L ampicillin or 50 mg/L chloramphenicol and supplemented with 1 mg/L thiamine. Cultures were incubated (37 °C, 250 RPM) until reaching OD600 = 0.4–0.6. MAR expression was induced using 500 μM IPTG and fed either 500 μM ccMA, ctMA or ttMA. Cultures were then transferred to anaerobic conditions (anaerobic tubes with argon headspace and sealed using chlorobutyl rubber stoppers and Parafilm) and incubated for 24 h at 30 °C, 250 RPM. After incubation, cultures were centrifuged and the supernatant analysed via LC–MS.

3.3.5 Cell lysate MAR activity on muconic acid isomers

Overnight cultures of *E. coli* Δ*iscR* transformed with MAR expression plasmids (pNK102, pNK146) or *E. coli* Δ*iscR* (control) were diluted to OD600 = 0.1 in 5 mL of LB media (Difco 240210: 10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) supplemented 100 mg/L ampicillin when needed. Cultures were incubated (37 °C, 250 RPM) until reaching

OD600 = 0.4–0.6. MAR expression was induced with 500 μM IPTG and cultures were transferred to anaerobic conditions and incubated for 24 h at 30 °C, 250 RPM. After incubation, 500 μL of 10× lysis solution (10 mg/mL lysozyme, 0.5% Tween-20, 10 mM DTT, 10 U/mL DNase, 25 mM MgCl2) was added to the culture using a needle through the rubber stopper to maintain low oxygen conditions within the tube. Cells were lysed for 1 h at 30°, 250 RPM. After lysis, 500 μL of 7 mM ccMA, ctMA or ttMA (final concentration 583 μM) was added using a needle through the rubber stopper. The cultures were incubated for 24 h (30 °C, 250 RPM), centrifuged, and the supernatant analysed via LC/MS. The control corroborated that neither media nor lysis solution reduced muconic acid to adipic acid.

3.3.6 Production of adipic acid analogs from lignin-derived aromatic monomers

Overnight cultures of *E. coli* Δ*iscR* transformed with CatA-MAR enzyme cascades (pNK111 or pNW1) were diluted to OD600 = 0.1 in 5 mL of M9 minimal media with 100 mg/L ampicillin or 50 mg/L chloramphenicol and supplemented with 1 mg/L thiamine and incubated at 37 °C, 250 RPM until reaching OD600 = 0.4–0.6. Co-expression of CatA and MAR was induced using 500 μM IPTG and fed either 1 mM catechol or substituted catechol (3-methyl-catechol, 4-methyl-catechol or 4-ethyl-catechol). Cultures were incubated for 2 h at 30 °C, 250 RPM before transferring to anaerobic conditions and cultured for 22 h. After incubation, the cultures were centrifuged and the supernatant was analyzed via LC/MS.

3.3.7 Chemical analysis

LC/MS analysis was completed using an Agilent 1100/1260 series system equipped with a 1260 ALS autosampler, a multi-wavelength detector (MWD) and a 6120 Single Quadrupole LC/MS with a Poroshell 120 SB-C18 3.0 mm × 50 mm × 2.7 µm column (Agilent) and an electrospray ion source. Column temperature was kept constant at 28 °C. LC method was based on method developed by Sun *et al.* [21]. LC conditions for muconic acid analysis: Solvent A—water with 0.1% formic acid, Solvent B—methanol. Gradient: 1 min 95%:5%:0.2 (%A:%B:flow rate in ml/min), 15 min ramp to 75%:25%:0.2, 1 min ramp to 95%:5%:0.2, 18 min 95%:5%:0.2. LC conditions for analogs of muconic acid and adipic acid were: Solvent A—water with 0.1% formic acid, Solvent B—acetonitrile. Gradient: 1 min 95%:5%:0.1, 25 min ramp to 75%:25%:0.1, 1 min ramp to 95%:5%:0.1, 18 min 95%:5%:0.1. For CatA screening, catechol and muconic acid were quantified using the MWD (274 nm catechol, 260 nm muconic acid). Quantification was done using the area under the peak for the corresponding MWD signal. For MAR screening, MS acquisition was done in negative ion mode for all samples and using selected ion monitoring (SIM). Catechol (*m/z* 109.1), ccMA (*m/z* 141.1), ctMA (*m/z* 141.1), ttMA (*m/z* 141.1), hexenedioic acid (*m/z* 143.1), adipic acid (*m/z* 145.1), 3-methylcatechol (*m/z* 123.1), 2-methyl-muconic acid (*m/z* 155.1), 2-methyl-hexenedioic acid (*m/z* 157.1), 2-methyl-adipic acid (*m/z* 159.1), 4-methyl-catechol (*m/z* 123.1), 3-methyl-muconic acid (*m/z* 155.1), 3-methyl-hexenedioic acid (*m/z* 157.1), 3-methyl-adipic acid (*m/z* 159.1), 4-ethyl-catechol (*m/z* 137.1), 3-ethyl-muconic acid (*m/z* 169.1), 3-ethyl-hexenedioic acid (*m/z* 171.1) and 3-ethyl-adipic acid (*m/z* 173.1). Quantification was done using the area under the peak for the corresponding SIM signal. Areas were converted to concentrations

for catechol, ccMA, ctMA, ttMA and adipic acid, using standard curves generated using commercial standards. There are no commercial standards for 2-methyl-muconic acid, 3-methyl-muconic acid, 3-ethyl-muconic acid, hexenedioic acid, 2-methyl-hexenedioic acid, 3-methyl-hexenedioic acid, 3-ethyl-hexenedioic acid and 2-methyl-adipic acid, MS counts are reported.

3.3.8 SDS-PAGE analysis of CatA expression

Overnight cultures of *E. coli* DH10B transformed with His-tagged CatA expression plasmids (pNK167-171) were diluted to OD600 = 0.1 in 5 mL of LB media with 100 mg/L ampicillin. Cultures were incubated (37°C, 250 RPM) until reaching OD600 = 0.4-0.6. CatA expression was induced using 500 µM IPTG and incubated for 4 hours (37°C, 250 RPM). After incubation, cultures were centrifuged and the pellets resuspended in 1x phosphate buffered saline (PBS: Teknova P0195 – 135 mM NaCl, 2.7 mM KCl, 4.3 mM Na2HPO4, 1.4 mM NaH2PO4, pH 7) to a cell density of 40 mg/mL. 1 mL of cell suspension was lysed using sonication. The sonication was performed on ice, and the sonication protocol was 8 on-off cycles of 30 seconds. After sonication, the lysed cells were centrifuged. The total protein concentration of the soluble fraction was measured by the absorbance at 280nm using a NanoDrop Lite (Thermo). Protein concentrations were standardized to 2 mg/mL total protein by diluting in PBS to a total volume of 50 µL. After addition of 10 µL of 6x SDS loading dye, the samples were heated at 95°C for 15 minutes. 20 µL of each sample was loaded to both 12% Bis-Tris SDS-PAGE gels (Invitrogen) along with the BenchMarkTM His-tagged Protein Standard (Invitrogen LC5606). Two identical

gels were run in parallel in MES SDS running buffer (Invitrogen) at 200 V for 45 minutes at 4°C. One gel was stained with Coomassie Blue and one was used for Western Blot.

3.3.9 Western blot analysis of CatA expression

Proteins in the SDS-PAGE gel were transferred to nitrocellulose paper using ThermoFisher iBlot™ Transfer Stack. The primary antibody was anti-polyhistidine antibody produced in mouse (Sigma H1029) and the secondary antibody was anti-mouse Fc produced in goat (Sigma A3638). After antibody incubation, the nitrocellulose paper was stained with BCIP/NBT (Sigma BC5655).

3.4 Results and Discussion

3.4.1 Catechol to muconic acid conversion

CatAs from *P. putida*, *Pseudomonas aeruginosa*, *Acinetobacter calcoaceticus*, *Acinetobacter baylyi*, and *A. baylyi* ADP1 have been used to produce muconic acid in *E. coli* [6,21–23]. We compared the *E. coli* performance of CatAs from *P. putida*, *A. baylyi* ADP1 (both dimers with moderate specific activity—*P. putida* 22.4 μM/min/mg), *Candida albicans* (dimer with a higher specific activity, 63 μM/min/mg), *Rhodococcus sp*. AN22 (a monomer) and *Rhodococcus opacus* (structurally well studied) [8,24–28] (Table 3.1). The CatAs were expressed from three promoters with different strengths in low and medium plasmid copy number (Figure 3.1b-d). All CatAs performed similarly independent of promoter strength and plasmid copy number, reaching almost 100% catechol to muconic acid conversion after 24 h (Figure 3.1d). Of note, *A. baylyi* ADP1 CatA showed < 10% conversion in half of the tested conditions despite similar protein expression (Figure 3.3).

Active site alignment of the ADP1 CatA and *P. putida* CatA revealed that at position 76 ADP1 CatA codes for a proline, while *P. putida* CatA codes for an alanine (Figure 3.4). A proline to alanine mutation at this position has been shown to increase ADP1 CatA specific activity 10 fold [27]. Therefore, P76 in ADP1 CatA may lead to the underperformance observed when compared to the other CatA homologs. For all subsequent experiments, we used *Rhodococcus sp*. AN22 CatA as it performed well and it is a monomer, reducing the metabolic load of the system. *E. coli* expressing AN22 CatA converts catechol to ccMA as the major product. However, we also detected the thermodynamically stable isomer ctMA, which is likely produced from ccMA after media acidification after 24 h growth, which reaches a pH 4.3 after starting at pH ~ 7 (Figure 3.1b,c). We measured the maximum ccMA to ctMA isomerization rate to be between pH 3–5 (Figure 3.5), with previous literature confirming the maximum at pH 4 [17]. No isomerization of ctMA to ttMA was observed, which is consistent with the literature [17].

Table 3.1. Rationale for CatA screening choice.

Organism	Quaternary structure	Specific Activity (µmol/min/mg)	Logic for Testing
Acinetobacter baylyi. ADP1	Dimer	Only Crude Reported [6,27]	Similar to used to achieve highest cis,cis-muconic acid titers in *Escherichia coli* to date [6]. Mutational study of active site [27].
Candida albicans	Dimer	63 [25]	CatA with the highest reported specific activity [25]
Pseudomonas putida	Dimer	22.4 [8]	Natively expressed in *Pseudomonas putida*, an organism used for lignin monomer upgrading [8].
Rhodococcus opacus	Dimer	n.d. [26]	Enables structural insight through crystal structures with a variety of substrates [26].
Rhodococcus sp. AN22	Monomer	35 [24]	CatA monomer with highest reported specific activity [24].

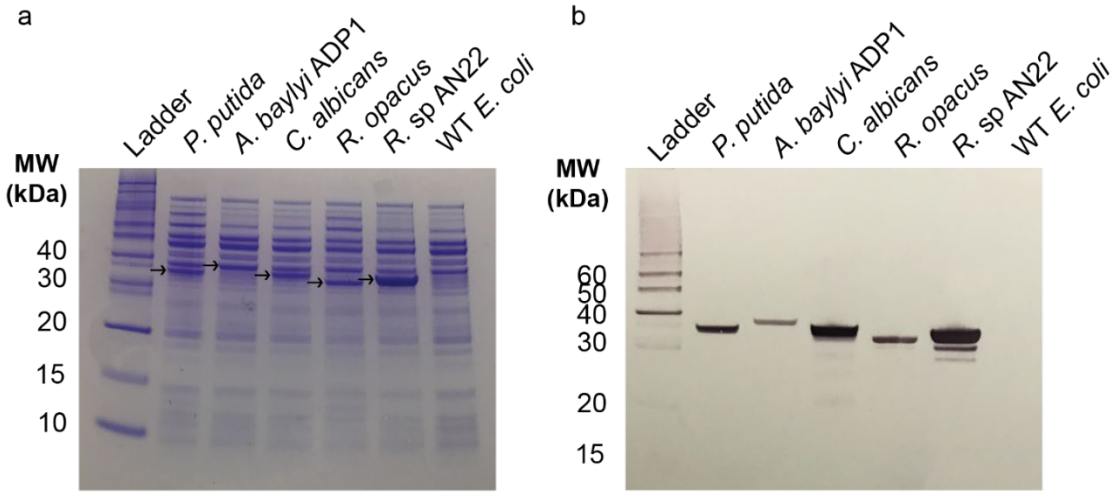

Figure 3.3. Expression of CatA homologs. (a) Coomassie stained SDS-PAGE gel for His6-CatA homologs expressed under pTrc promoter, in low copy vector. (b) Western Blot His6-CatA homologs. CatA Molecular Weight: *P. putida* – 34.3 kDa, *A. baylyi* ADP1 – 34.3 kDa, *C. albicans* – 33.8 kDa, *R. opacus* – 30.7 kDa, *R. sp.* AN22 – 31.6 kDa.

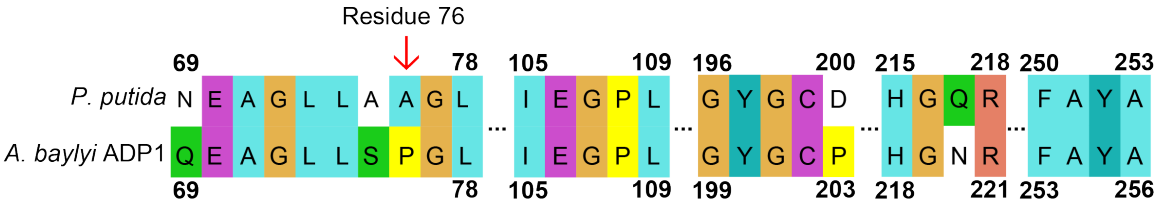

Figure 3.4. Active site alignment of *Pseudomonas putida* CatA and *Acinetobacter baylyi* ADP1 CatA. Active site residues from Han *et al.* [27]. Colors represent different classes of amino acid. Key difference at residue 76 may account for conversion differences when using *P. putida* CatA and *A. baylyi* ADP1 CatA.

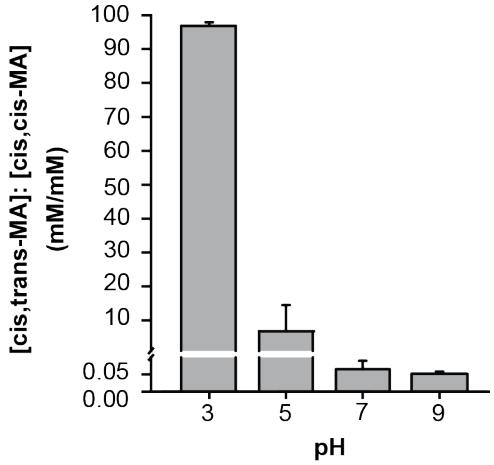

Figure 3.5. Isomerization of *cis, cis*-muconic acid to *cis, trans*-muconic acid as a function of pH. Ratio of concentration of ctMA to concentration of ccMA after 300 µM solution of ccMA was incubated for 24 h at 37°C and 250 RPM in M9 media of varying pH value.

3.4.2 Muconic acid to adipic acid conversion

We compared the *E. coli* performance of *Bacillus coagulans* MAR (MAR-BC) and *Clostridium acetobutylicum* MAR (MAR-CA) [14] expressed from two promoters with different strengths in low, medium and high plasmid copy number (Figure 3.6). MAR-BC achieved a 6.1% conversion of 500 μM ccMA to adipic acid using the pTrc promoter from a medium-copy plasmid after a 24 h anaerobic growth. This is lower than Sun *et al*. who used batch fermentation to convert 2.8 mM of ccMA to adipic acid (18.0% total conversion) using MAR-CA [15]. We rationalize our lower percent conversion in comparison by the use of minimal M9 media supplemented only with 0.5% glucose rather than a modified M9 medium, which contains 0.5% yeast extract, 0.25% glucose and 1% glycerol that aids in higher protein expression and increased cell density.

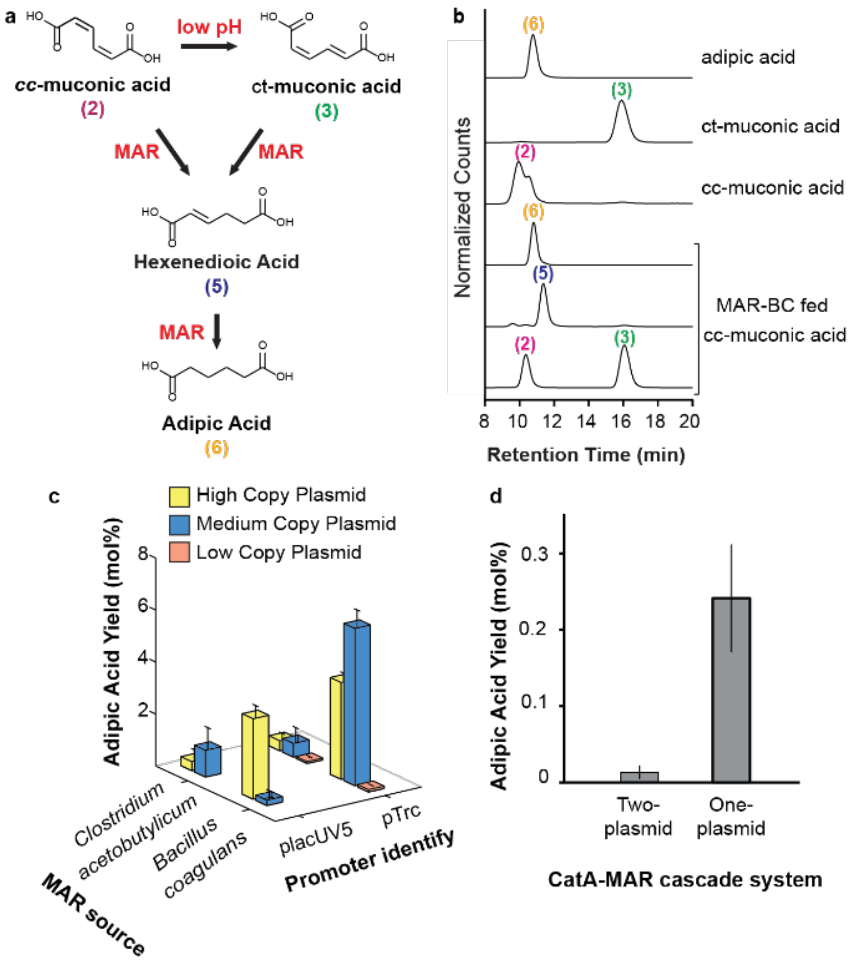

Figure 3.6. Adipic acid production from muconic acid and catechol. (a) Reduction of *cis, cis*-muconic acid (cc-muconic acid) and *cis, trans*-muconic acid (ct-muconic acid) to adipic acid by muconic acid reductase (MAR). (b) LC/MS chromatograms of *E. coli* expressing *Bacillus coagulans* MAR fed 500 µM ccMA after 24 h incubation results in adipic acid, hexenedioic acid as well as unreacted *cis, cis*-muconic acid and *cis, trans*-muconic acid. Also shown, standards for adipic acid (m/z 146; 10.8 min), *cis, trans*-muconic acid (m/z 142; 16.1 min), and *cis, cis*-muconic acid (m/z 142; 10.0 min). (c) Adipic acid yields (mol%) obtained by screening MARs from two different sources using two different promoter strengths and in low, medium and high copy plasmids. (d) Adipic acid yields obtained from feeding 500 mg/L catechol to *E. coli* co-expressing *Rhodococcus sp.* AN22 CatA and *Bacillus coagulans* MAR in a two and one plasmid system. For c and d, experiments were run in triplicate and error bars represent the standard deviation from the mean.

3.4.3 Adipic acid production from catechol

The experimental conditions for the CatA-MAR enzyme cascade required balancing the molecular oxygen requirement of the CatA reaction with the oxygen sensitivity of MAR. Thus, fermentations were run as a two-stage batch process with a 2 h aerobic stage followed by a 22-h anaerobic stage [29]. The CatA-MAR cascade was tested in a two- and one-plasmid system. As Figure 3.5d shows, the one-plasmid system resulted in 1.6 mg/L of adipic acid after 24 h, or a 0.241% molar yield from the fed 1 g/L catechol, an 18-fold improvement over the two-plasmid set up.

3.4.4 MAR muconic acid isomer preference

At pH < 7, ccMA isomerized to ctMA, and both isomers were present at a roughly equal molar ratio in the fermentation broth (Figure 3.6b). Hypothesizing that MAR may not reduce ctMA as efficiently as ccMA, we fed ccMA, ctMA and, for completion, ttMA to cells expressing MAR and to a cell lysate expressing MAR. In the cell-based experiment, both MARs showed a preference to reduce ctMA over ccMA with MAR-CA showing an eightfold preference. MAR-BC also showed a higher adipic acid yield than MAR-CA; 11.5-fold in the case of ccMA, and twofold in the case of ctMA (Figure 3.7a). In the cell lysate-based experiment, both MARs reduced ccMA and ctMA to the same extent. Interestingly, MAR-CA showed a higher adipic acid yield than MAR-BC (Figure 3.7b). MAR-CA also showed a preference for ttMA, which is consistent with previous literature [14]. Of note, the overall yields in the cell lysate experiment were lower than the cell-based experiment, likely due to enzyme deactivation from cell lysis solution components and the reduced co-factor concentration. Taken together, we rationalize the *in vivo* MAR substrate

preference for ctMA to be the result of increased membrane permeability for ctMA over ccMA. We rationalize the higher *in vivo* activity of MAR-BC over MAR-CA to the fact that once a muconic acid isoform enters the cell the MAR experiences a higher localized substrate concentration. Previously, it has been suggested that MAR-BC has a lower substrate affinity but higher catalytic activity than MAR-CA [14]. Such MAR-BC enzymatic characteristics would fit the observed results.

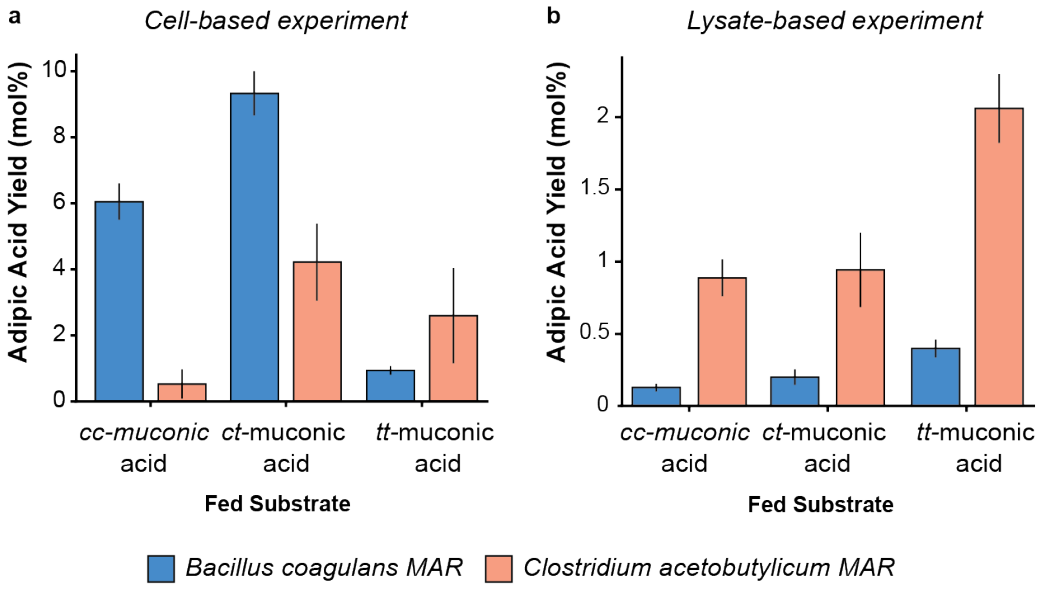

Figure 3.7. Muconic acid reductase substrate preference. (a) Adipic acid production of *Escherichia coli* expressing *Bacillus coagulans* muconic acid reductase (MAR) or *Clostridium acetobutylilcum* MAR when feeding 500 µM of *cis, cis*-muconic acid (cc-muconic acid), *cis, trans*-muconic acid (ct-muconic acid), or *trans, trans*-muconic acid (tt-muconic acid) isomers. (b) Adipic acid yield of lysed *E. coli* expressing *B. coagulans* MAR and *C. acetobutylicum* MAR fed 583 µM of muconic acid isomers. All experiments were performed in triplicate and error bars represent the standard deviation from the mean.

3.4.5 MAR muconic acid analog preference

Given that MAR reduces all muconic acid isomers, we investigated the extent to which the CatA-MAR cascade can produce branched adipic acid analogs from alkyl substituted catechols likely found in pretreated lignin [16]. We confined our analysis to commercially available catechols previously shown to be oxidized by CatA: 3-methyl-catechol (3MC), 4-methyl-catechol (4MC), and 4-ethylcatechol (4EC) [30]. CatA oxidized 3MC, 4MC and 4EC to 2-methyl-muconic acid, 3-methyl-muconic acid and 3-ethyl-muconic acid, respectively (Figure 3.8a-c). MARs singly reduced 3-methyl-muconic acid and 3-ethylmuconic acid to 3-methyl-hexenedioic acid and 3-ethyl-hexenedioic acid. Only 2-methyl-muconic was doubly reduced to produce 2-methyl-adipic acid by MAR (Figure 3.8a). While the substrate specificities of MAR-BC and MAR-CA were similar, MAR-BC led to higher yields of the singly and fully reduced muconic acid analogs (Figure 3.8). CatA oxidation of 3MC resulted in two 2-methyl-muconic acid peaks. Presumably, *cis, cis*-2-methyl-muconic acid isomerized to the *cis, trans* isomer over the course of the reaction (Figure 3.8a, Figure 3.9a). Interestingly, CatA oxidation of 4MC or 4EC resulted in a single peak for 3-methyl-muconic acid and 3-ethyl-muconic acid, respectively (Figure 3.8b-c, Figure 3.9b-c). Given the steric crowding between the ketone group and the alkyl chains at position 3 in muconic acid in the cis,cis configuration, we presume that only *cis,trans*-3-methylmuconic acid and *cis, trans*-3-ethyl-muconic acid are present in the media.

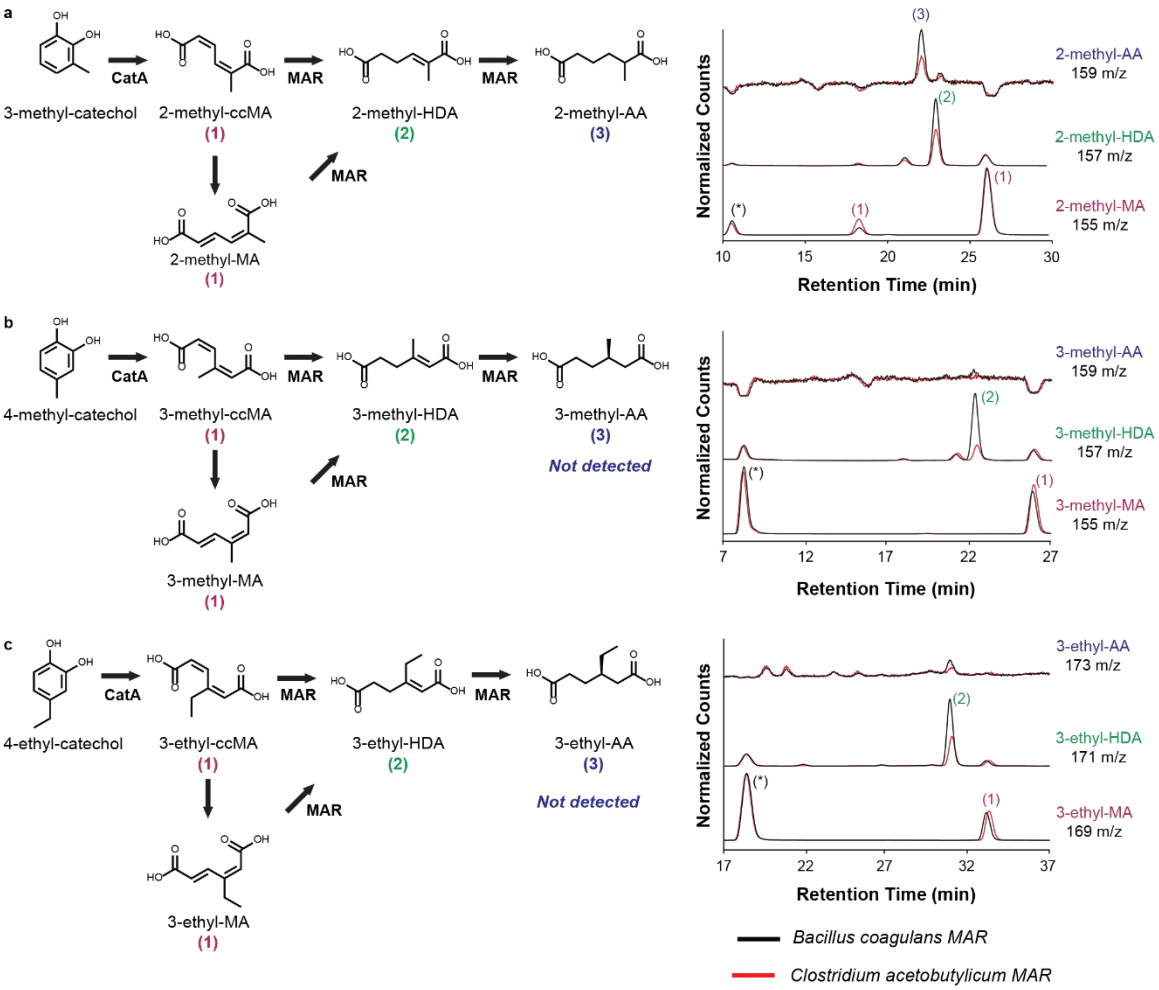

Figure 3.8. Production of muconic acid and adipic acid analogs. Reaction schematic and LC/MS chromatograms of *Escherichia coli* co-expressing *Rhodococcus sp.* AN22 CatA and either *B. coagulans* MAR or *C. acetobutylicum* MAR and fed 1 mM (a) 3-methyl-catechol results in 2-methyl hexenedioic acid and 2-methyl-adipic acid; (b) 4-methyl-catechol results in 3-methy hexenedioic acid, and (c) 4-ethyl-catechol results in 3-ethyl hexenedioic acid. *: Peaks visible in MS but not in UV channel, thus not muconic acid analog peaks (Figure 3.9).

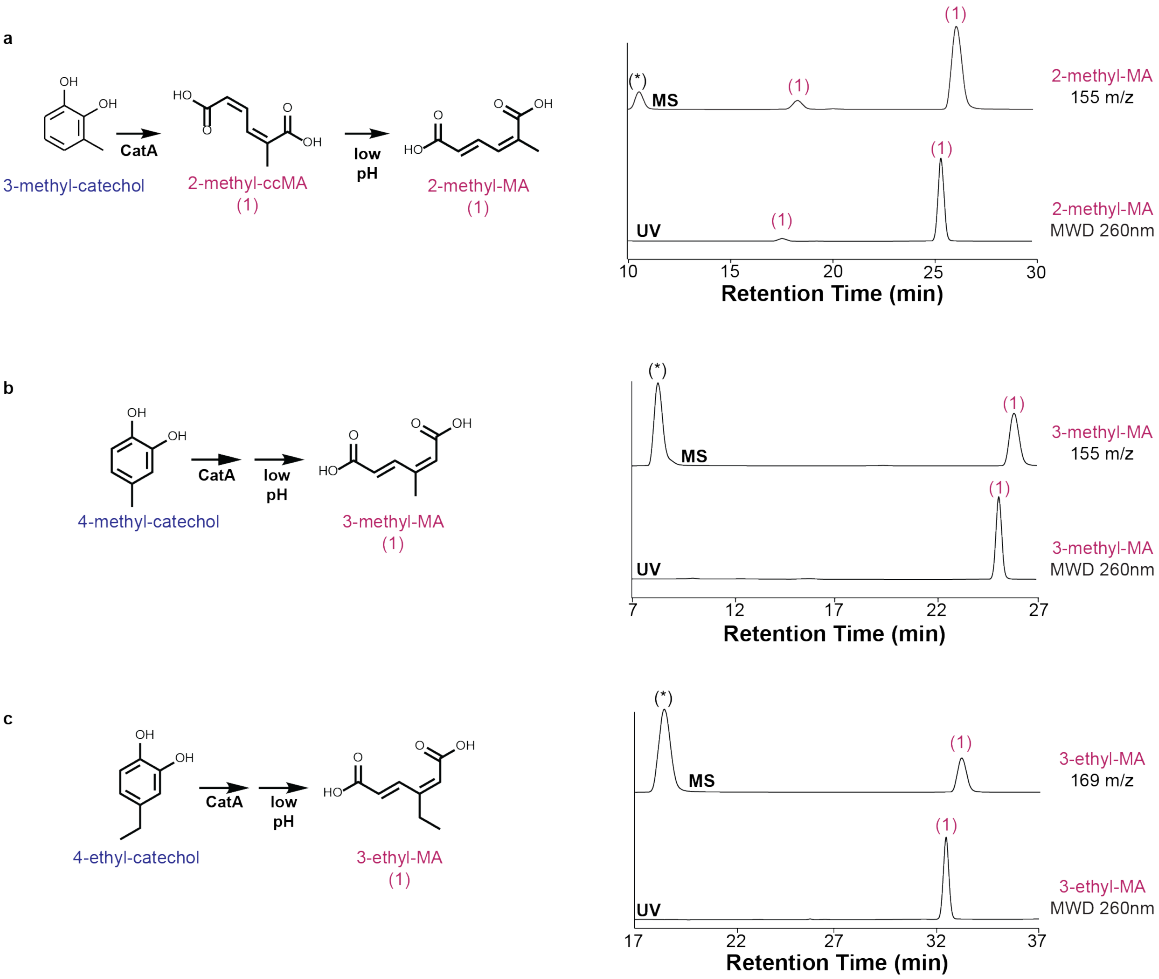

Figure 3.9. Comparison of MS and UV active peaks for Figure 3.8. (a) 2-methyl muconic acid (m/z 155) and UV trace. (b) 3-methyl-muconic acid (m/z 155) and UV trace. (c) 3-ethyl-muconic acid (m/z 169) and UV trace. All traces come from *E. coli* ΔiscR expressing CatA from *Rhodococcus sp.* AN22 and MAR from *Bacillus coagulans* fed 1 mM of substituted catechol after 2-hr aerobic and 22-hr anaerobic incubation. Peaks appearing in both SIM and UV channels are muconic acid analogs. *: Peaks only present in SIM (likely media components). Difference in peak times between UV trace and SIM due to volume (tubing) distance between MWD and MS detectors.

3.5 Conclusions

The fully biological production of adipic acid from catechol, a pretreated lignin monomer, was achieved by screening CatA from different sources and optimizing its co-expression with MAR in *E. coli* to produce adipic acid at 1.6 mg/L or a 0.241% molar yield. The muconic acid yields presented in this work were lower than in previous studies, which we attribute to the use of batch fermentation rather than a biocatalysis set up and the use of minimal media rather than rich media, which makes it problematic to calculate yields.

A closer study of oxygen sensitivity differences between MAR-BC and MAR-CA may help to engineer a more oxygen tolerant enzyme that will be useful in the production of adipic acid from lignin-derived monomers as oxygen is both a substrate (CatA) and an inhibitor (MAR) of the process. Engineering MAR for oxygen tolerance will help improve catechol to adipic acid yields in the future.

A key finding of this work is the use of MAR to produce branched adipic acid analogs. Thus, the CatA-MAR cascade could be used to convert lignin-derived monomers to chiral branched dicarboxylic acids that may give tunable properties to nylon-6,6 like polymers. Application of the enzyme cascade to a variety of lignin-derived monomers demonstrates increased utility as a lignin valorization approach.

3.6 References

1. Kruyer, N.S. & Peralta-Yahya, P. Metabolic engineering strategies to bio-adipic acid production. *Curr. Opin. Biotechnol*. **45,** 136-143 (2017).

2. Polen, T., Spelberg, M. & Bott, M. Toward biotechnological production of adipic acid and precursors from biorenewables. *J. Biotechnol*. **167,** 75-84 (2013).

3. Kohlstedt, M. *et al.* From lignin to nylon: cascaded chemical and biochemical conversion using metabolically engineered *Pseudomonas putida*. *Metab. Eng*. **47,** 279-293 (2018).

4. van Duuren, J.B.J.H. *et al.* Limited life cycle and cost assessment for the bioconversion of lignin-derived aromatics into adipic acid. *Biotechnol. Bioeng.* **117,** 1381-1393 (2020).

5. Salvachúa, D. *et al.* Bioprocess development for muconic acid production from aromatic compounds and lignin. *Green. Chem*. **20,** 5007-5019 (2018).

6. Niu, W., Draths, K.M. & Frost, J.W. Benzene-free synthesis of adipic acid, *Biotechnol. Progr*. **18,** 201-211 (2002).

7. Vardon, D.R. *et al.* cis,cis-Muconic acid: separation and catalysis to bio-adipic acid from nylon-6,6 polymerization. *Green. Chem*. **18,** 3397-3413 (2016).

8. Vardon, D.R. *et al.* Adipic acid production from lignin. *Energy. Environ. Sci*. **8,** 617-628 (2015).

9. Claypool, J.T. & Raman, D.R. Development and validation of a technoeconomic analysis tool for early-stage evaluation of bio-based chemical production processes. *Bioresour. Technol*. **150,** 486-495 (2013).

10. Gunukula, S. & Anex, R.P. Techno-economic analysis of multiple bio-based routes to adipic acid. *Biofuel. Bioprod. and Biorefin.* **11,** 897-907 (2017).

11. Zhao, M. *et al.* Metabolic engineering of *Escherichia coli* for producing adipic acid through the reverse adipate-degradation pathway. *Metab. Eng.* **47,** 254-262 (2018).

12. Johnson, C.W. *et al.* Innovative chemicals and materials from bacterial aromatic catabolic pathways. *Joule*. **3,** 1523-1537 (2019).

13. Niu, W. *et al.* Direct biosynthesis of adipic acid from lignin-derived aromatics using engineered *Pseudomonas putida* KT2440. *Metab. Eng*. **59,** 151-161 (2020).

14. Joo, J.C. *et al.* Alkene hydrogenation activity of enoate reductases for an environmentally benign biosynthesis of adipic acid. *Chem. Sci*. **8,** 1406-1413 (2017).

15. Sun, J., Raza, M., Saun, X. & Yuan, Q. Biosynthesis of adipic acid via microaerobic hydrogenation of *cis, cis*-muconic acid by oxygen sensitive enoate reductase, *J. Biotechnol*. **280,** 49-54 (2018).

16. Schutyser, W. *et al.* Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chem. Soc. Rev.* **47,** 852-908 (2018).

17. Carraher, J.M., Pfennig, T., Rao, R.G., Shanks, B.H. & Tessonnier, J.-P. cis,cis-Muconic acid isomerization and catalytic conversion to biobased cyclic-C6-1,4-diacid monomers. *Green. Chem*. **19,** 3042-3050 (2017).

18. Espah Borujeni, A. & Salis, H.M. Translation initiation is controlled by RNA folding kinetics via a ribosome drafting mechanism. *J. Am. Chem. Soc*. **138,** 7016-7023 (2016).

19. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol*. **27,** 946 (2009).

20. Lee, T.S. *et al.* BglBrick vectors and datasheets: a synthetic biology platform for gene expression. *J. Biol. Eng.* 12 (2011).

21. Sun, X.X. *et al.* A novel muconic acid biosynthesis approach by shunting tryptophan biosynthesis via anthranilate. *Appl. Environ. Microbiol.* **79,** 4024-4030 (2013).

22. Lin, Y., Sun, X., Yuan, Q. & Yan, Y. Extending shikimate pathway for the production of muconic acid and its precursor salicylic acid in *Escherichia coli*. Metab. *Eng*. **23,** 62-69 (2014).

23. Wu, W. *et al.* Lignin valorization: two hybrid biochemical routes for the conversion of polymeric lignin into value-added chemicals. *Scientific Rep*. **7,** 8420 (2017).

24. Matsumura, E., Ooi, S., Murakami, S., Takenaka, S. & Aoki, K. Constitutive synthesis, purification and characterization of catechol 1,2-dioxygenase from the aniline-assimilating bacterium *Rhodococcus sp.* AN-22. *J. Biosci. Bioeng*. **98,** 71-76 (2004).

25. Tsai, S.-C. & Li, Y.-K. Purification and characterization of a catechol 1,2-dioxygenase from a phenol degrading *Candida albicans* TL3. Arch. *Microbiol.* **187,** 199-206 (2007).

26. Matera, I. *et al.* Catechol 1,2-dioxygenase from the gram-positive *Rhodococcus opacus* 1CP: quantitative structure/activity relationship and the crystal structures of nativee and catechols adducts. *J. Struct. Biol*. **170,** 548-564 (2010).

27. Han, L. *et al.* Engineering catechol 1,2-dioxygenase by design for improving the performance of the cis, cis-muconic acid and synthetic pathway in *Escherichia coli, Scientific Rep*. **5,** 11 (2015).

28. Sanakis, Y., Mamma, D., Christakopoulos, P. & Stamatis, H. Catechol 1,2-dioxygenase from *Pseudomonas putida* in organic media – an electron paramagnetic resonance study. *Int. J. Biol. Macromol*. **33,** 101-106 (2003).

29. Raj, K. *et al.* Biocatalytic production of adipic acid from glucose using engineered *Saccharomyces cerevisiae*. *Metab. Eng. Commun*. **6,** 28-32 (2018).

30. Cha, C.-J. Catechol 1,2-dioxygenase from *Rhodococcus rhodochrous* N75 capable of metabolizing alkyl-substituted catechols. *J. Microbiol. Biotechn*ol. **16,** 778-785 (2006).

CHAPTER 4.

DESIGNING THE PRODUCTION OF A MARTIAN ROCKET PROPELLANT USING *IN SITU* RESOURCE UTILIZATION

4.1 Abstract

Recent success of the NASA Perseverance mission proves that the age of Martian exploration is upon us. However, the step from sending robots to sending humans to the surface of Mars requires further technological development towards the return journey from Mars to Earth. A critical piece of this journey is the launch of a Mars Ascent Vehicle from the surface of Mars into low Mars orbit that requires 30 tons of rocket propellant shipped from Earth. To reduce mission payload, we designed a biotechnology-enabled *in situ* resource utilization (ISRU) strategy in which photosynthetic cyanobacteria convert abundant Martian CO_2 into simple sugars that are extracted and upgraded into the novel rocket propellant, 2,3-butanediol, by engineered *Escherichia coli*. Our process model identified key optimization parameters such as the use of lighter-weight process materials and achievable biological improvements in cyanobacterial biomass productivity, enzymatic digestion of cyanobacterial biomass, and 2,3-butandeiol yield. Critically, an optimized bio-ISRU produces 30.6 tons of excess O_2, a key resource used for human life support and other ISRU processes in the O_2-poor Martian environment. This is a significant improvement on a chemical-enabled ISRU that produces no excess O_2 with comparable payload mass and power usage.

4.2 Introduction

For human presence beyond earth to become a reality [1], we will need to not only send humans far into space, but also bring them back home safely. A key challenge towards this vision is the origin of the fuel needed to bring spacecraft back to Earth. For example, shipping the ~30 tons of methane and liquid oxygen (LOX) needed to power a rocket launch from the surface of Mars into low orbit requires an initial payload of 500 tons, which would cost ~$8 billion to launch from Earth [2] (Figure 4.1a). An *in situ* resource utilization (ISRU) strategy would leverage Martian resources to produce the rocket propellant on Mars to bring spacecraft back to Earth [2].

Currently proposed ISRU strategies, including the NASA Human Exploration of Mars Design Reference Architecture (DRA 5.0), focus on methane (CH_4) as an alternative rocket propellant due to its favorable heating value and specific impulse (I_{sp}) compared to rocket propellant 1 (RP-1) [2] (Figure 4.1a). Specifically, DRA 5.0 proposes using the Sabatier reaction to combine Martian carbon dioxide (CO_2) with hydrogen (H_2) either shipped from Earth or produced from Martian water via electrolysis to produce CH_4. Challenges with the DRA 5.0 strategy include the large volume required for H_2 shipment, tripling the size of the delivery spacecraft, and the fact that the Sabatier reaction does not produce enough oxygen (O_2) to burn all the CH_4 produced, requiring supplemental O_2 production through solid oxide CO_2 electrolysis (SOCE). Oxygen production is critical because Martian atmospheric O_2 levels (0.13%, $9x10^{-3}$ hPa) are insufficient for rocket propellant burning [3]. Alternatively, CH_4 could be shipped from Earth with only the necessary O_2 being produced *in situ* using SOCE (Figure 4.1b). This O_2 only strategy (herein referred to as DRA 5.0 ISRU) reduces the payload mass from Earth from 30 tons

(for CH_4 and LOX) to 7.5 tons: ~6.5 tons of CH_4 and ~1 ton of equipment for O_2 production.

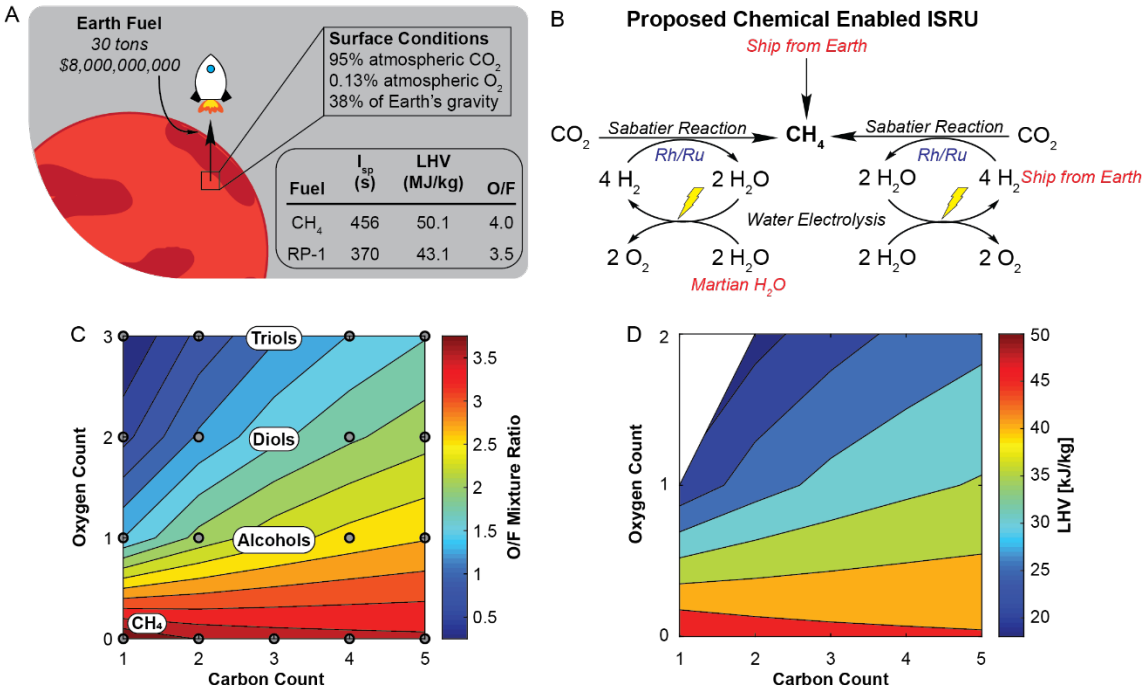

Figure 4.1.Chemical enabled ISRU fuel production (A) Martian ISRU for rocket propellant production can greatly reduce mission costs and take advantage of abundant Martian CO_2 (B) Current ISRU fuel strategies usually target CH_4 as the fuel of interest. CH_4 can be sent from Earth or produced using the Sabatier reaction using H_2 shipped from Earth or produced from harvested Martian H_2O using water electrolysis. The required LOX can come from O_2 produced during water electrolysis (resulting in excess CH_4) or supplemented with solid oxide CO_2 electrolysis (SOCE). (C) Adding oxygen atoms to the fuel compound reduces the ratio of mass of required LOX per mass of required fuel. An unoxygenated compound, ex. CH_4, requires 4 times as much O_2 as fuel, whereas oxygenated compounds, ex. diols, have a ratio around 2. (D) Adding oxygen atoms to the fuel compound reduces the lower heating value (LHV) from 50.1 kJ/kg for CH_4 to 25-30 kJ/kg for alcohols and 20-25 kJ/kg for diols.

The design of a Martian ISRU should take into account the differences between Mars and Earth, in particular gravity and atmospheric composition. Mars has 38% of Earth's gravity (3.73 m/s^2 and 9.81 m/s^2, respectively) [3], opening the doors for a more extensive array of chemicals to be considered as Martian rocket propellants due to a lowered energy density requirement. Further, the Martian atmosphere contains >20,000-fold less O_2 than on Earth (0.009 hPa on Mars and 210 hPa on Earth) accounting for only 0.13% of the Martian atmosphere [3]. This impacts rocket propellant design, as the amount of O_2 needed for fuel burning is dependent on the chemical composition of the fuel. Specifically, oxygen-containing compounds require less LOX than non-oxygen containing compounds (Figure 4.1c). For example, CH_4 with zero oxygen atoms requires a 4:1 O_2 to fuel mass ratio, whereas methanol and methanediol, with one and two oxygen atoms respectively, lower this ratio to 1.5:1 and 1:1. Further, oxygen-containing rocket propellants burn cleaner, potentially allowing engine reuse in multi-planetary missions. Although the presence of oxygen atoms in a compound lowers its heating value (Figure 4.1d), Mars' lower gravity enables some of these compounds to still be applicable as rocket propellants. Finally, Mars has 20 times more CO_2 than Earth (6.67 hPa and 0.38 hPa, respectively), accounting for ~95% of total atmospheric pressure, offering an excellent carbon source for hydrocarbon rocket propellant production.

Biological systems can convert CO_2 into chemicals, and ISRUs that leverage C1 metabolizing bacteria have been proposed for the production of fuels, materials, and astronaut nutrition [4]. For example, methanotrophs convert CO_2 to CH_4 at 96% purity with a productivity of 530 mmol/L/hr [5], meaning that a 5.66 L reactor provides the same hourly CH_4 productivity (48 g hr^{-1}) as the ~1.5 L Sabatier reactor currently in use on the

International Space Station [6]. Towards materials production, photosynthetic cyanobacteria that naturally accumulate polyhydroxyalkanoates could support *in situ* production of polymeric materials [7] from sunlight and CO_2, producing O_2 as a byproduct. Unlike the O_2 generated via SOCE, which produces carbon monoxide (CO) as a byproduct, biologically evolved O_2 does not need to be scrubbed for human life support.

Rather than having the C1 metabolizing microbe capture CO_2 and convert it to the final product, the autotroph can instead be grown as a feedstock for a heterotrophic microbe. Such a strategy enables the use of biosynthetic workhorses, such as *Escherichia coli* and *Saccharomyces cerevisiae*, which have been more extensively engineered for chemical production, to produce the Martian rocket propellant and other chemicals at high titers and yields [8]. *S. cerevisiae* can produce ethanol from cyanobacteria biomass without any additional nutrients [9, 10]. *E. coli* can produce 5.9 g/L of fusel alcohols with 57% yield from cyanobacterial sugars and 32% yield of cyanobacterial proteins [11]. Additionally, the shorter doubling time of heterotrophs (*S. cerevisiae*: 3 hours, *E. coli*: 30 minutes) when compared to autotrophs (cyanobacteria: 7 hours) allows the use of smaller cyanobacteria farms in conjunction with smaller and shorter fermenter production runs to produce the desired chemical quantities. After chemical separation, the heterotrophic microbe's biomass could be fed to the autotroph, maximizing nutrient use.

Here, we investigate the technical feasibility of producing a proposed Martian rocket propellant, 2,3-butanediol (2,3-BDO), using a biotechnology-enabled ISRU (bio-ISRU) strategy on Mars. Specifically, cyanobacteria capture Martian CO_2 and convert it to simple sugars and key nutrients, which an engineered *E. coli* converts to Martian rocket propellant in a continuous process. Key to this analysis are the constraints that a Martian

setting imposes, such as light intensity and water limitations. Our process model includes 1) the continuous cultivation of cyanobacteria, 2) the enzymatic digestion of cyanobacteria to recover the sugars, nitrogen and other trace nutrients, 3) the cultivation of *E. coli* for 2,3-BDO production from cyanobacterial glucose, and 4) the separation and purification of 2,3-BDO from the fermentation broth. When compared to the proposed DRA 5.0 ISRU, the bio-ISRU for rocket propellant production requires a 4.1-fold higher payload mass from Earth and only 1% higher power consumption, but generates 65 tons of excess O_2 that could be used for subsequent spacecraft launches or other aspects of Martian exploration, such as human colony formation, offering a distinct advantage over the chemical system. To make the technical viability of the bio-ISRU competitive with the proposed DRA 5.0 ISRU, doubling the cyanobacteria growth rate under Martian conditions, increasing the digester glucose release by 33%, increasing the *E. coli* yield of 2,3-BDO from glucose by 16%, and switching to lighter reactor component materials is needed.

4.3 Results

4.3.1 Challenges in producing rocket propellant on Mars

While fuel production unit operations are similar on Mars and on Earth, special considerations must be made for implementation on Mars. First, the average surface temperature on Mars is -55°C compared to 15°C on Earth [3]. While temperatures as low as -60°C do not affect microbial viability, microbes do not grow at those temperatures. Therefore, our analysis assumes a process temperature of 25°C, which can be achieved via a combination of reactor jacketing and implementation of a transparent dome-like structure to trap radiant solar heat akin to a greenhouse [12]. Second, the thinner Martian atmosphere

allows higher levels of ionizing and gamma radiation to reach the surface when compared to Earth. Fortunately, many cyanobacteria, including *Arthrospira platensis*, and *E. coli* are resistant to Martian radiation levels, maintaining viability for centuries [13, 14]. Of more concern is the level of UV radiation that reaches the Martian surface, which poses a risk for genetic mutations. To mitigate these risks, bags for cyanobacteria growth and/or the greenhouse dome should be made of a UV reflecting material while allowing transmission of photosynthetically active radiation (PAR). Importantly, the intensity of PAR that reaches the Martian surface is 57% lower than on Earth due to the increased distance from the sun [3]. This decrease in photon flux will result in a decreased cyanobacteria photosynthetic growth rate, thus biomass productivity is modeled using light as the growth-limiting factor. Third, biological processes require water. Although recent reports suggest that Martian water exists in sufficient amounts for the proposed bio-ISRU, harvested water may need to be pretreated to reduce salt content to enable microbial growth [15]. Fourth, microbial growth requires nitrogen, phosphorous, and trace metals in addition to carbon. While some trace metals (e.g. Mg, Ca, K and Na) could be obtained from Mars in a bioavailable form via electrochemistry, the majority of these nutrients will need to be shipped from Earth [16]. Of note, trace metals fed to the cyanobacteria will carry through to the hydrolysate fed to *E. coli* [17]. To complete the cycle, *E. coli* biomass can be fed back to the cyanobacteria after rocket propellant separation to effectively recycle the trace metals and nitrogen [18]. Fifth, Mars has a nitrogen (N_2) partial pressure of 0.189 hPa [3], 10-fold lower than that required for cyanobacterial N_2 fixation [19]. In our analysis, we have not considered N_2 fixation, and use a cyanobacteria, *A. platensis*, that does not fix N_2. Instead, we propose shipping diammonium phosphate ($(NH_4)_2PO_4$) and ammonia (NH_3) from

Earth. While implementation of a N_2 fixing cyanobacteria, such as *Anabaena cylindrica*, would reduce nutrient payload mass, it would likely require atmospheric N_2 concentration or cyanobacteria engineering to enable cyanobacterial N_2 fixation. Finally, the entire infrastructure, including nuclear power sources, must be shipped from Earth. This has direct implications for reactor design and whole process design as payload mass directly impacts mission cost. To account for these challenges, the proposed bio-ISRU analyzes water use, power requirement, and total infrastructure payload mass as key process metrics.

4.3.2 Martian rocket propellant design

Based on energy density, phase behavior and biological reachability, we propose short chain (C_3-C_4) hydrocarbons with two oxygen atoms (diols) as potential martial rocket propellants. As shown in Figure 4.1c, oxygen content within the propellant structure reduces the amount of LOX needed for combustion, i.e. the stoichiometric oxygen/fuel mixture ratio. A lower ratio is advantageous on Mars due to its low atmospheric O_2 content, making oxygen-containing compounds (e.g. alcohols, diols) preferable for Martian application. Compared to alcohols and C_1-C_2 diols, C_3-C_4 diols have the required energy content (> 14 MJ/L) and specific impulse (I_{sp}) to propel a human Mars Ascent Vehicle (MAV) with lower oxidizer need. Triols require even less oxidizer, however, their energy content is too low to propel a human MAV. Additionally, short chain diols have appropriate boiling and melting points to remain liquid or solid over Martian temperature (-153°C to 20°C), and thus have lower volumetric and energetic storage requirements (Table 4.1). Finally, towards bioreachability, several natural and engineered microbes produce short-chain diols in high titers and yields [20], whereas longer chain diols are only produced in

low yields and titers (C5) [21] or are undemonstrated to date (>C5). Specifically, we considered 1,2-propanediol (C3) [22], 1,3-butanediol (C4) [23], and 2,3-butanediol (C4) [24] as Martian rocket propellants, all of which can be biologically synthesized today. Of note, when compared to the proposed rocket propellant methane, diols require less than half the LOX mass per mass of fuel. Because the diols will be liquid or solid on Mars, the energy needed to compress, or refrigerate, gaseous methane to a liquid for storage and propulsion is sidestepped.

Table 4.1. Design criteria for proposed Martian rocket propellants

Propellant Candidate		**Boiling Point (°C)**	**Melting Point (°C)**	**I_{sp} (s)**	**O/F (kg/kg)**	**LHV (MJ/kg)**
RP-1	$C_{12}H_{26}$	216	-10	370	3.48	43.1
Methane	CH_4	-162	-187	456	4	50.1
1,2-Ethanediol	$C_2H_6O_2$	197	-12	395	1.29	18
1,3-Propanediol	$C_3H_8O_2$	214	-27	421	1.68	22.9
1,3-Butanediol	$C_4H_{10}O_2$	207	-50	429	1.96	26.2
2,3-Butanediol	$C_4H_{10}O_2$	184	19	429	1.96	26.2

4.3.3 2,3-butanediol as a Martian rocket propellant

To assess the bioreachability of each of the diols of interest, we calculated the *E. coli* maximum theoretical yield from glucose using the Constraint Based Reconstruction and Analysis toolbox [25] and the iML1515 *E. coli* genome scale model [26]. Theoretical yields for 1,2-PDO, 1,3-BDO, and 2,3-BDO are 0.615, 0.543, and 0.538 g diol/g glucose, respectively. Although 1,2-PDO has the highest theoretical yield, its lower theoretical I_{sp}

and lower O_2:fuel ratio means 1,2-PDO requires ~3 tons more glucose mass as 1,3-BDO and 2,3-BDO to produce sufficient propellant for a MAV launch (Table 4.2). Due to its higher technology readiness level, we focus the analysis on 2,3-BDO production, which has been produced in *E. coli* with a yield of 0.432 g/g glucose, titer of 73.8 g/L, and productivity of 1.17 g/L/hr [27]. The experimental yield of 2,3-BDO is 2.4- and 3.9-fold higher than 1,2-PDO and 1,3-BDO, respectively, and reaches 80% of the theoretical maximum [22, 23, 27]. Thus, production of 2,3-BDO leads to a reduced cyanobacteria farm footprint, a smaller microbial fermenter size, and less fermentation byproducts, streamlining the proposed rocket propellant purification. Experimental yields of 1,2-PDO (0.178 g/g glucose [22]) and 1,3-BDO (0.11 g/g glucose [23]) are likely limited by cofactor balancing challenges. Specifically, production of one 2,3-BDO generates one NADH, whereas production of one 1,2-PDO or one 1,3-BDO results in the net consumption of one NADH (Figure 4.2). Thus, the NADH produced during glycolysis is sufficient to drive 2,3-BDO synthesis without cofactor regeneration. Additionally, separations processes for 2,3-BDO from fermentation broth are well established [28] whereas separations process for 1,2-PDO and 1,3-BDO are still being explored.

Table 4.2. Bioreachability of short chain diols

Compound	Maximum Experimental Yield (g/g)	Maximum Theoretical Yield (g/g)	Propellant Required (tons)	Glucose Required at Max Theoretical Yield (tons)
1,2-Propanediol	0.178 [22]	0.615	14.84	24.13
1,3-Butanediol	0.11 [23]	0.543	11.39	20.97
2,3-Butanediol	0.432 [27]	0.538	11.39	21.17

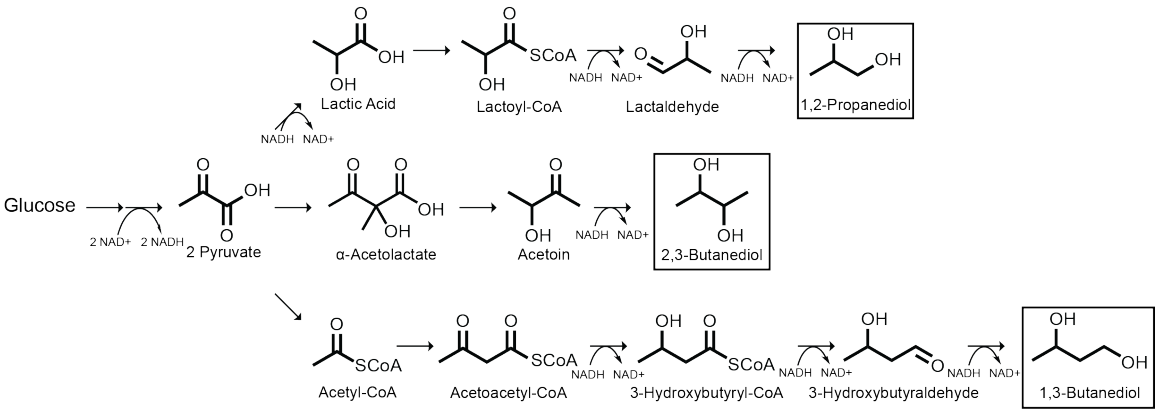

Figure 4.2. Metabolic pathways for short chain diols. Modeled pathways for conversion of glucose to 1,2-propanediol, 2,3-butanediol, and 1,3-butanediol in *E. coli*.

To set the mass production target for 2,3-BDO we used the ideal rocket equation and the calculated theoretical I_{sp} values. This equation indicates that the total required propellant (fuel plus LOX) varies by only 3% if using 2,3-BDO instead of methane. Based on the calculated I_{sp} and oxygen fuel ratios for 2,3-BDO, we require 11.39 tons of 2,3-BDO to power the MAV. To account for variation in actual I_{sp} values versus theoretical, we have increased this estimate to 15 tons as an upper end goal. Thus, the determined process metrics are based off of producing 15 tons of 2,3-BDO, requiring 29.4 tons of LOX, during a long-stay, 500-day human surface visit to Mars mission.

4.3.4 Process design overview

The bio-ISRU of 2,3-BDO involves four modules: 1) photosynthetic Martian CO_2 fixation to produce simple sugars from sunlight, water, and nitrogen using cyanobacteria, 2) cyanobacteria biomass concentration and enzymatic digestion to release sugars and trace nutrients into a usable form for the downstream microbe, 3) microbial fermentation to

convert the released sugars into 2,3-BDO, and 4) extraction and separation of 2,3-BDO to ~95% purity from the microbial broth (Figure 4.3).

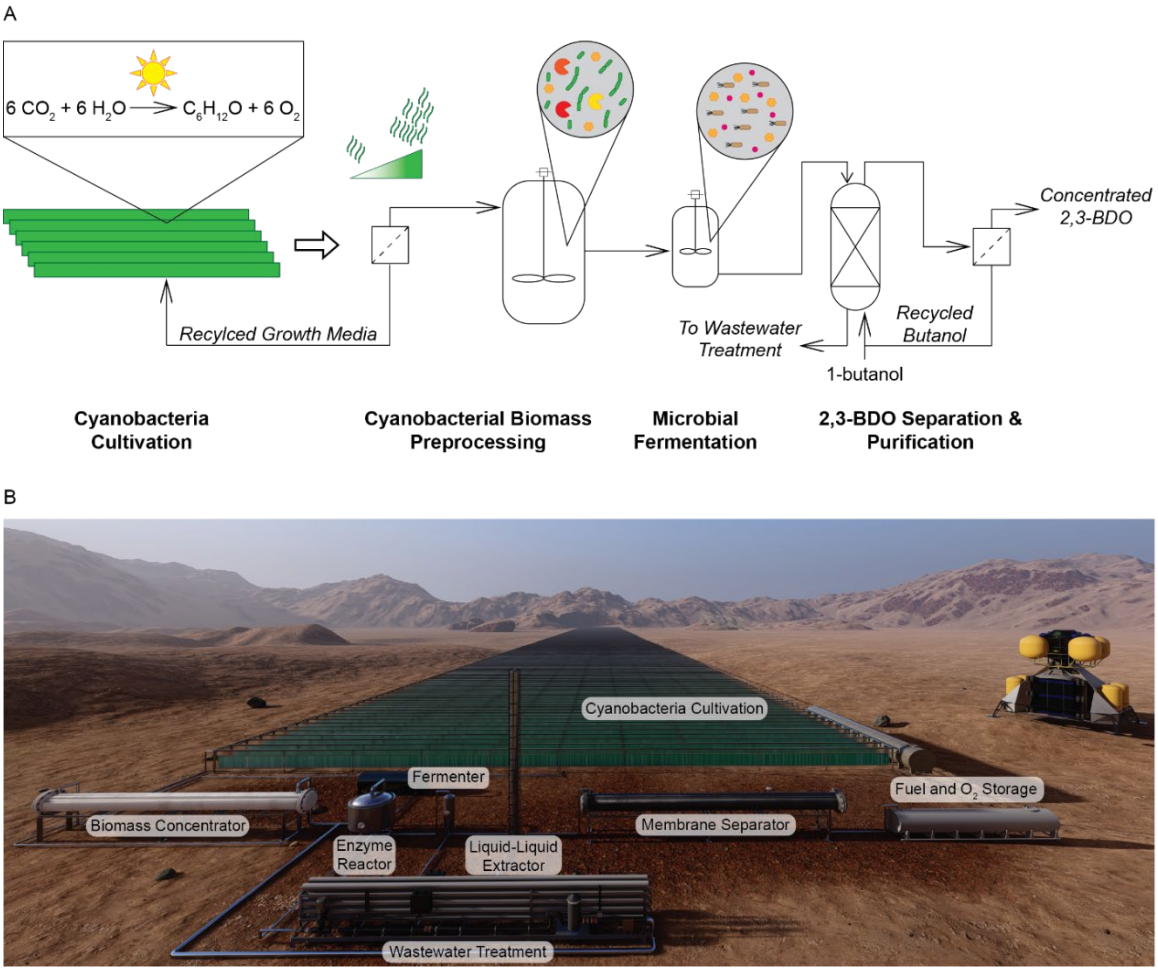

Figure 4.3. Process overview for continuous production of 2,3-BDO from cyanobacterial glucose. (A) The 2,3-BDO production process requires four modules. First, cyanobacterial cultivation for biomass production in photobioreactors or biofilm growth. Second, biomass preprocessing consisting of biomass concentration through membrane filtration and enzymatic breakdown in a stirred tank. Third, microbial fermentation of the cyanobacterial glucose to produce 2,3-BDO. Fourth, 2,3-BDO separation and purification using sequential liquid-liquid extraction and membrane separation. (B) A rendering of what the process might look like on Mars, with a Martian rover included for scale. Importantly, the cyanobacterial cultivation modules makes up the bulk of the material and land footprints.

4.3.5 Cyanobacteria growth

A. platensis (i.e. spirulina) will be used to fix Martian CO_2. In *A. platensis*, nitrogen limitation maximizes carbohydrate production, with glycogen reaching up to 60% of dry cell weight [9]. *A. platensis* grows well at 0.38 hPa of CO_2, thus the 6.67 hPa CO_2 present on Mars will enable *A. platensis* growth. In fact, higher CO_2 levels have been shown to improve cyanobacteria growth [29]. Beyond CO_2 concentration, a process concern is the negative effect of the <10 hPa total atmospheric pressure on Mars versus 1013 hPa on Earth. Elevated CO_2 levels help to mitigate this effect, with 50 hPa of CO_2 in a 100% CO_2 atmosphere showing uninhibited growth [30]. However, the Martian atmosphere still has 5-fold lower pressure than previously studied. Therefore, CO_2 pressurization may be needed to achieve optimal growth, and further study at these low pressures is needed to determine the full effects and validate our assumption that growth is indeed light limited and not CO_2 limited.

Cyanobacteria culturing requires addition of nitrogen, phosphorous, and trace elements (Ca, Cu, Fe, K, Mg, Mn, Na, and Zn) [31]. Of these, nitrogen and phosphorous will account for the largest contribution to payload mass. To reduce mission risk, all the required nitrogen and phosphorous to produce enough cyanobacterial biomass to reach the 15 tons of 2,3-BDO target will be shipped from Earth. Discussion about reducing payload mass through nitrogen and phosphorous recycling is included in the discussion below. Calculation of required nitrogen and phosphorous mass was based on the elemental composition of cyanobacteria ($C_{4.5}H_{8.2}N_{0.129}O_{2.219}S_{0.006}P_{0.007}$ [32]) and nutrients will be provided in 20% excess to allow for adequate uptake concentrations in the growth media [33]. Under these requirements 1.12 tons of $(NH_4)_2PO_4$, and 1.81 tons of anhydrous NH_3

are included in payload mass calculations. Alternative sources of nitrogen and phosphorous include fixing N_2 in situ via sunlight driven chemical catalysis [34] or harvesting phosphorous from Martian regolith [35, 36]. However, due to the crucial nature and large requirement of these nutrients, shipping from Earth helps to lower mission risk. Trace elements are provided based on their previously reported minimum requirements to support unhindered cyanobacterial growth and are recycled through the process [31].

4.3.6 Cyanobacteria growth modes

While Martian water is sufficient to sustain the bio-ISRU, harvesting and keeping the water in liquid form will require 1-2 kW per ton of water [2]. With this in mind, we explored two cyanobacteria cultivation methods: water-intensive suspended growth and less water-intensive biofilm growth (Figure 4.4a). For suspended growth, we propose using low density polyethylene (LDPE) photobioreactors (PBRs) [37], which lower contamination threats to the Martian environment, have low evaporation rates, allow growth rates to be highly controlled through mixing, enable CO_2 delivery via gas bubbling, and can be shipped in rolls to be inflated upon arrival. On Earth, suspended growth is used for the industrial production of cyanobacterial biomass for food or biodiesel [32]. Suspended growth rate and productivity were modeled using Monod growth kinetics based on lab- and pilot-scale growth of *A. platensis* [38]. A key assumption to the model is that growth is limited by light penetration rather than CO_2, nitrogen, or phosphorous sources (Figure 4.4b). This is a common assumption for modeling cyanobacterial growth on Earth [38], and is valid on Mars where photon flux is 57% lower than on Earth due to the increased distance from the sun [3].

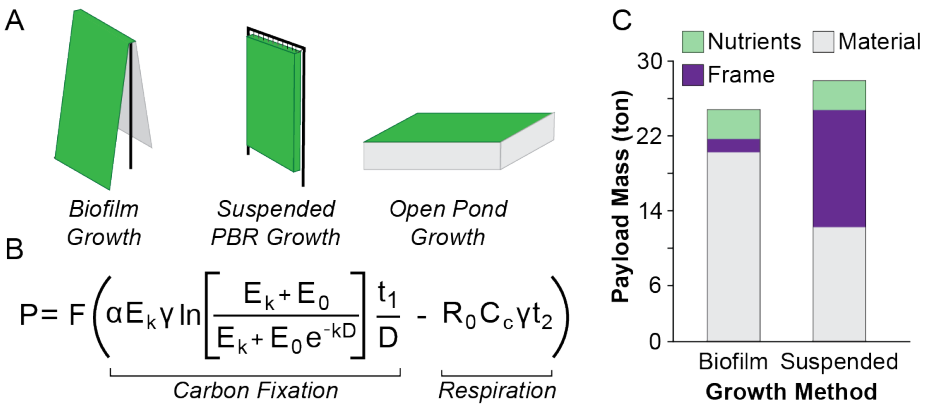

Figure 4.4. Comparison of biofilm and suspended cyanobacterial growth methods. (A) Representation of biofilm growth, where the cyanobacteria grow in a thin film on a porous substrate, suspended photobioreactor (PBR) growth where the cyanobacteria are grown in PBRs filled with water, and open pond growth where the cyanobacteria are grown in large horizontal, ponds open to the atmosphere. (B) Cyanobacterial growth model developed by Karemore *et al*. [38] to describe growth of *Arthrospira platensis* based on a light limited Monod growth model. The equation gives daily, areal productivity, taking into account both sunlight driven photosynthetic carbon fixation as well as light-independent respiration. (C) Total payload mass of the cyanobacterial cultivation unit for biofilm and suspended growth, taking into account the growth nutrients, support frame and growth substrate (biofilm) or PBR material (suspended).

Parameters for the model were determined as follows (Table 4.3). First, light penetration was maximized by spacing the 4.5 cm wide (model parameter D) PBR units 1 m apart [32] and using hanging bag PBRs instead of a horizontal system, increasing the ratio of light absorption area (PBR surface area) to land area (model paramter F) to 1.84 versus 1 for a horizontal system. The F ratio could be increased up to 15 by reducing the PBR spacing or increasing the PBR height, at the expense of reduced light intensity at the PBR surface due to shading from surrounding bags [39]. Second, biomass concentration is modeled at 1 g/L, the same as used on Earth, to allow for maximum light penetration into the culture (model parameter k). Finally, environmental (model parameters E_0, t_1, and t_2) [3] and species specific growth parameters (α, γ, E_k, R_0, and C_c) [38] were taken from

literature resulting in a modeled biomass productivity of 6.54 g/m^2/day, in terms of land area footprint of the cyanobacteria farm to allow for use of published power and mass correlations [32]. The cyanobacterial cultivation will be run as a continuous reactor where the calculated biomass productivity was used to determine the rate of constant biomass harvest to maintain the culture at the desired concentration of 1 g/L. The calculated productivity requires processing of 0.157 tons of biomass per day to reach our production goal and timeline. On Earth, cyanobacteria have growth productivities between 20 and 30 g/m^2/day at temperatures from 20-30°C for both suspended and biofilm growth [32, 40].

Table 4.3. Cyanobacteria growth model parameters

Model Parameter	Parameter Definition	Parameter Value Biofilm	Parameter Value Suspended	Source
α	moles carbon fixed per photon	0.061 mol C*mol photons^{-1}		[38]
E_k	photosaturation parameter, photon flux of max biomass growth (defined @ 25°C)	161.01 µmol photons*m^{-2}*s^{-1}		[38]
E_0	photon flux at reactor surface	142.8 µmol photons*m^{-2}*s^{-1}		[3]
k	light penetration parameter	8750 m^{-1}	175 m^{-1}	[38]
D	reactor diameter	0.00025 m	0.04 m	This work
γ	g biomass per mole of carbon fixed	22.2 g DCW*mol C^{-1}		[32]
t_1	time of light exposure, photosynthesis time	43200 s	43200 s	This work
R_0	carbon consumed in maintenance per mass of chlorophyll (defined @ 25°C)	0.001395 µmol C*mg Chl^{-1}*min^{-1}		[38]
C_c	mass of cholorphyll per reactor volume	135 mg Chl*m^{-3}	18000 mg Chl*m^{-3}	[38]
t_2	total time in unit of interest	24600 s	24600 s	This work

While cyanobacterial biofilm growth holds promise in improved gas delivery, light penetration, reduced water usage, and reduced energy demand due to elimination of culture mixing, it has only been performed in lab or pilot scale facilities, not industrially [40]. To support biofilm growth, we modeled a hydrophilic, porous, growth substrate [41], resembling a cotton towel [40], 0.3 mm thick to match the thickness of the LDPE PBR bags [37]. To model biofilm productivity, the same published growth model used for suspended growth was used. Only light penetration (model paramter k), film thickness (model parameter D), and areal chlorophyll concentration (model parameter Cc), which depend on the increased biomass concentration in biofilm growth (7.5 g/m2 versus 1 g/L) were changed. Other growth parameters, including photosynthetic efficiency (model paramter α), respiration rate (model paratmer R_0), and optimal photon flux (model parameter E_k), were assumed to be species dependent, and thus indpendent of growth method and were used unchanged. For direct comparison, the 1 m spacing and F ratio of 1.84 were also applied to biofilm growth, resulting in a modeled productivity of 6.64 $g/m^2/day$, similar to the value modeled for suspended growth.

Based on the modeled productivities, cyanobacteria biofilm growth reduces water consumption by 90% when compared to suspended growth: from 997,568 L to 94,933 L, based on required PBR volume and a biofilm requirement of 2.18 L/m2 of growth substrate [42]. This results in energy savings of up to 1800 kW based on previously reported metrics for water harvesting via soil evaporation [2]. From a payload mass perspective, the biofilm system mass is 11% less than that of the suspended system due mainly to the increased mass of support frame needed to hold the 10-fold greater mass of water, even considering the reduced system weight under Martian gravity. Of important consideration, the cotton

for biofilm growth substrate accounts for the majority, 82%, of the cyanobacterial growth system mass (Figure 4.4c, Table 4.4).

Table 4.4. Mass of cyanobacteria growth system components.

Growth Method	Growth Substrate/PBR Material	PBR Mass (tons)	Frame Mass (tons)	Nutrient Mass (tons)
Biofilm	0.3 mm thick cotton fabric	20.26	1.43	3.13
Suspended	0.3 mm thick LDPE [37]	12.25	12.53	3.14

4.3.7 Cyanobacteria biomass concentration

Cyanobacteria concentration prior to enzymatic digestion for sugar release was modeled using cross-flow filtration due to its lower payload mass when compared to centrifugation equipment, and more rapid and efficient separation than settling tanks [43]. A hydrophilic and neutrally charged membrane concentrates the biomass from 1g/L in the PBR to 20g/L for input into the enzymatic digester. The membrane is modelled to have a 40 kDa cut-off [44, 45] with long-term membrane flux approaching 40 L/m2/h and shown to be stable over the course of 6 weeks [46]. Cyanobacterial biomass collection from biofilm growth was set at a continuous 20 g/L, eliminating the need for a concentration unit operation for biofilm cultivation mode [9].

4.3.8 Cyanobacteria biomass digestion

Lysozyme, α-amylase, and glucoamylase digestion releases between 0.3 and 0.45 g of glucose/ g of biomass over 24hr and 48hr, respectively [9]. While a shorter digestion

time requires a smaller enzymatic reactor volume and reduced water usage, it also requires a larger cyanobacteria biomass to achieve the desired sugar output, thus a larger cyanobacteria farm. Increasing biomass digestion from 24 to 48 hours reduces total cyanobacterial biomass production by ~34%, leading to a ~30% reduction in water requirement and payload mass of the cyanobacterial cultivation unit for both biofilm and suspended growth systems (Figure 4.5, Table 4.5). Taken together, decreasing the size of the cyanobacterial growth operation is beneficial, even at the expense of increased size of the biomass digester.

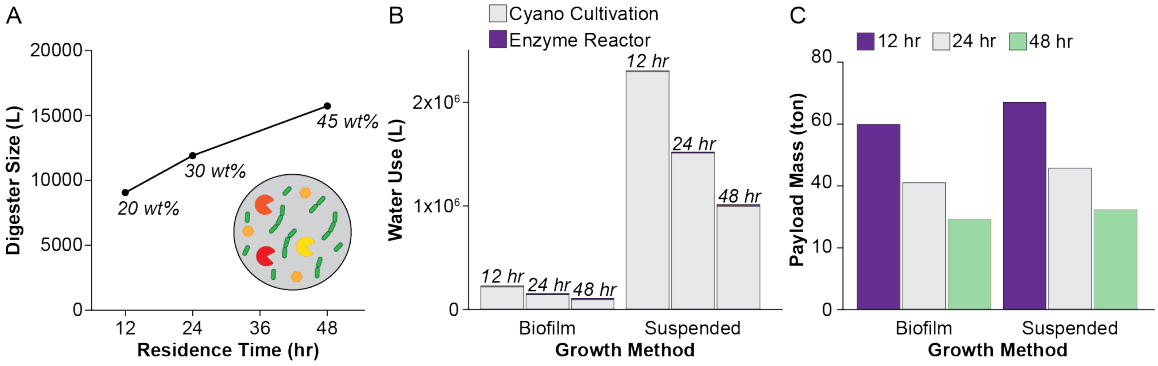

Figure 4.5. Optimization of digester residence time. (A) The size of the enzyme digester depends on the residence time, which also determines glucose yield from cyanobacterial biomass. (B) Increasing digester reactor time decreases process water used and (C) payload mass for both biofilm and suspended growth modes.

Table 4.5. Effect of enzyme digester size on water usage and payload mass.

Growth Method	Enzyme Treatment (hrs)	Cyano Land (m^2)	Cyano Water (L)	Enzyme Water (L)	Total Water (L)	Cyano Growth System Mass (tons)
Biofilm	12	54533	218743	9053	2.28×10^5	59.93
	24	35859	143838	11905	1.56×10^5	41.05
	48	23682	94993	15725	1.11×10^5	29.20
Suspended	12	55387	2297010	9053	2.31×10^6	67.04
	24	36422	1510493	11905	1.55×10^6	45.73
	48	24054	997568	15725	1.01×10^6	32.29

4.3.9 Microbial production of 2,3-butanediol on Mars

E. coli was modeled to convert cyanobacteria to 2,3-BDO as hydrolyzed cyanobacteria has previously been used as the only feedstock for chemical production in *E. coli* [11]. Furthermore, as Mars is devoid of life, using a non-pathogenic microbe reduces biosafety contamination concerns. Finally, *E. coli* has been engineered to produce 2,3-BDO, and for process design we used a steady state productivity of 1.17 g 2,3-BDO/L/hr and a yield of 0.432 g 2,3-BDO/g glucose, matching the *E. coli* state of the art [27]. Of note, O_2 is important for NADH regeneration during 2,3-BDO production [27]. Cyanobacterial cultivation produces excess O_2 to meet the ~18 ton requirement of the fermentation unit over the course of 500 sols in addition to the O_2 required for rocket propellant LOX.

4.3.10 Rocket propellant extraction and separation

To purify 2,3-BDO from the fermentation broth, we modeled liquid-liquid extraction (LLE) followed by membrane-based separation. In the LLE unit, hydrophilic 2,3-BDO is extracted from fermentation broth using butanol, a previously validated solvent [47], which does not theoretically reduce fuel performance if not fully removed from the final product based on similar heating value (33.9 MJ/kg) and specific impulse (439 s). Liquid-liquid extraction was modeled using the Phasepy python package [48] with binary interaction parameters regressed in AspenTech. To concentrate and dewater 2,3-BDO, a polydimethylsiloxane/polyvinyledenefluoride membrane was modeled using published flux differential equations [47]. Our model indicated that 0.5 m^2 of membrane area is required to reach 95% 2,3-BDO purity. Thus, the membrane separations unit accounts for

<<1% of total payload mass. As butanol must be shipped from Earth, the process includes a solvent recycle stream to minimize the overall solvent input into the system. In order to reduce water usage, the water rich stream coming out of the LLE unit will be purified and fed back to the *E. coli* fermenter. An important consideration will be the amount of butanol lost in this stream, as lost butanol will increase payload mass and introduce toxicity risks to the *E. coli*. Our analysis shows that the butanol concentration in water outlet stream will be ~0.01%, however careful consideration must be taken to prevent accumulation. Finally, to mitigate this risk, *E. coli* could be engineered for improved butanol tolerance [49].

4.3.11 Water recycling

Although the amount of water on Mars will be sufficient to support a Martian colony, as well as the bio-ISRU of 2,3-BDO [15], maximization of process efficiency and integration requires wastewater management and recycling. Implementation of reverse osmosis for wastewater purification will prevent buildup of salts and proteins over the course of the process. The concentrated waste stream can be reintroduced to the process in order to recycle trace nutrients, while the remainder of the rejected waste can be utilized in other ISRU applications, such as for fertilizer, as the bio-ISRU does not produce any toxic compounds. For this application, a reverse osmosis membrane with long-term stability and a lightweight material was modeled. Specifically as a thin-film composite membrane, 200 μM thick with water flux of 25 $L/m^2/hr$ [50]. Based on the flowrate of 5.46 L/min, a membrane area of 13.1 m^2 is required, bringing the mass of the required water treatment unit to 0.008 tons, 0.03% of total payload mass.

4.3.12 Full process analysis

Holistic analysis of the process was performed to determine water usage, power, and total payload mass requirements (Figure 4.6, Table 4.6). Of note, cyanobacteria biofilm growth requires 89% less water, 66% less power, and 11% less payload mass, a difference of 3.98 tons. The stark difference in energy consumption between biofilm and suspended growth (26.44 kW vs 77.60 kW) comes from the elimination of mixing during biofilm growth, which accounts for ~25% of the power usage even with reduced power demand due to reduced Martian gravity, and elimination of the biomass concentration unit operation, which uses 32.76 kW (42% of total). If the energy for water harvesting is included in the calculations, biofilm growth requires 9.6-fold less power than suspended growth.

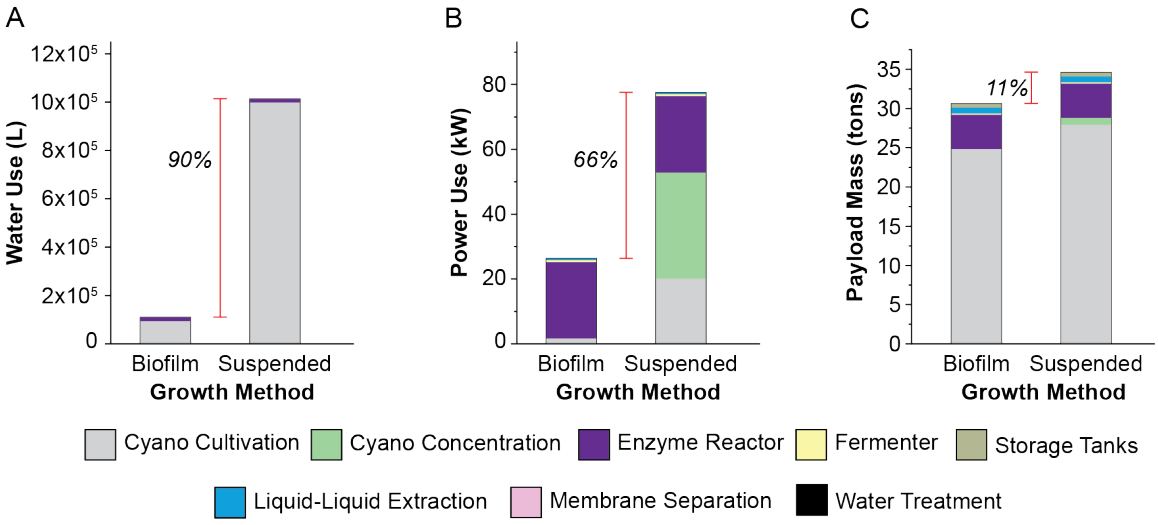

Figure 4.6. Process metrics for 2,3-BDO production from cyanobacterial glucose. Comparison of (A) water, (B) power, and (C) mass requirements of biofilm and suspended growth modes, accounting for all unit operations in the proposed process. Results are based on 15 tons of 2,3-BDO produced over 500 sols.

Table 4.6. Water, power, and payload mass requirement for bio-ISRU

Unit Operation	Water Requirement (L)		Power Requirement (kW)		Mass Requirement (tons)	
	Biofilm	**Suspend**	**Biofilm**	**Suspend**	**Biofilm**	**Suspend**
Cyano Cultivation	9.50×10^4	9.98×10^5	1.69	20.09	24.83	27.92
Cyano Concentration	N/A		N/A	32.76	N/A	0.89
Enzyme Reactor	15725		23.59		4.37	
Fermenter	393		0.59		0.17	
Liquid-Liquid Extraction	N/A		0.49		0.74	
Membrane Separation	N/A		0.075		6.23×10^{-4}	
Storage Tanks[a]	N/A		N/A		0.55	
Water Treatment	N/A		Not calculated		0.008	
Total	1.11×10^5	1.01×10^6	26.44	77.60	30.66	34.64

When compared to the DRA 5.0 ISRU, the bio-ISRU for rocket propellant production requires 4.1-fold greater payload mass and 1% greater power usage (Table 4.7). Importantly, the bio-ISRU produces 65.8 tons of excess O_2 over 500 sols (500 Mars days). This excess O_2 is ready for use in other applications, such as subsequent aircraft launches or crew life support systems. For context, a crew of 6 consumes 2 tons of O_2 for life support (e.g. breathing) over the course of 500 sols [2]. To expand upon this advantage, improvements in the bio-ISRU that reduce payload mass and power use are needed.

Table 4.7. Comparison of bio-ISRU with NASA DRA 5.0 ISRU

Method	Fuel	Excess O_2 (tons)	Payload Mass (tons)	Power Required (kW)	Source
NASA DRA 5.0 ISRU	CH_4	0	7.51	26.08	DRA 5.0
Biotechnology-enabled ISRU	2,3-BDO	65.82	30.66	26.44	This work
Optimized Biotechnology-enabled ISRU	2,3-BDO	30.56	9.32	16.15	This work

4.4 Discussion

4.4.1 Biological process improvement strategies

To determine engineering targets for improving the bio-ISRU, we focused on three biological strategies: 1) increasing cyanobacterial biomass productivity, 2) increasing glucose yield in the enzyme digester, and 3) increasing 2,3-BDO yield in the fermenter. As 81% of the payload mass comes from cyanobacterial growth, we first focused on improving cyanobacterial biomass productivity, currently at 6.64 g/m^2/day. Using process modeling, we analyzed water, power, and payload mass requirements for cyanobacterial growth productivities between 5 g/m^2/day and 15 g/m^2/day (Figure 4.7a). Improving productivity to 15 g/m^2/day reduces water usage by 48% and payload mass by 40%, but only reduced the modeled power consumption by 4%. This is expected as the enzyme digester dominates power consumption, and is relatively unchanged by cyanobacterial productivity.

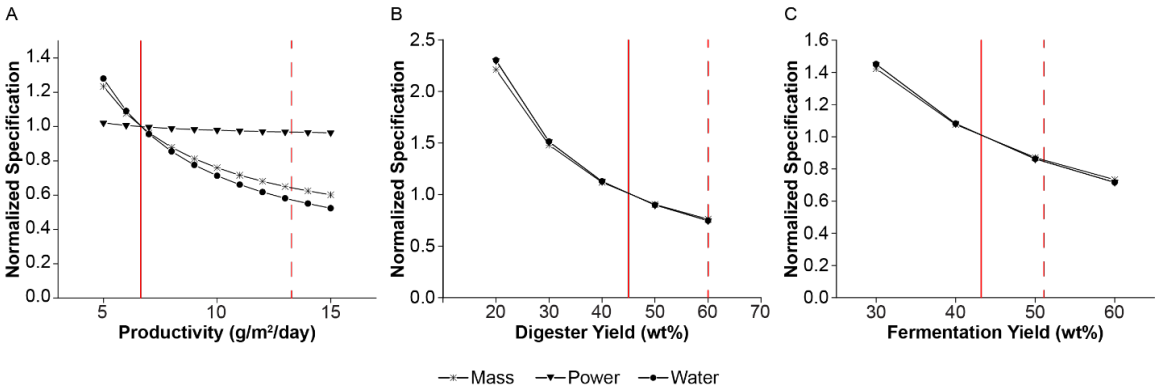

Figure 4.7. Impact of biological process improvements on process specifications. Payload mass, power use, and water use can all be reduced by biological improvements in (A) cyanobacterial growth productivity, (B) the yield of glucose from biomass in the enzyme digester, and (C) the yield of 2,3-BDO from glucose in the fermenter. Solid vertical lines represent productivity and yields used in the unoptimized model, dashed vertical lines represent engineering targets used in the optimized model.

Increasing cyanobacterial biomass productivity could be achieved by a) improving light input by supplementing sunlight with artificial lighting, b) improving cyanobacteria photosynthetic efficiency, e.g. by utilizing the wider range of the light spectrum that reaches Mars, or c) increasing the growth rate of cyanobacteria by adapting *A. platensis* to grow faster at 25ºC (35°C optimal) or using a cyanobacterial species with faster growth rate at low temperatures. If artificial lighting is used to double the photon flux, productivity is increased by 55% to 10.27 g/m^2/day. Nevertheless, this adds 250 tons of payload mass and 1700 kW of power due to the large land footprint of the cyanobacterial growth operation based on calculations for solar powered artificial lighting [51]. Integration of organic light emitting diodes (LED) [52] powered by photovoltaics [53] into the dome surrounding our process could provide a light-weight, future strategy for increasing light flux to the cyanobacteria, though the technology requires further developments. Thus, improving cyanobacteria photosynthetic efficiency is a more viable strategy. To model this,

we varied the α parameter in the cyanobacterial growth model which represents the moles of carbon fixed per mole of photons the cell receives. The theoretical maximum for this parameter is 0.125, representing 8 moles of photons per mole of fixed carbon. The current parameter value, 0.061, is 49% of the theoretical maximum. However, improvement in photosynthetic efficiency alone is not enough, as we would need to reach an α value of 0.117 (93% of theoretical max) to achieve a doubling in biomass productivity and improving photosynthesis in any autotrophic organism has been an difficult task [54]. Yet, strategies such as RuBisCo overexpression and reducing the size of the light harvesting antennae to reduce over-absorption of light energy have achieved ~55% improvements in cyanobacterial photosynthesis rate and biomass production, respectively [55].

The second biological improvement strategy focuses on the enzyme digester. As modeled, the enzymatic digestion has a 45% yield by weight of total biomass or 75% yield on total glucose [9]. Increasing the enzyme loading or implementation of an undigested biomass recycle stream, could increase the yield to take advantage of the additional available glucose. Assuming the enzymatic digestion could reach 60 wt% of total digested biomass (Figure 4.7b), payload mass and power requirement of the entire process would be reduced by 25%. Alternatively, cyanobacteria could be engineered to more easily release sugars by storing them as soluble sugars such as glucose, sucrose, glucosylglycerol, or glucosylglycerate rather than glycogen [56]. Previously, cyanobacteria have been engineered to secrete glucose [57]. Such an approach would eliminate the enzymatic digestion step and the need to ship digestive enzymes from earth. Another approach would be to utilize stored fatty acids instead of stored sugars. Cyanobacteria can accumulate fatty

acids around 40-45% of dry cell weight [58] and those fatty acids could be extracted and fed to *E. coli* or used as feedstock for chemical biodiesel production [59].

The final biological strategy, improving yield for 2,3-BDO from glucose in the microbial fermenter more limited room for improvement. The state-of-the-art yield, 0.432 g 2,3-BDO/ g glucose, is 80% of the theoretical maximum, 0.54 g/g, determined by flux balance analysis. Even so, improving yields by 18% to reach 95% of theoretical yield would reduce water, power and mass requirements by around 16% for all metrics (Figure 4.7c).

Taken altogether, doubling cyanobacterial biomass productivity to 13.28 g/m^2/day, increasing enzyme yield to 60 wt%, and increasing 2,3-BDO yield to 95% of theoretical yield reduces power consumption by 39% to 16.15 kW, 38% less than the DRA 5.0 ISRU. Furthermore, in this biological optimized scenario water use and payload mass are reduced by 64% and 57%, respectively, due to the 69% reduction in cyanobacterial farm size. The payload mass for the biological optimized, bio-ISRU is 13.08 tons, which is still 74% higher than the DRA 5.0 ISRU (Figure 4.8). Thus, a materials optimization is needed to make the bio-ISRU competitive with the DRA 5.0 ISRU.

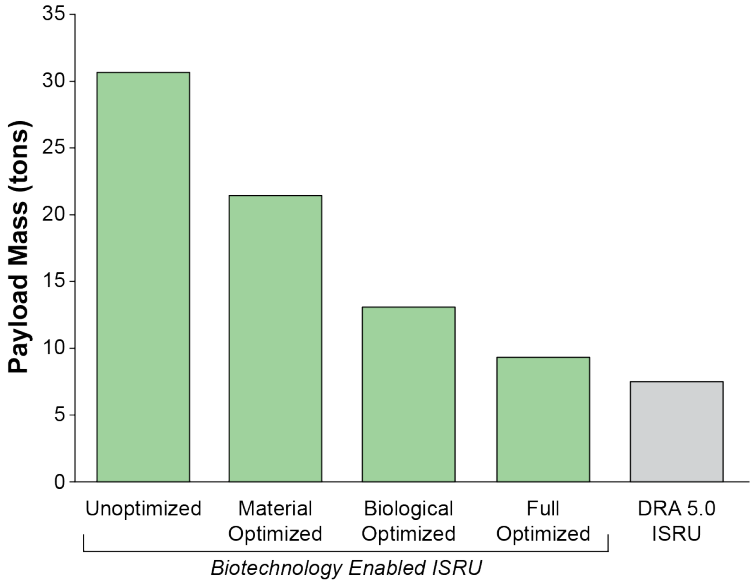

Figure 4.8. Process optimization to reduce payload mass. Biological and material optimization can reduce payload mass of the biotechnology-enabled ISRU process. Unoptimized is the original modelled scenario, material optimized uses HDPE and LDPE for reactor components instead of steel and cotton. Biological optimized improves cyanobacterial biomass productivity 2-fold to 13.28 g/m2/day, improves digester yield to 60 wt%, and improves fermenter yield to 95% of theoretical max. Full optimized implements both material and biological optimizations. The DRA 5.0 ISRU is the mass of shipping CH_4 as fuel from Earth along with process equipment to produce O_2 via SOCE in situ. This payload mass comes from NASA DRA 5.0 [2].

4.4.2 Materials process improvement strategies

To supplement the biological improvements, two key changes in materials are proposed. First, the cyanobacterial biofilm growth substrate accounts for 66% of the total payload mass. Replacing the cotton-like material with a material with similar density to the LDPE used for the PBR bags would reduce payload mass by 23% to 23.47 tons. Second, in the initial model, steel was chosen as the material for the microbial fermenter and other unit operations based on Earth requirements, including ease of sterilization and structural stability. On Mars, the lower gravity should allow for better performance of lighter weight materials. For example, using fermenters made of aluminum or high-density polyethylene

would further reduce process mass to 21.94 and 21.43 tons, respectively, if implemented along with the lighter biofilm growth substrate.

A process design strategy to reduce payload mass is to improve recycling of nitrogen and phosphorous. Our model indicates that in the biologically optimized scenario, the cyanobacterial growth will consume 1.82 kg of nitrogen and 0.28 kg of phosphorous per day. A three day supply of nutrients, giving ample time for nitrogen and phosphorus to recycle through the enzyme digester, fermenter, and water treatment unit operations reduces diammonium phosphate and ammonia requirement from 1.85 tons to 0.011 tons.

The fully optimized bio-ISRU will require both the biological and materials improvements suggested. Once implemented, the payload mass will be reduced to 9.32 tons, making the payload mass of the bio-ISRU only 24% greater than the DRA 5.0 ISRU (Figure 4.8). Importantly, the power usage of the optimized bio-ISRU would be 16.15 kW, 38% lower than the DRA5-ISRU (Table 4.7). In addition, the improved bio-ISRU still produces 30.6 tons of excess O2, something not achievable with the DRA 5.0 ISRU and a clear advantage of this system to aid Mars colonization. Finally, the lifetime of the majority of the process components is between 10-20 years under Earth conditions [60], allowing for reuse over multiple missions. If nutrients and butanol solvent can be successfully recovered, the only resupply needed will be the biofilm growth substrate which has a lifetime of 1-2 years [37], meaning subsequent missions will only require 4.12 tons of payload mass under the optimized conditions.

4.5 Conclusions

Taking into account the differences between Mars and Earth will be pivotal to developing efficient technologies for the ISRU production of fuel, food and chemicals on Mars. By keenly considering these differences, we propose that oxygen-containing hydrocarbons are potential Martian rocket propellants that have the energy to launch a human MAV into orbit from the surface Mars into low Mars orbit. This observation opened the door to a proposed biotechnology-enabled ISRU for the production of 2,3-BDO and the presented technical analysis of this process. Using state-of-the-art technology, the proposed bio-ISRU would be possible; however, higher payload mass and energy requirements do not make it competitive with previously proposed chemically-enabled ISRU, such as the DRA 5.0 ISRU for a single mission. Yet, the excess evolution of O_2 as a byproduct, which could help sustain human colony formation, and the possibility of quickly adapting the bio-ISRU to produce other chemicals by simply swapping the engineered heterologous microbe make the bio-ISRU a very attractive proposition. Indeed, the presented technical analysis suggests that by doubling the cyanobacteria growth productivity, improving the enzyme digester efficiency by 33%, increasing the heterotroph yield of 2,3-BDO by 16%, and implementing lighter reactor components the bio-ISRU requires only a 24% higher payload than the chemically-enabled ISRU while still producing 30.6 tons of excess oxygen. Thus, working on the biological and materials challenges presented in this work, has the potential to enable future human presence beyond Earth to become a reality.

4.6 References

1. National Aeronautics and Space Administration Authorization Act of 2005. S.1281, 109th Congress (2005).

2. Drake, B.G., Human Exploration of Mars Design Reference Architecture 5.0. *National Aeronautics and Space Administration* NASA/SP-2009-566 (2009).

3. Verseux, C. et al. Sustainable life support on Mars – the potential roles of cyanobacteria. *Int. J. Astrobiol.* **15,** 65-92 (2015).

4. Menezes, A.A., Cumbers, J., Hogan, J.A. & Arkin, A.P. Towards synthetic biological approaches to resource utilization on space missions. *J. R. Soc. Interface* **12,** 20140715 (2015).

5. Rittmann, S., Seifert, A. & Herwig, C. Essential prerequisites for successful bioprocess development of biological CH4 production from CO2 and H2. *Crit. Rev. Biotechnol.* **35,** 141-151 (2015).

6. Starr, S.O. & Muscatello, A.C. Mars in situ resource utilization: a review. *Planet. Space Sci.* **182,** 104824 (2020).

7. Arias, D.M., García, J. & Uggetti, E. Production of polymers by cyanobacteria grown in wastewater: current status, challenges and future perspectives. *New Biotechnol.* **55,** 46-57 (2020).

8. Knoot, C.J., Ungerer, J., Wangikar, P.P. & Pakrashi, H.B. Cyanobacteria: promising biocatalysts for sustainable chemical production. *J. Biol. Chem.* **293,** 5044-5052 (2018).

9. Aikawa, S. et al. Direct conversion of Spirulina to ethanol without pretreatment or enzymatic hydrolysis processes. *Energy Environ. Sci.* **6,** 1844-1849 (2013).

10. Möllers, K.B., Cannella, D., Jørgensen, H. & Frigaard, N.-U. Cyanobacterial biomass as carbohydrate and nutrient feedstock for bioethanol production by yeast fermentation. *Biotechnol. Biofuels* **7,** 64 (2014).

11. Liu, F. et al. Bioconversion of distillers' grains hydrolysates to advanced biofuels by an Escherichia coli co-culture. **Microb. Cell Fact.** *16,* 192 (2017).

12. Hublitz, I., Henninger, D.L., Drake, B.G. & Eckart, P. Engineering concepts for inflatable Mars surface greenhouses. *Adv. Space Res.* **34,** 1546-1551 (2004).

13. Badri, H., Monsieurs, P., Conninx, I., Wattiez, R. & Leys, N. Molecular investigation of the radiation resistance of edible cyanobacterium Arthrospira sp. PCC 8005. *Microbiologyopen* **4,** 187-207 (2015).

14. Dartnell, L.R., Desorgher, L., Ward, J.M. & Coates, A.J. Modelling the surface and subsurface Martian radiation environment: implications for astrobiology. *Geophys. Res. Lett.* **34,** L02207 (2007).

15. Lauro, S.E. et al. Multiple subglacial water bodies below the south pole of Mars unveiled by new MARSIS data. *Nat. Astron.* **5,** 63-70 (2021).

16. Cockell, C.S. Trajectories of Martian Habitability. Astrobiol. 14, 182-203 (2014).

17. Jena, U., Vaidyanathan, N., Chinnasamy, S. & Das, K.C. Evaluation of microalgae cultivation using recovered aqueous co-product from thermochemical liquefaction of algal biomass. *Bioresour. Technol.* **102,** 3380-3387 (2011).

18. Xu, Q., Bai, F., Chen, N. & Bai, G. Utilization of acid hydrolysate of recovered bacterial cell as a novel organic nitrogen source for L-tryptophan fermentation. *Bioengineered* **10,** 23-32 (2019).

19. McKay, C.P. & Marinova, M.M. The physics, biology, and environmental ethics of making Mars habitable. *Astrobiol.* **1,** 89-109 (2001).

20. Vivek, N. et al. Recent advances in microbial biosynthesis of C3-C5 diols: genetics and process engineering approaches. *Bioresour. Technol.* **322,** 124527 (2021).

21. Wang, J. et al. Bacterial synthesis of C3-C5 diols via extending amino acid catabolism. *Proc. Natl. Acad. Sci. USA* **117,** 19159 (2020).

22. Niu, W., Kramer, L., Mueller, J., Liu, K. & Guo, J. Metabolic engineering of Escherichia coli for the de novo stereospecific biosynthesis of 1,2-propanediol through lactic acid. *Metab. Eng. Commun.* **8,** e00082 (2019).

23. Kataoka, N., Vangnai, A.S., Tajima, T., Nakashimada, Y. & Kato, J. Improvement of (R)-1,3-butanediol production by engineered Escherichia coli. *J. Biosci. Bioeng.* **115,** 475-480 (2013).

24. Sathesh-Prabu, C., Kim, D. & Lee, S.K. Metabolic engineering of Escherichia coli for 2,3-butanediol production from cellulosic biomass by using glucose-inducible gene expression system. *Bioresour. Technol.* **309,** 12361 (2020).

25. Heirendt, L. et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat. Protoc.* **14,** 639-702 (2019).

26. Monk, J.M. et al. iML1515, a knowledgebase that computes Escherichia coli traits. *Nat. Biotechnol.* **35,** 904-908 (2017).

27. Xu, Y. et al. Systematic metabolic engineering of Escherichia coli for high-yield production of fuel bio-chemical 2,3-butanediol. *Metab. Eng.* **23,** 22-33 (2014).

28. Hakizimana, O., Matabaro, E. & Lee, B.H. The current strategies and parameters for the enhanced microbial production of 2,3-butanediol. *Biotechnol. Rep.* **25,** e00397 (2020).

29. Singh, S.P. & Singh, P. Effect of CO2 concentration on algal growth: a review. *Renew. Sustain. Energy Rev.* **38,** 172-179 (2014).

30. Murukesan, G. et al. Pressurized Martian-like pure CO2 atmosphere supports strong growth of cyanobacteria, and causes significant changes in their metabolism. *Orig. Life Evol. Biosph.* **46,** 119-131 (2016).

31. Cogne, G., Lehmann, B., Dussap, C.G. & Gros, J.B. Uptake of macrominerals and trace elements by the cyanobacterium Spirulina platensis (Arthrospira platensis PCC 8005) under photoautotrophic conditions: culture medium optimization. *Biotechnol. Bioeng.* **81,** 588-593 (2003).

32. Clippinger, J.N. & Davis, R.E. Techno-economic analysis for the production of algal biomass via closed photobioreactors: future cost potential evaluated across a range of cultivation system designs. *National Renewable Energy Laboratory* NREL/TP-5100-72716 (2019).

33. Davis, R.E. et al. Process design and economics for the production of algal biomass: algal biomass production in open pond systems and processing through dewatering for downstream conversion. *National Renewable Energy Laboratory* NREL/TP-5100-64772 (2016).

34. Comer, B.M. et al. Prospects and challenges for solar fertilizers. *Joule* **3,** 1578-1605 (2019).

35. Cannon, K.M., Britt, D.T., Smith, T.M., Fritsche, R.F. & Batcheldor, D. Mars global simulant MGS-1: a Rocknest-based open standard for basaltic Martian regolith simulants. *Icarus* **317,** 470-478 (2019).

36. Olsson-Francis, K. & Cockell, C.S. Use of cyanobacteria for in-situ resource use in space applications. *Planet. Space Sci.* **58,** 1279-1285 (2010).

37. Tredici, M.R., Rodolfi, L., Biondi, N., Bassi, N. & Sampietro, G. Techno-economic analysis of microalgal biomass production in a 1-ha Green Wall Panel (GWP®) plant. *Algal Res.* **19,** 253-563 (2016).

38. Karemore, A., Yuan, Y., Porubsky, W. & Chance, R. Biomass and pigment production for Arthrospira platensis via semi-continuous cultivation in photobioreactors: temperature effects. *Biotechnol. Bioeng.* **117,** 3081-3093 (2020).

39. Schultze, L.K.P. et al. High light and carbon dioxide optimize surface productivity in a twin-layer biofilm photobioreactor. *Algal Res.* **8,** 37-44 (2015).

40. Wang, J. et al. Field study on attached cultivation of Arthrospira (Spirulina) with carbon dioxide as carbon source. *Bioresour. Technol.* **283,** 270-276 (2019).

41. Zhang, L. et al. Attached cultivation for improving the biomass productivity of Spirulina platensis. *Bioresour. Technol.* **181,** 136-142 (2015).

42. Ozkan, A., Kinney, K., Katz, L. & Berberoglu, H. Reduction of water and energy requirement of algae cultivation using an algae biofilm photobioreactor. *Bioresour. Technol.* **114,** 542-548 (2012).

43. Al-Hattab, M., Ghaly, A. & Hammouda, A. Microalgae harvesting methods for industrial production of biodiesel: critical review and comparative analysis. *J. Fundam. Renew. Energy Appl*. **5,** 1000154 (2015).

44. Rossi, N., Jaouen, P., Legentilhomme, P. & Petit, I. Harvesting of cyanobacterium Arthrospira platensis using organic filtration membranes. *Food Bioprod. Process.* **82,** 244-250 (2004).

45. Rossi, N., Petit, I., Jaouen, P., Legentilhomme, P. & Derouiniot, M. Harvesting of cyanobacterium Arthrospira platensis using inorganic filtration membranes. *Sep. Sci. Technol.* **40,** 3033-3050 (2005).

46. Rossignol, N., Vandanjon, L., Jaouen, P. & Quéméneur, F. Membrane technology for the continuous separation microalgae/culture medium: compared performances of cross-flow microfiltration and ultrafiltration. *Aquac. Eng.* **20,** 191-208 (1999).

47. Shao, P. & Kumar, A. Recovery of 2,3-butanediol from water by a solvent extraction and pervaporation separation scheme. *J. Membr. Sci.* **329,** 160-168 (2009).

48. Chaparro, G. & Mejía, A. Phasepy: a Python based framework for fluid phase equilibria and interfacial properties computation. *J. Comput. Chem.* **41,** 2504-2526 (2020).

49. Liu, S., Qureshi, N. & Hughes, S.R. Progress and perspectives on improving butanol tolerance. *World J. Microbiol. Biotechnol.* **33,** 51 (2017).

50. Yang, Z. et al. A review on reverse osmosis and nanofiltration membranes for water purification. *Polymers* **11,** 1252 (2019).

51. Drysdale, A., Nakamura, T., Yorio, N., Sager, J. & Wheeler, R. Use of sunlight for plant lighting in a bioregenerative life support system – equivalent system mass calculations. *Adv. Space Res.* **42,** 1929-1943 (2008).

52. Xu, R.-P., Li, Y.-Q. & Tang, J.-X. Recent advances in flexible organic light-emitting diodes. *J. Mater. Chem. C* **4,** 9116-9142 (2016).

53. Pulli, E., Rozzi, E. & Bella, F. Transparent photovoltaic technologies: current trends towards upscaling. *Energy Convers. Manage.* **219,** 112982 (2020).

54. Bailey-Serres, J., Parker, J.E., Ainsworth, E.A., Oldroyd, G.E.D. & Schroeder, J.I. Genetic strategies for improving crop yields. *Nature* **575,** 109-118 (2019).

55. Hu, G., Li, Y., Ye, C., Liu, L. & Chen, X. Engineering microorganisms for enhanced CO2 sequestration. *Trends Biotechnol.* **37,** 532-547 (2019).

56. Xu, Y. et al. Altered carbohydrate metabolism in glycogen synthase mutants of Synechococcus sp. strain PCC 7002: cell factories for soluble sugars. *Metab. Eng.* **16,** 56-67 (2013).

57. Niederholtmeyer, H., Wolfstädter, B.T., Savage, D.F., Silver, P.A. & Way, J.C. Engineering cyanobacteria to synthesize and export hydrophilic products. *Appl. Environ. Microbiol.* **76,** 3462 (2010).

58. Karatay, S.E. & Dönmez, G. Microbial oil production from thermophile cyanobacteria for biodiesel production. *Appl. Energy* **88,** 3632-3635 (2011).

59. Farrokh, P., Sheikhpour, M., Kasaeian, A., Asadi, H. & Bavandi, R. Cyanobacteria as an eco-friendly resource for biofuel production: a critical review. *Biotechnol. Prog.* **35,** e2835 (2019).

60. Tredici, M.R. et al. Energy balance of algal biomass production in a 1-ha "Green Wall Panel" plant: how to produce algal biomass in a closed reactor achieving a high net energy ratio. *Appl. Engergy* **154,** 1103-1111 (2015).

CHAPTER 5.

MEMBRANE AUGMENTED CELL-FREE SYSTEMS: A NEW FRONTIER IN BIOTECHNOLOGY

Reproduced with permission from:

Kruyer, N.S., Sugianto, W., Tickman, B.I., Alba Burbano, D., Noireaux, V., Carothers, J.M., and Peralta-Yahya, P. Membrane augmented cell-free systems: a new frontier in biotechnology. *ACS Synth. Biol*. In Press.

5.1 Abstract

Membrane proteins are present in a wide array of cellular processes from primary and secondary metabolite synthesis to electron transport and single carbon metabolism. A key barrier to applying membrane proteins industrially is their difficult functional production. Beyond expression, folding, and membrane insertion, membrane protein activity is influenced by the physicochemical properties of the associated membrane, making it difficult to achieve optimal membrane protein performance outside the endogenous host. In this review, we highlight recent work on production of membrane proteins in membrane augmented cell-free systems (CFSs) and applications thereof. CFSs lack membranes and can thus be augmented with user-specified, tunable, mimetic membranes to generate customized environments for production of functional membrane proteins of interest. Membrane augmented CFSs would enable the synthesis of more complex plant secondary metabolites, the growth and division of synthetic cells for drug

delivery and cell therapeutic applications, as well as enable green energy applications including CH_4 capture and artificial photosynthesis.

5.2 Introduction

The functional, heterologous expression of membrane proteins is one of the missing puzzle pieces in establishing industrially relevant biological processes ranging from the production of medicinal compounds to the capture of methane (CH_4) to the bioremediation of heavy metal pollutants (Figure 5.1a). Plant-based medicinal compounds are synthesized via multi-enzyme cascades composed of several transmembrane cytochromes P450 (CYPs) that decorate the compounds' scaffolds [1, 2]. Particulate methane monooxygenase (MMO) oxidizes CH_4 to methanol [3-5], which enters C1 assimilation pathways in natural and synthetic methanotrophs [3], potentially able to convert the ~650 million tons of CO_2 equivalents produced in the U.S. [6] into high-density fuels to power trucks and airplanes. Heavy metals, such as uranium from nuclear waste, can be bioremediated using MtrCAB, which facilitates the transfer of electrons from the organism to the heavy metal [7]. Towards therapeutic applications, the robust functional heterologous expression of surface receptors would support the development of drug delivery vehicles and cell therapies. For instance, G protein-coupled receptors (GPCRs) are the target of more than 30% of FDA approved drugs. Routine heterologous expression of GPCRs would facilitate the development of high-throughput screening platforms for the discovery of new drugs or the study of signaling cascades in the absence of endogenous GPCRs [8]. Access to a wider array of functional receptors would also expand cell therapies beyond detection of cell surface antigens on cancer cells to the detection of soluble, small molecule ligands around

the tumor to improve targeting and reduce on-target off-tumor toxicity, i.e. targeting a non-tumor tissue expressing the same antigen [9]. For example, by using GPCRs, which mediate most cellular responses to small molecules [10]. Finally, transmembrane proteins are pivotal in primary metabolism, determining the biosynthetic performance of the production host. For instance, a network of transmembrane proteins synthesizes the phospholipids needed to build the cell's membranes. The oxidative phosphorylation pathway used to produce ATP in plants, bacteria, and humans is also composed of transmembrane proteins.

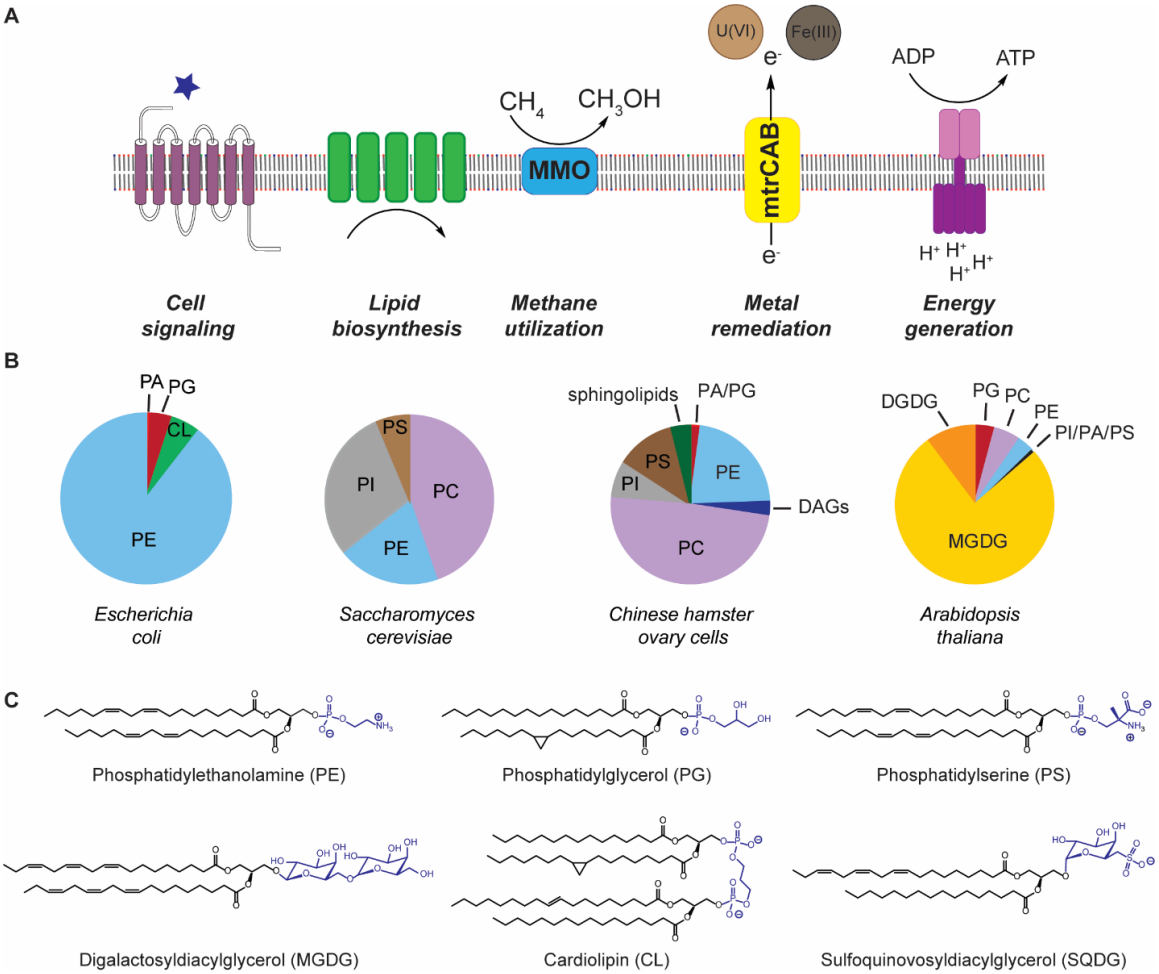

Figure 5.1. Biological roles of membrane proteins and cell membranes. (A) Cellular processes that occur at the cell membrane include cell signaling, lipid biosynthesis, methane utilization, metal remediation, and energy generation. (B) Membrane compositions in different organisms. PG: Phosphatidylglycerol, PA: Phosphatidic acid, PE: Phosphatidylethanolamine, CL: Cardolipin, PC: Phosphatidylcholine, PI: Phosphatidylinositol, PS: Phosphatidylserine, DAG: Diacylglycerol, MGDG: Monogalactosyldiacylglycerol, DGDG: Digalactosyldiacylglycerol. (C) Sample phospholipid structures. Phospholipid heads in blue.

Application of membrane proteins is hindered by their difficult production outside their endogenous host, with successful applications often requiring engineering of the transmembrane domain. For example, the microbial synthesis of plant natural products is limited to hosts amenable to plant transmembrane CYP production, such as *Saccharomyces cerevisiae*, in spite of other hosts, such as *Escherichia coli*, having achieved higher precursor yields [11]. The development of protein chimeras and truncation of plant CYPs to generate soluble variants has been successful [12], however, only one or two CYPs are usually engineered at a time, far from the five to ten CYPs required for many plant biosynthetic pathways.

Beyond transmembrane protein production, the physicochemical properties of the membrane, including composition, fluidity, curvature, and molecular crowding, influence the production and activity of membrane proteins [13, 14]. For example, expression of *Catharanthus roseus* geraniol 10-hydroxylase in *S. cerevisiae* has a 8.3-fold lower activity than the same protein synthesized in a plant membrane, likely due to decreased enzyme stability and suboptimal reductase pairing [15]. In addition, a molecular dynamics simulation of a human CYP, CYP3A4, showed that lipid composition and electrostatics impact membrane incorporation and membrane protein orientation [16]. Indeed, the membrane compositions of mammalian, microbial, and plant cells vary vastly from one another [17-20] (Figure 5.1b).

As non-living systems devoid of membranes, cell-free systems (CFSs) could be augmented with tailor-made membranes to fulfill specific membrane protein requirements and applications, thus functionally producing membrane-bound proteins that are challenging to synthesize using cells. Briefly, CFSs are composed of a cell lysate or

purified cell machinery (PURE) supplemented with the nucleotides, energy sources, amino acids, cofactors, and salts necessary for transcription and translation [21]. PURE systems are often preferred for more complex protein synthesis due to reduced background and optimized conditions. For example, bacteriorhodopsin, ATP synthase, and enzymes in the lipid biosynthesis pathway have all been synthesized in PURE systems [22]. On the other hand, preparation of PURE reaction mix is low throughput and expensive. Thus, for potential scale-up applications cell lysate-based systems are required. Among cell-lysate-based CFSs, *E. coli*-based CFSs are the most commonly used platform with applications to the production of therapeutics, genetic circuit engineering, and construction of synthetic cells, chemical biosynthesis, and protein production [23].

In membrane augmented CFSs, the user has complete control over composition, fluidity, and crowding of membranes in addition to curvature and vesicle size in the case of encapsulated CFSs. Lipids of different structure, length, saturation, and charge, a number of them commercially available, could be used to optimize the membrane composition (Figure 5.1c). Further, encapsulated CFSs could have different lipid compositions depending on the location of the bilayer leaflet. A membrane augmented CFS would open the doors to important bio-industrial applications, such as the development of hybrid chemical-biological processes, the use of organic solvents for in situ product extractions, and more efficient downstream processes for product separation. Developing genetic control systems to control both the lipid and protein composition of membrane augmented CFSs will be pivotal to achieve the high levels of membrane enzyme activity to enable these applications. For example in the bioremediation space, expression of

MtrCAB in a membrane augmented CFS would enable circumvention of cell toxicity issues to facilitate applications at high contaminant concentrations.

In this review, we highlight recent advances in membrane-based CFSs and their application in the heterologous production of membrane proteins. Further, we examine the potential for genetic control systems, such as those implemented with CRISPR-based transcriptional regulation, to improve the cell-free synthesis of membrane proteins for chemical production and for the study of protein-membrane interactions. Although we are not yet at the level of on-demand membrane augmented CFS generation, the potential advantages of such systems in terms of enabling new chemistry and improving chemical bioproduction processes make a new frontier in biotechnology.

5.3 Production of Membrane Proteins in CFSs

5.3.1 Membrane augmented CFSs

Membrane protein production in CFSs is often limited by self-aggregation [13], and addition of oil droplets to a PURE CFS has enabled the production of single-span transmembrane proteins, such as FasL and TRAIL used as anti-cancer therapeutics [24]. Oil droplets, however, are limited to the display of surface receptors, and cannot recapitulate the physicochemical properties of native lipids that support membrane protein activity [13]. Phospholipid-like additives, such as nanodiscs and liposomes, recapitulate membrane composition better (Figure 5.2) [25]. Nanodiscs are phospholipids stabilized by membrane scaffold proteins [26], which form a planar bilayer akin to a lipid raft. Both sides of a nanodisc are indistinguishable from one another and open to interact with the CFS environment [13]. Nanodiscs are commercially available and can be directly added to

CFSs. Nanodiscs, however, may not work to produce every transmembrane protein [13], and their composition cannot be easily changed [25]. The cost of Nanodiscs (~$1/μL of CFS reaction volume), currently limits this technology to mostly research applications. Liposomes are spherical phospholipid vesicles [13] that provide a tunable environment in terms of shape, size, and composition to mimic both prokaryotic (d= 0.1–5.0 μm) and eukaryotic (d= 10–100 μm) cells. The availability of a wide array of phospholipids allows the generation of liposomes with tunable composition and size. Given liposomes' large surface area, they are the preferred scaffold for co-expression of multiple membrane proteins. Liposomes can be added to a CFS reaction or the CFS can be encapsulated within the liposomes, albeit the yield of liposomes loaded with CFS is typically small (~100 liposomes per μL of reaction mix), limiting their application. Using commercial phospholipids, liposomes have a cost of <$0.01/μL of CFS reaction, enabling high-end biotechnology applications [27]. Finally, inverted vesicles formed during the CFS cell lysis procedure have been used to produce complex membrane proteins, including the oxidative phosphorylation pathway [28] and transmembrane oligosaccharyltransferases for protein glycosylation [29]. As used, inverted vesicles require production and insertion of membrane proteins at the cellular stage, which brings the usual challenges with heterologous production of membrane proteins. The cell used for protein production controls the membrane composition and vesicle size, making it difficult to explore the extent to which these variables have an effect on enzyme activity or control these variables for the desired application.

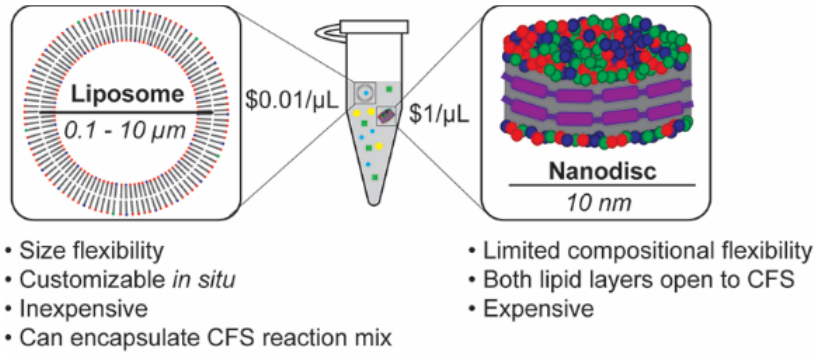

Figure 5.2. Membrane augmented cell-free systems. Mimetic membranes, in the form of liposomes and nanodiscs, are added directedly to CFSs to aid in membrane protein production.

In situ phospholipid synthesis by CFSs would reduce the cost of membrane augmented CFSs, opening the doors to bio-industrial applications. *In situ* phospholipid synthesis has been achieved by feeding glycerol-3-phosphate and acyl-CoA. Because the lipid biosynthesis enzymes are themselves membrane-bound proteins, the CFS must be inoculated with preformed liposomes [30]. To enable feeding of fatty acids as a substrate, FadD was expressed in the CFS leading to phospholipid production and observable vesicle growth [31]. Nevertheless, low fatty acid solubility and its detrimental impact on protein stability limited this work, making the case for using glucose as the starting material in the future. Of note, as the fatty acids are located on the outside of the liposome, the phospholipids are only incorporated to the outer leaflet of the bilayer, limiting the synthesis of asymmetric membranes and potentially disturbing the incorporation of transmembrane proteins.

5.3.2 Chemical composition of membrane augmented CFSs

Membrane proteins often rely on the phospholipids around them for activity. For example, the reaction rate, substrate affinity, and reductase coupling efficiency of the

human cardiomycete epoxygenase responsible for oxidation of fatty acids and xenobiotics (CYP2J2) was shown to be impacted by the concentrations of phosphatidylcholine (PC) and phosphatidylserine (PS) [32]. In another example, the bacterial structural protein MreB was polymerized when expressed in a liposome composed of PC and phosphatidylethanolamine-polyethylene glycol (PE-PEG) due to PEG's effect on membrane crowding, changing the shape of the liposome from spherical to rod-like [33]. Therefore, membrane protein activity can be altered not only by protein engineering, but also by engineering the membranes in which they are embedded.

Changing the chemical structure of the phospholipid heads or tails causes changes in the membrane's physical properties such as fluidity, thickness, and charge [34]. *In situ* changes in phospholipid composition have been achieved by using different acyltransferases to convert phosphatidic acid (PA) to either phosphatidylglycerol (PG) or PE [30]. Importantly, the PE:PG ratio was maintained to the *E. coli* membrane composition, 75% PE [35], both genetically, by placing PE and PG production under control of different promoters and dosing the respective polymerases, and enzymatically using PssA, which associates with PG rich membranes and catalyzes the synthesis of PE [36]. Finally, conversion of phosphatidylmonomethylethanolamine (PMME) into PC, the most prevalent phospholipid in eukaryotes, was achieved by expressing the methyltransferase Opi3 in nanodiscs [37].

5.3.3 Physical properties of membrane augmented CFSs

Liposome vesicle diameter, shape, curvature, and number of lamella can be optimized to create a native-like membrane environment. Optimizing the physical

properties of the mimetic membrane is key to protein folding, membrane incorporation, and activity. Using the multi-drug transporter EmrE as a model, it was shown that surface area-to-volume ratio in smaller vesicles improved membrane insertion and the ratio of incorporated ErmE to total synthesized ErmE was a function of liposome size, not of DNA concentration [38]. It remains to be seen if this correlation holds true for other membrane proteins, or in membranes composed of more than one type of lipid. Optimal membrane size may depend on other factors such as membrane fluidity and rigidity and may change depending on the transmembrane protein source organism. Another important membrane physical parameter is elasticity. Using diblock copolymers to increase the membrane elasticity, it was shown that decreasing the membrane area expansion modulus improved membrane folding of the mechanosensitive channel of large conductance (MscL) using the PURE system [39]. However, differing results for the model membrane protein channel rhodopsin (ChR2) indicated that this conclusion may not be generalizable to all membrane proteins and the number of transmembrane regions may influence the optimal membrane elasticity. Importantly, in addition to polymers, biosynthetic compounds, such as cholesterol, can tune the elasticity of membranes.

5.3.4 Membrane protein insertion into membrane augmented CFS

Chaperones facilitate the insertion of membrane proteins in cells; however, solely membrane composition can affect the efficiency and directionality of this process. In a PURE-based CFS lacking chaperones, the 12-transmembrane proteins LacY and XylE were expressed and spontaneously incorporated into the liposome membrane. The membrane protein incorporation was lipid dependent with 2.6- and 1.5-fold increases in

LacY and XylE incorporation, respectively, between the worst (100% PC) and best (100% PG) performing membrane compositions [40]. Importantly, lipid composition played a role in establishing the correct directionality for membrane insertion. While the 100% PG membrane had the highest protein incorporation, the lack of PE resulted in LacY being incorporated in an inverted membrane orientation. This result supported previous work that identified PE as a non-proteinaceous chaperone of LacY [41]. Non-spontaneous membrane protein insertion can be facilitated in a PURE system by the secYEG translocon, which successfully integrated YidC and LepB into an exogenous liposome [42]. Spontaneous co-translational integration of the multi-subunit secYEG [42] and ATP synthase [43] complexes into liposomes in CFSs suggests that many multi-subunit membrane proteins will be functional in CFSs without need for additional reaction components.

In encapsulated CFSs, orientation is critical for protein function, and membrane proteins often have to asymmetrically localize in the bilayer leaflet. Using SNAP-tag modified fluorescent proteins and liposomes composed of benzylguanine-modified phospholipids, membrane asymmetry was achieved by encapsulating a CFS expressing mCherry-SNAP within the modified liposome, and suspending the encapsulated CFS in a second CFS expressing GFP-SNAP. Fluorescence microscopy confirmed mCherry was localized to the inner membrane of the liposome while GFP localized to the outer membrane [44]. Controlling protein localization is key in applications that benefit from separating intermediates in metabolic pathways. For example, the production of CMP-N-acetylneuraminic acid was improved 2-fold by encapsulating the first pathway enzyme, N-acyl-D-glucosamine-2-epimerase, within a polymersome and attaching the rest of the

pathway on the outside of the polymersome, thus reducing inhibition of N-acyl-D-glucosamine-2-epimerase by a late pathway intermediate [45].

5.4 Applications of Membrane Augmented CFSs

Expanding the use of CFSs to include membrane proteins offers a wide range of applications from single enzyme assays to use of the membrane proteins as part of longer enzyme pathways. Table 5.1 offers an overview of current literature in which membrane proteins are applied in a CFS using nanodiscs, liposomes, or inverted vesicles, highlighting both in vitro synthesized and reconstituted proteins from multiple protein classes.

Table 5.1. Applications of membrane proteins in cell-free systems

Protein Type Key
Structural/Cell Maintenance
Transporter/Porin
Energy Production
Receptor
Secondary Metabolism

Protein	Description	Application	Membrane type	Membrane composition	Protein Synthesis	CFS Method	Citation
MreB	Cell shape-determining protein	Synthetic cells	Liposome	Egg PC PE-PEG	Cell-Free	Cell lysate	[33]
				E. coli total lipid extract	Purified enzymes	N/A	[77]
HyaA	Hydrogensase-1 small chain	Protein characterization	Liposome	DOPC DGS-NTA DMPE-RhoB	Cell-Free	PURE	[78]
OmpA	Outer membrane protein A						
YfbF	Glycosyl transferase	Protocol development	Liposome	DOPC	Cell-Free	PURE	[79]
CyoE	Protoheme IX farnesyltransferase						
FtsZ	Cell division protein	Synthetic cells	Liposome	POPC POPG DOPE-RhoB	Cell-Free	PURE	[80]
FtsA							
ZipA							
GPAT	Lipid biosynthesis	Synthetic cells	Liposome	DOPC DOPE DOPG Cardiolipin	Cell-Free	PURE	[30, 36]
LPAAT							
CdsA							
PgsA							

Table 5.1 (continued)

PgpABC PssA Psd	Lipid biosynthesis	Synthetic cells	Liposome	DOPC DOPE DOPG Cardiolipin	Purified enzymes	N/A	[31]
CYP2J2	Cytochrome P450	Protein characterization	Nanodisc	POPC POPS Cholesterol	Purified enzymes	N/A	[32]
CYP5A1	Cytochrome P450	Protein characterization	Nanodisc	POPC POPS POPE	Purified enzymes	N/A	[81]
CYP2B4	Cytochrome P450	Protein characterization	Nanodisc	DMPC POPC POPS	Purified enzymes	N/A	[82]
Opi3	Methyltransferase	Protein characterization	Nanodisc	DMPG DOPG DOPMME	Cell-Free	Cell lysate	[37]
OST	Oligosaccharyltransferases	Protein characterization	Nanodisc	POPC	Cell-Free	Cell lysate	[46]
		Protocol development	Inverted vesicle	N/A	Cell-Free	Cell lysate	[29]
SecYEG YidC LepB	Translocon Insertase Signal peptidase	Protocol development	Liposome	Soybean lipid extract	Cell-Free	PURE	[42]
MscL	Large-conductance mechanosensitive channel	Synthetic cells	Liposome	Egg PC	Cell-Free	Cell lysate	[83]
		Protein characterization	Liposome	DOPC + PEG diblock copolymer or detergent	Cell-Free	PURE	[39]

Table 5.1 (continued)

ErmE	Multidrug transporter	Protein characterization	Liposome	POPC	Cell-Free	PURE	[38]
		Protocol development	Liposome	POPC	Cell-Free	PURE	[84]
LacY	Sugar transporter	Protein characterization	Liposome	DMPC DOPC DOPE DOPG	Cell-Free	PURE	[40]
XylE							
α-hemolysin	Pore forming protein	Protein characterization	Liposome	POPC Cholesterol	Cell-Free	PURE	[85]
		Synthetic cells	Liposome	POPC Cholesterol	Cell-Free	Cell lysate	[60]
PFO	Pore protein perfringolysin O	Synthetic cells	Liposome	POPC Cholesterol	Cell-Free	Cell lysate	[59]
ATP synthase	ATP synthesizing protein complex	Synthetic cells	Liposome	DOPC DOPE DOPG	Cell-Free	Cell lysate	[43]
				POPC	Purified enzymes	N/A	[68]
					Cell-Free	PURE	
				POPC POPE POPG Cholesterol	Purified enzymes	N/A	[67]
		Protocol development	Inverted vesicle	N/A	*In vivo*[1]	N/A	[28]

Table 5.1 (continued)

Bacteriorhodopsin	Photoconverter	Synthetic cells	Liposome	POPC	Purified enzymes	N/A	[68]
					Cell-Free	PURE	
				POPC POPE POPG Cholesterol	Purified enzymes	N/A	[67]
Photosystem II	Photoconverter	Synthetic cells	Liposome	POPC POPE POPG Cholesterol	Purified enzymes	N/A	[67]
PsbS	Photosystem II subunit S	Synthetic cells	Liposome	Asolectin	Cell-Free	Cell lysate[3]	[86]
MtrCAB	Electron transfer pathway	Protein characterization	Liposome	PC	Purified enzymes	N/A	[87]
CX3CR1	Chemokine G-protein coupled receptors	Protein characterization	Liposome	POPC PE-PEG	Cell Free	PURE	[47]
CCR5			Nanodisc	POPC POPS Cholesterol			
CYP725A	Cytochrome P450	Protein characterization	Nanodisc	POPC	Purified enzymes	N/A	[88]

DGS-NTA = Dioleoylglycero-[(N-(5-amino-1-carboxypentyl) iminodiacetic acid) succinyl] (nickel salt);
DMPC = Dimyristoylphosphatidylcholine; DMPE-RhoB = Dimyristoylphosphatidylethanolamine-rhodamine B;
DMPG = Dimyristoylphosphatidylglycerol; DOPC = Dioleoylphosphatidylcholine; DOPE = Dioleoylphosphatidylethanolamine;
DOPE-RhoB = Dioleoylphosphatidylethanolamine-rhodamine B; DOPG = Dioleoylphosphatidylglycerol;
DOPMME = Dioleoylphosphatidylmonomethylethanolamine; PC = Phosphatidylcholine; PE-PEG = Phosphatidylethanolamine-polyethylene glycol; POPC = Palmitoyl-oleoyl-phosphatidylcholine; POPE = Palmitoyl-oleoyl-phosphatidylethanolamine;

Table 5.1 (continued)

POPG = Palmitoyl-oleoyl-phosphatidylglycerol; POPS = Palmitoyl-oleoyl-phosphatidylserine
In the protein synthesis column: 'Cell-Free' indicates that the enzyme was synthesize in the cell-free system. 'Purified enzymes' indicates that the protein was synthesized *in vivo*, purified, and reconstituted in CFS
[1] ATP synthase produced *in vivo* and recovered in CFS after cell lysis
[2] Cell lysate reaction mix contains added, purified T7 polymerase

5.4.1 Enzyme Assays

Transmembrane protein production in CFSs enables their study in the absence of endogenous metabolic pathways or metabolites that may confound the results. For example, oligosaccharyltransferase homologs have been synthesized in cell lysate based CFSs using nanodiscs and used to rapidly identify acceptor proteins [46], bypassing competition from an endogenous glycosylation system. Furthermore, the chemokine GPCRs CX3CR1 and CCR5 have been produced in CFSs using nanodiscs in a PURE system for structural (electron microscopy) and functional (surface plasmon resonance) studies [47]. As membrane augmented CFSs become more widely available, we expect other membrane protein classes to make use of this technology.

5.4.2 Energy Production

CFSs hold incredible promise for the production of fuels and chemicals due to their high productivity when compared to microbes, up to 815 mg/L/h in the case of mevalonate [48]. Cell-free production of butanol and hydrogen could be improved by extending the respective pathways using particulate MMO for the assimilation of CH_4 to enable use of a C1 feedstock [49]. Such a system would also take advantage of the improved separation and reduced effects of toxicity provided by using CFSs. It is worth noting that CFS production platforms are currently not cost competitive with microbial ones for biofuel production and have not been scaled to the volume or run as long as microbial cell factories [50]. The generation of electricity in microbial fuel cells would also benefit from improvements in membrane protein expression in CFSs by helping to overcome barriers in feedstock toxicity and limits on electric current production due to organism viability [51].

To this end, enzymatic fuel cells using purified enzymes can produce electricity from hydrogen, methanol, formate, pyruvate, and sugars, though without the need for membrane proteins [52]. However, these systems often have a lifetime of hours to days. Incorporation of these enzymes into a CFS would allow for enzyme and cofactor regeneration, extending the lifetime of the fuel cell. Furthermore, use of membrane augmented CFS could further improve enzyme stability, and expand the range of usable fuels to include CH_4.

5.4.3 Chemical Biosynthesis

Despite the now large array of chemicals synthesized using CFSs [50], none of them relies on a transmembrane protein for biosynthesis (Figure 5.3). Focusing on plant natural products, CFSs have been used to produce monoterpenes [53] and cannabinoids [54], but improved expression of membrane proteins would increase the structural diversity of reachable products. Specifically, expression of transmembrane plant CYPs, such as those found in monoterpene indole alkaloid (MIA) and benzylisoquinoline alkaloid (BIA) pathways [1], would allow CFS production of the anticancer drug Taxol, which requires 19 enzymatic steps after geranylgeranyl diphosphate with eight of those steps catalyzed by transmembrane CYPs [55]. Furthermore, synthesis of the MIA intermediate strictosidine requires four CYPs [56] while the synthesis of the BIA intermediate noscapine requires five [57]. Today, *S. cerevisiae* is used to produce plant CYPs, however, a long doubling time (3 hours) limits the rapid prototyping of CYPs and reductase partners. Highlighting the challenges associated with natural product production in *S. cerevisiae*, recent production of the tropane alkaloid scopolamine required N-terminal engineering of the membrane-bound littorine synthase from *Attropa belladonna* for functional expression in

S. cerevisiae [58]. Finally, the membrane compositions of yeast and plants are very different (Figure 5.1b), likely impacting transmembrane protein functionality. Membrane augmented CFSs offer a potential upgrade on *S. cerevisiae* in terms of productivity, reduced effects of product toxicity, and flexibility in membrane composition.

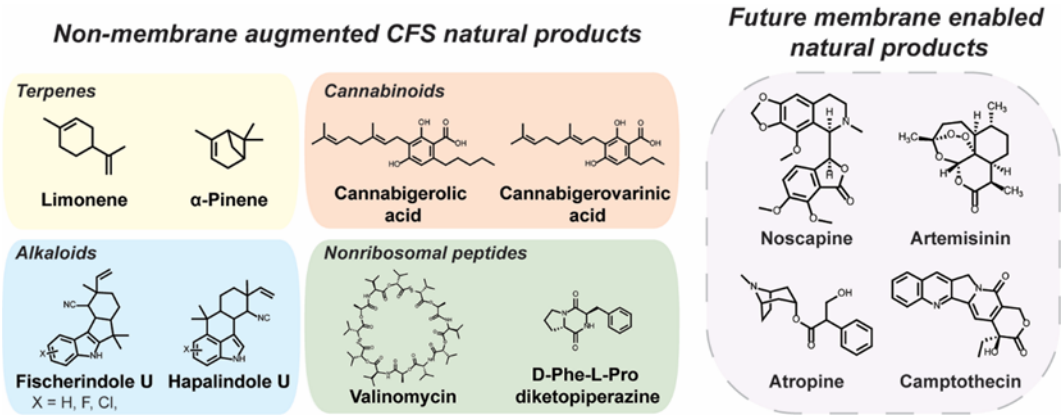

Figure 5.3. Production of natural products using cell-free systems. Compounds produced to date do not require membrane proteins or membrane augmented CFSs. In the dashed box, natural products that depend on the production of transmembrane proteins that could be synthesized by using a membrane augmented CFS.

5.4.4 Synthetic cells

Synthetic cells, particles that mimic biological cells, but have different characteristics, functions, or parts, can be generated by encapsulating CFSs in liposomes. Membrane proteins play pivotal roles in the development of synthetic cells: from growth and division via phospholipid biosynthesis and Z-ring proteins, to cell-to-cell communication mediated by cell surface receptors for the development of cell therapies, to the generation of NADH and ATP to extend their chemical bioproduction time.

In the context of cell-to-cell communication, a synthetic cell constitutively expressing a brain-derived neurotrophic factor was activated by addition of homoserine lactone, leading to production of the pore forming perfringolysin O, secretion of the neurotrophic factor, and differentiation in co-cultured stem cells [59]. Similarly, the pore forming protein alpha-hemolysin has been leveraged to control uptake or secretion of doxycycline and isopropyl-β-D-thiogalactoside (IPTG), which activated downstream luciferase expression as a proof of concept reporter gene [60].

In the context of chemical bioproduction, synthetic cells act as miniature bioreactors that, unlike cells, do not use the carbon for cell growth or maintenance, and essentially route all carbon for chemical production (Figure 5.4a). Thus, bioreactor synthetic cells act as immobilized enzyme catalysts, potentially enjoying easy re-use and separation, while allowing the enzyme inside the reactor to work within the simulated synthetic cell environment with limited loss of activity. Liposome volume and stability is particularly important for bioreactor synthetic cells. Decreased synthetic cell volume results in higher local substrate concentration, speeding up the reaction of low substrate affinity enzymes [61]. Nevertheless, high local product concentrations may limit yields through increased enzyme product inhibition. The stability of the liposome will determine the lifetime of the reactions that take place inside it, making it crucial to understand how each of these physical properties affects liposome stability, which is known to be affected by liposome structure [62] and synthesis method [63].

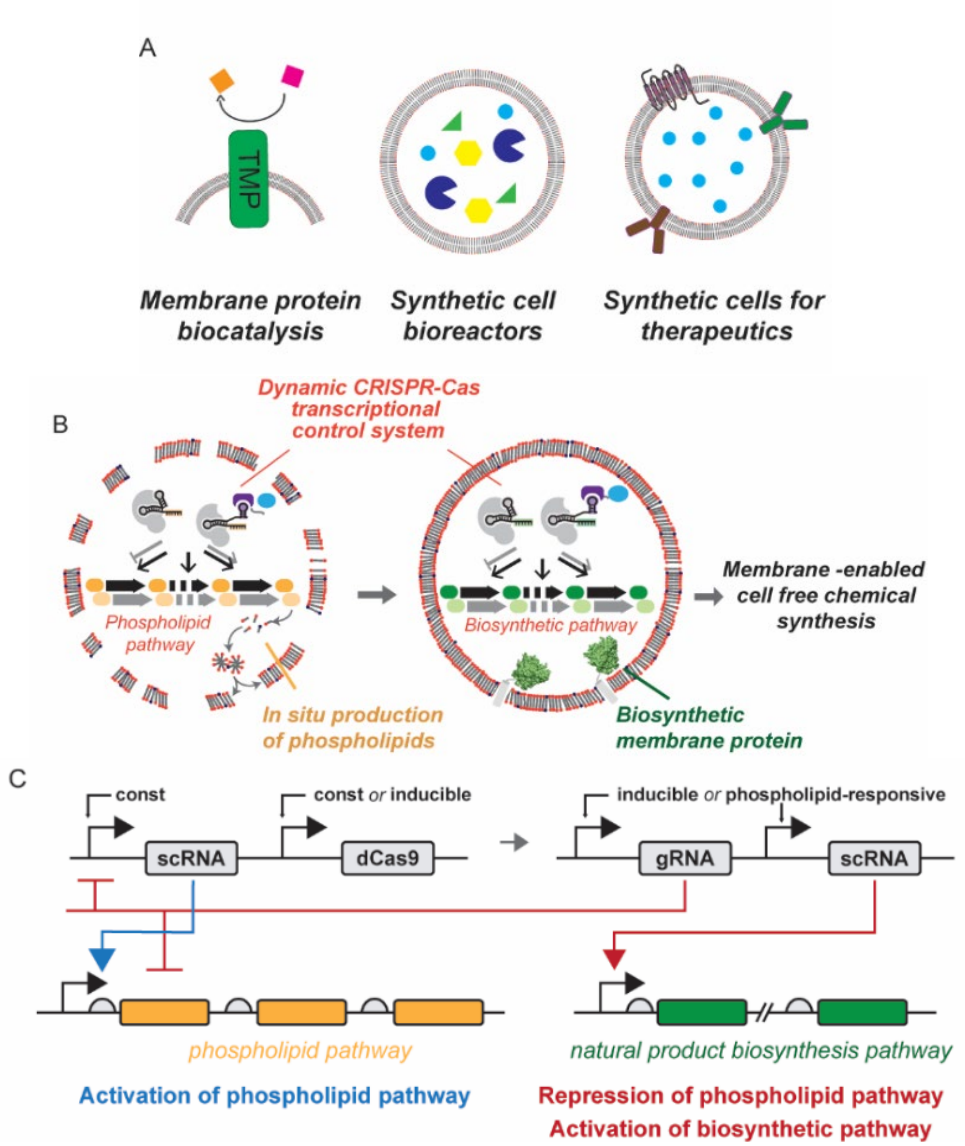

Figure 5.4. Applications of membrane augmented cell-free systems. (A) Potential bio-industrial applications of membrane proteins in membrane augmented cell-free sysytems. (B) Dynamic CRISPR-Cas control system can be implemented to stagger phospholipid and natural product biosynthesis, allowing for in situ liposome self-assembly prior to membrane protein production and chemical synthesis. (C) Sample two-stage control strategy for staggered phospholipid and natural product biosynthesis. First, constitutive (const) or induced expression of a modified guide RNA that enables the recruitment of a transcriptional activator (scaffold RNA, or scRNA) activates expression of the phospholipid pathway. Second, a user-supplied inducer or sufficient phospholipid concentration triggers expression of the natural product biosynthesis pathway (through targeted scRNA) as well as repression of the phospholipid pathway using a gRNA coupled with catalytically inactive Cas9 (dCas9). Suppression of lipid biosynthesis during stage two helps conserve precious CFS resources.

Membrane proteins often require prosthetic groups for activity. For instance, CYPs need heme, and membrane-bound glucose dehydrogenases used for NAD(P)H regeneration use pyrroloquinoline quinone (PQQ). Prosthetic groups need to be incorporated into proteins as they fold and are inserted in the membrane to achieve proper enzyme function. Often, CFSs lack the biosynthetic pathways to synthesize prosthetic groups and these compounds need to be exogenously added to the reaction. For scale-up applications, however, enriched CFSs from organisms where the prosthetic group biosynthetic pathways are expressed and/or upregulated will be needed for the functional expression of membrane proteins. Some inroads towards this goal have been made. For example, PQQ was synthesized by *Gluconobacter oxydans*-based CFS carrying the machinery to convert fed pqqA precursor to PQQ [64]. In another example, heme was biosynthesized in *E. coli* CFS via addition of purified 5-aminolevulinic acid synthase, and heme was successfully incorporated into P450 BM3 [65]. Of note, prosthetic group biosynthesis should be carefully regulated to avoid CFS poisoning or unnecessarily diverting carbon flux from the desired product. Additionally, chaperones could be introduced to help in the incorporation of prosthetic groups, such ferrochelatases for heme loading [66].

The final challenge in bioreactor synthetic cells is the need for reducing power (NADH) and energy (ATP) regeneration to drive reactions for extended periods of time to reduce process cost. Co-factor regeneration in CFSs has been achieved using glyceraldehyde-3-phosphate dehydrogenase and applied to monoterpene production [53]. Production of ATP in CFSs was achieved early on using oxidative phosphorylation [28]. More recently, efforts have moved to produce ATP from light using purified ATP synthase,

photosystem II, and proteorhodopsin reconstituted in liposomes [67], or using bacteriorhodopsin to generate the proton gradient necessary to produce ATP upon light induction [68]. Importantly, all ATP production mechanisms require membrane proteins to generate an electron gradient, which requires proteins to have the correct membrane orientation. This can be partially controlled by limiting spontaneous membrane integration through modulating cholesterol and diacylglycerol concentration [69]. However, the effect of these compounds is phospholipid dependent and requires continued study on more complex lipid mixtures [70].

5.5 Control Strategies for the Generation of Membrane Augmented CFSs

5.5.1 Genetic control of membrane properties

CFSs have a limited amount of resources for the formation of enzymes and metabolites as well as transcription and translation machinery. For synthetic cells to become industrially relevant, the cost effective synthesis of both phospholipids for membrane formation and membrane-bound actuating biomolecule(s) is needed. The actuating biomolecule(s) can be single enzymes for biocatalysis applications, multi-enzyme pathways for chemical bioproduction, or receptors for synthetic therapeutic applications. To enable these applications, a control system that can dynamically program gene expression of multiple units is needed. The control system should 1) have low overhead resource consumption, 2) be capable of turning genes both on and off to generate sequential phases of gene expression programs, 3) be tunable to precisely regulate gene expression dynamics, and 4) be scalable to allow for construction of increasingly complex gene regulatory systems [71, 72]. Such a control system would enable synthesis of

phospholipid-producing enzymes early in a reaction to form membranes, followed by a shift in production to membrane proteins to produce the chemicals, both with programmable stoichiometry and timing (Figure 5.4b,c). Control systems in liposomes will need to go beyond controlling the enzyme ratios [36], and dynamically control on/off gene expression. This capability would allow for controlled changes in membrane properties through the course of the reaction without outside intervention.

The control system should not only regulate lipid synthesis, but also balance the expression of the membrane bound actuating biomolecule to prevent aggregation while maximizing efficiency and rate of membrane protein insertion. Most of the current understanding on membrane protein insertion relies on expression rates to prevent saturation of membrane insertion machinery [14]. In the case of liposomes, hydrophobic interactions among membrane proteins lead to self-aggregation and interactions between membrane proteins and the lipid membrane play an outsized role in protein insertion, posing additional challenges to the control system. Implementation of a gene control system capable of delivering distinct gene expression profiles would develop a better understanding of how liposomes change over time, how liposomes adsorb proteins from the CFS, and how the density of proteins already in a liposome affects how much membrane protein uptake occurs.

5.5.2 CRISPRa/i – a control system for the expression of membrane proteins in membrane augmented CFS

While elementary gene regulation has been implemented in CFSs [73], efforts toward developing multi-gene control systems to provide precise regulation over gene

expression, membrane protein insertion, and function remain at an early stage. The CRISPR-Cas system provides a powerful suite of tools for multi-gene transcriptional control [71]. Briefly, catalytically inactive Cas9 protein can be directed to specific DNA sequences by guide RNAs that recognize target sequences based on predictable Watson-Crick base pairing to activate (CRISPRa) or repress (CRISPRi) gene expression. Although the rules governing CRISPRa from bacterial promoters are complex [74], a growing set of validated CRISPRa components enables the rapid construction of increasingly complex gene regulatory systems [71]. Multi-gene CRISPR circuits have been engineered through the regulated expression of up to seven distinct sgRNAs in the same system [75], and CRISPRi has been shown to operate efficiently in CFSs [76]. Thus, by combining new capabilities for CRISPRa with existing tools for CRISPRi in CFSs, it should be possible to engineer multi-gene programs for membrane augmented CFSs operating through the regulated expression of guide RNAs. Multi-guide RNA CRISPRa/i circuits could then be used to program distinct gene expression modes to enhance functional membrane protein expression. Consequently, CRISPRa/i circuits could provide an efficient mechanism for implementing dynamic multi-gene control, while preserving valuable CFS metabolites and cofactors.

A toolbox of pulse-generating CRISPRa/i circuits could be used to both investigate and optimize how membrane protein expression dynamics impact membrane insertion and function. Here, the network topology would specify the timing of gene expression pulses or regulatory functions. Further, incorporation of input-responsive pulses of gene expression or regulation into CRISPRa/i networks would extend tunable control to semi-continuous reactions and at any specific time within a membrane CFS reaction. Ultimately,

it may be possible to engineer CRISPRa/i programs as process controls for membrane augmented CFS bio-industrial applications that regulate protein expression, minimize the waste of valuable precursors and energy molecules and prevent the accumulation of destabilizing intermediates [72].

5.6 Future Directions

Membrane proteins play a pivotal role in bioenergy, biomedical, and bio-industrial applications, and our ability to harness their potential hinges on their functional production outside their endogenous host. Although to some extent heterologous membrane proteins can be engineered for optimal functional heterologous production, an alternate and now more and more feasible strategy is to engineer CFSs with tailor-made augmented membranes to ensure optimal transmembrane protein activity. Although we are far from on-demand membrane augmented CFS generation, the potential advantages of such systems in terms of enabling new chemistry and improved chemical bioproduction processes make this a worthwhile endeavor. The realization that we have not only protein engineering, but also membrane engineering in our toolkit when tackling transmembrane protein challenges should help us accelerate some of these applications.

The biggest step forward in this field will be moving from using membrane augmented CFSs for protein characterization and analysis to larger scale application for chemical biosynthesis, bioremediation or synthetic cells. This will require effective scale up of membrane protein synthesis and controlled expression of phospholipid biosynthesis enzymes and biosynthetic pathway enzymes. Due to cost and tunability, liposomes appear to be the better option for scaled up membrane protein synthesis in CFSs. This said, work

towards spontaneous assembly of liposomes in situ and without need for organic solvents would help lower process cost for liposome production. Furthermore, study and improvement on membrane protein-liposome stability will be necessary for widespread application. Metrics such as total turnover number and half-life may be useful to help quantify scalability of these systems.

Successful implementation of membrane protein synthesis in CFSs will open the door for enzymatic production of toxic chemicals in non-living systems as well as non-living biosensors and bioremediation tools that can be applied environmentally without risk of biological contamination, or loss of function due to environmental toxicity. Finally, more sophisticated synthetic cells aided by membrane proteins hold great promise in therapeutic applications for targeted drug therapies as well as communicating with cellular environments to make expression decisions based on external stimuli. If aided by CRISPRa/i, these decisions and logic gates can be made significantly more complex to respond to combinations of signals and give the synthetic cell temporal, on/off control over gene expression.

5.7 References

1. Cravens, A., Payne, J., and Smolke, C.D. (2019) Synthetic biology strategies for microbial biosynthesis of plant natural products. *Nat. Commun. 10,* 2142.

2. Ehrenworth, A.M., and Peralta-Yahya, P. (2017) Accelerating the semisynthesis of alkaloid-based drugs through metabolic engineeirng. *Nat. Chem. Biol. 13,* 249-258.

3. Clomburg, J.M., Crumbley, A.M., and Gonzalez, R. (2017) Industrial biomanufacturing: the future of chemical production. *Science 355,* aag0804.

4. Ross, M.O., MacMillan, F., Wang, J., Nisthal, A., Lawton, T.J., Olafson, B.D., Mayo, S.L., Rosenzweig, A.C., and Hoffman, B.M. (2019) Particulate methane monooxygenase contains only mononuclear copper centers. *Science 364,* 566-570.

5. Fei, Q., Guarnieri, M.T., Tao, L., Laurens, L.M.L., Dowe, N., and Pienkos, P.T. (2014) Bioconversion of natural gas to liquid fuel: opportunities and challenges. *Biotechnol. Adv. 32,* 596-614.

6. Desai, M., and Camobreco, V. (2020) Inventory of U.S. Greenhouse Gas Emissions and Sinks. *U.S. Environmental Protection Agency*, EPA 430-R-20-002.

7. Jiang, S., and Hur, H.-G. (2013) Effects of the anaerobic respiration of *Shewanella oneidensis* MR-1 on the stability of extracellular U(VI) nanofibers. *Microbes Environ. 28,* 312-315.

8. Yasi, E.A., Kruyer, N.S., and Peralta-Yahya, P. (2020) Advances in G protein-coupled receptor high-throughput screening. *Curr. Opin. Biotechnol. 64*, 210-217.

9. Hong, M.H., Clubb, J.D., and Chen, Y.Y. (2020) Engineering CAR-T cells for next-generation cancer therapy. *Cancer Cell 38,* 473-488.

10. Kipniss, N.H., Dingal, P.C.D.P., Abbott, T.R., Gao, Y., Wang, H., Dominguez, A.A., Labanieh, L., and Qi, L.S. (2017) *Nat. Commun. 8,* 2212.

11. Yang, D., Park, S.Y., Park, Y.S., Eun, H., and Lee, S.Y. (2020) Metabolic engineering of *Escherichia coli* for natural product biosynthesis. *Trends Biotechnol. 38,* 745-765.

12. Liu, X., Zhu, X., Wang, H., Liu, T., Cheng, J., and Jiang, H. (2020) Discovery and modification of cytochrome P450 for plant natural products biosynthesis. *Synth. Syst. Biotechnol. 5,* 187-199.

13. Sachse, R., Dondapati, S.K., Fenz, S.F., Schmidt, T., and Kubick, S. (2014) Membrane protein synthesis in cell-free systems from bio-mimetic systems to bio-membranes. *FEBS Lett. 588,* 2774-2781.

14. Guigas, G., and Weiss, M. (2016) Effects of protein crowding on membrane systems. *Biochim. Biophys. Acta. 1858,* 2441-2450.

15. Collu, G., Unver, N., Peltenburg-Looman, A.M.G., van der Heijden, R., Verpoorte, R., and Memelink, J. (2001) Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis. *FEBS Lett. 508, 2*15-220.

16. Navrátilová, V., Paloncýová, M., Berka, K., and Otyepka, M. (2016) Effect of lipid charge on membrane immersion of cytochrome P450 3A4. *J. Phys. Chem. B. 120,* 11205-11213.

17. Furse, S., Wienk, H., Boelens, R., de Kroon, A.I.P.M., and Killian, J.A. (2015) *E. coli* MG1655 modulates its phospholipid composition through the cell cycle. *FEBS Lett. 589*, 2726-2730.

18. Lindberg, L., Santos, A.X., Riezman, H., Olsson, L., and Bettiga, M. (2013) Lipidomic profiling of *Saccharomyces cerevisiae* and *Zygosaccharomyces bailii* reveals critical changes in lipid composition in response to acetic acid stress. *PLoS One 8*, e73936.

19. Symons, J.L., Cho, K.-J., Chang, J.T., Du, G., Waxham, M.N., Hancock, J.F., Levental, I., and Levental, K.R. (2021) Lipidomic atlas of mammalian cell membranes reveals hierarchical variation induced by culture conditions, subcellular membranes, and cell lineages. *Soft Matter 17*, 288-297.

20. Jia, Y., Tao, F., and Li, W. (2013) Lipid profiling demonstrates that suppressing *Arabidopsis* phospholipase Dδ retards ABA-promoted leaf senescence by attenuating lipid degradation. *PLoS One 8*, e65687.

21. Sun, Z.Z., Hayes, C.A., Shin, J., Caschera, F., Murray, R.M., and Noireaux, V. (2013) Protocols for implementing an *Escherichia coli* based TX-TL cell-free expression system for synthetic biology. *J. Vis. Exp. 16,* e50762.

22. Kuruma, Y., and Ueda, T. (2015) The PURE system for the cell-free synthesis of membrane proteins. *Nat. Protoc. 10,* 1328-1344.

23. Silverman, A.D., Karim, A.S., and Jewett, M.C. (2020) Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet. 21,* 151-170.

24. Yunker, P.J., Asahara, H., Hung, K.-C., Landry, C., Arriaga, L.R., Akartuna, I., Heyman, J., Chong, S., and Weitz, D.A. (2016) One-pot system for synthesis, assembly, and display of functional single-span membrane proteins on oil-water interfaces. *Proc. Natl. Acad. Sci. USA 113,* 608-613.

25. Henrich, E., Dötsch, V., and Bernhard, F. (2015) Screening for lipid requirements of membrane proteins by combining cell-free expression with nanodiscs. *Methods Enzymol. 556,* 351-369.

26. Bayburt, T.H., Grinkova, Y.V., and Sligar, S.G. (2002) Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins. *Nano. Letters 2,* 853-856.

27. Garamella, J., Garenne, D., and Noireaux, V. (2019) TXTL-based approach to synthetic cells. *Methods Enzymol. 617,* 217-239.

28. Jewett, M.C., Calhoun, K.A., Voloshin, A., Wuu, J.J., and Swartz, J.R. (2008) An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol. Syst. Biol. 4,* 220.

29. Jaroentomeechai, T., Stark, J.C., Natarajan, A., Glasscock, C.J., Yates, L.E., Hsu, K.J., Mrksich, M., Jewett, M.C., and DeLisa, M.P. (2018) Single-pot glycoprotein biosynthesis using a cell-free transcription-translation system enriched with glycosylation machinery. *Nat. Commun. 9,* 2686.

30. Scott, A., Noga, M.J., de Graaf, P., Westerlaken, I., Yildirim, E., and Danelon, C. (2016) Cell-free phospholipid biosynthesis by gene-encoded enzymes reconstituted in liposomes. *PLoS One 11,* e0163058.

31. Exterkate, M., Caforio, A., Stuart, M.C.A., and Driessen, A.J.M. (2018) Growing membranes *in vitro* by continuous phospholipid biosynthesis from free fatty acids. *ACS Synth. Biol. 7,* 153-165.

32. Huff, H.C., Maroutsos, D., and Das, A. (2019) Lipid composition and macromolecular crowding effects on CYP2J2-mediated drug metabolism in nanodiscs. *Protein Sci. 28,* 928-940.

33. Garenne, D., Libchaber, A., and Noireaux, V. (2020) Membrane molecular crowding enhances MreB polymerization to shape synthetic cells from spheres to rods. *Proc. Natl. Acad. Sci. USA 117,* 1902-1909.

34. Renne, M.F., and de Kroon, A.I.P.M. (2018) The role of phospholipid molecular species in determining the physical properties of yeast membranes. *FEBS Lett. 592,* 1330-1345.

35. Sohlenkamp, C., and Geiger, O. (2016) Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiol. Rev. 40,* 133-159.

36. Blanken, D., Foschepoth, D., Serrão, A.C., and Danelon, C. (2020) Genetically controlled membrane synthesis in liposomes. *Nat. Commun.* 2020, **11**: 4317.

37. Henrich, E., Löhr, F., Pawlik, G., Peetz, O., Dötsch, V., Morgner, N., de Kroon, A.I., and Bernhard, F. (2018) Lipid conversion by cell-free synthesized phospholipid methyltransferase Opi3 in defined nanodisc membranes supports an *in trans* mechanism. *Biochemistry 57,* 5780-5784.

38. Soga, H., Fujii, S., Yomo, T., Kata, Y., Watanabe, H., and Matsuura, T. (2014) *In vitro* membrane protein synthesis inside cell-sized vesicles reveals the dependence of membrane protein integration on vesicle volume. *ACS Synth. Biol. 3,* 372-379.

39. Jacobs, M.L., Boyd, M.A., and Kamat, N.P. (2019) Diblock copolymers enhance folding of a mechanosensitive membrane protein during cell-free expression. *Proc. Natl. Acad. Sci. USA 116,* 4031-4036.

40. Harris, N.J., Pellowe, G.A., and Booth, P.J. (2020) Cell-free expression tools to study co-translational folding of alpha helical membrane transporters. *Sci. Rep. 10,* 9125.

41. Nagamori, S., Vázquez-Ibar, J.L., Weinglass, A.B., and Kaback, H.R. (2003) *In vitro* synthesis of lactose permease to probe the mechanism of membrane insertion and folding. *J. Biol. Chem. 278,* 14820-14826.

42. Matsubayashi, H., Kuruma, Y., and Ueda, T. (2014) Cell-free synthesis of secYEG translocon as the fundatmental protein transport machinery. *Origins Life Evol. Biospheres 44,* 331-334.

43. Matthies, D., Haberstock, S., Joos, F., Dötsch, V., Vonck, J., Bernhard, F., and Meier, T. (2011). Cell-free expression and assembly of ATP synthase. *J. Mol. Biol. 413,* 593-603.

44. Uyeda, A., Watanabe, T., Hohsaka, T., and Matsuura, T. (2018) Different protein localizations on the inner and outer leaflet of cell-sized liposomes using cell-free protein synthesis. *Synth. Biol. 3,* ysy007.

45. Klermund, L., Poschenrieder, S.T., and Castiglione, K. (2017) Biocatalysis in polymersomes: improving multienzyme cascades with incompatible reaction steps by compartmentalization. *ACS Catalysis 7,* 3900-3904.

46. Schoborg, J.A., Hershewe, J.M., Stark, J.C., Kightlinger, W., Kath, J.E., Jaroentomeechai, T., Natarajan, A., DeLisa, M.P., and Jewett, M.C. (2018) A cell-free platform for rapid synthesis and testing of active oligosaccharyltransferases. *Biotechnol. Bioeng. 115,* 739-750.

47. Gessesse, B., Nagaike, T., Nagata, K., Shimizu, Y., and Ueda, T. (2018) G-protein coupled receptor protein synthesis on a lipid bilayer using a reconstituted cell-free protein synthesis system. *Life 8,* 54.

48. Dudley, Q.M., Anderson, K.C., and Jewett, M.C. (2016) Cell-free mixing of *Escherichia coli* crude extracts to prototype and rationally engineer high-titer mevalonate synthesis. *ACS Synth. Biol. 5,* 1578-1588.

49. Zhang ,Y.-H.P. (2015) Production of biofuels and biochemicals by *in vitro* synthetic biosystems: opportunities and challenges. *Biotechnol. Adv. 33,* 1467-1483.

50. Bowie, J.U., Sherkhanov, S., Korman, T.P., Valliere, M.A., Opgenorth, P.H., and Liu, H. (2020) Synthetic biochemistry: the bio-inspired cell-free approach to commodity chemical production. *Trends Biotechnol. 38,* 766-778.

51. Santoro, C., Arbizzani, C., Erable, B., and Ieropoulos, I. (2017) Microbial fuel cells: from fundamentals to applications. A reivew. *J. Power Sources 356,* 225-244.

52. Xiao, X., Xia, H.-q., Wu, R., Bai, L., Yan, L., Magner, E., Cosnier, S., Lojou, E., Zhu, Z., and Liu, A. (2019) Tackling the challenges of enymatic bio(fuel) cells. *Chem. Rev. 119,* 9509-9558.

53. Korman, T.P., Opgenorth, P.H., and Bowie, J.U. (2016) A synthetic biochemistry platform for cell free production of monoterpenes from glucose. *Nat. Commun. 8,* 15526.

54. Valliere, M.A., Korman, T.P., Woodall, N.B., Khitrov, G.A., Taylor, R.E., Baker, D., and Bowie, J.U. (2019) *Nat. Commun. 10,* 565.

55. Jennewein, S., Wildung, M.R., Chau, M., Walker, K., and Croteau, R. (2004) Random sequencing of an induced *Taxus* cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proc. Natl. Acad. Sci. USA 101*, 9149-9154.

56. Brown, S., Clastre, M., Courdavault, V., and O'Connor, S.E. (2015) De novo production of the plant-derived alkaloid strictosidine in yeast. *Proc. Natl. Acad. Sci. USA 112,* 3205-3210.

57. Li, Y., Li, S., Thodey, K., Trenchard, I., Cravens, A., and Smolke, C.D. (2018) Complete biosynthesis of noscapine and halogenated alkaloids in yeast. *Proc. Natl. Acad. Sci. USA 115,* E3922-E3931.

58. Srinivasan, P., and Smolke, C.D. (2020) Biosynthesis of medicinal tropane alkaloids in yeast. *Nature 585,* 614-619.

59. Toparlak, Ö.D., Zasso, J., Bridi, S., Serra, M.D., Macchi, P., Conti, L., Baudet, M.-L., and Mansy, S.S. (2020). Artificial cells drive neural differentiation. *Sci. Adv. 6,* eabb4920.

60. Adamala, K.P., Martin-Alarcon, D.A., Guthrie-Honea, K.R., and Boyden, E.S. (2017) Engineering genetic circuit interactions within and between synthetic minimal cells. *Nat. Chem. 9,* 431-439.

61. Tsitkov, S., and Hess, H. (2019) Design principles for a compartmentalized enzyme cascade reaction. *ACS Catal. 9,* 2432-2439.

62. Rideau, E., Dimova, R., Schwille, P., Wurm, F.R., and Landfester, K. (2018) Liposomes and polymersomes: a comparative review towards cell mimicking. *Chem. Soc. Rev. 47*, 8572.

63. Has, C. and Sunthar, P. (2020) A comprehensive review on recent preparation techniques of liposomes. *J. Liposome Res. 30,* 336-365.

64. Wang, G., Zhou, Y., Ma, K., Zhang, F., Ye, J., Zhong, G., and Yang, X. (2021) Bioconversion of recombinantly produced precursor peptide pqqA into pyrroloquinoline quinone (PQQ) using a cell-free *in vitro* system. *Protein Expr. Purif. 178,* 105777.

65. Kwon, Y.-C., Oh, I.-S., Lee, N., Lee, L.-H., Yoon, Y.J., Lee, E.Y., Kim, B.-G., and Kim, D.-M. (2013) Inegrating cell-free biosyntheses of heme prosthetic group and apoenzymes for the synthesis of functional P450 monooxygenase. *Biotechnol. Bioeng. 110,* 1193-1200.

66. Sudhamsu, J., Kabir, M., Airola, M.V., Patel, B.A., Yeh, S.-R., Rousseau, D.L., and Crane, B.R. (2010) Co-expression of ferrochelatase allows for complete heme

incorporation into recombinant proteins produces in *E. coli*. *Protein Expr. Purif. 73,* 78-82.

67. Lee, K.Y., Park S.-J., Lee, K.A, Kim, S.-H., Kim, H., Meroz, Y., Mahadevan, L., Jusng, K.-H., Ahn, T.K., Parker, K.K., and Shin, K. (2018) Photosynthetic artificial organelles sustain and control ATP-dependent reactions in a protocellular system. *Nat. Biotechnol. 36,* 530-535.

68. Berhanu, S., Ueda, T., and Kuruma, Y. (2019) Artificial photosynthetic cell producing energy for protein synthesis. *Nat. Commun. 10*, 1325.

69. Nakamura, S., Suzuki, S., Saito, H., and Nishiyama, K.-I. (2018) Cholesterol blocks spontaneous insertion of membrane proteins into liposomes of phosphatidylcholine. *J. Biochem. 163,* 313-319.

70. Nomura, K., Yamaguchi, T., Mori, S., Fujikawa, K., Nishiyama, K.-I., Shimanouchi, T., Tanimoto, Y., Morigaki, K., and Shimamoto, K. (2019) Alteration of membrane physicochemical properties by two factors for membrane protein integration. *Biophys. J. 117,* 99-110.

71. Fontana, J., Sparkman-Yager, D., Zalatan J.G., and Carothers, J.M. (2020) Challenges and opportunities with CRISPR activation in bacteria for data-driven metabolic engineering. *Curr. Opin. Biotechnol. 64,* 190-198.

72. Fontana, J., Voje, W.E., Zalatan J.G., and Carothers, J.M. (2018) Prospects for engineering dynamic CRISPR-Cas transcriptional circuits to improve bioproduction. *J. Ind. Microbiol. Biotechnol. 45,* 481-490.

73. Caschera, F., and Noireaux, V. (2016) Compartmentalization of an all *E. coli* cell-free expression system for the construction of a minimal cel. *Artif. Life. 22,* 185-195.

74. Fontana, J., Dong, C., Kiattisewee, C., Chavali V.P., Tickman, B.I., Carothers, J.M., and Zalatan, J.G. (2020) Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nat. Commun. 11,* 1618.

75. Gander, M.W., Vrana, J.D., Voje, W.E., Carothers, J.M., and Klavins, E. (2017) Digital logic circuits in yeast with CRISPR-dCas9 NOR gates. *Nat. Commun. 8,* 15459.

76. Marshall, R., Maxwell, C.S., Collins, S.P., Jacobsen, T., Luo M.L., Begemann, M.B., Gray, B.N., January, E., Singer, A., He, Y., Beisel, C.L., and Noireaux, V. (2018) Rapid and scalable characterization of CRISPR technologies using an *E. coli* cell-free transcription-translation system. *Mol. Cell. 69,* 146-157.

77. Salje, J., van den Ent, F., de Boer, P., and Löwe, J. (2011) Direct membrane binding by bacterial actin MreB. *Mol. Cell. 43,* 478-487.

78. Ando, M., Schikula, S., Sasaki, Y., and Akiyoshi, K. (2018) Proteoliposome engineering with cell-free membrane protein synthesis: control of membrane protein sorting into liposomes by chaperoning systems. *Adv. Sci. 5,* 1800524.

79. Niwa, T., Sasaki, Y., Uemura, E., Nakamura S., Akiyama, M., Ando, M., Shinichi, S., Mukai, S.-a., Ueda, T., Taguchi, H., and Akiyoshi, K. (2015) Comprehensive study of liposome-assisted synthesis of membrane proteins using a reconstituted cell-free translation system. *Sci. Rep. 5,* 18025.

80. Furusato, T., Horie, F., Matsubayashi, H.T., Amikura, K., Kuruma, Y., and Ueda, T. (2018) De novo synthesis of basal bacterial cell division proteins FtsZ, FtsA, and ZipA inside giant vesicles. *ACS Synth. Biol. 7,* 953-961.

81. Das, A., Varma, S.S., Mularczyk, C., and Meling, D.D. (2014) Functional investigations of thromboxane synthase (CYP5A1) in lipid bilayers of nanodiscs. *ChemBioChem. 15,* 892-899.

82. Ravula, T., Barnaba, C., Mahajan, M., Anantharamaiah, G.M., Im, S.-C., Waskell, L., and Ramamoorthy, A. (2017) Membrane environment drives cytochrome P450's spin transition and its interaction with cytochrome b5. *Chem. Commun. 53,* 12798-12801.

83. Majumder, S., Garamella, J., Wang, Y.-L., DeNies, M., Noireaux, V., and Liu, A.P. (2017) Cell-sized mechanosensitive and biosensing compartment programmed with DNA. *Chem. Commun. 53,* 7349-7352.

84. Ohta, N., Kato, Y., Watanabe, H., Mori, H., and Matsuura, T. (2016) *In vitro* membrane protein synthesis inside Sec translocon-reconstituted cell-sized liposomes. *Sci. Rep. 6,* 36466.

85. Fujii, S., Matsuura, T., Sunami, T., Kazuta, Y., and Yomo, T. (2013) *In vitro* evolution of α-hemolysin using a liposome display. *Proc. Natl. Acad. Sci. USA 110,* 16796-16801.

86. Krishnan, M., de Leeuw, T.J.J.F., and Pandit, A. (2019) Cell-free soluble expression of the membrane protein PsbS. *Protein Expr. Purif. 159,* 17-20.

87. White, G.F., Shi, Z., Shi, L., Dohnalkova, A.C., Fredrickson, J.K., Zachara, J.M., Butt, J.N., Richardson D.J., and Clarke, T.A. (2012) Development of a proteoliposome model to probe transmembrane electron-transfer reactions. *Biochem. Soc. Trans. 40,* 1257-1260.

88. Biggs, B.W., Rouch, J.E., Kambalyal, A., Arnold, W., Lim, C.G., Mey, M.D., O'Neil-Johnson, M., Starks, C.M., Das, A., and Ajikumar, P.K. (2016) Orthogonal assays clarify the oxidative biochemistry of Taxol P450 CYP725A4. *ACS Chem. Biol. 11,* 1445-1451.

CHAPTER 6.

FUTURE DIRECTIONS

6.1 Conclusions and Future Directions

6.1.1 Expanding use of lignin as a microbial feedstock

In Chapters 2 and 3, we explored the use of lignin-derived monomers for microbial production of adipic acid in *Escherichia coli*, demonstrating proof of concept, but highlighting the need for yield improvements to make this metabolic engineering application industrially relevant. The work presented here, as well as other places in literature, makes it clear that the final enzymatic step to convert muconic acid to adipic acid must be improved to reach industrial relevance [1]. Towards this, a deeper exploration of muconic acid reductase is required. Based on the differences in whole-cell versus cell-lysate yields in MAR homologs from *Bacillus coagulans* and *Clostridium acetobutylicum*, we suggest that MAR-BC has a better turnover rate (k_{cat}) while MAR-CA has better substrate binding (K_m). Deeper analysis of the active sites of these enzymes, as well as substrate bound crystal structures, may reveal key residues responsible for substrate binding and catalysis, providing residues for rational protein design of a better MAR. In addition, structural analysis outside the active site may reveal that improved cofactor binding or electron shuttling from cofactors to substrate is responsible for improved turnover in MAR-BC, providing more potential mutation sites for rational design.

On the other hand, rational protein design could be put aside in favor of a directed evolution based approach. Previously a transcription-factor based biosensor for adipic acid

was developed with tetracycline resistance as the reporter gene [2,3]. If this dependence on adipic acid for survival is integrated into the *E. coli* genome, we can then use this engineered strain for directed evolution of an improved MAR. This would open the door to less-rational protein engineering strategies such as error-prone PCR.

In addition to enzyme engineering, an adipic acid biosensor could be used for an expanded enzyme search to identify a more stable and more active MAR. Enzyme candidates would be identified using sequence similarity searches starting from known MARs. Sequence similarity networks would be used to identify a diverse pool of MAR targets from a variety of organism sources. Important sequence similarity considerations would be making sure the screened enzymes have the required cofactor binding motifs for muconic acid reduction. Interestingly, the enzyme candidates could be screened for non-iron-sulfur cluster containing enzymes. This would help eliminate one source of O_2 sensitivity, thus improving scale up potential for this enzymatic process. A high-throughput screen would allow an enzyme candidate pool in the 1000's, with cloning support by liquid handling automation.

Finally, outside of adipic acid, lignin-derived aromatics offer advantageous feedstocks for many compounds [4]. However, cost-effective use of this heterogeneous mixture requires further development of metabolic funneling approaches where a number of feedstocks with a common core structure (ex. an aromatic ring) can be funneled to a single core metabolite that can then be upgraded into the final product [5,6]. While this aromatic core is often present in many active pharmaceutical ingredients, lignin is typically only considered as a feedstock for commodity chemicals and materials. Investigation of aromatic tailoring enzymes (such as cytochrome P450s) and advanced research into lignin

separation to establish pipelines for pure aromatics, could fuel incorporation of lignin-derived monomers into active pharmaceutical ingredients.

6.1.2 Engineering improved cyanobacterial growth

Chapter 4 presented a detailed process design for production of 2,3-butanediol using Martian CO_2 on the surface of Mars. Given the relatively slow growth of cyanobacteria on Mars compared to Earth, the analysis showed that the mass, water, and power required to drive this biotechnology process could be greatly reduced by reducing the size of the cyanobacterial growth farm. A key way to do this without adding additional infrastructure, such as artificial lighting, is to engineer more efficient photosynthesis in cyanobacteria. In the context of Mars, an intriguing strategy towards improved photosynthesis would be to take advantage of the extended range of wavelengths of light that reach the surface of Mars compared to Earth [7]. Expanding the range of photosyntheticallly active radiation could improve photosynthesis by increasing the number of photons usable by the microbe. One can imagine increasing the wavelengths of absorbed light by adding new light absorbing pigments into the cyanobacteria. This could be coupled with existing engineering strategies towards improving photosynethic efficiency such as reducing the rate of photon uptake such that less photons are lost and the photosynthetic mechanism can funnel all absorbed light towards carbon fixation [8].

On the process side, this biotechnology approach could be expanded to produce any fermentation reachable compound, from materials to pharmaceuticals. An interesting analysis would look at the amount of cyanobacteria it would take to produce all the required compounds to support a Martian colony. Finally, a detailed life cycle analysis of the

process would help better characterize the environmental impacts of implementing this process including biological waste production, as well as infrastructure waste from the reactor materials.

6.1.3 Applications of membrane associated proteins in cell-free systems

In Chapter 5, we provide a literature review of the use of membrane proteins in cell-free systems (CFSs) with focus on how this technology could be used towards metabolic engineering applications. The use of membrane proteins in CFSs will open up a wealth of opportunities for chemical production, especially towards producing natural products. This will be especially advantageous as many natural products are toxic to microbes, and thus cannot be produced at high titers [9]. With CFSs, this challenge is bypassed. The first step along this pathway will be successful expression of a functional membrane-associated enzyme in a CFS as part of a biosynthetic pathway. The next steps will be to use CFSs as a rapid screening method for pathway elucidation. As a quick expression platform without any side reactions, potential enzymes could be rapidly screened with low background. Also, enzymes from multiple sources could be screened simultaneously by adding CFS encapsulated in liposomes of different compositions into the same reaction as a method for combinatorial pathway building.

Further, as a complement to pathway elucidation, combinations of enzymatic and heterogeneous catalytic pathways could be implemented. For unknown steps in a metabolic pathway, there may be chemical catalysts known to perform this reaction. The use of a CFS, as opposed to a living organism, allows for *in situ* combinations of biological and

chemical catalysis to build a highly optimized pathway that switches between the two catalytic methods, without the need for costly separation steps in between.

6.2 References

1. Raj, K. et al. Biocatalytic production of adipic acid from glucose using engineered *Saccharomyces cerevisiae*. *Metab. Eng. Commun.* **6,** 28-32 (2018).

2. Thompson, M.G. et al. Robust characterization of two distinct glutarate sensing transcription factors of *Pseudomonas putida* l-lysine metabolism. *ACS Synth. Biol.* **8,** 2385-2396 (2019).

3. Dietrich, J.A., Shis, D.L., Alikhani, A. & Keasling, J.D. Transcription factor-based screens and synthetic selections for microbial small-molecule biosynthesis. *ACS Synth. Biol.* **2,** 47-58 (2013).

4. Yaguchi, A.L, Lee, S.J. & Blenner, M.A. Synthetic biology towards engineering microbial lignin biotransformation. *Trends Biotechnol.* (2021)

5. Linger, J.G. et al. Lignin valorization through integrated biological funneling and chemical catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 12013-12018 (2014).

6. Mallinson, S.J.B. et al. A promiscuous cytochrome P450 aromatic O-demethylase for lignin bioconversion. *Nat. Commun.* **9,** 2487 (2018).

7. Verseux, C. et al. Sustainable life support on Mars – the potential roles of cyanobacteria. *Int. J. Astrobiol.* **15,** 65-92 (2015).

8. Hu, G., Li, Y., Ye, C., Liu, L. & Chen, X. Engineering microorganisms for enhanced CO_2 sequestration. *Trends Biotechnol.* **37,** 532-547 (2019).

9. Liu, X., Ding, W. & Jiang, H. Engineering microbial cell factories for the production of plant natural products: from design principles to industrial-scale production. *Microb. Cell Fact.* **16,** 125 (2017).

APPENDIX A.

TABLE OF PLASMIDS

A.1 Plasmids used in this work

Table A.1. Plasmids used in Chapter 3

Strain Number	Plasmid Name	Description	Source
PPY2109	pBbA1a	p15a (medium copy), pTrc, ampR	Lee et al [1]
PPY2110	pBbA1c	p15a, pTrc, CmR	Lee et al [1]
PPY2111	pBbS1a	SC101 (low copy), pTrc, ampR	Lee et al [1]
PPY2112	pBbB1c	BBR1 (high copy), pTrc, CmR	Lee et al [1]
PPY2113	pBbA5c	p15a, lacUV5, CmR	Lee et al [1]
PPY2114	pBbS5c	SC101, lacUV5, CmR	Lee et al [1]
PPY2115	pBbB5a	BBR1, lacUV5, ampR	Lee et al [1]
PPY2116	pBbA7a	p15a, T7, ampR	Lee et al [1]
PPY2117	pBbS7a	SC101, T7, ampR	Lee et al [1]
PPY1521	pNK45	pBbS1a-CatA-AN22*	This Study, Chapter 3
PPY1626	pNK53	pBbA1a-CatA-putida	This Study, Chapter 3
PPY1627	pNK54	pBbA1a-CatA-albicans	This Study, Chapter 3
PPY1628	pNK55	pBbA1a-CatA-opacus	This Study, Chapter 3
PPY1629	pNK56	pBbA1a-CatA-AN22	This Study, Chapter 3
PPY1630	pNK57	pBbA1a-CatA-ADP1	This Study, Chapter 3
PPY1631	pNK58	pBbS1a-CatA-putida	This Study, Chapter 3
PPY1632	pNK59	pBbS1a-CatA-albicans	This Study, Chapter 3
PPY1633	pNK60	pBbS1a-CatA-opacus	This Study, Chapter 3

Table A.1 (continued)

PPY1634	pNK61	pBbS1a-CatA-AN22	This Study, Chapter 3
PPY1635	pNK62	pBbS1a-CatA-ADP1	This Study, Chapter 3
PPY1636	pNK63	pBbA5c-CatA-putida	This Study, Chapter 3
PPY1637	pNK64	pBbA5c-CatA-albicans	This Study, Chapter 3
PPY1638	pNK65	pBbA5c-CatA-opacus	This Study, Chapter 3
PPY1639	pNK66	pBbA5c-CatA-AN22	This Study, Chapter 3
PPY1640	pNK67	pBbA5c-CatA-ADP1	This Study, Chapter 3
PPY1641	pNK68	pBbS5c-CatA-putida	This Study, Chapter 3
PPY1642	pNK69	pBbS5c-CatA-albicans	This Study, Chapter 3
PPY1643	pNK70	pBbS5c-CatA-opacus	This Study, Chapter 3
PPY1644	pNK71	pBbS5c-CatA-AN22	This Study, Chapter 3
PPY1645	pNK72	pBbS5c-CatA-ADP1	This Study, Chapter 3
PPY1646	pNK73	pBbA7a-CatA-putida	This Study, Chapter 3
PPY1647	pNK74	pBbA7a-CatA-albicans	This Study, Chapter 3
PPY1648	pNK75	pBbA7a-CatA-opacus	This Study, Chapter 3
PPY1649	pNK76	pBbA7a-CatA-AN22	This Study, Chapter 3
PPY1650	pNK77	pBbA7a-CatA-ADP1	This Study, Chapter 3
PPY1651	pNK78	pBbS7a-CatA-putida	This Study, Chapter 3
PPY1652	pNK79	pBbS7a-CatA-albicans	This Study, Chapter 3

Table A.1 (continued)

PPY1653	pNK80	pBbS7a-CatA-opacus	This Study, Chapter 3
PPY1654	pNK81	pBbS7a-CatA-AN22	This Study, Chapter 3
PPY1655	pNK82	pBbS7a-CatA-ADP1	This Study, Chapter 3
PPY1774	pNK102	pBbA1a-MAR-BC	This Study, Chapter 3
PPY1776	pNK111	pBbA1a-MAR-BC_CatA-AN22**	This Study, Chapter 3
PPY1967	pNW1	pBbA1a-MAR-CA_CatA-AN22**	This Study, Chapter 3
PPY1973	pNW5	pBbB1c-MAR-BC	This Study, Chapter 3
PPY1976	pNW8	pBbB5a-MAR-BC	This Study, Chapter 3
PPY1983	pNK146	pBbA1a-MAR-CA	This Study, Chapter 3
PPY1984	pNK147	pBbB5a-MAR-CA	This Study, Chapter 3
PPY1985	pNK148	pBbB1c-MAR-CA	This Study, Chapter 3
PPY1986	pNK149	pBbA5c-MAR-CA	This Study, Chapter 3
PPY1987	pNW10	pBbA5c-MAR-BC	This Study, Chapter 3
PPY2087	pNK159	pBbA1c-MAR-BC	This Study, Chapter 3
PPY2088	pNK160	pBbS5c-MAR-BC	This Study, Chapter 3
PPY2098	pNK161	pBbS5c-MAR-CA	This Study, Chapter 3
PPY2090	pNK162	pBbS1a-MAR-BC	This Study, Chapter 3
PPY2091	pNK163	pBbS1a-MAR-CA	This Study, Chapter 3
PPY2104	pNK167	pBbS1a-6His-CatA-putida	This Study, Chapter 3

Table A.1 (continued)

PPY2105	pNK168	pBbS1a-6His-CatA-albicans	This Study, Chapter 3
PPY2106	pNK169	pBbS1a-6His-CatA-opacus	This Study, Chapter 3
PPY2107	pNK170	pBbS1a-6His-CatA-ADP1	This Study, Chapter 3
PPY2122	pNK171	pBbS1a-6His-CatA-AN22	This Study, Chapter 3

All plasmids have RBS1 (see Table D.1) unless indicated with *
* = single gene that had RBS2 (see Table D.1)
** = operon plasmid, MAR had RBS1 and CatA had RBS2.

A.2 References

1. Lee, T.S. et al. BglBrick vectors and datasheets: a synthetic biology platform for gene expression. *J. Biol. Eng.* 12 (2011).

APPENDIX B.

TABLE OF STRAINS

B.1 Strains used in this work

Table B.1. Strains used in Chapter 3

Strain #	Description	Source
PPY252	*Escherichia coli* DH10B	Invitrogen
PPY2108	*Escherichia coli* BL21 (DE3)	NEB
PPY1535	*Escherichia coli* K-12 BW25113 Δ*iscR*	Baba et al [1]

In chapter 3, PPY252 (*E. coli* DH10B) was used for cloning and muconic acid production for pTrc and lacUV5 containing plasmids. PPY2108 (*E. coli* BL21 (DE3)) was used for muconic acid production for pT7 containing plasmids. PPY1535 (*E. coli* Δ*iscR*) was used for all adipic acid producing strains.

B.2 References

1. Baba, T. et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2, 1-11 (2006).

APPENDIX C.

PRIMER TABLES

C.1 Primers used in this work

Table C.1. Primers used in Chapter 3

Primer Name	Sequence
NK18	gacaattaatcatccggctcg
NK19	CACTTTATGCTTCCGGCTCG
NK20	cgcgaaattaatacgactcactatag
NK22	ccgacaaacaacagataaaacg
NK26	GGATCCTTAACCTTCTTGCAGTGCGC
NK27	GGATCCTTATAACTTAATCTCCGCATCCTGACG
NK28	GGATCCTTATGCCGGATCTAAAACAAAATTATAGG
NK29	GGATCCTTATGCTTCCGGGTCCAGC
NK30	GGATCCTTACACTGCCAGGCGCG
NK31	AGATCTATGCACCATCACCATCACCATACTGTTAAAATCAGCCATACAGC
NK32	AGATCTATGCACCATCACCATCACCATTCACAGGCGTTTACAGAAAGCG
NK33	AGATCTATGCACCATCACCATCACCATACCACAACAGAAAGTCCTACAGC
NK35	AGATCTATGCACCATCACCATCACCATGAGGTTAAGATTTTTAATACCCAAGATG
NK40	AGATCCAGACGATAGCGACCTACCCAGGAGGAAATAACATGACTACAATGGAGAATCCGACC
NK41	GAACCCAACATTTATTATAATAAGGAGTTAAACAATGACTACAATGGAGAATCCGACC
NK43	CGGATAACAATTTCAGAATTCAAAAGATCTAGATCCAGACGATAGCGACC
NK47	GCCTGGAGATCCTTACTCGAGTTTGGGGATCCTTATGCTTCCGGG
NK44	CGGATAACAATTTCAGAATTCAAAAGATCTGAACCCAACATTTATTATAATAAGG

Table C.1 (continued)

NK63	AGATCCAGACGATAGCGACCTACCCAGGAGGAAATAACATGCACCATCACCATCACC
NK64	AGATCCAGACGATAGCGACCTACCCAGGAGGAAATAACATGACTGTTAAAATCAGCCATACAGC
NK65	GCCTGGAGATCCTTACTCGAGTTTGGATCCTTAACCTTCTTGCAGTGC
NK66	AGATCCAGACGATAGCGACCTACCCAGGAGGAAATAACATGTCACAGGCGTTTACAGAAAGC
NK67	GCCTGGAGATCCTTACTCGAGTTTGGATCCTTATAACTTAATCTCCGC
NK68	AGATCCAGACGATAGCGACCTACCCAGGAGGAAATAACATGACCACAACAGAAAGTCC
NK69	GCCTGGAGATCCTTACTCGAGTTTGGATCCTTATGCCGGATCTAAAACAAAATTATAGG
NK70	AGATCCAGACGATAGCGACCTACCCAGGAGGAAATAACATGGAGGTTAAGATTTTTAATACCC
NK71	GCCTGGAGATCCTTACTCGAGTTTGGATCCTTACACTGCCAGG
NK75	GCCTGGAGATCCTTACTCG
NK82	AGATCCAGACGATAGCGACCTACCCAGGAGGAAATAACATGAAATACAAAAAGCTATTTGAAACTGTG
NK83	AGATCCAGACGATAGCGACCTACCCAGGAGGAAATAACATGAACAAATACAAGAAATTATTTGAACCAATC
NK86	GTGAGCGGATAACAATTTCAGAATTC
NK88	GGTCTGTTTCCTGTGTGAAATTGTTATCCTTACAGATTTGCCGCGACTTCG
NK89	AGGATAACAATTTCACACAGGAAACAGACCGAACCCAACATTTATTATAATAAGGAGTTAAACAATG
NK153	GGATCCAAACTCGAGTAAGGATC
NK154	CTTTTGAATTCTGAAATTGTTATCCGC

APPENDIX D.

DNA SEQUENCES

D.1 Chapter 3 genes

D.2.1 Pseudomons putida CatA (WP_010954549)

ATGACTGTTAAAATCAGCCATACAGCAGATATTCAGGCATTCTTTAATAGAGT

TGCAGGTTTAGATCATGCAGAAGGCAATCCGCGTTTTAAACAGATTATTCTGC

GCGTTTTACAAGATACAGCACGTTTAATTGAAGATTTAGAAATTACCGAAGA

TGAATTTTGGCATGCAGTGGATTATCTGAACCGTCTTGGCGGACGCAACGAA

GCCGGTTTACTGGCAGCAGGCCTTGGTATTGAACATTTTCTGGATCTGTTACA

GGACGCAAAAGACGCAGAAGCAGGTCTGGGCGGTGGTACCCCGCGCACCAT

TGAAGGCCCGCTGTATGTTGCAGGTGCACCGCTGGCACAGGGCGAAGCACGC

ATGGATGATGGTACCGATCCGGGCGTTGTTATGTTTCTGCAAGGTCAGGTATT

TGACGCAGATGGTAAACCGCTGGCAGGCGCAACCGTTGATCTGTGGCATGCG

AATACCCAGGGTACCTATAGCTATTTTGATAGCACCCAGAGCGAATTTAACCT

GCGCCGTCGTATTATTACCGATGCAGAAGGCCGTTATCGTGCACGCAGCATT

GTGCCGAGCGGCTATGGCTGTGATCCGCAGGGTCCGACCCAGGAATGTCTGG

ATCTGCTGGGCCGTCATGGTCAGCGCCCGGCACATGTTCATTTCTTTATTAGC

GCACCGGGTCATCGCCATCTGACCACCCAGATTAATTTTGCAGGTGATAAAT

ATCTGTGGGATGATTTTGCGTATGCAACCCGCGATGGTCTGATTGGTGAACTG

CGCTTTGTGGAAGATGCGGCAGCAGCACGCGATCGCGGTGTGCAGGGTGAAC

GTTTTGCGGAACTGAGTTTTGACTTTCGCCTGCAAGGCGCTAAAAGCCCGGAT

GCAGAAGCACGCAGCCATCGCCCGCGCGCACTGCAAGAAGGTTAA

D.2.2 Acinetobacter baylyi ADP1 CatA (CAG68305)

ATGGAGGTTAAGATTTTTAATACCCAAGATGTGCAAGACTTCCTGCGCGTTGC

AAGCGGCCTGGAACAAGAAGGCGGTAACCCGCGTGTGAAACAAATTATTCAT

CGCGTTCTGAGTGATTTATATAAAGCAATTGAAGACTTAAACATTACCAGTG

ATGAATATTGGGCAGGTGTGGCATATCTGAACCAGCTGGGCGCAAACCAGGA

AGCAGGCCTGTTAAGCCCTGGCCTGGGCTTTGATCATTATTTAGATATGCGCA

TGGATGCAGAAGATGCAGCACTGGGCATTGAAAACGCAACACCGCGCACCA

TTGAAGGCCCGTTATATGTGGCAGGCGCACCTGAAAGTGTGGGTTACGCACG

CATGGATGATGGCAGCGATCCGAATGGTCATACCCTGATCCTGCACGGCACC

ATTTTTGATGCAGATGGCAAACCTCTGCCGAACGCAAAAGTTGAAATCTGGC

ACGCAAACACCAAGGGCTTCTATAGCCATTTTGATCCTACCGGCGAACAACA

GGCATTTAACATGCGTCGCAGCATTATTACAGATGAGAACGGTCAGTATCGC

GTTCGCACCATTCTGCCGGCAGGCTATGGCTGCCCGCCGGAAGGCCCGACCC

AACAGCTGTTAAACCAGTTAGGTCGCCACGGTAATCGCCCGGCACATATTCA

TTATTTTGTGAGCGCAGATGGCCATCGCAAACTGACCACCCAGATTAACGTTG

CAGGCGATCCGTATACCTATGATGATTTTGCATACGCAACCCGCGAGGGCTT

AGTTGTGGATGCAGTGGAACATACCGATCCGGAAGCAATTAAAGCAAACGAT

GTGGAAGGCCCGTTTGCAGAAATGGTGTTCGATCTGAAACTGACACGCCTGG

TGGATGGTGTGGATAACCAGGTGGTTGATCGCCCGCGCCTGGCAGTGTAA

D.2.3 Candida albicans CatA (KCQ97177)

ATGTCACAGGCGTTTACAGAAAGCGTGAAAACAAGCCTGGGTCCGAACGCGA

CCCCTCGCGCGAAGAAGCTGATTGCAAGTTTAGTGCAGCATGTGCATGACTTT

GCACGCGAAAACCATCTGACAACAGAAGATTGGCTTTGGGGTGTGGATTTTA

TTAATCGCATTGGTCAAATGAGCGATAGTCGCCGCAATGAAGGCATTTTAGT

GTGCGATATTATTGGTCTGGAAACCCTGGTGGATGCCCTGACCAACGAAAGT

GAACAAAGCAACCATACAAGTAGCGCGATTCTGGGTCCGTTTTATTTACCGG

ACAGCCCGGTGTATCCGAATGGCGGTAGCATTGTTCAGAAGGCAATTCCGAC

CGACGTTAAATGTTTTGTGCGCGGTAAAGTTACCGACACCGAAGGTAAACCG

TTAGGCGGTGCACAGCTGGAAGTTTGGCAGTGTAATAGCGCAGGTTTCTATA

GCCAACAGGCAGATCATGATGGCCCGGAATTTAATCTGCGTGGTACATTTATT

ACCGATGATGAAGGTAATTATAGCTTTGAATGCCTGCGGCCTACCAGCTACC

CGATTCCGTATGATGGCCCGGCAGGTGACCTGCTGAAAATTATGGATCGCCA

TCCGAATCGTCCTAGCCACATTCATTGGCGCGTGTCGCATCCGGGTTATCATA

CCCTGATTACCCAGATTTATGATGCGGAATGTCCGTATACCAACAACGATAG

CGTGTATGCGGTTAAAGATGATATTATTGTGCATTTTGAAAAGGTGGATAAC

GATGATAAAGACCTGGTGGGTAAAGTGGAATATAAACTGGATTATGATATTA

GCCTGGCGACCGAAAGCAGTATTCAGGAAGCGCGCGCGGCAGCGAAAGCGC

GTCAGGATGCGGAGATTAAGTTATAA

D.2.4 Rhodococcus sp. AN22 CatA (BAH56722)

ATGACTACAATGGAGAATCCGACCGCGCATGGTAGCGGCAATGCGGCAACC

GATAAATTTAAAAGTGAACGCGTTACAAGCGACACAAGCGTTGAACGCGCGA

GCGCAATTTATAAAGATTTACTGGATGCGCTGGCAGGCATTGTTGATAAACA

TCAGGTGACCTATGATGAATATAGAGTGTTAAAACAATGGCTGATTGATGTG

GGAGAATATGGTGAATGGCCTTTATGGTTAGATGTGTTTCTTGAACATGAGAT

TGAGAAAGTTCATTACAATCGCAAAGGTTTTACCGGCACCAAAGGTAGTATT

GAAGGCCCTTATTATGTGCCGGATAGCCCTAAATTACCGAGTAAATGCACCA

TGCCTATGCGCGAGAAGGATAAAGTGGCGCCGCCTCTGGTGTTTAAAGGTCA

GGTGACCGATTTAGAAGGCAATGGTCTGCCGGGCGCAACCGTGGAACTGTGG

CATGCGGATGAAGAGGGCTTTTACAGCCAGTTTGCGCCGGGCATTCCTGAGT

GGAACCTGCGCGGCACAGTGGAAGTGGATGAAAACGGCAACTTTGAAATTAC

AACACTGAAACCGGCGCCGTATCAGATTCCGAGCGATGGCCCGACCGGCTGG

TTTATCAAAAGTTATGGCGGCCACCCGTGGAGACCGGCGCATCTGCATCTTAT

GGTGAAAGCACCTGGCAAACGCACCATTACCACCCAACTTTATTTTCAAGGC

GGCGAATGGGTCGAAGATGATGTGGCAACAGCGGTTAAACCGGAACTGATTC

TTGACCCGCAACCGAACGCGGATGGCGTGGCGGAAGTGACCTATACCTTCGT

GCTGGACCCGGAAGCATAA

D.2.5 Rhodococcus opacus CatA (3HGI_A)

ATGACCACAACAGAAAGTCCTACAGCAGCAGGTTCAGGTAGCGCAGCAACA

GATAAATTCAAAGCTGAACGCGCAACAGCAGATACAAGCCCGGAACGCCTG

GCAGCAATTGCAAAAGATGCATTAGGCGCACTGAACGATGTGATTTTAAAAC

ATGGCGTTACCTATCCGGAATATCGCGTGTTTAAACAATGGCTGATTGATGTG

GGCGAAGGCGGTGAATGGCCATTATTTCTGGATGTGTTTATTGAACATAGCGT

GGAAGAAGTGCTGGCACGCAGTCGCAAAGGTACAATGGGTTCTATCGAAGGC

CCGTATTATATTGAAAACAGCCCGGAACTGCCTAGCAAATGCACATTACCGA

TGCGCGAAGAAGATGAGAAAATTACCCCGCTGGTGTTTAGCGGCCAGGTGAC
CGATCTGGATGGCAATGGCCTGGCAGGCGCGAAAGTGGAACTGTGGCATGCA
GATAACGATGGTTATTATAGCCAGTTTGCGCCTCATTTACCGGAATGGAACCT
GCGCGGCACCATCATTGCGGATGAAGAAGGTCGTTATGAAATTACCACCATT
CAGCCGGCACCGTATCAGATTCCGACCGATGGCCCGACCGGCCAGTTTATTG
AAGCGCAGAACGGTCATCCGTGGCGCCCGGCACATCTGCATCTGATTGTGAG
CGCACCGGGCAAGGAAAGCGTTACCACCCAGCTGTATTTTAAAGGCGGTGAA
TGGATTGATAGCGATGTGGCAAGCGCGACCAAACCGGAACTGATTCTGGACC
CAAAGACCGGCGATGATGGCAAGAACTATGTGACCTATAATTTTGTTTTAGA
TCCGGCATAA

D.2.6 Bacillus coagulans MAR (AEO99944)

ATGAAATACAAAAAGCTATTTGAAACTGTGAAAATAAGGAATGTGGAACTCA
AAAATCGTTATGCAATGGCACCAATGGGTCCGCTGGGTCTTGCCGATGCAGA
AGGCGGTTTCAACCAGCGCGGGATTGAGTATTATACAGCCCGTGCGCGCGGG
GGAACCGCTCTGATTATTACCGGCGTCACTTTCGTTGATAATGAAGTGGAAG
AGCACGGAATGCCAAACGTACCTTGCCCGACCCATAACCCTGTCCATTTTGTC
CGGACTTCCAAAGAAATGACAGAGCGCATCCATGCATATGATTCGAAAATTT
TTCTGCAAATGAGCGCCGGTTTTGGCCGGGTGACGATCCCGACAAACCTTGG
CGAGTACCCGCCGGTTGCACCGTCGCCAATCCCGCATCGCTGGCTGGATAAA
ACATGTCGCGAACTGACAGTTGAAGAAATTCATTCCATTGTCCGCAAATTCG
GGGATGGGGCGTTCAATGCGAAGCGCGCCGGATTTGACGGGGTGCAAATCCA
TGCTGTGCACGAAGGCTATTTGCTCGACCAGTTTGCGATTGCGTTTTTCAACA

AACGTACCGATGCATACGGTGGCCCGCTTGAAAATCGCCTTCGTTTTGCCCGG

GAAATTGTCGAGGAAATTAAACAGCGCTGTGGCGAAGATTTTCCTGTGACGC

TCCGCTTCAGCCCGAAAAGTTTTATCAAGGATTGGCGGGAAGGGGCACTGCC

TGGCGAGGAGTTTGAAGAAAAAGGCCGCGATTTGGATGAAGGCATCGAGGC

AGCAAAGCTGCTCGTTTCCTACGGCTATGATGCTCTGGACGTCGATGTTGGTT

CTTATGATTCATGGTGGTGGAGCCATCCTCCGATGTACCAGAAGAAGGGGCT

TTACATTCCGTATGCCAGGCTGGTGAAGGAAGCTGTCGATGTGCCTGTCCTTT

GCGCGGGCCGCATGGACAATCCGGATCTTGCACTTGCCGCACTGGAAGACGG

AGCATGTGATATTATCAGCTTGGGCCGCCCGTTATTGGCTGACCCGGATTACG

TCAATAAGCTCCGAATCGGGCAGGTTGCCGATATCCGCCCGTGTCTGTCATGC

CATGAAGGCTGCATGGGTCGGATCCAGGAGTATTCTTCCTTAGGCTGCGCAG

TGAATCCGGCTGCCTGTCGAGAAAAAGAAGCAGCATTGACACCTGCTTTAAA

AAAGAAACGCGTACTGATTGCAGGCGGCGGCGTGGCCGGATGCGAAGCTGC

CCGTGTGCTTGCATTGCGCGGCCATGAACCGGTCATTTTTGAAAAATCGAACC

GTTTAGGCGGCAACTTGATCCCTGGCGGCGCACCTGATTTTAAAGAAGATGA

CCTGGCGCTTGTTGCCTGGTATGAGCATACGTTGGAACGCCTTGGCGTAGAA

ATTCATTTGAATACTGCATTGACAAAAGAAGAAATTTTGGCTGCAAACGTGG

ATGCCGTGCTGATTGCAACGGGTTCGAATCCGAAAATTTTGCCGCTCGACGG

AAAAAACAAAGTATTTACAGCAGAAGATGTTTTGCTCGATAAAGTGGATGCC

GGGCAACATGTTGTCATTGTCGGCGGCGGTCTTGTCGGCTGCGAACTGGCTTT

GAACCTTGCAGAAAAAGGAAAAGATGTCTCGCTTGTGGAAATGCAGGACAA

ACTGCTGGCAGTTAATGGTCCGCTTTGCCACGCTAACTCGGACATGCTGGAA

AGACTCGTACCGTTTAAAGGTGTTCAAGTCTACACTTCTTCAAAAATAGTAGA

TACGACAGAAAAGACAGCCGTTGTGGATGTTGACGGCGAATTGCGTGAAATT
GAAGCAGACAGCATTGTGCTCGCAGTCGGCTACTCGGCTGAAAAATCACTCT
ATGAAGATTTAAAGTTTGAGGTTGCCGATCTTCATGTGGTTGGCGATGCCCGC
AAGGTCGCAAACATCATGTATGCCATCTGGGATGCTTACGAAGTCGCGGCAA
ATCTGTAA

D.2.7 Clostridium acetobutylicum MAR (AEI32805)

ATGAACAAATACAAGAAATTATTTGAACCAATCAAAATTGGAAAATGTGAAA
TCAAAAACCGTTTTGCATTAGCTCCAATGGGCCCTTTAGGACTAGCTGATAGT
GAAGGTGGTTTCAACCAAAGAGGAATAGACTACTATACTGAAAGAGCAAAA
GGTGGCACAGGATTAATAATAACAGGAGTTACCTTTGTAGATAATGAAGTTG
AAGAACACGGAATGCCTAATTGTCCTTGTCCAACACATAATCCAGTTCAATTC
GTAAGAACTGGTAGAGAAATGACTGAAAGAATACACGCATACAATTCTAAA
GTATTTTTACAAATGTCAGGTGGATTTGGTAGAGTTACTATACCTACTAACTT
AGGAGAATTTCCTCCAGTTGCCCCATCTCCAATTCAACATAGATGGCTTGACA
AAACTTGTCGTGAACTTACAGTAGATGAAATTAAATCAATAGTTAAAAAATT
TGGTGAAGGAGCTTTTAATGCTAAAAGGGCCGGCTTTGATGGAGTTCAAATT
CATGCTGTTCATGAAGGATACCTTATAGATCAATTTGCTATTTCATTATTTAAT
CATAGAACCGATGAATACGGCGGAAGCTTAGAAAATAGACTTCGCTTTGCAA
GAGAAATCGTTGAAGAAATTAAAAATCGCTGTGGAGAAGATTTCCCTGTAAC
ACTTAGATATTCACCAAAAAGCTTTATTAAAGATCTTAGAGATGGAGCACTTC
CTGGTGAAGAATTCGTTGAAAAGGGAAGAGACCTTGACGAAGGTGTTGAGGC
TGCAAAACTTCTTGTATCTTATGGATATGATGCTTTAGATACAGATGTTGGTT

CTTATGATTCATGGTGGTGGAGTCATCCGCCTATGTACCAGGAAAAAGGCTT

ATATAGAAAATACGCTAAATTAATGAAGGATACTGTTGATGTTCCAGTTATTT

GCGCTGGAAGAATGGATGATCCTGATATGGCCTTAGAAGCTGTAGAAAATGG

AACCTGCGATGTTATAAGTCTAGGAAGACCTCTTCTTGCAGACCCTGACTACG

TAAATAAGTTAAGAAGTAATAAATGCAAATCAATAAGACCTTGTATTTCCTG

TCAAGAAGGTTGTATGGGACGTGTTCAACATTACTCAATGTTAAACTGCGCTG

TAAACCCTCAAGCTTGTAAGGAAAGAGCTAACTCACTTACTCCAATAATTAA

AAGCAAAAAAGTATTAATAGTTGGAGGAGGAGTTGCTGGCTGTGAAGCTGCT

AGAGTTCTAGCTCTTAGAGGTCATGAACCTGTACTTTATGAAAAGAGCAATA

GATTAGGCGGAAATCTTATACCTGGTGGAGCACCAAGCTTTAAAGAAGATGA

CATAGCATTAGCTGATTGGTATACAAATACCTTAAAAGAGCTAAACGTTGAA

GTCAACTTAAATAGCGAGGTTACAAAAGAACAAATTTTAAATTCCAAGTTTG

ATACAGTAATCGTAGCAACAGGATCAACTCCAAAGGTTTTCCCACTTGGAGA

TGACGAAAAAGTATTCACCGCTGCTGAAGTATTACTAGGACAAAAAGATCCT

GGAGAAACAACTGTTGTAGTTGGAGGAGGTCTAGTAGGCTGCGAATTAGCAT

TAGATCTTGCTAAAAAAGGCAAAAAGGTAACTATTGTTGAAGCCTTAAATAA

AATACTAGCTTTAAATGGTCCTTTATGTTCTGCAAACAGCGAAATGCTTCAAA

AATTAATACCTTTTAATGGCATCGATGTAAAGGCAAATTCAAAAGTAAAAGG

ATACAAAAATGGATTGCTTAAAATGGAAACAGAAAACGGAATAGAAGAATT

ACCATGTGATTCAGTAATATTATCTGTTGGATATAAAGAAGAAAACTCCTTAT

ACAAGGAATTAGAATTTGAAATTCCAGAAATCTACCTTCTAGGAGATGCTCG

TAAGGTATCTAATATCATGTATGGTATTTGGGATGCTTTTGAAGTTGCAAACC

ATATATAA

D.2 Ribosomal binding site (RBS) sequences

Table D.1. RBS sequences

Name	Nucleotide Sequence	Predicted Translation Initiation Rate (a.u.) [1-2]
RBS1	agatccagacgatagcgacctacccaggaggaaataac	6161.40
RBS2	gaacccaacatttattataataaggagttaaaca	74751.05

D.3 References

1. Espah Borujeni, A. & Salis, H.M. Translation initiation is controlled by RNA folding kinetics via a ribosome drafting mechanism. *J. Am. Chem. Soc.* 138, 7016-7023 (2016).

2. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* 27, 946 (2009).

APPENDIX E.

FLUX BALANCE ANALYSIS OF METABOLIC PATHWAYS

E.1 Flux Balance Analysis of Adipic Acid Producing Pathways

To calculate the maximum theoretical yields for each pathway we used flux balance analysis. The calculations were done using Constraint-Based Reconstruction and Analysis (COBRA) toolbox for Matlab [1]. Genome scale models for *Escherichia coli* [2], *Pseudomonas putida* [3], and *Saccharomyces cerevisiae* [4] were edited to include the desired pathways. The equations added for each reaction are presented in the following subsections. For each reaction, the enzyme name and E.C. number used to determine cofactor use is provided.

The model was run with glucose uptake rate (EX_glc__D(e)) was set to -10 for *E. coli* and *S. cerevisiae* models. For the reactions in *P. putida*, the PCA dioxygenase reaction was eliminated using the removeRxns formula in the COBRA Toolbox. Furthermore, in order to ensure that all muconic acid production came from aromatic precursors, the glucose uptake rate in the *P. putida* models was set to 0. Finally, we attempted to prevent further conversion of muconic acid by eliminating the MUCCYCI reaction in the *P. putida* model. However, when this was attempted the solver did not converge.

E.1.1 Muconic acid production from dehydroxyshikimic acid

DHS → PCA → catechol → muconic acid

3DhskD: 3dhsk[c] → pca[c] + h2o[c]

(3DhskD= 3-dehydroshikimate dehydratase - EC 4.2.1.118)

PcaDC: pca[c] → cat[c] + h2o[c]

(PcaDC= protocatechuate decarboxylase – EC 4.1.1.63)

CatDO: cat[c] + o2[c] → mua[c]

(CatDO= catechol-1,2-dioxygenase - EC 1.13.11.1)

MUAe: mua[c] →mua[e]

EX_MUA: mua[e] →

(Exchange Reaction)

E.1.2 Muconic acid production from anthranilate

Anthranilate → catechol → muconic acid

AnthD: anth[c] + o2[c] + nadh[c] + 2h[c] → cat[c] + co2[c] + nad[c] + nh4[c]

(AnthD= anthranilate-1,2-dioxygenase - EC 1.14.12.1)

CatDO: cat[c] + o2[c] → mua[c]

(CatDO= catechol-1,2-dioxygenase - EC 1.13.11.1)

MUAe: mua[c] →mua[e]

EX_MUA: mua[e] →

(Exchange Reaction)

E.1.3 Muconic acid production from p-hydroxybenzoate

PHB → PCA → catechol → muconic acid

PHBO: 4hbz[c] + o2[c] + nadph[c] + 2h[c] → pca[c] + nadp[c] + h2o[c]

(PHBO= p-hydroxybenzoate monooxygenase - EC 1.14.13.2)

PcaDC: pca[c] → cat[c] + h2o[c]

(PcaDC= protocatechuate decarboxylase – EC 4.1.1.63)

CatDO: cat[c] + o2[c] → mua[c]

(CatDO= catechol-1,2-dioxygenase - EC 1.13.11.1)

MUAe: mua[c] →mua[e]

EX_MUA: mua[e] →

(Exchange Reaction)

E.1.4 Muconic acid production from isochorismate

Isochorismate → salicylate → catechol → muconic acid

ICL: ichor[c] → sca[c] + pyr[c]

(ICL= isochorismate lyase - EC 4.2.99.21)

ScaO: sca[c] + o2[c] + nadh[c] + h[c] → cat[c] + co2[c] + nad[c] + h2o[c]

(ScaO= salicylate 1-monooxygenase - EC 1.14.13.1)

CatDO: cat[c] + o2[c] → mua[c]

(CatDO= catechol-1,2-dioxygenase - EC 1.13.11.1)

MUAe: mua[c] →mua[e]

EX_MUA: mua[e] →

(Exchange Reaction)

E.1.5 Muconic acid production from 2,3-dihyroxybenzoate

2,3-DHB → catechol → muconic acid

23DhbDC: 23dhb[c] → cat[c] + co2[c]

(23DhbDC= 2,3-Dihydroxybenzoate decarboxylase - EC 4.1.1.46)

CatDO: cat[c] + o2[c] → mua[c]

(CatDO= catechol-1,2-dioxygenase - EC 1.13.11.1)

MUAe: mua[c] →mua[e]

EX_MUA: mua[e] →

(Exchange Reaction)

E.1.6 Glucaric acid production from glucose

Glucose → glucaric acid

INO1: g6p[c] → mi1p__D[c]

(INO1 = myo-Inositol-1-phosphate synthase – EC 5.5.1.4)

MIOX: inost[c] + o2[c] → glcur[c] + h2o[c]

(MIOX= myo-inositol oxygenase - EC 1.13.99.1)

UDH: glcur[c] + nad[c] + h2o[c] → glcr[c] + nadh[c] + h[c]

(UUDH= Uronate dehydrogenase - EC 1.1.1.203)

E.1.7 Adipic acid production from TCA cycle intermediates – route 1

AcCoA + SuccCoA → adipyl-CoA → adipic acid

3OxT: accoa[c] + succoa[c] → oxacoa[c] + coa[c]

(3OxT = 3-oxoadipyl-CoA thiolase - EC 2.3.1.74)

3OxD: oxacoa[c] + nadh[c] + h[c] → hoacoa[c] + nad[c]

(3OxD = 3-oxoadipyl-CoA dehydrogenase - EC 1.1.1.35)

3HoaH: hoacoa[c] → dhacoa[c] + h2o[c]

(3HoaH = 3-hydroxyadipyl-CoA hydratase - EC 4.2.1.17)

23DhaH: dhacoa[c] + nadph[c] → adicoa[c] + nadp[c]

(23DhaH = adipyl-CoA dehydrogenase - EC 1.3.1.38) *Reverse

AdCoAK: adicoa[c] + amp[c] + 2pi[c] → ada[c] + coa[c] + atp[c]

(AdCoAK = adipate-CoA ligase - EC 6.2.1.-) *Reverse

ADAe: ada[c] → ada[e]

ExADA: ada[e] →

(Exchange Reaction)

E.1.8 Adipic acid production from TCA cycle intermediates – route 2

AcCoA + SuccCoA → adipyl-CoA → adipyl-phosphate → adipic acid

3OxT: accoa[c] + succoa[c] → oxacoa[c] + coa[c]

(3OxT = 3-oxoadipyl-CoA thiolase - EC 2.3.1.74)

3OxD: oxacoa[c] + nadh[c] + h[c] → hoacoa[c] + nad[c]

(3OxD = 3-oxoadipyl-CoA dehydrogenase - EC 1.1.1.35)

3HoaH: hoacoa[c] → dhacoa[c] + h2o[c]

(3HoaH = 3-hydroxyadipyl-CoA hydratase - EC 4.2.1.17)

23DhaH: dhacoa[c] + fadh2[c] → adicoa[c] + fad[c]

(23DhaH = adipyl-CoA dehydrogenase - EC 1.3.99.-) *Reverse

PBT: adicoa[c] + pi[c] → adipi[c] + coa[c]

(PBT= phosphate butrylyltransferase - EC 2.3.1.19)

BTK: adipi[c] + adp[c] → ada[c] + atp[c]

(BTK= butyryl kinase - EC 6.2.1.2)

ADAe: ada[c] → ada[e]

ExADA: ada[e] →

(Exchange Reaction)

E.1.9 Muconic acid production from lignin-derived p-coumaric acid

p-Coumaric acid → muconic acid

Added reactions:

PcaD: 34dhbz[c] → catechol[c] + co2[c]

(PcaD = protocatechuate decarboxylase – EC 4.1.1.63)

MUAe: ccmuac[c] → ccmuac[e]

EX_MUA: ccmuac[e] →

(Exchange Reaction)

Eliminated reaction:

PcaO: o2[c] + 34dhbz[c] -> 2 h[c] + CCbuttc[c]

(PcaO = protocatechuate 3,4-dioxygenase – EC 1.13.11.3)

E.1.10 Muconic acid production from lignin-derived ferulic acid

Ferulic acid → muconic acid

Added reactions:

PcaD: 34dhbz[c] → catechol[c] + co2[c]

(PcaD = protocatechuate decarboxylase – EC 4.1.1.63)

MUAe: ccmuac[c] → ccmuac[e]

EX_MUA: ccmuac[e] →

(Exchange Reaction)

Eliminated reaction:

PcaO: o2[c] + 34dhbz[c] -> 2 h[c] + CCbuttc[c]

(PcaO = protocatechuate 3,4-dioxygenase – EC 1.13.11.3)

E.1.11 Muconic acid production from lignin-derived benzoate

Benzoate → muconic acid

MUAe: ccmuac[c] → ccmuac[e]

EX_MUA: ccmuac[e] →

(Exchange Reaction)

E.2 Flux Balance Analysis of Short Chain Diols

Theoretical yields of the analyzed compounds were calculated using the Constraint-Based Reconstruction and Analysis (COBRA) Toolbox v3.0 in Matlab [5]. Engineered pathways for 1,2-propanediol [6], 1,3-butanediol [7] and 2,3-butanediol [8] were added to the genome scale model for *E. coli* MG1655, iML1515 [9]. Endogenous metabolites were used when available and pathways started from lactic acid, 3-hydroxybutyryl-CoA and α-acetolactone for 1,2-PDO, 1,3-BDO and 2,3-BDO respectively. For engineered 1,2-PDO biosynthesis, reactions present in the base model which produce methylglyoxal were removed (simulated gene deletion) to ensure that 1,2-PDO flux was calculated through the engineered pathway and not confounded by the native pathway. Glucose uptake rate (EX_glc_D_e) was fixed at -10, and O2 uptake rate (EX_o2_e) was fixed at -15 for all pathway simulations. Theoretical yields were obtained using the optimizeCbModel

function with the objective function set for the exchange reaction of the final product. The reactions and metabolites added for each pathway are presented below. Full pathway description is found in Chapter 4 (Figure 4.

E.2.1 1,2-Propanediol production from glucose

Metabolites Added:

Lacoyl-CoA: lac_coa_c

Reactions Added:

Lactoyl-CoA Transferase: lac__D_c + accoa_c → lac_coa_c + ac_c

Lactaldehyde Dehydrogenase: lac_coa_c + nadh_c → lald__D_c + nad_c + coa_c

Reactions Removed:

MGSA: dhap_c -> pi_c + mthgxl_c

AACTOOR: o2_c + aact_c + h2o_c -> nh4_c + mthgxl_c + h2o2_c

E.2.2 1,3-Butanediol production from glucose

Metabolites Added:

3-hydroxybutyraldehyde: 3hbald_c

1,3-butanediol, cytosolic: 13bdo_c

1,3-butanediol, extracellular: 13bdo_e

Reactions Added:

Butyraldehyde Dehydrogenase: 3hbcoa_c + nadh_c → 3hbald_c + nad_c + coa_c

Butyraldehyde Dehydrogenase: 3hbald_c + nadh_c → 13bdo_c + nad_c

1,3-BDO Transport: 13bdo_c <=> 13bdo_e

1,3-BDO Exchange: 13bdo_e <=>

E.2.3 2,3-Butanediol production from glucose

Metabolites Added:

R-acetoin: acetoin_c

2,3-butanediol, cytosolic: 23bdo_c

2,3-butanediol, extracellular: 23bdo_e

Reactions Added:

α-Acetolactate Decarboxylase: alac__S_c → acetoin_c','co2_c

2,3-Butanediol Dehydrogenase: acetoin_c + nadh_c → 23bdo_c + nad_c

2,3-BDO Transport: 23bdo_c <=> 23bdo_e

2,3-BDO Exchange: 23bdo_e <=>

E.3 References

1. Schellenberger, J. et al. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.* **6,** 1290-1307 (2011).

2. Feist, A.M. et al. Model-driven evaluation of the production potential for growth-coupled products of Escherichia coli. *Metab. Eng.* **12,** 173-186 (2010).

3. Nogales, J., Palsson, B.Ø., & Thiele, I. A genome-scale metabolic reconstruction of Pseudomonas putida KT2440: iJN746 as a cell factory. *BMC Syst. Biol.* **2,** 79-99 (2008).

4. Mo, M.L., Palsson, B.Ø. & Herrgård, M.J. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *BMC Syst. Biol.* **3,** 37-54 (2009).

5. Heirendt, L. et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat. Protoc.* **14,** 639-702 (2019).

6. Niu, W., Kramer, L., Mueller, J., Liu, K. & Guo, J. Metabolic engineering of Escherichia coli for the de novo stereospecific biosynthesis of 1,2-propanediol through lactic acid. *Metab. Eng. Commun.* **8,** e00082 (2019).

7. Kataoka, N., Vangnai, A.S., Tajima, T., Nakashimada, Y. & Kato, J. Improvement of (R)-1,3-butanediol production by engineered Escherichia coli. *J. Biosci. Bioeng.* **115,** 475-480 (2013).

8. Xu, Y. et al. Systematic metabolic engineering of Escherichia coli for high-yield production of fuel bio-chemical 2,3-butanediol. *Metab. Eng.* **23,** 22-33 (2014).

9. Monk, J.M. et al. iML1515, a knowledgebase that computes Escherichia coli traits. *Nat. Biotechnol.* **35,** 904-908 (2017).

APPENDIX F.

UNIT OPERATION CALUCLATIONS FOR BIOTECHNOLOGY-ENABLED *IN SITU* RESOURCE UTILIZATION

F.1 Unit Operation Specifications

Material Density

- Stainless steel = 7.85 g/cm^3
- Aluminum = 2.7 g/cm^3
- Low density polyethylene (LDPE) = 1 g/cm^3
- High density polyethylene (HDPE) = 1 g/cm^3
- Cotton = 1.55 g/cm^3
- Polyvinylchloride (PVC) = 1.38 g/cm^3

Suspended Growth PBR

- 4.5 cm thick, 1 m tall, 45 m wide
- PBR material: 0.3 mm thick LDPE

Biofilm Growth Substrate

- 1 m tall
- Growth substrate material: 0.3 mm thick cotton

Biomass Concentrator

- Reactor material = 6 cm thick steel

Enzyme Digester

- 20% headspace on volume determined by flowrate and desired residence time
- Rector height:reactor radius ratio = 2.5
- Reactor material = 6 mm thick stainless steel
- Reactor residence time = 48 hours

E. coli Fermentation

- 20% headspace on volume determined by flowrate and desired residence time
- Rector height:reactor radius ratio = 2.5
- Reactor material = 6 mm thick stainless steel
- Reactor residence time = 72 minutes

Liquid-Liquid Equilibrium Extraction

- 20% headspace on volume determined by flowrate and desired residence time
- Rector height:reactor radius ratio = 2.5
- Reactor material = 6 mm thick stainless steel
- Reactor residence time = 60 minutes

Membrane Separation Unit

- Reactor material = 6 cm thick stainless steel
- Membrane = polydimethylsiloxane/polyvinyledenefluoride

Storage Tanks

- 2,3-BDO storage: 16 m^3 storage tank made of 6 cm thick HDPE
- O_2 storage: 28 m^3 storage tank made of 6 cm thick HDPE

F.2 Propellant Property Calculation

F.2.1 Theoretical specific impulse

Theoretical values for specific impulse (I_{sp}) were calculated assuming that all stored chemical energy in the fuel is converted to kinetic energy, allowing a simplified calculation of exhaust velocity. A simplified calculation fits our analysis, as the I_{sp} values were used for comparison only and all compared fuels are hydrocarbons, meaning they have similar combustion temperatures and exhaust compositions if we assume complete combustion. Thus the calculation of I_{sp} (Equation F.1) was approximated using Equation F.2, where v_e is the exhaust velocity, g is the gravitational force, ΔH is the enthalpy of combustion for the fuel of interest, n_{H2O} is the number of moles of H2O produced from complete combustion of 1 mole of fuel, ΔH_{vap} is the heat of vaporization of water, and m_e is the molar mass of the exhaust determined based on complete combustion of the fuel of interest.

$$I_{sp} = \frac{v_e}{g} \qquad \text{Equation F.1}$$

$$I_{sp} = \frac{\sqrt{\frac{2(\Delta H - n_{H2O}\Delta H_{vap})}{m_e}}}{g} \qquad \text{Equation F.2}$$

F.2.2 Lower heating value

The lower heating value of the fuel candidates was calculated as the enthalpy of formation of the compound of interest minus the combined enthalpy of formation of the CO_2 and H_2O formed during fuel combustion on a per mass basis.

F.2.3 Required oxygen to fuel ratio

The required O_2:fuel ratio is the mass of O_2 required for complete combustion of the fuel, on a per mass basis.

F.3 Process Water Use Calculation

The water requirement of the system was calculated based on the water usage of the cyanobacterial growth unit, the enzyme digester, and the fermenter. For suspended growth, water use was based on the sizing and spacing of the photobioreactors which led to a requirement of 41.47 L/m^2 land area. For biofilm growth, water use was based off of 2.18 L/m^2 of growth substrate area required in previous literature [1]. The water requirement for the digester and fermenter were based off the calculated flowrate out of the cyanobacterial concenter (or from the growth unit in the case of biofilm growth) and the residence time of the digester (48 hours) and fermenter (72 minutes). No water loss was accounted for and 100% of required water was recycled.

F.4 Process Power Use Calculation

F.4.1 Suspended cyanobacteria growth

Power requirement came from culture mixing and pumping. Mixing was assumed to be directly proportional to gravity with a power input of 52 W/m^3 of culture media while pumping power was taken from literature at 0.058 kWh/m^3 with the volume corresponding to the flowrate coming out of the cyanobacterial growth operation [2].

F.4.2 Biofilm cyanobacteria growth

Here, power requirement came only from pumping required for nutrient resupply, and was based on the pilot scale value of 0.038 W/m^2 culture area [1].

F.4.3 Cyanobacteria concentrator

The power requirement of the cross flow filtration unit was assumed to be 5 kWh/m^3 of volume processed based on previous literature [3].

F.4.4 Digester, fermenter, and liquid-liquid extraction units

Here, power requirement was calculated based on the power required for reactor agitation via impeller mixing. Since Mars has decreased gravity compared to Earth, the power requirement for mixing would be lower. As a conservative estimate, we estimated a 25% reduction in power requirement, thus making the mixing requirement 1.5 kW/m^3 of reactor volume [4].

F.4.5 Membrane separation unit

Here, calculation of power requirement was based off the energy model developed in previous literature for membrane separation of 2,3-butanediol from butanol [5], the same system used in our process model.

F.5 Process Mass Calculation

F.5.1 Cyanobacteria growth

Payload mass for the cyanobacterial growth comprised of mass of the growth substrate, mass of the support frame, and mass of the nutrients required for cyanobacterial growth. Growth substrate mass for suspended growth was based on the dimensions of the PBR and a 0.3 cm thick LDPE material. For biofilm growth, the mass was based on the total growth substrate area and a 0.3 cm thick cotton material

Required nutrient mass was based on the requirement of 0.029 and 0.002 moles of nitrogen and phosphorous required per mole of biomass. Phosphorous was entirely provided by diammonium phosphate, which also provides two moles of nitrogen. The rest of the nitrogen was provided by ammonia. Enough nitrogen and phosphorous was shipped to supply the feed the process during the entire 500 day mission with 20% excess, assuming no nitrogen or phosphorous recycling. In addition, trace elements were provided for a media composition of 1 g/L NaCl, 0.03 g/L CaCl2, 1 g/L K2SO4, 0.008 g/L MgSO4, 0.08 g/L EDTA, 0.01 g/L FeSO4, 0.00023 g/L MnCl2, 0.00011 g/L ZnSO4, and 0.00003 g/L CuSO4 [6].

Required mass of the polyvinylchloride (PVC) support frame was estimated based on the 78 tons/hectare of steel support structure used in pilot scale suspended growth [2]. Frame mass was corrected for Martian gravity which will reduce frame support needed by reducing the weight of the algal cultivation reactors. Mass for biofilm growth was further reduced due to the reduced water content that makes the total mass of the biofilm reactors 11% of the suspended system. Total mass needed was based on volume of support frame and a PVC density of 1.38 tons/m^3.

F.5.2 Cyanobacteria concentrator

The cyanobacteria concentrator mass was calculated based on requiring 155.61 m^2 of organic membrane housed in a steel cylinder 2 m long with a 1 m radius.

F.5.3 Enzyme digester

The enzyme reactor was designed to be a steel vessel with a 2.5 height to radius ratio, with 20% excess volume to the volume required to reach the desired residence time. In addition, the mass of enzymes needed for biomass digestion was included in this calculation. Here we assumed enzyme concentration of 1 g/L lysozyme, 1 g/L amylase, and 1 g/L glucoamylase [7]. Enough enzyme was shipped to allow the enzymes to be replaced every 10 days.

F.5.4 Fermenter

The enzyme reactor was designed to be a steel vessel with a 2.5 height to radius ratio, with 20% excess volume to the volume required to reach the desired residence time.

F.5.5 Liquid-liquid extractor

The enzyme reactor was designed to be a steel vessel with a 2.5 height to radius ratio, with 20% excess volume to the volume required to reach the desired residence time. In addition, the mass of butanol needed for extraction was included in this calculation. The mass of butanol was modeled based on a pure butanol input flowrate of 0.001 L/min, this input rate gave us our desired 95% 2,3-butanediol purity.

F.5.6 Membrane separator

The membrane separator mass was calculated based on requiring 0.5 m^2 of polydimethylsiloxane/polyvinyledenefluoride membrane housed in a steel cylinder 0.5 m long with a 0.007 m radius.

F.5.7 Storage tanks

The 2,3-butanediol and O_2 storage tanks are 16 m^3 and 28 m^3, respectively, composed of 0.006 m thick HDPE.

F.5.8 Water treatment

The water treatment mass was calculated based on requiring 13.1 m^2 of thin-film composite membrane housed in a steel cylinder 0.5 m long with a 0.036 m radius.

F.6 Photosynthetic Oxygen Production

Excess O_2 production was calculated based on 1 mole of O_2 produced for every mole of CO_2 fixed into biomass. O_2 was consumed in the fermenter at a rate of 0.12 mol O_2/L/hr at a biomass concentration of 6 g/L [8].

F.7 References

1. Ozkan, A., Kinney, K., Katz, L. & Berberoglu, H. Reduction of water and energy requirement of algae cultivation using an algae biofilm photobioreactor. *Bioresour. Technol.* **114,** 542-548 (2012).

2. Tredici, M.R. et al. Energy balance of algal biomass production in a 1-ha "Green Wall Panel" plant: how to produce algal biomass in a closed reactor achieving a high net energy ratio. *Appl. Energy* **154,** 1103-1111 (2015).

3. Al-Hattab, M., Ghaly, A. & Hammouda, A. Microalgae harvesting methods for industrial production of biodiesel: critical review and comparative analysis. *J. Fundam. Renew. Energy Appl.* **5,** 1000154 (2015).

4. Koutinas, A.A. et al. Techno-economic evaluation of a complete bioprocess for 2,3-butanediol production from renewable resources. *Bioresour. Technol.* **204,** 55-64 (2016).

5. Shao, P. & Kumar, A. Process energy efficiency in pervaporative and vacuum membrane distillation separation of 2,3-butanediol. *Can. J. Chem. Eng.* **89,** 1255-1265 (2011).

6. Cogne, G., Lehmann, B., Dussap, C.G. & Gros, J.B. Uptake of macrominerals and trace elements by the cyanobacterium Spirulina platensis (Arthrospira platensis PCC 8005) under photoautotrophic conditions: culture medium optimization. *Biotechnol. Bioeng.* **81,** 588-593 (2003).

7. Aikawa, S. et al. Direct conversion of Spirulina to ethanol without pretreatment or enzymatic hydrolysis processes. *Energy Environ. Sci.* **6,** 1844-1849 (2013).

8. Andersen, K.B. & von Meyenburg, K. Are growth rates in Escherichia coli in batch cultures limited by respiration? *J. Bacteriol.* **144,** 114-123 (1980).

APPENDIX G.

ADVANCING THE POTENTIAL FOR THE PRODUCTION OF CHEMICALS FROM CARBON DIOXIDE IN *ESCHERICHIA COLI*

Reproduced with permission from:

Kruyer, N.S. & Peralta-Yahya, P. Advancing the potential for the production of carbon dioxide in *Escherichia coli*. *Biochemistry* **59,** 731-732 (2020).

G.1 Perspective

Biological production of chemicals from alternative carbon feedstocks, such as carbon dioxide (CO_2), is pivotal to the development of sustainable chemical processes. Today, a majority of biologically produced chemicals are obtained via engineered heterotrophs, such as the workhorse organism *Escherichia coli* that converts biomass feedstock to fuels and chemicals [1]. An implicit assumption in this process is that plants will harvest energy from sunlight and fix CO_2 from the environment to generate the biomass that *E. coli* will use for chemical production. Therefore, biological production of chemicals from renewable biomass is dependent on the price of plant biomass (e.g., corn bushel price), the organism's growth rate, and its productivity. Engineering photoautotrophs, such as cyanobacteria and microalgae, which fix CO_2 using sunlight, permits circumvention of the reliance on plant feedstock and has been used to produce biofuels. However, *E. coli* currently has several advantages over photoautotrophs,

including a growth rate that is 5 times faster than that of cyanobacteria, a more sophisticated synthetic biology toolbox for engineering metabolic pathways, and independence from light, which not only reduces the cost of the bioreactor but also increases the productivity of the system.

One of the Grand Challenges in metabolic engineering has been to port the CO_2 fixing capabilities of autotrophs to a workhorse heterotroph, such as *E. coli*. An autotrophic *E. coli* could leverage the knowledge of more than 100 metabolic pathways already optimized for the production of fuels and chemicals and produce them from CO_2. Milo group's 2016 attempt to port CO_2 fixation into *E. coli* achieved a 35% CO_2 fixation into biomass [2]. Three years later, in their 2019 contribution, Gleizer *et al*. engineered an *E. coli* capable of using CO_2 as its sole carbon source [3]. The authors harnessed metabolic engineering to build an *E. coli* with disrupted glycolysis and pentose phosphate pathways, capable of fixing CO_2 via the Calvin cycle, and using formate as the electron source. This strain was then optimized for CO_2 fixation via adaptive evolution, over 350 days, resulting in mutations that regulate gene expression and pathway regulation. The introduction of formate into the growth media, and formate dehydrogenase as a method for producing reducing power, proved to be the key additions for improving from 35% to 100% of biomass from CO_2.

The adaptive evolution step was fundamental to achieve an *E. coli* that uses CO_2 as its sole carbon source. To rationally engineer such an *E. coli* would have required the design, building, and testing of an extremely large number of constructs, as this problem requires replacing a central metabolic pathway (the glycolysis/pentose phosphate pathway) with a heterologous one (the Calvin cycle). This is orders of magnitude more difficult than

engineering a chemical production pathway to siphon central metabolism intermediates (e.g., pyruvate) for chemical production. In the latter, one need only up- or downregulate the central metabolism, not replace it. More and more, we are seeing adaptive (or directed) evolution being applied to metabolic engineering challenges, and Gleizer *et al*. show that, as long as we can link the desired activity to biomass accumulation, we can replace the central metabolism in biological systems.

An exciting application of this work is the availability of a new, more efficient, autotrophic host for chemical production. The increased efficiency comes from the use of formate, as opposed to light, to generate reducing power. As a light-independent organism, the engineered *E. coli* is not limited by the low efficiency of light utilization [4] and is instead capable of utilizing all of the exogenously fed formate toward biomass production through carbon fixation [3] Albeit, the current strain exhales more CO_2 than it consumes (Figure G.1).

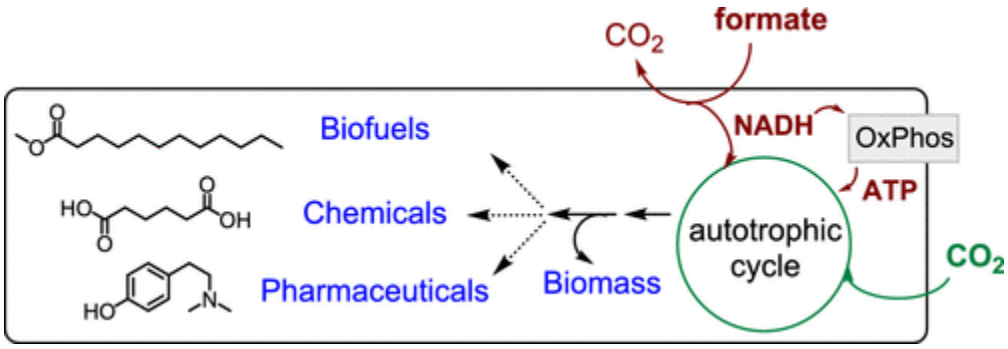

Figure G.1. Advancing the potential for the production of chemicals from carbon dioxide in *E. coli*.

Several steps remain before the autotrophic *E. coli* can be implemented for chemical production at scale. The autotrophic *E. coli* doubling time will need to be shortened (18 h, as compared to 30 min for the wild type). Also, the Type II RuBisCo used by Gleizer *et al*. has a high k_{cat} yet low selectivity for CO_2 over O_2, which at the 10% CO_2 concentration used in the chemostat pushed the selectivity toward the carbon fixation reaction. In bioreactors, a high CO_2 concentration can be achieved via fed gas. However, for applications under atmospheric conditions or in cases in which CO_2 inhibits the reaction, a Type I RuBisCO, with a higher selectivity for CO_2 and intermediate k_{cat} values, will need to be used. Another way to improve carbon fixation in a Type II RuBisCO system is through a carbon-concentrating mechanism. In bacteria, this often comes in the form of a carboxysome or carboxysome-like bacterial microcompartment (BMCs). While they are not natively found in *E. coli*, recent advancements in heterologous expression and customized cargo loading of BMCs suggest that this is a viable solution for this system [5]. Increasing the local CO_2 concentration can compensate for lower selectivity while taking advantage of high turnover numbers in Type II RuBisCO.

Increasing CO_2 levels in Earth's atmosphere present the challenge of how to capture this greenhouse gas, with a goal of addressing the effects of global climate change. The ability to convert the captured carbon into a value-added product is appealing for carbon capture and sequestration as well as for sustainable chemical production and reduction of our dependence on nonrenewable fossil fuels. The successful engineering of a fully autotrophic E. coli strain opens the door for the use of a fast-growing model organism in carbon capture and production of chemicals and biofuels, offering distinct advantages over

currently available microbes such as algae and cyanobacteria. Gleizer *et al*.'s accomplishment presents a first, but crucial step, on the path toward sustainable production of value-added chemicals from CO_2.

G.2 References

1. Sarria, S., Kruyer, N.S. & Peralta-Yahya, P. Microbial synthesis of medium-chain chemicals from renewables. *Nat. Biotechnol.* **35,** 1158-1166 (2017).

2. Antonovsky, N. et al. Sugar synthesis from CO_2 in *Escherichia coli. Cell* **166,** 115-125 (2016)

3. Gleizer, S. et al. Conversion of *Escherichia coli* to generate all biomass carbon from CO_2. *Cell* **179,** 1255 (2019).

4. Hannon, M., Gimpel, J., Tran, M., Rasala, B. & Mayfield, S. Biofuels from algae: challenges and potential. *Biofuels* **1,** 763-784 (2010).

5. Bonacci, W. et al. Modularity of carbon-fixing protein organelle. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 478-483 (2012).