Expanding the frontiers of plants utilization:

The metabolic power of *Novosphingobium aromaticivorans* for lignin valorization

By

Jose Miguel Pérez

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Civil and Environmental Engineering)

at the

UNIVERSITY OF WISCONSIN-MADISON

2020

Date of final examination: 10/27/2020

The dissertation is approved by the following members of the Final Oral Committee:

Daniel R. Noguera, Professor, Civil and Environmental Engineering

Timothy J. Donohue, Professor, Bacteriology

Katherine D. McMahon, Professor, Civil and Environmental Engineering

Gregory W. Harrington, Professor, Civil and Environmental Engineering

John Ralph, Professor, Biochemistry

Abstract

Expanding the frontiers of plants utilization:
The metabolic power of *Novosphingobium aromaticivorans* for lignin valorization

by

Jose Miguel Pérez
Doctor of Philosophy – Civil and Environmental Engineering
University of Wisconsin-Madison

Professor Daniel R. Noguera

Plant biomass is the most abundant terrestrial source of renewable carbon that has the potential to replace fossil fuels to produce fuels and chemicals. For decades, the strategies for the utilization of its sugar-based fraction have involved the combination of physicochemical and biochemical techniques with the utilization of microorganisms, obtaining promising results. One of the major components of plant biomass is lignin, a phenolic heteropolymer that presents additional challenges because of its high recalcitrance and heterogeneity. As lignin valorization is predicted to be critical for the economic viability of a plant-based biorefinery, the development of appropriate tools is required. To make higher value products from lignin, recent studies have proposed the integration of chemical lignin deconstruction with biological upgrade by genetically engineered bacteria.

In this work, we explored the potential of *Novosphingobium aromaticivorans* DSM12444 as a platform organism for lignin valorization in the context of a biomass-to-product pipeline that integrates chemical deconstruction with biological transformation. We used genetic tools to engineer the bacterium's metabolic mechanisms for production of 2-pyrone-4,6-dicarboxylic acid (PDC) and demonstrated its ability to utilize a broad range of phenolic compounds produced by chemical lignin depolymerization. The engineered strain was able to produce PDC with high yields using lignin depolymerized with different techniques and from multiple plant sources. Finally, we explored *N. aromaticivorans* enzymatic mechanisms for phenolic compound degradation and we were able to identify, using genomic, genetic and enzymology tools, enzymes responsible for critical reactions such as aromatic *O*-demethylation and

aromatic ring opening. As a result, we created a 2nd generation engineered strain with improved product conversion yields with respect to the first engineered strain.

Overall, this work provides valuable knowledge about the metabolic potential of bacteria for lignin valorization in the context of a biomass-to-product biorefinery and expands the portfolio of potentially useful platform microorganisms. Ongoing and future research on increasing the knowledge of *N. aromaticivorans* mechanisms for phenolic compound degradation, compatibility with different chemical biomass deconstruction techniques, expanding the portfolio of valuable products, and developing industrially relevant culture techniques will provide valuable information to fully understand *N. aromaticivorans* potential for lignin valorization.

Acknowledgements

The day I sat at Dan's office to talk about the possibility of joining his research group at the Great Lakes Bioenergy Research Center (GLBRC), he highlighted that they tried to answer research questions that required the combined effort from people of different fields. After more than five years of work, I can confirm that the results presented in this manuscript would not exist without a continuous collaboration with many great researchers. It is hard to remember all names to list them here, but anyone who ever contributed to this work should know that I partially owe you these results and I am greatly thankful for your generosity.

First of all, I want to thank my advisor Dr. Dan Noguera for believing in me when I asked him to be part of his research team and for giving me all the freedom to explore every idea that came to my mind. I also want to thank Dr. Tim Donohue for supporting my work and for welcoming in his lab at the GLBRC. Also a member of the GLBRC, I want to thank Dr. John Ralph for his contribution to my understanding of the chemical nature of lignin. Finally, to Dr. Trina McMahon and Dr. Greg Harrington. I took classes with both and the knowledge gained in microbiology and water treatment were essential to my research.

I joined the Donohue-Noguera lab at the GLBRC as an agriculture engineer with a little knowledge in microbiology and chemistry. As the challenge of trying to use microbes to make value from something so little known in the world as lignin was beyond my capacities at that time, I required a significant amount of help from experienced researchers. I found two of such researchers in the GLBRC community, Dr. Wayne Kontur and Dr. Steve Karlen who generously adopted the role of mentors and helped me when I was in great need. I don't have enough words to thank both of them for the incredible amount of time they invested in me.

I would also thank members of the Donohue-Noguera lab who were essential to perform the large amount of laboratory work that this thesis required. I want to thank German Umana, Yanjun Ma, Carson Gehl, Derek Gille, and Alyssa Niles for the great work at running bacterial growth experiments, creating

deletion mutants, and purifying enzymes. Other actual and former students/researchers from the Donhue-Noguera lab I want to thank are Zach Oshlag, Rachelle Lemke, Matt Scarborough, Alexandra Linz, Kevin Walters, Nathan Fortney, Kevin Myers, and Bryan Lakey. From the Noguera Lab, I want to thank Francisco Moya and Jackie Bastyr Cooper.

I also want to thank researchers from other research groups at UW-Madison. From the John Ralph lab, Canan Sener and Yanding Li. With Yanding we conceived the third chapter of this thesis and Canan provided depolymerized lignin for the whole study. From Shannon Stahl lab, I want to thank Manar Alherech who provided the first batch of depolymerized lignin with which we demonstrated the success of the first PDC-producing mutant strain. From the Josh Coon lab, I want to thank Alan Highbee for his great help in teaching me mass spectrometry and providing HPLC-MS analysis.

There were also other people at the GLBRC who contributed to this work and I want to thank them. Mick Mcgee performed HPLC-RID analysis; Jason Coplien and Dennis Haak provided GVL lignin for the study presented in chapter three. In addition, I want to thank the GLBRC staff who in different ways helped me during these years: Eva Ziegelhoffer, Tina Nielsen, Laura Ketterhagen, and Emily Hickey. I also want to thank the collaboration of Eric Hegg from GLBRC and Michigan State University who provided the Cu-AHP lignin used in chapter two of this work.

Finally, I want to thank people who are not researchers but were close to me during all these years. First, my wife Cait, who supported me through this great challenge. She patiently suffered long nights of my absence while I was performing endless experiments in the lab and always waited for me with a smile. Also, my family of origin founded by my parents Enrique and Maria Teresa, and Cait's family who welcomed me in this country, provided an incredible support to both my PhD studies and my marriage.

Finally, I want to thank Mary, the Mother of God. These years I've been in a deep spiritual journey and this thesis is only one of its products, probably a little one. She has always been at my side and I know She will always be.

Table of contents

Abstract ……..... i

Acknowledgements …….. iii

Table of contents ……v

List of figures …… x

List of tables ……. xvii

1. Introduction ……... 1

1.1. Motivation …….. 1

1.2. The challenge of lignocellulose valorization …… 3

1.3. Lignin …… 4

1.4. Proposed strategy to valorize lignin …….. 5

1.5. Novosphingobium aromaticivorans DSM12444 …… 7

1.6. Identification of research needs …… 8

1.7. Research objectives …… 9

1.8. References …….. 10

2. Funneling aromatic products of chemically depolymerized lignin into 2-pyrone-4-6-dicarboxylic acid with *Novosphingobium aromaticivorans* …….. 16

2.1 Abstract …….. 16

2.2 Introduction ……..... 17

2.3 Results …… 19

2.3.1 Model of aromatic metabolism by N. aromaticivorans DSM12444 and justification of experimental approach …….. 19

2.3.2 Construction of a *N. aromaticivorans* mutant that accumulates PDC from G and H aromatics ……..…..... 22

2.3.3 Construction of a *N. aromaticivorans* mutant that accumulates PDC from S aromatics ……..... 28

2.3.4 Construction of a *N. aromaticivorans* mutant that accumulates PDC from S, G, and H aromatics ……..... 30

2.3.5 The fate of unconverted aromatic carbon ……... 30

2.3.6 Production of PDC from chemically depolymerized lignin …… 31

2.3.7 Production of PDC from vanillic acid and vanillin in a fed-batch reactor ……. 36

2.4 Discussion …….. 36

2.5 Materials and methods …….. 40

2.6 Conclusions …… 45

2.7 Acknowledges …… 46

2.8 Supplementary figures and tables ……. 47

2.9 Supplementary methods ……. 56

2.9.1 Construction of plasmids for deleting genes Saro_2819 or Saro_2864/5 ……... 56

2.9.2 Deletion of genes Saro_2819 and Saro_2864/5 ……... 56

2.9.3 Purification of PDC ……. 57

2.9.4 Steps in the synthesis of S diketone …… 57

2.9.5 Steps in the synthesis of G diketone …… 59

2.10 References …….. *61*

3. Integrating lignin depolymerization and microbial funneling processes using agronomically relevant feedstocks ……... 65

3.1 Abstract …….. *65*

3.2 Introduction ……..... *66*

3.3 Results ……..... *69*

3.3.1 Conversion of single compounds into PDC ……..... 69

3.3.2 Lignin isolation and depolymerization ……..... 73

3.3.3 Production of PDC from lignin hydrogenolysis products …….... 76

3.4 Discussion …….. *79*

3.5 Conclusions …… *83*

3.6 Materials and methods ……... *84*

3.7 Acknowledgements ……. *87*

3.8 Supplementary figures and tables …….. *88*

3.9 References …….. *92*

4. Functional pathway redundancy in the metabolism of plant-derived phenolics by *Novosphingobium aromaticivorans* …….. 96

4.1 Abstract ……. *96*

4.2 Introduction ……... *97*

4.3 Results …….. *101*

4.3.1 Identification of putative *O*-demethylases in *N. aromaticivorans* …… 101

4.3.2 Identification of putative aromatic ring opening dioxygenases in *N. aromaticivorans* ……... 111

4.4 Discussion …… 118

4.4.1 *O*-demethylation reactions …… 119

4.4.2 Aromatic ring opening reactions …….. 123

4.4.3 Implications for PDC production from lignin-derived aromatics by *N. aromaticivorans* ……125

4.4.4 Concluding remarks ……... 126

4.5 Materials and methods …… 127

4.6 Acknowledgements ……132

4.7 Supplementary materials ……... 133

4.7.1 Supplementary methods ……133

4.7.2 Supplemental results ……... 134

4.7.3 Supplemental figures and tables …… 136

4.8 References …… 147

5. Major findings and future directions ……. 151

5.1 Summary …..…… 151

5.2 Future directions ……... 152

5.2.1 Mechanisms for monomeric and oligomeric aromatic compound degradation in *N. aromaticivorans* ……... 152

5.2.2 Compatibility with chemical depolymerization techniques ……. 152

5.2.3 Product diversity …… 153

5.2.4 PDC production in an industrially relevant scale …….. 153

5.3 References …… 154

List of figures

Figure 1.1. Global energy consumption by primary source from 1800 to 2018. The data used to generate this figure was obtained from Smil, 2017 …….. 2

Figure 1.2. Concentration of CO_2 in the earth's atmosphere from March 1958 to August 2020. Source: NOAA Earth System Research Laboratory …… 3

Figure 1.3. General microbial mechanism for lignin catabolism …… 6

Figure 1.4. Proposed lignin valorization strategy ……7

Figure 2.1. Predicted pathways of S unit (syringic acid), G unit (vanillic acid), and H unit (p-hydroxybenzoic acid) metabolism in *N. aromaticivorans* DSM12444, based on work in *Sphingobium sp*. SYK 6 and *N. aromaticivorans* DSM12444. In this model, deletions of the genes *ligI* (Saro_2819), *desC* (Saro_2864), and *desD* (Saro_2865) are hypothesized to enable the funneling (represented by light blue arrows) of S, G, and H lignocellulosic biomass derived aromatic compounds into 2-pyrone-4,6-dicarboxylic acid (PDC). Abbreviations: 3-methylgallate, 3-MGA; 4-carboxy-2-hydroxy-6-methoxy-6-oxohexa-2,4-dienoate, CHMOD; 4-carboxy-2-hydroxy-cis,cis-muconate-6-semialdehyde, CHMS; 4-oxalomesaconate, OMA ……. 20

Figure 2.2. Cell density of representative *N. aromaticivorans* cultures grown on 3 mM vanillic acid (panel A) or 3 mM p-hydroxybenzoic acid (Panel B). Parent strain 12444*Δ1879* represented by squares and dashed line; strain 12444ΔligI represented by circles ...…… 23

Figure 2.3. Cell density and extracellular metabolite concentration of representative *N. aromaticivorans* strain 12444ΔligI cultures grown on a combination of 3 mM vanillic acid and 3 mM glucose (panels A and C) or a combination of 3 mM p-hydroxybenzoic acid and 3 mM glucose (panels B and D) ……24

Figure 2.4. Cell density and extracellular metabolite concentrations of representative cultures of *N. aromaticivorans* strains 12444*Δ1879* (panels A and C) and 12444*ΔligI* (panels B and D) grown in media containing syringic acid ……. 27

Figure 2.5. Cell densities and extracellular metabolite concentrations of *N. aromaticivorans* strains 12444*ΔdesCD* (left hand side panels) and 12444*ΔligIΔdesCD* (right hand side panels) grown on 3 mM syringic acid (panels A and B) or a combination of 3 mM syringic acid and 3 mM glucose (panels C to F) …… 29

Figure 2.6. Cell density (panels A and B) and extracellular metabolite concentrations (panels C to F) of representative cultures of *N. aromaticivorans* strains 12444*ΔligIΔdesCD* (left hand side panels) and 12444*Δ1879* (right hand side panels) grown on formic acid induced depolymerized poplar lignin supplemented with glucose. Panels C and D show extracellular concentrations of lignin derived aromatic compounds and PDC as a product, and panels E and F show extracellular concentrations of glucose and formic acid. Formic acid is present in the low molecular weight products of chemical depolymerization, whereas glucose was added to enhance bacterial cell growth ……... 33

Figure 2.7. Cell density and extracellular metabolite concentrations of representative *N. aromaticivorans* strain 12444*ΔligIΔdesCD* cultures grown on minimal media supplemented with S-diketone and glucose (panels A and C) or G-diketone and glucose (panels B and D) …… 35

Figure 2.S1. Cell density and extracellular metabolite concentrations of *N. aromaticivorans* strains 12444*ΔligI* (solid circles) or 12444*ΔligIΔdesCD* (solid triangles) grown on a combination of glucose and vanillin (A), *p*-hydroxybenzaldehyde (B), ferulic acid (C), *p*-coumaric acid (D), and syringaldehyde (E)..47

Figure 2.S2. Cell density and extracellular metabolite concentrations of representative *N. aromaticivorans* strain 12444*ΔdesCD* cultured on 3 mM vanillic acid …… 48

Figure 2.S3. GC-MS peaks of compounds identified in media containing glucose plus the products of formic-acid-induced depolymerization of oxidized poplar lignin; before inoculation (A), after growth of *N.*

aromaticivorans strain 12444*ΔligIΔdesCD* (B), after growth of *N. aromaticivorans* strain 12444*Δ1879* (C). Only strain 12444*ΔligIΔdesCD* accumulates PDC in the growth medium. Panel D shows the absence of additional peaks in an abiotic control experiment ……. 49

Figure 2.S4. GPC chromatogram of media containing glucose plus the products of formic-acid-induced depolymerization of oxidized poplar lignin; before inoculation (A), abiotic control after 78 hours of incubation (B), after growth of *N. aromaticivorans* strain 12444*Δ1879* (C), after growth of *N. aromaticivorans* strain 12444*ΔligIΔdesCD* (D) …… 50

Figure 2.S5. GPC chromatogram of the "oligomers" range at ʎ=254 of media containing glucose plus the products of formic-acid-induced depolymerization of oxidized poplar lignin. Mw: weight average molecular weight; Mn: number average molecular weight; Mw/Mn: dispersity index …… 51

Figure 2.S6. Extracellular metabolite concentrations of a *N. aromaticivorans* strain 12444*ΔligIΔdesCD* culture fed with a concentrated mixture of vanillic acid, vanillin, and glucose. A maximum PDC concentration of 26.7 mM was observed after 48 hours of cultivation …… 51

Figure 2.S7. GC-MS chromatogram (A) and MS spectrum (B) of PDC isolated from a culture of PDC-producing *N. aromaticivorans* strain 12444*ΔligI* grown on a mixture of 3 mM vanillic acid and 3 mM glucose …… 54

Figure 2.S8. ^{1}H NMR spectra in acetone-d_6 of PDC isolated from a culture of PDC-producing *N. aromaticivorans* strain 12444*ΔligI* grown on a mixture of 3mM vanillic acid and 3 mM glucose …….. 55

Figure 3.1. Schematics of a biomass-to-PDC pipeline that integrates GVL-biorefinery producing furfural and dissolving pulp, with lignin hydrogenolysis and biological funneling …….. 69

Figure 3.2. Probing the ability for *N. aromaticivorans* strain PDC cultures to convert the major hydrogenolysis products to PDC in minimum media supplemented with glucose. Cell density and extracellular metabolite concentration of cultures supplemented with dihydrosinapyl alcohol (DSA) (A),

dihydroconiferyl alcohol (DCA) (B), 7,8-dihydroferulic acid methyl ester (Me-DHFA) (C), propyl syringol (PS) (D), propyl guaiacol (PG) (E), 7,8-dihydro-*p*-coumaric acid methyl ester (Me-DH*p*CA) (F), methyl syringol (MS) (G), methyl guaiacol (MG) (H), or *p*-hydroxybenzoic acid methyl ester (Me-*p*HBA) (I). Values correspond to the average of three biological replicates ……71

Figure 3.3. Representative *N. aromaticivorans* strain PDC cultures grown in minimal media supplemented with glucose and depolymerized GVL-lignin from maple (A), poplar (B), sorghum (C), and switchgrass (D). (Top row of plots) Cell density, (middle row of plots) glucose, total aromatics and PDC production, (bottom row of plots) extracellular metabolite concentration of primary monophenolic products from hydrogenolysis. Each data point corresponds to the average of three biological replicates using the lignin hydrogenolysis products from one of three depolymerization reactions …….. 77

Figure 3.S1. Cell density and extracellular metabolite concentration of *N. aromaticivorans* strain PDC cultures in minimum media supplemented with glucose and ethylsyringol (ES) (A), ethylguaiacol (EG) (B), syringol (S) (C), or guaiacol (G) (D). Values correspond to the average of three biological replicates ... 88

Figure 3.S2. UV/vis chromatogram of samples from experiments with ES and EG collected at the beginning of the experiment (A and B, respectively), after 18 hours (C and D), and abiotic control after 18 hours of incubation (E and F). ES and EG disappear after 18 hours of experiment with *N. aromaticivorans* strain PDC and new peaks are detected a RT = 4.79 min (C) and TR = 3.49 min (D). Abiotic controls show the presence of ES (E) and EG (F) and the absence of new peaks …….. 89

Figure 4.1. Pathways for the metabolism of S, G, and H phenolics that have been proposed for *Sphingobium sp*. SYK-6 (21-23) and *Novosphingobium aromaticivorans* (9, 15), showing the location of *O*-demethylation and aromatic ring opening steps ……... 100

Figure 4.2. Growth of indicated strains of *N. aromaticivorans* in minimal media supplemented with syringic acid (panel A) or vanillic acid (panel B). Top panels show data for parent strain and single deletion mutants, whereas bottom panels show growth curves for strains with multiple gene deletions …….104

Figure 4.3. Cell densities and extracellular compound concentrations for the indicated *N. aromaticivorans* strains grown in minimal media containing glucose and syringic acid. Values represent the average of three biological replicates ……..105

Figure 4.4. Growth and extracellular aromatic compound concentrations of different strains of *N. aromaticivorans* in minimal media supplemented with glucose and 3-MGA. Values represent the average of three biological replicates …… 108

Figure 4.5. Growth and extracellular aromatic compound concentrations of different strains of *N. aromaticivorans* cultures in minimal media supplemented with glucose and vanillic acid. Values represent the average of three biological replicates …… 109

Figure 4.6. Substrate and product concentration during in vitro enzyme assays of LigM (Saro_2861) (panels A, C, and E) and DesA (Saro_2404) (panels B, D, and F) with syringic acid (panels A and B), 3-MGA (panels C and D), and vanillic acid (panels E and F). Concentrations in reactions lacking recombinant proteins are shown in dashed line. Reported concentrations are the average of three replicate assays. The low amounts of 3-MGA measured in the assays with LigM and syringic acid (panel A) was a trace contaminant in the syringic acid used in these assays ……. 111

Figure 4.7. Growth of different strains of *N. aromaticivorans* in minimal media supplemented with syringic acid (panel A) or vanillic acid (panel B) …….. 113

Figure 4.8. Growth and extracellular compound concentration of indicated *N. aromaticivorans* cultures in minimal media supplemented with glucose and syringic acid ……114

Figure 4.9. Growth and extracellular compound concentration of the indicated strains of *N. aromaticivorans* in minimal media supplemented with glucose and vanillic acid. Values represent the average of three biological replicates ……. 116

Figure 4.10. Substrate and product concentration of in vitro assays of LigAB (Saro_2812/3) (panels A, C, and E) and LigAB2 (Saro_1233/4) (panels B, D, and F) on 3-MGA (panels A and B), GA (panels C and D), and PCA (panels E and F). Concentration of substrates in assays performed in the absence of added recombinant protein are shown in dashed line and the same symbol and color as condition with enzyme. Samples were analyzed with HPLC-MS/MS and concentrations correspond to the average of three replicates …….. 117

Figure 4.11. Updated pathway for metabolism of S, G, and H phenolics by *N. aromaticivorans*, with specific assignment of *O*-demethylase and aromatic ring opening dioxygenase activities based on the in vivo and in vitro experiments described in this study ……. 120

Figure 4.S1. Vanillic acid concentration during in vitro enzyme assays of LigM (Saro_2861) (A) and DesA (Saro_2404) (B) with 1 mM or 0.5 mM H_4folate. In conditions with 0.5 mM H_4folate, the substrate decreases slower than in conditions with 1 mM H_4folate …… 136

Figure 4.S2. Total ion chromatogram of samples collected after 1 hour of reaction time of in vitro assays with LigAB and LigAB2 incubated in the presence of 3-MGA …….. 136

Figure 4.S3. Average mass spectrum (m/z) of PDC (RT = 16.36 m – 16.40 m) and 3-MGA (RT = 16.60 m – 16.63 m) of samples collected after 1 hour of reaction time of in vitro assays with LigAB and LigAB2 incubated in the presence of 3-MGA ……... 137

Figure 4.S4. Total ion chromatogram of samples collected after 24 hour of reaction time of in vitro assays with LigAB and LigAB2 incubated in the presence of GA ……. 137

Figure 4.S5. Average mass spectrum (m/z) of PDC-TMS (RT = 16.36m – 16.40 m), GA-TMS (RT = 16.86 m – 16.89 m), and enol OMA-TMS (RT = 16.97 m – 17.03 m) of samples collected after 24 hours of reaction time of in vitro assays with LigAB and LigAB2 incubated in the presence of GA …….. 138

Figure 4.S6. UV-vis chromatogram and spectra of authentic PDCA standard (A and B), ammonium sulfate treated samples collected at 24 hours from in vitro assays with LigAB and PCA (C and D), LigAB2 and PCA (E and F), and control without enzyme (G) ……139

Figure 4.S7. Mass balance and CHMOD relative concentration from in vitro assays of LigAB (Saro_2812/3) (A) and LigAB2 (Saro_1233/4) (B) with 3-MGA as a substrate. Mass balance was calculated as the sum of the molar concentrations of 3-MGA and PDC, divided by the concentration of 3-MGA in a sample lacking enzyme. Results are expressed in percentage (%). Dotted lines show 100%. Samples were analyzed with HPLC-MS, values correspond to the integrated area of peaks generated by the EIC with m/z = 215, attributed to CHMOD and concentrations correspond to the average of three replicates …….140

Figure 4.S8. Binary gradient profile of HPLC-MS analysis …… 140

List of tables

Table 2.1. PDC yield from different aromatic compounds by *N. aromaticivorans strains* 12444*ΔligI* (*p*-coumaric acid, *p*-hydroxybenzaldehyde, *p*-hydroxybenzoic acid, ferulic acid, vanillin, and vanillic acid) and 12444*ΔligIΔdesCD* (G diketone, syringaldehyde, syringic acid, and S diketone) into PDC. Numbers in parenthesis represent one standard deviation of the average from 3 replicate cultures……. 25

Table 2.2. Comparison of c ell densities and extracellular concentrations at stationary phase of *N. aromaticivorans* strain 12444*ΔligIΔdesCD* cultures grown on glucose or glucose plus protocatechuic acid. Data shown represents the average of 3 biological replicates. Error represent one standard deviation … 31

Table 2.S1. Bacterial strains and plasmids used in this study …….. 52

Table 2.S2. Primers used in this study …….. 52

Table 2.S3. Multiple reaction module (MRM) conditions for HPLC-MS quantification of compounds used in this study …… 53

Table 3.1. Yield of PDC (mol) at 19-24 h reaction time from the aromatic substrates screened in this study. Standard error was determined from n=3 technical replicates …….. 72

Table 3.2. Characterization of the GVL-lignin and hydrogenolysis product mixture from the four representative biomasses, maple (S/G hardwood), poplar (S/G hardwood with *p*HBA esters), and two grasses sorghum and switchgrass (with different levels of *p*CA and FA esters. Standard error was determined from n=3 technical replicates …… 75

Table 3.3. PDC produced by *N. aromaticivorans* strain PDC from the crude hydrogenolysis product mixture of the four representative biomasses, maple (S/G hardwood), poplar (S/G hardwood with *p*HBA esters), and two grasses sorghum and switchgrass (with different levels of *p*CA and FA esters). Standard error was determined from n=3 technical replicates ……..78

Table 3.S1. Yield of PDC (mol) at 19-24 h reaction time from experiments with syringol (Sy), methylsyringol (MS), ethylsyringol (ES), guaiacol (Gu), methylguaiacol (MG), or ethylguaiacol (EG). Standard error was determined from n=3 technical replicates …….. 90

Table 3.S2. Ratio of quantified hydrogenolysis monomers determined by GC-FID analysis of the crude filtered methanol product solution …… 90

Table 3.S3. Predicted PDC yield based on convertible monomers quantified after hydrogenolysis assuming a stoichiometric conversion of convertible monomers (DSA, DCA, Me-*p*HBA, Me-DH*p*CA, MeDHFA, *p*HBA, MS, and MG) ……91

Table 3.S4. MRM transitions used to quantify aromatic monomers in the *N. aromaticivorans* bioconversion assays ……. 91

Table 4.1. List of *N. aromaticivorans* parent and mutant strains with deletions of putative *O*-demethylases used in this study …….. 103

Table 4.2. PDC yields from *N. aromaticivorans* strains in the presence of glucose plus the indicated aromatic substrate ……107

Table 4.3. List of *N. aromaticivorans* mutant strains with deletions of putative aromatic ring cleavage dioxygenases used in this study ……112

Table 4.4. PDC yields from *N. aromaticivorans* strains in the presence of glucose plus the indicated aromatic substrate ……115

Table 4.S1. Primers used in this study …… 141

Table 4.S2. Plasmids used in this study ……... 145

Table 4.S3. List of *E. coli* strains …… 146

Table 4.S4. Multiple reaction module conditions for HPLC-MS quantification of compounds used in this study …… 146

1. Introduction

1.1 Motivation

The construction of the modern world has been possible in part thanks to two revolutionary events that begun in north-western Europe during the last two decades of the 18th century, and resulted in the complete restructuring of the preceding world's order: the industrial and the French revolutions. Whereas the rather political French revolution provided the foundations for a new social and political order, the industrial (British) revolution provided the nature-transforming tools necessary to build the infrastructure that allowed the development of a new society of free people (1). One of the main consequences of the establishment of the modern society was the increased demand for energy and materials to fuel both the construction and the functioning of the new infrastructure (2). As the sources of energy used before the revolutionary times were insufficient to meet the ever increasing demand, the utilization of fossil carbon sources such as coal, oil and natural gas have become by far the dominant raw material consumed (2) (Figure 1.1). In addition, the manufacturing of new synthetic polymeric materials and chemicals such as pharmaceuticals, pesticides and fertilizers developed by the chemicals industry relies on similar sources of fossil carbon (3).

One of the main negative consequences of the extensive utilization of fossil carbon sources has been the disruption of the global carbon cycle. Large amounts of carbon that were fixed by ancient organisms, stored in marine sediments, and transformed by geological forces during millions of years into what today we know as fossil fuels (4) have been extracted, processed, consumed, and ultimately converted into carbon dioxide that accumulated in the atmosphere (Figure 1.2). This net accumulation of carbon dioxide in the atmosphere is proposed by many scientists to be the main contributor to the anthropogenic global warming phenomenon (5).

One potential alternative to fossil carbon that can provide energy and raw materials for the production of fuels and chemicals is lignocellulosic biomass (or lignocellulose, main structural component of plant cell

wall). Its production by plants rely on energy from solar radiation and the fixation of carbon dioxide from the atmosphere (6, 7) and represents one of the largest available sources of renewable carbon on earth (8, 9). In this work, I explore techniques to utilize lignocellulose as a raw material to produce chemicals.

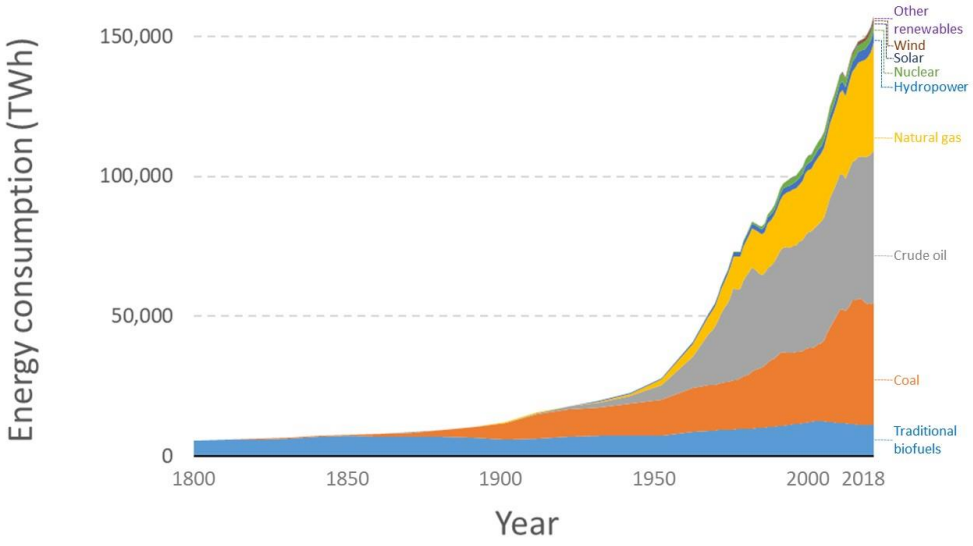


Figure 1.1. Global energy consumption by primary source from 1800 to 2018. The data used to generate this figure was obtained from Smil, 2017 (2).

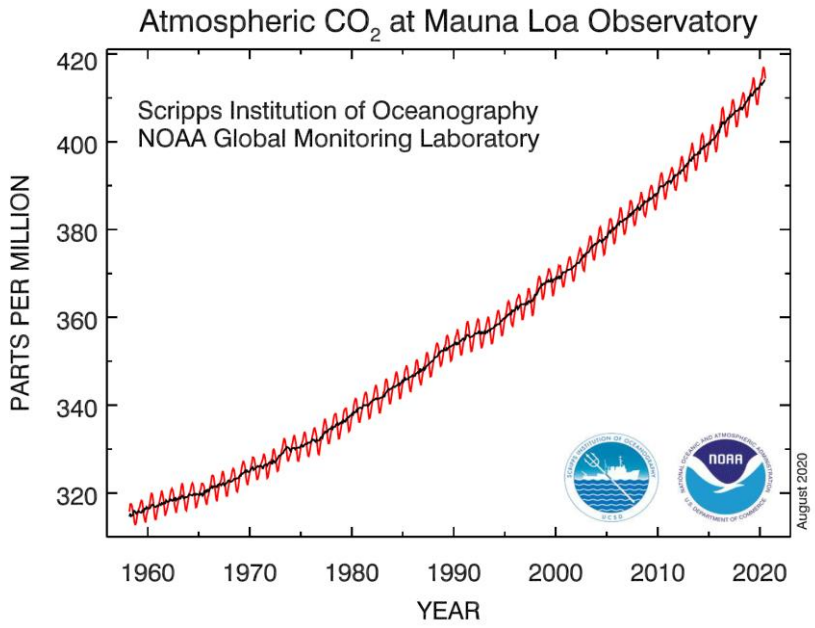


Figure 1.2. Concentration of CO2 in the earth's atmosphere from March 1958 to August 2020. Source: NOAA Earth System Research Laboratory (10).

1.2 The challenge of lignocellulose valorization

The main structural material that constitutes the plant cell wall, lignocellulose, is composed of three interlinked organic polymers: cellulose, hemicellulose, and lignin (11). Cellulose is a polysaccharide formed by glucose subunits interconnected by β-1-4 bonds and organized in crystalline microfibrils that provide the main structure and shape of the cell wall (11, 12). Hemicellulose is a group of amorphous heteropolymers of C5 and C6 sugars that play an important role in strengthening the cell wall structure (13). Lignin is an aromatic amorphous heteropolymer that, similar to hemicelluloses, plays a critical role in the integrity of the cell wall and in addition provides waterproofing properties and protection against pathogens (14).

The complex and heterogeneous nature of lignocellulose presents significant challenges to efficiently use it as a raw material for the production of fuels and chemicals. One approach that has shown promising

results deconstruction of lignocellulose by chemical techniques to separate the carbohydrate polymers (cellulose and hemicellulose) from lignin, followed by enzymatic depolymerization (saccharification) of the sugar fraction into its C5 and C5 monomeric constituents and convert them into alcohols via fermentation with microorganisms (15). However, use of the lignin fraction for a source of more than heat and power is still an unresolved challenge and developing appropriate techniques for its valorization is proposed by some to be crucial to make the lignocellulose biorefinery economically viable (16-18). This work focuses specifically on techniques to valorize lignin and other plant phenolics in the context of a lignocellulose biorefinery.

1.3 Lignin

As one of the major components of the plant cell wall, lignin can account for up to 30% of the total biomass weight (19). Plants produce lignin by extracellular radical coupling of aromatic compounds called monolignols. The most abundant monolignols used by plants to produce lignin are sinapyl alcohol, coniferyl alcohol and *p*-coumaryl alcohol. When incorporated into the lignin polymer, these monolignols produce different phenylpropanoid units that differ in the number of methoxy groups attached to the main phenolic structure: syringyl (S), with two methoxy groups; guaiacyl (G), one methoxy group; and *p*-hydroxyphenyl (H), with no methoxy groups (14, 20). The aromatic units in lignin are connected by different types of interlinkages and their relative abundance is affected by the relative abundance of the different type of monolignols (14). The most abundant lignin interlinkage is the β-O-4 (β-aryl ether) bond, which is also the most labile and thus the main target for lignin deconstruction strategies (21, 22). Some plant species have *p*-coumarate and ferulate (grasses) (23, 24) or p-hydroxybenzoate (*Populus sp.*) (25) esters as pendant groups attached as side chains to the main lignin structure.

1.4 Proposed strategy to valorize lignin

Multiple chemical strategies have been proposed to valorize lignin beyond its use as a source of energy for the refinery. Some strategies are focused on the utilization of isolated lignins without altering its polymeric nature (26), whereas other strategies rely on its depolymerization into low molecular weight compounds that can then be converted into valuable fuels and chemicals (16, 27). The second approach has gained significant attention with the emergence of multiple chemical techniques for lignocellulose deconstruction and lignin depolymerization that produce a high yield of monomers (21, 22, 27, 28). Other approaches for lignin depolymerization rely on biological mechanisms such as extracellular oxidative fungal enzymes (29) or *in vitro* enzymatic reactions that target specific lignin linkages (30).

Although depolymerization techniques can be effective in breaking down the lignin structure, they produce heterogeneous mixtures of multiple monomeric and oligomeric aromatic compounds that present significant challenges for further upgrade and purification. A strategy to handle this heterogeneity that has gained recent attention is the simultaneous transformation of multiple aromatic compounds into a single product using engineered bacteria in what has been called "biological funneling" (31-34). Some soil bacteria and fungi that thrive on decaying plant biomass and microorganisms that live in the digestive tract of lignocellulose-eating animals such as termites have evolved strategies to break down lignin and catabolize its aromatic constituents to produce energy and cell biomass building blocks (35, 36). The general biological strategy to utilize lignin as a food source consists of four main steps (Figure 1.3): 1) Lignin depolymerization performed by oxidative mechanisms involving enzymes such as laccases and peroxidases, the reduction of O_2 or hydrogen peroxide, and often mediators such as Mn^{3+} (35); 2) Metabolism of depolymerized materials by "upper pathways" to transform multiple monomeric and dimeric aromatic compounds released from the lignin polymer via convergent pathways into a few central aromatic intermediates that are compatible with aromatic ring opening mechanisms (37-40); 3) Aromatic ring opening under either anaerobic-anoxic (reductive) or aerobic (oxidative) conditions (37); 4) "Lower

pathways" for the conversion of ring opening products into central carbon metabolites such as acetyl-CoA or succinyl-CoA (37).

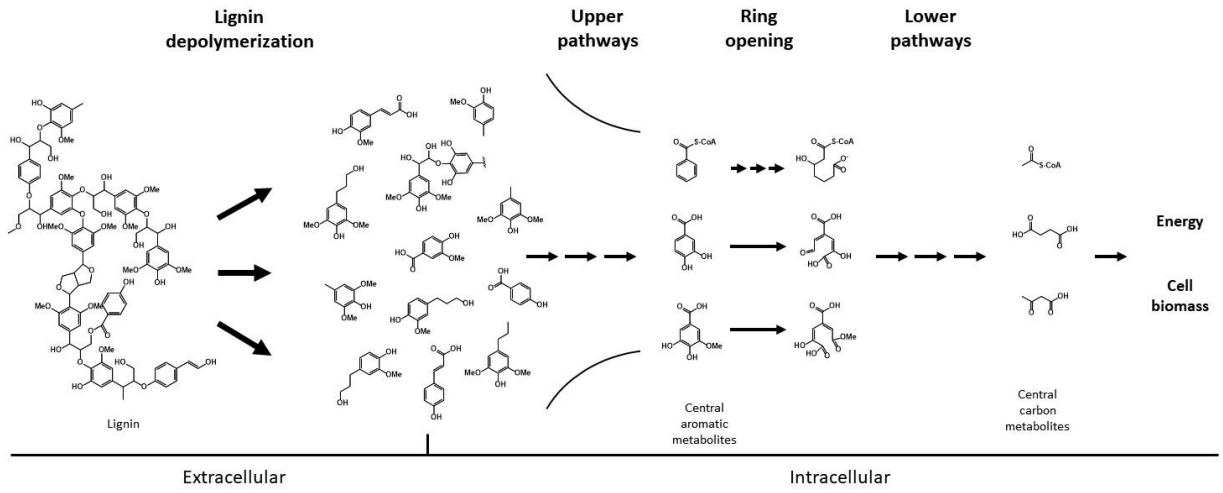


Figure 1.3. General microbial mechanism for lignin catabolism (35, 37)

The upper pathways of microbial aromatic metabolism can be repurposed in engineered microbes to simultaneously transform multiple aromatic compounds present in deconstructed biomass into a single valuable product. Some bacteria species such as *Pseudomonas putida* (32), *Rhodococcus jostii* (41), *Rhodopseudomonas palustris* (42), *Escherichia coli* (43), and yeast species such as *Rhodosporidium turuloides* (44) have been used as platform organisms for biological funneling. In specific examples, some of these microbes have been engineered to transform plant derived phenolics into aromatic compounds such as vanillin (45), gallate (46), pyrogallol (46), benzoic acid (42), aromatic dicarboxylic acids (45, 47) and non-aromatic compounds such as *cis*,*cis*-muconate (46, 48-51), 3-carboxy muconate (43), β-keto adipate (52), muconolactone (52), pyridine-2,4-dicarboxylic acid (53-58), and polyhydroxyalkanoates (32).

In this work, I propose to test the utility of a lignin valorization strategy that integrates chemical biomass deconstruction and lignin depolymerization with biological funneling to produce a single valuable product

(Figure 1.4). An ideal microbe to be used as a chassis for biological funneling of depolymerized lignin products should be capable of naturally metabolizing a wide variety of S, G, and H type aromatic compounds, compatible with biorefinery culture conditions and stress tolerant, and genetically tractable (31). I am interested in investigating *Novosphingobium aromaticivorans* DSM12444 as a potential platform organism for biological funneling of lignocellulose-derived aromatic compounds.

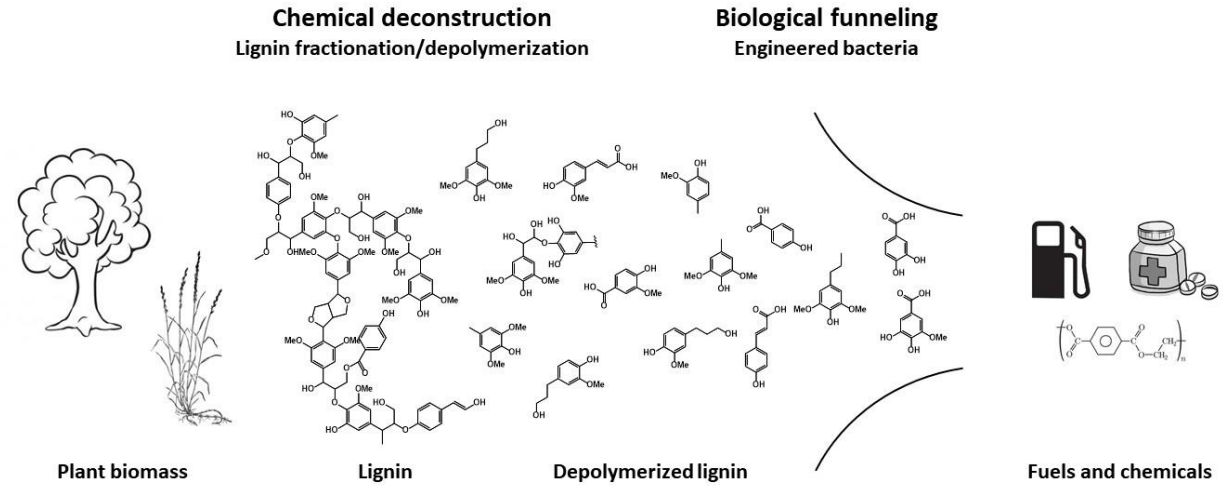


Figure 1.4. Proposed lignin valorization strategy

1.5 Novosphingobium aromaticivorans DSM12444

The gram negative *N. aromaticivorans* DSM12444 is an alphaproteobacterium first isolated from a deep sediment of the Atlantic Coastal Plain beneath the Savannah River Site (59-61) capable of catabolizing multiple aromatic compounds aerobically (62). In addition, this microorganism is capable of breaking β-O-4 lignin linkages of model aromatic dimers (40, 63, 64). It has been proposed that it can degrade multiple lignin derived aromatic compounds containing S structures via the central aromatic metabolite 3-methoxygallic acid (3-MGA), and G and H type aromatic compounds via protocatechuic acid (PCA) (62). Its proposed metabolic routes for aromatic compounds has numerous similarities to the mechanisms

found in the model organism *Sphingobium* sp. SYK-6 (38, 65). In addition, *N. aromaticivorans* is able to grow on non-aromatic substrates such as glucose (57). In previous work researchers have obtained genomic-scale data sets for this organism (62) and made this microbe amenable for genetic modifications (64).

1.6 Identification of research needs

Previous work with engineered bacteria to funnel lignocellulose-derived aromatic compounds, have included substantial efforts in expanding the spectrum of products possible to be made by a few different bacterial species. In some cases, such as *Pseudomonas putida* KT2440, the existing metabolic capabilities in the host microbe have been manipulated to reroute the metabolites fluxes into the product *cis,cis*-muconate (48). However, the range of aromatic substrates that can be metabolized by *P. putida* KT2440 includes G and H type but not S type. In other cases, such as *Pseudomonas putida* PpY1100, the platform organism does not have aromatic compounds metabolic capabilities so researchers imported them by adding a set of enzymes that allowed the engineered strain to funnel a few compounds into PDC (55). In addition, in most studies the engineered microbes have been shown to transform aromatic compounds with a limited variety of side-chains. As the different lignin chemical deconstruction techniques often produce compounds that retain the main aromatic structures found in lignin (S, G, or H) but contain a wide variety of side-chains (21), expanding the knowledge of bacterial aromatic metabolic capabilities and exploring new platform organisms is necessary. In addition, as numerous new chemical lignin depolymerization techniques emerge, it is necessary to investigate their compatibility with biological funneling in the context of a biomass-to-product pipeline or biorefinery (Figure 1.4). Thus, despite significant efforts by several groups, we still lack a microbial chassis capable of metabolizing the suite of aromatic compounds we expect to be derived from biomass, that is also amenable to the engineering improvements to produce valuable products from the heterogeneous material that is derived from one or more deconstruction methods

1.7 Research objectives

In this dissertation, I addressed these research needs by studying *N. aromaticivorans* as a potential new bacterial platform for biological funneling of depolymerized lignin. In Chapter 2, I describe the engineering of *N. aromaticivorans* to simultaneously funnel multiple S, G, and H lignin-derived aromatic compounds into PDC. I investigated the effect of different gene deletions in blocking the metabolic pathways and forcing the accumulation of the metabolic intermediate PDC. I calculated conversion yields from S, G, and H type aromatic compounds containing different side-chains and demonstrated the compatibility of the engineered strain with real lignin depolymerized by an oxidative method (66).

In the Chapter 3, I describe research on the compatibility of γ-valerolactone (GVL) (67) lignin isolation and reductive depolymerization (hydrogenolysis) (68) with biological funneling by the engineered strain of *N. aromaticivorans* described in Chapter 2 to produce PDC in the context of a biomass-to-PDC pipeline. First, a number of different aromatic compounds commonly found in lignin hydrogenolysis products were tested for PDC production by the engineered strain of *N. aromaticivorans*. Based on the results, I selected appropriate hydrogenolysis conditions to favor the production of compounds that showed high compatibility with biological funneling. Then, biomass from maple, poplar, sorghum, and switchgrass were independently pretreated and fractionated using GVL and the isolated lignin was depolymerized via hydrogenolysis using the selected conditions. The resulting depolymerization products were then tested for biological funneling. The results obtained in this study provided key information to improve the compatibility of lignin hydrogenolysis depolymerization with biological funneling into PDC by the engineered strain of *N. aromaticivorans*.

In Chapter 4, I describe research on the metabolism of S type aromatic compounds in *N. aromaticivorans* to determine the key reactions that affect its conversion yield into PDC. I investigated multiple aromatic *O*-demethylation and aromatic ring opening reactions involved in the metabolism of syringic acid and

vanillic acid. I created a set of gene deletion mutants targeting single aromatic *O*-demethylases and aromatic ring opening dioxygenases or combinations of two or more deletions and observed growth phenotypes and extracellular metabolite accumulation. Through these experiments, I found a potential new aromatic *O*-demethylase that I called DmtS. In addition, I expressed and purified the *O*-demethylases DesA and LigM, and the dioxygenases LigAB and LigAB2, and performed *in vitro* enzyme assays to determine their substrate specificity. As a result, I was able to create an improved PDC-producing engineered strain of *Novosphingobium aromaticivorans* that can generate PDC from syringic acid with stoichiometric yield.

In Chapter 5, I summarize the main findings and conclusions of this work and propose future directions for research.

1.8 References

1. Hobsbawm EJ. 1962. The age of revolution, 1789-1848, First Vintage Books edition ed. Vintage Books, New York.

2. Smil V. 2017. Energy transitions: Global and national perspectives, Second edition ed. Praeger.

3. Nagalakshmaiah M, Afrin S, Malladi RP, Elkoun S, Robert M, Ansari MA, Svedberg A, Karim Z. 2019. Biocomposites: Present trends and challenges for the future, p 197-215. *In* Koronis G, Silva A (ed), Green Composites for Automotive Applications doi:https://doi.org/10.1016/B978-0-08-102177-4.00009-4. Woodhead Publishing.

4. Fanchi JR. 2004. Chapter five - Origin of Fossil Fuels, p 126-158. *In* Fanchi JR (ed), Energy Technology and Directions for the Future doi:https://doi.org/10.1016/B978-0-12-248291-5/50007-9. Academic Press, Boston.

5. Anderson TR, Hawkins E, Jones PD. 2016. CO2, the greenhouse effect and global warming: from the pioneering work of Arrhenius and Callendar to today's Earth System Models. Endeavour 40:178-187.

6. Lee J. 1997. Biological conversion of lignocellulosic biomass to ethanol. J Biotechnol 56:1-24.

7. Perlack RD, Stokes BJ. 2011. U.S. billion-ton update: Biomass supply for a bioenergy and bioproducts industry. Oak Ridge National Laboratory, Oak Ridge, TN,

8. Tuck CO, Perez E, Horvath IT, Sheldon RA, Poliakoff M. 2012. Valorization of biomass: Deriving more value from waste. Science 337:695-699.

9. Agrawal R, Singh NR, Ribeiro FH, Delgass WN. 2007. Sustainable fuel for the transportation sector. Proc Natl Acad Sci 104:4828-4833.

10. ESRL. 2020. Trends in atmospheric carbon dioxide. https://www.esrl.noaa.gov/gmd/ccgg/trends/. Accessed Aug-26-2020.

11. Cosgrove DJ. 2005. Growth of the plant cell wall. Nat Rev Mol Cell Biol 6:850-861.

12. Carpita NC. 2011. Update on mechanisms of plant cell wall biosynthesis: How plants make cellulose and other (1→4)-β-d-Glycans. Plant Physiol 155:171-184.

13. Scheller HV, Ulvskov P. 2010. Hemicelluloses. Annu Rev Plant Biol 61:263-289.

14. Boerjan W, Ralph J, Baucher M. 2003. Lignin biosynthesis. Annu Rev Plant Biol 54:519-546.

15. Menon V, Rao M. 2012. Trends in bioconversion of lignocellulose: Biofuels, platform chemicals & biorefinery conceptMenon, Vishnu; Rao, Mala. Prog Energy Combust Sci 38:522-550.

16. Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon RA, Gilna P, Keller M, Langan P, Naskar AK, Saddler JN, Tschaplinski TJ, Tuskan GA, Wyman CE. 2014. Lignin valorization: Improving lignin processing in the biorefinery. Science 344.

17. Corona A, Biddy MJ, Vardon DR, Birkved M, Hauschild MZ, Beckham GT. 2018. Life cycle assessment of adipic acid production from lignin. Green Chem 20:3857-3866.

18. Davis R, Tao L, Tan ECD, Biddy MJ, Beckham GT, Scarlata C, Jacobson J, Cafferty K, Ross J, Lukas J, Knorr D, Schoen P. 2013. Lignocellulosic biomass to hydrocarbons: Dilute-acid and enzymatic deconstruction of biomass to sugars and biological conversion of sugars to hydrocarbons. NREL Tech Rep doi:10.2172/110.470:88-101.

19. Ralph J, Brunow G, Boerjan W. 2007. Lignins. eLS.

20. Ralph J, Lapierre C, Boerjan W. 2019. Lignin structure and its engineering. Curr Opin Biotechnol 56.

21. Schutyser W, Renders T, Van den Bosch S, Koelewijn SF, Beckham GT, Sels BF. 2018. Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. Chem Soc Rev 47:852-908.

22. Sun Z, Fridrich B, de Santi A, Elangovan S, Barta K. 2018. Bright side of lignin depolymerization: Toward new platform chemicals. Chem Rev 118:614-678.

23. Harris PJ, Hartley RD. 1976. Detection of bound ferulic acid in cell walls of the Gramineae by ultraviolet fluorescence microscopy. Nature 259:508.

24. Harris PJ, Hartley RD. 1980. Phenolic constituents of the cell walls of monocotyledons. Biochem Syst Ecol 8:153-160.

25. Smith DCC. 1955. Ester groups in lignin. Nature 176:267.

26. Luo H, Abu-Omar M. 2017. Chemicals from lignin doi:10.1016/B978-0-12-409548-9.10235-0.

27. Zakzeski J, Bruijnincx PCA, Jongerius AL, Weckhuysen BM. 2010. The catalytic valorization of lignin for the production of renewable chemicals. Chem Rev 110:3552-3599.

28. Galkin MV, Samec JSM. 2016. Lignin valorization through catalytic lignocellulose fractionation: A fundamental platform for the future biorefinery. ChemSusChem 9:1544-1558.

29. Martínez AT, Speranza M, Ruiz-Dueñas FJ, Ferreira P, Camarero S, Guillén F, Martínez MJ, Gutiérrez A, del Río JC. 2005-9. Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin. Int Microbiol 8:195-204.

30. Gall DL, Ralph J, Donohue TJ, Noguera DR. 2017. Biochemical transformation of lignin for deriving valued commodities from lignocellulose. Curr Opin Biotechnol 45:120-126.

31. Beckham GT, Johnson CW, Karp EM, Salvachua D, Vardon DR. 2016. Opportunities and challenges in biological lignin valorization. Curr Opin Biotechnol 42:40-53.

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

33. Xu Z, Lei P, Zhai R, Wen Z, Jin M. 2019. Recent advances in lignin valorization with bacterial cultures: microorganisms, metabolic pathways, and bio-products. Biotechnol Biofuels 12:32.

34. Li X, Zheng Y. 2019. Biotransformation of lignin: Mechanisms, applications and future work. Biotechnol Prog 36:e2922.

35. Bugg TD, Ahmad M, Hardiman EM, Rahmanpour R. 2011. Pathways for degradation of lignin in bacteria and fungi. Nat Prod Rep 28:1883-96.

36. Cragg SM, Beckham GT, Bruce NC, Bugg TDH, Distel DL, Dupree P, Etxabe AG, Goodell BS, Jellison J, McGeehan JE, McQueen-Mason SJ, Schnorr K, Walton PH, Watts JEM, Zimmer M. 2015. Lignocellulose degradation mechanisms across the tree of life. Curr Opin Chem Biol 29:108-10819.

37. Fuchs G, Boll M, Heider J. 2011. Microbial degradation of aromatic compounds - from one strategy to four. Nat Rev Microbiol 9:803-16.

38. Kamimura N, Takahashi K, Mori K, Araki T, Fujita M, Higuchi Y, Masai E. 2017. Bacterial catabolism of lignin-derived aromatics: New findings in a recent decade: Update on bacterial lignin catabolism. Environ Microbiol Rep 9:679-705.

39. Kontur WS, Bingman CA, Olmsted CN, Wassarman DR, Ulbrich A, Gall DL, Smith RW, Yusko LM, Fox BG, Noguera DR, Coon JJ, Donohue TJ. 2018. Novosphingobium aromaticivorans uses a Nu-class glutathione S -transferase as a glutathione lyase in breaking the β-aryl ether bond of lignin. J Biol Chem 293:4955-4968.

40. Kontur WS, Olmsted CN, Yusko LM, Niles AV, Walters KA, Beebe ET, Vander Meulen KA, Karlen SD, Gall DL, Noguera DR, Donohue TJ. 2019. A heterodimeric glutathione S-transferase that stereospecifically breaks lignin's β(R)-aryl ether bond reveals the diversity of bacterial β-etherases. J Biol Chem 294:1877-1890.

41. Sainsbury PD, Hardiman EM, Ahmad M, Otani H, Seghezzi N, Eltis LD, Bugg TDH. 2013. Breaking down lignin to high-value chemicals: The conversion of lignocellulose to vanillin in a gene deletion mutant of Rhodococcus jostii RHA1. ACS Chem Biol 8:2151-2156.

42. Austin S, Kontur WS, Ulbrich A, Oshlag JZ, Zhang W, Higbee A, Zhang Y, Coon JJ, Hodge DB, Donohue TJ, Noguera DR. 2015. Metabolism of multiple aromatic compounds in corn stover hydrolysate by Rhodopseudomonas palustris. Environ Sci Technol 49:8914.

43. Gosling A, Fowler SJ, O'Shea MS, Straffon M, Dumsday G, Zachariou M. 2011. Metabolic production of a novel polymer feedstock, 3-carboxy muconate, from vanillin. Appl Microbiol Biotechnol 90:107-116.

44. Yaegashi J, Kirby J, Ito M, Sun J, Dutta T, Mirsiaghi M, Sundstrom ER, Rodriguez A, Baidoo E, Tanjore D, Pray T, Sale K, Singh S, Keasling JD, Simmons BA, Singer SW, Magnuson JK, Arkin AP, Skerker JM, Gladden JM. 2017. Rhodosporidium toruloides: a new platform organism for conversion of lignocellulose into terpene biofuels and bioproducts. Biotechnol Biofuels 10.

45. Mycroft Z, Gomis M, Mines P, Law P, Bugg TDH. 2015. Biocatalytic conversion of lignin to aromatic dicarboxylic acids in Rhodococcus jostii RHA1 by re-routing aromatic degradation pathways. Green Chem 17:4974-4979.

46. Wu W, Dutta T, Varman A, Eudes A, Manalansan B, Loqué D, Singh S. 2017. Lignin valorization: Two hybrid biochemical routes for the conversion of polymeric lignin into value-added chemicals. Sci Rep 7:8420-8420.

47. Shindo K, Osawa A, Nakamura R, Kagiyama Y, Sakuda S, Shizuri Y, Furukawa K, Misawa N. 2004. Conversion from arenes having a benzene ring to those having a picolinic acid by simple growing cell reactions using Escherichia coli that expressed the six bacterial genes involved in biphenyl catabolism. J Am Chem Soc 126:15042-15043.

48. Vardon D, Franden MA, Johnson C, Karp E, Guarnieri M, Linger J, Salm M, Strathmann T, Beckham G, Ferguson G. 2015. Adipic acid production from lignin. Energy Environ Sci 8:617-628.

49. Salvachua D, Johnson CW, Singer C, Rohrer H, Peterson D, Black BA, Knapp A, Beckham G. 2018. Bioprocess development for muconic acid production from aromatic compounds and lignin. Green Chem 20:5007-5019.

50. Sonoki T, Takahashi K, Sugita H, Hatamura M, Azuma Y, Sato T, Suzuki S, Kamimura N, Masai E. 2018. Glucose-free cis,cis-muconic acid production via new metabolic designs corresponding to the heterogeneity of lignin. ACS Sustainable Chem Eng 6:1256-1264.

51. Shinoda E, Takahashi K, Abe N, Kamimura N, Sonoki T, Masai E. 2019. Isolation of a novel platform bacterium for lignin valorization and its application in glucose-free cis,cis-muconate production. J Ind Microbiol Biotechnol doi:10.1007/s10295-019-02190-6:1-10.

52. Okamura-Abe Y, Abe T, Nishimura K, Kawata Y, Sato-Izawa K, Otsuka Y, Nakamura M, Kajita S, Masai E, Sonoki T, Katayama Y. 2016. Beta-ketoadipic acid and muconolactone production from a lignin-related aromatic compound through the protocatechuate 3,4-metabolic pathway. J Biosci Bioeng 121:652-658.

53. Otsuka Y, Nakamura M, Shigehara K, Sugimura K, Masai E, Ohara S, Katayama Y. 2006. Efficient production of 2-pyrone 4,6-dicarboxylic acid as a novel polymer-based material from protocatechuate by microbial function. Appl Microbiol Biotechnol 71:608-614.

54. Nakajima M, Nishino Y, Tamura M, Mase K, Masai E, Otsuka Y, Nakamura M, Sato K, Fukuda M, Shigehara K, Ohara S, Katayama Y, Kajita S. 2009. Microbial conversion of glucose to a novel chemical building block, 2-pyrone-4,6-dicarboxylic acid. Metab Eng 11:213-220.

55. Qian Y, Otsuka Y, Sonoki T, Mukhopadhyay B, Nakamura M, Jellison J, Goodell B. 2016. Engineered microbial production of 2-pyrone-4,6-dicarboxylic acid from lignin residues for use as an industrial platform chemical. BioResources 11:6097-6109.

56. Luo Z, Kim W, Lee S. 2018. Metabolic engineering of Escherichia coli for efficient production of 2-pyrone-4,6-dicarboxylic acid from glucose. ACS Synth Biol 7:2296-2307.

57. Perez JM, Kontur WS, Alherech M, Coplien J, Karlen SD, Stahl SS, Donohue TJ, Noguera DR. 2019. Funneling aromatic products of chemically depolymerized lignin into 2-pyrone-4-6-dicarboxylic acid with Novosphingobium aromaticivorans. Green Chem 21:1340-1350.

58. Suzuki Y, Okamura-Abe Y, Nakamura M, Otsuka Y, Araki T, Otsuka H, Navarro RR, Kamimura N, Masai E, Katayama Y. 2020. Development of the production of 2-pyrone-4,6-dicarboxylic acid from lignin extracts, which are industrially formed as by-products, as raw materials. J Biosci Bioeng doi:10.1016/j.jbiosc.2020.02.002.

59. Fredrickson JK, Brockman FJ, Workman DJ, Li SW, Stevens TO. 1991. Isolation and characterization of a subsurface bacterium capable of growth on toluene, naphthalene, and other aromatic compounds. Appl Environ Microbiol 57:796-803.

60. Fredrickson JK, Balkwill DL, Drake GR, Romine MF, Ringelberg DB, White DC. 1995. Aromatic-degrading Sphingomonas isolates from the deep subsurface. Appl Environ Microbiol 61:1917.

61. Fredrickson JK, Balkwill DL, Romine MF, Shi T. 1999. Ecology, physiology, and phylogeny of deep subsurface Sphingomonas sp. J Ind Microbiol Biotechnol 23:273-283.

62. Cecil JH, Garcia DC, Giannone RJ, Michener JK. 2018. Rapid, parallel identification of pathways for catabolism of lignin-derived aromatic compounds in Novosphingobium aromaticivorans. Appl Environ Microbiol 84:1-13.

63. Gall DL, Ralph J, Donohue T, Noguera D. 2014. A Group of Sequence-Related Sphingomonad Enzymes Catalyzes Cleavage of beta-Aryl Ether Linkages in Lignin beta-Guaiacyl and beta-Syringyl Ether Dimers. Environ Sci Technol 48:12454-12463.

64. Kontur WS, Bingman CA, Olmsted CN, Wassarman DR, Ulbrich A, Gall DL, Smith RW, Yusko LM, Fox BG, Noguera DR, Coon JJ, Donohue TJ. 2018. Novosphingobium aromaticivorans uses a Nu-class glutathione S-transferase as a glutathione lyase in breaking the β-aryl ether bond of lignin. J Biol Chem 293:4955-4968.

65. Masai E, Katayama Y, Fukuda M. 2007. Genetic and biochemical investigations on bacterial catabolic pathways for lignin-derived aromatic compounds. Biosci, Biotechnol, Biochem 71:1-15.

66. Rahimi A, Arne U, Joshua JC, Shannon SS. 2014. Formic- acid- induced depolymerization of oxidized lignin to aromatics. Nature 515:249-252.

67. Luterbacher JS, Azarpira A, Motagamwala AH, Lu F, Ralph J, Dumesic JA. 2015. Lignin monomer production integrated into the γ-valerolactone sugar platform. Energy Environ Sci 8:2657-2663.

68. Li Y, Shuai L, Kim H, Motagamwala AH, Mobley JK, Yue F, Tobimatsu Y, Havkin-Frenkel D, Chen F, Dixon RA, Luterbacher JS, Dumesic JA, Ralph J. 2018. An “ideal lignin” facilitates full biomass utilization. Sci Adv 4.

2. Funneling aromatic products of chemically depolymerized lignin into 2-pyrone-4-6-dicarboxylic acid with *Novosphingobium aromaticivorans*

A version of this chapter has been published as:

Perez, Jose. M., Wayne S. Kontur, Manar Alherech, Jason Coplien, Steven D. Karlen, Shannon S. Stahl, Timothy J. Donohue, and Daniel R. Noguera. 2019. 'Funneling aromatic products of chemically depolymerized lignin into 2-pyrone-4-6-dicarboxylic acid with Novosphingobium aromaticivorans', Green Chemistry, 21: 1340-50.

The contribution from each co-author is as follows:

Jose M. Perez: Conceived the Project, created bacterial mutant strains, planned and performed the experiments, performed HPLC-MS analysis, wrote the manuscript

Wayne S. Kontur: Created bacterial mutant strains, wrote the manuscript

Manar Alerech: Produced depolymerized lignin, synthesized aromatic diketones, wrote the manuscript

Jason Coplien: Synthesized aromatic diketones

Steven D. Karlen: Performed ^{1}H NMR and GPC analyses, wrote the manuscript

Shannon S. Stahl: Wrote the manuscript

Timothy J. Donohue: Wrote the manuscript

Daniel R. Noguera: Wrote the manuscript

2.1 Abstract

Lignin is an aromatic heteropolymer found in plant biomass. Depolymerization of lignin, either through biological or chemical means, invariably produces heterogeneous mixtures of low molecular weight aromatic compounds. Microbes that can metabolize lignin-derived aromatics have evolved pathways that funnel these heterogeneous mixtures into a few common intermediates before opening the aromatic ring. In this work, I engineered *Novosphingobium aromaticivorans* DSM12444, via targeted gene deletions, to use its native funneling pathways to simultaneously convert plant-derived aromatic compounds containing syringyl (S), guaiacyl (G), and *p*-hydroxyphenyl (H) aromatic units into 2-pyrone-4,6-dicarboxylic acid (PDC), a potential polyester precursor. In batch cultures containing defined media, the

engineered strain converted several of these depolymerization products, including S-diketone and G-diketone (non-natural compounds specifically produced by chemical depolymerization), into PDC with yields ranging from 22% to 100%. In batch cultures containing a heterogeneous mixture of aromatic monomers derived from chemical depolymerization of poplar lignin, 59% of the measured aromatic compounds were converted to PDC. Overall, our results show that *N. aromaticivorans* has ideal characteristics for its use as a microbial platform for funneling heterogeneous mixtures of lignin depolymerization products into PDC or other commodity chemicals.

2.2 Introduction

The impact of fossil carbon utilization on the global environment has encouraged the search for sustainable strategies to convert renewable resources into fuels and chemicals. Biorefining, the industrial activity of deriving fuels and chemicals from plant biomass in a sustainable and economically viable manner, is essential to reduce the proportion of fossil fuels that powers the global economy. Plant biomass, the most abundant renewable organic resource on Earth, is primarily composed of sugars and phenolic compounds (3, 4). While there are already established approaches to derive fuels from the sugar components of plant biomass (5), effective methods for biomass deconstruction to recover and valorize the phenolic components are only starting to emerge (6, 7). One source of phenolic compounds is lignin, an alkyl-aromatic heteropolymer that is interlinked with cellulose and hemicellulose in plant cell walls and accounts for up to 30% of the total lignocellulosic biomass weight (8). There are other sources of phenolics in plant biomass, such as arabinofuranosides in grasses (9, 10) or lignin bound *p*-hydroxybenzoate in some hardwoods (11). I am interested in evaluating the potential for bio-based production of valuable chemicals from the phenolic components of plant biomass.

The most abundant biomass-derived phenolics can be classified based on the number of methoxy groups attached to the main phenyl structure; these are syringyl (S; two methoxy groups), guaiacyl (G; one

methoxy group), and *p*-hydroxyphenyl (H; no methoxy groups) units (12). Several approaches have been recently described for biomass deconstruction and lignin depolymerization that result in recovery of S, G, and H aromatic units (6). However, the heterogeneity of the resulting mixtures presents a major challenge for conversion into commodity chemicals because of the low quantity of valuable marketable compounds in deconstructed lignin samples and the technical limitations for their separation or purification from other components (7).

I am exploring microbial strategies for the conversion of deconstructed lignin into commodity chemicals since microorganisms have evolved strategies to gain energy from the degradation of a large variety of aromatics compounds (13, 14). In general, microbial transformation of aromatic compounds occurs by a combination of upper metabolic pathways, which convert multiple compounds into key aromatic intermediates (13) in what has been called "biological funneling" (16), and a central aromatic pathway that breaks the aromaticity and renders metabolic products that enter central carbon metabolism (13, 14). Biological funneling has been recently described for the conversion of plant-derived phenolics to aromatic compounds such as vanillin (17) and benzoic acid (18), and to non-aromatic compounds, such as *cis,cis*-muconate (19), β-keto adipate (20), muconolactone (20), 2-pyrone-4,6-dicarboxylic acid (PDC) (21, 22). pyridine-2,4-dicarboxylic acid (23), and polyhydroxyalkanoates (16). Some of these approaches require extensive metabolic re-routing and introduction of foreign pathways (19, 22), while others rely on a small number of mutations that redirect aromatic metabolism to the product of interest (17, 18).

Here I report on the impact of gene deletions in the central aromatic catabolic pathways of *Novosphingobium aromaticivorans* DSM12444, an organism known to degrade aromatic compounds (24) and to break down interlinkages in lignin (25) that allow it to funnel a large diversity of plant-derived phenolics into PDC, a potential bioplastic and epoxy adhesives precursor (26). A complete genome sequence is available for this *α*-proteobacterium (GenBank NC_007794.1), and the organism is amenable to genetic and genomic techniques needed to test the role of individual genes in aromatic metabolism, and model, engineer, or improve its pathways (25). Specifically, I show that by using a defined set of

mutations, *N. aromaticivorans* can be engineered to simultaneously produce PDC from all three major types of plant-derived phenolic compounds (S, G, and H). In addition, I found that this organism can metabolize aromatics simultaneously with the use of other organic carbon sources (such as glucose), a feature that allows mutant strains to excrete compounds derived from the incomplete metabolism of the aromatics. This work represents a valuable advance in using bacteria to funnel aromatic compounds into defined single commodities and shows that *N. aromaticivorans* could be an ideal microbial chassis for valorization of lignin and other plant-derived aromatics.

2.3 Results

2.3.1 Model of aromatic metabolism by N. aromaticivorans DSM12444 and justification of experimental approach

N. aromaticivorans DSM12444, a bacterium isolated from a polyaromatic hydrocarbon-contaminated sediment in the deep subsurface, aerobically utilizes a variety of aromatic compounds as sole carbon and energy sources for growth (24). Based on its genome content, a recent analysis of *N. aromaticivorans* aromatic metabolism using a transposon library (2), and the known metabolism of lignin-derived aromatics in the related α-proteobacterium *Sphingobium sp*. SYK-6 (1), I propose a model for the degradation pathways of plant-derived aromatic compounds in this organism (Figure 2.1). Consistent with the predicted pathways in *N. aromaticivorans* and *Sphingobium sp*. SYK-6, I propose that G and H aromatic units are degraded via protocatechuic acid (Figure 2.1), with ring opening by LigAB, a 4,5 extradiol dioxygenase that yields 4-carboxy-2-hydroxy-cis,cis-muconate-6-semialdehyde (CHMS). CHMS is then oxidized to PDC by the dehydrogenase LigC. LigI is predicted to hydrolyze PDC to produce 4-oxalomesaconate (OMA) (27), which is further transformed to the central carbon metabolites pyruvate and oxaloacetate (Figure 2.1).

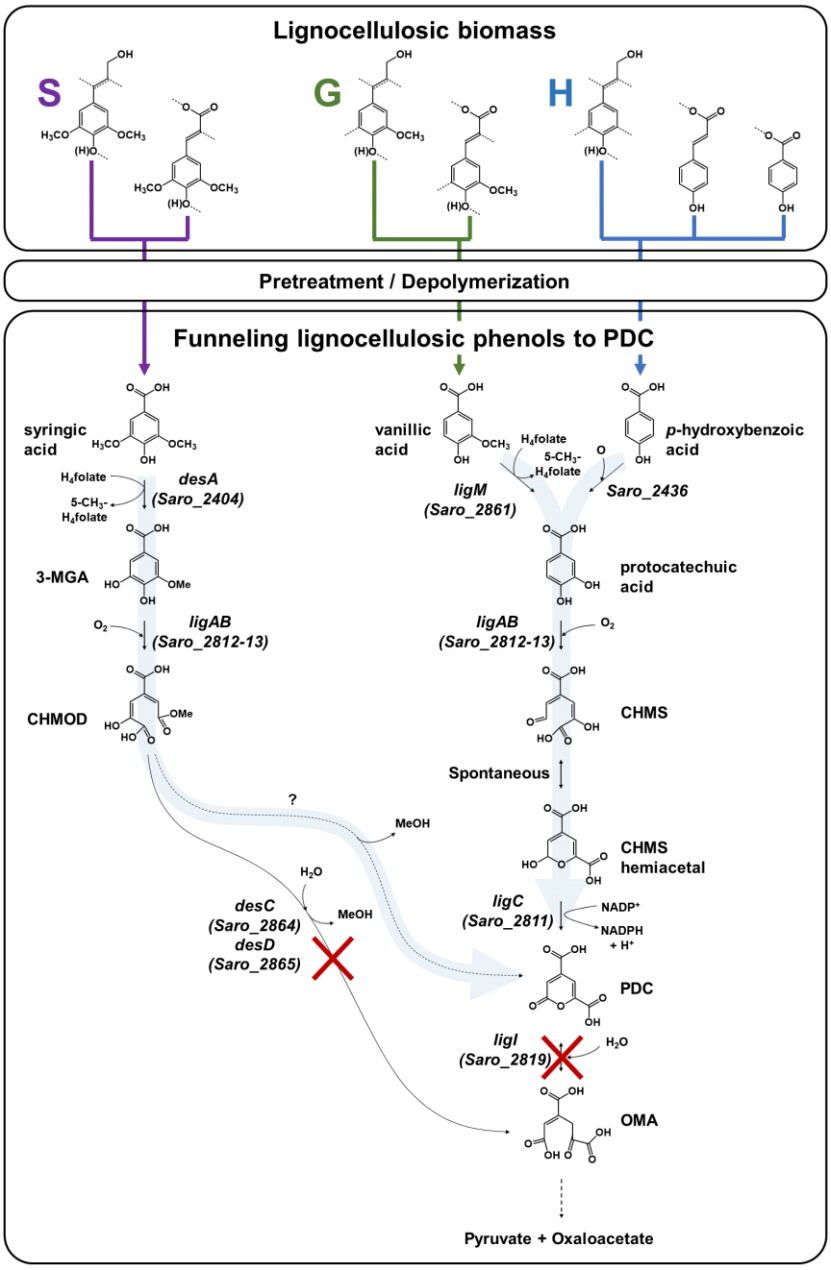


Figure 2.1. Predicted pathways of S unit (syringic acid), G unit (vanillic acid), and H unit (*p*-hydroxybenzoic acid) metabolism in *N. aromaticivorans* DSM12444, based on work in *Sphingobium sp*. SYK 6 (1) and *N. aromaticivorans* DSM12444 (2). In this model, deletions of the genes *ligI* (Saro_2819), *desC* (Saro_2864), and *desD* (Saro_2865) are hypothesized to enable the funneling (represented by light blue arrows) of S, G, and H lignocellulosic biomass derived aromatic compounds into 2-pyrone-4,6-dicarboxylic acid (PDC). Abbreviations: 3-methylgallate, 3-MGA; 4-carboxy-2-hydroxy-6-methoxy-6-oxohexa-2,4-dienoate, CHMOD; 4-carboxy-2-hydroxy-*cis*,*cis*-muconate-6-semialdehyde, CHMS; 4-oxalomesaconate, OMA.

Dimethoxylated aromatics (S aromatics) are predicted to be degraded via a separate pathway, with demethylation of syringic acid to 3-methylgallate (3-MGA) carried out by the *O*-demethylase DesA (Figure 2.1). In *N. aromaticivorans*, LigAB has been proposed to catalyze ring opening to produce a mixture of stereoisomers of 4-carboxy-2-hydroxy-6-methoxy-6-oxohexa-2,4-dienoate (CHMOD); a *cis-trans* isomerase, DesD, isomerizes one of the stereoisomers, and the methylesterase DesC completes demethylation of CHMOD to OMA (2). Two other routes of 3-MGA degradation are proposed in *Sphingobium sp*. SYK-6, one requiring ring opening by the 3,4-dioxygenase DesZ and cyclization to PDC and another one requiring *O*-demethylation to gallate by LigM followed by ring opening by the dioxygenase DesB (1). While LigM is present in *N. aromaticivorans*, homologues of DesZ and DesB are not encoded in its genome (2). In addition, the LigAB of *Sphingobium sp*. SYK-6 has been shown to produce a combination of CHMOD and PDC when 3-MGA is the substrate (28), and there are reports of slow abiotic transformation of CHMOD to PDC (29). Therefore, in our model (Figure 2.1), I hypothesize that the main enzymatic route of 3-MGA degradation in *N. aromaticivorans* is via CHMOD to OMA, but that PDC may also be a product of enzymatic or abiotic CHMOD transformation.

I used the above model to hypothesize which disruptions in the aromatic degradation pathways in *N. aromaticivorans* would lead to accumulation of specific pathway intermediates. I chose to focus on creating mutations that could lead to accumulation of PDC (Figure 2.1), which is of particular interest since this dicarboxylic acid has been shown to be a suitable precursor for polyesters (30). I hypothesized that a disruption of the proposed G and H degradation pathway via the deletion of the *ligI* gene (Figure 2.1) would prevent PDC degradation and lead to its accumulation in cultures fed G and H aromatics as substrates. Furthermore, I predicted that this metabolic disruption would result in strains with limited ability to grow on G and H aromatics, since most of the carbon in these compounds would remain in the PDC molecule. If this latter prediction is correct, then the addition of another substrate would be needed to support growth of cells on G or H aromatics lacking a functional *ligI* gene. In addition, given the possibility of PDC production from CHMOD (Figure 2.1), I also hypothesized that deleting the *desCD*

genes would result in accumulation of upstream intermediates and redirection of metabolism via PDC (Figure 2.1).

Below I describe how we tested these hypotheses and how the defined mutations lead to PDC accumulation from (i) G and H units, (ii) S, G, and H units, and (iii) aromatics that are present in depolymerized lignin.

2.3.2 Construction of a N. aromaticivorans mutant that accumulates PDC from G and H aromatics

We constructed strain 12444*ΔligI* by deleting the *ligI* gene and cultured it initially in minimal media containing glucose since this gene was not predicted to be necessary for glucose metabolism. To test the role of this gene in metabolism of aromatic compounds, I attempted to grow strain 12444*ΔligI* on minimal media containing 3 mM vanillic acid or 3 mM *p*-hydroxybenzoic acid as representative of G and H aromatics, respectively. As expected, strain 12444*ΔligI* was unable to grow on either of these substrates as sole carbon sources (Figure 2.2). When glucose was provided in addition to vanillic acid or *p*-hydroxybenzoic acid, strain 12444*ΔligI* was able to grow (Figure 2.3 panels A and B) and, in both cases, glucose and the aromatic substrate were removed from the media, a small amount of protocatechuic acid transiently accumulated, and PDC accumulated as the final product of the transformations (Figure 2.3 panels C and D). The PDC yield from vanillic acid and *p*-hydroxybenzoic acid by strain 12444*ΔligI* in these cultures were 81% (± 17%) and 73% (± 1.7%), respectively (Table 2.1).

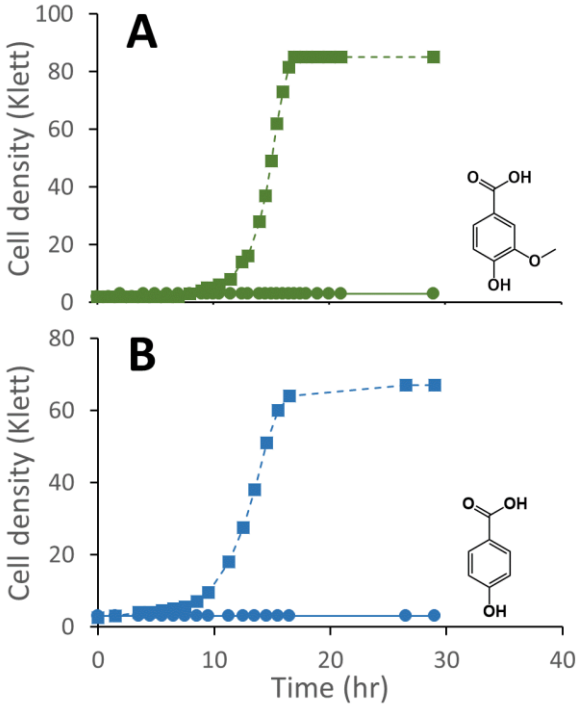


Figure 2.2. Cell density of representative *N. aromaticivorans* cultures grown on 3 mM vanillic acid (panel A) or 3 mM *p*-hydroxybenzoic acid (Panel B). Parent strain 12444*Δ1879* represented by squares and dashed line; strain 12444*ΔligI* represented by circles.

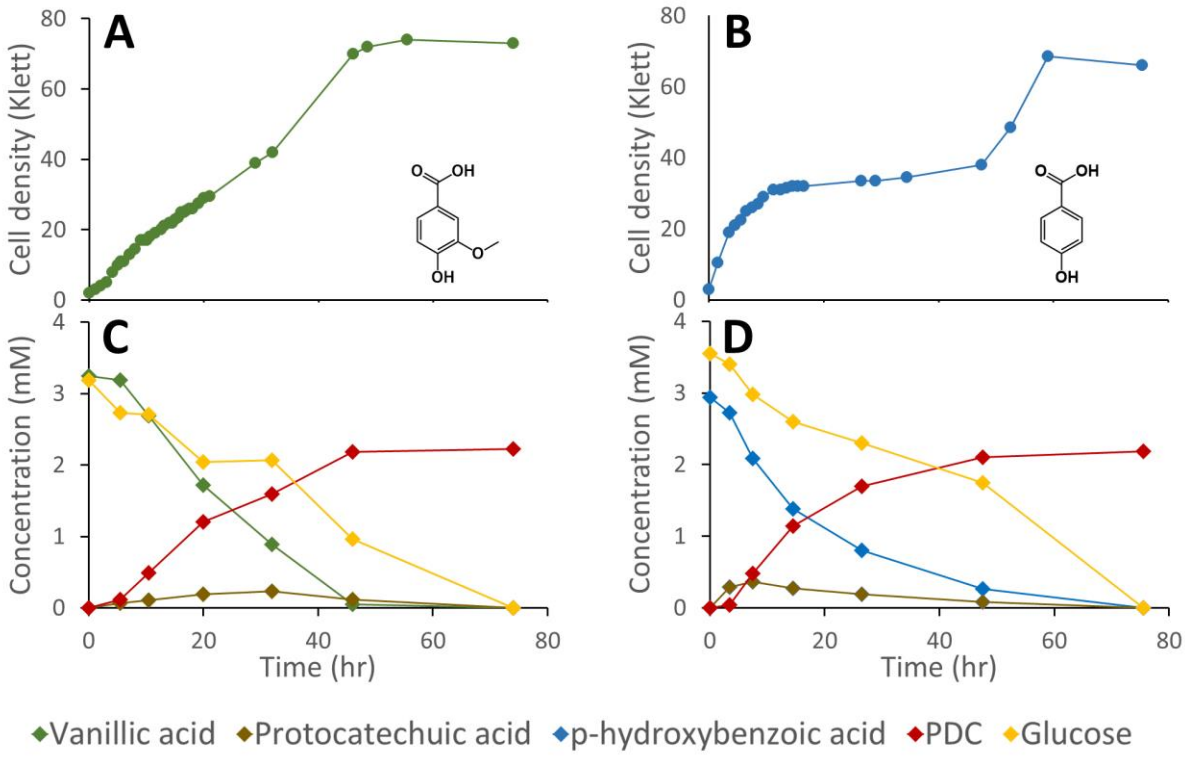
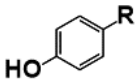
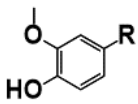
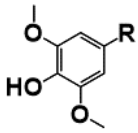


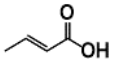
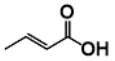
Figure 2.3. Cell density and extracellular metabolite concentration of representative *N. aromaticivorans* strain 12444*ΔligI* cultures grown on a combination of 3 mM vanillic acid and 3 mM glucose (panels A and C) or a combination of 3 mM *p*-hydroxybenzoic acid and 3 mM glucose (panels B and D).

Table 2.1. PDC yield from different aromatic compounds by *N. aromaticivorans* strains 12444*ΔligI* (*p*-coumaric acid, *p*-hydroxybenzaldehyde, *p*-hydroxybenzoic acid, ferulic acid, vanillin, and vanillic acid) and 12444*ΔligIΔdesCD* (G diketone, syringaldehyde, syringic acid, and S diketone) into PDC. Numbers in parenthesis represent one standard deviation of the average from 3 replicate cultures.







Compound	R	Yield (%)
p -coumaric acid		84 (± 5.4)
p -hydroxybenzaldehyde		79 (± 2.0)
p -hydroxybenzoic acid		73 (± 1.7)
ferulic acid		76 (± 10.0)
vanillin		100 (± 0.1)
vanillic acid		81 (± 17.0)
G-diketone		107 (± 1.6)
syringaldehyde		90 (± 7.2)
syringic acid		66 (± 13.2)
S-diketone		22 (± 0.7)

In theory, other G and H aromatics metabolized by *N. aromaticivorans* would also produce PDC when fed to strain 12444*ΔligI* (Figure 2.1). I tested this prediction with the G aromatics vanillin and ferulic acid and the H aromatics *p*-hydroxybenzaldehyde and *p*-coumaric acid (Figure 2.S1 and Table 2.1). Cultures grown on minimum media with 3 mM vanillin plus 3 mM glucose showed transient accumulation of vanillic acid (Figure 2.S1A), then a nearly stoichiometric accumulation of PDC. In the cultures grown with glucose and *p*-hydroxybenzaldehyde (Figure 2.S1B), a transient accumulation of extracellular *p*-

hydroxybenzoic acid and protocatechuic acid was observed, then accumulation of PDC with a 79% (± 2%) yield (Table 2.1). Cultures grown on ferulic acid plus glucose showed a transient accumulation of vanillic acid and protocatechuic acid (Figure 2.S1C), then accumulation of PDC with a 76% (± 10%) yield (Table 2.1). Similarly, the cultures grown with p-coumaric acid and glucose transiently accumulated extracellular p-hydroxybenzoic and protocatechuic acids (Figure 2.S1D), then accumulated PDC with an efficiency of 84% (± 5.4%) (Table 2.1).

These results are consistent with transformation of G and H aromatics via the predicted pathway of Figure 2.1. The observed PDC yields (Table 2.1) suggest that PDC is the main intermediate that accumulates, and that disruption of the *ligI* gene is sufficient to prevent PDC catabolism.

The inability of 12444*ΔligI* to metabolize PDC is not predicted to affect the degradation of aromatics containing S units, since the metabolism of these compounds would follow the 3-MGA, CHMOD, OMA pathway (Figure 2.1). In agreement with this hypothesis, when strain 12444*ΔligI* was fed 3 mM syringic acid as the sole carbon source, growth of this mutant reached final cell densities similar to those of parent strain 12444*Δ1879* and this aromatic was metabolized to a similar extent in both strains (Figure 2.4). This observation confirms that LigI is not necessary for syringic acid degradation. However, these experiments also showed that PDC accumulates in the growth media in both cases, representing 28% (0.97 mM) and 26% (0.89 mM) of the initial concentration of syringic acid for strains 12444*Δ1879* and 12444*ΔligI*, respectively.

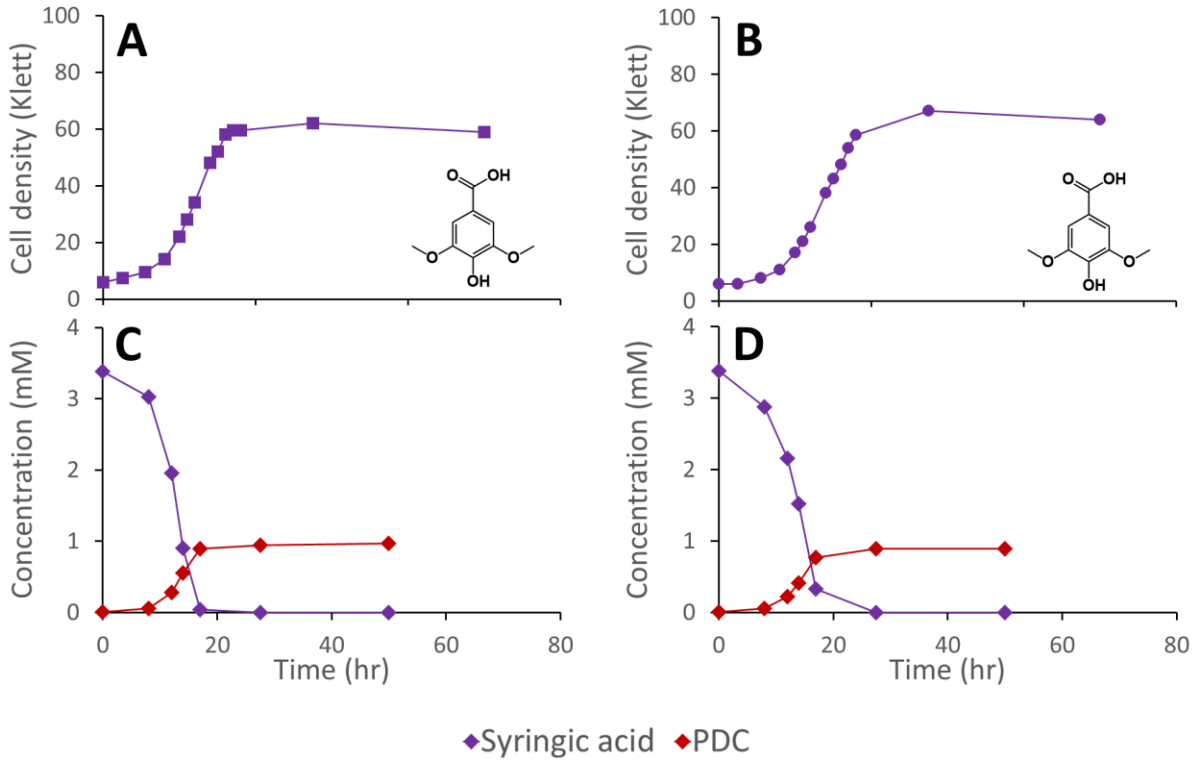


Figure 2.4. Cell density and extracellular metabolite concentrations of representative cultures of *N. aromaticivorans* strains 12444*Δ1879* (panels A and C) and 12444*ΔligI* (panels B and D) grown in media containing syringic acid.

2.3.3 Construction of an N. aromaticivorans mutant that accumulates PDC from S aromatics

Dimethoxylated phenolics, such as syringic acid, are predicted to be degraded by *N. aromaticivorans* via the 3-MGA, CHMOD, OMA pathway (Figure 2.1). Based on this prediction, I hypothesize that deleting the *desCD* genes would disrupt the degradation of S aromatics (Figure 2.1), leading to the accumulation of the intermediate CHMOD. However, this mutation may not be sufficient to prevent growth of *N. aromaticivorans* on S aromatics because CHMOD may undergo abiotic or enzymatic transformation to PDC (29), which could then be hydrolyzed by LigI. Thus, to test these hypotheses, we constructed strain 12444*ΔdesCD* by deleting the *desCD* genes from strain 12444*Δ1879*.

Growth was not observed when strain 12444*ΔdesCD* was cultured in minimal media with 3 mM syringic acid as the sole carbon source (Figure 2.5A), indicating that either *desC*, *desD* or both genes are essential for growth on syringic acid, in agreement with observations reported previously (2). To test the 12444*ΔdesCD* strain for a defect in S aromatic metabolism when growth was occurring, I inoculated the strain into media containing both 3 mM glucose and 3 mM syringic acid (Figure 2.5C). The 12444*ΔdesCD* strain grew, with consumption of both syringic acid and glucose, and with increased PDC accumulation compared to strain 12444*Δ1879*, converting 49% (± 0.9%) of the syringic acid into PDC (versus 28% for 12444*Δ1879*; Figure 2.4C). This suggests that increased cyclization of CHMOD to PDC took place, although this observation is not sufficient to determine whether the reaction is abiotic or enzymatic. Growth of 12444*ΔdesCD* on vanillic acid as the only carbon source demonstrated that the disruption in *desCD* does not affect the catabolism of G units and does not lead to detectable PDC accumulation (Figure 2.S2).

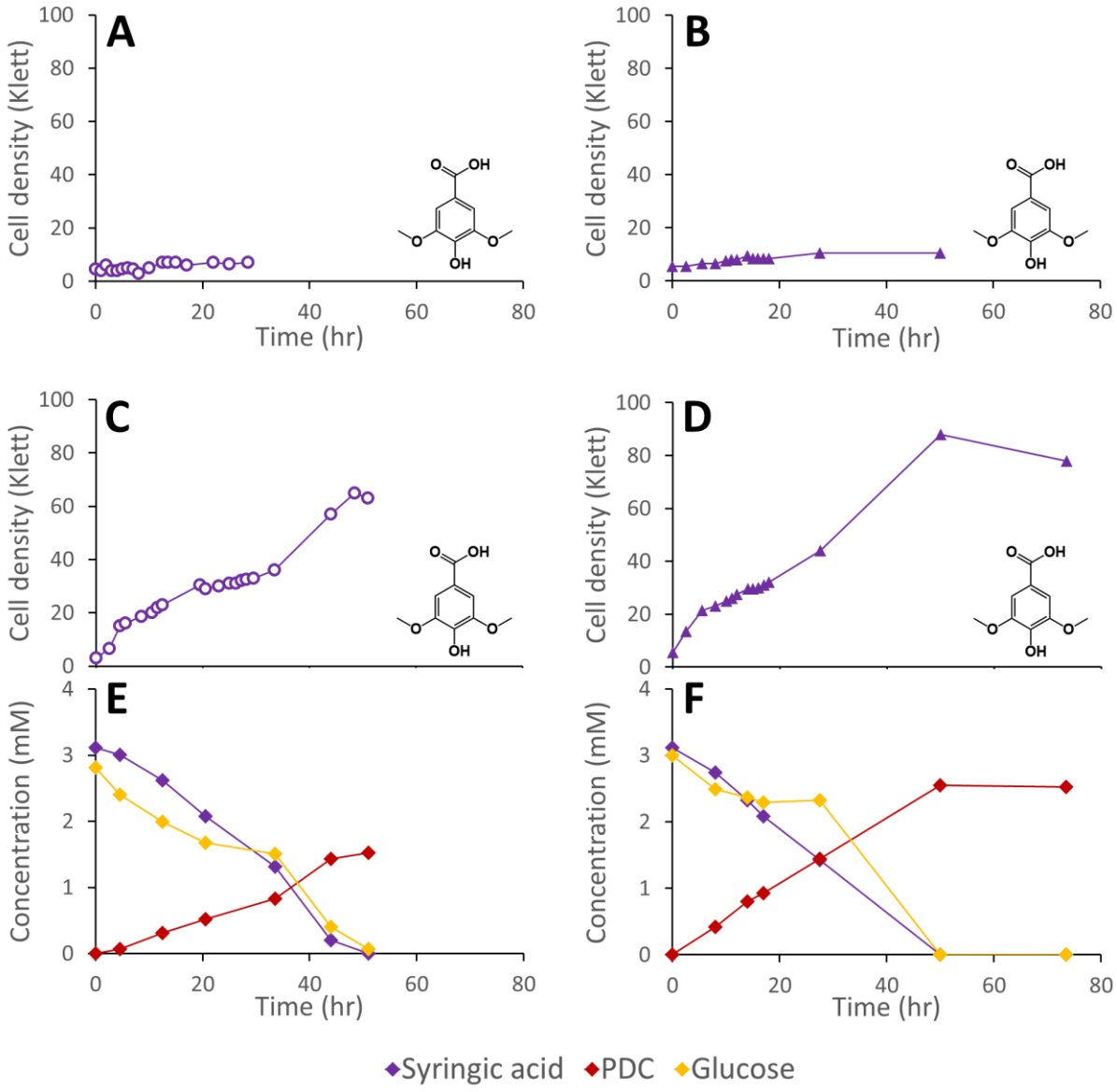


Figure 2.5. Cell densities and extracellular metabolite concentrations of *N. aromaticivorans* strains 12444*ΔdesCD* (left hand side panels) and 12444*ΔligIΔdesCD* (right hand side panels) grown on 3 mM syringic acid (panels A and B) or a combination of 3 mM syringic acid and 3 mM glucose (panels C to F).

2.3.4 Construction of an N. aromaticivorans mutant that accumulates PDC from S, G, and H aromatics

Based on the observations with strains 12444*ΔligI* and 12444*ΔdesCD*, I hypothesized that a mutant missing *ligI* and *desCD* would be able to produce a higher yield of PDC from S aromatics. We generated this strain (12444*ΔligIΔdesCD*) and found that when it was cultured in minimal media with 3 mM syringic acid as the sole carbon source, it did not grow, as expected from previously presented data (Figure 2.5B). When glucose was added to the growth media, strain 12444*ΔligIΔdesCD* grew (Figure 2.5D), glucose and syringic acid were removed from the media, and PDC accumulated (Figure 2.5F). Indeed, the PDC yield of 12444*ΔligIΔdesCD* (66% ± 13%), was higher than that of 12444*ΔdesCD* (49% ± 0.9%) (Table 2.1).

PDC production from syringaldehyde by strain 12444*ΔligIΔdesCD* was also tested. When this strain was grown on 1 mM syringaldehyde plus 3 mM glucose (Figure 2.S1E), syringaldehyde disappeared from the growth media, syringic acid was transiently detected, and PDC accumulated with a 90% (± 7%) yield (Table 2.1).

2.3.5 The fate of unconverted aromatic carbon

Since PDC yields were typically less than 100%, it is possible that some aromatic compounds are degraded via alternative routes not blocked by the *ΔligI* and *ΔdesCD* mutations, and therefore, a fraction of aromatics may be still used as carbon and energy sources for growth in strain 12444*ΔligIΔdesCD*. To evaluate this hypothesis, I compared cell yields in 12444*ΔligIΔdesCD* cultures grown on either 3 mM glucose or 3 mM glucose plus 3 mM protocatechuic acid. The cultures grown on glucose reached a final density of 165 (± 1) Klett units and no glucose or PDC was detected in the culture media (Table 2.2). The cultures receiving glucose plus protocatechuic acid reached a final cell density of 202 (± 2) Klett units (Table 2.2). In these cultures, all glucose was consumed and 0.2 mM (± 0.03) protocatechuic acid remained in the growth media (Table 2.2). The calculated yield of PDC based on the consumed protocatechuic acid was 85% (± 1%) (Table 2.2). Since in both conditions the same amount of glucose

was provided, the higher cell density observed in the cultures containing glucose plus protocatechuic acid can be explained by the use of a fraction of protocatechuic as a carbon and energy source for cell growth, presumably via a less efficient alternative pathway. The absence of PDC in the cultures containing only glucose shows that strain 12444*ΔligIΔdesCD* does not produce PDC from glucose.

Table 2.2. Comparison of c ell densities and extracellular concentrations at stationary phase of *N. aromaticivorans* strain 12444*ΔligIΔdesCD* cultures grown on glucose or glucose plus protocatechuic acid. Data shown represents the average of 3 biological replicates. Error bars represent one standard deviation.

	Glucose	Glucose + protocatechuic acid
Maximim cell density (Klett)	165.3 (± 0.58)	201.7 (± 2.08)
Metabolites concentration immediately after inoculation		
Glucose (mM)	3.1 (± 0.02)	3.1 (± 0.04)
Protocatechuic acid (mM)	0.0	2.9 (± 0.02)
PDC (mM)	0.0	0.0
Metabolites concentration at stationary phase		
Glucose (mM)	0.0	0.0
Protocatechuic acid (mM)	0.0	0.2 (± 0.03)
PDC (mM)	0.0	2.3 (± 0.04)
PDC yield (%)	0.0	85 (± 1.10)

2.3.6 Production of PDC from chemically depolymerized lignin

Lignocellulosic biomass pretreatment and chemical depolymerization of lignin typically result in heterogeneous mixtures of aromatics with variable molar yields of monomers recovered (6, 7). Based on the above results, a strain lacking both LigI and DesCD activity might also be able to simultaneously convert all three classes (S, G, and H) of plant-derived aromatics into PDC. To test the ability of strain 12444*ΔligIΔdesCD* to produce PDC simultaneously from multiple S, G, and H aromatic compounds, I cultured it in glucose-containing media supplemented with the products of depolymerized poplar lignin (4), which contained a mixture of S, G, and H aromatic compounds (Figure 2.6). For comparison, strain

12444*Δ1879* was cultured in the same media. In addition, a flask containing the same media without cells was incubated as an abiotic control. A large proportion of the aromatic compounds present in this type of depolymerized lignin are S and G type diketones (4) and no information has been previously reported about the ability of bacteria to degrade them. Thus, in the experiments below we also tested for metabolism of the S and G diketones and their potential conversion into PDC.

In the abiotic control, none of the aromatic compounds were transformed after 77.5 h of incubation (Figure 2.S3D, Figure 2.S4B, Figure 2.S5). In the inoculated cultures, both strains grew, and, in both cases, all the major aromatic compounds (G-diketone, S-diketone, *p*-hydroxybenzoic acid, vanillin, vanillic acid, syringaldehyde, and syringic acid) disappeared from the growth media (Figure 2.6 and Figure 2.S3 panels B and C). PDC only accumulated in the 12444*ΔligIΔdesCD* cultures, reaching a concentration of 0.49 mM (±0.02), which corresponds to a molar yield of 59% (±1.9%) assuming that all of the above aromatics were used as a source of this compound (Figure 2.6).

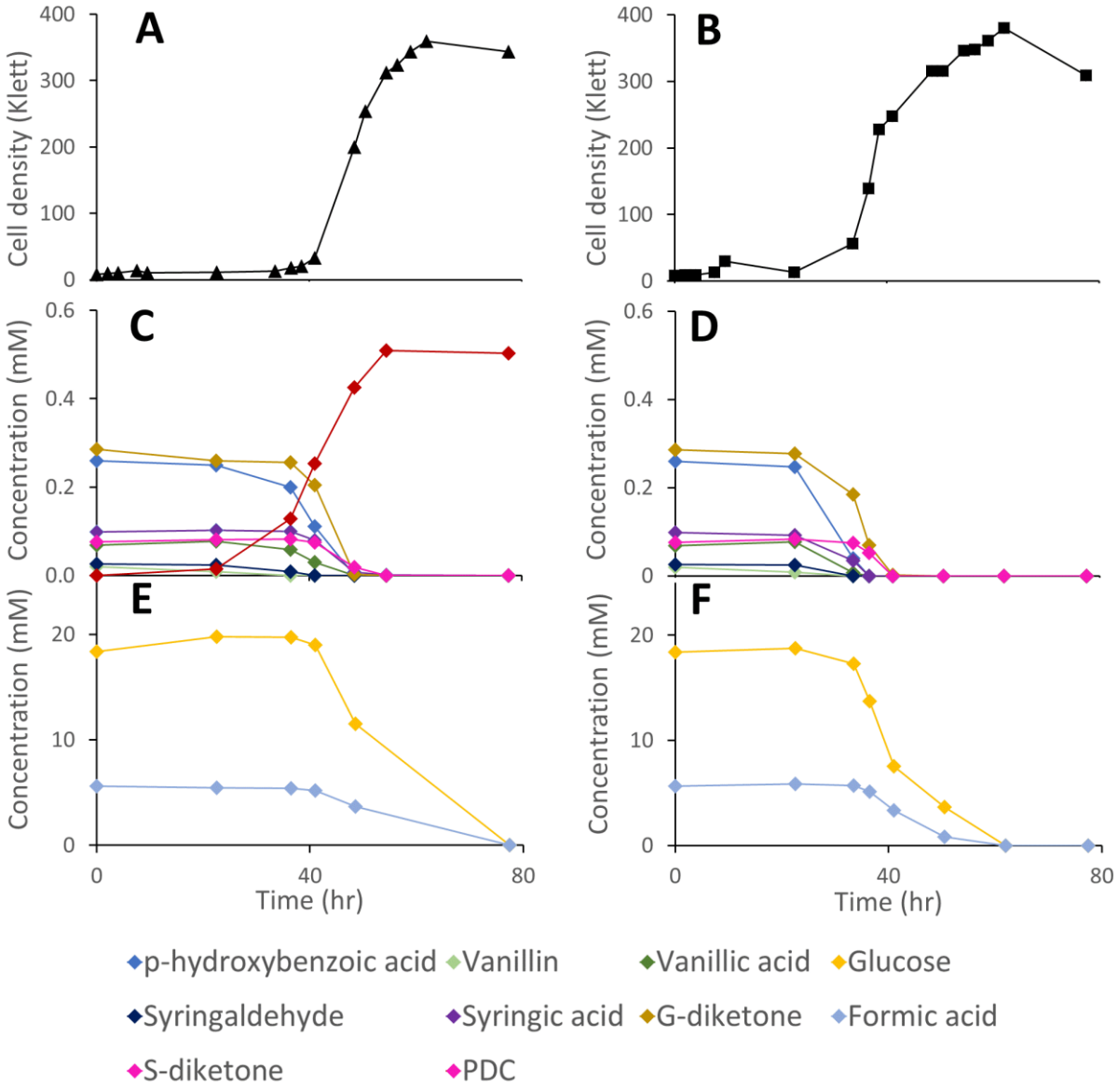


Figure 2.6. Cell density (panels A and B) and extracellular metabolite concentrations (panels C to F) of representative cultures of *N. aromaticivorans* strains 12444*ΔligIΔdesCD* (left hand side panels) and 12444*Δ1879* (right hand side panels) grown on formic acid induced depolymerized poplar lignin supplemented with glucose. Panels C and D show extracellular concentrations of lignin derived aromatic compounds and PDC as a product, and panels E and F show extracellular concentrations of glucose and formic acid. Formic acid is present in the low molecular weight products of chemical depolymerization, whereas glucose was added to enhance bacterial cell growth.

Gel permeation chromatography (GPC) was performed to determine the presence of, and evaluate changes in, oligomeric lignin fragments found in these depolymerized lignin samples (Figures 2.S4 and 2.S5). This analysis showed presence of compounds with a wide range of molecular weights (Mw), which we grouped in 2 ranges (See Materials and Methods). Based on the analysis of standards, compounds eluting between 17.0 and 22.7 min corresponded to oligomeric lignin fragments, while compounds eluting after 22.7 min are dimeric and monomeric compounds. An abiotic control showed that during 78 hours of incubation there was an observable increase in low Mw oligomers, likely from reactive monomer condensation, that resulted in an average Mw decrease from 857 to 722 Da (Figure 2.S5). Both microbial cultures showed a decrease in the dimeric and monomeric compounds (signals eluting after 22.7 min) compared to the abiotic control sample. As with the sample before incubation, both microbial cultures showed the decrease in oligomer Mw attributed to reactive monomer condensation, but not as much as in the abiotic control (Figure 2.S5). Accumulation of PDC in experiments with 12444*ΔligIΔdesCD* was observable by a peak at 23.55 min (Figure 2.S4D), corresponding to that of the PDC standard, which was not observed in the abiotic control or the experiment with the parent strain 12444*Δ1879*.

While the above data suggest that 12444*ΔligIΔdesCD* is able to convert the G, S, and H units found in depolymerized lignin into PDC, the lack of stoichiometric conversion into PDC makes it difficult to assess how well each substrate is metabolized and converted into this product. To specifically test PDC production from the S and G aromatic diketones, I grew cultures of *N. aromaticivorans* strain 12444*ΔligIΔdesCD* on minimum media supplemented with chemically synthesized S-diketone plus glucose or G-diketone plus glucose (see Supplementary Information for aromatic diketone synthesis procedures). In the cultures containing S-diketone, 12444*ΔligIΔdesCD* grew, glucose and the aromatic diketone disappeared from the growth media, and PDC accumulated with a yield of 22.0% (± 0.7%) (Table 2.1, Figure 2.7, panels A and C). On the other hand, in the cultures supplemented with G-diketone (which contained small amounts of vanillic acid and vanillin as impurities from the synthesis method) both glucose and the aromatic substrates disappeared and PDC accumulated (Figure 2.7, panels B and D),

with a nearly stoichiometric yield (107% ± 1.6%, Table 2.1) for G-diketone (assuming a 100% yield from the vanillic acid and vanillin impurities). From this, I conclude that strain 12444*ΔligIΔdesCD* metabolizes these S and G diketones, using pathways that are also involved in degradation of the S, G, and H aromatics normally found in lignin, and converts them into PDC, albeit at different efficiencies.

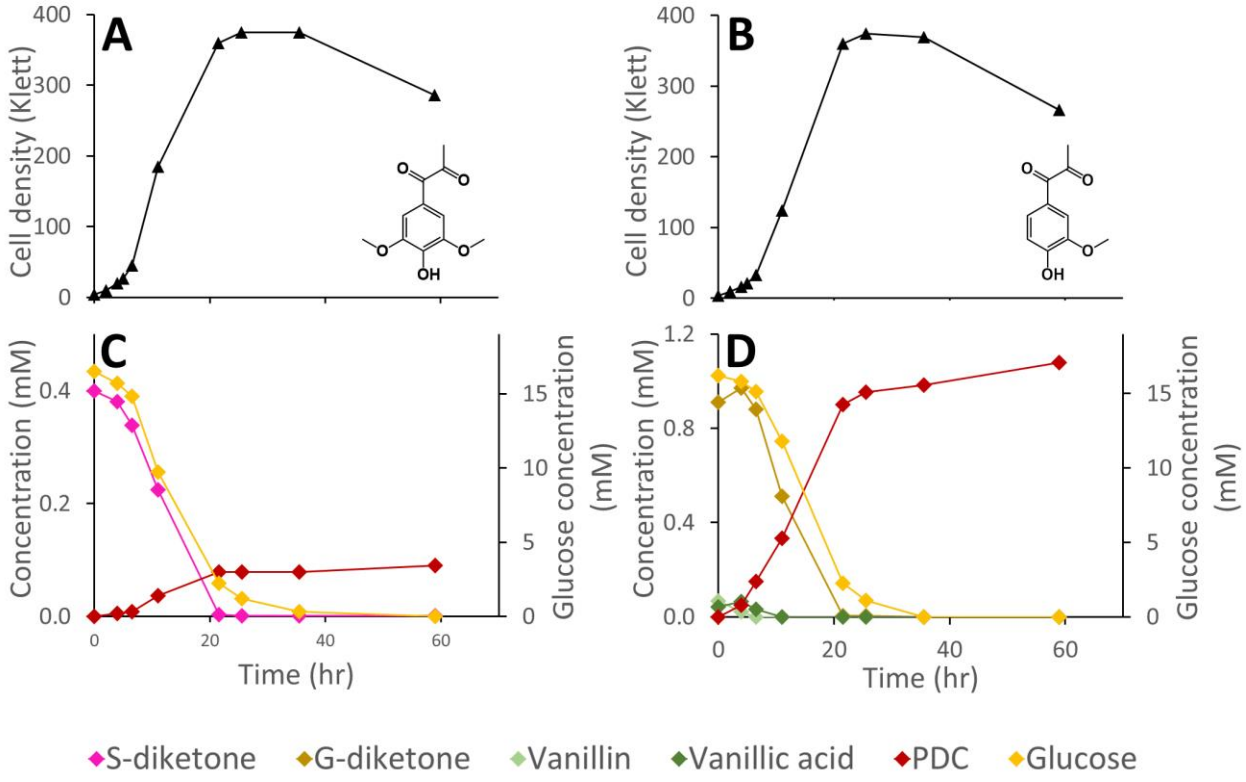


Figure 2.7. Cell density and extracellular metabolite concentrations of representative *N. aromaticivorans* strain 12444*ΔligIΔdesCD* cultures grown on minimal media supplemented with S-diketone and glucose (panels A and C) or G-diketone and glucose (panels B and D).

2.3.7 Production of PDC from vanillic acid and vanillin in a fed-batch reactor

To study the feasibility of PDC production by strain 12444*ΔligIΔdesCD* at titers higher than those observed in batch cultures, I cultured the mutant strain in a pH-controlled fed-batch reactor in which a concentrated solution containing vanillic acid, vanillin, and glucose was intermittently fed. In this experiment, a maximum concentration of 26.7 mM (4.9 g/L) of PDC was reached after 48 hours of incubation (Figure 2.S6), which represents a more than 8 times higher concentration than observed in the batch experiments reported here. As the reaction progressed, an accumulation of glucose, vanillic, and protocatechuic acid was observed. Further experiments are necessary to optimize bioreactor conditions to increase PDC titer and minimize substrate and intermediate metabolite accumulation.

2.4 Discussion

The economic and environmental viability of producing fuels and chemicals from lignocellulose is tightly connected to the efficiency of its utilization. New methods are needed to efficiently utilize the recalcitrant aromatic fractions, such as lignin (31). Multiple chemical approaches have shown promising results for breaking down the complex lignin polymer into small molecule aromatic units (6, 7). However, the heterogeneous nature of the depolymerization products obtained pose challenges for further upgrading to valuable products (32). One successful strategy to address the chemical heterogeneity is to funnel the mixture of compounds through convergent aromatic biodegradation pathways into one valuable product by interruption and/or redirection of the metabolic flow to a pathway intermediate (19, 22, 23). These studies suggest that a mixed approach that integrates chemical and biological tools has the potential to be an effective strategy to maximize the yield of desired products from lignin transformation. Some of the major challenges in biological funneling are the transformation of unnatural products resulting from chemical depolymerization for which microbial metabolic capabilities are unknown, the maximization of target product yield while minimizing the accumulation of undesired intermediates or end products, and

the identification of industrially useful target molecules that could most readily be produced from lignin components via known metabolic pathways (15).

The present study addresses each of these issues using mutant strains of *N. aromaticivorans* DSM12444, a microbe naturally capable of degrading S, G, and H type aromatic compounds, as a well as lignin derived aromatic dimers (25, 33). I chose *N. aromaticivorans* DSM12444 due to its known or predicted ability to grow in the presence of multiple aromatic compounds, its suitability for genetic analysis and modification, its ability to co-metabolize aromatics in the presence of other organic compounds (such as sugars, which are another plentiful product of plant biomass degradation), and the potential to produce single valuable products using defined mutants.

The efficiency of carbon recovery in valuable compounds depends on factors such as the target product, the minimization of undesired metabolic byproducts, and number or amount of substrates being metabolized by the bacterium. Products derived from metabolic intermediates in the upper aromatic catabolic pathways of bacteria like *N. aromaticivorans* DSM12444 should yield higher carbon recovery than products derived from lower pathways, where more carbon may have already been lost during degradation. We selected PDC as the target product for this study because, in addition to its proven potential as a polyester precursor (26), it is the earliest compound in which the degradation pathways for S, G, and H aromatic compounds were predicted to converge in defined *N. aromaticivorans* mutants (Figure 2.1).

The observation of PDC accumulation when strain 12444*Δ1879* was grown on syringic acid (28%; Figure 2.4C) was surprising, since I had predicted that the majority of the syringic acid would follow the 3-MGA, CHMOD, OMA pathway (Figure 2.1) when the pathway was not altered by mutation. Furthermore, we had predicted that any PDC formed during syringic acid degradation in this strain would be oxidized by LigI to OMA (Figure 2.1). The sequential increase in PDC yield in strains 12444*ΔdesCD* (49%; Figure 2.5E) and 12444*ΔligIΔdesCD* (66%; Figure 2.5F) confirms the participation of DesC, DesD, and LigI in the degradation of S type aromatics in *N. aromaticivorans* and suggests that a large

fraction of the syringic acid is naturally channeled through PDC. Since PDC does not accumulate in 12444*Δ1879* cultures grown on the products from chemically depolymerized lignin (Figure 2.6D) we offer two alternative hypotheses that would need to be tested in the future. First, it is possible that G or H substrates regulate expression of LigI in *N. aromaticivorans*. Thus, LigI would be poorly or not expressed when S type aromatics are the sole carbon source, allowing for some PDC accumulation by strain 12444*Δ1879* grown on syringic acid. On the other hand, LigI would be expressed when 12444*Δ1879* is grown on the mixtures of S, G, and H aromatics present in depolymerized lignin, preventing PDC accumulation. Alternatively, since it is not known whether CHMOD transformation to PDC is abiotic or enzymatic, it may be possible that CHMOD is secreted into the growth media where it undergoes spontaneous cyclization, resulting in extracellular PDC accumulation. Higher PDC yields by 12444*ΔdesCD* and 12444*ΔligIΔdesCD* could then be explained by increased CHMOD secretion when the aromatic degradation pathways are blocked.

We observe nearly stoichiometric conversion of vanillin and G-diketone into PDC, without extracellular accumulation of other aromatics. However, conversion of *p*-coumaric acid, *p*-hydroxybenzaldehyde, *p*-hydroxybenzoic acid, ferulic acid, vanillic acid, syringaldehyde, syringic acid, and S-diketone to PDC was found to have somewhat lower efficiencies (Table 2.1). The non-stoichiometric conversion of these aromatic compounds into PDC by *N. aromaticivorans* is not due to accumulation of intermediate metabolites such as syringic acid, vanillic acid, p-hydroxybenzoic acid and protocatechuic acid, since they only accumulated transiently. Instead, the lower conversion efficiencies could potentially be explained by the presence of alternative, less efficient, and poorly studied pathways for the degradation of those compounds. For instance, the *N. aromaticivorans* genome contains multiple genes annotated as aromatic ring cleavage dioxygenases for which specificity has not yet been established (34). The presence of a catechol degradation pathway in *N. aromaticivorans* that uses 2,3-cleavage of the aromatic ring has been suggested as a possible alternative pathway for protocatechuic acid degradation (2). Such alternative non-specific reaction of a catechol dioxygenase could explain the observed lower efficiencies in the

transformation of some G and H aromatics to PDC. This hypothesis is supported by the increased cell density observed in cultures of strain 12444*ΔligIΔdesCD* grown in media containing glucose plus protocatechuic acid compared to cultures only fed glucose (Table 2.2). Another enzyme with low substrate specificity appears to be the *O*-demethylase LigM, included in our model as catalyzing the demethylation of vanillic acid (Figure 2.1). In *Sphingobium sp*. SYK-6, LigM is also predicted to catalyze *O*-demethylation of 3-MGA to gallate (1), which is then proposed to be oxidized to OMA by either LigAB, a dioxygenase with broad specificity (Figure 1), or DesB, an enzyme not present in *N. aromaticivorans*. Although this route for degradation of S aromatics is not predicted to be important in *N. aromaticivorans* (2), LigM activity with 3-MGA and LigAB activity with gallate could contribute to lowering the efficiency of PDC formation from S aromatics by bypassing the blockage in S aromatic degradation intended with the *desCD* mutation. Thus, future identification and analysis of additional pathways involved in aromatic metabolism by *N. aromaticivorans* DSM12444 could provide useful information for further increasing the yield of PDC or other target chemicals by preventing aromatic substrates from being degraded by alternative routes.

Fed-batch experiments in a pH-controlled bioreactor showed an increase of up to 8.7 times in PDC titers with respect to titers obtained in batch experiments. These results show a promising potential for production of PDC from aromatic compounds. However, in this experiment, a progressive accumulation of aromatic substrates and glucose was observed. Additional research will be necessary to optimize culture conditions.

The efficiency of lignin conversion to a desired product is also impacted by the nature of the aromatic compounds that result from chemical lignin depolymerization, which may be different from natural products of environmental lignin depolymerization. Therefore, the existence of microbial pathways to metabolize these products could be crucial to increase product recovery. For example, formic-acid-induced depolymerization of oxidized lignin produces a high proportion of aromatic diketones (4), compounds that have also been reported to be present in lignocellulose dilute acid hydrolysates (35).

Biological sources of these or structurally related compounds have not been reported, so it was previously unknown whether *N. aromaticivorans* DSM12444 could metabolize these products or convert them into PDC or other valuable materials. In this study, we found that *N. aromaticivorans* can convert both S- and G-type diketones into PDC, indicating that they are also degraded via the predicted aromatic degradation pathways (Figure 2.1). However, the upper pathway enzymes that transform the diketones to known intermediates in the aromatic degradation pathways remain unknown.

Finally, chemically depolymerized lignin yields a variety of higher molecular weight lignin derived products in addition to monomeric units (4). Sphingomonad bacteria, such as *N. aromaticivorans* DSM12444, are known or predicted to be capable of breaking most of the linkages found between aromatic subunits in natural lignin in defined ways that yield predictable mono-aromatic products that can be further metabolized (1, 36), *N. aromaticivorans*, specifically, is known to be capable of degrading model aromatic dimers containing β-aryl-ether bonds (25) and its genome contains homologs of genes that code for the degradation of other aromatic dimers in *Sphingobium sp*. SYK-6.1 This is an unexplored, but potentially important aspect of employing *N. aromaticivorans* as a platform microbe for valorization of mixtures of low molecular weight aromatic compounds generated from chemical depolymerization of lignin.

2.5 Materials and Methods

Bacterial strains, growth media and culturing conditions. A variant of *N. aromaticivorans* DSM12444 (strain 12444*Δ1879*) that lacks the gene Saro_1879 (putative *sacB*; SARO_RS09410 in the recently reannotated genome in NCBI) (25) was used as a parent strain to create the deletion mutant strains 12444*ΔligI* (lacks gene Saro_2819; SARO_RS14300), 12444*ΔdesCD* (lacks the genes Saro_2864 and Saro_2865; SARO_RS14525 and SARO_RS14530), and 12444*ΔligIΔdesCD* (lacks genes Saro_2819, Saro_2864, and Saro_2865). All genetic modifications used a variant of the plasmid pk18mobsacB (37),

which contains *sacB* and a kanamycin resistance gene. A detailed procedure for constructing strains with gene deletions is contained in the Supplementary Information. All bacterial strains and plasmids used in this study are listed in Table 2.S1. Primers used in the construction of the mutant strains are listed in Table 2.S2.

Escherichia coli cultures were grown in LB medium containing 50 μg/mL kanamycin at 37 °C. *N. aromaticivorans* cultures were grown in SISnc-V0 media supplemented with the indicated carbon source at 30 °C. SISnc-V0 media is a modification of Sistrom's minimal media (38) in which succinate, L-glutamate, L-aspartate, and vitamins were omitted. For routine culture and storage, the growth media was supplemented with 1 g/L glucose. For gene modifications, the growth media was supplemented with 1 g/L glucose and 50 μg/mL kanamycin, or 1 g/L glucose and 10% sucrose.

***N. aromaticivorans* growth experiments.** Cell cultures were grown overnight in SISnc-V0 media supplemented with 1 g/L glucose, then diluted 1:1 with fresh SISnc-V0 containing 1 g/L glucose and incubated for one hour. Then, 2 ml of the growing culture was spun for 5 min at 5000 rpm, and the cell pellets were resuspended into fresh SISnc-V0 media containing no added carbon source. The resuspended cells were diluted 1:100 into SISnc-V0 media supplemented with the indicated carbon source, then shaken at 200 rpm and 30°C. Cell growth was monitored by measuring cell density using a Klett-Summerson photoelectric colorimeter with a red filter. For *N. aromaticivorans*, 1 Klett unit (KU) is equal to ~8x106 cfu/ml (25). Culture samples (1 mL) were collected at various time points, spun for 5 min at 5000 rpm and 4 °C, and the supernatants were filtered through 0.22 μM nylon syringe tip filters (Fisher Scientific), then stored at -20 °C. Each culture was grown at least three times and the data shown corresponds to the results obtained from a representative culture. Conversion efficiency of aromatics to product was calculated by dividing the total amount of product by the total amount of aromatic substrates consumed. Conversion efficiencies reported correspond to the average and standard deviation of the efficiencies calculated for all replicates.

Production of PDC in a fed-batch bioreactor. A 250-ml bioreactor (Infors, model Multifors 2) containing 130 ml minimum media with 12 mM glucose was inoculated with 2 ml of *N. aromaticivorans* strain 12444*ΔligIΔdesCD* culture that had been pre-grown overnight with glucose. After 7.5 hrs of batch incubation, the bioreactor was intermittently fed media containing 226 mM vanillic acid, 34 mM vanillin, 550 mM glucose, 15 g/L ammonium sulfate, and 5% (v/v) DMSO. Culture pH was controlled by the addition of 1M KOH when needed, to maintain pH 7. Temperature was maintained at 30°C and the stirrer speed between 250 and 320 rpm. Air was used to deliver oxygen at a flow rate of 1 L/min. During 50 hours of operation, a total of 29 ml of feed solution was added.

Analysis of extracellular metabolites. Metabolite identification was performed by gas chromatography-mass spectrometry (GC-MS) of filtered culture supernatants. Sample aliquots (150 µL) were combined with 70 µL of 1 mM m-coumaric acid in water (internal standard), acidified with HCl to pH < 2, and ethyl acetate extracted (3x500 µL). The three ethyl acetate extractions were combined, dried under a stream of N2 at 40 °C, and derivatized by the addition of 150 µL of pyridine and 150 µL of N,O-Bis(trimethylsilyl)trifluoro-acetamide with trimethylchlorosilane (99:1, w/w, Sigma) and incubated at 70 °C for 45 min. The derivatized samples were analyzed on an Agilent GC-MS (GC model 7890A, MS Model 5975C) equipped with a (5% phenyl)-methylpolysiloxane capillary column (Agilent model HP-5MS). The injection port temperature was held at 280 °C and the oven temperature program was held at 80 °C for 1 min, then ramped at 10°C/min to 220 °C, held for 2 min, ramped at 20°C/min to 310 °C, and held for 6 min. The MS used an electron impact (EI) ion source (70 eV) and a single quadrupole mass selection scanning at 2.5 Hz, from 50 to 650 m/z. The data was analyzed with Agilent MassHunter software suite, using m-coumaric acid as internal standard.

Quantitative analysis of glucose and formic acid were performed on an Agilent 1260 infinity HPLC equipped with a refractive index detector (HPLC-RID) (Agilent Technologies, Inc., Palo Alto, CA) and an Aminex HPX-87H with Cation-H guard column (BioRad, Inc. Hercules, CA). The mobile phase was 0.02 N sulfuric acid at a flow rate of 0.5 ml/min.

Quantitative analysis of aromatic compounds and PDC were performed on a Shimadzu triple quadrupole liquid chromatography mass spectrometer (LC-MS) (Nexera XR HPLC-8045 MS/MS). The mobile phase was a binary gradient consisting of solvent A (water) and solvent B (0.1% formic acid in a 2:1 mixture of acetonitrile and methanol, v/v). The stationary phase was a Phemonenex Kinetex F5 column (2.6 µm pore size, 2.1mm ID, 150mm length, P/N: H18-105937). All compounds were detected by multiple-reaction-monitoring (MRM) and quantified using the strongest MRM transition (Table 2.S3).

^{1}H-NMR analysis. Nuclear magnetic resonance (NMR) spectroscopy was performed on a Bruker Biospin (Billerica, MA) Avance 500 MHz spectrometer equipped with a 5-mm quadruple-resonance ^{1}H/^{31}P/^{13}C/^{15}N QCI gradient cryoprobe with inverse geometry (proton coils closest to the sample). Samples were prepared as ~1 mg in 600 µL acetone-d_6.

Gel permeation chromatography (GPC) analysis. Analytical GPC was performed on a Shimadzu LC20 with a photodiode array detector (SPD-M20A). Separation was performed using a PSS PolarSil linear S column (7.8 mm x 30 cm, 5µm) at 35°C. The mobile phase was 5.2 mM sodium phosphate buffer at pH 8, pumped at 0.5 mL/min, 60-min run time. The molecular weight distribution was calibrated at ʎ=254 nm using PDC (184 g/mol, 23.55 min) and Poly(styrene sulfonate) sodium salts, Mp (retention time): 976 kDa (13.20 min), 258 kDa (13.55 min), 65.4 kDa (14.78 min), 47 kDa (16.07 min), 9.74 kDa (17.96 min), 4.21 kDa (19.433 min), and 2.18 kDa (20.35 min) from the PSS-psskit (Polymer Standards Service-USA, Inc, Amherst, MA, USA). Monomer standards were also ran to establish the lower threshold of the column and confirmed that some of them interact with the stationary phase in the alkaline-water mobile phase, these were: rosmarinic acid (360 g/mol, 21.49 min), ferulic acid (194 g/mol, 26.63 min), p-coumaric acid (164 g/mol, 24.96 min), vanillic acid (168 g/mol, 24.22 min), p-hydroxybenzoic acid (138 g/mol, 24.87 min), and guaiacol (124 g/mol, 39.82 min). Compounds eluting from 17.0–22.7 min correspond to oligomeric lignin, while compounds eluting after 22.7 min, correspond to dimeric and monomeric compounds. It should be noted that no Mw values were calculated for peaks detected after 22.7 min, as they were outside the calibration range of the GPC column. In the control samples there were

strong monomer signals eluting after 26.0 min, especially a pair of signals at ~30 min with an absorption band at 375 nm. Most of these monomer signals were not present, or were much weaker, in the inoculated samples after 78 hours of incubation.

Preparation of media containing depolymerized lignin products. Lignin was isolated by acid precipitation from pretreatment liquor of poplar biomass that had been pretreated by the copper alkaline hydrogen peroxide method (AHP-Cu) (39-41). The lignin was depolymerized using an adaptation of the oxidative methods described previously (4). Depolymerization products were recovered by ethyl acetate extraction, followed by solvent evaporation. This material was re-dissolved in water while adjusting the pH to 7.0 to favor solubilization of aromatic compounds. Consistent with reported products of oxidative depolymerization,4 quantitative HPLC-MS analysis showed concentrations of 1 mM G-diketone, 0.35 mM S-diketone, 0.37 mM syringic acid, 0.12 mM syringaldehyde, 0.44 mM vanillic acid, 0.1 mM vanillin, and 0.93 mM *p*-hydroxybenzoic acid in the final aqueous solution. For experiments with *N. aromaticivorans*, aliquots of this solution (25 mL) were mixed with concentrated (5X) SISnc-V0 media containing 1 g/L glucose (20 mL) and water (55 mL).

Chemicals. Syringic acid, syringaldehyde, ferulic acid, vanillic acid, vanillin, *p*-coumaric acid, *p*-hydroxybenzoic acid, *p*-hydroxy-benzaldehyde, and protocatechuate were purchased from Sigma-Aldrich (St. Louis, MO). G- and S- diketones were synthesized according to the methods described in the ESI. PDC was produced by culturing *N. aromaticivorans* 12444*ΔligI* in 1 L of SISnc-V0 media supplemented with 3 mM vanillic acid and 0.5 g/L (2.8 mM) glucose, and purified following a simplified version of published methods (42), obtaining a >97% pure chemical standard for GC-MS and LC-MS quantifications. Specific details of these procedures are detailed in the ESI. The identity of PDC was confirmed by comparing the GC-MS spectrum of TMS derivatives (Figure 2.S7) and the 1H-NMR spectrum (Figure 2.S8) with those reported previously (43).

2.6 Conclusions

A promising path to produce valuable products from the abundant and renewable raw material lignin is to integrate chemical and biological strategies to chemically depolymerize lignin into heterogeneous mixtures of compounds that are then funneled into a single valuable product using microbial catalysts. An ideal microbial catalyst would be capable of simultaneously converting aromatic compounds containing S, G, and H structures, including non-natural compounds generated by chemical depolymerization, into a single compound with high efficiency.

Here, I focused on the microbial production of PDC from aromatic products known to be generated by chemical methods of lignin depolymerization. PDC has been shown to have potential as a precursor for polyesters and there is growing interest in using microbes to generate it from lignin (21, 22). However, the range of lignin-derived aromatic substrates that could be converted into PDC was limited (21, 22). This study expanded the range, as we demonstrated how we could take advantage of *N. aromaticivorans*' natural ability to degrade plant-derived aromatics to create mutant strains that simultaneously funnel a wider range of lignocellulose-derived aromatic compounds (including S, G, and H units, and the non-natural S- and G- diketones) into PDC. It is also important that *N. aromaticivorans* naturally produces PDC via its native metabolic pathways, and therefore, creating PDC-producing strains did not require extensive genetic engineering and optimization steps. Future improvement in PDC yields would require identification of alternative pathways that may be contributing to aromatic degradation in this organism. Ultimately, the information and strategies developed here and in future optimizations of PDC production by *N. aromaticivorans* DSM12444 could potentially be used to develop this and other microbes into platforms for producing a wide range of additional valuable compounds from lignin.

2.7 Acknowledgements

This work was supported by U.S. Department of Energy (DOE) Great Lakes Bioenergy Research Center grants (DOE Office of Science BER DE-FC02-07ER64494 and DE-SC0018409). Additional funding from the Chilean National Commission for Scientific and Technological Research (CONICYT) as a fellowship to Jose M. Perez is also acknowledged. I thank Erik Hegg for providing isolated poplar lignin from AHP-Cu pretreatment and Mick McGee for HPLC-RID analyses.

2.8 Supplementary figures and tables

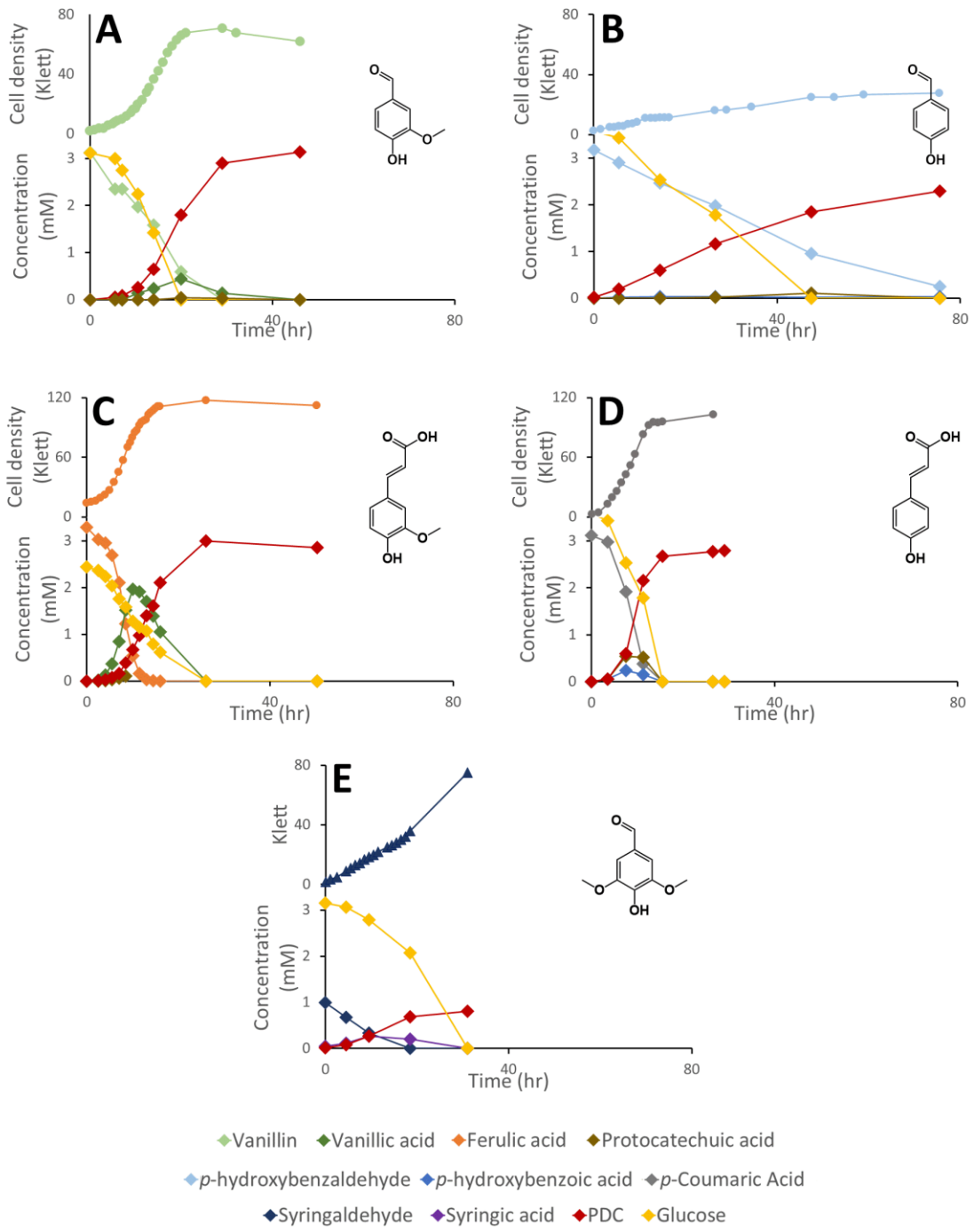


Figure 2.S1. Cell density and extracellular metabolite concentrations of *N. aromaticivorans* strains 12444*ΔligI* (solid circles) or 12444*ΔligIΔdesCD* (solid triangles) grown on a combination of glucose and vanillin (A), *p*-hydroxybenzaldehyde (B), ferulic acid (C), *p*-coumaric acid (D), and syringaldehyde (E).

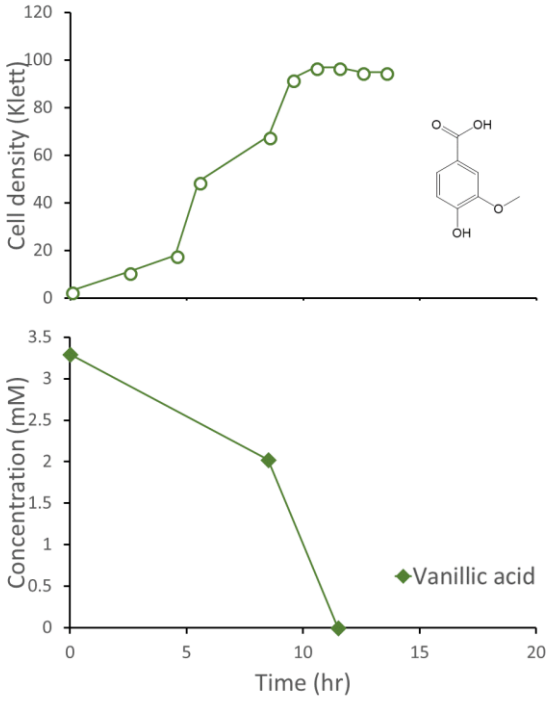


Figure 2.S2. Cell density and extracellular metabolite concentrations of representative *N. aromaticivorans* strain 12444*ΔdesCD* cultured on 3 mM vanillic acid.

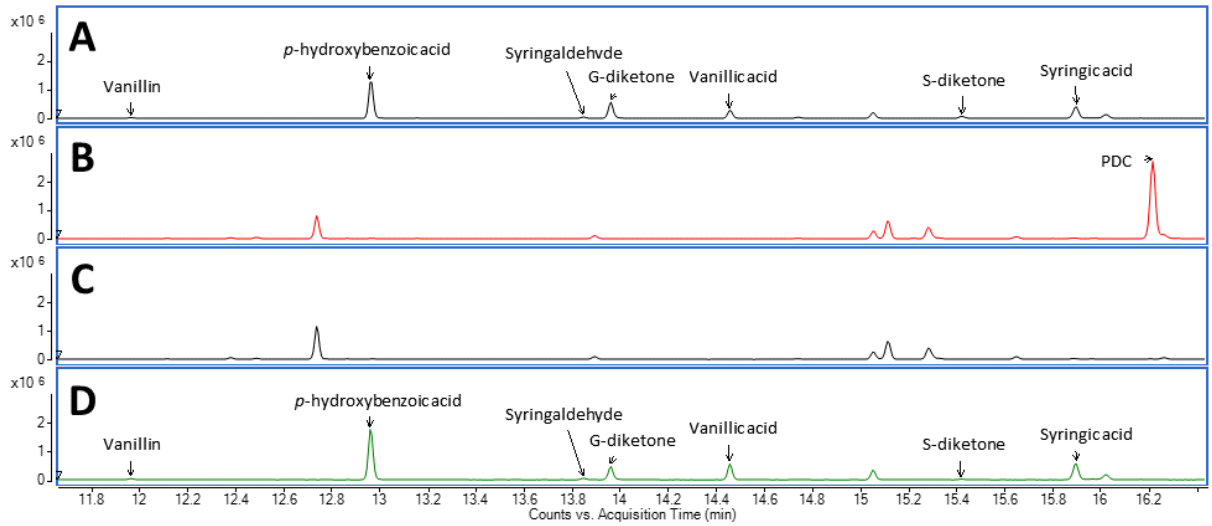


Figure 2.S3. GC-MS peaks of compounds identified in media containing glucose plus the products of formic-acid-induced depolymerization of oxidized poplar lignin; before inoculation (A), after growth of *N. aromaticivorans* strain 12444*ΔligIΔdesCD* (B), after growth of *N. aromaticivorans* strain 12444*Δ1879* (C). Only strain 12444*ΔligIΔdesCD* accumulates PDC in the growth medium. Panel D shows the absence of additional peaks in an abiotic control experiment.

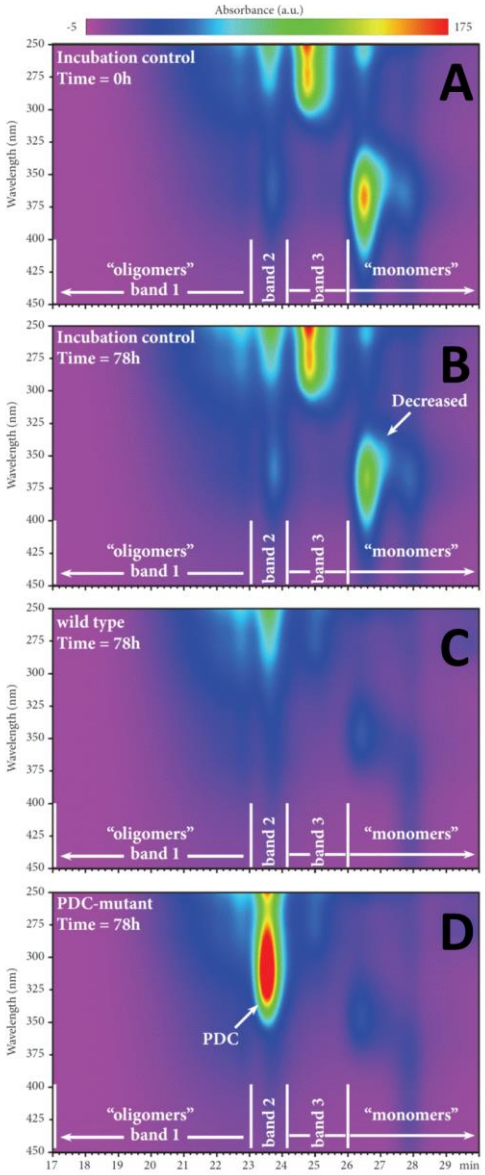


Figure 2.S4. GPC chromatogram of media containing glucose plus the products of formic-acid-induced depolymerization of oxidized poplar lignin; before inoculation (A), abiotic control after 78 hours of incubation (B), after growth of *N. aromaticivorans* strain 12444*Δ1879* (C), after growth of *N. aromaticivorans* strain 12444*ΔligIΔdesCD* (D).

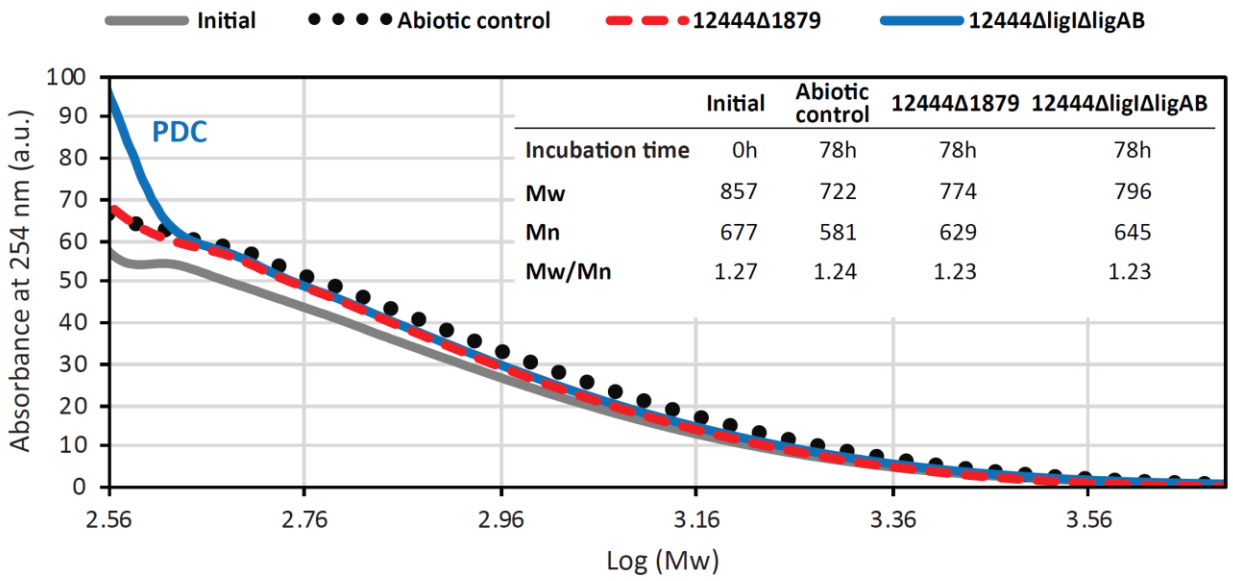


Figure 2.S5. GPC chromatogram of the "oligomers" range at λ=254 of media containing glucose plus the products of formic-acid-induced depolymerization of oxidized poplar lignin. Mw: weight average molecular weight; Mn: number average molecular weight; Mw/Mn: dispersity index.

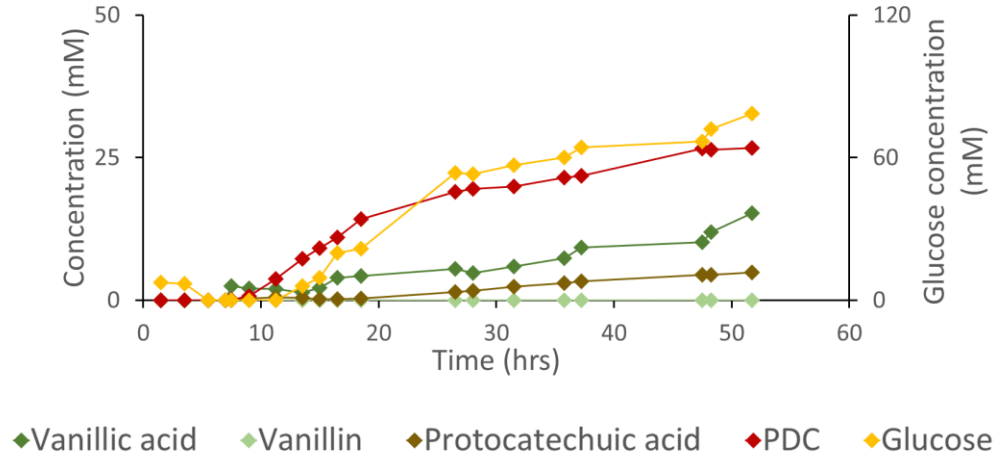


Figure 2.S6. Extracellular metabolite concentrations of a *N. aromaticivorans* strain 12444*ΔligIΔdesCD* culture fed with a concentrated mixture of vanillic acid, vanillin, and glucose. A maximum PDC concentration of 26.7 mM was observed after 48 hours of cultivation.

Table 2.S1. Bacterial strains and plasmids used in this study

Strains	Details	Reference
Novosphingobium aromaticivorans strains		
12444Δ1879	DSM12444 (WT) ΔSaro1879	(25)
12444*ΔligI*	12444Δ1879 ΔSaro2819	This study
12444*ΔdesC/D*	12444Δ1879 ΔSaro2864/5	This study
12444*ΔligIΔdesC/D*	12444Δ1879 ΔSaro2819 ΔSaro2864/5	This study
Escherichia coli strains		
DH 5α	F- Φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rK–, mK+) phoA supE44 λ– thi-1 gyrA96 relA1	Bethesda Research Laboratories (44)
S17-1	recA pro hsdR RP4-2-Tc::Mu-Km::Tn7	(45)
Plasmids		
pK18*mobsacB*	pMB1ori sacB kanR mobT oriT(RP4) lacZα	(46)
pK18msB/ΔSaro2819	pK18*mobsacB* containing genomic regions flanking Saro2819	This study
pK18msB/ΔSaro2864/5	pK18*mobsacB* containing genomic regions flanking Saro2864/5	This study

Table 2.S2. Primers used in this study

Name	Sequence	Obs
Saro2819_Del-R	5'-GCGCCAATCCATACCACGGATTATGCGAATACTACTCCATCCATCAGCTTG-3'	
Saro2819-pK18_Amp-F	5'-**CGATTCATTAATGCAGCTGGCACGACAG**GAGCGAATGGCATGAGTTCACATTCAGC-3'	Region in bold matches sequence in pK18msB
Saro2819_Del-F	5'-GCTGATGGATGGAGTAGTATTCGCATAATCCGTGGTATGGATTGGCGCATG-3'	
Saro2819-pK18_Amp-R	5'-**GTTTCTGCGGACTGGCTTTCTAGATGTTC**CTGCATGGTCTGGTCCTGTTCAAGCAG-3'	Region in bold matches sequence in pK18msB
Saro2864-5_Del_R	5'-GGGTAGTCTGGATCATTCAGACTCGCATGGTGCCGAG-3'	
Saro2864-5-pK18_Amp_F	5'-**CGATTCATTAATGCAGCTGGCACGACAG**CAGGTCGGCTTCAAGGAGGAAGTTCTG-3'	Region in bold matches sequence in pK18msB
Saro2864-5_Del_F	5'-CCATGCGAGTCTGAATGATCCAGACTACCCGCCGTTATC-3'	
Saro2864-5-pK18_Amp_R	5'-**GTTTCTGCGGACTGGCTTTCTAGATGTTC**GACCACTATGCAATGGAATGGAACCTGC-3'	Region in bold matches sequence in pK18msB
Saro2865_Start-SNP_F	5'-GGCATGCTCGGCACCATGCG-3'	
Saro2865_Start-SNP_R2	5'-GCCGTCGACCGCGAGAGCTTG-3'	

Table 2.S3. Multiple reaction module (MRM) conditions for HPLC-MS quantification of compounds used in this study.

Compound	MW (g/mol)	Parent (-) m/z	Transition 1	Transition 2	Transition 3
PDC	184.103	183	183 -> 139.05 CE11	183 -> 111 CE14	183 -> 94.95 CE12
Protocatechuic acid	154.12	153	153 -> 108.95 CE14	153 -> 107.95 CE25	153 -> 90.95 CE27
p -hydroxybenzoic acid	138.12	137	137 -> 93 CE15	137 -> 65 CE30	
Vanillic acid	168.15	167	167 -> 123.05 CE15	167 -> 108 CE21	167 -> 152.05 CE18
p -hydroxybenzaldehyde	122.12	121.2	121.2 -> 92.05 CE26	121.2 -> 93.10 CE22	121.2 -> 41 CE49
Syringic acid	198.17	197	197 -> 121.05 CE18	197 -> 153.10 CE15	197 -> 182.10 CE15
Vanillin	152.15	151	151 -> 136 CE17	151 -> 92 CE22	151 -> 108 CE24
p -Coumaric acid	164.16	163	163 -> 119.05 CE15	163 -> 93 CE31	163 -> 116.95 CE33
Syringaldehyde	182.18	181	181 -> 166.10 CE16	181 -> 151 CE22	181 -> 123 CE28
Ferulic acid	194.19	193	193 -> 149 CE13	193 -> 134 CE16	193 -> 133 CE27
G-diketone	194.19	193	193 -> 178.10 CE20	193 -> 136 CE21	193 -> 107 CE31
S-diketone	224.21	223	223 -> 208.10 CE19	223 -> 193.10 CE20	223 -> 165.10 CE27

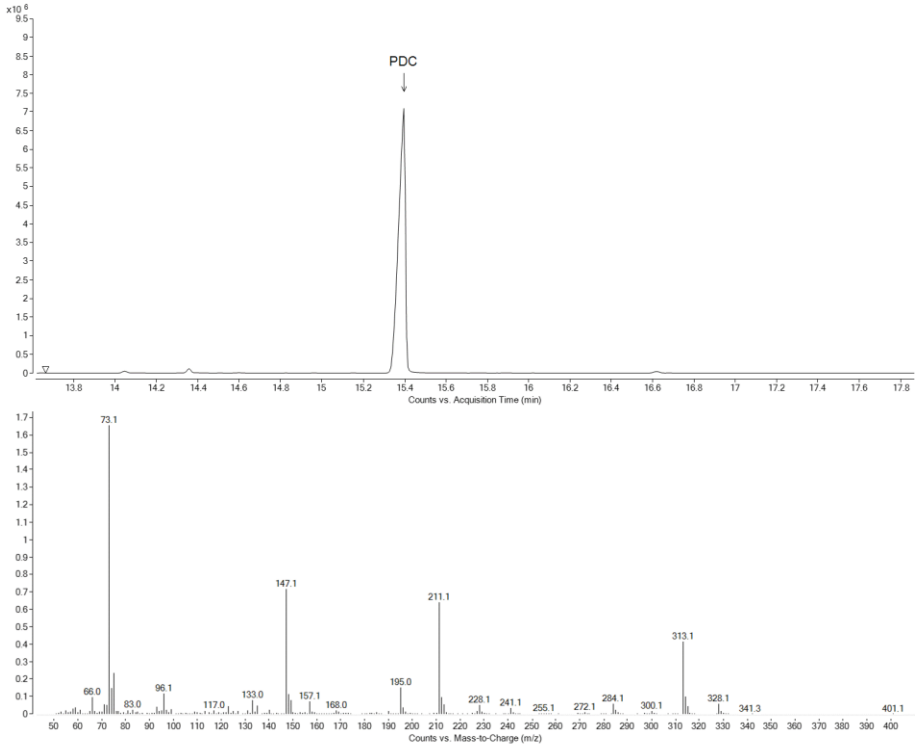


Figure 2.S7. GC-MS chromatogram (A) and MS spectrum (B) of PDC isolated from a culture of PDC-producing *N. aromaticivorans* strain 12444*ΔligI* grown on a mixture of 3 mM vanillic acid and 3 mM glucose.

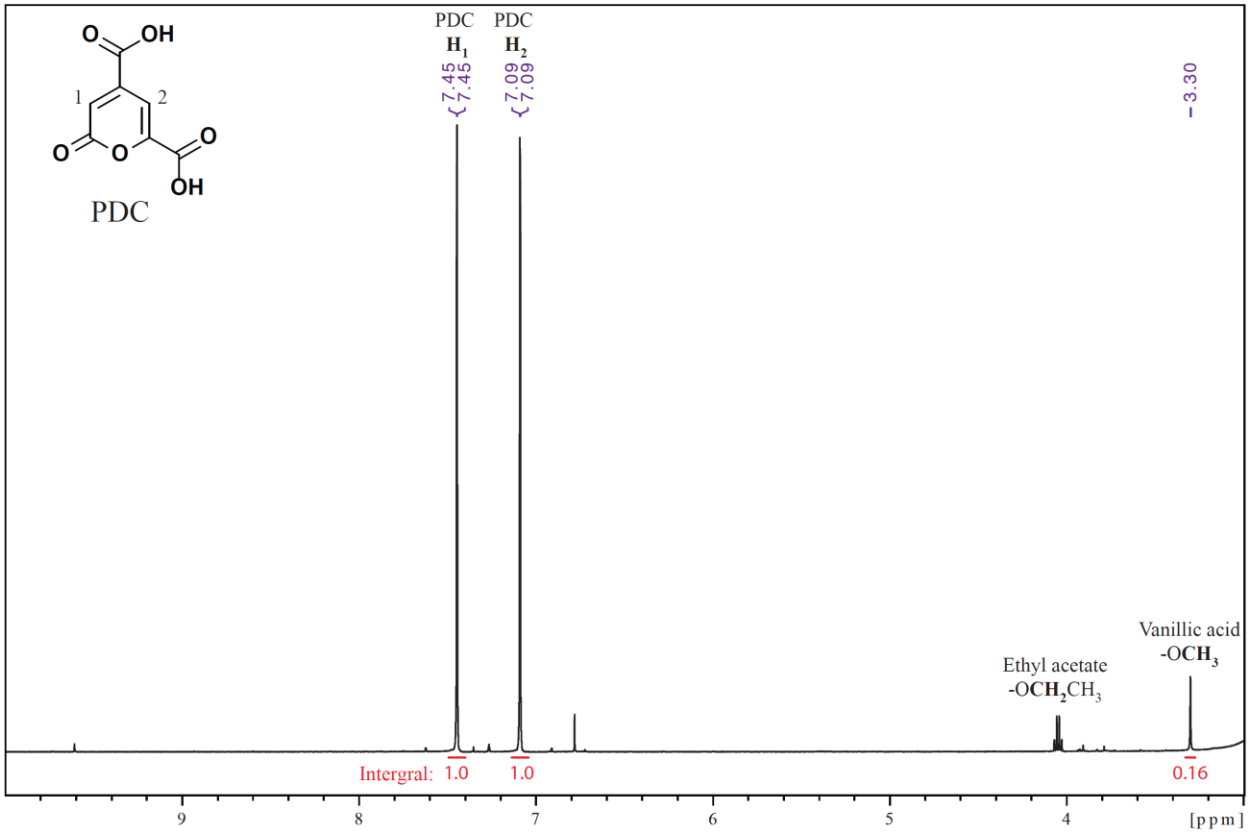


Figure 2.S8. ^{1}H NMR spectra in acetone-d_6 of PDC isolated from a culture of PDC-producing *N. aromaticivorans* strain 12444*ΔligI* grown on a mixture of 3mM vanillic acid and 3 mM glucose.

2.9 Supplementary methods

2.9.1 Construction of plasmids for deleting genes Saro_2819 or Saro_2864/5.

Regions of *N. aromaticivorans* genomic DNA containing ~1100 bp upstream and downstream of Saro_2819 or Saro_2864/5 were PCR amplified separately using the pairs of primers Saro2819_Del-R / Saro2819-pK18_Amp-F and Saro2819_Del-F / Saro2819-pK18_Amp-R for Saro_2819, and Saro2864-5_Del_R / Saro2864-5-pK18_Amp_F and Saro2864-5_Del_F / Saro2864-5-pK18_Amp_R for Saro_2864/5 (Table S2). The pairs of DNA amplified flanking regions for each gene were combined with linearized pK18msB using NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs, Ipswich, MA) to produce the plasmids pK18msB/ΔSaro2819 and pK18msB/ΔSaro2864/5, respectively. A 32 bp region of Saro_2865 (including the start codon) is predicted to overlap with Saro_2866. To prevent transcription of this region of Saro_2865, this putative start codon of Saro_2865 was mutated by replacing a T by a C at position 3088561 in the genome (in addition to deleting the sequence of Saro_2865 downstream of the Saro_2866 stop codon). To mutate the Saro_2865 start site, PCR was performed on plasmid pK18msB/ΔSaro2864/5 using the primers Saro2865_Start-SNP_F and Saro2865_Start-SNP_R2, which were previously phosphorylated with polynucleotide kinase from Promega (Madison, WI). The amplified product was circularized with T4 DNA ligase from New England Biolabs to obtain the circular plasmid pK18msB/ΔSaro2864/5. The plasmids were then transformed into NEB 5-alpha competent *E. coli* (New England Biolabs). The transformed *E. coli* cells were then cultured in LB media + kanamycin and the plasmids purified using a Quiagen® Plasmid Maxi Kit (Quiagen, Germany).

2.9.2 Deletion of genes Saro_2819 and Saro_2864/5.

The purified plasmids were then transformed into competent *E. coli* S17-1 and subsequently mobilized into N. aromaticivorans strain 12444*Δ1879* or 12444*ΔligI* cells via conjugation. Transconjugant cells of *N. aromaticivorans* (single cross overs) were isolated on SISnc-V0 plates containing 1 g/L glucose and 50 ug/mL kanamycin. To select for cells that eliminated the plasmid via a second instance of homologous recombination (double crossovers), single crossover cells were cultured on SISnc-V0 media containing 1 g/L glucose and 10% sucrose. Double crossover cells were isolated on SISnc-V0 plates containing 1 g/L

glucose and 10% sucrose. PCR amplified regions of the target genes were sequenced to verify the deletions.

2.9.3 Purification of PDC

PDC was biologically produced by culturing *Novosphingobium aromaticivorans* strain 12444*ΔligI* in SISnc-V0 media supplemented with 3 mM vanillic acid and 3 mM glucose. Cells were grown to stationary phase and the culture media spun at 5000 RPM for 10 minutes and then filtered using a 500 ml Rapid – Flow bottle top filter with 0.2 μM SFCA membrane (Thermo Scientific). The filtrate (~900 mL) was transferred to a large 2 L separatory funnel and prepared for extraction of the acidic PDC by dilution with 50 mL brine (saturated sodium chloride) and 20 mL concentrated hydrogen chloride. The acidified PDC was extracted with ethyl acetate (4x100 mL). The combined ethyl acetate fraction (~400 mL) was extracted with 0.1 M sodium hydroxide (4x50 mL). The combined sodium hydroxide fraction was acidified with 2 M hydrogen chloride (20 mL) and brine (50 mL), then extracted with ethyl acetate (3x100 mL). The combined ethyl acetate fraction was dried using anhydrous sodium sulfate, filtered through a qualitative cellulose filter (VWR 28320-100), and the solvent removed on a rotatory evaporator giving 297 mg of PDC as a light orange solid. A TMS derivatized sample of the isolated PDC was characterized by GC-MS (method described in materials and methods section), which showed that PDC was the only peak, indicating a fairly high purity. The identity and purity of the PDC was confirmed by comparison of the ^{1}H NMR data to previously published values. The NMR and GC-MS spectra indicated the purity of PDC to be approximate 97%.

2.9.4 Steps in the synthesis of S diketone

Synthesis of 4-acetyl syringaldehyde. To a 100 mL round bottom flask with stir bar were added syringaldehyde (3.296 g, 18.09 mmol), acetic anhydride (3.2 mL, 33.85 mmol), diisopropyl ethyl amine (1 mL, 5.74 mmol), potassium carbonate (793 mg, 5.74 mmol), and dichloromethane (50 mL). The solution was allowed to stir at room temperature. After 24 hours, the reaction was added to a separatory

funnel, washed with saturated sodium bicarbonate (3 x 100 mL), and concentrated in vacuo to yield 4-acetyl syringaldehyde as an off-white solid (3.812 g, 17.00 mmol, 94% yield). ^{1}H NMR (400 MHz, Chloroform-d) δ 9.91 (s, 1H), 7.16 (s, 2H), 3.91 (s, 6H), 2.37 (s, 3H).

Synthesis of 1-(4-acetoxy-3,4-dimethoxyphenyl)-1-propene. An oven dried, 100 mL round bottom flask with stir bar was charged with ethyltriphenylphosphonium bromide (7.0 g, 18.85 mmol), outfitted with a rubber septum, and the atmosphere within it purged with nitrogen. Freshly distilled THF (50 mL) was added via syringe and cooled to -78° C. While stirring, a solution of 2.0 M lithium diisopropyl amide (9.5 mL, 19 mmol) was added to generate ethenyltriphenylphosphonium bromide. While this solution stirred for 30 minutes, an oven dried, 250 mL round bottom flask with stir bar was charged with 4-acetyl-syringaldehyde (3.812 g, 17.0 mmol), sealed with a rubber septum, and purged with nitrogen. Freshly distilled THF (50 mL) was added via syringe and cooled to -78° C. Once the aldehyde was fully dissolved, the ethenyltriphenylphosphonium bromide solution was transferred by cannula and positive pressure to the 4-acetyl-syringaldehyde solution in a dropwise manner over the course of 45 minutes. Upon completion, the reaction was allowed to stir at -78° C for an hour. The reaction was then brought to room temperature and stirred for two hours. The solution was quenched with saturated aqueous ammonium chloride and concentrated under reduced pressure. The remaining solution was diluted with water and extracted with ethyl acetate (3 x 100 mL). The organic layer was then evaporated leaving behind a pale yellow solid. The crude was purified by flash silica chromatography (5:1 hexanes/ethyl acetate). Fractions corresponding to the desired product were combined and evaporated, leaving behind 1-(4-acetoxy-3,4-dimethoxyphenyl)-1-propene as a white powder (1.2 g, 5.36 mmol, 32 % yield, 1.08:1 cis/trans). ^{1}H NMR (400 MHz, Chloroform-d) δ 6.54 (s, 2H), 6.35 (dq, J = 11.6, 1.9, 1H), 5.79 (dq, J = 11.6, 7.2 Hz, 1H), 3.82 (s, 6H), 2.34 (s, 3H), 1.92 (dd, J = 7.2, 1.9 Hz, 3H).

Synthesis of 1-(4-acetoxy-3,4-dimethoxyphenyl)-1,2-propane dione. To a 100 mL round bottom flask with stir bar were added 1-(4-acetoxy-3,4-dimethoxyphenyl)-1-propene (720 mg, 3.05 mmol), dichloro(p-cymene)Ru(II) dimer (69.2 mg, 0.042 mmol), tetrabutylammonium iodide (336.4 mg, 0.91 mmol), tert-

butyl hydroperoxide (70% solution in water, 3.6 mL), toluene (20 mL), acetonitrile (20 mL), and water (2.2 mL). The solution was allowed to stir at room temperature for 30 minutes then quenched with an excess of saturated aqueous sodium thiosulfate. The organic layer was isolated, concentrated in vacuo to a thick residue, and then purified by flash silica chromatography (4:1 hexanes/ethyl acetate). The resulting bright yellow fractions corresponding to the product were combined and evaporated to yield 1-(4-acetoxy-3,4-dimethoxyphenyl)-1,2-propane dione as a bright yellow solid (445 mg, 1.67 mmol, 55% yield). ^{1}H NMR (400 MHz, Chloroform-d) δ 7.33 (s, 2H), 3.88 (s, 6H), 2.53 (s, 3H), 2.36 (s, 3H).

Synthesis of 1-(4-hydroxy-3,4-dimethoxyphenyl)-1,2-propane dione (S-diketone). To a 250 mL round bottom flask were added 1-(4-acetoxy-3,4-dimethoxyphenyl)-1,2-propane dione (445 mg, 1.67 mmol), 3 M HCl (35 mL), and methanol (75 mL). The solution stirred at room temperature and reaction progress was monitored by TLC. Upon completion, the reaction was concentrated, diluted with saturated sodium bicarbonate, and washed with ethyl acetate. The aqueous layer was acidified with dilute ammonium chloride and extracted with ethyl acetate (3 x 50 mL). The resulting organic layer was concentrated and purified by flash silica chromatography (4:1 hexanes/ethyl acetate). The desired fractions were combined and evaporated to yield 1-(4-hydroxy-3,4-dimethoxyphenyl)-1,2-propane dione (S-diketone) as a bright yellow solid (259 mg, 1.16 mmol, 69% yield). ^{1}H NMR (400 MHz, Chloroform-d) δ 7.34 (s, 2H), 6.11 (s, 1H), 3.95 (s, 3H), 2.53 (s, 3H).

2.9.5 Steps in the synthesis of G diketone

Synthesis of isoeugenyl acetate. To a 100 mL round bottom flask with stir bar were added isoeugenol (2.6 mL, 17.10 mmol), acetic anhydride (3.00 mL, 31.73 mmol), diisopropyl ethyl amine (1 mL, 5.74 mmol), potassium carbonate (793 mg, 5.74 mmol), and dichloromethane (500 mL). The solution was allowed to stir at room temperature. After 24 hours, the reaction was added to a separatory funnel, washed with saturated sodium bicarbonate (3 x 100 mL), and concentrated in vacuo. The resulting off white powder was recrystallized from hot acetone to yield isoeugenyl acetate as white crystals (2.292 g, 11.11 mmol, 65% yield). ^{1}H NMR (400 MHz, Chloroform-d) δ 6.95 (d, J = 8.1 Hz, 1H), 6.92 (d, J = 1.8 Hz,

1H), 6.89 (dd, J = 8.1, 1.9 Hz, 1H), 6.36 (dq, J = 15.6, 1.7 Hz, 1H), 6.18 (dq, J = 15.7, 6.6 Hz, 1H), 3.84 (s, 3H), 2.30 (s, 3H), 1.88 (dd, J = 6.6, 1.6 Hz, 3H).

Synthesis of 1-(4-acetoxy-3-methoxyphenyl)-1,2-propane dione. To a 250 mL round bottom flask with stir bar were added isoeugenyl acetate (2.060 g, 9.99 mmol), dichloro(p-cymene)Ru(II) dimer (69.2 mg, 0.11 mmol), tetrabutylammonium iodide (1.12 g, 3.03 mmol), tert-butyl hydroperoxide (70% solution in water, 10 mL), toluene (30 mL), acetonitrile (30 mL), and water (7 mL). The solution was allowed to stir at room temperature for 45 minutes then quenched with an excess of saturated aqueous sodium thiosulfate. The organic layer was isolated, concentrated in vacuo to a thick residue, and then purified by flash silica chromatography (4:1 hexanes/ethyl acetate). The resulting bright yellow fractions corresponding to the product were combined and evaporated to yield 1-(4-acetoxy-3-methoxyphenyl)-1,2-propane dione as a bright yellow solid (1.28 g, 5.42 mmol, 54% yield). ^{1}H NMR (400 MHz, Chloroform-d) δ 7.66 (d, J = 1.9 Hz, 1H), 7.64 (dd, J = 8.1, 1.9 Hz, 1H), 7.16 (d, J = 8.2 Hz, 1H), 3.90 (s, 3H), 2.52 (s, 3H), 2.34 (s, 3H).

Synthesis of 1-(4-hydroxy-3-methoxyphenyl)-1,2-propane dione (G-diketone). To a 500 mL round bottom flask were added 1-(4-acetoxy-3-methoxyphenyl)-1,2-propane dione (1.00 g, 4.23 mmol), 3 M HCl (90 mL), and methanol (190 mL). The solution was stirred at room temperature and reaction progress was monitored by TLC. Upon completion, the reaction was concentrated, diluted with saturated sodium bicarbonate, and washed with ethyl acetate. The aqueous layer was acidified with dilute ammonium chloride and extracted with ethyl acetate (3 x 100 mL). The resulting organic layer was concentrated and purified by flash silica chromatography (4:1 hexanes/ethyl acetate). The desired fractions were combined and evaporated to yield 1-(4-hydroxy-3-methoxyphenyl)-1,2-propane dione as a bright yellow, viscous oil (526 mg, 2.71 mmol, 64% yield). ^{1}H NMR (400 MHz, Chloroform-d) δ 7.61 (dd, J = 8.3, 1.9 Hz, 1H), 7.58 (d, J = 1.9 Hz, 1H), 6.98 (d, J = 8.3 Hz, 1H), 6.21 (s, 1H), 3.97 (s, 3H), 2.51 (s, 3H).

2.10 References

1. Kamimura N, Takahashi K, Mori K, Araki T, Fujita M, Higuchi Y, Masai E. 2017. Bacterial catabolism of lignin-derived aromatics: New findings in a recent decade: Update on bacterial lignin catabolism. Environmental Microbiology Reports 9:679-705.

2. Cecil JH, Garcia DC, Giannone RJ, Michener JK. 2018. Rapid, parallel identification of pathways for catabolism of lignin-derived aromatic compounds in *Novosphingobium aromaticivorans*. Appl Environ Microbiol 84:1-13.

3. Vogel J. 2008. Unique aspects of the grass cell wall. Curr Opin Plant Biol 11:301-307.

4. Rahimi A, Ulbrich A, Coon JJ, Stahl SS. 2014. Formic- acid- induced depolymerization of oxidized lignin to aromatics. Nature 515.

5. Jørgensen H, Kristensen JB, Felby C. 2007. Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. Biofuels, Bioprod Biorefin 1(2):119-134.

6. Sun Z, Fridrich B, de Santi A, Elangovan S, Barta K. 2018. Bright side of lignin depolymerization: Toward new platform chemicals. Chem Rev 118:614-678.

7. Schutyser W, Renders T, Van den Bosch S, Koelewijn SF, Beckham GT, Sels BF. 2018. Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. Chem Soc Rev 47:852-908.

8. Ralph J, Brunow G, Boerjan W. 2007. Lignins. eLS.

9. Harris PJ, Hartley RD. 1976. Detection of bound ferulic acid in cell walls of the Gramineae by ultraviolet fluorescence microscopy. Nature 259:508.

10. Harris PJ, Hartley RD. 1980. Phenolic constituents of the cell walls of monocotyledons. Biochem Syst Ecol 8:153-160.

11. Smith DCC. 1955. Ester Groups in Lignin. Nature 176:267.

12. Vanholme R, Morreel K, Darrah C, Oyarce P, Grabber JH, Ralph J, Boerjan W. 2012. Metabolic engineering of novel lignin in biomass crops. New Phytol 196:978-1000.

13. Fuchs G, Boll M, Heider J. 2011. Microbial degradation of aromatic compounds - from one strategy to four. Nat Rev Microbiol 9:803-16.

14. Bugg TD, Ahmad M, Hardiman EM, Rahmanpour R. 2011. Pathways for degradation of lignin in bacteria and fungi. Nat Prod Rep 28:1883-96.

15. Beckham GT, Johnson CW, Karp EM, Salvachua D, Vardon DR. 2016. Opportunities and challenges in biological lignin valorization. Curr Opin Biotechnol 42:40-53.

16. Linger JG, Vardon DR, Guarnieri MT, Karp EM, Hunsinger GB, Franden MA, Johnson CW, Chupka G, Strathmann TJ, Pienkos PT, Beckham GT. 2014. Lignin valorization through integrated biological funneling and chemical catalysis. Proc Natl Acad Sci U S A 111:12013-12018.

17. Sainsbury PD, Hardiman EM, Ahmad M, Otani H, Seghezzi N, Eltis LD, Bugg TDH. 2013. Breaking down lignin to high-value chemicals: The conversion of lignocellulose to vanillin in a gene deletion mutant of *Rhodococcus jostii* RHA1. ACS Chem Biol 8:2151-2156.

18. Austin S, Kontur WS, Ulbrich A, Oshlag JZ, Zhang W, Higbee A, Zhang Y, Coon JJ, Hodge DB, Donohue TJ, Noguera DR. 2015. Metabolism of multiple aromatic compounds in corn stover hydrolysate by *Rhodopseudomonas palustris*. Environ Sci Technol 49:8914.

19. Vardon D, Franden MA, Johnson C, Karp E, Guarnieri M, Linger J, Salm M, Strathmann T, Beckham G, Ferguson G. 2015. Adipic acid production from lignin. Energy Environ Sci 8:617-628.

20. Okamura-Abe Y, Abe T, Nishimura K, Kawata Y, Sato-Izawa K, Otsuka Y, Nakamura M, Kajita S, Masai E, Sonoki T, Katayama Y. 2016. Beta-ketoadipic acid and muconolactone production from a lignin-related aromatic compound through the protocatechuate 3,4-metabolic pathway. J Biosci Bioeng 121:652-658.

21. Otsuka Y, Nakamura M, Shigehara K, Sugimura K, Masai E, Ohara S, Katayama Y. 2006. Efficient production of 2-pyrone 4,6-dicarboxylic acid as a novel polymer-based material from protocatechuate by microbial function. Appl Microbiol Biotechnol 71:608-614.

22. Qian Y, Otsuka Y, Sonoki T, Mukhopadhyay B, Nakamura M, Jellison J, Goodell B. 2016. Engineered microbial production of 2-pyrone-4,6-dicarboxylic acid from lignin residues for use as an industrial platform chemical. BioResources 11:6097-6109.

23. Mycroft Z, Gomis M, Mines P, Law P, Bugg TDH. 2015. Biocatalytic conversion of lignin to aromatic dicarboxylic acids in *Rhodococcus jostii* RHA1 by re-routing aromatic degradation pathways. Green Chem 17:4974-4979.

24. Fredrickson JK, Balkwill DL, Drake GR, Romine MF, Ringelberg DB, White DC. 1995. Aromatic- degrading Sphingomonas isolates from the deep subsurface. Appl Environ Microbiol 61:1917.

25. Kontur WS, Bingman CA, Olmsted CN, Wassarman DR, Ulbrich A, Gall DL, Smith RW, Yusko LM, Fox BG, Noguera DR, Coon JJ, Donohue TJ. 2018. *Novosphingobium aromaticivorans* uses a Nu-class glutathione S -transferase as a glutathione lyase in breaking the β-aryl ether bond of lignin. J Biol Chem 293:4955-4968.

26. Shikinaka K, Otsuka Y, Nakamura M, Masai E, Katayama Y. 2018. Utilization of lignocellulosic biomass via novel sustainable process. J Oleo Sci 67:1059-1070.

27. Masai E, Shinohara S, Hara H, Nishikawa S, Katayama Y, Fukuda M. 1999. Genetic and biochemical characterization of a 2-Pyrone-4,6-dicarboxylic acid hydrolase involved in the protocatechuate 4,5-cleavage pathway of *Sphingomonas paucimobilis* SYK-6. J Bacteriol 181:55-62.

28. Kasai D, Masai E, Miyauchi K, Katayama Y, Fukuda M. 2004. Characterization of the 3-O-Methylgallate dioxygenase gene and evidence of multiple 3-O-Methylgallate catabolic pathways in *Sphingomonas paucimobilis* SYK-6. J Bacteriol 186:4951-4959.

29. Sze IS, Dagley S. 1987. Degradation of substituted mandelic acids by meta fission reactions. J Bacteriol 169:3833-5.

30. Michinobu T, Hishida M, Sato M, Katayama Y, Masai E, Nakamura M, Otsuka Y, Ohara S, Shigehara K. 2008. Polyesters of 2-pyrone-4,6-dicarboxylic acid (PDC) obtained from a metabolic intermediate of lignin. Polym J (Tokyo, Jpn) 40:68-75.

31. Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon RA, Gilna P, Keller M, Langan P, Naskar AK, Saddler JN, Tschaplinski TJ, Tuskan GA, Wyman CE. 2014. Lignin valorization: Improving lignin processing in the biorefinery. Science 344.

32. Zakzeski J, Bruijnincx PCA, Jongerius AL, Weckhuysen BM. 2010. The catalytic valorization of lignin for the production of renewable chemicals. Chem Rev 110:3552-3599.

33. Gall DL, Ralph J, Donohue TJ, Noguera DR. 2017. Biochemical transformation of lignin for deriving valued commodities from lignocellulose. Curr Opin Biotechnol 45:120-126.

34. D'argenio V, Notomista E, Petrillo M, Cantiello P, Cafaro V, Izzo V, Naso B, Cozzuto L, Durante L, Troncone L, Paolella G, Salvatore F, Di Donato A. 2014. Complete sequencing of Novosphingobium sp. PP1Y reveals a biotechnologically meaningful metabolic pattern.(Research article)(Report). Acta Vet Scand 15:384.

35. Mitchell VD, Taylor CM, Bauer S. 2014. Comprehensive analysis of monomeric phenolics in dilute acid plant hydrolysates. BioEnergy Res 7:654-669.

36. Masai E, Katayama Y, Fukuda M. 2007. Genetic and biochemical investigations on bacterial catabolic pathways for lignin-derived aromatic compounds. Biosci, Biotechnol, Biochem 71:1-15.

37. Schäfer A, Tauch A, Jäger W, Kalinowski J, Thierbach G, Pühler A. 1994. Small mobilizable multi- purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene 145:69-73.

38. Sistrom WR. 1962. The kinetics of the synthesis of photopigments in *Rhodopseudomonas spheroides*. J Gen Microbiol 28:607-616.

39. Bhalla A, Bansal N, Stoklosa R, Fountain M, Ralph J, Hodge D, Hegg E. 2016. Effective alkaline metal-catalyzed oxidative delignification of hybrid poplar. Biotechnol Biofuels 9:34.

40. Das A, Rahimi A, Ulbrich A, Alherech M, Motagamwala AH, Bhalla A, da Costa Sousa L, Balan V, Dumesic JA, Hegg EL, Dale BE, Ralph J, Coon JJ, Stahl SS. 2018. Lignin conversion to low-molecular-weight aromatics via an aerobic oxidation-hydrolysis sequence: Comparison of different lignin sources. ACS Sustainable Chemistry & Engineering 6:3367-3374.

41. Li Z, Chen CH, Liu T, Mathrubootham V, Hegg EL, Hodge DB. 2013. Catalysis with Cu II (bpy) improves alkaline hydrogen peroxide pretreatment. Biotechnol Bioeng 110:1078-1086.

42. Michinobu T, Bito M, Yamada Y, Katayama Y, Noguchi K, Masai E, Nakamura M, Ohara S, Shigehara K. 2007. Molecular properties of 2-pyrone-4,6-dicarboxylic acid (PDC) as a stable metabolic intermediate of lignin isolated by fractional precipitation with Na+ ion. Bull Chem Soc Jpn 80:2436-2442.

43. Kasai D, Masai E, Katayama Y, Fukuda M. 2007. Degradation of 3-O-methylgallate in *Sphingomonas paucimobilis* SYK- 6 by pathways involving protocatechuate 4,5-dioxygenase. FEMS Microbiol Lett 274:323-328.

44. Taylor RG, Walker DC, McLnnes RR. 1993. E.coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. Nucleic Acids Research 21:1667-1678.

45. Simon R, Priefer U, Pühler A. 1983. A broad host range mobilization system for in vivo genetic engineering: Transposon mutagenesis in gram negative bacteria. Bio/Technology 1:784.

46. Schäfer A, Tauch A, Jäger W, Kalinowski J, Thierbach G, Pühler A. 1994. Small mobilizable multi- purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene 145:69-73.

3. Integrating lignin depolymerization and microbial funneling processeses using agronomically relevant feedstocks

A version of this chapter has been prepared for submission to a peer-reviewed journal. The contribution of each co-author is as follows:

Jose M. Perez: Conceived the project, planned and performed the experiments, performed HPLC-MS analysis, wrote the manuscript

German Umana: Planned and performed the experiments, performed HPLC-MS analysis, wrote the manuscript

Canan Sener: Performed lignin hydrogenolysis, performed GC-FID analysis, wrote the manuscript

Jason Coplien: Produced GVL lignin, wrote the manuscript

Dennis Haak: Produced GVL lignin

Yanding Li: Conceived the project

Steven D. Karlen: Conceived the project, performed ^{1}H NMR and GC-FID analyses, wrote the manuscript

John Ralph: Wrote the manuscript

Timothy J. Donohue: Wrote the manuscript

Daniel R. Noguera: Wrote the manuscript

3.1 Abstract

Plant biomass, or lignocellulose, is one of the most abundant renewable sources of organic carbon with the potential to replace fossil fuels for the production of fuels and chemicals. It is composed of sugar-based polymers such as cellulose and hemicellulose, and lignin, an amorphous aromatic heteropolymer. Although numerous techniques have been developed to utilize the sugar fraction, lignin utilization still remains an unresolved challenge and its valorization is critical for the economic feasibility of future lignocellulose-based biorefineries. One strategy to valorize valorize lignin in the context of a biorefinery is the combination of chemical techniques for lignocellulose fractionation and lignin depolymerization followed by biological funneling of the resulting monomeric phenolic compounds into a single valuable product using engineered bacteria. In previous work, we investigated the integration of γ-valerolactone (GVL) biomass fractionation with hydrogenolysis lignin depolymerization. We separately developed an engineered strain of the bacterium *Novosphingobium aromaticivorans* capable of simultaneously

converting multiple lignin-derived phenolic compounds into 2-pyrone-4,6-dicarboxylic acid (PDC), a potential bioplastics and epoxy adhesives precursor. In this work, we demonstrate the feasibility of producing PDC in the context of a biomass-to-PDC pipeline, integrating GVL pretreatment, hydrogenolysis lignin depolymerization, and biological funneling with *N. aromaticivorans* strain PDC. We first investigated the ability of the engineered bacterium to produce PDC form single compounds using an array of different compounds commonly found in lignin depolymerized by hydrogenolysis using different catalysts. We found that phenolic compounds with a propanol, methyl, propanoic methyl ester, or a carboxylic acid methyl ester side-chains were converted into PDC with molar yields ranging from 28% to nearly stoichiometric. On the other hand, compounds with propyl or ethyl side-chains, syringol, or guaiacol, did not lead to the production of PDC. Then, we extracted lignin from four types of agronomically relevant sources of lignocellulose, namely maple, poplar, sorghum, and switchgrass using GVL, and depolymerized by hydrogenolysis using Pd/C, a catalyst known to favor the production of phenolic compounds with a propanol as the side-chain. The depolymerized lignin was then transformed into PDC by the engineered strain. For all lignin sources PDC was produced with yields of 88.4, 139.1, 103.0, and 79.2 g PDC/kg lignin for maple, poplar, sorghum, and switchgrass, respectively. Our results show the feasibility of the integration of the different techniques and provides valuable information for the future optimization of the biomass-to-PDC pipeline.

3.2 Introduction

Plant biomass is a promising raw material for the production of fuels and chemicals that enable the development of a sustainable and carbon neutral economy (1). The major structural component of plant biomass is the cell wall, which is composed of sugar polymers such as cellulose and hemicellulose, and lignin, a phenolic heteropolymer. Whereas numerous successful techniques have been developed to valorize the sugar fraction, the challenges derived from the natural recalcitrance and heterogeneity of lignin, combined with its tendency to condense under classical pulping conditions, have often limited its

utilization beyond a low value source of energy (2) or as filler material in construction applications (3). As the economic feasibility of biorefineries producing fuels and chemicals from plant biomass depend on the efficient utilization of all polymeric components (4, 5), additional efforts are necessary to develop appropriate techniques for the valorization of the underutilized phenolic fraction.

One approach to valorize lignin in the context of a biorefinery is the combination of chemical depolymerization followed by microbial funneling of the mixture of aromatics present in depolymerized lignin into a single valuable product (6). One of the target products of biological funneling is 2-pyrone-4,6-dicarboxylic acid (PDC), a stable intermediate in the bacterial metabolism of multiple types of lignocellulose-derived aromatic compounds (7, 8). Potentially, PDC could be used as a precursor to bioplastics and epoxy adhesives (9). In previous work, we showed that an engineered strain of *Novosphingobium aromaticivorans* DSM12444 (strain PDC) was able to produce PDC from multiple aromatic compounds commonly found after lignin depolymerization and demonstrated its application with depolymerized lignin from poplar (8). Although other studies have shown that PDC can be biologically produced from single aromatic compounds (10), plant biomass extracts (11), industrial lignin extracts (11, 12), and non-lignin derived substrates such as glucose (13), a systematic evaluation of integrating chemical lignin depolymerization with microbial funneling using different agronomically relevant species is lacking. A wide variety of potential lignin depolymerization techniques are currently under investigation (14), with all of them capable of breaking down the β-aryl-ether (β-O-4) bond, which is the most abundant polymer backbone linkage. Most lignin depolymerization strategies that target cleavage of the β-aryl-ether bonds can be performed on isolated lignin streams or directly on whole biomass. The efficiency of most of these depolymerization processes, in terms of the yield of aromatic monomers, is proportional to the fraction of the sidechain linkages that are β-aryl-ether, which depends on the source and quality of the lignin. Whereas, highly condensed lignins (e.g., Kraft pulp) typically have very few remaining β-aryl-ether bonds and produce very few monomers, native-like lignins produce a higher monomer yield (15, 16).

We have previously investigated integrating catalytic hydrogenolysis of lignin into the γ-valerolactone (GVL) sugar stream production biorefinery (17). The GVL process is an acidic organosolv technique that utilizes 80% GVL/20% water with 100–500 mM sulfuric acid. A biorefinery based on the GVL process could produce a product portfolio consisting of dissolvable pulp, furfural, and a high quality lignin stream (18). The reported benefits of the GVL process over other biomass deconstruction strategies are that it is biomass agnostic and depolymerizes the cellulose and hemicellulose into C6 and C5 sugars without the need for enzymatic hydrolysis (19).

We are interested in creating integrated lignin depolymerization and microbial funneling processes that maximize the conversion of plant-derived phenolics to valuable products. Most lignin depolymerization processes that have been proposed for application in biorefineries are performed using heterogeneous catalysts under either reductive or oxidative conditions (14, 20). We have reported the production of PDC from the crude product mixture of an oxidative process (8). Reductive depolymerization processes produce a completely different set of aromatic products, which have fewer degrees of unsaturation and have lower oxygen content than found in products from oxidative processes. Catalytic hydrogenolysis is a typical reductive processes that produces high yields of phenolic monomers and oligomers under many different catalytic conditions (14). When performed under milder conditions, the catalyst cleaves the aryl-ether bonds and saturates non-aromatic double bonds producing a viscous oil with complex chemical composition (20, 21). Under harsher conditions, the catalyst can reduce all the double bonds and deoxygenate the hydrocarbons, producing light hydrocarbon (cracking) (22). Catalytic depolymerization conditions with Pd/C, Ni/C, Ni/Al_2O_3, Rh/C, and other heterogeneous catalysts can be tuned to favor products that retain the main S, G, or H phenolic substructure with propanol as the side-chain. Under slightly different conditions these same catalysts, as well as other catalysts (e.g., solid supported ruthenium, platinum, zinc, and copper catalysts) can be tuned to favor propane as the side-chain or the fully saturated propyl-cyclohexanes (14, 20).

In this study, we expanded our previous efforts in integrating the GVL-biorefinery producing furfural and dissolving pulp (18), with lignin hydrogenolysis (17) to include biological funneling of the reductive hydrogenolysis product mixture in a biomass-to-PDC pipeline (Figure 3.1). With the large array of possible monomers that can be produced by hydrogenolysis, we first evaluated which compounds could be used by the PDC-producing *N. aromaticivorans* strain (8). With this new knowledge, we chose a catalyst that favors product mixtures predicted to produce PDC at high yield. We applied these pipeline conditions to four industrially relevant biofuel feedstocks that represent both grasses and hardwoods, namely poplar, switchgrass, sorghum, and maple. This comparative evaluation of PDC production from different feedstocks offers important observations needed to optimize the feasibility of a biomass-to-product pipeline that uses GVL deconstruction followed by hydrogenolysis lignin depolymerization and biological funneling to PDC.

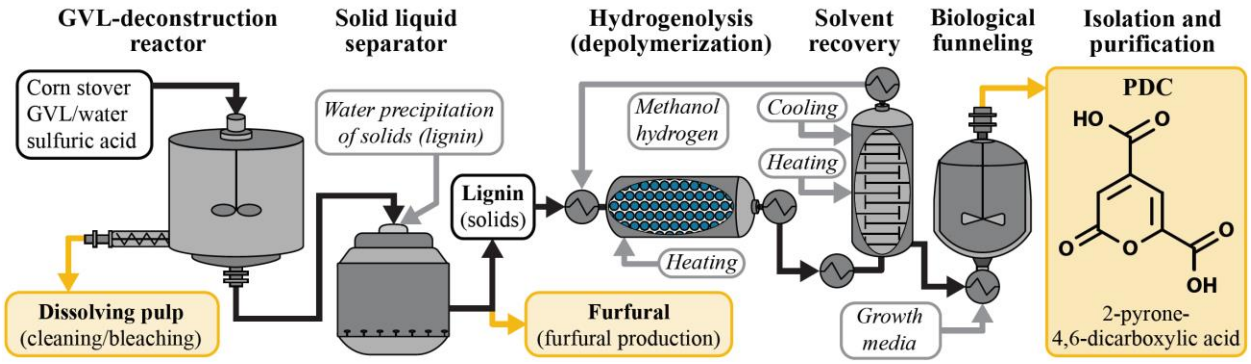


Figure 3.1. Schematics of a biomass-to-PDC pipeline that integrates GVL-biorefinery producing furfural and dissolving pulp, with lignin hydrogenolysis and biological funneling.

3.3 Results

3.3.1 Conversion of single compounds into PDC

Lignin hydrogenolysis performed under mild conditions produces a wide variety of aromatic compounds that retain the methoxy groups of the phenolic subunits and have different side-chains. The ratio and

composition of the dominant sidechains are tuned by the reaction condition (e.g., catalyst type and loading, solvent, pressure, lignin concentration, and reaction temperature profile) and catalyst selection (14, 20). Under mild conditions, the major phenolic monomer products of hydrogenolysis are phenolic syringyl (S) or guaiacyl (G) units with 3-propanol (dihydrosinapyl alcohol (DSA) and dihydroconiferyl alcohol (DCA)) or propyl sidechains (propylsyringol (PS) and propylguaiacol (PG)), *p*-hydroxybenzoic acid methyl ester (Me-*p*HBA, major product when the biomass is from *Populus* hardwoods such as, poplar, aspen, and willow), 7,8-dihydro-*p*-coumaric acid methyl ester (Me-DH*p*CA) and 7,8-dihydroferulic acid methyl ester (Me-DHFA) (major products when the biomass is from grasses such as sorghum, switchgrass, and corn stover) (20). We tested PDC production from these known hydrogenolysis products when the PDC strain was grown in minimal medium containing glucose (needed to support cell growth (8)) plus each individual aromatic compound (Figure 3.2, Table 3.1). We also tested PDC production from other aromatic products known to be minor products of hydrogenolysis, such as guaiacol (Gu), syringol (Sy), and the phenolic S and G units with either an ethyl (ethylsyringol (ES) and ethylguaiacol (EG)) or methyl (methylsyringol (MS) and methylguaiacol (MG)) sidechain (Figure 3.S1, Table 3.S1).

All the major products of hydrogenolysis tested (Figure 3.2) were consumed by *N. aromaticivorans*, but not all of them produced significant quantities of PDC. PDC accumulation was observed from DSA, DCA, Me-DHFA, Me-DH*p*CA, and Me-*p*HBA, whereas experiments with PS and PG did not show significant PDC accumulation. The rates at which the aromatic substrates disappear from the medium varied, but by the end of the experiments (19 or 24 hr), only PG and Me-*p*HBA were detectable in the culture medium. The PDC yield from the aromatics that were transformed to PDC was somewhat correlated to the number of methoxy groups in the compound, with highest yields observed for compound with zero or one methoxy group (Me-DHFA, Me-DH*p*CA, Me-*p*HBA) and the lower yields obtained with DSA (Table 3.1).

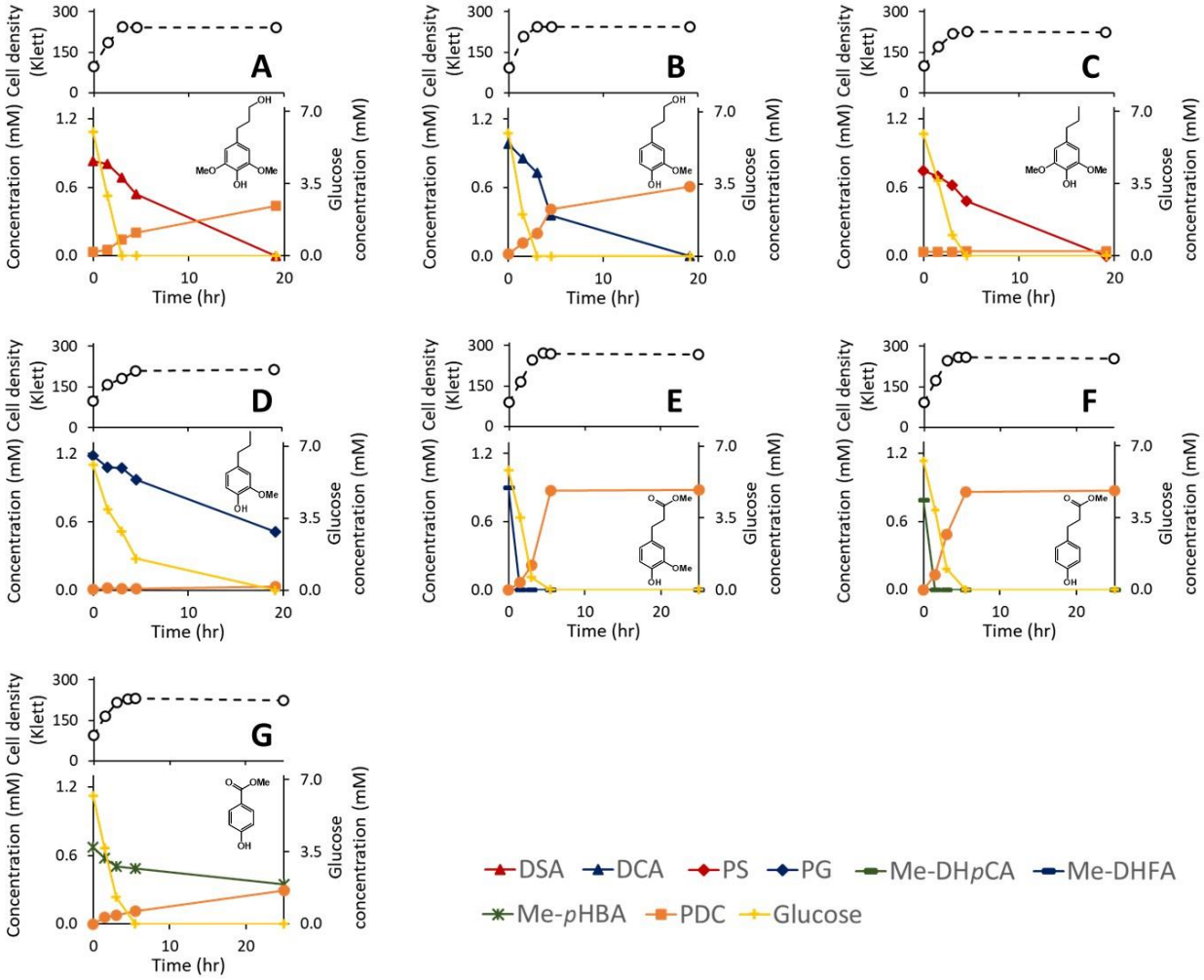
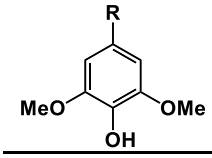
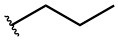
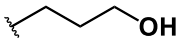
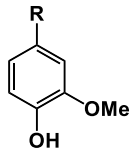
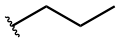
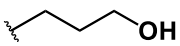
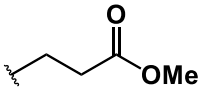
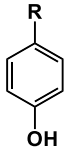
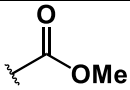
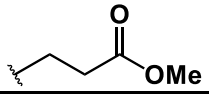


Figure 3.2. Probing the ability for *N. aromaticivorans* strain PDC cultures to convert the major hydrogenolysis products to PDC in minimum media supplemented with glucose. Cell density and extracellular metabolite concentration of cultures supplemented with dihydrosinapyl alcohol (DSA) (A), dihydroconiferyl alcohol (DCA) (B), 7,8-dihydroferulic acid methyl ester (Me-DHFA) (C), propyl syringol (PS) (D), propyl guaiacol (PG) (E), 7,8-dihydro-*p*-coumaric acid methyl ester (Me-DH*p*CA) (F), methyl syringol (MS) (G), methyl guaiacol (MG) (H), or *p*-hydroxybenzoic acid methyl ester (Me-*p*HBA) (I). Values correspond to the average of three biological replicates.

In cultures supplemented with aromatics containing fully dehydrated and saturated propane side-chains, PS and PG (Figure 3.2, panels C and D), the substrate was completely or partially consumed during the experiment, but PDC did not accumulate. Traces of PDC were detected in these cultures at the beginning of the experiment but did not increased as the aromatic substrate was consumed, suggesting that its detection was not related to the consumption of the corresponding supplemented compound, but likely the result of impurities in the substrates used.

Among the aromatics expected to be minor products during hydrogenolysis (Figure 3.S1), the cultures supplemented with MS or MG were able to degrade partially or completely the aromatic substrate during the experiment (Figure S1, panels C and D). In both cases, a consistent increase in PDC concentration was observed (Table 3.S1).

Table 3.1. Yield of PDC (mol) at 19-24 h reaction time from the aromatic substrates screened in this study. Standard error was determined from n=3 technical replicates.

	Compound	R	PDC molar yield (%)
	Propyl syringol (PS)		0.3% ± 0.5%
	Dihydrosinapyl alcohol (DSA)		48.6% ± 3.2%
	Propyl guaiacol (PG)		3.5% ± 1.1%
	Dihydroconiferyl alcohol (DCA)		60.5% ± 1.4%
	7,8-dihydroferulic acid methyl ester (Me-DHFA)		97.9% ± 0.3%
	*p*-hydroxybenzoic acid methyl ester (Me-*p*HBA)		91.6% ± 6.2%
	7,8-dihydro-*p*-coumaric acid methyl ester (Me-DH*p*CA)		110.2% ± 0.3%

In cultures supplemented with syringol (S), a small reduction in its concentration was observed while there was glucose still present and cell were growing (Figure 3.S1E). On the other hand, cultures supplemented with guaiacol (G) showed a consumption of the aromatic substrate even after all glucose was consumed and the compound was completely degraded at the end of the experiment (Figure 3.S1F). In both cases, traces of PDC accumulated in the growth medium during the first 2 hours of reaction and did not increase later in the experiments, indicating that the detection of PDC was not related to its production for the aromatic substrate, but from impurities in the substrates used. Experiments with ES and EG also showed consumption of the aromatic compound, but no PDC production.

3.3.2 Lignin isolation and depolymerization

The lignin hydrogenolysis product screening showed that the *N. aromaticivorans* PDC strain can produce PDC from aromatics containing a propanol, methyl, or methyl ester side-chains and does not produce detectable levels of PDC from aromatics containing a propane or ethyl side chain, nor aromatics without an alkyl side-chain. Scanning across the literature, Pd/C (23-26), Rh/C (27), Ni/C (25) and Ni/Al_2O_3 (25) all produce more propanol than propane side-chains. Of these, Pd/C is an attractive catalyst, as it is fairly air stable and retains the high propanol selectivity over a range of reaction conditions (14). Therefore, we selected to use commercially available Pd/C catalyst for the hydrogenolysis performed in this study. To ensure a very high quality (native like) isolated lignin, we performed the GVL process under the mild conditions previously reported to minimize lignin condensation (17), although this resulted in lower lignin extraction efficiencies compared to harsher conditions that are required to liquefy the cellulose portion of the biomass (17).

In this study, we selected four plant species as the biomass source; namely maple, poplar, switchgrass, and energy sorghum. The maple is representative of most hardwood species, it has a simple lignin composition of mostly S and G units (28). HSQC NMR analysis indicated an S/G ratio of 1.48 (58:42 S/G) and that the side-chains were 62% β-O-4, 14% phenylcoumaran, and 25% resinol (Table 3.2). Poplar (and all *Populus* species) are a unique clade of hardwoods characterized by the presence of *p*-

hydroxybenzoate pendent groups on their lignin S units (29). The NM6 poplar used in this study has been reported (30) to have an S/G ratio of 1.5 (60:40 S/G) and side-chains containing 82% β-O-4 linkages, 10% phenylcoumaran, and 8% resinol (Table 3.2). Switchgrass and sorghum are both grasses, which are known to process easily due to their hemicellulose containing high levels of arabinose-bound ferulate cross linkages and low total lignin content (31). These grasses have different growth patterns, with switchgrass growing in clumps that are 1–2.7 m in height and consisting of mostly leaf tissue and thin hollow flowering stems. Sorghum also grows in clumps that are 1-4 m tall but consist primarily of a thick walled woody stem tissue filled with a starchy white pith. The lignin composition of the two grasses is also different, with switchgrass having a higher S content and an S/G ratio of 1.38 (58:42 S/G), while sorghum is higher in G content with an S/G ratio of 0.82 (45/55 S/G) (Table 3.2). The changes in the sidechain composition reflect these changes with switchgrass consisting of 62% β-O-4 linkages, 27% phenylcoumaran, and 7% resinol, while sorghum was comprised of 57% β-O-4 linkages, 29% phenylcoumaran, and 10% resinol (Table 3.2).

Applying the mild GVL-process (17) described in the Materials and Methods section to the four feedstocks resulted in isolated lignin (GVL-lignins) with yields of 15.2% (maple), 22.3% (poplar), 35.0% (energy sorghum), and 41.0% (switchgrass), with respect to the total Klason lignin values. Experiments with these four feedstocks provide insight into the efficiency of isolating native like GVL-lignins and how lignin composition alters the downstream processes of hydrogenolysis monomer yield and microbial conversion to PDC, as discussed below.

Table 3.2. Characterization of the GVL-lignin and hydrogenolysis product mixture from the four representative biomasses, maple (S/G hardwood), poplar (S/G hardwood with *p*HBA esters), and two grasses sorghum and switchgrass (with different levels of *p*CA and FA esters. Standard error was determined from n=3 technical replicates.

GVL-lignin	**Hard maple**	**Poplar (NM6)**	**Energy Sorghum**	**Switchgrass**
wt% Klason lignin	24.9%	27%	28%	39.7%
wt% GLV-lignin of whole biomass	3.8%	5.4%	5.6%	7.9%
wt% GVL-lignin of Klason lignin	15.2%	22.3%	35.0%	41.0%
NMR analysis of the GVL-lignins				
S:G	60:40	60:40	45:55	58:42
S/G ratio	1.5	1.5	0.82	1.38
Sidechains:				
β-*O*-4:β-5:β-β	62:14:25	82:10:8	57:29:10	62:27:7
(β-*O*-4 : phenylcoumaran : resinol)				
Hydrogenolysis				
(g/kg GVL-lignin ± SEM)				
H monolignols	3.6±0.8	7.9±0.2	7.2±0.1	5.9±0.1
G monolignols	36.5±1.9	58.2±2.6	23.9±0.4	27.3±0.3
S monolignols	48.5±1.2	115.5±5.6	23.0±0.5	20.5±0.2
p-Hydroxybenzoate	0.2±0.1	15.8±0.6	2.8±0.0	1.4±0.0
p-Coumarate	0.8±0.3	0.6±0.0	48.9±1.2	20.3±0.7
Ferulate	0.2±0.0	0.6±0.0	11.2±0.3	10.5±0.2
Total	89.8±1.9	198.7±8.9	117.0±2.4	86.0±1.1
wt% distribution of monomers*				
Propanol side-chain	76% (78%)	63% (72%)	23% (50%)	37% (58%)
Propane side-chain	7% (8%)	18% (20%)	16% (34%)	16% (25%)
Ethyl side-chain	5% (5%)	2% (3%)	4% (8%)	5% (9%)
Methyl side-chain	7% (7%)	3% (3%)	2% (5%)	4% (6%)
Phenols	3% (3%)	1% (2%)	1% (3%)	2% (3%)
γ-esters	2%	13%	54%	36%
%GVL-lignin became monomers	9.0±0.2%	19.9±0.9%	11.7±0.2%	8.6±0.1%
%GVL-lignin that were esters	0.1±0.1%	1.7±0.1%	6.3±0.2%	3.2±0.1%

* Values in parenthesis correspond to the wt% of the total monolignol derived monomers (excluding γ-esters).

A wide array of solvents have been shown to be compatible in hydrogenolysis and the choice of solvent can alter product distributions (14). As a starting point for integrating hydrogenolysis with microbial funneling, we chose to use methanol due to its low boiling point, low cost, and fairly low toxicity for an organic solvent. Methanol also helps to partition the hydrophobic hydrogenolysis products into the aqueous media. The Pd/C catalyst was used as received from the manufacturer to provide a baseline for

the process. The composition of the crude filtered methanol product solutions formed under our reaction conditions, were analyzed via GC/FID. As expected from previous studies (14, 23), the monomer product distribution heavily favored the production of propanol over the propane or truncated (ethyl or methyl) sidechains, as shown in Table 3.2 (Table 3.S2 shows the product distribution of individual compounds). The maple and poplar had 78% and 72% of the monolignol derived monomers (excluding products from esters units) become propanols, while the grasses showed much lower selectivity with sorghum at 50% and switchgrass at 58% (Table 3.2). These differences in propanol vs. propane/truncated sidechains are masked somewhat by the amount of esters released. In the case of poplar, *p*HBA comprised 12% of the quantified monomers (monolignol derived + ester units). For sorghum and switchgrass, the sums of *p*CA and FA were, 54% and 36% of the quantified monomers, respectively. There is a direct correlation between γ-ester release (content) and the observed decrease in selectivity of aryl-propanol over aryl-propane products, presumably a result of addition of H_2 across the Cγ-O ester bond. Overall, the fraction of aromatic monomers with PDC production potential by the engineered PDC strain was 85%, 79%, 79%, and 77% for maple, poplar, sorghum, and switchgrass, respectively.

3.3.3 Production of PDC from lignin hydrogenolysis products

To test for metabolism and PDC production from these deconstructured biomass fractions, the crude filtered methanol product solution was dissolved in minimal media (See Materials and Methods) and supplemented with glucose to provide a known and constant carbon source for bacterial growth. This media, containing 20 ml/L of the filtered hydrogenolysis product solution, was mixed with a fresh culture of the PDC strain at a ratio of 1:1 v/v to initiate the biocatalysis experiments. An abiotic control was prepared by mixing media containing the hydrogenolysis product with sterile media at a ratio of 1:1 v/v. The composition of the media in the different biocatalysis conditions and the abiotic control was monitored during the experiments using HPLC/MS.

In most cultures, the known aromatic compounds were consumed within the first 6.5 hours of incubation (Figure 3.3), except for cultures supplemented with depolymerized lignin from poplar (Figure 3.3B). In

all cases, PDC accumulated in the culture media, reaching maximum concentrations of 0.12±0.01mM, 0.19±0.01mM, 0.14±0.01mM, and 0.11±0.01mM for maple, poplar, sorghum, and switchgrass, respectively (Figure 3.3).

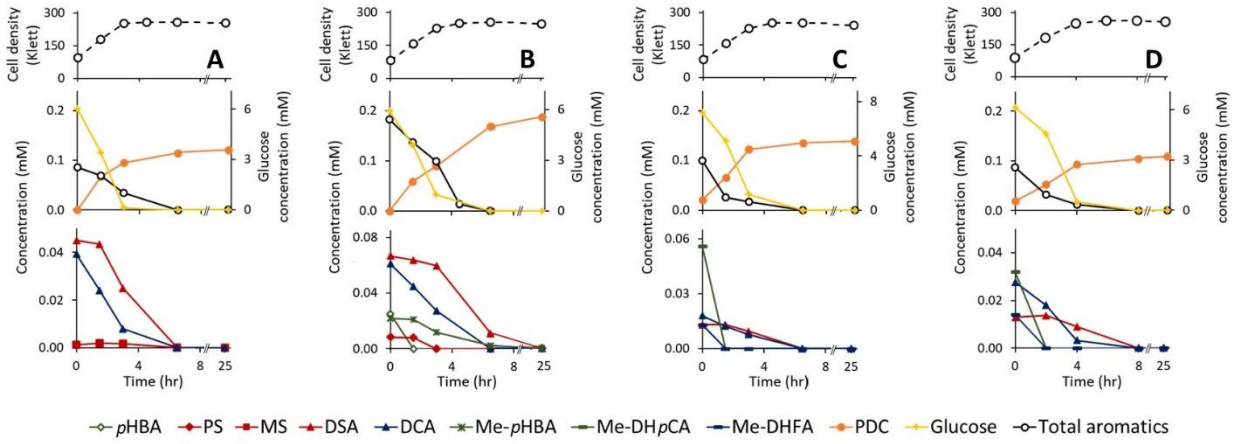


Figure 3.3. Representative *N. aromaticivorans* strain PDC cultures grown in minimal media supplemented with glucose and depolymerized GVL-lignin from maple (A), poplar (B), sorghum (C), and switchgrass (D). (top row of plots) Cell density, (middle row of plots) glucose, total aromatics and PDC production, (bottom row of plots) extracellular metabolite concentration of primary monophenolic products from hydrogenolysis. Each data point corresponds to the average of three biological replicates using the lignin hydrogenolysis products from one of three depolymerization reactions.

The calculated yields of PDC from GVL-lignin were 88.4±3.8, 139.1±0.3, 103.3±4.6, and 79.2±0.4 g PDC/kg GVL-lignin for maple, poplar, sorghum, and switchgrass, respectively (Table 3.3). These yields are higher than the yields that would be predicted based on the concentration of aromatic monomers known to produce PDC that were present in the methanol-soluble fraction (Figure 3.2). Using the observed PDC yields for each individual compound, we anticipated to see yields of 38.0 g PDC/kg GVL-lignin for maple, 86.7 g PDC/kg GVL-lignin for poplar, 53.3 g PDC/kg GVL-lignin for switchgrass, and

85.3 g PDC/kg GVL-lignin for sorghum (Table 3.3). The magnitude of this difference seems to be inversely proportional to the fraction of aromatic monomer products derived from γ-esters (Me-*p*HBA, Me-DH*p*CA, and Me-DHFA). However, when the estimation of PDC production potential is calculated assuming a stoichiometric conversion of aromatic monomers to PDC (for DSA, DCA, Me-*p*HBA, Me-DH*p*CA, DHFA, *p*HBA, MS, and MG), the predicted yields are closer to the observed values (Table 3.S3).

When the GVL lignin extraction process is taken into account, the PDC yields observed were 3.4±0.01, 7.5±0.02, 5.8±0.26, and 6.29±0.03 g PDC/kg whole cell wall (wcw) for maple, poplar, sorghum, and switchgrass, respectively (Table 3.3).

Table 3.3. PDC produced by *N. aromaticivorans* strain PDC from the crude hydrogenolysis product mixture of the four representative biomasses, maple (S/G hardwood), poplar (S/G hardwood with *p*HBA esters), and two grasses sorghum and switchgrass (with different levels of *p*CA and FA esters). Standard error was determined from n=3 technical replicates.

PDC yield	Hard maple	Poplar (NM6)	Energy Sorghum	Switchgrass
g/kg GVL-lignin ± SEM	88.4±3.8	139.1±0.3	103.3±4.6	79.2±0.4
g/kg wcw ± SEM	3.4±0.01	7.5±0.02	5.8±0.26	6.29±0.03
wt% PDC from GVL-lignin	8.8%	13.9%	10.3%	7.9%
wt% from wcw biomass	0.3%	0.8%	0.6%	0.6%
Predicted PDC yield based on monomers quantified end measured individual aromatic conversion yields (g PDC/kg GVL-lignin) *	38.0	95.9	85.3	53.3

* PDC conversion yield of *p*HBA was obtained from Perez et al. (2019) (8).

3.4 Discussion

One approach to make valuable chemicals from lignin is to deconstruct it using chemical depolymerization techniques that produce mixtures of aromatic compounds and subsequently funnel them into a single product using engineered bacteria. In this work, we integrated the GVL-biorefinery producing furfural and dissolving pulp (18), with catalytic hydrogenolysis depolymerization of the high quality lignin fraction, and biological funneling of the hydrogenolysis product mixture to PDC using the engineered PDC strain (Figure 3.1). Through this integration, we identified critical factors that affect the overall yield of PDC from dedicated bioenergy crops in 3 of the main components of the pipeline: catalytic hydrogenolysis, biological funneling, and GVL lignin extraction (Figure 3.1), as discussed below:

Conversion feasibility of aromatic monomers. To assess the compatibility of the hydrogenolysis product with biological funneling to PDC, we explored the feasibility of conversion for the major phenolic products of catalytic hydrogenolysis under conditions that retain the aromatic phenol. Our results show that compounds containing propanol (DSA and DCA), methyl (MS and MG), or carboxylic acid methyl esters (Me-DH*p*CA, Me-DHFA, and Me-*p*HBA), can be metabolized into PDC with variable yields by the engineered strain (Table 3.1, Table 3.S1). Methyl esters (Me-*p*HBA, Me-DH*p*CA, and Me-DHFA) were converted into PDC with yields close to stoichiometric, whereas aromatics with a propanol side-chain (DHSA and DHCA) or methyl side-chain (MS and MG) exhibited lower yields ranging from 28.0% to 66.3% (Table 3.1 and Table 3.S1). In addition, S type compounds were converted into PDC with yields lower than G type compounds. These observations can be explained by the potential existence of competing metabolic pathway that lead to the complete oxidation of the aromatic substrate and is consistent with the lower yields previously reported for S aromatics respect to G aromatics (8). Although the production of PDC suggests that those compounds are metabolized via central aromatic metabolites such as protocatechuic acid (PCA) or 3-methoxygallic acid (3-MGA) (7, 8), the data obtained does not provide insight about the mechanisms involved in the side-chain processing.

On the other hand, compounds with a propyl side-chain (PS and PG) and ethyl side chain (ES and EG), are also consumed, no detectable PDC is produced, but in all cases we observed the formation of other compounds by HPLC-UV. For example, in the experiments with ES and EG, *N. aromaticivorans* strain PDC consumes the aromatic substrate in 18 hours of experiment and new peaks with retention times of 4.79 min and 3.49 min were detected, respectively (Figure 3.S2). We speculate that *N. aromaticivorans* can partially transform the side-chain but is unable to metabolize them into central aromatic compounds such as 3-MGA or PCA. We did not pursue the identification of these compounds nor the mechanisms involved in this study. A potential strategy to broaden the engineered bacterium's substrate specificity is to engineer foreign metabolic pathways for side-chain processing to produce central aromatic metabolites such as 3-MGA and PCA (6).

Compounds without a side-chain (Sy and Gu) were completely or partially consumed but no PDC production was observed. Although metabolism of these compounds have not been investigated before, Gu metabolism via catechol has been studied in *Pseudomonas putida* where they are used as a substrate for the production of muconic acid via biological funneling (32). In this organism and others, Gu is metabolized via a route that does not involve the production of PDC. Also, protein engineering of enzymes capable of *O*-demethylation of Sy has enabled *P. putida* to metabolize this compound (33). We speculate that *N. aromaticivorans* can catabolize Gu and Sy by promiscuous enzymes with higher affinity to Gu than Sy and that their metabolic pathway does not involve the production of PDC. A potential strategy to enable *N. aromaticivorans* to convert Sy and Gu into PDC could involve enzyme engineering of *O*-demethylases and the introduction of foreign enzymes for the production of central aromatic metabolites via carboxylation of the aromatic ring to form 3-MGA or PCA. However to our knowledge, only anaerobic mechanisms for aromatic carboxylation have been described so far and potential aerobic aromatic decarboxylases would need high concentrations of CO_2 to work in reverse (34).

Catalytic hydrogenolysis. As the screening for PDC production from different types of compounds formed by catalytic hydrogenolysis of lignins showed that only some of them (DHSA, DHCA, MS, MG,

and the phenolic esters) are compatible with biological funneling using *N. aromaticivorans* strain PDC, the choice of catalysts was narrowed down to those that favor production of the propanol side chain over the fully reduced propyl side-chain. As mentioned previously, this list includes Pd/C (23-26), Rh/C (27), Ni/C (25) and Ni/Al2O3 (35). To keep the results of this study as reproducible as possibly for future optimization studies, we chose to use Pd/C from a commercially available source. Pd/C is a very robust and common catalyst that is only slightly air and light sensitive, unlike some of the more engineered catalysts. We chose to use methanol as the solvent as it is fairly inexpensive, has a low toxicity for organic solvents, and has shown to produce mostly the propanol side-chain when used in combination with Pd/C, H_2 at 30 bar, and 200 °C (20). The analysis of the crude filtered methanol product solution showed that, as the literature studies indicated, for all lignin sources the majority of the identified monomeric products obtained were compounds that can be funneled into PDC (Table 3.1), although a minor but still significant fraction of the products correspond to compounds that *N. aromaticivorans* strain PDC is unable to convert into PDC (Table 3.2). Future efforts to increase the compatibility of catalytic hydrogenolysis with biological funneling can include increasing the product selectivity of the depolymerization process towards "convertible" compounds.

Biological funneling. For all biomass sources, the engineered strain was able to produce PDC from deconstructed GVL-lignin showing their compatibility with the proposed valorization method. The overall efficiency of producing PDC from using the hydrogenolysis products of GVL-lignin was calculated to be greatest for poplar (139.1±0.3 g PDC/kg GVL lignin) followed by sorghum (103.3±4.6 g PDC/kg GVL lignin), then maple (88.4±3.8 g PDC/kg GVL lignin) and switchgrass (79.2±0.4 g PDC/kg GVL lignin) (Table 3.3). This trend was correlated to the yield of monomers from the respective GVL-lignins depolymerized by catalytic hydrogenolysis; poplar (19.9%), sorghum (11.7%), maple (9.0%), and switchgrass (8.6%). The chemical composition of the GVL-lignins positively couples to the monomer yield trend to further increase PDC conversion efficiencies. Poplar lignin contains high levels of *p*HBA

(12 wt% of the lignin) (29), whereas sorghum and switchgrass have significant quantities of *p*CA (10 wt% of the lignin).

Notably, in all cases the PDC yield was higher than predicted when the amount of different aromatic compounds and the observed conversion yields of individual compounds was considered (Table 3.3). This observation can be explained by three different but not exclusive hypothesis. First, the conversion yields for each compound in the reaction vessel could be higher than the yields observed in the single compound experiments. The concentrations for each of the lignin derived aromatics quantified in the reaction vessels were all below 0.08 mM, whereas in the single compound experiments the concentration of aromatic compounds were more than 10 times higher. This difference in concentration could affect the expression of alternative aromatic degradation pathways that do not involve the production of PDC as it has been hypothesized that *N. aromaticivorans* can utilize more than one route for the degradation of the central aromatic metabolites 3-MGA and PCA (8). Second, it is possible that the catalytic hydrogenolysis process produces additional monomers that were not identified and quantified and *N. aromaticivorans* strain PDC was able to convert them into PDC. In fact, when the prediction of PDC yield is based in the assumption that all convertible monomers identified in the depolymerized lignin methanol solution are converted with stoichiometric yields (Table 3.S3), the observed PDC yield is closer to the prediction but still higher for maple, sorghum, and switchgrass, although not for poplar. Third, it is possible that aromatic substrates other than monomeric compounds present in the depolymerized lignin mixture were converted into PDC. Previous studies have demonstrated that *N. aromaticivorans* can metabolize aromatic dimers, for example those interliked via β-O-4 bonds (36-38).

GVL extraction. When the GVL-lignin extraction process is also incorporated into the yield calculation, the order changes significantly. Poplar was still the most efficient at 7.5±0.02 g PDC/kg biomass followed by switchgrass (6.29±0.03 g PDC/kg biomass) and sorghum (5.8±0.26 g PDC/kg biomass), and then maple (3.4±0.01 g PDC/kg biomass). The lower PDC yield from maple can be explained by the lack of phenolic ester pendent groups on its lignin compared to the other biomass sources tested. Genetic

modifications that target amount of ester pendant groups incorporated into the lignin (39-41) could have a positive impact in the PDC yield. In addition, these yield values are strongly affected by the GVL extraction efficiencies (maple, 15.2%; poplar, 22.3%; sorghum, 35.0%; switchgrass, 41.0%), suggesting that improvements in lignin extraction can have a significant impact in the overall PDC yield. A deconstruction approach that incorporates lignin depolymerization from whole cell wall biomass such as RCF (42) would allow the whole lignin fraction to be depolymerized without the losses from the extraction process.

3.5 Conclusions

In this study, we demonstrated the feasibility of integrating reductive chemical depolymerization of isolated lignin with upgrade to PDC via biological funneling by an engineered strain of *N. aromaticivorans* (strain PDC). We also demonstrated that the strategy is technically feasible for lignin of multiple sources. The selectivity for chemical depolymerization products that can be metabolized by the selected engineered bacterium is a critical factor that determines which depolymerization process and specific operation conditions of those processes that have the potential to become a successful strategy. Finally, understanding the microbial mechanisms for metabolism of aromatic compounds could provide valuable tools to improve other potential bacterial platforms for biological funneling of depolymerized lignin and production of high value commodity chemicals.

3.6 Materials and methods

Chemicals. All chemicals were purchased from Sigma-Aldrich (St Louis, MO). PDC was produced and purified according to the procedure described in J. M. Perez *et al*. (8).

Bacterial strains and culture conditions. The engineered strain of *Novosphingobium aromaticivorans* DSM12444 (strain PDC) lacking the genes Saro_1879 (*sacB*), Saro_2819 (*ligI*), and Saro_2864/5 (*desC/desD*) (8), was used in this study. Cultures were grown in SMB media supplemented with the indicated carbon source at 30 °C. SMB media contains 20 mM Na_2HPO_4, 20 mM KH_2PO_4, 7.5 mM $(NH_4)_2SO_4$, 0.167 mM $ZnSO_4$, 0.125 mM $FeSO_4$, 0.028 mM $MnSO_4$, 0.006 mM $CuSO_4$, 0.009 mM $Co(NO_3)_2$, 0.016 mM $Na_2B_4O_7$, 24.319 mM $MgSO_4$, 1.667 mM $CaCl_2$, 0.013 mM $(NH_4)_6Mo_7O_{24}$. For routine culture and storage, the growth media was supplemented with 1 g/L glucose.

Bioconversion of single compounds. *N. aromaticivorans* strain PDC cultures were grown 10 ml of SMB media supplemented with 1 g/L glucose to reach stationary phase. The experiments were initiated by diluting the cultures with 10 ml of fresh SMB media supplemented with 2 g/L glucose and the corresponding aromatic compounds (dissolved in pure methanol), to a concentration of 2 mM. The final concentration of methanol in the growth media was less than 0.5 %. Triplicate cultures were grown in flasks, shaken at 200 rpm, and incubated at 30 °C. Bacterial cell density was monitored using a Klett-Summerson photoelectric colorimeter with a red filter. Samples were collected regularly, filtered with a 0.2 μm PES syringe filter and stored at -20 °C. The experiments were terminated after 20 hours of reaction. Each compound was tested in three biological replicates.

GVL-lignin extraction. GVL-Lignin was extracted from ground biomass (17, 43). Briefly, 185 g of biomass was mixed with 1665 g of a solvent system comprising 80 wt% GVL, 19 wt% water, and 100 mM sulfuric acid. In a typical run, this mixture was treated at 90 °C for 90 min in a high-pressure reactor fitted with a custom-designed impeller system capable of mixing biomass at high solid concentration (44). Under these reaction conditions, a portion of the lignin was partially depolymerized and dissolved. The

liquid was cooled to room temperature and the solid fraction was removed by filtration. The GVL-lignin was precipitated by diluting the filtrate with water (1:9 v/v, filtrate:water). The suspension was left to settle for two days, then the GVL-lignin was pelleted by centrifugation and the supernatant decanted. The pellet was washed three times with boiling DI water, by suspending the solids and filtration. The wet solids were freeze-dried to obtain dry lignins that were used in the hydrogenolysis depolymerization without further purification.

Hydrogenolysis of GVL-lignin. The GVL-lignin was deconstructed by hydrogenolysis using hydrogen over palladium on carbon (5 wt% Pd/C) catalyst in methanol. In a 50 mL Hastelloy Parr reactor, equipped with a mechanical stirrer and heating mantle, 750 mg of GVL-lignin, 375 mg Pd/C, 30 mL methanol, and 9 mg (65 nmol) 1,2-dimethoxybenzene (DMB, an internal standard for determining monomer yields) were added. The reactor was sealed, purged and pressurized with hydrogen gas to 30 bar. The reaction vessel was then heated to 200 °C at a ramp rate of 6 °C/min and held there for 3 h (200 °C, 65 bar). Once complete the heating mantle was removed and the reactor was rapidly cooled back to room temperature using compressed air stream to assist in cooling. Once at room temperature, the reaction vessel was slowly depressurized to minimize the loss of volatile product. The catalyst and any residual solids were removed by a 0.1 µm PTFE (polytetrafluoroethylene) filter. The product mixture was stored at -5 °C for until it was used in the biocatalyst experiments. This mixture provided the same GC-FID determined composition and PDC production from microbial digestion after >6 months storage in the freezer.

Note: The DMB internal standard is similar to guaiacol and syringol in that it is not able to be converted to PDC through the microbial funneling.

Analysis of hydrogenolysis aromatic monomers. Quantitative analysis of the aromatic monomers was performed on a Shimadzu GC-2010Plus equipped with an AOC-20i autosampler and a Phenomenex Xebron column (ZB-5HT 15 mx0.25 mmx0.25µm), helium gas mobile phase held at a constant linear velocity of 35 cm/sec, and hydrogen/compressed air FID gases. The injection port was set to 250 °C with a split ratio of 10:1. The temperature program was: 50 °C for 0.5 min, then ramped to 100 °C at 20

°C/min, the ramped to 200 °C at 5 °C/min, the ramped to 305 °C at 30 °C/min, and held at 305 °C for 3 min. Calibration curves for the quantified monomers were determined by a seven point calibration curve relative to DMB. The calculated values were determined as: mmol/mg GVL-lignin relative to 100% recovery of 65 nmol DMB.

Depolymerized lignin bioconversion experiments. *N. aromaticivorans* strain PDC cultures were grown 10 ml of SMB media supplemented with 6 mM glucose to reach stationary phase. The experiments were initiated by diluting the cultures with 10 ml of fresh SMB media supplemented with 12 mM glucose and 20 ml/L methanol solution containing the corresponding lignin depolymerization products at a concentration of 250 mg/L. The cultures were grown in triplicate in flasks, shaken at 200 rpm, and incubated at 30 °C. Bacterial cell density was monitored using a Klett-Summerson photoelectric colorimeter with a red filter. Samples were collected regularly, filtered with a 0.2 μm PES syringe filter and stored at -20 °C. The experiments were terminated after 25 hours of reaction. Each of the three hydrogenolysis products obtained from each plant tissue was tested in three biological replicates.

Analysis of extracellular metabolites. Quantitative analysis of extracellular aromatic compounds and PDC was performed on a Shimadzu triple quadrupole liquid chromatography mass spectrometer (LC-MS, Nexera XR HPLC-8045 MS/MS). The mobile phase was a binary gradient consisting of solvent A (0.2% formic acid in water) and solvent B (methanol) at a flow rate of 0.5 ml/min. The column was conditioned at 5% B, the elution program was 5% B hold 0.1 min, ramp to 20% B at 0.5 min, ramp to 30% B at 3.5 min, ramp to 50% B at 5 min, ramp to 95% B at 5 min and hold for 1.5 min to wash the column, then reset the column by returning to 5% B at 7 min and holding for 2.5 min to equilibrate the column for the next injection. The stationary phase was a Kinetex F5 column (Phemonenex, 2.6 μm pore size, 2.1 mm ID, 150 mm length, P/N: 00F-4723-AN). All compounds were detected by multiple-reaction-monitoring (MRM) and quantified using the strongest MRM transition, Table 3.S4.

Quantitative analysis of glucose was performed on an Agilent 1260 infinity HPLC equipped with a refractive index detector (HPLC-RID) (Agilent Technologies, Inc., Palo Alto, CA) and an Aminex HPX-

87H with Cation-H guard column (BioRad, Inc. Hercules, CA). The mobile phase was 0.02 N sulfuric acid at a flow rate of 0.5 ml/min.

3.7 Acknowledgements

This work was supported by U.S. Department of Energy (DOE) Great Lakes Bioenergy Research Center grant (DOE Office of Science BER DE-SC0018409). Additional funding from the Chilean National Commission for Scientific and Technological Research (CONICYT) as a fellowship to Jose M. Perez is also acknowledged.

3.8 Supplementary figures and tables

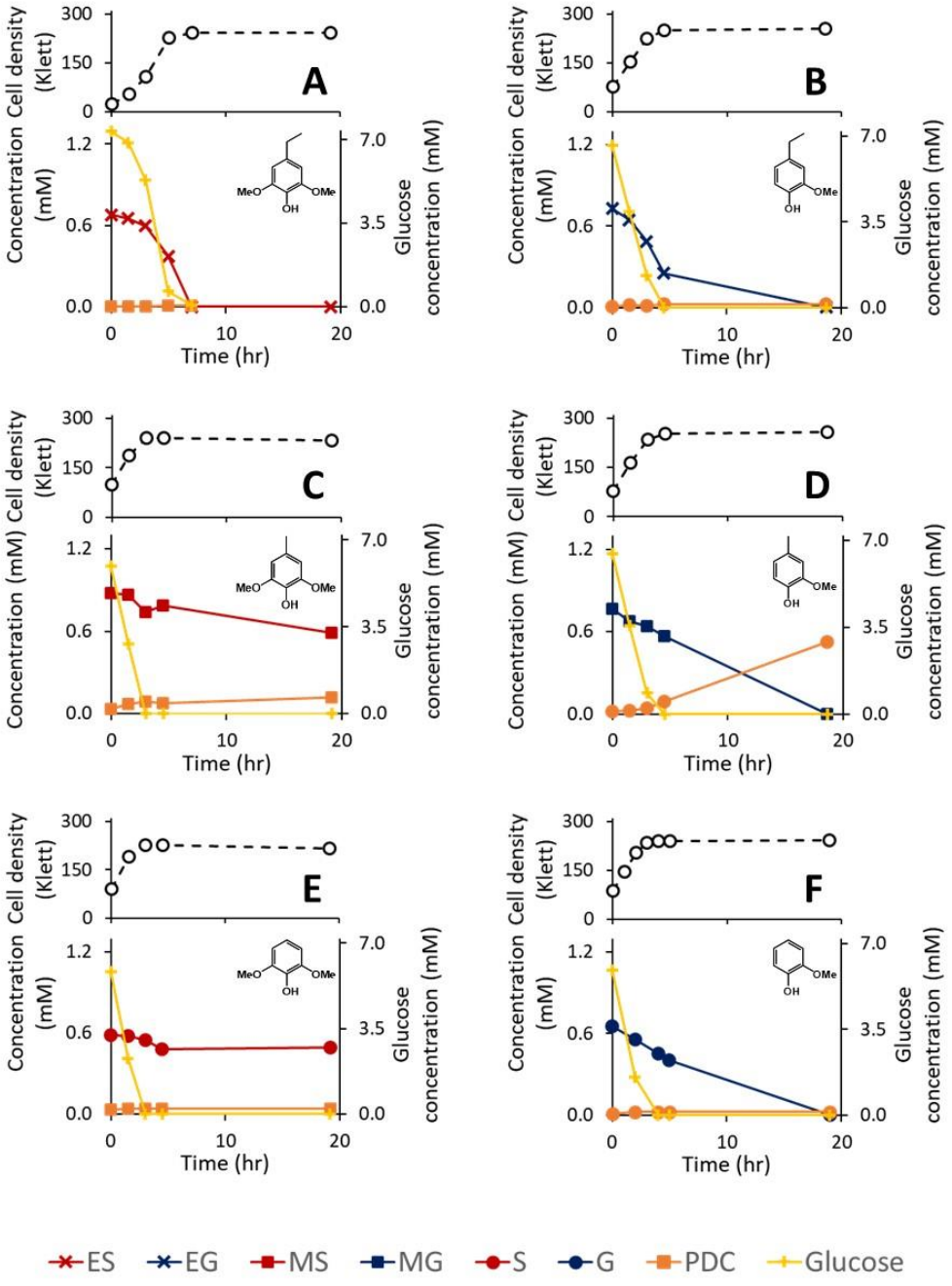


Figure 3.S1. Cell density and extracellular metabolite concentration of *N. aromaticivorans* strain PDC cultures in minimum media supplemented with glucose and ethylsyringol (ES) (A), ethylguaiacol (EG) (B), syringol (S) (C), or guaiacol (G) (D). Values correspond to the average of three biological replicates.

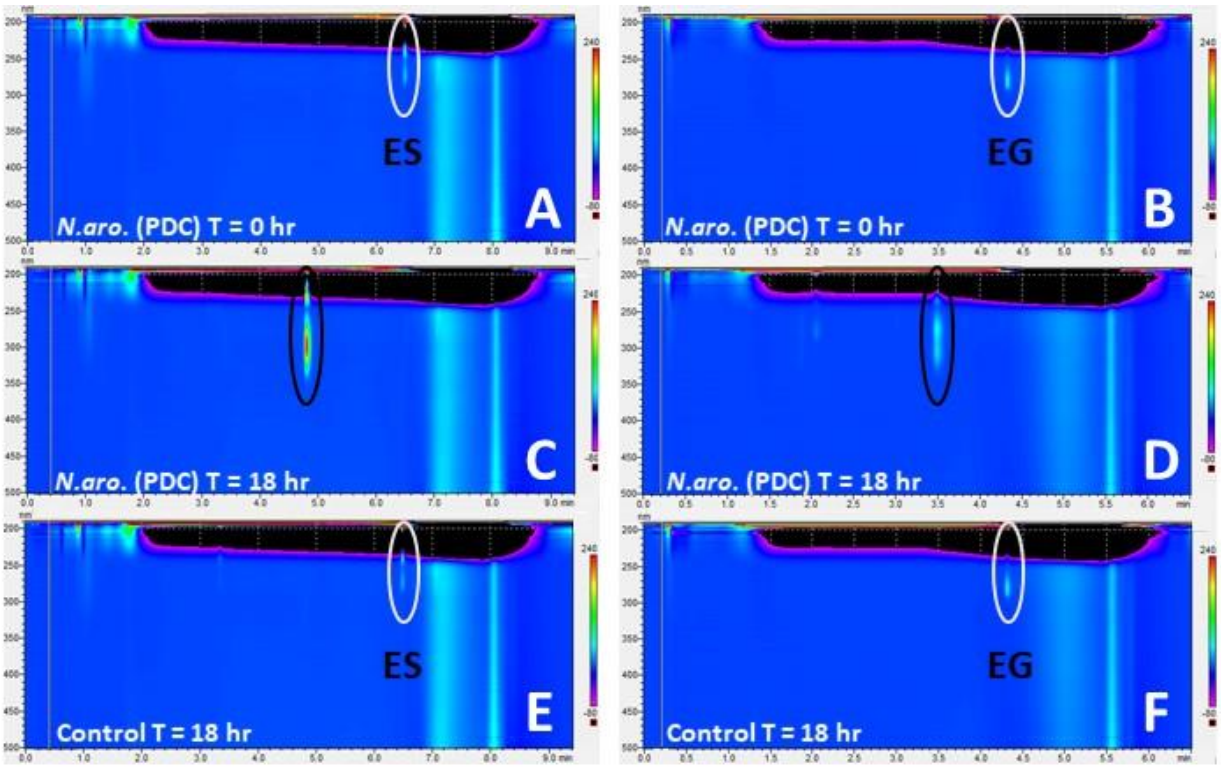


Figure 3.S2. UV/vis chromatogram of samples from experiments with ES and EG collected at the beginning of the experiment (A and B, respectively), after 18 hours (C and D), and abiotic control after 18 hours of incubation (E and F). ES and EG disappear after 18 hours of experiment with *N. aromaticivorans* strain PDC and new peaks are detected a RT = 4.79 min (C) and TR = 3.49 min (D). Abiotic controls show the presence of ES (E) and EG (F) and the absence of new peaks.

Table 3.S1. Yield of PDC (mol) at 19-24 h reaction time from experiments with syringol (Sy), methylsyringol (MS), ethylsyringol (ES), guaiacol (Gu), methylguaiacol (MG), or ethylguaiacol (EG). Standard error was determined from n=3 technical replicates.

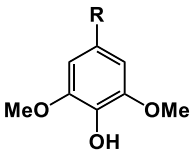
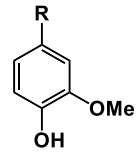
	Compound	R	PDC molar yield (%)
	Ethyl syringol (ES)		0.9% ± 0.0%
	Methyl syringol (MS)		28.0% ± 2.7%
	Syringol (Sy)	H	5.4% ± 3.5%
	Ethyl guaiacol (EG)		1.7% ± 1.1%
	Methyl guaiacol (MG)		66.3 % ± 2.6 %
	Guaiacol (Gu)	H	2.3% ± 1.6%

Table 3.S2. Ratio of quantified hydrogenolysis monomers determined by GC-FID analysis of the crude filtered methanol product solution.

Monomer Ratio	Hard maple	Poplar (NM6)	Energy Sorghum	Switchgrass
Propylphenol	4%	4%	9%	9%
Guaiacol	1%	0%	1%	1%
Methylguaiacol	4%	1%	2%	3%
Ethylguaiacol	4%	2%	4%	5%
Propylguaiacol	2%	1%	2%	2%
Syringol	2%	1%	1%	1%
Methylsyringol	4%	2%	1%	1%
Ethylsyringol				
Propylsyringol	2%	12%	5%	5%
7,8-dihydro-*p*-hydroxycinnamyl alcohol	24%	12%	3%	6%
7,8-dihydrosinapyl alcohol	26%	28%	8%	10%
7,8-dihydroconiferyl alcohol	26%	24%	12%	21%
Methyl *p*-hydroxybenzoate	1%	12%	4%	2%
Methyl 7,8-dihydro-*p*-coumarate	1%	0%	43%	24%
Methyl 7,8-dihydroferulate	0%	0%	7%	9%

Table 3.S3. Predicted PDC yield based on convertible monomers quantified after hydrogenolysis assuming a stoichiometric conversion of convertible monomers (DSA, DCA, Me-*p*HBA, Me-DH*p*CA, MeDHFA, *p*HBA, MS, and MG).

PDC yield	Hard maple	Poplar (NM6)	Energy Sorghum	Switchgrass
Predicted PDC yield based on monomers quantified (g/kg GVL-lignin)	70.4	159.3	99.3	68.8

Table 3.S4. MRM transitions used to quantify aromatic monomers in the *N. aromaticivorans* bioconversion assays.

Mass Spectrometer	**Shimadzu LCMS-8045**									
ESI Source	300ºC Ionization mode: DUIS ESI/APCI									
Operation Mode	Multiple Reaction Monitoring (MRM) Argon gas, 230 kPa									
MRM Transition Details Positive (+) mode										
Compound Name	**Retention Time**	**Transition 1**	**CE 1**	**Rel. int.**	**Transition 2**	**CE 2**	**Rel. int.**	**Transition 3**	**CE 3**	**Rel. int.**
Guaiacol (G)	4.84 min	125>65	-21	100	125>93	-15	75			
7,8-dihydroconiferyl alcohol (DCA)	4.92 min	183>137	-13	100	183>133	-12	45	183>105	-20	45
7,8-dihydrosinapyl alcohol (DSA)	5.30 min	213>195	-20	100	213>167	-13	81	213>107	-6	46
Syringol (S)	5.49 min	155>123	-23	144	155>95	-30	100	155>77	-45	72
p-hydroxybenzoic acid methyl ester (Me-*p*HBA)	6.09 min	153>121	-20	100	153>93	-25	48	153>65	-32	62
Methylguaiacol (MG)	6.33 min	139>107	-15	100	139>79	-20	87	139>77	-26	55
7,8-dihydro-*p*-coumaric acid methyl ester (Me-DH*p*CA)	6.4 min	181>107	-26	100	181>149	-10	0			
Methylsyringol (MS)	6.58 min	169>137	-26	100	169>109	-32	80	169>91	-40	72
Ethylsyringol (ES)	6.61 min	183>65	-32	100	183>95	-35	90	183>91	-14	88
7,8-dihydro-ferulic acid methyl ester (Me-DHFA)	6.62 min	211>137	-20	100	211>179	-9	48	211>122	-34	22
Ethylguaiacol (EG)	6.88 min	153>77	-26	100	153>121	-15	97	153>91	-23	46
Propylguaiacol (PG)	7.05 min	167>125	-14	100	167>107	-16	184	167>135	-13	99
Propylsyringol (PS)	7.05 min	197>123	-18	100	197>79	-47	31	197>107	34	26
MRM Transition Details Negative (-) mode										
PDC	1.44 min	183>111	13	100	183>139	11	95	183>95	11	67
p-hydroxybenzoic acid (*p*HBA)	2.68 min	137>93	14	100	137>65	29	10			

3.9 References

1. Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon RA, Gilna P, Keller M, Langan P, Naskar AK, Saddler JN, Tschaplinski TJ, Tuskan GA, Wyman CE. 2014. Lignin valorization: Improving lignin processing in the biorefinery. Science 344.

2. Zakzeski J, Bruijnincx PCA, Jongerius AL, Weckhuysen BM. 2010. The Catalytic Valorization of Lignin for the Production of Renewable Chemicals. Chem Rev 110:3552-3599.

3. Tsvetkov M, V., Salganskii E, A. 2018. Lignin: Applications and Ways of Utilization (Review). Russ J Appl Chem 91:1129-1136.

4. Davis R, Tao L, Tan ECD, Biddy MJ, Beckham GT, Scarlata C, Jacobson J, Cafferty K, Ross J, Lukas J, Knorr D, Schoen P. 2013. Lignocellulosic biomass to hydrocarbons: Dilute-acid and enzymatic deconstruction of biomass to sugars and biological conversion of sugars to hydrocarbons. NREL Tech Rep doi:10.2172/110.470:88-101.

5. Corona A, Biddy MJ, Vardon DR, Birkved M, Hauschild MZ, Beckham GT. 2018. Life cycle assessment of adipic acid production from lignin. Green Chem 20:3857-3866.

6. Beckham GT, Johnson CW, Karp EM, Salvachua D, Vardon DR. 2016. Opportunities and challenges in biological lignin valorization. Curr Opin Biotechnol 42:40-53.

7. Cecil JH, Garcia DC, Giannone RJ, Michener JK. 2018. Rapid, parallel identification of pathways for catabolism of lignin-derived aromatic compounds in Novosphingobium aromaticivorans. Appl Environ Microbiol 84:1-13.

8. Perez JM, Kontur WS, Alherech M, Coplien J, Karlen SD, Stahl SS, Donohue TJ, Noguera DR. 2019. Funneling aromatic products of chemically depolymerized lignin into 2-pyrone-4-6-dicarboxylic acid with Novosphingobium aromaticivorans. Green Chem 21:1340-1350.

9. Shikinaka K, Otsuka Y, Nakamura M, Masai E, Katayama Y. 2018. Utilization of lignocellulosic biomass via novel sustainable process. J Oleo Sci 67:1059-1070.

10. Otsuka Y, Nakamura M, Shigehara K, Sugimura K, Masai E, Ohara S, Katayama Y. 2006. Efficient production of 2-pyrone 4,6-dicarboxylic acid as a novel polymer-based material from protocatechuate by microbial function. Appl Microbiol Biotechnol 71:608-614.

11. Qian Y, Otsuka Y, Sonoki T, Mukhopadhyay B, Nakamura M, Jellison J, Goodell B. 2016. Engineered microbial production of 2-pyrone-4,6-dicarboxylic acid from lignin residues for use as an industrial platform chemical. BioResources 11:6097-6109.

12. Suzuki Y, Okamura-Abe Y, Nakamura M, Otsuka Y, Araki T, Otsuka H, Navarro RR, Kamimura N, Masai E, Katayama Y. 2020. Development of the production of 2-pyrone-4,6-dicarboxylic acid from lignin extracts, which are industrially formed as by-products, as raw materials. J Biosci Bioeng doi:10.1016/j.jbiosc.2020.02.002.

13. Nakajima M, Nishino Y, Tamura M, Mase K, Masai E, Otsuka Y, Nakamura M, Sato K, Fukuda M, Shigehara K, Ohara S, Katayama Y, Kajita S. 2009. Microbial conversion of glucose to a novel chemical building block, 2-pyrone-4,6-dicarboxylic acid. Metab Eng 11:213-220.

14. Sun Z, Fridrich B, de Santi A, Elangovan S, Barta K. 2018. Bright Side of Lignin Depolymerization: Toward New Platform Chemicals. Chem Rev 118:614-678.

15. Rinaldi R, Jastrzebski R, Clough MT, Ralph J, Kennema M, Bruijnincx PCA, Weckhuysen BM. 2016. Paving the way for lignin valorization: Recent advances in bioengineering, biorefining and catalysis. Angew Chem, Int Ed Engl 55:8164-8215.

16. Galkin MV, Samec JSM. 2016. Lignin valorization through catalytic lignocellulose fractionation: A fundamental platform for the future biorefinery. ChemSusChem 9:1544-1558.

17. Luterbacher JS, Azarpira A, Motagamwala AH, Lu F, Ralph J, Dumesic JA. 2015. Lignin monomer production integrated into the γ-valerolactone sugar platform. Energy Environ Sci 8:2657-2663.

18. Alonso DM, Hakim SH, Zhou S, Won W, Hosseinaei O, Tao J, Garcia-Negron V, Motagamwala AH, Mellmer MA, Huang K, Houtman CJ, Labbé N, Harper DP, Maravelias C, Runge T, Dumesic JA. 2017. Increasing the revenue from lignocellulosic biomass: Maximizing feedstock utilization. Sci Adv 3:e1603301.

19. Luterbacher JS, M. RJ, Alonso DM, Han J, Youngquist JT, Maravelias CT, Pfleger BF, Dumesic JA. 2014. Nonenzymatic sugar production from biomass using biomass-derived γ-valerolactone. Science 343:277-280.

20. Schutyser W, Renders T, Van den Bosch S, Koelewijn SF, Beckham GT, Sels BF. 2018. Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. Chem Soc Rev 47:852-908.

21. Harris EE, D'Ianni J, Adkins H. 1938. Reaction of hardwood lignin with hydrogen. J Am Chem Soc 60:1467-1470.

22. Weitkamp J. 2012. Catalytic hydrocracking—Mechanisms and versatility of the process. ChemCatChem 4:292-306.

23. Torr KM, van de Pas DJ, Cazeils E, Suckling ID. 2011. Mild hydrogenolysis of in-situ and isolated Pinus radiata lignins. Bioresour Technol 102:7608-7611.

24. Van den Bosch S, Schutyser W, Koelewijn S-F, Renders T, Courtin CM, Sels BF. 2015. Tuning the lignin oil OH-content with Ru and Pd catalysts during lignin hydrogenolysis on birch wood. Chem Commun (Cambridge, U K) 51.

25. Renders T, Van den Bosch S, Vangeel T, Ennaert T, Koelewijn S-F, Van den Bossche G, Courtin CM, Schutyser W, Sels BF. 2016. Synergetic effects of alcohol/water mixing on the catalytic reductive fractionation of poplar wood. ACS Sustainable Chem Eng 4:6894-6904.

26. Schutyser W, Van den Bosch S, Renders T, De Boe T, Koelewijn S-F, Dewaele A, Ennaert T, Verkinderen O, Goderis B, Courtin CM, Sels BF. 2015. Influence of bio-based solvents on the catalytic reductive fractionation of birch wood. Green Chem 17:5035-5045.

27. Pepper JM, Supathna P. 1978. Lignin and related compounds. VI. A study of variables affecting the hydrogenolysis of spruce wood lignin using a rhodium-on-charcoal catalyst. Can J Chem 56:899-902.

28. Nimz HH, Robert D, Faix O, Nemr M. 1981. Carbon-13 NMR spectra of lignins, 8. Structural differences between lignins of hardwoods, softwoods, grasses and compression wood. Holzforschung 35:0018-383016-26.

29. Smith DCC. 1955. p-Hydroxybenzoate groups in the lignin of aspen (populus tremula). J Chem Soc doi:10.1039/JR9550002347:2347-2351.

30. Bhalla A, Bansal N, Pattathil S, Li M, Shen W, Particka CA, Karlen SD, Phongpreecha T, Semaan RR, Gonzales-Vigil E, Ralph J, Mansfield SD, Ding S-Y, Hodge DB, Hegg EL. 2018. Engineered lignin in poplar biomass facilitates Cu-catalyzed alkaline-oxidative pretreatment. ACS Sustainable Chem Eng 6:2932-2941.

31. Hatfield RD, Marita JM, Frost KM, Grabber JM, Ralph J, Lu FM, Kim HM. 2009. Grass lignin acylation: p-coumaroyl transferase activity and cell wall characteristics of C3 and C4 grasses. Planta 229:1253-1267.

32. Vardon D, Franden MA, Johnson C, Karp E, Guarnieri M, Linger J, Salm M, Strathmann T, Beckham G, Ferguson G. 2015. Adipic acid production from lignin. Energy Environ Sci 8:617-628.

33. Machovina MM, Mallinson SJB, Knott BC, Meyers AW, Garcia-Borràs M, Bu L, Gado JE, Oliver A, Schmidt GP, Hinchen DJ, Crowley MF, Johnson CW, Neidle EL, Payne CM, Houk KN, Beckham GT, McGeehan JE, DuBois JL. 2019. Enabling microbial syringol conversion through structure-guided protein engineering. Proc Natl Acad Sci 116:13970-13976.

34. Payer SE, Marshall SA, Bärland N, Sheng X, Reiter T, Dordic A, Steinkellner G, Wuensch C, Kaltwasser S, Fisher K, Rigby SEJ, Macheroux P, Vonck J, Gruber K, Faber K, Himo F, Leys D, Pavkov-Keller T, Glueck SM. 2017. Regioselective para-carboxylation of catechols with a prenylated flavin dependent decarboxylase. Angew Chem, Int Ed Engl 56:13893-13897.

35. Van den Bosch S, Renders T, Kennis S, Koelewijn S, Van den Bossche G, Vangeel T, Deneyer A, Depuydt D, Courtin C, Thevelein J, Schutyser W, Sels B. 2017. Integrating lignin valorization and bio-ethanol production: on the role of Ni-Al2O3 catalyst pellets during lignin-first fractionation. 19:3313-3326.

36. Gall DL, Ralph J, Donohue T, Noguera D. 2014. A Group of Sequence-Related Sphingomonad Enzymes Catalyzes Cleavage of beta-Aryl Ether Linkages in Lignin beta-Guaiacyl and beta-Syringyl Ether Dimers. Environ Sci Technol 48:12454-12463.

37. Kontur WS, Bingman CA, Olmsted CN, Wassarman DR, Ulbrich A, Gall DL, Smith RW, Yusko LM, Fox BG, Noguera DR, Coon JJ, Donohue TJ. 2018. Novosphingobium aromaticivorans uses a Nu-class glutathione S -transferase as a glutathione lyase in breaking the β-aryl ether bond of lignin. J Biol Chem 293:4955-4968.

38. Kontur WS, Olmsted CN, Yusko LM, Niles AV, Walters KA, Beebe ET, Vander Meulen KA, Karlen SD, Gall DL, Noguera DR, Donohue TJ. 2019. A heterodimeric glutathione S-transferase that stereospecifically breaks lignin's β(R)-aryl ether bond reveals the diversity of bacterial β-etherases. J Biol Chem 294:1877-1890.

39. Withers S, Lu F, Kim H, Zhu Y, Ralph J, Wilkerson CG. 2012. Identification of grass-specific enzyme that acylates monolignols with p-coumarate. J Biol Chem 287:8347-8355.

40. Smith RA, Gonzales-Vigil E, Karlen SD, Park J-Y, Lu F, Wilkerson CG, Samuels L, Ralph J, Mansfield SD. 2015. Engineering monolignol p-Coumarate conjugates into poplar and arabidopsis lignins. Plant Physiol 169:2292-3001.

41. Petrik DL, Karlen SD, Cass CL, Padmakshan D, Lu F, Liu S, Le Bris P, Antelme S, Santoro N, Wilkerson CG, Sibout R, Lapierre C, Ralph J, Sedbrook JC. 2014. p-Coumaroyl-CoA:monolignol transferase (PMT) acts specifically in the lignin biosynthetic pathway in Brachypodium distachyon. Plant J 77:713-726.

42. Van den Bosch S, Schutyser W, Vanholme R, Driessen T, Koelewijn S-F, Renders T, De Meester B, Huijgen WJJ, Dehaen W, Courtin CM, Lagrain B, Boerjan W, Sels BF. 2015. Reductive lignocellulose fractionation into soluble lignin-derived phenolic monomers and dimers and processable carbohydrate pulps. Energy Environ Sci 8.

43. Motagamwala AH, Won W, Maravelias CT, Dumesic JA. 2016. An engineered solvent system for sugar production from lignocellulosic biomass using biomass derived γ-valerolactone. Green Chem 18:5756-5763.

44. Klingenberg DJ, Root TW, Burlawar S, Scott CT, Bourne KJ, Gleisner R, Houtman C, Subramaniam V. 2017. Rheometry of coarse biomass at high temperature and pressure. Biomass Bioenergy 99:69-78.

4. Functional pathway redundancy in the metabolism of plant-derived phenolics by *Novosphingobium aromaticivorans*

A version of this chapter has been prepared for submission to a peer-reviewed journal. The contribution of each co-author is as follows:

Jose M. Perez: Conceived the project, created deletion mutant strains, expressed and purified recombinant proteins, planned and performed the experiments, performed HPLC-MS analysis, wrote the manuscript

Wayne S. Kontur: Conceived the project, created deletion mutant strains, expressed and purified recombinant proteins, wrote the manuscript.

Carson Gehl: Created gene deletion mutant strains, performed the experiments

Derek M. Gille: Created gene deletion mutant strains, expressed and purified recombinant proteins

Yanjun Ma: Created gene deletion mutant strains

Alyssa V. Niles: Expressed and purified recombinant proteins

German Umana: Performed the experiments

Timothy J. Donohue: Wrote the manuscript

Daniel R. Noguera: Wrote the manuscript

4.1 Abstract

Lignin is one of the major components of plant biomass and is a heteropolymer composed of phenolic subunits. Because of its heterogeneity and recalcitrance, the development of efficient methods for its valorization still remains an open challenge. One approach to utilize lignin as a feedstock to produce valuable chemicals is its chemical deconstruction into mixtures of monomeric phenolic compounds followed by biological funneling into a single product. *Novosphingobium aromaticivorans* DSM12444 is a sphingomonad bacterium that has been proposed as a promising platform organism for biological funneling and it has been previously engineered to produce 2-pyrone-4,6-dicarboxylic acid (PDC). It can simultaneously metabolize multiple syringyl (S), guaiacyl (G), and *p*-hydroxyphenyl (H) lignin-derived aromatics through convergent routes involving the central aromatic intermediates 3-methoxygallic acid (3-MGA) and protocatechuic acid (PCA). Two critical reactions that are common in *N. aromaticivorans*

metabolism of S, G, and H type phenolics are aromatic *O*-demethylation and oxidative ring opening. I used genetic and enzymology tools to investigate enzymes predicted to be responsible for these reactions in *N. aromaticivorans*. The results obtained show the involvement of DesA in *O*-demethylation of syringic acid and vanillic acid; LigM, in *O*-demethylation of vanillic acid and 3-MGA; and DmtS, a potentially new *O*-demethylase involved in the conversion of 3-MGA into gallic acid (GA). In addition, I found LigAB to be the main aromatic ring opening dioxygenase involved in 3-MGA, PCA, and GA ring opening reactions, and LigAB2, a previously uncharacterized dioxygenase with high activity on GA and a minor role in 3-MGA and PCA ring openings. Our results also show evidence of a metabolic route for 3-MGA catabolism in *N. aromaticivorans* not characterized previously that involves the production of the aromatic intermediate GA and channels ~15% of the metabolic flow in addition to the predicted primary route via 3-MGA ring opening to produce CHMOD. The new knowledge obtained in this study allowed for the creation of an improved engineered strain for the funneling of S-type aromatic compounds into PDC that exhibits stoichiometric conversion yield of the model compound syringic acid.

4.2 Introduction

Lignocellulosic plant biomass, composed of the polysaccharides cellulose and hemicellulose and the phenolic-based polymer lignin, is the most abundant organic material on the planet with potential to support a sustainable economy based on renewable feedstocks (1). Numerous studies predict that the economic and environmental viability of lignocellulosic biomass utilization for fuel and chemical production will be increased by the utilization of as much of these polymers as possible, including the use of lignin for production of chemicals (2-4). I am interested in deciphering the bacterial metabolism of phenolic compounds to engineer bacterial hosts to convert biomass-derived lignin into chemicals.

Lignin is an amorphous heteropolymer containing mainly syringyl (S; two methoxy groups), guaiacyl (G; one methoxy group), and *p*-hydroxyphenyl (H; no methoxy groups) phenolic structures that differ in the

number of methoxy groups attached to the aromatic ring (5). One approach to valorizing lignin that has gained significant attention is to first use chemical techniques to deconstruct plant biomass and generate mixtures containing a large fraction of lower molecular weight phenolic compounds that could then be transformed by engineered microbes to a single valuable product (6). This funneling of phenolic mixtures to single compounds has been demonstrated with engineered strains of *Pseudomonas putida* (6, 7), *Rhodococcus jostii* (8), and *Novosphingobium aromaticivorans* (9). In addition, other microbes, such as the yeast *Rhodosporidium toruloides* (10) and the photoheterotrophic bacterium *Rhodopseudomonas palustris* (11-13) have been extensively studied for their ability to transform the plant-derived phenolic compounds often present in deconstructed plant biomass.

Among the desirable features for a microbial strain to be used as a chassis for the development of microbial lignin valorization strategies are an ability to metabolize the majority of the biomass-derived phenolic compounds and to funnel them into native convergent metabolic pathways (14). I am studying the sphingomonad bacterium *N. aromaticivorans* as a platform microbe for lignin valorization because it efficiently and simultaneously utilizes a large variety of S, G, and H phenolics (9, 15) and because it and other sphingomonads have enzymes to cleave different inter-subunit bonds in the lignin polymer (16-20). These features, when combined with the genetic and genomic information on sphingomonads like *N. aromaticivorans* could support strategies to maximize the number and type of lignin depolymerization products (e.g., phenolic monomers, oligomers) that can be microbially transformed into a range of valuable chemicals.

Metabolic pathways for the degradation of phenolics in sphingomonads have been proposed for *Sphingobium sp*. SYK-6, based on experiments with mutant strains and purified enzymes (21-23), and for *N. aromaticivorans*, based on analysis of a genome-scale transposon library and a set of targeted deletion mutants (9, 15). These studies have revealed several commonalities in the phenolic metabolism pathways of both organisms (Figure 4.1), but there remain significant knowledge gaps that limit the engineering of strains with increased transformation of phenolics to desired products.

To illustrate some of these knowledge gaps, syringic acid, which is an abundant biomass component that has been analyzed as a model S phenolic in both organisms, is demethylated to 3-methoxygallic acid (3-MGA), after which, multiple pathways have been proposed for 3-MGA transformation by *Sphingobium sp*. SYK-6 (21, 22). In one proposed pathway, the demethylation of 3-MGA produces gallic acid (GA), whose aromatic ring is cleaved by a dioxygenase to produce 4-oxalomesaconate (OMA) (24). Alternatively, it is proposed that the aromatic ring of 3-MGA is cleaved by a dioxygenase to produce 4-carboxy-2-hydroxy-6-methoxy-6-oxohexa-2,4-dienoate (CHMOD), which could be converted to OMA by an unknown enzyme (21, 22). Furthermore, *in vitro* experiments with a purified dioxygenase from *Sphingobium sp*. SYK-6 (LigAB) and 3-MGA as the substrate showed rapid production of 2-pyrone-4,6-dicarboxylic acid (PDC) in addition to the ring cleavage product CHMOD, a result that led to the hypothesis that LigAB catalyzes the transformation of 3-MGA to both CHMOD and PDC (22). In *N. aromaticivorans*, the only pathway thus far proposed for syringic acid metabolism (15) is via demethylation to 3-MGA, ring cleavage to CHMOD, and conversion to OMA. However, a *N. aromaticivorans* deletion mutant lacking the proposed enzymes for CHMOD conversion to OMA (DesC and DesD) resulted in accumulation of PDC, suggesting the possibility for cyclization of CHMOD to PDC that is independent of these enzymes (9).

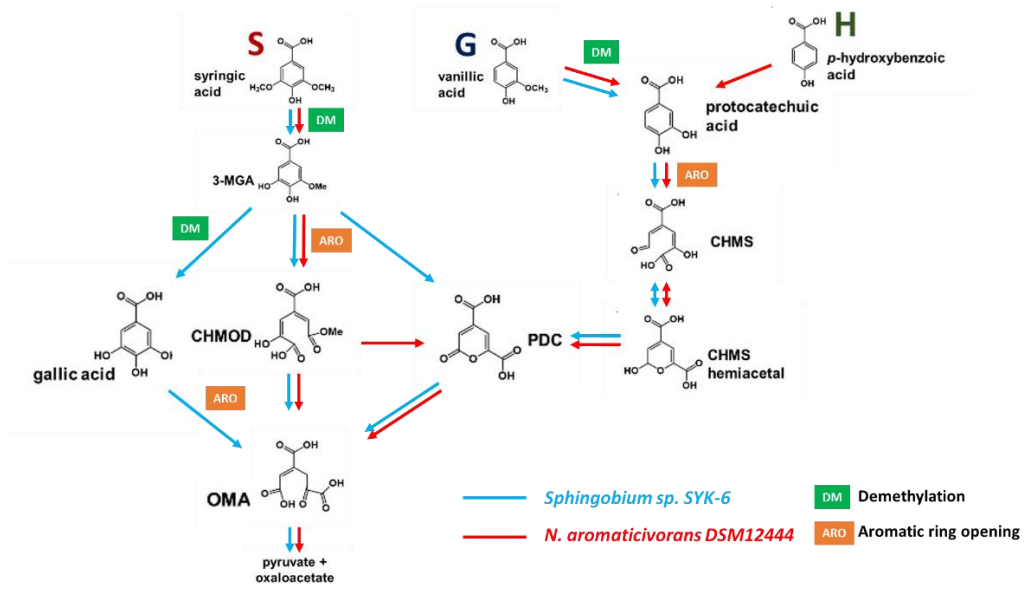


Figure 4.1. Pathways for the metabolism of S, G, and H phenolics that have been proposed for *Sphingobium sp*. SYK-6 (21-23) and *Novosphingobium aromaticivorans* (9, 15), showing the location of *O*-demethylation and aromatic ring opening steps.

There are also knowledge gaps in the bacterial metabolism of the other major G and H phenolic substituents of plant cell walls. In the case of the G phenolic vanillic acid, its metabolism by both *Sphingobium sp*. SYK-6 and *N. aromaticivorans* is proposed (15, 16) to entail demethylation to protocatechuic acid (PCA), ring cleavage to 4-carboxy-2-hydroxy-cis,cis-muconate-6-semialdehyde (CHMS), oxidation to PDC, and hydrolysis to OMA (Figure 1). Degradation of the H phenolic *p*-hydroxybenzoic acid has only been studied in *N. aromaticivorans* (9), with experimental evidence suggesting transformation to PCA as the initial step to enter the described pathway for G phenolics (Figure 4.1). However, it is unclear if the S and G/H branches of the phenolic degradation pathways in sphingomonads have common or pathway-specific *O*-demethylation and aromatic ring cleavage enzymes (Figure 4.1). Indeed, it has been proposed that some of these demethylase enzymes have broad substrate specificity and are active in multiple branches. For example, one dioxygenase (LigAB) in *N. aromaticivorans* has been proposed to be active in the ring cleavage of 3-MGA and PCA (15), and in

Sphingobium sp. SYK-6, an *O*-demethylase (LigM) has been proposed to be active on both 3-MGA and vanillic acid (25).

To address knowledge gaps on the reactions and enzymes involved in S, G, and H phenolic metabolism by *N. aromaticivorans*, I analyzed putative *O*-demethylases and aromatic ring opening dioxygenases that are predicted to be involved in metabolism of these plant-derived aromatics. I present results of experiments with purified enzymes and with deletion mutants that provide evidence for functionally redundant pathways and for enzymes with broad substrate specificity in this organism. Our studies led us to identify an aromatic ring-opening dioxygenase (LigAB2) and an *O*-demethylase (DmtS), which have not been previously shown to have these activities in sphingomonads. In addition, the newly acquired knowledge on enzyme redundancy and substrate specificity allowed us to engineer a second-generation *N. aromaticivorans* DSM 12444 strain with improved yields of PDC from plant-derived phenolics.

4.3 Results

Lignocellulosic plant biomass, composed of the polysaccharides cellulose and hemicellulose and the phenolic-based polymer lignin, is the most abundant organic material on the planet with potential to support a sustainable economy based on renewable feedstocks (1). Numerous studies predict that the economic and environmental viability of lignocellulosic biomass utilization for fuel and chemical production will be increased by the utilization of as much of these polymers as possible, including the use of lignin for production of chemicals (2-4). I am interested in deciphering the bacterial metabolism of phenolic compounds to engineer bacterial hosts to convert biomass-derived lignin into chemicals.

4.3.1 Identification of putative aromatic O-demethylases in N. aromaticivorans

Based on growth properties of *N. aromaticivorans* mutants, gene products with significant amino acid sequence identity to *O*-demethylases, encoded by Saro_2861 (*ligM*) and Saro_2404 (*desA*), have been proposed to be involved in metabolism of vanillic acid and syringic acid, respectively (15). These two

proteins share ~78% and ~71% amino acid sequence identity with known or proposed *Sphingobium sp*. SYK-6 *O*-demethylases SLG_12740 (*ligM*) (25) and SLG_25000 (*desA*) (26). In addition, since *O*-demethylation of vanillic acid in *Pseudomonas* (27, 28) is catalyzed by VanAB, I analyzed the *N. aromaticivorans* gene Saro_1872 (hereafter called *dmtS*), which encodes a protein with the closest amino acid sequence identity to the VanA subunit of this enzyme (e.g., ~27% amino acid identity with VanA of *Pseudomonas sp*. strain HR199 (29)).

Effect of deleting putative *O*-demethylase genes on growth of *N. aromaticivorans*. To begin evaluating the involvement of *ligM*, *desA*, and *dmtS* in the degradation of S and G phenolics in *N. aromaticivorans* we generated mutants (Table 4.1) containing combinations of deletions of these 3 genes in a parent strain (12444*Δ1879*) and in a strain (12444PDC) in which deletions in PDC- and CHMOD-degradation genes led to accumulation of PDC from S, G, and H aromatics (9).

Figure 4.2 shows growth curves for the parent strain and for its corresponding set of mutant strains. When cultured in the presence of only syringic acid, the parent strain and the mutant strains 12444*ΔligM*, 12444*ΔdmtS*, and 12444*ΔligMΔdmtS* had similar growth, whereas all mutant strains lacking *desA* were unable to grow (Figure 4.2A). This suggests that *desA* is essential for *N. aromaticivorans* growth on syringic acid, whereas *ligM* and *dmtS* are not.

When cultured in the presence of vanillic acid as the sole carbon source, three distinct growth patterns were observed for this set of mutants (Figure 4.2B). First, all mutant strains with an intact *ligM* (12444*ΔdesA*, 12444*ΔdmtS*, and 12444*ΔdesAΔdmtS*) exhibited a similar growth pattern as the parent strain, suggesting that the DesA and DmtS enzymes are not essential for *N. aromaticivorans* growth on vanillic acid. Second, strains 12444*ΔligM* and 12444*ΔligMΔdmtS* showed lower growth rates and a lower final cell density compared to the parent strain, and third, strains simultaneously lacking *desA* and *ligM*, 12444*ΔdesAΔligM* and 12444*ΔdesAΔligMΔdmtS* were unable to grow in the presence of vanillic acid. Overall, these results suggest that the LigM enzyme is sufficient (strains with an intact *ligM* always grow similarly to wild type cells on vanillic acid) but not necessary for *N. aromaticivorans* growth on vanillic

acid, and that DesA may replace the function of LigM in vanillic acid metabolism, but cannot fully compensate for the loss of LigM. On the other hand, it follows from this analysis that *dmtS* does not appear to make a significant contribution to *N. aromaticivorans* growth on vanillic acid.

Table 4.1. List of *N. aromaticivorans* parent and mutant strains with deletions of putative *O*-demethylases used in this study.

Strain	Background used for strain construction	Relevant Characteristics
Parent strain (12444*Δ1879*)*	DSM12444	12444Δ1879
PDC-producing strain (12444PDC)*	12444*Δ1879*	12444Δ1879 ΔSaro_2819 ΔSaro_2864 ΔSaro_2865
12444*ΔligM*	12444*Δ1879*	12444Δ1879 ΔSaro_2861
12444*ΔdesA*	12444*Δ1879*	12444Δ1879 ΔSaro_2404
12444*ΔdmtS*	12444*Δ1879*	12444Δ1879 ΔSaro_1872
12444*ΔligMΔdesA*	12444*ΔligM*	12444Δ1879 ΔSaro_2861 ΔSaro_2404
12444*ΔligMΔdmtS*	12444*ΔligM*	12444Δ1879 ΔSaro_2861 ΔSaro_1872
12444*ΔdesAΔdmtS*	12444*ΔdmtS*	12444Δ1879 ΔSaro_2404 ΔSaro_1872
12444*ΔligMΔdesAΔdmtS*	12444*ΔligMΔdmtS*	12444Δ1879 ΔSaro_2861 ΔSaro_2404 ΔSaro_1872
12444PDC*ΔligM*	12444PDC	12444PDC ΔSaro_2861
12444PDC*ΔdesA*	12444PDC	12444PDC ΔSaro_2404
12444PDC*ΔdmtS*	12444PDC	12444PDC ΔSaro_1872
12444PDC*ΔligMΔdesA*	12444PDC*ΔligM*	12444PDC ΔSaro_2861 ΔSaro_2404
12444PDC*ΔligMΔdmtS*	12444PDC*ΔligM*	12444PDC ΔSaro_2861 ΔSaro_1872
12444PDC*ΔdesAΔdmtS*	12444PDC*ΔdmtS*	12444PDC ΔSaro_2404 ΔSaro_1872
12444PDC*ΔligMΔdesAΔdmtS*	12444PDC*ΔligMΔdmtS*	12444PDC ΔSaro_2861 ΔSaro_2404 ΔSaro_1872

*Construction of the parent and PDC-producing strains is described in Perez et al. (9). The parent strain has a deletion of Saro_1879 (*sacB*) to render it susceptible to genetic manipulation using *sacB* as a counterselective marker (9).

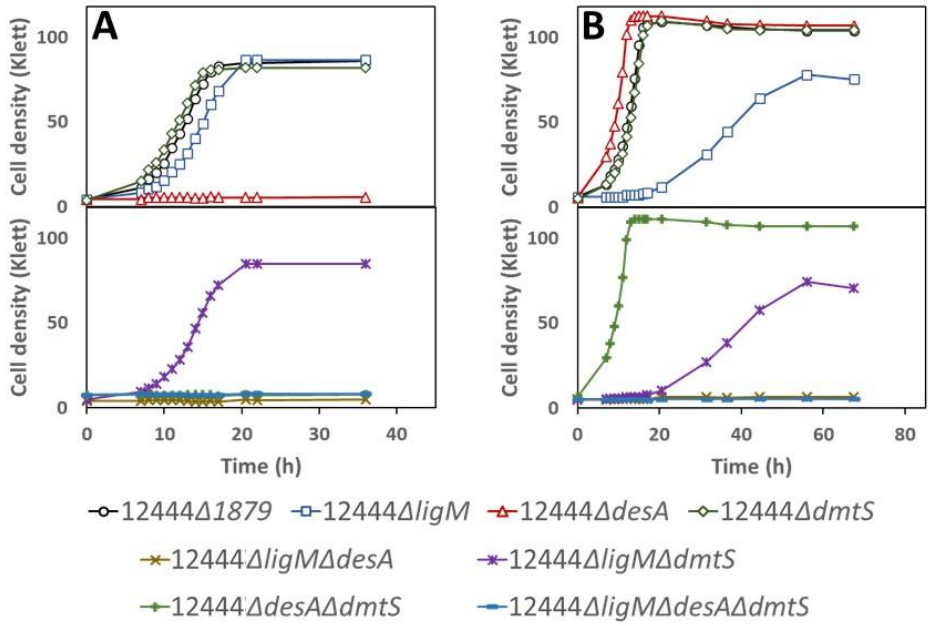


Figure 4.2. Growth of indicated strains of *N. aromaticivorans* in minimal media supplemented with syringic acid (panel A) or vanillic acid (panel B). Top panels show data for parent strain and single-deletion mutants, whereas bottom panels show growth curves for strains with multiple gene deletions.

Because the growth defects caused by gene deletions in the parent strain do not necessarily reveal which specific intracellular reactions are being affected, and can obscure the roles of genes with redundant functions, we also used the PDC-producing strain as a background to construct mutants lacking the same combinations of putative *O*-demethylase genes (Table 4.1). With this second set of mutants, which require a non-aromatic carbon source for growth, we used PDC production as a proxy to elucidate the metabolic pathways affected by the gene deletions.

In the presence of glucose and syringic acid as an aromatic carbon source, the original PDC-producing strain (12444PDC) completely removed the syringic acid from the medium and produced PDC with a yield of ~85% (Figure 4.3A; Table 4.2). Strains with deletions of *desA* plus at least one more putative *O*-demethylase (12444PDC*ΔdesAΔligM*, 12444PDC*ΔdesAΔdmtS*, and 12444PDC*ΔdesAΔligMΔdmtS*) were not able to degrade syringic acid (Figure 4.3E, 4.3F, 4.3H), whereas the strain with the single *desA*

deletion consumed syringic acid and produced PDC at a ~88% yield (Figure 4.3B). As three out of the four mutants containing a deletion of *desA* were unable to consume syringic acid, these results support a role for DesA in syringic acid metabolism by *N. aromaticivorans*, which is proposed to be its demethylation to 3-MGA (Figure 4.1). This predicted role for DesA in syringic acid metabolism is in agreement with the impact of deleting *desA* in the parent strain (Figure 4.2) and with the prior fitness defects caused by transposon insertions in this gene when a mutant library is grown in the presence of this aromatic compound (15). However, PDC production by single mutants lacking only the *desA* gene suggests that, when *N. aromaticivorans* is not depending on aromatics for growth, LigM and DmtS may substitute for DesA activity in the metabolism of syringic acid.

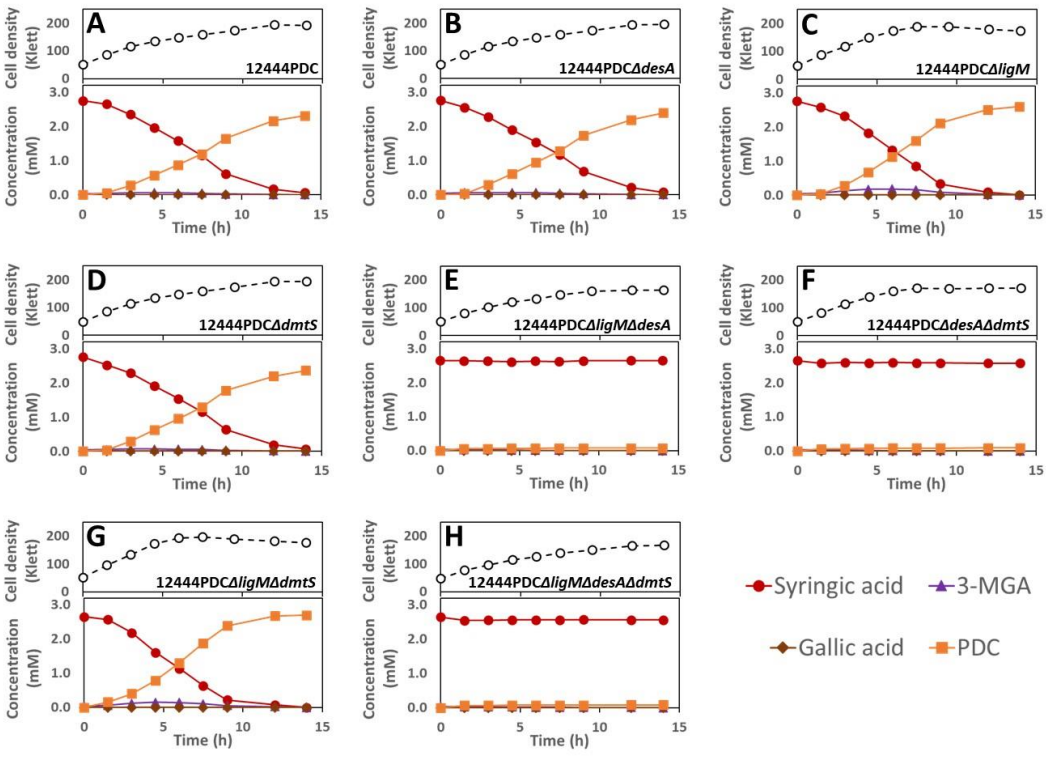


Figure 4.3. Cell densities and extracellular compound concentrations for the indicated *N. aromaticivorans* strains grown in minimal media containing glucose and syringic acid. Values represent the average of three biological replicates.

Furthermore, all the mutant strains with intact *desA* that were constructed in the PDC-producing, 12444PDC, background (12444PDC*ΔligM*, 12444PDC*ΔdmtS*, and 12444PDC*ΔligMΔdmtS*) consumed syringic acid and accumulated PDC. Notably, the PDC yield from syringic acid by strain 12444PDC*ΔligMΔdmtS* was stoichiometric and reproducibly higher than the yield of this product obtained in the other strains (Table 4.2). This suggests that the simultaneous deletion of both *ligM* and *dmtS* blocks another previously unknown pathway for syringic acid metabolism that normally detracts from PDC production in the 12444PDC strain. Since the PDC yield by the 12444PDC strain was ~85%, I interpret these results as indicating that only a small fraction of the syringic acid is metabolized via this previously unknown pathway.

I predict that this alternative pathway for syringic acid metabolism involves *O*-demethylation (by LigM and/or DmtS) of 3-MGA to GA, since that pathway has been proposed to occur in *Sphingobium sp*. SYK-6 (Figure 4.1). An observed small and transient accumulation of 3-MGA in the culture media when 12444PDC*ΔligM* and 12444PDC*ΔligMΔdmtS* (Figure 4.3C, 4.3G) are grown in the presence of syringic acid supports this hypothesis. To further test this hypothesis, We grew the 12444PDC strain and its corresponding set of putative *O*-demethylase deletion mutants in the presence of glucose and 3-MGA (Figure 4.4). I posited that using 3-MGA as the aromatic substrate would provide a growth substrate that only requires one demethylation reaction compared to growth with syringic acid that is predicted to require the removal of two methyl groups. In this set of experiments, strain 12444PDC was able to completely remove 3-MGA from the medium, converting ~94% of it into PDC. Moreover, in contrast with the experiments using syringic acid (Figure 4.3), each of the putative *O*-demethylase mutant derivatives of 12444PDC were able to completely remove 3-MGA from the media (Figure 4.4). This result indicates that none of these putative *O*-demethylases is essential for conversion of 3-MGA to PDC. It also suggests that 3-MGA is primarily metabolized via ring cleavage in *N. aromaticivorans*, a process that does not require *O*-demethylation (Figure 4.1). In addition, all strains with deletions of *ligM* (12444PDC*ΔligM*, 12444PDC*ΔdesAΔligM*, 12444PDC*ΔligMΔdmtS*, and 12444PDC*ΔdesAΔligMΔdmtS*)

reproducibly produced more PDC than strain 12444PDC (Table 4.2), a result that suggests the involvement of LigM in the proposed alternative pathway of syringic acid metabolism. Notably, stoichiometric conversion of 3-MGA to PDC was achieved in the mutants with simultaneous deletion of *ligM* and *dmtS* (12444PDC*ΔligMΔdmtS,* and 12444PDC*ΔdesAΔligMΔdmtS*), which, in agreement with the experiments using syringic acid, suggests that the loss of LigM and DmtS blocks the alternative pathway for syringic acid metabolism in this bacterium.

Table 4.2. PDC yields from *N. aromaticivorans* strains in the presence of glucose plus the indicated aromatic substrate.

Aromatic Substrate	Strain*	Deleted genes			PDC Yield (%)**
		Saro_2404 (*desA*)	Saro_2861 (*ligM*)	Saro_1872 (*dmtS*)	
Syringic acid	12444PDC				84.8 ± 0.6
	12444PDC*ΔligM*		X		93.0 ± 0.8
	12444PDC*ΔdesA*	X			88.1 ± 1.4
	12444PDC*ΔdmtS*			X	86.5 ± 1.5
	12444PDC*ΔligMΔdesA*	X	X		N/C***
	12444PDC*ΔligMΔdmtS*		X	X	100.4 ± 2.5
	12444PDC*ΔdesAΔdmtS*	X		X	N/C***
	12444PDC*ΔligMΔdesAΔdmtS*	X	X	X	N/C***
3-MGA	12444PDC				93.7 ± 0.3
	12444PDC*ΔligM*		X		97.2 ± 0.3
	12444PDC*ΔdesA*	X			93.8 ± 0.4
	12444PDC*ΔdmtS*			X	94.0 ± 0.1
	12444PDC*ΔligMΔdesA*	X	X		98.1 ± 1.0
	12444PDC*ΔligMΔdmtS*		X	X	100.1 ± 0.8
	12444PDC*ΔdesAΔdmtS*	X		X	90.9 ± 0.7
	12444PDC*ΔligMΔdesAΔdmtS*	X	X	X	100.7 ± 0.7
Vanillic acid	12444PDC				95.9 ± 1.5
	12444PDC*ΔligM*		X		92.8 ± 2.6
	12444PDC*ΔdesA*	X			98.0 ± 3.0
	12444PDC*ΔdmtS*			X	92.8 ± 1.8
	12444PDC*ΔligMΔdesA*	X	X		47.1 ± 6.7
	12444PDC*ΔligMΔdmtS*		X	X	97.8 ± 3.0
	12444PDC*ΔdesAΔdmtS*	X		X	99.4 ± 1.3
	12444PDC*ΔligMΔdesAΔdmtS*	X	X	X	34.5 ± 4.3

* Strains defined in Table 1.

** Yield reported as average and standard deviation of three biological replicates.

*** PDC yield was not calculated (N/C) for conditions in which the aromatic substrate was not consumed.

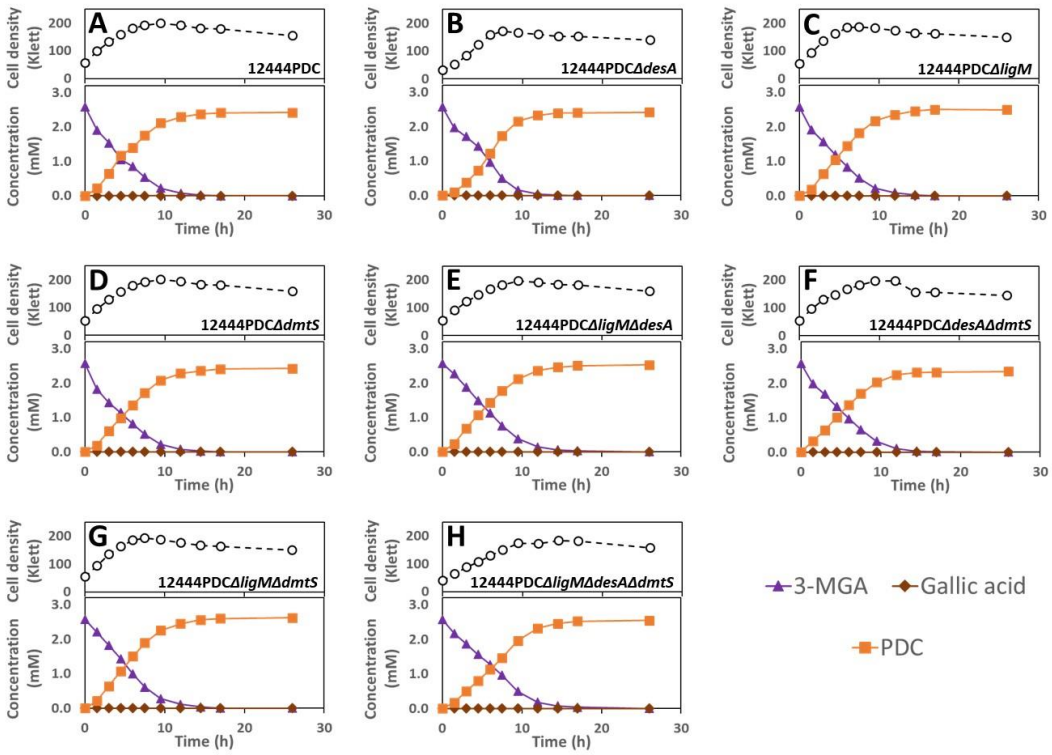


Figure 4.4. Cell densities and extracellular compounds concentrations for the indicated *N. aromaticivorans* strains grown in minimal media containing glucose and 3-MGA. Values represent the average of three biological replicates.

In an additional experiment with the PDC-producing strain and its derivatives, I analyzed the role of the putative *O*-demethylases in metabolism of vanillic acid. In the presence of glucose and vanillic acid, strain 12444PDC completely removed the vanillic acid from the medium and converted it into PDC with ~96% yield (Table 4.2). Transient extracellular accumulation of a small amount of PCA during the course of the experiment (Figure 4.5A) supports the predicted role of demethylation in the degradation of vanillic acid (Figure 4.1). Analysis of the mutants lacking single putative *O*-demethylases (Figure 4.5B, 4.5C, 4.5D) revealed a decrease in the vanillic acid consumption rate when *ligM* was deleted (Figure 4.5C), compared to the other two strains, suggesting a role for LigM in vanillic acid metabolism. In the set of double *O*-demethylase mutants (Figure 4.5E, 4.5F, 4.5G), no effect was seen when *desA* and *dmtS* were deleted (Figure 4.5F) but a decrease in vanillic acid consumption rates was evident with all double

mutants which lacked the *ligM* gene. This observation is also consistent with the reduced rate of vanillic acid degradation in the single *ligM* deletion mutant (Figure 4.5C). Notably, the rate of vanillic acid consumption decreased the most when *ligM* and *desA* were both deleted (Figure 4.5E), suggesting that DesA can partially substitute for LigM in the degradation of vanillic acid. Finally, minimal vanillic acid degradation was observed in a mutant that lacks all 3 of the putative *O*-demethylase genes (Figure 4.5H), consistent with the lack of LigM and DesA activity in this strain. Taken together, the data suggest that under the conditions tested, vanillic acid demethylation in *N. aromaticivorans* is primarily dependent on LigM, but DesA can also demethylate vanillic acid, and that DmtS does not have an important role in vanillic acid degradation, consistent with our findings with the *O*-demethylase mutants made from the 12444*Δ1879* parent strain (Figure 4.2).

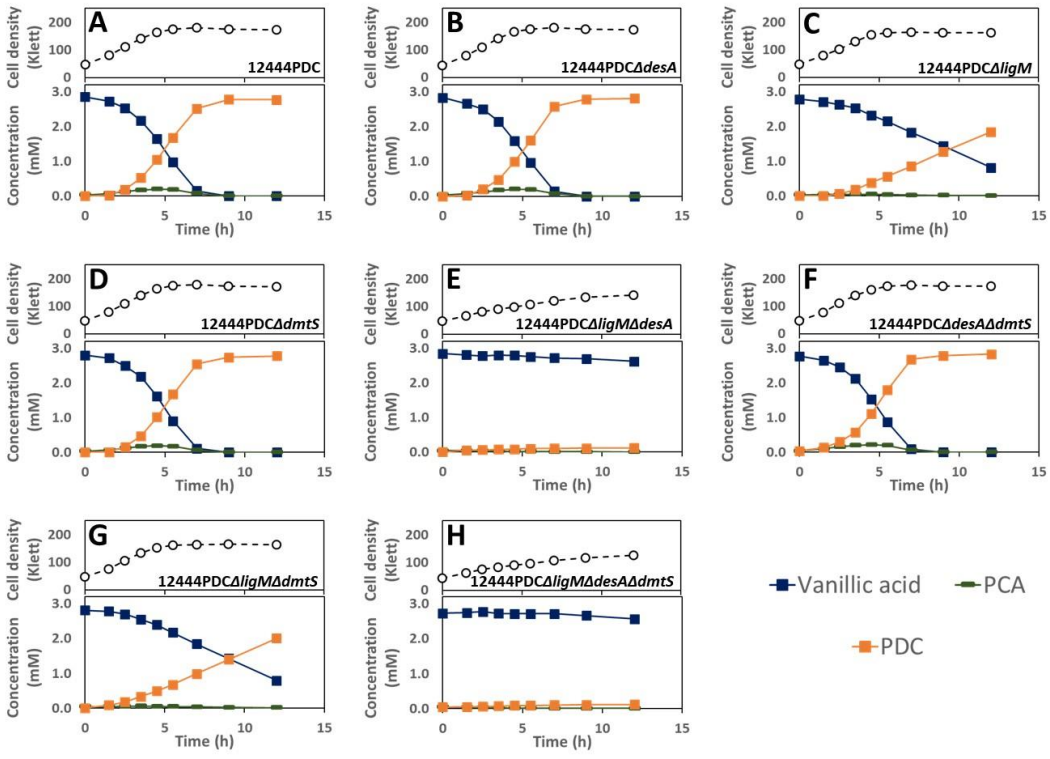


Figure 4.5. Growth and extracellular aromatic compound concentrations of different strains of *N. aromaticivorans* cultures in minimal media supplemented with glucose and vanillic acid. Values represent the average of three biological replicates.

Activity of LigM and DesA with aromatic substrates. *In vitro* experiments were performed with purified recombinant LigM and DesA proteins to test their activity with the methylated aromatic substrates. These assays were only performed with LigM and DesA since we have thus far been unsuccessful in purifying an active recombinant DmtS protein. The LigM and DesA homologues from *Sphingobium sp*. SYK-6 have been shown to be tetrahydrofolate (H_4folate) dependent *O*-demethylases (25, 26), and *in vitro* assays with recombinant LigM and DesA of *N. aromaticivorans* confirmed their dependency on H_4folate for demethylation of the substrates tested (Figure 4.S1).

I found that the recombinant LigM protein of *N. aromaticivorans* was able to convert 3-MGA into GA and vanillic acid into PCA at comparable rates under our assay conditions (Figure 4.6C, 4.6E). However, under identical conditions, the recombinant LigM protein was unable to convert a detectable amount of syringic acid into 3-MGA (Figure 4.6A). These results are consistent with the observations in growth experiments with mutant strains, which predicted LigM's involvement in vanillic acid and syringic acid metabolism; in the case of syringic acid, these *in vitro* data suggest that during syringic acid metabolism, LigM does not catalyze the demethylation of syringic acid to 3-MGA, but instead demethylates 3-MGA, resulting in the small amount of syringic acid that is metabolized via GA.

I also found that the recombinant DesA of *N. aromaticivorans* demethylated both syringic acid and, somewhat more slowly, vanillic acid (Figure 4.6B, 4.6F), but was not active in demethylating 3-MGA (Figure 4.6D). These results support the critical role of DesA in the demethylation of syringic acid that was predicted by analyzing growth of mutant strains (Figure 4.2) and its potential role in vanillic acid transformation when *ligM* is deleted (Figure 4.5), and offer new evidence supporting the hypothesis that this enzyme does not participate in 3-MGA demethylation.

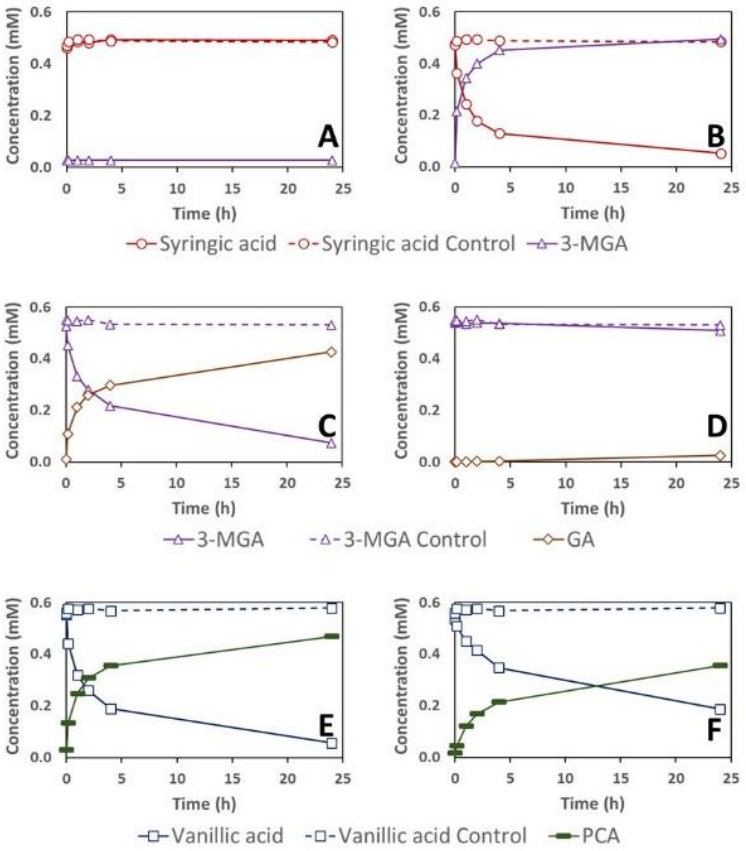


Figure 4.6. Substrate and product concentration during *in vitro* enzyme assays of LigM (Saro_2861) (panels A, C, and E) and DesA (Saro_2404) (panels B, D, and F) with syringic acid (panels A and B), 3-MGA (panels C and D), and vanillic acid (panels E and F). Concentrations in reactions lacking recombinant proteins are shown in dashed line. Reported concentrations are the average of three replicate assays. The low amounts of 3-MGA measured in the assays with LigM and syringic acid (panel A) was a trace contaminant in the syringic acid used in these assays.

4.3.2 Identification of putative aromatic ring opening dioxygenases in N. aromaticivorans

The opening of the aromatic ring is an essential step in the assimilation of plant-derived phenolic compounds into intermediary metabolism (Figure 4.1). Three ring opening dioxygenases with activity in the metabolism of S and G phenolics have been described in *Sphingobium sp*. SYK-6: SLG_12510 and

SLG_12500 (LigAB), SLG_19030 (DesZ), and SLG_03330 (DesB) (16). However, only one aromatic ring opening dioxygenase has been implicated in the metabolism of these compounds in *N. aromaticivorans*. This is a LigAB homologue, encoded by Saro_2813 (*ligA*) and Saro_2812 (*ligB*) (15), whose subunits have ~67% and ~70% amino acid sequence identity with LigA and LigB of *Sphingobium sp*. SYK-6, respectively. A search of the *N. aromaticivorans* genome reveals genes that could encode another hetero-oligomeric aromatic ring opening dioxygenase encoded by Saro_1233 and Saro_1234 (hereafter referred to as ligA2 and ligB2, respectively), whose subunits have amino acid sequences that are ~33% and ~42% identical to the LigA and LigB of *N. aromaticivorans*. I sought to test the roles of *N. aromaticivorans* LigAB and LigAB2 in the metabolism of aromatic compounds derived from plant biomass.

Table 4.3. List of *N. aromaticivorans* mutant strains with deletions of putative aromatic ring cleavage dioxygenases used in this study.

Strain	Background used for strain construction	Relevant Characteristics
12444*ΔligAB*	12444*Δ1879**	12444Δ1879 ΔSaro_2812 ΔSaro_2813
12444*ΔligAB2*	12444*Δ1879*	12444Δ1879 ΔSaro_1233 ΔSaro_1234
12444*ΔligABΔligAB2*	12444*ΔligAB*	12444Δ1879 ΔSaro_2812 ΔSaro_2813 ΔSaro_1233 ΔSaro_1234
12444PDC*ΔligAB*	12444PDC*	12444PDC ΔSaro_2812 ΔSaro_2813
12444PDC*ΔligAB2*	12444PDC	12444PDC ΔSaro_1233 ΔSaro_1234
12444PDC*ΔligABΔligAB2*	12444PDC*ΔligAB*	12444PDC ΔSaro_2812 ΔSaro_2813 ΔSaro_1233 ΔSaro_1234

*Construction of the parent strain (12444*Δ1879*) is described in Kontur et al. (2018) (17) and the PDC-producing strain (12444PDC) is described in Perez et al. (2019) (9). See also footnote in Table 1.

Effect of deleting genes encoding putative aromatic ring opening dioxygenases on growth of *N. aromaticivorans*. To evaluate the roles of LigAB and LigAB2 in the degradation of S and G phenolics by *N. aromaticivorans*, we tested growth and aromatic metabolism by mutants containing combinations of

deletions in *ligAB* and *ligAB2* in the parent strain (12444*Δ1879*) and the PDC-producing strain (12444PDC) (Table 4.3). When cultured in the presence of either syringic acid or vanillic acid, the parent strain and strain 12444*ΔligAB2* both grew well, whereas strains with deletion of ligAB (12444*ΔligAB* and 12444*ΔligABΔligAB2*) were unable to grow (Figure 4.7). This indicates that LigAB is necessary for *N. aromaticivorans* growth on both syringic and vanillic acids, whereas LigAB2 is not.

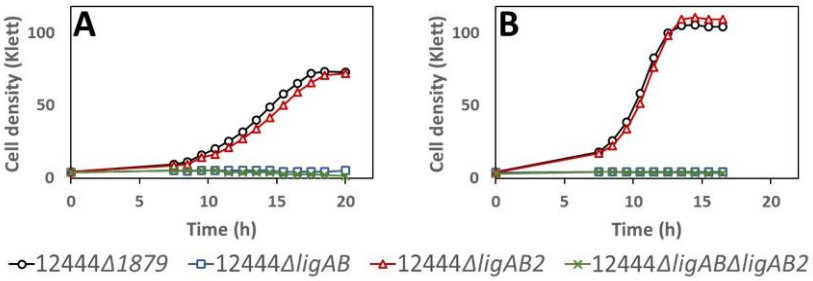


Figure 4.7. Growth of different strains of *N. aromaticivorans* in minimal media supplemented with syringic acid (panel A) or vanillic acid (panel B).

Investigating the metabolism of syringic acid by mutants containing deletions in *ligAB* and *ligAB2* in the PDC-producing strain (12444PDC) (Figure 4.8), I observed that all strains consumed syringic acid, but only the mutant with intact LigAB (Figure 4.8B) produced PDC with a high yield (~84%; Table 4.4), supporting the hypothesis that LigAB plays an important role in PDC production. The proposed pathway for PDC production from syringic acid includes ring cleavage of 3-MGA to CHMOD, which is then converted to PDC (Figure 4.1). The high PDC yield observed in cells in which LigAB is present indicates that this enzyme plays a major role in the transformation of 3-MGA. It also suggests that in *N. aromaticivorans*, 3-MGA ring opening is the primary route for 3-MGA metabolism, while demethylation of 3-MGA to GA is a secondary route responsible for only a small fraction of the metabolized 3-MGA. In support of this, the tests with the two mutants that were missing LigAB (Figure 4.8A, 4.8C) showed

slower rates of syringic acid consumption and accumulation of 3-MGA and GA, the predicted metabolic intermediates in this secondary metabolic route. Strain 12444PDC*ΔligAB* also converted only ~2% of the syringic acid to PDC (Table 4.4), suggesting that LigAB2 may have some role in 3-MGA metabolism, whereas strain 12444PDC*ΔligABΔligAB2* did not produce any detectable PDC (Table 4.4), suggesting a complete interruption of 3-MGA aromatic ring opening in the mutant that lacked both of the putative dioxygenases, LigAB and LigAB2.

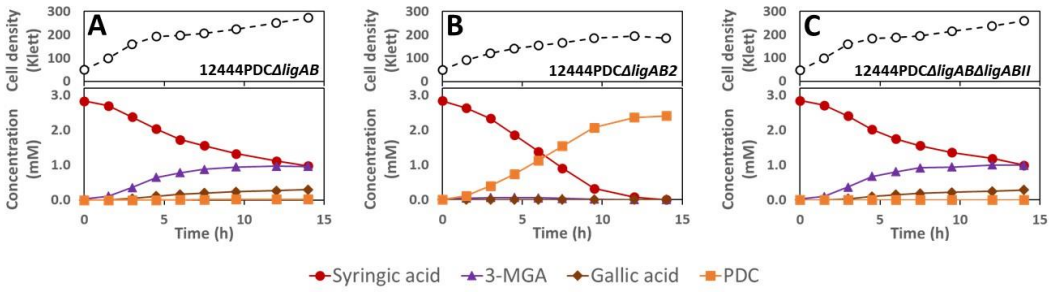


Figure 4.8. Growth and extracellular compound concentration of indicated *N. aromaticivorans* cultures in minimal media supplemented with glucose and syringic acid.

Notably, in the two experiments where I observed GA accumulation (Figure 4.8A, 4.8C), the media darkened in color, which could be due to a rapid non-enzymatic transformation of GA under the conditions used for these experiments. Since the sum of the molar concentration of unreacted syringic acid plus the accumulated metabolites does not add up to 100% in either experiment, these results also suggests that the GA produced was partially degraded either abiotically or biologically.

Table 4.4. PDC yields from *N. aromaticivorans* strains in the presence of glucose plus the indicated aromatic substrate.

Aromatic Substrate	Strain*	Deleted genes		PDC Yield (%)**
		Saro_2812/3 (*ligAB*)	Saro_1233/4 (*ligAB2*)	
Syringic acid	12444PDC***			84.8 ± 0.6
	12444PDC*ΔligAB*	X		2.4 ± 0.6
	12444PDC*ΔligAB2*		X	84.0 ± 2.1
	12444PDC*ΔligABΔligAB2*	X	X	0.0 ± 0.0
Vanillic acid	12444PDC***			95.9 ± 1.5
	12444PDC*ΔligAB*	X		11.6 ± 1.4
	12444PDC*ΔligAB2*		X	93.3 ± 2.8
	12444PDC*ΔligABΔligAB2*	X	X	0.0 ± 0.0

* Strains defined in Table 1.
** Yield reported as average and standard deviation of three biological replicates.
*** Results for strain 12444PDC are the same are reported in Table 2 and copied here to facilitate comparison.

The role of the putative ring-opening dioxygenases in vanillic acid degradation was also tested using the PDC-producing strain deletion mutants (Figure 4.9). Both the LigAB and LigAB2 mutant derivatives of the PDC-producing strain showed degradation of vanillic acid, but only the mutant with intact *ligAB* genes produced high yields of extracellular PDC (Figure 4.9B, Table 4.4). The strains with deleted *ligAB* genes, 12444PDC*ΔligAB* (Figure 4.9A) and 12444PDC*ΔligABΔligAB2* (Figure 4.9C), had a slower consumption of vanillic acid, only removing ~80% of it from the media and converting ~48% of it into extracellular PCA over the course of the experiments. These observations suggest that LigAB has a significant role in PCA ring opening in *N. aromaticivorans*. Notably, strain 12444PDC*ΔligAB* also converted ~12% of the vanillic acid into PDC, whereas 12444PDC*ΔligABΔligAB2* did not accumulate any detectable PDC in the media, suggesting that in the absence of LigAB, LigAB2 may also function in PCA ring opening, while the absence of both LigAB and LigAB2 completely eliminates the PCA ring opening activity in *N. aromaticivorans*.

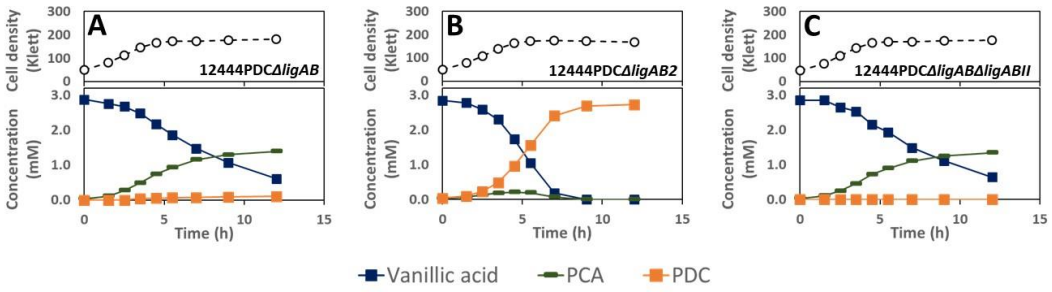


Figure 4.9. Growth and extracellular compound concentration of the indicated strains of *N. aromaticivorans* in minimal media supplemented with glucose and vanillic acid. Values represent the average of three biological replicates.

Activity of LigAB and LigAB2 with aromatic substrates. To investigate the predicted activities of LigAB and LigAB2 in aromatic ring opening, we purified recombinant forms of the proteins and tested them for activity *in vitro* (Figure 4.10). With 3-MGA as the substrate, LigAB completely removed the 3-MGA from the reaction mixture within two hours (Figure 4.10A), whereas LigAB2 only removed ~20% of the 3-MGA over the course of 24 hours (Figure 4.10B). With each purified enzyme, two products that transiently accumulated in the assay were identified as isomers of CHMOD (See Supplementary Information), and a third product was identified as PDC. These results show that LigAB catalyzes the transformation of 3-MGA to CHMOD, and that CHMOD is subsequently transformed to PDC, as it has been proposed previously (9). They also show for the first time that LigAB2 can also convert 3-MGA to CHMOD, albeit at a slower rate than with purified LigAB under the assay conditions used, which is consistent with results from the *in vivo* analyses of defined mutants in each of the corresponding genes (Figures 4.7 and 4.8).

Both LigAB and LigAB2 also showed activity when tested with GA as the substrate (Figure 4.10C, 4.10D). One compound that accumulated in these assays has been identified as OMA, the expected product of this ring opening reaction (See Supplementary Information). Under identical assay conditions,

LigAB2 degraded ~31% of the GA after 2 h of reaction (Figure 4.10D), whereas LigAB degraded only ~11% (Figure 4.10C). These findings suggest that of these two ring-opening dioxygenases, *N. aromaticivorans* LigAB2 has higher catalytic activity with GA than purified LigAB.

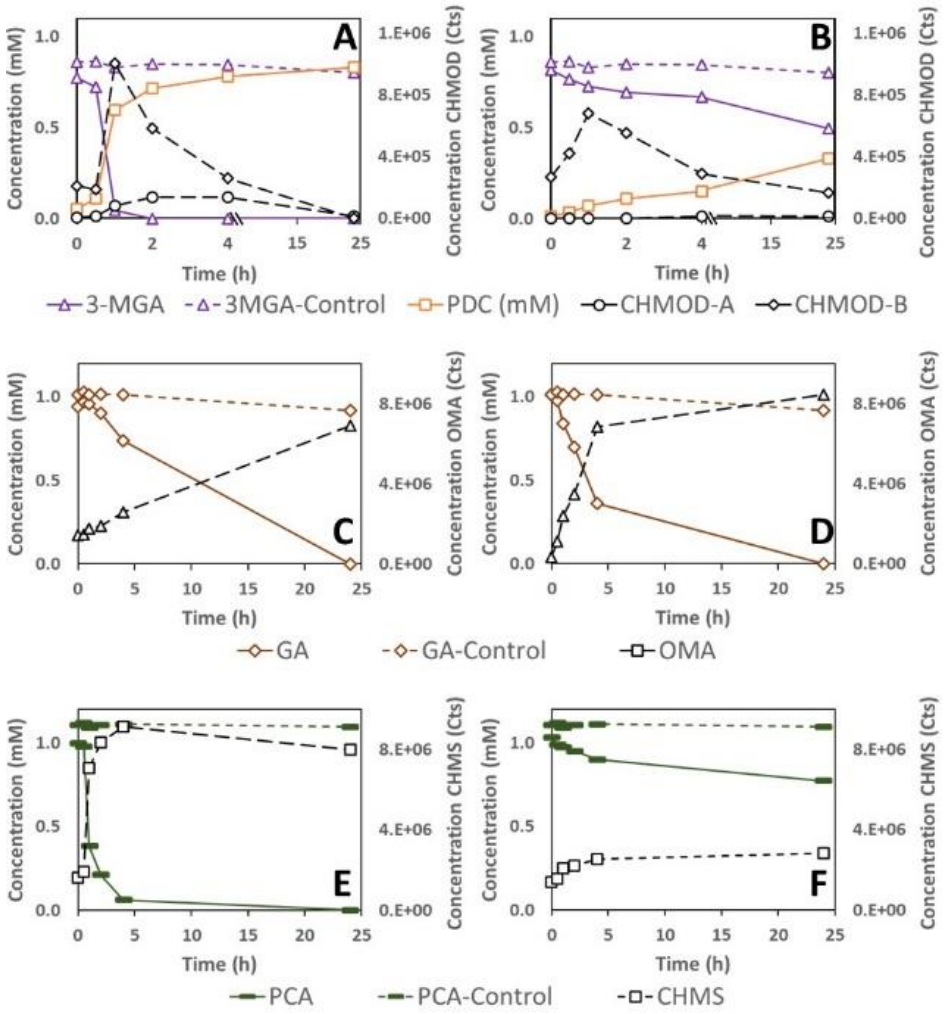


Figure 4.10. Substrate and product concentration of *in vitro* assays of LigAB (Saro_2812/3) (panels A, C, and E) and LigAB2 (Saro_1233/4) (panels B, D, and F) on 3-MGA (panels A and B), GA (panels C and D), and PCA (panels E and F). Concentration of substrates in assays performed in the absence of added recombinant protein are shown in dashed line and the same symbol and color as condition with enzyme. Samples were analyzed with HPLC-MS/MS and concentrations correspond to the average of three replicates.

When tested with PCA as the substrate under identical assay conditions, LigAB removed ~81% of the PCA from the reaction mixture after two hours (Figure 4.10E), while LigAB2 only removed ~14% of the PCA over the same time span (Figure 4.10F). The reaction product from both assays was identified as CHMS (See Supplementary Information), consistent with the expected ring opening product of these enzymes with PCA as a substrate. The relative rates of PCA disappearance in assays using LigAB and LigAB2 suggests that, under the assay conditions, LigAB is more catalytically active with PCA, which is consistent with the *in vivo* experiments with the individual mutant strains (Figure 4.9).

4.4 Discussion

Strategies to successfully engineer efficient microbial catalysts that produce valuable compounds from chemically depolymerized lignocellulosic biomass have several requirements, including the need for the microorganisms to funnel a heterogeneous mixture of plant-derived phenolic compounds through central pathways, and the ability to genetically engineer the microorganisms to direct flow of carbon from central aromatic metabolic pathways to the production of valuable compounds. To develop these strategies, it is necessary to acquire a detailed understanding of the native aromatic metabolic pathways in the microorganisms to be used as chasses for lignin valorization. This study focuses on advancing the knowledge of native aromatic metabolism in *N. aromaticivorans*, a sphingomonad of interest because it efficiently degrades the major S, G, and H phenolic substituents of plant biomass (9, 15) and because it is equipped with metabolic pathways for breaking down interunit linkages in lignin (17-20). The ability of *N. aromaticivorans* to efficiently degrade many aromatic compounds may be linked to having functionally redundant aromatic degradation pathways and enzymes with broad substrate specificity. While having redundant pathways and enzyme promiscuity may confer this microorganism with ecological advantages in nature, these features can create challenges or opportunities when engineering such microorganisms to produce high yields of a desired compound. For example, I reported earlier that the 12444PDC strain of *N. aromaticivorans* is able to funnel multiple lignin-derived aromatic compounds

into PDC, but the PDC yields from S phenolics were lower than those from G and H phenolics (9). A possible explanation for this observation is pathway redundancy (16), which would allow *N. aromaticivorans* to channel a fraction of the S phenolics through one or more uncharacterized pathways that were not blocked in the 12444PDC strain.

Based on the published analysis of aromatic metabolism of *Sphingobium sp*. SYK-6 (Figure 4.1), I posited that one uncharacterized step could be the *O*-demethylation of 3-MGA to form GA with subsequent aromatic ring opening to produce OMA (Figure 4.1). Because *O*-demethylation and aromatic ring opening are also involved in other branches of the aromatic metabolism pathways, in this work I systematically evaluated these two functions in *N. aromaticivorans*. Below I discuss the new information derived from the integration of *in vivo* experiments with mutant strains and *in vitro* experiments with purified enzymes (summarized in Figure 4.11). The knowledge gained from these experiments helps us define roles for enzymes not previously described in *N. aromaticivorans* or other sphingomonads, identify functional pathway redundancy in the metabolism of S phenolics by this organism, and refine predictions of the substrate specificity of key *N. aromaticivorans* enzymes.

4.4.1 O-demethylation reactions

The two H_4folate-dependent *O*-demethylases, DesA and LigM, and the *O*-demethylase DmtS first described in this study, were each shown to have a role in the metabolism of S and G phenolics in *N. aromaticivorans* (Figure 4.11). While I obtained genetic evidence for the role of DmtS as an *O*-demethylase, we were not able to purify recombinant DmtS, so I lack direct information about its substrate specificity. Of known *O*-demethylases, DmtS is most similar in its predicted amino acid sequence (~27% identical) to the monooxygenase component (subunit A) of the vanillic acid *O*-demethylase VanAB, whose function has been demonstrated in other bacteria (29, 30) and plays a key role on aromatic metabolism in *P. putida* (27, 28). In *P. putida* and other bacteria, the *vanA* gene is found in an operon with *vanB*, which codes for the putative reductase, VanB. However, *dmtS* is not contained in a gene cluster in *N. aromaticivorans* with a gene that encodes a protein with amino acid sequence identity

to a known VanB reductase. Nevertheless, I was able to use mutants lacking dmtS to confirm its role in aromatic metabolism in *N. aromaticivorans*.

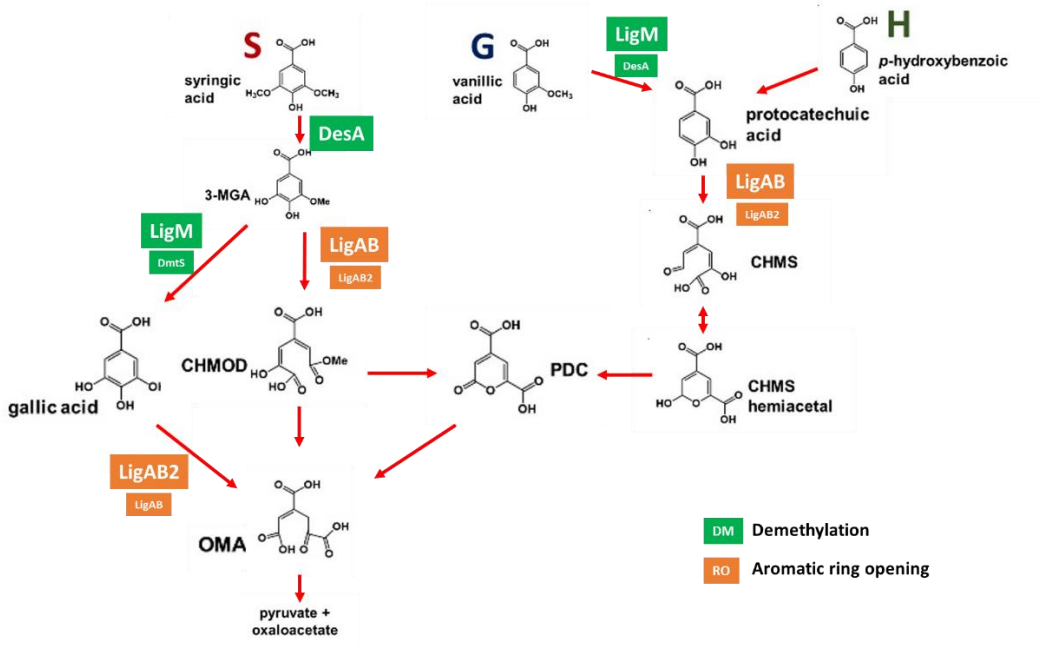


Figure 4.11. Updated pathway for metabolism of S, G, and H phenolics by *N. aromaticivorans*, with specific assignment of *O*-demethylase and aromatic ring opening dioxygenase activities based on the *in vivo* and *in vitro* experiments described in this study.

***O*-demethylation of syringic acid.** Several of our results support the role of the H_4folate-dependent *O*-demethylase DesA in demethylation of syringic acid to 3-MGA. First, the *in vitro* assays with DesA showed stoichiometric conversion of syringic acid to 3-MGA, and the *in vitro* assays with LigM preparations that were active with other aromatic compounds showed no detectable activity with this substrate (Figure 4.6). Second, *in vivo* experiments with the set of putative *O*-demethylase mutants showed that DesA was essential for *N. aromaticivorans* growth on syringic acid (Figure 4.2). Third, the experiments with the set of putative *O*-demethylase mutants constructed in the previously reported PDC-

producing strain (12444PDC) showed that three out of the four mutants lacking *desA* did not degrade syringic acid (Figure 4.3).

To our knowledge, no enzymes other than DesA have been reported to demethylate syringic acid and produce 3-MGA in any sphingomonad. Indeed, in *Sphingobium sp*. SYK-6, DesA has been proposed to be the only enzyme responsible for syringic acid *O*-demethylation and, similar to *N. aromaticivorans*, the deletion of the *desA* gene prevents the mutant strain from growing on syringic acid as the sole substrate (25). Thus, I was surprised to find PDC production when *desA* was the only gene deleted in the 12444PDC background (Figure 4.3B). To explain this result, I hypothesize that, when *N. aromaticivorans* does not use syringic acid as the growth substrate (glucose was the growth substrate with the PDC-producing strain and its derivatives), an *O*-demethylase other than DesA may perform this function. However, given the results from the *in vitro* assays (Figure 4.6), I predict that LigM is not this other *O*-demethylase, and there was no genetic evidence to link DmtS with this function.

***O*-Demethylation of 3-MGA.** Although *O*-demethylation of 3-MGA to GA has been described in *Sphingobium sp*. SYK-6 (26), evidence for enzymes involved in this reaction in *N. aromaticivorans* was lacking before our studies. Here, I provide several lines of evidence that 3-MGA is converted to GA in *N. aromaticivorans*, but it is not the major route of 3-MGA metabolism. Instead, our results support the hypothesis that aromatic ring opening of 3-MGA to CHMOD is the major route for the metabolism of syringic acid. First, the non-stoichiometric conversion of syringic acid to PDC in the 12444PDC strain (Figure 4.3A) is evidence for the existence of a secondary pathway that supports the degradation of a small fraction (~15%) of the syringic acid in the PDC-producing strain. Second, stoichiometric conversion of syringic acid to PDC is achieved in the mutant that has an intact *desA* (necessary for the conversion of syringic acid to 3-MGA) and lacks both *ligM* and *dmtS* (Figure 4.3G), leading to the hypothesis that both LigM and DmtS are active in this secondary pathway. This hypothesis was further confirmed in experiments using 3-MGA as the aromatic substrate (Figure 4.4), where stoichiometric conversion of 3-MGA to PDC is only achieved by simultaneous deletion of *ligM* and *dmtS* (Table 4.2).

Third, the *in vitro* assays with purified recombinant LigM (Figure 4.6) showed that GA was stoichiometrically produced from 3-MGA. Finally, when comparing PDC yields from the 12444PDC strain and the strains with deletion of either *dmtS* or *ligM* (Table 4.2), a small but reproducible increase in PDC yield occurs when *ligM* is deleted, suggesting that *in vivo* LigM is more active in 3-MGA demethylation than DmtS.

Based on a comparison of our results to that from other labs, the role of 3-MGA *O*-demethylation in the metabolism of syringic acid appears to be different in *N. aromaticivorans* and *Sphingobium sp*. SYK-6. For instance, inactivation of *ligM* in *N. aromaticivorans* did not cause a detectable effect in growth on syringic acid (Figure 4.2A), whereas inactivation of this gene in *Sphingobium sp*. SYK-6 has a detrimental effect on both growth rate and final cell density when cells use syringic acid (25). This suggests that between the two functionally redundant pathways for converting 3-MGA to OMA in *N. aromaticivorans*, ring opening to CHMOD carries more flow of carbon than demethylation to GA, which is the opposite of what was reported for *Sphingobium sp*. SYK-6. In addition, our results also suggest redundancy in the enzymes that demethylate 3-MGA to GA in *N. aromaticivorans*, with both LigM and DmtS having this activity. *Sphingobium sp*. SYK-6 appears to have less enzyme redundancy in this reaction, as deletion of *ligM* eliminated 3-MGA conversion to GA in cell extract experiments (25).

***O*-demethylation of vanillic acid.** Based on our data, I propose that LigM plays a major role in demethylation of vanillic acid, but that DesA can perform this reaction, although with reduced efficiency (Figure 4.11). First, the *in vitro* assays showed that both these enzymes could convert vanillic acid to PCA, with LigM having a faster degradation rate (Figure 4.6). Second, *in vivo* experiments with the set of *O*-demethylase mutants in the wild-type background showed that all mutants with an intact *ligM* gene grew as well as the parent strain, strains lacking *ligM* but with intact *desA* showed a detectable growth defect, and strains lacking both *ligM* and *desA* could not grow in the presence of vanillic acid (Figure 4.2). Consistent with these findings, our *in vivo* experiments with the set of mutants constructed in the 12444PDC background confirmed that deleting *ligM* slowed down the rates of vanillic acid degradation,

deleting *desA* alone did not have an effect, but that deleting both *ligM* and *desA* prevented vanillic acid degradation (Figure 4.5). Finally, a role for DmtS in vanillic acid *O*-demethylation was not evident in the results of any of the experiments. These observations are consistent with the predicted role of LigM and DesA in the metabolism of vanillic acid by *Sphingobium sp*. SYK-6 (25).

4.4.2 Aromatic ring opening reactions

Aromatic ring opening of phenolic compounds in sphingomonads is predicted to be catalyzed by extradiol dioxygenases (cleaving at the 4,5 position), which often exhibit broad substrate specificity (16, 31, 32). In *Sphingobium sp*. SYK-6, at least three dioxygenases are reported to be involved in metabolism of S and G compounds: LigAB, with highest activity on PCA (33), DesZ, with highest activity on 3-MGA and GA (21), and DesB, specific for GA (24). In contrast, previous analysis of *N. aromaticivorans* has implicated only one extradiol dioxygenase, LigAB, in aromatic ring opening of phenolic compounds (15). In this study, I showed that two aromatic ring opening enzymes in *N. aromaticivorans*, LigAB and LigAB2, catalyzed extradiol cleavage of 3-MGA, GA, and PCA, the three intermediates the pathways for metabolism of S, G, and H aromatics whose rings are capable of being enzymatically opened by a 4,5-dioxygense (Figure 4.11). In addition, our results show that LigAB has greater activity with 3-MGA and PCA than LigAB2, whereas LigAB2 has higher activity with GA than LigAB.

Aromatic ring opening of 3-MGA. Our results support the hypothesis that LigAB is the primary enzyme for 3-MGA ring opening in *N. aromaticivorans* (15). They also provide new evidence that LigAB2 has activity with 3-MGA and could partially substitute for the role of LigAB.

The *in vitro* assays using recombinant LigAB and LigAB2 proteins with 3-MGA as a substrate showed accumulation and disappearance of the known stereoisomers of CHMOD, and accumulation of PDC (Figure 4.10A, 4.10B). *In vitro* accumulation of PDC from 3-MGA has also been observed in enzyme assays with the LigAB of *Sphingomonas sp*. SYK-6 (22). Rapid PDC accumulation in those experiments led to the proposal that PDC is a product of LigAB activity in *Sphingomonas sp*. SYK-6. Several results

from our reactions of LigAB and LigAB2 from *N. aromaticivorans* with 3-MGA lead us to the different conclusion that CHMOD is the direct product of the LigAB and LigAB2 reactions with 3-MGA, and that CHMOD is non-enzymatically converted to PDC under the conditions of the assays. First, CHMOD is unstable in aqueous solution, and non-enzymatic cyclization of CHMOD to PDC at neutral pH has been reported (34). Unfortunately, the CHMOD appears to cyclize under our reaction conditions too fast for us to be able to purify this compound for determination of its molar concentration in our assays. However, the sum of the molar 3-MGA and PDC concentrations over the course of the enzymatic reactions inversely correlated with the accumulation of the CHMOD isomers (Figure 4.S7), suggesting CHMOD is an intermediate in the conversion of 3-MGA to PDC. Third, at the end of the LigAB reaction, when 3-MGA and CHMOD are not detectable (Figure 4.10A), PDC accumulation reaches the initial concentration of 3-MGA in the assay, indicating the absence of any other potential products of CHMOD degradation. Finally, the results of *in vitro* assays of LigAB and LigAB2 with other substrates (Figure 4.10) produced products consistent with the proposed function of these enzymes as extradiol aromatic ring opening dioxygenases, and CHMOD is the expected product of such reaction when 3-MGA is the substrate (Figure 4.11).

GA ring opening. The *in vitro* evidence obtained in this study indicates that LigAB2 reacts more rapidly with GA than with 3MGA or PCA, and confirms OMA as the product of GA ring opening (Figure 4.10). The experiments with the set of deletion mutants in the 12444PDC background (Figure 4.8) also provided evidence that either LigAB or LigAB2 could be responsible for metabolizing the fraction of syringic acid that is normally channeled through GA in the 12444PDC strain. However, under the conditions of our studies, the GA that accumulated in the medium apparently converted abiotically to an unknown product. Thus, further research is needed to specifically ascertain the role of LigAB and LigAB2 on GA transformation *in vivo*.

PCA ring opening. The results of *in vitro* (Figure 4.10) and *in vivo* (Figure 4.9) experiments with vanillic acid provided evidence that the ring opening of PCA is primarily catalyzed by LigAB, but that LigAB2

could substitute *in vivo* when *ligAB* was deleted. Furthermore, deleting both sets of genes eliminated growth (Figure 4.7) and PDC production (Figure 4.9), indicating that no other enzyme in *N. aromaticivorans* could catalyze the ring opening of PCA under our growth and media conditions.

4.4.3 Implications for PDC production from lignin-derived aromatics by N. aromaticivorans

I have previously tested PDC production from plant-derived phenolics by *N. aromaticivorans* (9) because this native pathway intermediate is a potential building block for bio-based plastic and epoxy adhesives (35). I showed that strain 12444PDC, which contains mutations that block the conversion of PDC and CHMOD to OMA , could transform S, G, and H phenolics to PDC, but exhibited lower PDC yields from S phenolics than from G or H phenolics (9). Based on this observation I proposed that strain 12444PDC contained additional, previously unidentified, gene products that could metabolize S phenolics (9). In this study, I identified these previously unknown enzymes in the proposed alternative pathway for metabolism of syringic acid in *N. aromaticivorans* that *O*-demethylates 3-MGA to GA (Figure 4.11). I further showed that these same enzymes catalyze ring opening of GA to OMA (Figure 4.11). The activity of the enzymes in this previously unknown alternative pathway helps explain why strain 12444PDC had lower PDC yields from S phenolics, since our data suggest that ~15% of the 3-MGA follows the GA-OMA route and does not contribute to PDC production in strain 12444PDC. Although deleting the genes in this alternative pathway could increase the yield of PDC, the broad substrate specificity of the other *N. aromaticivorans O*-demethylases described in this study implies that single gene deletions may not be sufficient to maximize PDC yields. As predicted, stoichiometric production of PDC from either syringic acid (Figure 4.3G) or 3-MGA (Figure 4.4G) was only possible with the deletion of both *ligM* and *dmtS* from the 12444PDC background. However, since *O*-demethylases are also needed in the demethylation of G phenolics, a potential side effect of *ligM* deletion (the main enzyme I found to be responsible for vanillic acid demethylation) could be the reduced conversion of G phenolics into PDC. Indeed, I observed the predicted negative effect in the rate of vanillic acid conversion when *ligM* was deleted (Figures 4.5C) although the PDC yield was not affected when this gene was inactivated in the 12444PDC strain (Table

4.2). Since DesA was shown to have activity with vanillic acid, one potential strategy to compensate for the absence of LigM could be to engineer PDC-producing strains that overexpress *desA*, as has been done with overexpression of *vanAB* in *P. putida* to accelerate vanillic acid degradation (36).

The results of our studies also make a new prediction that deleting enzymes responsible for the aromatic ring-opening step does not appear to be a productive strategy for reducing the flow of substrates through this alternative *N. aromaticivorans* pathway due to the broad substrate specificity of these dioxygenases. For example, while *in vitro* assays with the newly discovered LigAB2 showed its preference for GA over 3-MGA and PCA, its inactivation did not produce measurable effects on PDC yield (Table 4.4). Furthermore, inactivation of both LigAB and LigAB2 would also prevent aromatic ring opening with other biomass-derived phenolic substrates. From our data, I propose that the main function of LigAB2 *in vivo* is in ring opening reactions of aromatic compounds other than S, G, and H phenolics. While *Sphingobium sp*. SYK-6 has a predicted homologue of *N. aromaticivorans* LigAB2 (encoded by SLG_37520 and SLG_37530), I are not aware of any studies that investigate its role in S, G, or H-aromatic metabolism.

4.4.4 Concluding Remarks

N. aromaticivorans has many properties that make it attractive for use as a bacterial chassis for lignin valorization from monomeric and oligomeric mixtures of plant-derived aromatics. The diversity and redundancy of pathways for aromatic metabolism, and the broad substrate specificity of key enzymes in these pathways, allows it to degrade a large variety of S, G, and H phenolics (9) and other aromatic substrates (37). In addition, when presented with mixtures of phenolics from depolymerized lignin, *N. aromaticivorans* can simultaneously degrade these aromatic substrates (9). This organism also has pathways to degrade oligomeric lignin products, and pathway redundancies have also been identified within those networks (17, 18). Despite this, the lack of prior knowledge of aromatic and other metabolic pathways in this and other sphingomonads currently limits strategies to engineer strains and synthetic pathways that produce valuable products from biomass-derived phenolics.

In this study, I identified a previously unknown alternative route for syringic acid catabolism, described new enzymes that are involved in this pathway (DmtS and LigAB2), and performed a systematic genetic and enzymatic analysis of the *O*-demethylases and aromatic ring opening dioxygenases that function in *N. aromaticivorans*. The studies revealed that these *O*-demethylases and dioxygenases have broad substrate specificity suggesting they are able to participate in different reactions within individual pathways (Figure 4.11). The new knowledge on the metabolism of aromatic compounds by *N. aromaticivorans* obtained in this work has allowed us to design a strain (12444PDC*ΔligMΔdmtS*) with increased yield of PDC from syringic acid. It also will enable better predictions of metabolic routes that will facilitate engineering strains for improved yields of other desirable products from biomass-derived aromatics. While the interruption of 3-MGA *O*-demethylation (by deleting *ligM* and *dmtS*) in strain 12444PDC resulted in stoichiometric PDC production from syringic acid and vanillic acid, this new strain exhibited a reduced rate of vanillic acid degradation. This observation illustrates additional steps in aromatic acid degradation that are targets for future strain improvement.

4.5 Materials and methods

Bacterial strains, growth media and culturing conditions. Two variants of *N. aromaticivorans* DSM12444, strains 12444*Δ1879*, which lacks the gene Saro_1879 (putative *sacB*; Saro_RS09410) (17), and strain 12444*ΔligIΔdesCD* (also called 12444PDC), which lacks the genes Saro_1879, Saro_2819 (Saro_RS14300), Saro_2864 (Saro_RS14525), and Saro_2865 (Saro_RS14530) (9) were used as parent strains to generate the mutant strains shown in Table 4.1 and Table 4.4. All genetic modifications were carried out using a variant of the plasmid pK18*mobsacB* (38), which contains *sacB* and a kanamycin resistance gene. The gene deletions were performed as previously described (9), with the details of the processes used here described in the Supplementary Information. All primers, plasmids, and *Escherichia coli* strains used for cloning and protein expression are listed in supplemental Tables 4.S1, 4.S2, and 4.S3, respectively.

E. coli cultures were grown in LB media containing 50 μg ml-1 kanamycin at 37 °C. *N. aromaticivorans* cultures were grown in SMB media (see Supporting Information of Kontur et al. (2018) (17) for recipe) supplemented with the indicated carbon source at 30 °C. For routine culture and storage, SMB was supplemented with 1 g L^{-1} glucose. For constructing mutants, SMB was supplemented with 1 g L^{-1} glucose and either 50 μg ml^{-1} kanamycin, or 100 g L^{-1} sucrose as necessary. Cell density was monitored using a Klett-Summerson photoelectric colorimeter with a red filter.

***N. aromaticivorans* growth experiments.** *N. aromaticivorans* cultures were grown overnight in SMB medium supplemented with 1 g L^{-1} glucose. Cultures were diluted 1:1 with fresh medium containing 1 g L^{-1} glucose and incubated for 1 hour to resume cell growth. Cells from 2 ml of the growing culture were pelleted (5 min at 2,300 × g), washed with 1 ml SMB media without carbon, then resuspended into 600 μl SMB media with no added carbon. Growth experiments were initiated by adding 80 μl of the cell suspension into 8 ml fresh SMB media supplemented with either 3 mM syringic acid or 3 mM vanillic acid. Cultures were grown aerobically in 20 ml test tubes, shaken at 200 rpm at 30°C. Each experiment was repeated 3 times.

***N. aromaticivorans* extracellular metabolite analysis.** Bacterial cell cultures were grown overnight in 20 ml SMB media supplemented with 1 g L^{-1} glucose, then reactivated by adding 2 ml fresh media containing 1 g L^{-1} glucose and incubated for 1 hour. Experiments were initiated by inoculating 5 ml active culture into 15 ml fresh SMB media supplemented with 5 mM glucose and either 4 mM syringic acid, 4 mM 3-MGA, or 4 mM vanillic acid. Cultures were grown aerobically in 125 mL conical growth flasks, shaken at 200 rpm at 30°C. Samples were collected periodically, filtered (through 0.22 μm pores) to remove cells, and immediately analyzed by HPLC-MS to monitor the extracellular aromatic compounds.

Recombinant enzyme expression and purification. Genes Saro_2812/3 (*ligAB*), Saro_1233/4 (*ligAB2*), Saro_2861 (*ligM*), and Saro_2402 (*desA*) from *N. aromaticivorans* were independently cloned into the plasmid pVP302K (19), which incorporates a His_8-tag to the N-terminus of the translated transcripts connected via a tobacco etch virus (TEV) protease recognition site (see Supplementary Information for

plasmid construction details). The expression plasmids were transformed into *E. coli* B834 (39, 40) containing plasmid pRARE2 (Novagen, Madison, WI) and transformants grown in ZYM-5052 Autoinduction Medium (41) containing 50 μg ml^{-1} kanamycin and 20 μg ml^{-1} chloramphenicol. Recombinant proteins were purified by passing crude *E. coli* lysates through a nickel-nitrilotriacetic acid (Ni-NTA) column as described previously (17). His_8 tags were cleaved from recombinant proteins using TEV protease, and proteins were passed again though a Ni-NTA column to remove the cleaved His_8 tag and the TEV protease (which contains its own His tag).

***In vitro* aromatic ring-opening dioxygenase assays.** Preliminary experiments with recombinant LigAB (Saro_2812/3) and LigAB2 (Saro_1233/4) purified in the presence of air suggested that the enzymes were catalytically inactive, as reported for other homologues of these proteins (42). I thus separately mixed LigAB and LigAB2 with Reactivation Buffer (Buffer A, containing 20 mM Na2HPO4, 20 mM KH2PO4, 1 mM Fe2SO4 and 1 mM ascorbic acid, and prepared anaerobically) to a concentration of 2 μM enzyme active sites under anaerobic conditions, and incubated them anaerobically at 30°C for 21 hours to reactivate the enzymes. Three solutions containing Buffer B (20 mM Na2HPO4, 20 mM KH_2PO_4, and 1 mM ascorbic acid) and either 2 mM 3-MGA, 2 mM PCA or 2 mM GA were prepared in the presence of air. Enzyme assays were performed in 2ml polypropylene vials inside an anaerobic chamber at 30°C by mixing 750 μL of reactivated enzyme in Buffer A with 750 μL of aromatic substrate in Buffer B. As a control, I mixed 750 μL of Buffer A without enzyme with 750 μL each aromatic substrate in Buffer B. After 30 min, the vials were exposed to the ambient atmosphere outside of the anaerobic chamber for 10 min to expose the reactions to the O_2 predicted to be a substrate for the ring-opening reaction, then transferred back into the anaerobic chamber. Exposure to O_2 was repeated at 2 hours and 4 hours after the assay was initiated. Samples (250 μL) were collected at time zero, at 30 minutes and at 1, 2, 4, and 24 hours, and immediately mixed with 50 μL of 1N HCl to terminate the reaction, then analyzed with HPLC-MS to quantify substrate disappearance and formation of any products. Identification of OMA, CHMOD,

and PDC were performed by GC-MS of samples extracted with ethyl acetate immediately after collection. CHMS was identified by UV-vis spectrophotometry. Assays were performed in triplicate.

Aromatic *O*-demethylase enzyme assays. Enzyme assays were performed in triplicate under anaerobic conditions at 30°C (because of the expected O_2 sensitivity of H_4folate). Recombinant LigM (Saro_2861) and DesA (Saro_2404) were separately mixed with Buffer C (20mM Na_2HPO_4, 20mM KH_2PO_4, and 2mM H_4folate, and prepared anaerobically) to concentrations of 2 μM or 1 μM enzyme active sites under anaerobic conditions. Three independent solutions containing Buffer D (20mM Na_2HPO_4 and 20mM KH_2PO_4) and either 1 mM 3-MGA, 1 mM PCA or 1 mM GA were prepared aerobically. Enzyme reactions were initiated in 2 ml polypropylene vials inside an anaerobic chamber at 30°C by mixing 750 μL of an enzyme in Buffer C with 750 μL an aromatic substrate in Buffer D. A control was run by mixing 750 μL of Buffer C without enzyme with 750 μL each aromatic substrate in Buffer D. Samples (250 μL) were collected at time zero, at 30 minutes and at 1, 2, 4, and 24 hours, and immediately mixed with 50 μL of 1N HCl to stop the reaction, then analyzed with HPLC-MS to quantify substrate disappearance and product formation.

Analysis of extracellular metabolites and enzyme reaction products. All culture supernatants and *in vitro* enzyme assay samples were filtered (0.2 μm) prior to chemical analysis. Quantitative analyses of aromatic compounds were performed on a Shimadzu triple quadrupole liquid chromatography mass spectrometer (LC-MS) (Nexera XR HPLC-8045 MS/MS). Reverse-phase HPLC was performed using a binary gradient mobile phase consisting of Solvent A (0.2% formic acid in water) and methanol (gradient profile shown in Figure 4.S8), and a Kinetex F5 column (2.6 μm pore size, 2.1 mm ID, 150 mm length, P/N: 00F-4723-AN). All compounds were detected by multiple-reaction-monitoring (MRM) and quantified using the strongest MRM transition (Table 4.S4).

Identification of OMA and PDC was performed by gas chromatography-mass spectrometry (GC-MS). Sample aliquots (150 μL) were acidified with HCl to pH < 2, and ethyl acetate extracted (3 × 500 μL). The three extraction samples were combined, dried under a stream of N_2 at 40 °C, and derivatized by the

addition of 150 µL of pyridine and 150 µL of N,O-bis(trimethylsilyl)trifluoro- acetamide with trimethylchlorosilane (99 : 1, w/w, Sigma) and incubated at 70 °C for 45 min. The derivatized samples were analyzed on an Agilent GC-MS (GC model 7890A, MS Model 5975C) equipped with a (5% phenyl)-methylpolysiloxane capillary column (Agilent model HP-5MS). The injection port temperature was held at 280 °C and the oven temperature program was held at 80 °C for 1 min, then ramped at 10 °C min^{-1} to 220 °C, held for 2 min, ramped at 20 °C min^{-1} to 310 °C, and held for 6 min. The MS used an electron impact (EI) ion source (70 eV) and a single quadrupole mass selection scanning at 2.5 Hz, from 50 to 650 m/z. The data was analyzed with Agilent MassHunter software suite.

The product of PCA aromatic ring opening, predicted to be CHMS, was analyzed by its conversion into 2,4-pyridinedicarboxylic acid (PDCA). 100 µl of sample were pH neutralized by the addition of 10 µl NaOH 1.67 N solution. In addition, 5 µl $(NH_4)_2SO_4$ 10% solution was added and then incubated at room temperature for 24 h. Samples were analyzed by HPLC-UV using the same HPLC conditions described above. Eluent was analyzed for light absorbance between 190 and 400 nm by a Shimadzu SPD-M20A spectrophotometer.

Chemicals. All SMB media reagents, gallic acid, and vanillic acid were purchased from SigmaAldrich (St Louis, MO). Syringic acid was purchased from TCI (Tokyo Chemical Industry)-America (Portland, OR). 3-MGA was purchased from Carbosynth (Berkshire, UK). Protocatechuic acid was purchased from Sigma Aldrich (St Louis, MO).

4.6 Acknowledgements

This work was supported by U.S. Department of Energy (DOE) Great Lakes Bioenergy Research Center grant (DOE Office of Science BER DE-SC0018409). Additional funding from the Chilean National Commission for Scientific and Technological Research (CONICYT) as a fellowship to Jose M. Perez is also acknowledged.

4.7 Supplementary materials

4.7.1 Supplemental methods

Construction of plasmids for generating in-frame deletions of Saro_2861, Saro_2404, Saro_1872, Saro_2812/3, or Saro_1233/4. Regions of *N. aromaticivorans* genomic DNA containing ~1100 bp upstream and downstream of Saro_2861, Saro_2404, Saro_1872, Saro_2812/3, or Saro_1233/4 were PCR amplified separately using the pairs of primers Saro_2861-pK18_Amp-R/Saro_2861_Del-F, Saro_2861-pK18_Amp-F/Saro_2861_Del-R, Saro_2404-pK18_Amp-R/Saro_2404_Del-F, Saro_2404-pK18_Amp-F/Saro_2404_Del-R, Saro_1872-pK18_Amp-R/Saro_1872_Del-F, Saro_1872-pK18_Amp-F/Saro_1872_Del-R, Saro_2812/3-pK18_Amp-R/Saro_2812/3_Del-F, Saro_2812/3-pK18_Amp-F/Saro_2812/3_Del-R, Saro_1233/4-pK18_Amp-R/Saro_1233/4_Del-F, and Saro_1233/4-pK18_Amp-F/Saro_1233/4_Del-R (Table 4.S1). The pairs of DNA amplified flanking regions for each gene were combined with linearized pK18msB-MCS1 (17) using NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs, Ipswich, MA) to produce the plasmids pK18msB/ΔSaro2861, pK18msB/ΔSaro2404, pK18msB/ΔSaro1872, pK18msB/ΔSaro2812/3, and pK18msB/ΔSaro1233/4, respectively. The plasmids were then transformed into NEB 5-alpha competent *E. coli* (New England Biolabs). The transformed *E. coli* cells were then cultured in LB media + kanamycin, the plasmids purified using a Quiagen® Plasmid Maxi Kit (Qiagen, Germany), and DNA sequencing was used to confirm the presence of the desired deletion.

Deletion of Saro_2861, Saro_2404, Saro_1872, Saro_2812/3, or Saro_1233/4 from *N. aromaticivorans*. The plasmids constructed above were individually transformed into competent *E. coli* S17-1 and subsequently mobilized into different strains of *N. aromaticivorans* (see Table 4.1) via bacterial conjugation. Transconjugant cells of *N. aromaticivorans* (single cross overs) were isolated on SMB plates containing 1 g/L glucose and 50 ug/mL kanamycin. To generate and isolate cells that subsequently eliminated the plasmid via a second instance of homologous recombination (double crossovers), single crossover cells were cultured on solid SMB media containing 1 g/L glucose and 10%

sucrose. PCR amplified regions of the target genes were sequenced to verify the colonies in which the desired genes were deleted.

Construction of protein expression plasmids. Saro_2861, Saro_2404, Saro_2812/3, and Saro_1233/4 were individually amplified from *N. aromaticivorans* DSM12444 genomic DNA with the pairs of primers Saro_2861_pVP-F and Saro_2861_pVP-R, Saro_2404_pVP-F and Saro_2404_pVP-R, Saro_2813-2_pVP-HiFi_start and Saro_2813-2_pVP-HiFi_stop, and Saro_1233/4_pVP-F and Saro_1233/4_pVP-R, respectively (Table 4.S3). Each resulting DNA fragment was combined with linearized pVP302K (19) using NEBuilder® HiFi DNA Assembly Master Mix, as previously described (18), generating plasmids pVP302K/Saro2861, pVP302K/Saro2404, pVP302K/Saro2812/3, and pVP302K/Saro1233/4, which consist of a T5 promoter followed by coding sequences for a His_8-tag, a tobacco etch virus (TEV) protease recognition site, and the coding sequence of the *N. aromaticivorans* gene(s).

4.7.2 Supplemental results

Identification of products from experiments with recombinant LigAB and LigAB2. To investigate the activities of LigAB and LigAB2 and identify reaction products using different substrates (3-methoxygallic acid (3-MGA), gallic acid (GA), and protocatechuic acid (PCA)), we purified recombinant forms of the proteins and tested them for activity *in vitro*. We know from studies with homologs of these enzymes that they are prone to inactivation by oxidation of an active site Fe^{2+} (42). To minimize O_2 inactivation, while also providing the O_2 that is expected to be one of the enzyme substrates, I performed these assays using a hybrid anaerobic/aerobic method consisting of exposing the reaction vials to air for periods of 5 min after 10 min, 30 min, 2 h, and 4 h have elapsed in the reaction. Samples (250 μL) were collected before adding enzyme, at indicated times, and at 24 h after reaction initiation, and mixed with 50 μL 1N HCl to inactivate the enzyme before analysis. For assays using 3-MGA or GA as a substrate, samples were analyzed by HPLC-MS to monitor disappearance of the substrate and the formation of CHMOD, PDC, and OMA. GC-MS was used to verify the identity of PDC and OMA. For assays using PCA, all samples were analyzed by HPLC-MS to monitor the substrate disappearance and the formation

of CHMS. To verify the identity of CHMS, it was converted into 2,4-pyridinedicarboxylic acid (PDCA) and analyzed by HPLC-UV. Its light absorbance spectrum was compared with authentic PDCA.

When tested with 3-MGA as the substrate, HPLC-MS analysis showed the transient accumulation of 2 compounds with m/z = 215 in the negative mode, which is consistent with them being stereoisomers of CHMOD (MW = 216.15). GC-MS analysis of the TMS-derivatized ethyl acetate extracts of samples collected after 1 hour of reaction, only showed the presence of TMS-PDC (Figures 4.S2 and 4.S3), suggesting that any remaining CHMOD present in the samples cyclized into PDC during sample storage and the derivatization process before analysis.

When tested with GA as the substrate, HPLC-MS analysis showed the accumulation of a compound with m/z = 201 in the negative mode, which is consistent with it being OMA (MW = 202.12). GC-MS analysis of the TMS-derivatized ethyl acetate extracts of samples collected after 24 hour of reaction showed a compound with a fragmentation pattern identical to that previously found for the TMS derivative of the enol form of OMA (24) (Figures 4.S4 and 4.S5). This analysis also showed the presence of a TMS derivative of PDC, attributed to the production of PDC from the reaction of OMA with HCl, which was added to quench the reaction (43).

When tested with PCA as the substrate, HPLC-MS analysis showed the accumulation of a compound with m/z = 185 in the negative mode, which is consistent with it being CHMS (MW = 186.12). Samples collected after 24 hours of reaction and treated with $(NH_4)_2SO_4$ were analyzed by HPLC-UV. Samples from *in vitro* assays with LigAB showed the presence of a peak with identical retention time and light absorbance spectrum (Figures 4.S6C and 4.S6D) of that of authentic PDCA (Figures 4.S6A and 4.S6B). Samples from *in vitro* assays with LigAB2 showed a peak similar to PDCA and a peak corresponding to unreacted PCA (Figures 4.S6E and 4.S6F). Samples from control condition without enzyme only showed the presence of PCA (Figure 4.6G).

4.7.3 Supplemental figures and tables

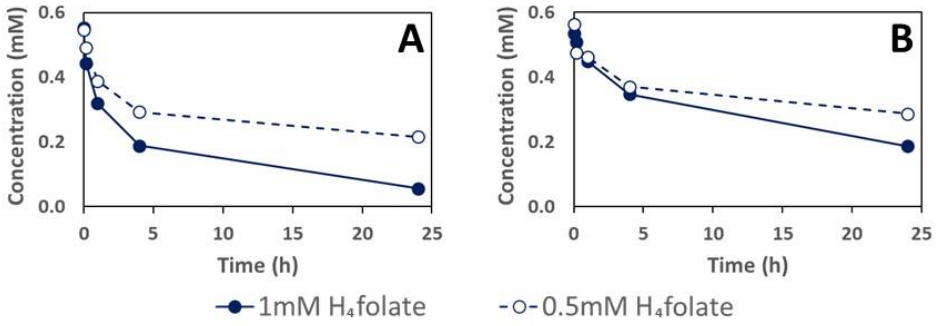


Figure 4.S1. Vanillic acid concentration during *in vitro* enzyme assays of LigM (Saro_2861) (A) and DesA (Saro_2404) (B) with 1mM or 0.5mM H_4folate. In conditions with 0.5 mM H_4folate, the substrate decreases slower than in conditions with 1 mM H_4folate.

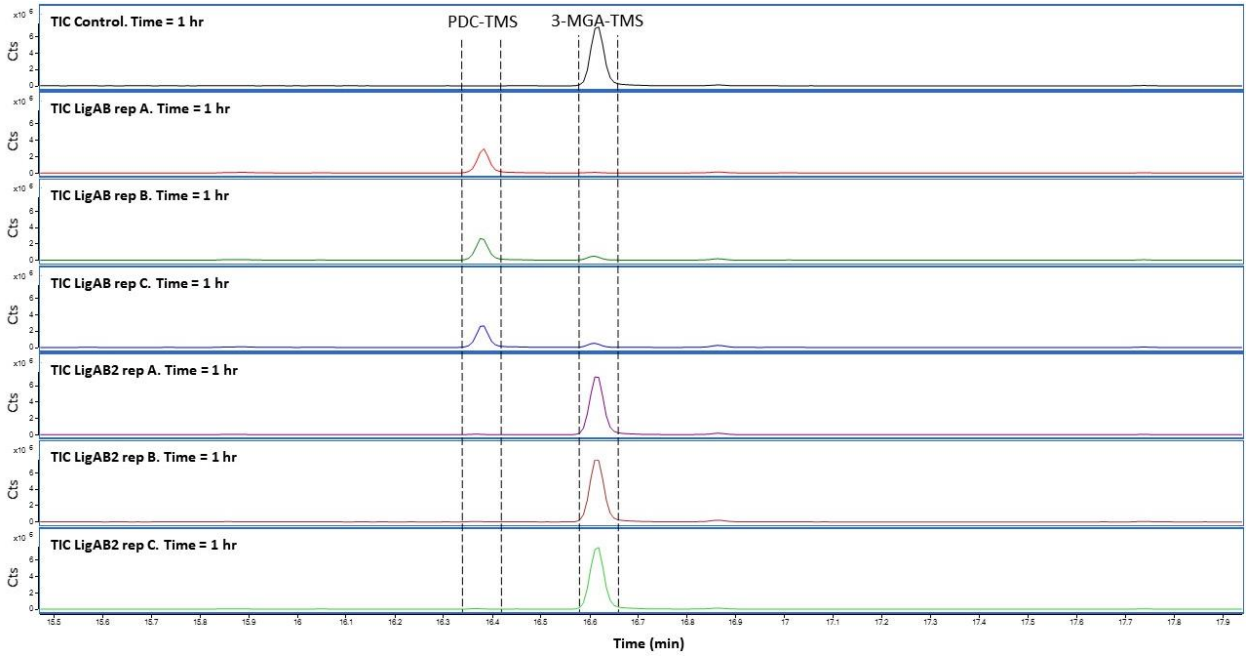


Figure 4.S2. Total ion chromatogram of samples collected after 1 hour of reaction time of *in vitro* assays with LigAB and LigAB2 incubated in the presence of 3-MGA.

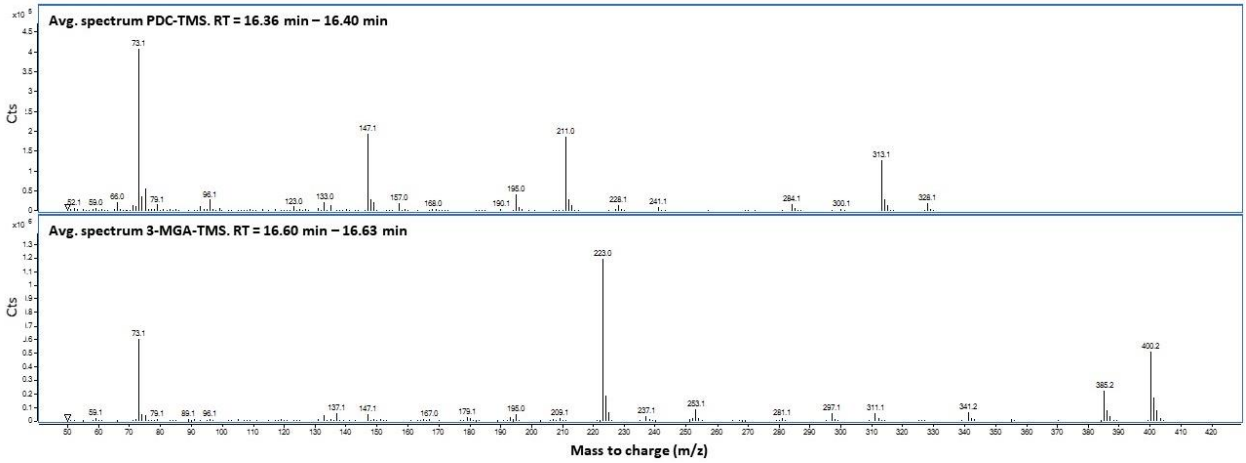


Figure 4.S3. Average mass spectrum (m/z) of PDC (RT = 16.36 m – 16.40 m) and 3-MGA (RT = 16.60 m – 16.63 m) of samples collected after 1 hour of reaction time of *in vitro* assays with LigAB and LigAB2 incubated in the presence of 3-MGA.

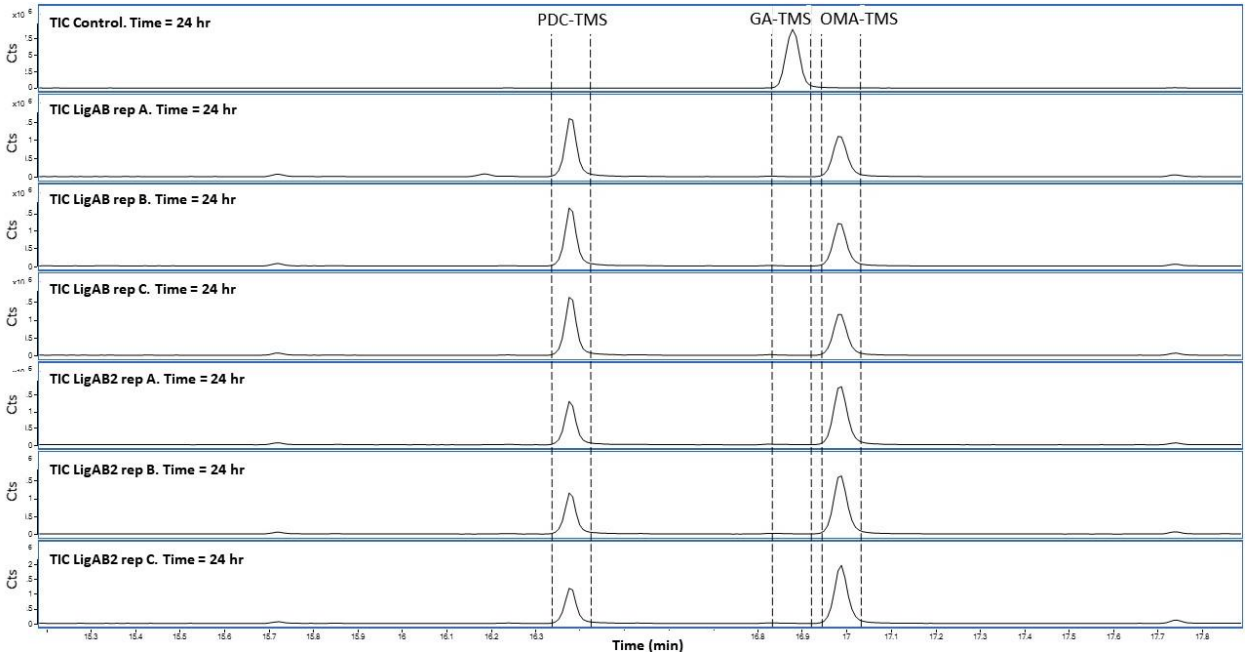


Figure 4.S4. Total ion chromatogram of samples collected after 24 hour of reaction time of *in vitro* assays with LigAB and LigAB2 incubated in the presence of GA.

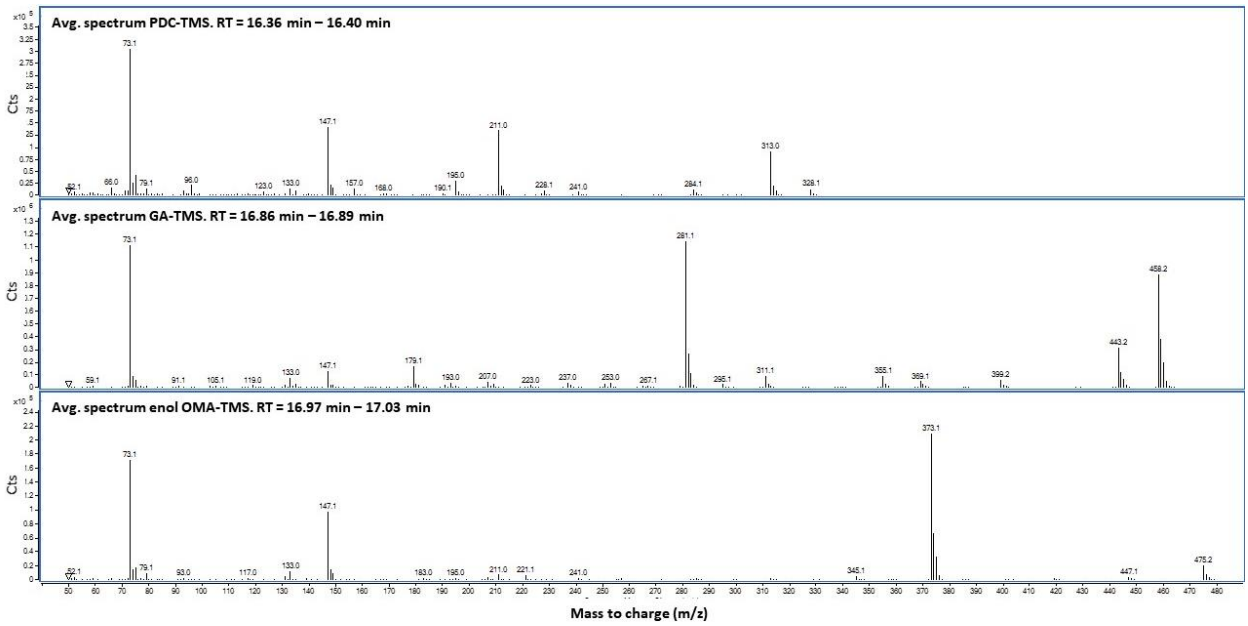


Figure 4.S5. Average mass spectrum (m/z) of PDC-TMS (RT = 16.36m – 16.40 m), GA-TMS (RT = 16.86 m – 16.89 m), and enol OMA-TMS (RT = 16.97 m – 17.03 m) of samples collected after 24 hours of reaction time of *in vitro* assays with LigAB and LigAB2 incubated in the presence of GA.

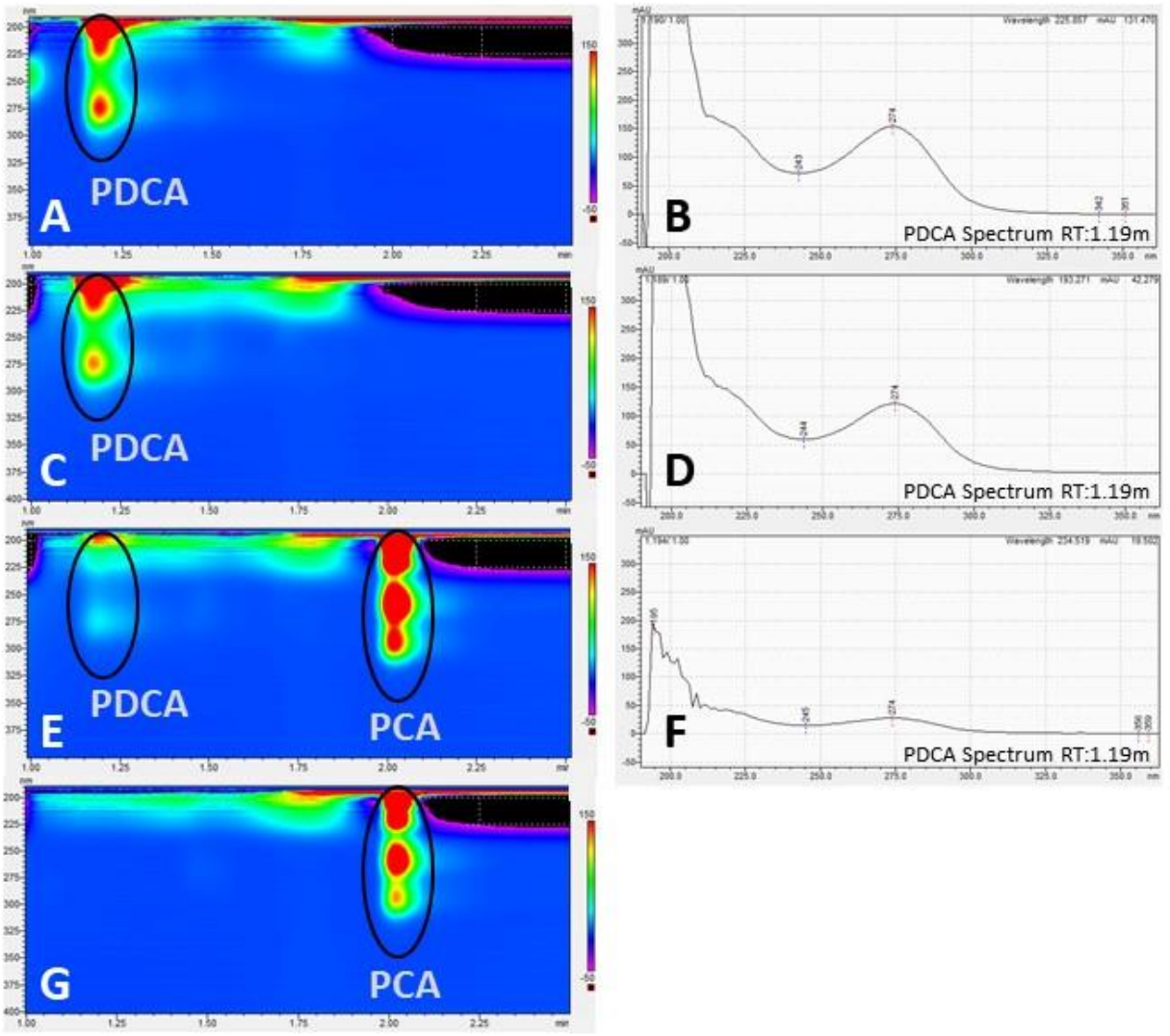


Figure 4.S6. UV-vis chromatogram and spectra of authentic PDCA standard (A and B), ammonium sulfate treated samples collected at 24 hours from *in vitro* assays with LigAB and PCA (C and D), LigAB2 and PCA (E and F), and control without enzyme (G).

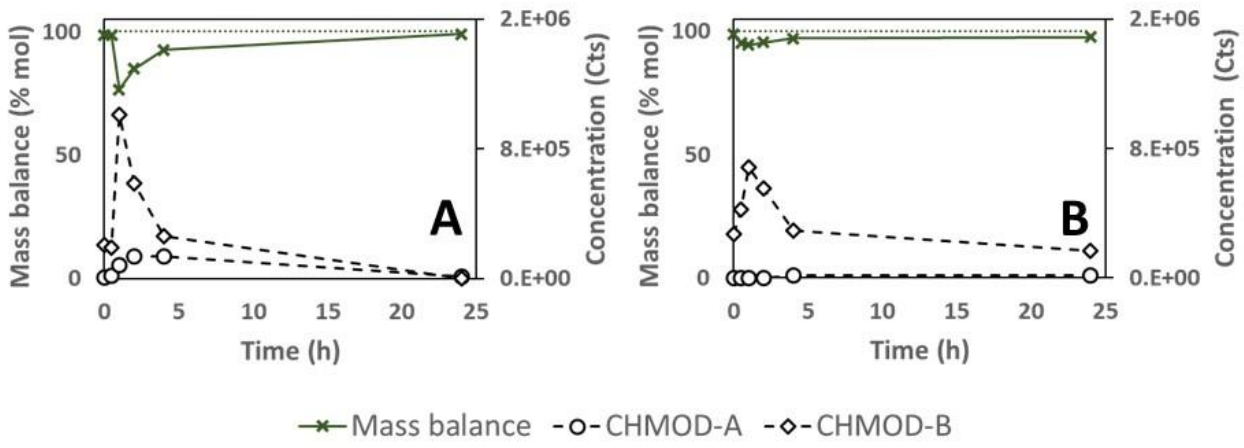


Figure 4.S7. Mass balance and CHMOD relative concentration from *in vitro* assays of LigAB (Saro_2812/3) (A) and LigAB2 (Saro_1233/4) (B) with 3-MGA as a substrate. Mass balance was calculated as the sum of the molar concentrations of 3-MGA and PDC, divided by the concentration of 3-MGA in a sample lacking enzyme. Results are expressed in percentage (%). Dotted lines show 100%. Samples were analyzed with HPLC-MS, values correspond to the integrated area of peaks generated by the EIC with m/z = 215, attributed to CHMOD and concentrations correspond to the average of three replicates.

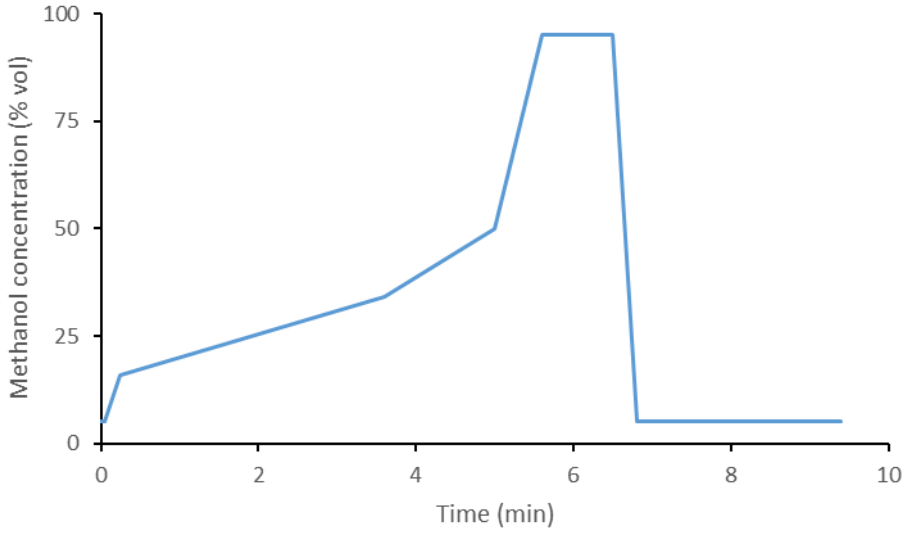


Figure 4.S8. Binary gradient profile of HPLC-MS analysis

Table 4.S1. Primers used in this study.

Name	Sequence	Notes
Saro_2861-pK18_Amp-R	5'-**GTTTCTGCGGACTGGCTTTCTAGATGTTC**CGTTGACGTGCATGTCTGTCCTCTC-3'	Region in bold matches sequence in pK18msB-MCS1
Saro_2861_Del-F	5'-**CTTCTCGTTGTACGAGTAGCCG**GACTGACTCTCCCGACTTGAAAATGG-3'	Region in bold matches sequence in Saro_2861_Del-R
Saro_2861_Del-R	5'-**CAAGTCGGGAGAGTCAGTC**CGGCTACTCGT ACAACGAGAAGCAG-3'	Region in bold matches sequence in Saro_2861_Del-F
Saro_2861-pK18_Amp-F	5'-**CGATTCATTAATGCAGCTGGCACGACAG**GACGAGATCAGCCTTTAGCCGATCC-3'	Region in bold matches sequence in pK18msB-MCS1
Saro_2404-pK18_Amp-R	5'-**GTTTCTGCGGACTGGCTTTCTAGATGTT**CCCCCACCCGTTTACTTCTTCATGC-3'	Region in bold matches sequence in pK18msB-MCS1
Saro_2404_Del-F	5'-**GCATCGGGAGAGCTTGAG**CAAGAAGTAAGGCTACAGGTGCGAACC-3'	Region in bold matches sequence in Saro_2404_Del-R
Saro_2404_Del-R	5'-**CACCTGTAGCCTTACTTCTTG**CTCAAGCTCT CCCGATGCGATTTC-3'	Region in bold matches sequence in Saro_2404_Del-F
Saro_2404-pK18_Amp-F	5'-**CGATTCATTAATGCAGCTGGCACGACAG**GCAATCTCGGAACTCGGCATCTACC-3'	Region in bold matches sequence in pK18msB-MCS1

Saro_1872-pK18_Amp-R	5'-**GTTTCTGCGGACTGGCTTTCTAGATGTT**CCATGGGGTTTCCTTACTTGTACTTGC-3'	Region in bold matches sequence in pK18msB-MCS1
Saro_1872_Del-F	5'-**CTATCTCTGAGGGCAGTGGC**GAATTCGCGTTCGAGCAGGAAGACATG-3'	Region in bold matches sequence in Saro_1872_Del-R
Saro_1872_Del-R	5'-**CTTCCTGCTCGAACGCGAATTC**GCCACTGCCCTCAGAGATAGAGAG-3'	Region in bold matches sequence in Saro_1872_Del-F
Saro_1872-pK18_Amp-F	5'-**CGATTCATTAATGCAGCTGGCACGACAG**GACAAGGCAGAAATCGTCGATGTGC-3'	Region in bold matches sequence in pK18msB-MCS1
Saro_2812/3-pK18_Amp-R	5'-**GTTTCTGCGGACTGGCTTTCTAGATGTTC**GCAAGACCATAGAGTTCAAACTTGAGAG-3'	Region in bold matches sequence in pK18msB-MCS1
Saro_2812/3_Del-F	5'-**CTACTCCTGTCTGGTCAG**CTACAGGCCCCTC TCCTTCA GC-3'	Region in bold matches sequence in Saro_2812/3_Del-R; underlined bases were modified from the genomic sequence to inactivate the Saro_2813 start codon and maintain the Saro_2814 stop codon
Saro_2812/3_Del-R	5'-**GAAGGAGAGGGGCCTGTAG**CTGACCAGACAGGAGTAGTACCCATG-3'	Region in bold matches sequence in Saro_2812/3_Del-F
Saro_2812/3-pK18_Amp-F	5'-**CGATTCATTAATGCAGCTGGCACGACAG**GT	Region in bold matches sequence in pK18msB-MCS1

	GCATTCAATTCATTCGTCTTTGCG ATGAG	
Saro_1233/4-pK18_Amp-R	5'-**GTTTCTGCGGACTGGCTTTCTA GATGTTC**G TGGTCCTCCTCGACATCAACATGC -3'	Region in bold matches sequence in pK18msB-MCS1
Saro_1233/4_Del-F	5'-**GTGTTGCATATGAAATGTCCGT C**CACGAGG TCAGCCGGAACTACCATATC-3'	Region in bold matches sequence in Saro_1233/4_Del-R
Saro_1233/4_Del-R	5'-**GTTCCGGCTGACCTCGTG**GACG GACATTTC ATATGCAACACATACGAATTTTCC -3'	Region in bold matches sequence in Saro_1233/4_Del-F
Saro_1233/4-pK18_Amp-F	5'-**CGATTCATTAATGCAGCTGGCA CGACAG**CA GCTCTGTGTTGTAGCGTTGCTGTC-3'	Region in bold matches sequence in pK18msB-MCS1
Saro_2861_pVP-F	5'-**CTAACTTTGTTATTTTCGGCTTT CTG**GGATC CTCTGTTGATGAAACCGGTC-3'	Region in bold matches sequence in pVP302K
Saro_2861_pVP-R	5'-**GTATTTTCAGAGCGCGATCGCA GGA**ATGG CGGCAAAGAACCTCGAAGAG-3'	Region in bold matches sequence in pVP302K
Saro_2404_pVP-F	5'-**GTATTTTCAGAGCGCGATCGCA GGA**ATGTG CCAGACCCTAGAGCAGGTC-3'	Region in bold matches sequence in pVP302K
Saro_2404_pVP-R	5'-**CTAACTTTGTTATTTTCGGCTTT CTG**GGTTC GCACCTGTAGCCTTACTTCTTG-3'	Region in bold matches sequence in pVP302K

Saro_2813-2_pVP-HiFi_start	**5'-GTATTTTCAGAGCGCGATCGCAGGA**GTGA CTGACAACAGCTCGACCGATAAG	Region in bold matches sequence in pVP302K
Saro_2813-2_pVP-HiFi_stop	**5'-CTAACTTTGTTATTTTCGGCTTTCTG**GTACT ACTCCTGTCTGGTCAGTCAGTCC	Region in bold matches sequence in pVP302K
Saro_1233/4_pVP-F	5'-**GTATTTTCAGAGCGCGATCGCAGGA**ATGAC ACCTGAAGGAAACCGCGAG-3'	Region in bold matches sequence in pVP302K
Saro_1233/4_pVP-R	5'-**CTAACTTTGTTATTTTCGGCTTTCTG**CAGCA TATGTGGCGGAGCCGTC-3'	Region in bold matches sequence in pVP302K

Table 4.S2. Plasmids used in this study.

Plasmid	Details	Reference
pK18*mobsacB*	pMB1ori *sacB kanR mobT* oriT(RP4) lac*Z*α	Schäfer A, et al. (1994) (38)
pVP302K	lac promoter lacI, *rtxA* (V. cholera), *kanR*; coding sequences for 8×His-tags, TEV protease cleavage site	Gall et al. (2014) (19)
pRARE2	p15a ori *camR*; tRNA genes for 7 rare codons in *E. coli*	Novagen
pK18msB-MCS1	pK18*mobsacB* lacking the multiple cloning site, with a new XbaI site introduced	Kontur et al. (2018) (17)
pK18msB/ΔSaro2861	pK18msB-MCS1 containing genomic regions flanking Saro_2861	This study
pK18msB/ΔSaro2404	pK18msB-MCS1 containing genomic regions flanking Saro_2404	This study
pK18msB/ΔSaro1872	pK18msB-MCS1 containing genomic regions flanking Saro_1872	This study
pK18msB/ΔSaro2812/3	pK18msB-MCS1 containing genomic regions flanking Saro_2812/3	This study
pK18msB/ΔSaro1233/4	pK18msB-MCS1 containing genomic regions flanking Saro_1233/4	This study
pVP302K/Saro2861	pVP302K containing *ligM* downstream of coding sequences for 8xHis-tag and TEV protease cleavage site	This study
pVP302K/Saro2404	pVP302K containing desA downstream of coding sequences for 8xHis-tag and TEV protease cleavage site	This study
pVP302K/Saro2812/3	pVP302K containing *ligAB* downstream of coding sequences for 8xHis-tag and TEV protease cleavage site	This study
pVP302K/Saro1233/4	pVP302K containing *ligA2B2* downstream of coding sequences for 8xHis-tag and TEV protease cleavage site	This study

Table 4.S3. List of *E. coli* strains.

Strain	Relevant characteristics	Reference
DH 5α	F- Φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rK-,mk+) phoA supE44 λ-thi-gyrA96 relA1	Bethseda Research Laboratories
S17-1	recA pro hsdR RP4-2-Tc::Mu-Km::Tn7	Simon et al. (1983) (44)
B834	F- *hsdS metE gal ompT*	Wood (1966) (40); Doherty (1995) (39)

Table 4.S4. Multiple reaction module conditions for HPLC-MS quantification of compounds used in this study.

Compound	MW (g/mol)	Parent (-) m/z	Transition 1	Transition 2	Transition 3
PDC	184.103	183.2	183.2 -> 111.1 CE13	183.2 -> 139.2 CE11	183.2 -> 95.1 CE11
Gallic acid	170.12	169.2	169.2 -> 125.1 CE16	169.2 -> 79.1 CE24	169.2 -> 97.1 CE21
Protocatechuic acid	154.12	153.1	153.1 -> 109.1 CE16	153.1 -> 108.1 CE24	153.1 -> 91.1 CE26
3-MGA	184.15	183.1	183.1 -> 168.2 CE16	183.1 -> 139.2 CE15	183.1 -> 123.1 CE27
Vanillic acid	168.15	167.1	167.1 -> 152.2 CE16	167.1 -> 108.1 CE17	167.1 -> 123.2 CE13
Syringic acid	198.17	197.1	197.1 -> 182.2 CE15	197.1 -> 95.1 CE29	197.1 -> 123.1 CE24

4.8 References

1. Perlack RD, Stokes BJ. 2011. U.S. billion-ton update: Biomass supply for a bioenergy and bioproducts industry. Oak Ridge National Laboratory, Oak Ridge, TN,

2. Davis R, Tao L, Tan ECD, Biddy MJ, Beckham GT, Scarlata C, Jacobson J, Cafferty K, Ross J, Lukas J, Knorr D, Schoen P. 2013. Lignocellulosic biomass to hydrocarbons: Dilute-acid and enzymatic deconstruction of biomass to sugars and biological conversion of sugars to hydrocarbons. NREL Tech Rep doi:10.2172/110.470:88-101.

3. Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon RA, Gilna P, Keller M, Langan P, Naskar AK, Saddler JN, Tschaplinski TJ, Tuskan GA, Wyman CE. 2014. Lignin valorization: Improving lignin processing in the biorefinery. Science 344.

4. Corona A, Biddy MJ, Vardon DR, Birkved M, Hauschild MZ, Beckham GT. 2018. Life cycle assessment of adipic acid production from lignin. Green Chem 20:3857-3866.

5. Vanholme R, Morreel K, Darrah C, Oyarce P, Grabber JH, Ralph J, Boerjan W. 2012. Metabolic engineering of novel lignin in biomass crops. New Phytol 196:978-1000.

6. Linger JG, Vardon DR, Guarnieri MT, Karp EM, Hunsinger GB, Franden MA, Johnson CW, Chupka G, Strathmann TJ, Pienkos PT, Beckham GT. 2014. Lignin valorization through integrated biological funneling and chemical catalysis. Proc Natl Acad Sci U S A 111:12013-12018.

7. Vardon D, Franden MA, Johnson C, Karp E, Guarnieri M, Linger J, Salm M, Strathmann T, Beckham G, Ferguson G. 2015. Adipic acid production from lignin. Energy Environ Sci 8:617-628.

8. Mycroft Z, Gomis M, Mines P, Law P, Bugg TDH. 2015. Biocatalytic conversion of lignin to aromatic dicarboxylic acids in *Rhodococcus jostii* RHA1 by re-routing aromatic degradation pathways. Green Chem 17:4974-4979.

9. Perez JM, Kontur WS, Alherech M, Coplien J, Karlen SD, Stahl SS, Donohue TJ, Noguera DR. 2019. Funneling aromatic products of chemically depolymerized lignin into 2-pyrone-4-6-dicarboxylic acid with *Novosphingobium aromaticivorans*. Green Chem 21:1340-1350.

10. Yaegashi J, Kirby J, Ito M, Sun J, Dutta T, Mirsiaghi M, Sundstrom ER, Rodriguez A, Baidoo E, Tanjore D, Pray T, Sale K, Singh S, Keasling JD, Simmons BA, Singer SW, Magnuson JK, Arkin AP, Skerker JM, Gladden JM. 2017. *Rhodosporidium toruloides*: a new platform organism for conversion of lignocellulose into terpene biofuels and bioproducts. Biotechnol Biofuels 10.

11. Oshlag JZ, Ma Y, Morse K, Burger BT, Lemke RA, Karlen SD, Myers KS, Donohue TJ, Noguera DR. 2020. Anaerobic degradation of syringic acid by an adapted strain of *Rhodopseudomonas palustris*. Appl Environ Microbiol 86.

12. Austin S, Kontur WS, Ulbrich A, Oshlag JZ, Zhang W, Higbee A, Zhang Y, Coon JJ, Hodge DB, Donohue TJ, Noguera DR. 2015. Metabolism of multiple aromatic compounds in corn stover hydrolysate by *Rhodopseudomonas palustris*. Environ Sci Technol 49:8914.

13. Harwood CS, Gibson J. 1997. Shedding light on anaerobic benzene ring degradation: a process unique to prokaryotes? J Bacteriol 179:301-9.

14. Beckham GT, Johnson CW, Karp EM, Salvachua D, Vardon DR. 2016. Opportunities and challenges in biological lignin valorization. Curr Opin Biotechnol 42:40-53.

15. Cecil JH, Garcia DC, Giannone RJ, Michener JK. 2018. Rapid, parallel identification of pathways for catabolism of lignin-derived aromatic compounds in *Novosphingobium aromaticivorans*. Appl Environ Microbiol 84:1-13.

16. Kamimura N, Takahashi K, Mori K, Araki T, Fujita M, Higuchi Y, Masai E. 2017. Bacterial catabolism of lignin-derived aromatics: New findings in a recent decade: Update on bacterial lignin catabolism. Environ Microbiol Rep 9:679-705.

17. Kontur WS, Bingman CA, Olmsted CN, Wassarman DR, Ulbrich A, Gall DL, Smith RW, Yusko LM, Fox BG, Noguera DR, Coon JJ, Donohue TJ. 2018. *Novosphingobium aromaticivorans* uses a Nu-class glutathione S-transferase as a glutathione lyase in breaking the β-aryl ether bond of lignin. J Biol Chem 293:4955-4968.

18. Kontur WS, Olmsted CN, Yusko LM, Niles AV, Walters KA, Beebe ET, Vander Meulen KA, Karlen SD, Gall DL, Noguera DR, Donohue TJ. 2019. A heterodimeric glutathione S-transferase that stereospecifically breaks lignin's β(R)-aryl ether bond reveals the diversity of bacterial β-etherases. J Biol Chem 294:1877-1890.

19. Gall DL, Ralph J, Donohue T, Noguera D. 2014. A group of sequence-related Sphingomonad enzymes catalyzes cleavage of β-aryl ether linkages in lignin β-Guaiacyl and β-Syringyl ether dimers. Environ Sci Technol 48:12454-12463.

20. Gall DL, Kontur WS, Lan W, Kim H, Li Y, Ralph J, Donohue TJ, Noguera DR. 2018. In vitro enzymatic depolymerization of lignin with release of syringyl, guaiacyl, and tricin units. Appl Environ Microbiol 84.

21. Kasai D, Masai E, Miyauchi K, Katayama Y, Fukuda M. 2004. Characterization of the 3-O-Methylgallate dioxygenase gene and evidence of multiple 3-O-methylgallate catabolic pathways in *Sphingomonas paucimobilis* SYK-6. J Bacteriol 186:4951-4959.

22. Kasai D, Masai E, Katayama Y, Fukuda M. 2007. Degradation of 3-O-methylgallate in *Sphingomonas paucimobilis* SYK- 6 by pathways involving protocatechuate 4,5-dioxygenase. FEMS Microbiol Lett 274:323-328.

23. Masai E, Shinohara S, Hara H, Nishikawa S, Katayama Y, Fukuda M. 1999. Genetic and biochemical characterization of a 2-pyrone-4,6-dicarboxylic acid hydrolase involved in the protocatechuate 4,5-cleavage pathway of *Sphingomonas paucimobilis* SYK-6. J Bacteriol 181:55-62.

24. Kasai D, Masai E, Miyauchi K, Katayama Y, Fukuda M. 2015. Characterization of the gallate dioxygenase gene: Three distinct ring cleavage dioxygenases are involved in syringate degradation by *Sphingomonas paucimobilis* SYK-6. J Bacteriol 187:5067-5074.

25. Abe T, Masai E, Miyauchi K, Katayama Y, Fukuda M. 2005. A tetrahydrofolate-dependent O-demethylase, LigM, is crucial for catabolism of vanillate and syringate in *Sphingomonas paucimobilis* SYK-6. J Bacteriol 187:2030-2037.

26. Masai E, Sasaki M, Minakawa Y, Abe T, Sonoki T, Miyauchi K, Katayama Y, Fukuda M. 2004. A novel tetrahydrofolate-dependent *O*-demethylase gene is essential for growth of *Sphingomonas paucimobilis* SYK-6 with syringate. J Bacteriol 186:2757-2765.

27. Jimenez JI, Minambres B, Garcia JL, Diaz E. 2002. Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440. Environ Microbiol 4:824-841.

28. Hibi M, Sonoki T, Mori H. 2005. Functional coupling between vanillate-*O*-demethylase and formaldehyde detoxification pathway. FEMS Microbiol Lett 253:237-242.

29. Priefert H, Rabenhorst J, Steinbuchel A. 1997. Molecular characterization of genes of *Pseudomonas sp*. strain HR199 involved in bioconversion of vanillin to protocatechuate. J Bacteriol 179:2595-2607.

30. Sudtachat N, Ito N, Itakura M, Masuda S, Eda S, Mitsui H, Kawaharada Y, Minamisawa K. 2009. Aerobic vanillate degradation and C1 compound metabolism in *Bradyrhizobium japonicum*. Appl Environ Microbiol 75.

31. Barry KP, Ngu A, Cohn EF, Cote JM, Burroughs AM, Gerbino JP, Taylor EA. 2015. Exploring allosteric activation of LigAB from *Sphingobium sp*. strain SYK-6 through kinetics, mutagenesis and computational studies. Arch Biochem Biophys 567:35-45.

32. Vaillancourt FH, Bolin JT, Eltis LD. 2006. The ins and outs of ring-cleaving dioxygenases. Crit Rev Biochem Mol Biol 41.

33. Barry KP, Taylor EA. 2013. Characterizing the promiscuity of LigAB, a lignin catabolite degrading extradiol dioxygenase from *Sphingomonas paucimobilis* SYK-6. Biochemistry 52:6724-6736.

34. Sze IS, Dagley S. 1987. Degradation of substituted mandelic acids by meta fission reactions. J Bacteriol 169:3833-5.

35. Shikinaka K, Otsuka Y, Nakamura M, Masai E, Katayama Y. 2018. Utilization of lignocellulosic biomass via novel sustainable process. J Oleo Sci 67:1059-1070.

36. Lin L, Cheng Y, Pu Y, Sun S, Li X, Jin M, Pierson EA, Gross DC, Dale BE, Dai SY, Ragauskas AJ, Yuan JS. 2016. Systems biology-guided biodesign of consolidated lignin conversion. Green Chemistry 18:5536-5547.

37. Fredrickson JK, Brockman FJ, Workman DJ, Li SW, Stevens TO. 1991. Isolation and characterization of a subsurface bacterium capable of growth on toluene, naphthalene, and other aromatic compounds. Appl Environ Microbiol 57:796-803.

38. Schäfer A, Tauch A, Jäger W, Kalinowski J, Thierbach G, Pühler A. 1994. Small mobilizable multi- purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene 145:69-73.

39. Doherty AJ, Ashford SR, Brannigan JA, Wigley DB. 1995. A superior host strain for the over-expression of cloned genes using the T7 promoter based vectors. Nucleic Acids Res 23:2074-2075.

40. Wood WB. 1966. Host specificity of DNA produced by Escherichia coli: Bacterial mutations affecting the restriction and modification of DNA. J Mol Biol 16:118-133.

41. Studier FW. 2005. Protein production by auto-induction in high-density shaking cultures. Protein Expression Purif 41:207-234.

42. Ono KN, Mitsuhiro; Hayaishi, Osamu. 1970. Purification and some properties of protocatechuate 4,5-dioxygenase. Biochim Biophys Acta 220:224-238.

43. Maruyama K. 1983. Purification and properties of 2-pyrone-4,6-dicarboxylate hydrolase. J Biochem 93:557-565.

44. Simon R, Priefer U, Pühler A. 1983. A broad host range mobilization system for in vivo genetic engineering: Transposon mutagenesis in gram negative bacteria. Bio/Technology 1:784.

5. Major findings and future directions

5.1 Summary

In this work, I investigated the bacterium *Novosphingobium aromaticivorans* DSM12444 as a potential microbial platform for biological production of chemicals from depolymerized lignin. I first demonstrated that by inactivating key reactions in its central aromatic metabolism, an engineered strain was able to produce PDC from multiple aromatic compounds derived from the three most abundant types of lignin monomeric units (S, G, and H) with high molar yields (Chapter 2). I also demonstrated the feasibility of producing PDC from depolymerized lignin using oxidative (Chapter 2) or reductive (Chapter 3) methods. I investigated the feasibility of using isolated lignin from different agronomically relevant plant biomass sources such as maple, poplar, sorghum, and switchgrass in the context of a biomass-to-PDC biorefinery that combines GVL lignin extraction, hydrogenolysis depolymerization, and biological funneling by the engineered strain of *N. aromaticivorans.* The results obtained allowed us to identify critical factors that affect the biomass to PDC conversion yield in the context of a biorefinery (Chapter 3). Finally, I used this and genomic information to identify two reactions in *N. aromaticivorans* aromatic metabolism – *O*-demethylation and aromatic ring opening – that are performed by enzymes with broad substrate specificity (Chapter 4). In one case, I found a metabolic divergence in the downstream catabolism of 3-MGA and identified enzymes involved in its conversion into gallic acid, a reaction not previously detected in *N. aromaticivorans*. I also found a new aromatic ring opening dioxygenase, LigAB2, and a potentially new aromatic *O*-demethylase, DmtS. Finally, I used information on these newly characterized enzymes to create an improved engineered strain of *N. aromaticivorans* that can produce PDC from the S model aromatic compound syringic acid with stoichiometric yields. Overall, this work shows the potential of understudied bacteria to be used for engineering as effective and efficient biocatalysts for the valorization of plant biomass.

5.2 Future directions

Although in this work I answered a number of relevant questions about the aromatic compounds metabolism in *N. aromaticivorans* and its potential for lignin valorization, a larger number of new important questions were identified. Four important future research areas that I identified are detailed below.

5.2.1 Mechanisms for monomeric and oligomeric aromatic compounds degradation in *N. aromaticivorans*

In this study, I found that *N. aromaticivorans* can metabolize compounds containing S, G, and H aromatic structures and a variety of side-chains. The metabolism of some of those side-chains are well described in *N. aromaticivorans* (1) and other organisms such as the model aromatic degrader *Sphingobium sp*. SYK-6 (2). However, the reactions involved in the side-chain processing of compounds that I found to be metabolized via the central metabolites 3-MGA and PCA, such as S and G diketones, DSA, DCA, MS, MG, Me-DH*p*CA, Me-DHFA, and Me-*p*HBA, remain unknown. Identifying the enzymes involved in those reactions would provide valuable tools to improve other existing bacterial funneling platform organisms. In addition, our results show that the PDC producing strain of *N. aromaticivorans* can degrade some compounds with a propyl or ethyl side-chain, as well as syringol and guaiacol but it does not produce PDC from these biomass products. Future research is encouraged to investigate the mechanisms involved and the products of their degradation.

5.2.2 Compatibility with chemical depolymerization techniques

In this work, I explored the compatibility of one oxidative (formic acid induced depolymerization of oxidized lignin (3)) and one reductive chemical lignin depolymerization technique (hydrogenolysis (4)) with microbial funneling by *N. aromaticivorans*. A large number of other chemical lignin depolymerization techniques have been developed that can generate monomeric products of unknown potential for biological funneling (5, 6). I encourage future research to expand the knowledge of *N.*

aromaticivorans potential to metabolize those compounds and its compatibility with depolymerized lignin produced using those techniques.

5.2.3 Product diversity

In this work, I investigated the production of PDC with *N. aromaticivorans* as a reporter to assess its potential as a platform organism for biological funneling of aromatic compounds. Other lignin-degrading bacteria have been engineered to produce a products such as *cis,cis*-muconate (7), β-keto adipate (8), or muconolactone (8), metabolic intermediates of the catabolism of catechol; or pyridine-2,4-dicarboxylic acid (PDCA) (9), a compound that can be produced from the product of the aromatic ring opening of PCA by a 4,5-extradiol dioxygenase such as LigAB. The *N. aromaticivorans* genome contains a gene (*aroY*) that codes for a putative PCA decarboxylase (10), enzyme that has been described in other bacteria to perform the conversion of PCA into catechol. PCA is the central aromatic metabolite to which the catabolism of numerous G and H type aromatic compounds converge. Interrupting the aromatic ring opening of PCA and altering the expression of *aroY* would allow the rerouting of G and H type aromatic compounds through catechol. Then the addition of appropriate genes into *N. aromaticivorans* genome can potentially lead an engineered strain to produce *cis,cis*-muconate, muconolactone, or β-keto adipate. On the other hand, the interruption of CHMS (product of PCA 4,5 ring opening, see Chapter 4) dehydrogenase activity can potentially lead toextracellular accumulation of this metabolite when an engineered strain is fed compounds known to be metabolized via PCA. When secreted to the growth media and in the presence of NH_4^+, CHMOD can cyclize into PDCA, as shown in previous works (9).

5.2.4 PDC production at an industrially relevant scale

In this work, I demonstrated the feasibility of producing PDC from a variety of single aromatic compounds and from real depolymerized lignin using batch cultures in media containing small amounts of aromatic substrates. To assess the potential of *N. aromaticivorans* as a microbial platform for the production of chemicals in an industrially relevant context, a new set of questions need to be answered.

Parameters such as substrate and product concentration tolerance, maximum product titers, productivity, type and amount of co-substrates needed to fuel the conversion, long-term stability of the engineered strains, etc., are essential to determine *N. aromaticivorans* potential in a biorefinery context. In Chapter 2, I showed that the PDC producing strain of *N. aromaticivorans* could accumulate up to 26.7 mM PDC in a non-optimized fed-batch bioreactor. I encourage future research to optimize a bioreactor by exploring configurations such as fed-batch, continuous, or continuous with cell recycling.

5.3 References

1. Cecil JH, Garcia DC, Giannone RJ, Michener JK. 2018. Rapid, parallel identification of pathways for catabolism of lignin-derived aromatic compounds in Novosphingobium aromaticivorans. Appl Environ Microbiol 84:1-13.

2. Kamimura N, Takahashi K, Mori K, Araki T, Fujita M, Higuchi Y, Masai E. 2017. Bacterial catabolism of lignin-derived aromatics: New findings in a recent decade: Update on bacterial lignin catabolism. Environmental Microbiology Reports 9:679-705.

3. Rahimi A, Ulbrich A, Coon JJ, Stahl SS. 2014. Formic- acid- induced depolymerization of oxidized lignin to aromatics. Nature 515.

4. Luterbacher JS, Azarpira A, Motagamwala AH, Lu F, Ralph J, Dumesic JA. 2015. Lignin monomer production integrated into the γ-valerolactone sugar platform. Energy Environ Sci 8:2657-2663.

5. Sun Z, Fridrich B, de Santi A, Elangovan S, Barta K. 2018. Bright side of lignin depolymerization: Toward new platform chemicals. Chem Rev 118:614-678.

6. Schutyser W, Renders T, Van den Bosch S, Koelewijn SF, Beckham GT, Sels BF. 2018. Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. Chem Soc Rev 47:852-908.

7. Vardon D, Franden MA, Johnson C, Karp E, Guarnieri M, Linger J, Salm M, Strathmann T, Beckham G, Ferguson G. 2015. Adipic acid production from lignin. Energy Environ Sci 8:617-628.

8. Okamura-Abe Y, Abe T, Nishimura K, Kawata Y, Sato-Izawa K, Otsuka Y, Nakamura M, Kajita S, Masai E, Sonoki T, Katayama Y. 2016. Beta-ketoadipic acid and muconolactone production from a lignin-related aromatic compound through the protocatechuate 3,4-metabolic pathway. J Biosci Bioeng 121:652-658.

9. Mycroft Z, Gomis M, Mines P, Law P, Bugg TDH. 2015. Biocatalytic conversion of lignin to aromatic dicarboxylic acids in *Rhodococcus jostii* RHA1 by re-routing aromatic degradation pathways. Green Chem 17:4974-4979.

10. Vardon D, Franden MA, Johnson C, Karp E, Guarnieri M, Linger J, Salm M, Strathmann T, Beckham G, Ferguson G. 2015. Adipic acid production from lignin. Energy Environ Sci 8:617-628.