# Aromatic dimer dehydrogenases from *Novosphingobium aromaticivorans* reduce aromatic diketones

Alexandra M. Linz1,2, Yanjun Ma1,2, Jose M. Perez1,2,3, Kevin S. Myers1,2, Wayne S. Kontur1,2, Daniel R. Noguera1,2,3, Timothy J. Donohue1,2,4,\*

1DOE Great Lakes Bioenergy Research Center, Univ. of Wisconsin-Madison, Madison, WI, USA

2Wisconsin Energy Institute, Univ. of Wisconsin-Madison, Madison, WI, USA

3Department of Civil and Environmental Engineering, Univ. of Wisconsin-Madison, Madison, WI, USA.

4Department of Bacteriology, Univ. of Wisconsin-Madison, Madison, WI, USA

**Running Title: Reduction of aromatic diketones**

\*Direct correspondence to: Timothy J. Donohue

email: [tdonohue@bact.wisc.edu](mailto:tdonohue@bact.wisc.edu)

phone: 6082624663

Keywords: (if needed?)

# Abstract

Lignin is a potential source of valuable aromatics, but its chemical depolymerization results in a heterogeneous mixture of products. Microbes could valorize depolymerized lignin by converting multiple aromatic substrates into one or a small number of products. In this study, we describe the ability of *Novosphingobium aromaticivorans* to metabolize 1-(4-hydroxy-3-methoxyphenyl)-1,2-propanedione (G-diketone), an aromatic Hibbert diketone which is produced during formic acid-catalyzed lignin depolymerization. By assaying genome-wide transcript levels from *N. aromaticivorans* during growth on G-diketone, 2-hydroxy-1-(4-hydroxy-3-methoxyphenyl)propan-1-one (GP-1) and other chemically-related aromatics, we hypothesized that the Lig dehydrogenases, previously characterized as oxidizing β-O-4 linkages in aromatic dimers, were involved in G-diketone metabolism by *N. aromaticivorans*. Using purified *N. aromaticivorans* Lig dehydrogenases, we found that LigL, LigN, and LigD each reduced G-diketone *in vitro* but with different substrate specificities and rates. All three Lig dehydrogenases reduced the Cα ketone of the G-diketone side-chain, while LigL, but not LigN or LigD, reduced the Cα ketone when provided GP-1 as a substrate. The newly identified activity of these Lig dehydrogenases expands the potential range of substrates utilized by *N. aromaticivorans* beyond what has been previously recognized. This is beneficial both for metabolizing a wide range of natural and non-native depolymerized lignin substrates and for engineering enzymes that are active with a broader range of aromatic compounds.

**Importance**

Lignin is a major plant polymer composed of aromatic units that have value as chemicals. However, the structure and composition of lignin has made it difficult to use this polymer as a renewable source of industrial chemicals. Bacteria like *Novosphingobium aromaticivorans* have the potential to make chemicals from lignin because of their ability to convert a variety of lignin-derived aromatics into valuable products. In this work, we report a newly discovered activity of previously characterized dehydrogenase enzymes with a chemically-modified byproduct of lignin depolymerization. We propose that the activity of *N. aromaticivorans* enzymes with both native lignin aromatics and those produced by chemical depolymerization will expand opportunities for producing industrial chemicals from the heterogenous components of this abundant plant polymer.

# Introduction

It is estimated that approximately 30% of organic carbon in the biosphere is comprised of lignin (1). Although lignin is abundant, the heterogeneous structure and chemical composition of this aromatic polymer has prevented its widespread use as a source of products for industrial applications (2). Currently, only a small fraction of available lignin is converted into chemicals, leaving the majority of the material either burned for energy or unused as a renewable source of products (3). Overcoming the inherent barriers to valorizing lignin would provide a renewable source for many chemicals currently produced from non-renewable resources (4). We seek to understand bacterial aromatic metabolism in order to biologically convert deconstructed lignin into valuable chemical products.

Guaiacyl (G)- and syringyl (S)-phenylpropanoids are the most abundant monoaromatics in the lignin polymer, with hydroxylbenzoyl (H)-aromatics as common pendent groups (5). Consequently, all existing chemical methods to depolymerize lignin produce mixtures of G-, S- and H-aromatic monomers, along with dimers or oligomers containing combinations of these aromatic units (6). The heterogeneity of depolymerized lignin presents a challenge to the production of single lignin-derived products, but this could potentially be achieved using microbes to funnel diverse lignin-derived aromatic compounds through a central metabolic pathway (7, 8). We study the ability of *Novosphingobium aromaticivorans* DSM12444 (formerly *Sphingomonas aromaticivorans* F199 (9)), an Alphaproteobacterium capable of metabolizing a diverse array of aromatic compounds, to convert depolymerized lignin aromatics into potentially valuable products (10). *N. aromaticivorans* is genetically tractable and degrades many aromatic compounds completely and quickly, making it an excellent organism for studying the metabolism of lignin-derived products (11, 12).

In this study, we investigate how *N. aromaticivorans* metabolizes an aromatic G-diketone (1-(4-hydroxy-3-methoxyphenyl)-1,2-propane dione) that is a chemical byproduct of both a formic acid-catalyzed lignin depolymerization process (13) and dilute acid hydrolysis of several potential lignocellulosic crops (14). Previous work has shown that *N. aromaticivorans* grew on this G-diketone, and that an engineered strain transformed it into 2-pyrone-4-6-dicarboxylic acid (PDC), a potentially valuable lignin-derived product (10). However, the enzymes that participate in G-diketone metabolism have yet to be identified. To identify enzymes and pathways that contribute to metabolism of the G-diketone, we measured global transcript patterns during growth on the G-diketone to other G-type monoaromatics. Based on these data, we hypothesized that the enzymes LigLNDO*,* which encode pyridine nucleotide-dependent dehydrogenases that initiate degradation of β-O-4 linked aromatic dimers in the closely related *Sphingobium* sp. SYK-6 (15), begin the process of metabolizing the G-diketone by reducing one or both of its ketone groups. To test this hypothesis, we purified LigL, LigN, and LigD and monitored the reduction of the G-diketone *in vitro* using NADH as a cofactor. The results of these experiments reveal how G-diketone and its metabolic byproducts enter the known aromatic metabolism pathways of *N. aromaticivorans*, demonstrate an alternative reductive function for dehydrogenases previously proposed to oxidize aromatic β-O-4 linked dimers, and illustrate the potential for microbial conversion of non-native products of chemical lignin depolymerization as a renewable source of valuable chemicals.

# Methods

## Analyzing growth with and metabolism of G-diketone

﻿ *N. aromaticivorans* cultures were grown aerobically in Standard Mineral Broth (SMB; DSMZ Medium 1185) without tryptone or yeast extract (16), supplemented with either glucose, an aromatic substrate, or both, shaken at 200 rpm and incubated at 30 °C. We measured cell density using a Klett-Summerson photoelectric colorimeter with a red filter; one Klett unit (KU) is equivalent to 8x106 CFU/mL for *N. aromaticivorans* (11).

To assay metabolism of the G-diketone, cultures of *N. aromaticivorans* were grown overnight with 1 g/L D-glucose *(Sigma-Aldrich, St. Louis, MO, USA)* before adding an equal volume of SMB containing 1 g/L glucose and incubating for one hour. We then harvested 2 ml of the culture (3075 rcf, 5 min, room temperature) and resuspended the cells into 0.5 mL fresh SMB with no added carbon source. The resuspended cells were inoculated (1% inoculation ratio by volume) into 50 mL SMB containing either 1g/L glucose or 0.5 g/L glucose plus an amount of G-diketone equivalent to the theoretical chemical oxygen demand (COD) of 0.5 g/L glucose (0.418 g/L G-diketone). 1-mL samples of these cultures were collected as a function of time and harvested (3075 rcf, 5 min, 4 °C). The supernatants of these samples were filtered through 0.22 μm nylon syringe tip filters *(Fisher Scientific, Hampton, NH, USA)*, and analyzed by HPLC-MS, HPLC-UV, and GC-MS.

## RNA-Seq culture conditions

The strain used for RNA analysis was *N. aromaticivorans* DSM12444 containing a deletion of the gene *sacB* (SARO\_RS09410, Saro\_1879), which conveys sucrose sensitivity (11) (Table S1). All cultures were grown in SMB with 0.5 g/L of D-glucose plus an indicated aromatic, as cultures grown on some aromatic compounds as the sole organic substrate produced insufficient biomass for RNA analysis. Amounts of aromatic substrate in the culture were normalized to have a theoretical COD equivalent to 0.5 g/L. The aromatic substrates were dissolved in dimethyl sulfoxide (DMSO) (5 μL per 10 mL of the final culture) prior to addition to the media. Cells grown in the presence of 1 g/L of glucose were used as the control culture for RNA analysis. Aromatic substrates used were protocatechuic acid (PCA), vanillic acid, vanillin, ferulic acid *(Sigma-Aldrich, St. Louis, MO, USA),* 2-hydroxy-1-(4-hydroxy-3-methoxyphenyl)propan-1-one (GP-1) *(Key Organics, Camelford, UK)*, and G-diketone, synthesized as previously described (10).

## RNA preparation

A 25-mL overnight culture grown in the presence of SMB plus glucose provided the inoculum for all cultures used for RNA extraction. Before shifting cells to media containing an aromatic substrate, we added 25 mL of SMB plus glucose and waited 40 min to allow cells to re-enter growth phase. At this time, 2-mL aliquots of this culture were harvested (6000 rpm, 5 min, 21 °C), washed in 1 mL of SMB with no added carbon source, and resuspended into 100 μL of SMB with no added carbon source. Equal volumes of this cell suspension were added to 25-mL triplicate cultures containing either glucose and an aromatic compound or only glucose as the sole carbon source at concentrations described above. These cultures were incubated overnight and then amended with 25 mL of fresh identical media to ensure logarithmic growth. After 40 min, cell growth was terminated by adding 5.7 mL of cold 95% ethanol + 5% acid phenol:chloroform to the entire 50-mL culture. Cells from each culture were harvested (6000 rpm, 12 min, 4 °C), the pellets were flash-frozen in ethanol and dry ice, and the supernatant from each culture was filtered and stored at -80 °C for subsequent HPLC analysis (see above).

Thawed cell pellets were lysed using an SDS/EDTA-based buffer at 65 °C (11). Genomic DNA was removed and RNA purified using the Qiagen RNEasy Kit *(Qiagen, Hilden, Germany)*. An on-column DNAse digestion *(Qiagen, Hilden, Germany)* was performed to remove remaining DNA. RNA was eluted in 50 μL nuclease-free water *(Life Technologies, Carlsbad, CA, USA)* and stored at -80 °C until preparation for sequencing. RNA yield was assessed using a Qubit Broad-Range RNA assay *(Life Technologies, Carlsbad, CA, USA)*.

## RNA sequencing and bioinformatics

RNA-Seq library preparation and sequencing was performed at the Department of Energy Joint Genome Institute (*Berkeley, CA, USA*). Libraries were created using the Illumina TruSeq Stranded Total RNA kit *(Illumina, San Diego, CA, USA)* following the standard protocol, which included ribosomal RNA depletion using the RiboZero bacterial kit *(Illumina, San Diego, CA, USA).* Four RNA-seq libraries were sequenced per lane on an Illumina NextSeq *(Illumina, San Diego, CA, USA)* in 2x151 reads using the manufacturer’s standard protocol.

After sequence analysis, the paired-end FASTQ files were split into those containing forward and reverse reads, and forward read files were retained for further analysis. Sequence reads were trimmed using Trimmomatic version 0.3 (17) with the default settings except for a HEADCROP of 5, LEADING of 3, TRAILING of 3, SLIDINGWINDOW of 3:30, and MINLEN of 36. After trimming, the sequence reads were aligned to the *N. aromaticivorans* genome sequence (GenBank accession NC\_007794.1) using Bowtie2 version 2.2.2 (18) with default settings except the number of mismatches was set to 1. Aligned sequence reads were mapped to gene locations using HTSeq version 0.6.0 (19) with default settings except that the “reverse” strandedness argument was used. The software edgeR version 3.26.8 (20) was used to identify significantly differentially expressed genes from pairwise analyses, using a Benjamini and Hochberg false discovery rate (FDR) less than 0.05 as a significance threshold (21). Raw sequencing reads were normalized using the reads per kilobase per million mapped reads (RPKM).

## Construction of N. aromaticivorans mutants

Strains containing individual in-frame deletions of *ligL, ligN, ligD*, and *ligO* (SARO\_RS09390, SARO\_RS03965, SARO\_RS01025, and SARO\_RS03960, respectively) were generated in a *sacB* mutant of wild-type strain DSM12444 using previously described plasmids and methods (22). Briefly, strains of the *Escherichia coli* WM3064 containing non-replicating pAK405 vectors with approximately 450 base pairs of DNA flanking each desired gene deletion were conjugated on solid medium in a 1:5 ratio with the *N. aromaticivorans sacB* recipient(23)*.* Conjugates in which the plasmid was incorporated into the genome via homologous recombination were selected by plating onto LB with kanamycin, then were grown overnight in 5-mL cultures of LB and plated onto LB with streptomycin to select for strains from which the plasmid was excised via a second round of homologous recombination. Mutants with desired gene deletions were confirmed via the ability to grow in the presence of streptomycin and inability to grow in the presence of kanamycin, as well as via PCR of genomic DNA with gene-specific primers (Table S1). Growth of confirmed deletion mutants and the *sacB* parent strain was tested in 25-mL triplicate flask cultures containing 1 mmol/L glucose + 1 mmol/L G-diketone in SMB + 0.05% DMSO, as well as additional cultures grown in 2 mmol/L glucose as a control.

## In vitro enzyme assays

The genes *ligL* (SARO\_RS09390)*, ligN* (SARO\_RS03965)*,* and *ligD* (SARO\_RS01025) were amplified from *N. aromaticivorans* DSM12444 and separately cloned into plasmid pVP302K (24) containing an N-terminal His8 tag. The expression plasmids were transformed into *E. coli* B834 containing pRARE2 plasmid (*Novagen, Gibbston, NJ, USA*). Protein expression was induced by growing the *E. coli* strains at 25 °C for 25 h in ZYM-5052 autoinduction medium containing kanamycin and chloramphenicol (25). Cells were harvested by centrifugation. The resulting pellet was resuspended in lysis buffer (50 mM NaH2PO4, 100 mM NaCl, 5 mM imidazole, 10% glycerol, 0.5 mM TCEP, and 1% TritonX-100) and lysed by sonication. Following centrifugation of the lysates, the supernatant was filtered through a 0.22-μm, 33-mm diameter, polyethersulfone filter *(MilliporeSigma, Burlington, MA, USA)* and passed through a gravity column packed with Ni2+-NTA resin (*Qiagen, Hilden, Germany*), washed with wash buffer (50 mM NaH2PO4, 200 mM NaCl, 25 mM imidazole, and 0.5 mM TCEP) and eluted in elution buffer (50 mM NaH2PO4, 300 mM NaCl, 500 mM imidazole, and 0.5mM TCEP). The eluted proteins were concentrated using Amicon® Ultra-15 centrifugal filter units (*MilliporeSigma, Burlington, MA, USA*) and dialyzed into dialysis buffer (50 mM NaH2PO4, 100 mM NaCl, and 0.5 mM TCEP), then flash frozen and stored at -80 C until further study. Protein concentrations were determined using the Bradford method.

These purified enzymes were used for *in vitro* enzymeassays using 0.5 mM G-diketone as a potential aromatic substrate and with or without NADH *(Sigma-Aldrich, St. Louis, MO, USA)* as a cofactor, in a buffer containing 25 mM Tris-HCl (pH 8) and 25 mM NaCl *(Sigma-Aldrich, St. Louis, MO, USA)*. An additional control was run with no enzyme added to assess spontaneous degradation of the G-diketone. In some cases, concentrations of the G-diketone were measured at 0, 1, 2, 3, and 24 hours using HPLC-MS. We tested for the presence of additional reaction products by GC-MS using material from the 24-hour timepoint.

We also performed spectrophotometric NADH oxidation/NAD+ reduction assays using LigL, LigN, and LigD incubated in the presence of G-diketone, GP-1, guaiacyl-glycerol-β-guaiacyl ether (GGE), or threo-1-(4-hydroxy-3-methoxyphenyl)propane-1,2-diol (threo-GD) (*BioCrick Co. Ltd., Chengdu, Sichuan, China*). Assay conditions were set as previously described, with NADH/NAD+ added immediately prior to measurement of optical density at 340 nm on an Olis DW-2000 spectrophotometer *(OLIS, Inc., Athens, GA, USA).* An assay with no cofactor was used as the reference for each run. Substrate concentrations were varied to determine the kinetic parameters of each enzyme.

*Chemical analysis*

Quantitative analyses of G-diketone, GP-1, and vanillic acid, were performed on a Shimadzu triple quadrupole liquid chromatography mass spectrometer (LC-MS) (Shimadzu model 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 solvent B (methanol) at a flow rate of 0.4 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 (2.6 μm pore size, 2.1 mm ID, 150 mm length, P/N: 00F-4723-AN). All compounds listed above were detected by multiple-reaction-monitoring (MRM) and quantified using the strongest MRM transition (Table NN). Threo-GD was analyzed by HPLC-UV (Shimadzu model SPD-M20A spectrophotometer) using the same HPLC conditions described above. Threo-GD was quantified by light absorbance at λ = 280 nm.

Chemical analysis to identify GP-1, GP-2, threo-GD, and erythro-GD were 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 N2 at 40 °C, 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. GC-MS spectrum and retention times for GP-1 and threo-GD were compared with authentic standards. Identity of GP-2 was confirmed by comparison with published GC-MS spectrum (14). Identity of erythro-GD was elucidated by GC-MS spectrum comparison with the one produced by threo-GD authentic standard.

# Results

## Extracellular products are transiently accumulated during G-diketone utilization

We previously demonstrated that *N. aromaticivorans* can grow on a mixture of either G-diketone or 1-(4-hydroxy-3,4-dimethoxyphenyl)-1,2-propane dione (S-diketone) and glucose, and that these compounds were each converted to PDC when an engineered strain was grown in their presence (10). This finding predicted that both of these diketones were metabolized through a pathway that included PDC as an intermediate. However, the pathway and enzymes involved in aromatic diketone degradation have yet to be described. We focused this study on dissecting the metabolism of G-diketone, which was the diketone with the highest conversion yield to PDC (10).

We grew *N. aromaticivorans* cells in media containing both the G-diketone and glucose to obtain sufficient biomass for our studies. Under these conditions, HPLC-MS analyses of the extracellular fractions showed a time-dependent disappearance of G-diketone from the culture media, as expected if it is imported and metabolized by *N. aromaticivorans* (Figure 1, Figure S1). HPLC-MS and HPLC-UV analysis of extracellular materials also quantified the time-dependent accumulation and subsequent loss of GP-1 and threo-GD via comparison to authentic standards (Figure 1, Figure S1). One interpretation of these observations is that GP-1 and threo-GD are intermediates in the metabolism of the G-diketone. To identify compounds that appear transiently during growth in the presence of the G-diketone, we analyzed the supernatant from samples taken at 75.5 hr and 98 hr (greatest concentrations of GP-1 and threo-GD, respectively) by GC-MS (Figure S2). The GC trace confirmed the presence of GP-1 and threo-GD (Figure S2). We also observed a peak with an MS spectrum similar to the previously published spectrum of GP-2 (14), and another peak with an identical MS spectrum to threo-GD was observed (Figure S2). We hypothesize that this second peak corresponds to the erythro isomer of threo-GD. The identity of the observed peaks suggested that they could be byproducts or metabolic intermediates of the G-diketone in which either one or both of the ketones were reduced to a hydroxyl group.

## Genome scale changes in transcript abundance during growth in the presence of G-diketone and other G-type aromatics

To identify candidates for gene products involved in metabolism of the G-diketone, we compared transcript abundances between *N. aromaticivorans* cells growing in the presence of glucose alone, or with the G-diketone or another G-type aromatic (PCA, vanillic acid, vanillin, GP-1, or ferulic acid) (Dataset S1). We found that the abundance of several hundred transcripts was altered in cells growing in the presence of the G-diketone and the other G-type aromatics (Base G in Figure 2) when compared to cells grown in the presence of glucose as a sole organic carbon course. Additional differences in transcript abundance were observed between cultures grown in the presence of G-diketone or GP-1 and the other G aromatics (Figure 2).

Since the number of transcripts with different abundance levels in these comparisons was in the hundreds (Figure 2, Dataset S1), we focused on genes encoding enzymes known to participate in aromatic metabolism by *N. aromaticivorans* (Figure 3). Trends in transcript abundance for most of these genes were as expected; for example, *ligAB,* which encode ring-opening dioxygenases, had increased transcript abundance in the presence of all tested aromatics compared to glucose, while *ferAB,* which encode enzymes that act on the sidechain of ferulic acid, were only differentially expressed in the presence of ferulic acid (with the exception of slightly elevated but significant levels of *ferA* in the presence of vanillin) (Figure 3). Many genes encoding enzymes involved in dimer degradation were more abundant in the presence of the G-diketone or GP-1 (Figure 3). These differently expressed transcripts included those from genes encoding NAD-dependent dehydrogenases (*ligD, ligO, ligL, ligN*) (15, 24), glutathione S-transferases (*ligE, ligF, baeAB*) (26), and a glutathione lyase (*NaGSTNu*) (11) that are known or predicted to be involved in breaking the β-O-4 aromatic linkage in lignin and that are increased in transcript abundance in the presence of the model dimeric β-O-4 compound GGE (11). Most of these *lig* genes are not co-localized on the genome (Figure S3), so the increase in abundance of these transcripts likely results from an increase in expression of several transcription units rather than regulation of a single unit.

We also analyzed the supernatants of these same cultures via HPLC for the presence of aromatic compounds. Cultures grown in the presence of glucose and either PCA, vanillic acid, vanillin, or ferulic acid contained significant amounts of the aromatic carbon source provided and lower concentrations of known intermediates in their metabolism (Dataset S2). Both vanillic acid and vanillin were found at low levels in the media of cells grown on GP-1 (Dataset S2). Since neither vanillic acid nor vanillin are not present at detectable levels in our GP-1 stock (Table S2), the presence of these two aromatics in the media of cells grown on GP-1 suggests they are byproducts of cellular metabolism of this aromatic compound.

*In vitro activity of Lig dehydrogenases with G-diketone*

Based on significant changes in transcript abundance and their ability to act on ketone moieties, we identified four genes encoding Lig dehydrogenases that have previously been shown to oxidize β-O-4 linked aromatic dimers as candidates in a pathway for utilization of G-diketone and GP-1. To investigate a potential role of *N. aromaticivorans* Lig dehydrogenases with G-diketone, we successfully purified three of the four recombinant Lig dehydrogenases (LigL, LigN, and LigD) and tested their activity *in vitro.* We were not able to obtain a recombinant LigO protein that was active when incubated with GGE, so no further assays were performed with this enzyme.

When testing recombinant LigL, LigN and LigD for activity with G-diketone *in vitro,* we found that this aromatic compound was reduced in a time-dependent manner in the presence of NADH (Figure 4). The time-dependent reduction of G-diketone required the presence of NADH (Figure 4), suggesting that NADH is a cofactor for this activity with all three dehydrogenases. To better understand the reductive activity of these dehydrogenases with G-diketone, we analyzed the aromatic products of the *in vitro* reactions by GC-MS (Figure 5). When each enzyme was incubated with G-diketone and NADH for 24 hours, GP-2 was observed in all reactions while the G-diketone had disappeared (Figure 5). GP-2 was not present in the reactions at the initial timepoint (Figure 5). Only LigL was able to reduce GP-1, producing both threo- and erythro-GD (Figure 5). No GD or depletion of GP-2 was observed with LigN or LigD (Figure 5). These results demonstrated that LigL, LigN and LigD each have the ability to reduce the Cα ketone of G-diketone when NADH is provided as a source of reducing power, and that LigL is additionally capable of reducing the Cα ketone when the Cβ ketone is already reduced.

Since homologues of these enzymes have been studied for their NAD-dependent dehydrogenase activity on β-O-4 linked aromatic dimers (27), we next measured the kinetic parameters of these recombinant Lig proteins via spectrophotometric NADH oxidation/NAD+ reduction assays with guaiacyl-glycerol-β-guaiacyl ether (GGE), a model β-O-4 linked aromatic dimer as substrate, compared to G-diketone and threo-GD (Figure 6). We attempted to measure kinetic parameters with GP-1 and NADH as the substrates, but absorbance of a putative reaction product at 340 nm interfered with our ability to monitor NADH oxidation. Recombinant LigL, LigN, and LigD showed relatively similar turnover frequencies (kcat) and Michaelis-Menten half-saturation constants (Km) when incubated with GGE and NAD+ (Figure 6), as expected given previous reports of their homologues from *Sphingobium* SYK-6 with this substrate (27). However, the kcat value for LigL with the G-diketone (1.44) was over 10-fold than the kcat value for LigL with GGE (0.12), while conversely LigN and LigD both had significantly lower kcat values for the G-diketone than GGE (0.1 and 0.02, respectively) (Figure 6). This finding suggests that LigL has a faster turnover with the G-diketone than with GGE compared to LigN and LigD. For LigD, we observed the opposite trend; the kcat value for LigD with the G-diketone was 10-fold lower (0.02) than the kcat value with GGE (0.2), suggesting that LigD acts more rapidly on GGE than on the G-diketone. However, when LigL was incubated with threo-GD, reduction of NAD+ was observed with a relatively low kcat value (0.03), although we were unable to identify a product (Figure 6). No reduction of NAD+ was observed when LigN or LigD were incubated with threo-GD (Figure 6).

*Genetic analysis of the role of Lig aromatic dehydrogenases in G-diketone utilization*

The results of the above *in vitro* enzyme assays indicate that LigL, LigN and LigO can each reduce the G-diketone. If each of the Lig dehydrogenases can contribute to G-diketone metabolism *in vivo*, it predicts that loss of any single Lig dehydrogenase would not have a significant impact on the growth or the production of microbial biomass of *N. aromaticivorans* in the presence ofthis aromatic compound. To test this hypothesis, we created strains with individual in-frame deletions of *ligL*, *ligN*, *ligD*,and *ligO* and tested their growth in media containing either glucose alone or glucose and G-diketone. Of the four individual mutants, none showed any significant defects in either growth rate or total biomass produced when grown on glucose as the sole carbon source or on glucose plus the G-diketone (Figure S4). These data show that none of these individual dehydrogenases is required for growth in media containing glucose and G-diketone and it predicts that each enzyme supports sufficient reduction of this aromatic to support wild type growth rates*.*

# Discussion

Microbial funneling of the products of depolymerized lignin into valuable compounds has the potential to overcome an existing challenge caused by the heterogeneous chemical composition of this polymer. Unlike many abiotic methods, microbes can simultaneously convert mixtures of aromatic monomers to one or a few desired products, either naturally or when engineered (10, 12). Several bacteria are potential chassis organisms for microbial funneling of depolymerized lignin. Sphingomonad bacteria such as *N. aromaticivorans* and *Sphingobium* SYK-6 have the native ability to cleave β-O-4 ether linkages in lignin-derived aromatic dimers (24) and can be engineered to metabolize a wide range of substrates to produce a valuable compound such as PDC(10, 12). Similarly, the Gammaproteobacterium *Pseudomonas putida* can either naturally or be engineered to produce valuable chemicals from one or more aromatic substrates (28–30).Other potential bacterial products of depolymerized lignin include *cis-cis* muconic acid, lipids, or polyhydroxyalkonates (31–35).

The success of microbial funneling is likely dependent on both the organism and the lignin depolymerization method. Some microbes may be better suited to funnel both native and chemically-modified products of lignin depolymerization, either before or after the addition of metabolic functions from another host. In this study, we examined how *N. aromaticivorans* consumes G-diketone, a phenylpropanone that is abundant in the products formed from a formic acid-induced lignin depolymerization method (13) and is detected in the products of dilute acid hydrolysis of several potential lignocellulosic biofuel feedstocks (14). This G-diketone belongs to a group of compounds known as Hibbert ketones (14), for which information on microbial degradation pathways is lacking. Below, we discuss the implications of our findings on the enzymes involved in the process of G-diketone degradation by *N. aromaticivorans*.

Previously characterized Lig dehydrogenases reduce G-diketone*.*

Based on genome-wide transcript analysis of cells grown in the presence of G-diketone, we hypothesized that the known aromatic dehydrogenases LigLNDOplayed a role in G-diketone metabolism, possibly by reducing one or both ketones on the aromatic sidechain. To test this hypothesis, we compared enzyme activity for recombinant LigL, LigN, and LigD incubated with the G-diketone or GGE, a model aromatic dimer. *In vitro* enzyme assays confirmed that LigL, LigN, and LigD were all able to reduce the Cα ketone of G-diketone in a reaction that required NADH as a source of reductant to produce GP-2. Each of these Lig dehydrogenases bound G-diketone and GGE with a similar affinity, since the Michaelis-Menten constants for the monoaromatic and dimeric substrates were comparable under identical reaction conditions (Figure 6). This predicts that the presence or absence of a second aromatic ring does not make a major contribution to binding of this substrate. Previous analysis of a Lig β-etherase that participates in GGE metabolism has shown that this enzyme is active with both β-O-4 linked aromatic dimers and monomers that contain different side chains (24). This, along with our observations, indicates that Lig enzymes analyzed in the past may have a broader role in the metabolism of biologically derived aromatics as well as those produced during chemical treatment of lignin and other aromatic-containing substrates. Indeed, our experiments show that *N. aromaticivorans* LigL, LigN and LigD are each able to oxidize the side chain of an aromatic dimer like GGE and reduce the Cα-side chain ketone in G-diketone to generate GP-2, and that LigL reduces both of the Cα- and Cβ-ketones to produce GD (Figure 5).

In this regard, we found that LigN and LigD had higher turnover frequencies when oxidizing the Cα bond in the aromatic dimer GGE compared to reducing the Cα ketone in the G-diketone (Figure 6). In contrast, LigL had a higher turnover frequency when reducing the Cα ketone in the G-diketone than when oxidizing the Cα position of GGE (Figure 4). These results predict there may be active site differences in these three aromatic dehydrogenases that impact their ability to oxidize or reduce individual substrates. Additionally, we found that LigL was able to reduce the Cα ketone to form GD when provided GP-1 as a substrate (Figure 6), while LigN and LigD could not reduce GP-1 (Figure 6). This suggests that LigL catalyzes another step in G-diketone degradation, resulting in a fully reduced side chain that is a substrate for subsequent cleavage to produce vanillin. Hibbert ketones such as GP-2 and GP-1 are known to spontaneously interconvert, so it is possible that, under cellular conditions, GP-2 can isomerize to GP-1 (14). Deletion of *ligL* did not result in a growth defect on the G-diketone, so it is possible that another as of yet unknown enzyme is capable of reducing GP-2 to GD. Lig dehydrogenases from other sphingomonads have been found to be stereospecific (27) or stereoselective (36) for the stereoconfiguration of Cα in GGE, so it is likely that stereochemistry may play an important role in enzymatic processing of the G-diketone sidechain. However, sources of erythro-GD are needed to assess the relative activity of LigD, LigL, and LigN with this compound.

In vivo degradation of the G-diketone

Our in vitro enzyme assays predicted that G-diketone degradation is initiated by the NADH-dependent reduction of the Cα ketone, producing GP-2. When *N. aromaticivorans* was grown in the presence of the G-diketone + glucose, we observed the transient extracellular accumulation of GP-1 and a compound predicted to be GP-2 (Figure 1), but were unable to separate the two predicted compounds in this material to determine which ketone was reduced.

Based on our findings, we can generate a model for metabolism of G-diketone by *N. aromaticivorans* (Figure 7)*.* This model predicts that the Cα ketone of the G-diketone is reduced by one of several redundant Lig dehydrogenases to GP-2, which can then isomerize to GP-1. GP-1 is further reduced by LigL to GD. We further propose that the sidechain of GD is cleaved by an as of yet uncharacterized enzyme(s), producing an unknown two-carbon product and vanillin, which is oxidized by LigV, which converts vanillin into vanillate in Sphingobium SYK-6 (37) (Figure 7). Consistent with this model, transcripts derived from *N. aromaticivorans* ligV were more abundant during growth on both G-diketone and GP-1 than during growth on glucose alone (Figure 3) (37), suggesting that vanillin is an intermediate in the G-diketone pathway. Unfortunately, our G-diketone preparations contain trace amounts of vanillin, making it difficult to assess the source of low extracellular levels of this compound in cells grown in the presence of G-diketone (Table S2). However, the transient extracellular accumulation of vanillin and vanillic acid when cells are grown in the presence of GP-1 (Dataset S2) supports the predicted pathway for *N. aromaticivorans* G-diketone metabolism*.*

*Substrate specificity of Lig pathway enzymes*

When taken together, this work expands our knowledge of the ability of *N. aromaticivorans* to metabolize and funnel a diversity of aromatic compounds into a common pathway. While this work focused on metabolism of G-diketone, previous studies have shown that cells can grow on S-diketone (10) and that several Lig enzymes are active with S-type aromatic substrates as well (24, 38). We propose that the use of multi-functional enzymes may provide *N. aromaticivorans* with the flexibility to consume a broad mix of monomeric and oligomeric aromatic substrates.

There are other examples of *N. aromaticivorans* enzymes that can act on both aromatic monomers and dimers. NaLigF2 is a β-etherase that cleaves aromatic dimers in the presence of glutathione (24). However, based on fitness analysis of a mutant library of *N. aromaticivorans*, this same gene product was proposed to be a subunit of a DesCD isomerase that converts 4-carboxy-2-hydroxy-6-methoxy-6-oxohexa-2,4-dienoate (CHMOD) to (4 E)-oxalomesaconate (OMA) during syringic acid degradation (22). As predicted, mutations that inactivate *desD* resulted in a loss of growth on syringic acid and accumulation of CHMOD, and *in vitro* enzyme assays showed that purified DesCD produced OMA (22). Additionally, the syringate o-demethylase DesA can also demethylate vanillic acid, albeit at a slower rate, and a second ring-opening dioxygenase, LigAB2, has been identified in *N. aromaticivorans* (12). Thus, it is possible that broad substrate specificity may be an underappreciated feature of *N. aromaticivorans* aromatic metabolism that is more common than previously thought.

The presence of overlapping, multi-functional, aromatic-degrading enzymes may allow *N. aromaticivorans* to adapt to, metabolize, and persist in the presence of a wide set of aromatic substrates that it might find in nature. This enzymatic flexibility is also potentially advantageous from an industrial production standpoint, since *N. aromaticivorans* could be well suited for metabolism of both naturally occurring aromatics and those that are chemical byproducts of lignin or other aromatic substrates. The relatively broad substrate specificity of *N. aromaticivorans* enzymes may also provide scaffolds for improvement in catalytic rate or utilization of aromatic substrates.

In conclusion, we coupled metabolite analyses, transcriptomics and enzyme assays to predict and identify proteins responsible for the degradation of the G-diketone in *N. aromaticiovorans.* These studies showed that three Lig dehydrogenases previously described for a role in oxidation of a Cα carbon in β -O-4 linked aromatic dimers reduce the Cα ketone as a first step in the metabolism of G-diketone. Based on this and other published studies, we propose that redundant multi-functional enzymes are a key feature of *N. aromaticivorans’* aromatic metabolism, allowing it to degrade diverse aromatic compounds, including ones found in nature and those generated as a byproduct of chemical deconstruction of lignin. This finding expands our understanding of microbial aromatic metabolism, and it furthers our ability to identify and engineer a microbial strain capable of valorizing lignin or chemical products of its deconstruction as part of biorefinery pipeline.

# Data availability

Raw sequencing reads are available through the DOE Joint Genome Institute Genome Portal with Project ID 1233250 (<https://genome.jgi.doe.gov/portal/Novarocriptomics_FD/Novarocriptomics_FD.info.html>

). Code and processed data files used in this project are available via the GitHub repo <https://github.com/GLBRC/AromaticDiketones>

GEO accession:

**Acknowledgements**

This material is based upon work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We thank Dr. Joshua Michener for providing plasmid vectors used to generate the *ligLNDO* deletion strains. We also thank Dr. Avery Vilbert for providing advice and expertise regarding the *in vitro* enzymatic assays.

# References

1. Boerjan W, Ralph J, Baucher M. 2003. Lignin Biosynthesis. Annu Rev Plant Biol 54:519–546.

2. Cao Y, Chen SS, Zhang S, Ok YS, Matsagar BM, Wu KC-W, Tsang DCW. 2019. Advances in lignin valorization towards bio-based chemicals and fuels: Lignin biorefinery. Bioresour Technol 291:121878.

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 (80- ). American Association for the Advancement of Science.

4. Becker J, Wittmann C. 2019. A field of dreams: Lignin valorization into chemicals, materials, fuels, and health-care products. Biotechnol Adv.

5. Ralph J, Lapierre C, Boerjan W. 2019. Lignin structure and its engineering. Curr Opin Biotechnol 56:240–249.

6. 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 Sustain Chem Eng 6:3367–3374.

7. Kamimura N, Sakamoto S, Mitsuda N, Masai E, Kajita S. 2019. Advances in microbial lignin degradation and its applications. Curr Opin Biotechnol 56:179–186.

8. Davis K, Moon TS. 2020. Tailoring microbes to upgrade lignin. Curr Opin Chem Biol 59:23–29.

9. Romine MF, Stillwell LC, Wong KK, Thurston SJ, Sisk EC, Sensen C, Gaasterland T, Fredrickson JK, Saffer JD. 1999. Complete sequence of a 184-kilobase catabolic plasmid from Sphingomonas aromaticivorans F199. J Bacteriol 181:1585–1602.

10. 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.

11. 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 -aryl ether bond of lignin. J Biol Chem 293:4955–4968.

12. Perez JM, Kontur WS, Gehl C, Gille DM, Ma Y, Niles A V., Umana G, Donohue TJ, Noguera DR. 2021. Redundancy in aromatic O-demethylation and ring opening reactions in Novosphingobium aromaticivorans and their impact in the metabolism of plant derived phenolics . Appl Environ Microbiol 1–23.

13. Rahimi A, Ulbrich A, Coon JJ, Stahl SS. 2014. Formic-acid-induced depolymerization of oxidized lignin to aromatics. Nature 515:249–252.

14. Mitchell VD, Taylor CM, Bauer S. 2014. Comprehensive Analysis of Monomeric Phenolics in Dilute Acid Plant Hydrolysates. Bioenergy Res 7:654–669.

15. Pereira JH, Heins RA, Gall DL, McAndrew RP, Deng K, Holland KC, Donohue TJ, Noguera DR, Simmons BA, Sale KL, Ralph J, Adams PD. 2016. Structural and biochemical characterization of the early and late enzymes in the lignin β-aryl ether cleavage pathway from sphingobium sp. SYK-6. J Biol Chem 291:10228–10238.

16. Stanier RY, Palleroni NJ, Doudoroff M. 1966. The aerobic pseudomonads: a taxonomic study. J Gen Microbiol 43:159–271.

17. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

18. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359.

19. Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31:166–169.

20. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140.

21. Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B 57:289–300.

22. Cecil JH, Garcia DC, Giannone RJ, Michener JK. 2018. Rapid, Parallel Identification of Catabolism Pathways of Lignin-Derived Aromatic Compounds in Novosphingobium aromaticivorans. Appl Environ Microbiol 84:e01185-18.

23. Kaczmarczyk A, Vorholt JA, Francez-Charlot A. 2012. Markerless gene deletion system for sphingomonads. Appl Environ Microbiol 78:3774–3777.

24. Gall DL, Ralph J, Donohue TJ, Noguera DR. 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.

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

26. Kontur WS, Olmsted CN, Yusko LM, Niles A V., 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 (R)-aryl ether bond reveals the diversity of bacterial -etherases. J Biol Chem 294:1877–1890.

27. Sato Y, Moriuchi H, Hishiyama S, Otsuka Y, Oshima K, Kasai D, Nakamura M, Ohara S, Katayama Y, Fukuda M, Masai E. 2009. Identification of three alcohol dehydrogenase genes involved in the stereospecific catabolism of arylglycerol-β-aryl ether by Sphingobium sp. strain SYK-6. Appl Environ Microbiol 75:5195–5201.

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

29. Elmore JR, Dexter GN, Salvachúa D, Martinez-Baird J, Hatmaker EA, Huenemann JD, Klingeman DM, Peabody GL, Peterson DJ, Singer C, Beckham GT, Guss AM. 2021. Production of itaconic acid from alkali pretreated lignin by dynamic two stage bioconversion. Nat Commun 12.

30. 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.

31. Johnson CW, Salvachúa D, Khanna P, Smith H, Peterson DJ, Beckham GT. 2016. Enhancing muconic acid production from glucose and lignin-derived aromatic compounds via increased protocatechuate decarboxylase activity. Metab Eng Commun 3:111–119.

32. 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.

33. Shields-Menard SA, AmirSadeghi M, Green M, Womack E, Sparks DL, Blake J, Edelmann M, Ding X, Sukhbaatar B, Hernandez R, Donaldson JR, French T. 2017. The effects of model aromatic lignin compounds on growth and lipid accumulation of Rhodococcus rhodochrous. Int Biodeterior Biodegrad.

34. He Y, Li X, Ben H, Xue X, Yang B. 2017. Lipid Production from Dilute Alkali Corn Stover Lignin by Rhodococcus Strains. ACS Sustain Chem Eng 5:2302–2311.

35. Liu ZH, Olson ML, Shinde S, Wang X, Hao N, Yoo CG, Bhagia S, Dunlap JR, Pu Y, Kao KC, Ragauskas AJ, Jin M, Yuan JS. 2017. Synergistic maximization of the carbohydrate output and lignin processability by combinatorial pretreatmentf. Innov Green Process Eng Sustain Energy Environ 2017 - Top Conf 2017 AIChE Annu Meet 2017-Octob:248–264.

36. Palamuru S, Dellas N, Pearce SL, Warden AC, Oakeshott JG, Pandey G. 2015. Phylogenetic and kinetic characterization of a suite of dehydrogenases from a newly isolated bacterium, strain SG61-1L, that catalyze the turnover of guaiacylglycerol-β-guaiacyl ether stereoisomers. Appl Environ Microbiol 81:8164–8176.

37. Masai E, Yamamoto Y, Inoue T, Takamura K, Hara H, Kasai D, Katayama Y, Fukuda M. 2007. Characterization of ligV essential for catabolism of vanillin by Sphingomonas paucimobilis SYK-6. Biosci Biotechnol Biochem 71:2487–2492.

38. 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.