β-Etherase and benzoyl-CoA pathway enzymes mediate biodegradation of lignin-related aromatic compounds

By

Daniel Leo Gall

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Civil & Environmental Engineering)

at the

UNIVERSITY OF WISCONSIN-MADISON

2015

Date of final oral examination: 5/4/2015

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

Daniel R. Noguera, Professor, Civil & Environmental Engineering
Timothy J. Donohue, Professor, Bacteriology
John Ralph, Professor, Biochemistry
Brian G. Fox, Professor, Biochemistry
Katherine D. McMahon, Professor, Civil & Environmental Engineering
Marc A. Anderson, Professor, Civil & Environmental Engineering

© Copyright by Daniel Leo Gall 2015

All Rights Reserved

ABSTRACT

Lignocellulose is composed of earth's three most abundant biopolymers: cellulose, hemicelluloses, and lignin. Lignin's recalcitrance presents major obstacles to deriving commercially desirable entities from lignocellulosic biomass and is largely attributable to the various types of racemic units with several types of inter-unit covalent bonds through which the polymer's aromatic monomers (coniferyl alcohol that produces guaiacyl units and sinapyl alcohol that produces syringyl units) are linked by combinatorial radical coupling reactions. Cleavage of the β-ether bonds, the most prominent type of inter-unit linkage, is thus crucial to lignin degradation processes. The use of purified recombinant "Lig" enzymes from *Sphingobium* sp. strain SYK-6, *Novosphingobium* sp. strain PP1Y, and *Novosphingobium aromaticivorans* strain DSM12444 in *in vitro* biochemical assays with glutathione and lignin model β-ether-linked aromatic dimers as co-substrates revealed that each organism possesses both the β(*R*)- and β(*S*)-specific β-etherase enzymes that are required to cleave the racemic β-ether linkages found in lignin. I also found that each Lig enzyme exhibits β-etherase activity with all four types of β-ether-linked dimer substructures found in nature: guaiacyl-β-guaiacyl, syringyl-β-guaiacyl, guaiacyl-β-syringyl, and syringyl-β-syringyl. Further investigation of the Lig enzymes from strain SYK-6 indicates that β-etherase catalysis involves cleavage of the β-ether bond resulting in a glutathione-conjugated β-S-thioether linkage and that this reaction causes inversion of the β-chirality (*i.e.*, conversion of β(*R*)-substrates to β(*S*)-products and β(*S*)-substrates to β(*R*)-products). In sum, these findings demonstrate that β-etherases are strictly stereoselective with respect to the β-configuration of their products, either β(*R*)- or β(*S*)-stereospecific amongst racemic substrates, and yet, non-specific with regard to the substrate's guaiacyl/syringyl composition. In a parallel study, I also found that metabolism of *meta*-hydroxy-aromatic acids (protocatechuate and *m*-hydroxy-benzoate) in the bacterium *Rhodopseudomonas palustris* can be induced via the benzoyl-CoA pathway by

metabolism of other pathway substrates (*e.g.*, benzoate). Although *meta*-hydroxy-aromatic acids were poor sole carbon sources, they were readily degraded in the presence of benzoate. *R. palustris hbaBCD* and *badDEFG* mutant phenotypes demonstrated that dehydroxylase HbaBCD and dearomatizing enzyme BadDEFG are required for protocatechuate *m*-hydroxy-benzoate metabolism, respectively. These results indicate that *meta*-substituted monoaromatic compounds, similar to those likely to be derived from lignin via β-ether cleavage, can be reductively transformed *in vivo* via benzoyl-CoA pathway enzymes, yielding potentially valuable commodities as pathway intermediates.

ACKNOWLEDGMENTS

I would like to express my sincerest gratitude to the many people who helped me through my academic journey. My advisor, Dr. Daniel Noguera, gave me every freedom imaginable to pursue the topics I found interesting as my project developed. This is also true of my co-advisors Dr. Timothy Donohue and Dr. John Ralph, who gladly took on the unenviable tasks of teaching microbiology and chemistry to an environmental engineering student. Dan, Tim, and John were, excellent academic advisors, but more importantly, caring life mentors, as they helped me through some difficult times that arose during graduate school. Thank you guys! And my efforts in the laboratory would have been completely fruitless without the organic synthesis training I received from Hoon Kim, Fachuang Lu, Matt Regner, Dharshana Padmakshan, Ruili Gao, Yuki Tobimatsu (the entire Ralph lab, honestly!), and Sally Ralph, or without the biochemical guidance from Dr. Brian Fox. And thanks to Dan N., Tim D., John R., Brian F., Dr. Katherine McMahon, and Dr. Marc Anderson for serving on my Ph.D. committee... I'm only on preliminary page *iii*, but I fear this is going to be a long one!

Second in sequence, but certainly not in importance, I thank my family. My sister Mandy has become what my grandfather "Boompa" always said she would: one of my best friends. Sometimes I miss the days when all 4 of my grandparents, Leo, Agnes, Chester, and Peggers could be with us up North, but now, our generation's trips to Cool Sha Nagh with Mandy and Nate's family (Carter, Lily, and Griffin) are pretttaaayyy good too! And of course, thanks to Dennis and Roberta Gall – my parents! There were times when I felt like I had missed out on time with family because I took school too seriously. But my parents never let me feel that way. They were always so supportive and understanding... even now, as I conclude the 24th grade, yikes! Thanks Mom – I wish more than anything that you were still with us. I miss how proud of me you always were! That pride helped me through every stage of life. And thanks, Dad – you are the best father, and my best friend in the world.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 5

LIST OF TABLES

Chapter 4

Chapter 5

CHAPTER 1: Introduction

Lignocellulosic biopolymers as feedstocks for production of valued commodities

Lignocellulosic biomass is composed of Earth's three most abundant terrestrial biopolymers (cellulose, hemicelluloses, and lignin) and accounts for the majority of organic carbon in the global budget (Argyropoulos and Menachem 1998). Both abundant and renewable, polysaccharides (cellulose and hemicelluloses, Figure 1-1) are attractive feedstocks for the production of paper, ethanol, and next-generation fuels (Simmons *et al.* 2010). Converting lignocellulose to valued commodities is, however, made challenging by the physiochemical constraints placed on polysaccharides by a third biopolymer, lignin (Figure 1-1). Comprising between 15 and 30% (dry weight) of vascular plant cell walls (Higuchi 1980; Lewis and Yamamoto 1990), lignin is a complex and combinatorial polymer consisting of aromatic monomers linked together via heterogeneous chemical bonds, and to which plant polysaccharides are covalently bound, particularly in grasses (Iiyama *et al.* 1994; Ralph *et al.* 2004a; Chen and Dixon 2007; Ralph 2010). Further, processes designed to simply remove lignin and convert the resulting cellulosic and hemicellulosic fibers to valued commodities (Lynd *et al.* 1991; Dale *et al.* 1993; Lynd *et al.* 2002; Teymouri *et al.* 2004; Yang and Wyman 2004; Lloyd and Wyman 2005; Murnen *et al.* 2007; Yang and Sen 2010) often result in the release of so-called "lignotoxins" that are often inhibitory to microbial conversion processes (Palmqvist and Hahn-Hägerdal 2000b; Palmqvist and Hahn-Hägerdal 2000a). The use of lignin as a feedstock by biomanufacturers is limited by the polymer's complexity and recalcitrance, and thus, lignin is often a waste byproduct of biorefineries or is incinerated to generate power for other processes (Tuck *et al.* 2012; Ragauskas *et al.* 2014). Because of the need for bioconversion processes that may (a) derive useful products from lignin, and (b) detoxify the aromatic contaminants that are often found in the polysaccharide fractions after crude processing of lignocellulosic biomass, my thesis investigates microbial metabolic pathways that enable derivation

of monomeric aromatic units from oligomeric lignin, and subsequent conversion of the lignin-derived monomers to potential fuel precursors.

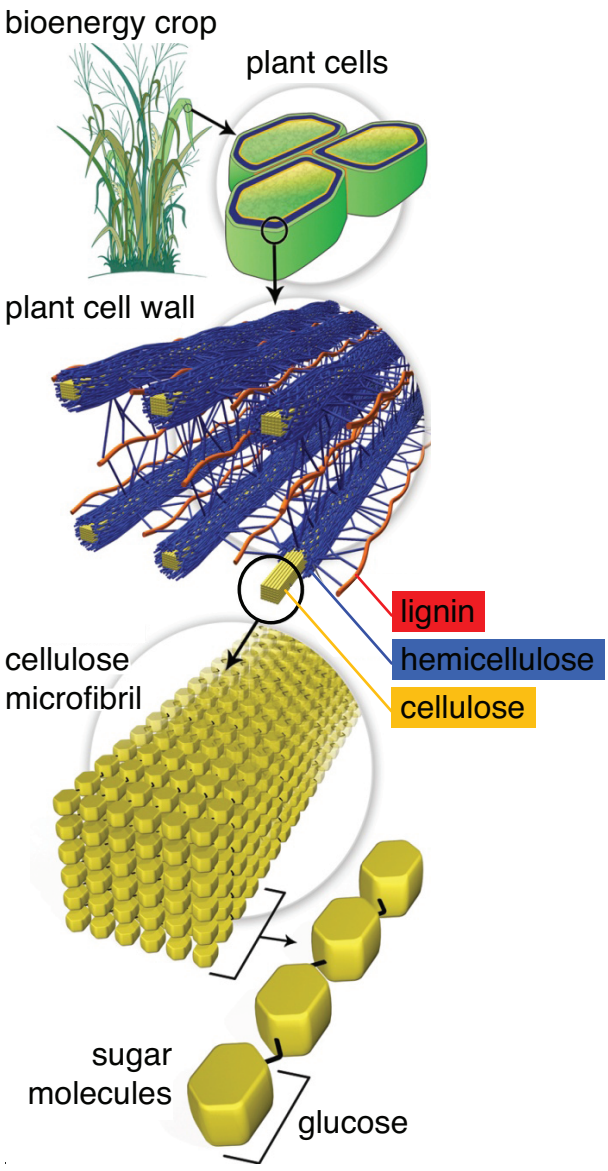

Figure 1-1. Spatial relationship between cellulose (yellow), hemicellulose (blue), and lignin (red) polymers in lignocellulose. Graphic adapted from the United States Department of Energy website: genomicscience.energy.gov/biofuels. [Note that lignin is clearly underrepresented and is overly connected to polysaccharides in this model].

In this introductory chapter, I explain the structure of lignin on a molecular level, discuss the type of inter-unit linkages that are targeted by the *Sphingobium* sp. strain SYK-6 β-etherase degradation pathway, and illustrate how this pathway's reactions lead to the derivation of monomeric units from oligomeric lignin. I then explain how the purple non-sulfur photo-heterotrophic bacterium, *Rhodopseudomonas palustris* was used as a biocatalyst for the reductive conversion of monomeric aromatic compounds to potential fuel precursors.

Lignin polymers – monomeric unit composition and inter-unit linkages

Lignin is an abundant and recalcitrant plant biopolymer that plays a role in water and nutrient transport (Boyce *et al.* 2004) and confers resistance to mechanical stress, photodegradation, and various modes of infectious attacks (Hejnowicz 2005; Bhuiyan *et al.* 2009). These heterogeneous and combinatorial polymers are produced via oxidative coupling of the 'monolignols,' primarily coniferyl and sinapyl alcohols (Figure 1-2A) (Boerjan *et al.* 2003; Ralph *et al.* 2004b). Radical condensation of monolignols and *p*-coumaroylated monolignols (Grabber *et al.* 1996) gives rise to lignin polymers that still have guaiacyl, and syringyl monoaromatic units, but that have *p*-coumarate pendant moieties decorating the lignin polymer (Figure 1-2A) (Higuchi 1980; Hahlbrock and Scheel 1989; Lewis and Yamamoto 1990). These units are not to be confused with the typically minor (and therefore rarely mentioned again here) *p*-hydroxyphenyl units in lignin that derive from the other monolignol, *p*-coumaryl alcohol. *p*-Hydroxyphenyl and *p*-coumarate units are *p*-hydroxylated (*para* to the aliphatic sidechain), guaiacyl units bear a *p*-hydroxyl and one *m*-methoxy group, and syringyl units are *p*-hydroxylated and *m,m*-dimethoxylated (to aid the reader with figures throughout this document, *p*-coumarate units, and the guaiacyl, and syringyl monoaromatic units are shown in green, blue, and red, respectively).

The inherent structural rigidity of lignin is due to the covalent bonding nature of the guaiacyl and syringyl unit cross-linkages that emerge during polymerization. Lignin biosynthesis is well

established to be caused by enzymatic radicalization of monolignols and subsequent non-enzymatic coupling of those radical cations in the plant apoplast (Freudenberg 1959; Boerjan *et al.* 2003), and thus, unlike other biopolymers, lignin's guaiacyl and syringyl monomers are not linked together by a single type of chemical bond. Rather, as monolignol radicals couple with the growing polymer (radical), the bond structure manifested between two monomers will be one of several inter-unit linkage types typically found in lignin (Figure 1-2B): resinols (β–β), 4–O–5 di-aryl ethers, β–O–4-aryl ethers, and phenylcoumarans (β–5).

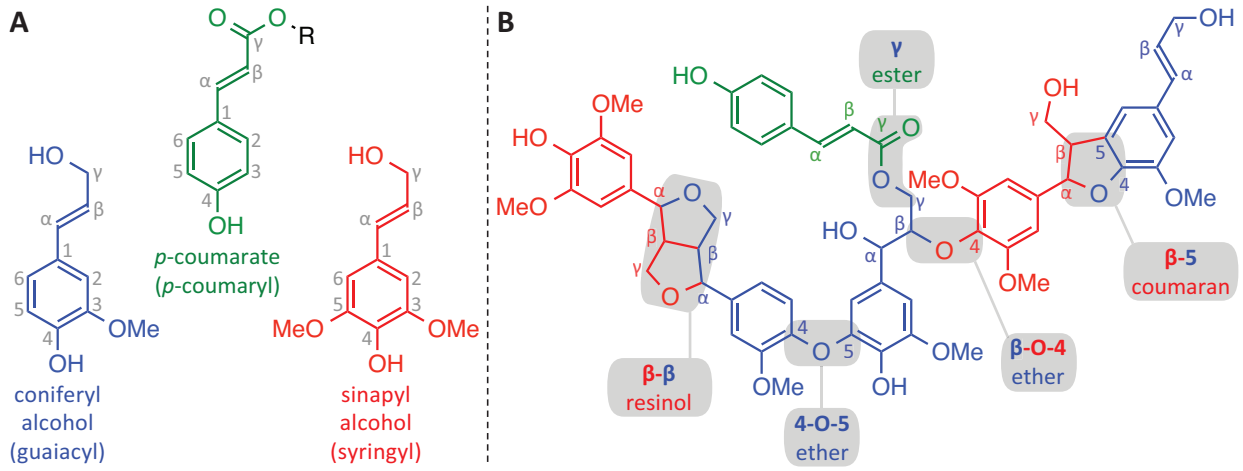

Figure 1-2. The metabolic precursors and dimeric substructures, guaiacyl (blue), and syringyl (red) lignin units. *p*-Coumarates (green), are attached to monolignols before polymerization and are not themselves involved in radical coupling. **(A)** The biosynthetic precursors of guaiacyl (coniferyl alcohol), syringyl (sinapyl alcohol), and *p*-coumarate, where R is the γ-carbon of guaiacyl or syringyl units) lignin monomeric units. **(B)** Common inter-unit linkages created by radical dimerization of lignin monomers. Inter-unit bonds are bolded and relevant carbon assignments creating inter-unit bonds are shown (*e.g.*, bound ring positions and chain positions α, β, and γ).

β–O–4-Aryl ether (termed β-ether hereafter) bonds account for 50–70% of all inter-unit linkages in lignin (Adler and Eriksoo 1955; Adler 1957; Adler 1977), and thus, cleavage of β-ether linkages is critical to lignin biodegradation. Thus, my thesis research investigated the routes through which microbial metabolic pathways can derive aromatic monomers from β-ether-linked lignin compounds (Masai *et al.* 2003; Sato *et al.* 2009) and reductively transform the monomeric

aromatic compounds into energy-rich aliphatic intermediates of potential commercial use (Harrison and Harwood 2005). To achieve this, I investigated various aspects of β-ether cleavage via the *Sphingobium* sp. strain SYK-6 β-etherase pathway and monoaromatic acid degradation in *Rhodopseudomonas palustris* strain CGA009 via the benzoyl-CoA pathway.

The *Sphingobium* sp. strain SYK-6 β-etherase pathway

In contrast to the oxidative and non-selective microbial methods for degrading lignin (Otsuka *et al.* 2003; Martinez *et al.* 2005), the *Sphingobium* sp. strain SYK-6 β-etherase pathway is collectively non-oxidative and is exclusively selective towards monomeric aromatic compounds as its products. Given that the primary goal of this research was to use to use microbial pathways to convert lignin to reduced intermediates that may serve as precursors in biofuel production pipelines, the selective and non-oxidative nature of the β-etherase pathway was attractive because of its expectedly high yield of monoaromatic products and its ability to derive these compounds while maintaining the reduction-oxidation potential of its substrates and products.

The cleavage of β-ether linkages in lignin model dimers such as guaiacylglycerol-β-guaiacyl ether (Figure 1-3A) reportedly occurs via three successive enzymatic reactions in strain SYK-6 (Fukuzumi and Katayama 1977; Katayama and Fukuzumi 1978; Katayama and Fukuzumi 1979a; Katayama and Fukuzumi 1979b; Masai *et al.* 1989; Masai *et al.* 1991; Sonoki *et al.* 2002; Masai *et al.* 2003; Masai *et al.* 2007; Sato *et al.* 2009; Tanamura *et al.* 2010): (1) nicotinamide adenine dinucleotide (NAD^+) is used as a cosubstrate of Cα-dehydrogenases LigD, LigO, LigL, and LigN in the oxidation of guaiacylglycerol-β-guaiacyl ether to α-keto-guaiacylglycerol-β-guaiacyl ether, yielding NADH (*i.e.*, reduced NAD^+) as a coproduct; (2) β-etherases LigE, LigF, and LigP reductively cleave the β-ether linkages in α-keto-guaiacylglycerol-β-guaiacyl ether using glutathione (GSH) as a reductive cosubstrate, resulting in separation of the residual polymer (shown as ROH) from mono-aromatic derivatives (believed to be GS-HPV) that are conjugated to glutathione through a thioether

linkage; and (3) GSH lyase, LigG, catalyzes the GSH-dependent cleavage of the GSH-conjugated substrate, yielding monoaromatic HPV and glutathione disulfide (GSSG) as coproducts. Intriguingly, the β-etherase pathway appears to catalyze chemical reactions that are biologically equivalent to the zinc-dependent reductive cleavage methods used for *in vitro* analyses of lignin substructures (Lu and Ralph 1997b; Lu and Ralph 1997a).

Although many of the pathway intermediates for degradation of guaiacylglycerol-β-guaiacyl (henceforth just termed guaiacyl-β-guaiacyl) ether have been identified, whether or not the pathway encodes for reactions that catalyze degradation of the three other types of β-ether linkages found in nature (guaiacyl-β-syringyl [Figure 1-3B], syringyl-β-guaiacyl [Figure 1-3C], and syringyl-β-syringyl [Figure 1-3D]) is a topic that has not yet been explored. If β-ether linkages between either guaiacyl or syringyl monomers can be cleaved via the β-etherase pathway, the two expected monoaromatic products of derivatization are HPV and HPS (Figure 1-3E).

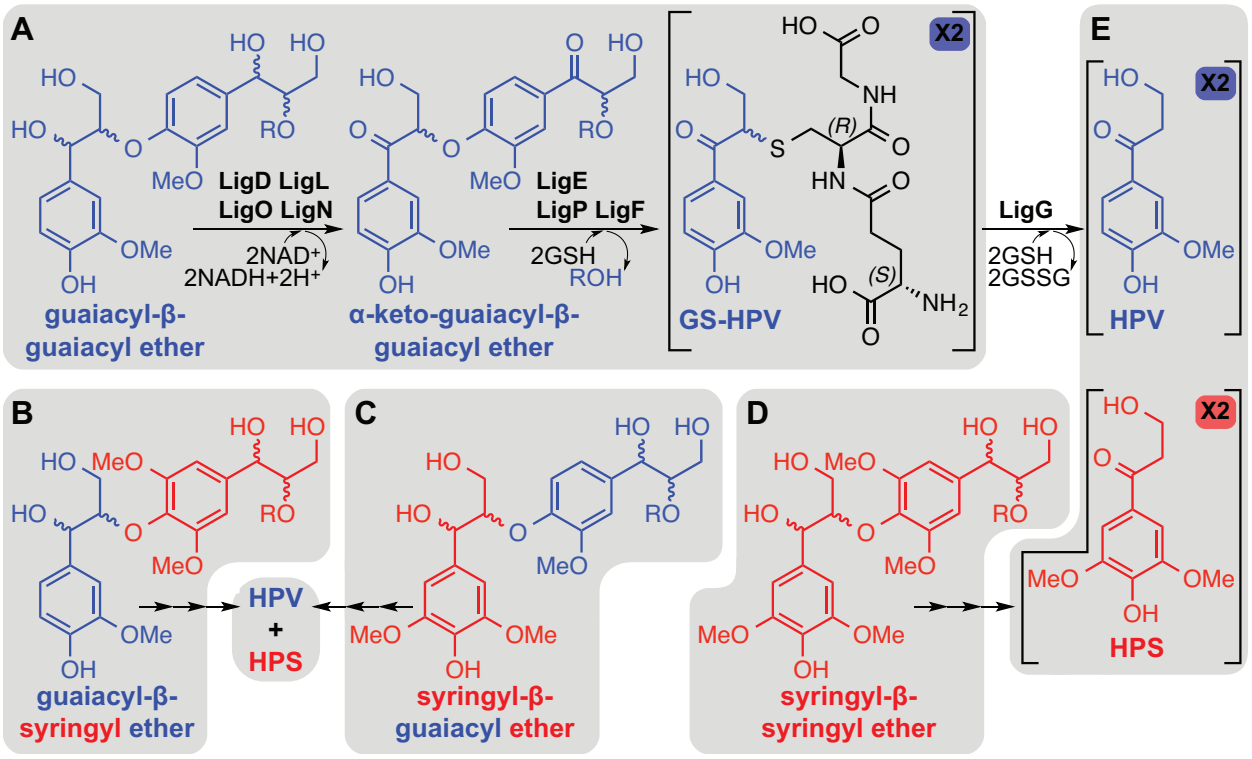

Figure 1-3. (A) The *Sphingobium* sp. strain SYK-6 β-etherase pathway catalyzing degradation of guaiacyl-β-guaiacyl ether. The structures of **(B)** guaiacyl-β-syringyl ether, **(C)** syringyl-β-guaiacyl ether, **(D)** syringyl-β-syringyl ether, and **(E)** HPV and HPS.

Because lignin is synthesized via radical condensation of coniferyl and sinapyl alcohol (Figure 1-2A) monolignols (*i.e.*, inter-unit linkages are formed non-enzymatically), guaiacyl and syringyl monoaromatic units invariably couple at their β-positions with the growing lignin polymer, creating chiral centers at their α- and β-positions with each chain-extending linkage (Higuchi 1980; Hahlbrock and Scheel 1989; Lewis and Yamamoto 1990). These non-stereoselective linkage reactions result in the formation of racemic lignin polymers (Ralph *et al.* 1999; Akiyama *et al.* 2002; Sugimoto *et al.* 2002), and thus, guaiacyl-β-guaiacyl ether (Figure 1-3A) bonds may be found in four possible configurations in nature: α(*S*)-guaiacyl-β(*R*)-guaiacyl ether, α(*R*)-guaiacyl-β(*R*)-guaiacyl ether, α(*S*)-guaiacyl-β(*S*)-guaiacyl ether, and α(*R*)-guaiacyl-β(*S*)-guaiacyl ether (Figure 1-4).

The existence of four guaiacyl-β-guaiacyl ether isomers in nature is overcome by strain SYK-6 because, for each step of the β-etherase pathway, multiple stereospecific enzymes are encoded in the organism's genome that confer both *R*- and *S*-stereospecificities. Of the Cα-dehydrogenases, LigD and LigO catalyze oxidation of the α(*R*)-isomers (α(*R*)-guaiacyl-β(*S*)-guaiacyl ether and α(*R*)-guaiacyl-β(*R*)-guaiacyl ether) whereas the α(*S*)-isomers (α(*S*)-guaiacyl-β(*S*)-guaiacyl ether and α(*S*)-guaiacyl-β(*R*)-guaiacyl ether) are oxidized by LigL and LigN (Figure 1-4) (Sato *et al.* 2009). β-Etherases LigE and LigP cleave one of the two β-enantiomeric ketones, α-keto-guaiacyl-β(*R*)-guaiacyl ether, whereas α-keto-guaiacyl-β(*S*)-guaiacyl ether is cleaved by LigF (Masai *et al.* 2003; Tanamura *et al.* 2010). LigG (and a yet-unidentified presumed stereochemical complement of LigG) stereospecifically cleaves the glutathione-conjugated product from LigF (believed to be β(*R*)-GS-HPV) to coproducts HPV and GSSG.

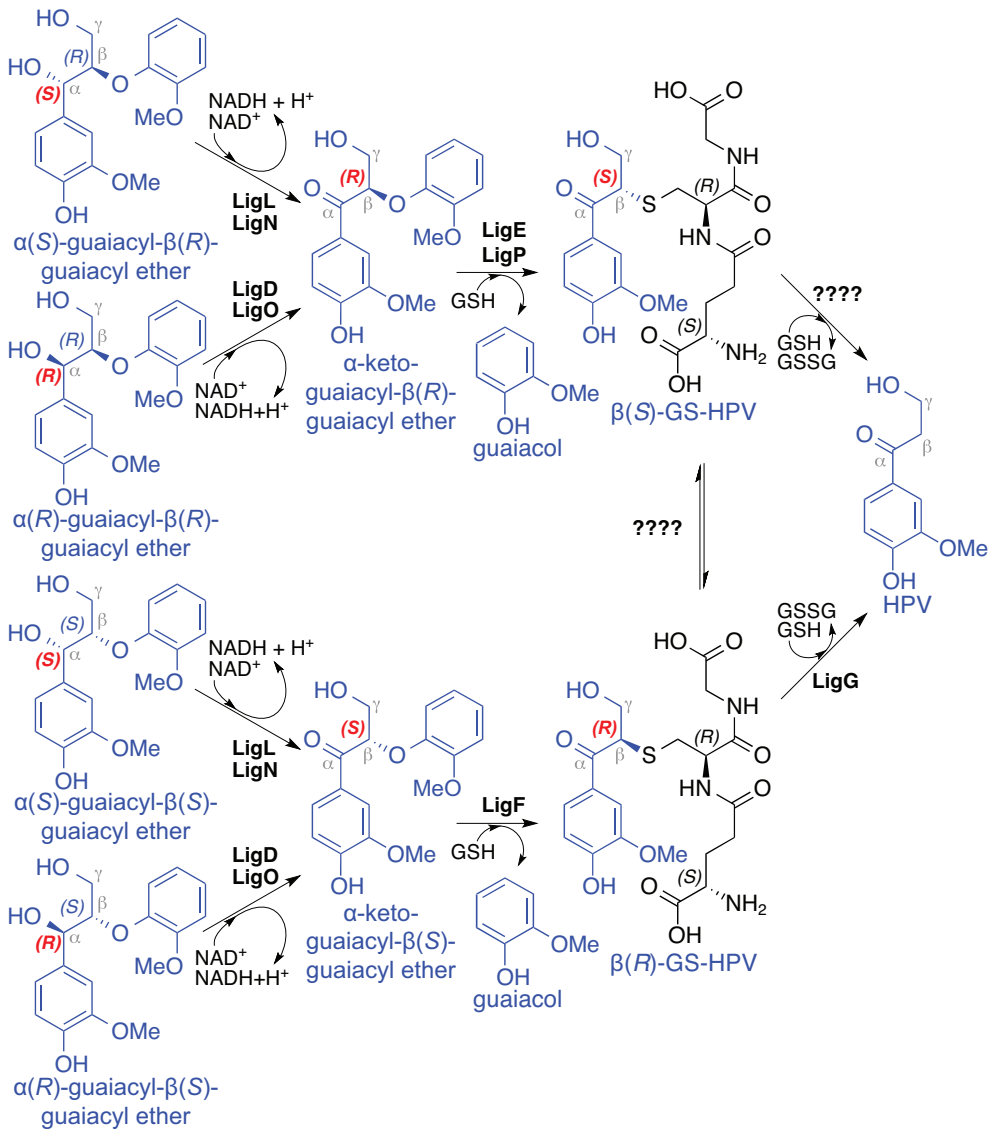

Figure 1-4. The *Sphingobium* sp. strain SYK-6 β-etherase pathway catalyzing degradation of each of the four isomers of guaiacyl-β-guaiacyl ether. Enantiomeric configuration labels for chiral carbons at which stereospecific reactions occur are shown in red.

In the following chapters, I elaborate on investigations carried out for determining the identities of the β-etherase pathway intermediates β(*R*)-GS-HPV and β(*S*)-GS-HPV (Figure 1-4). Further, I describe the study of β-etherase pathway enzymes from other organisms and investigations of α-keto-guaiacyl-β-syringyl ether (Figure 1-3B), α-keto-syringyl-β-guaiacyl ether

(Figure 1-3C), and α-keto-syringyl-β-syringyl ether (Figure 1-3D) as substrates in the pathway. Taken together, one can expect that, if β-etherase pathway reactions can be applied to each of the four types β-ether units found in nature, HPV and HPS (Figure 1-3) will comprise the resulting monoaromatic derivatives of β-etherase pathway catalysis with lignin.

The *Rhodopseudomonas palustris* strain CGA009 benzoyl-CoA pathway

Given that the expected monoaromatic compounds to be derived from lignin via catalysis of β-etherase pathway enzymes are HPV and HPS (Figure 1-3), I investigated metabolism of similar compounds in the bacterium *Rhodopseudomonas palustris*. Because *R. palustris* is known to reductively transform related monoaromatic compounds to long-chain dicarboxylic acids that are potential precursors of biofuel, I sought to determine whether or not these same metabolic products could be reached using lignin-derived aromatics as feedstocks (Gibson and Gibson 1992; Gibson *et al.* 1997; Pelletier and Harwood 2000).

R. palustris uses the benzoyl-CoA pathway for growth on aromatic compounds such as *p*-OH-benzoate and benzoate (Figure 1-5A). *p*-OH-Benzoate and benzoate require activation via Coenzyme A (CoA) ligation, and *p*-OH-benzoyl-CoA is reduced via enzyme HbaBCD to the central intermediate of the pathway, benzoyl-CoA. In *R. palustris*, benzoyl-CoA undergoes a further 4-electron reduction and dearomatization via enzyme BadDEFG to aliphatic the thioester (and potential biofuel precursor), pimeloyl-CoA. Additionally, the phenylpropanoid *p*-coumarate, which is found acylating aromatic monomers in monocot lignins (Figure 1-2), is metabolized via the CouA and CouB enzymes and the benzoyl-CoA pathway (Figure 1-5A). *p*-Coumarate metabolism is accomplished by derivation of acetyl-CoA and *p*-OH-benzaldehyde, and subsequent oxidation of *p*-OH-benzaldehyde to *p*-OH-benzoate.

Similarly, guaiacyl aromatic compounds are partially metabolized via CouA and CouB enzymatic activities (Harwood and Gibson 1988; Pan *et al.* 2008; Hirakawa *et al.* 2012). Like *p*-

coumarate, ferulate (a biosynthetic precursor of coniferyl alcohol in plants) metabolism follows the steps of CoA ligation, hydration, and carbon-carbon bond cleavage, yielding coproducts acetyl-CoA and vanillin that is ultimately oxidized to vanillate (Figure 1-5B). Despite previous reports to the contrary (Harwood and Gibson 1988), my preliminary experiments showed that *R. palustris* was unable to carry out ring fission and growth while using vanillate as a sole source of reducing power.

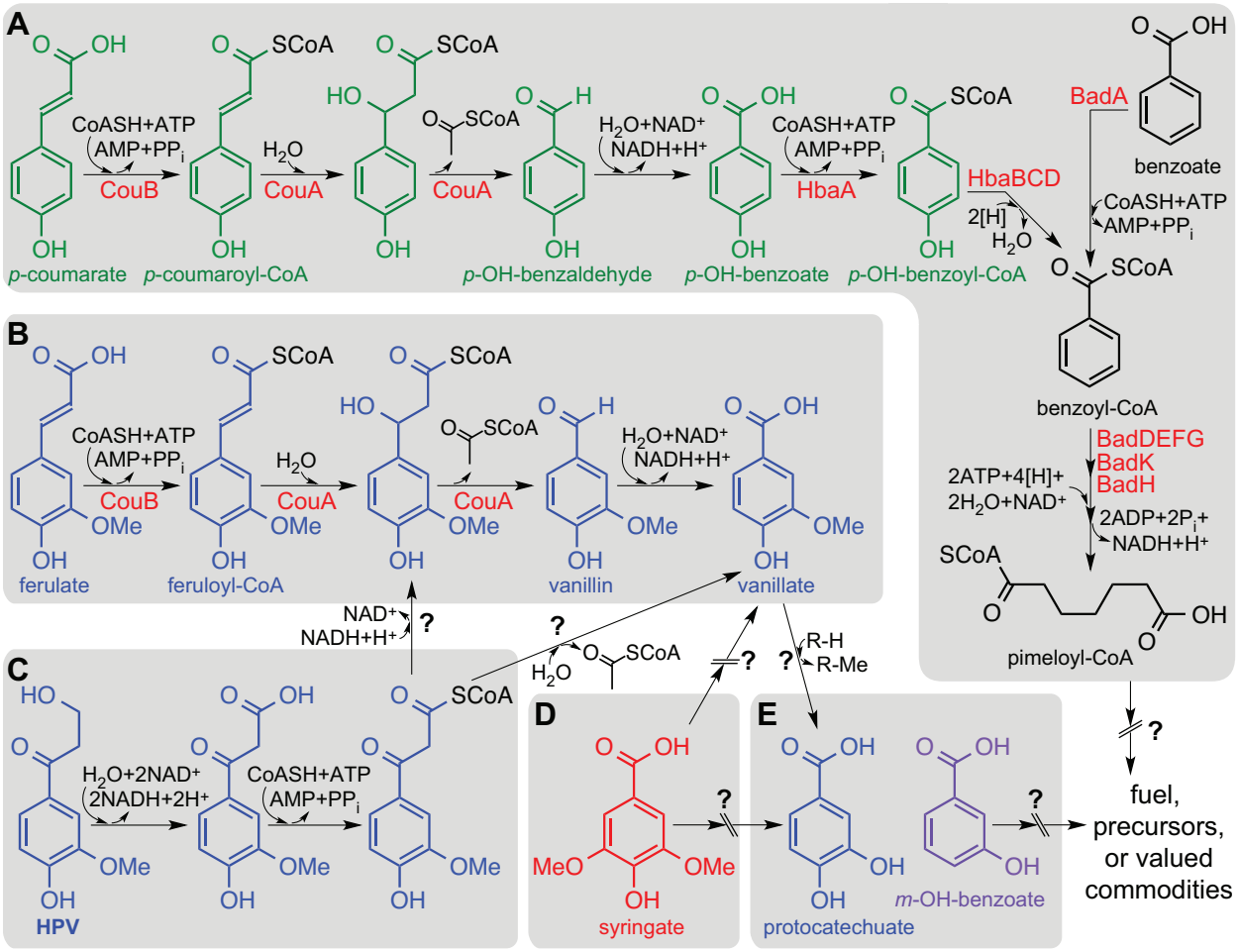

Figure 1-5. Benzoyl-CoA pathway-mediated transformation of aromatic acids. Previously characterized reactions for which genes are known, the encoded enzymes are shown (red) (Egland *et al.* 1997; Larimer *et al.* 2004; Hirakawa *et al.* 2012). Reactions that have yet to be demonstrated *in vivo* or *in vitro* are denoted with a question mark (?). **(A)** Benzoyl-CoA pathway functions used in metabolism of *p*-coumarate, *p*-OH-benzoate, and benzoate (Pan *et al.* 2008). Auxiliary pathways yielding vanillate from phenylpropanoids **(B)** coniferyl alcohol and ferulate (Hirakawa *et al.* 2012) or **(C)** HPV. Structures of **(D)** syringate, as well as **(E)** protocatechuate and *m*-OH-benzoate.

In another photoheterotrophic growth experiment, HPV, one of the expected lignin-derived monoaromatic products of β-etherase catalysis, was partially metabolized by *R. palustris*, apparently by using the aliphatic moiety of HPV for growth while exporting the aromatic moiety, vanillate (Figure 1-5C), to the culture medium. Anticipating that HPS may be metabolized in the same fashion in *R. palustris*, I envisioned that metabolism of the HPV and HPS aliphatic sidechains would result in aromatic carboxylates, vanillate (Figure 1-5B) and syringate (Figure 1-5D), respectively. After carrying out photoheterotrophic growth assays with vanillate and syringate, I found that neither these, nor any *m,m*-dimethoxylated, *m,m*-dihydroxylated, *m*-monomethoxylated, *m*-monomethoxylated or *m*-monohydroxylated aromatic carboxylates were utilized by *R. palustris* as sole sources of reducing power. However, in a later chapter, I describe methods for inducing expression of the benzoyl-CoA pathway enzymes, thereby enabling metabolism of protocatechuate and *m*-OH-benzoate (Figure 1-5E) in *R. palustris* under photoheterotrophic growth conditions.

Thesis structure

The focus of my thesis research was to identify and evaluate pathways that catalyze cleavage of β-ether linkages in lignin oligomers (covered in Chapters 2–4) and reductively dearomatize the monoaromatic derivatives to aliphatic intermediates (covered in Chapter 5) that may serve as precursors for biofuel production. Thus, I organized my thesis into four research chapters, each of which was written as a manuscript for publication in peer-reviewed journals and the described work resulted, in part, from the research I conducted for my dissertation.

First, I present findings from my investigation of the β-etherase pathway Lig enzymes. Previously, the stereospecificities of Cα(*R*)- (LigD and LigO) and Cα(*S*)-dehydrogenases (LigL and LigN), as well as the β(*R*)- (LigE and LigP) and β(*S*)-etherases (LigF) were reported (Masai *et al.* 2003; Sato *et al.* 2009; Tanamura *et al.* 2010). These findings indicated that Cα-dehydrogenation oxidatively eliminates α-chirality without affecting enantiomeric configuration at carbon β, thereby

stereoselectively yielding the two α-ketones that serve as substrates for the β-etherases (Figure 1-4). The mechanism and means of β-etherase catalysis, however, remained in question because the chemical identities of the presumed GSH-conjugated reaction products had not been determined. Because GSH is used as a cosubstrate of the β-etherases and the GSH thiol is commonly used as a nucleophile in enzymatic reactions (Armstrong 1997; Vuilleumier 1997; Sheehan *et al.* 2001; Hayes *et al.* 2005; Oakley 2005; Allocati *et al.* 2009), I hypothesized that β-catalysis is achieved through nucleophilic attack on the substrate (or an enzyme-substrate intermediate) by glutathione, and that β-ether cleavage is coupled to the formation of a β-S-thioether linkage. To test this, I synthesized GS-HPV (Figure 1-3) as a mixture of its two β-epimers (in addition to β(*R*)- and β(*S*)-epimerically pure analogs) and compared their nuclear magnetic resonance (NMR) spectra with those of isolated enzymatic reaction products formed by LigE-, LigP-, and LigF-catalyzed reactions with α-keto-guaiacyl-β-guaiacyl ether as a substrate. In support of my hypothesis, I found that LigE and LigP converted α-keto-guaiacyl-β(*R*)-guaiacyl ether to β(*S*)-GS-HPV and conversely, LigF derived β(*R*)-GS-HPV from α-keto-guaiacyl-β(*S*)-guaiacyl ether. Hence, these results revealed that glutathione-dependent β-etherase catalysis involves formation of a β-S-thioether linkage and is both substrate-stereospecific and product-stereoselective (Chapter 2).

The ever-expanding wealth of genomic sequences in public data banks eventually revealed homologous amino acid sequences to the Lig β-etherases encoded in the *Sphingobium* sp. strain SYK-6 genome (D'Argenio *et al.* 2011; Luo *et al.* 2012; Masai *et al.* 2012). Hypothesizing that this pathway (or possible variations thereof) may exist in other organisms, I cloned genes encoding homologs of LigE, LigP, and LigF from *Novosphingobium aromaticivorans* strain DSM12444 and *Novosphingobium* sp. PP1Y, and in addition to those cloned previously from strain SYK-6, I expressed and purified the encoded recombinant enzymes in order to test them for β-etherase catalysis. Given that, of the four types of β-ether linkages found in nature, only guaiacyl-β-guaiacyl-linked compounds (Figure 1-3A) had been tested as substrates of this pathway in strain SYK-6, I

also investigated the abilities of each homologous enzyme to cleave guaiacyl-β-syringyl (Figure 1-3B), syringyl-β-guaiacyl (Figure 1-3C), and syringyl-β-syringyl ether linkages (Figure 1-3D) in synthesized model lignin compounds. This study revealed for the first time that the β-etherase pathway is encoded in other organisms besides strain SYK-6 and demonstrated that each enzyme catalyzes cleavage of each of the four types of β-ether linkages found in lignin (Figure 1-3A–D). Additionally, these experiments showed that, while indifferent to the guaiacyl/syringyl composition of the substrate, each LigE/LigP homolog exhibited β(*R*)-stereospecificity whereas β(*S*)-enantiomers were exclusively cleaved by LigF and its homologs (Chapter 3).

The study of Lig β-etherases resulted in a collaboration with crystallographers in Dr. George Phillips laboratory (Rice University, formerly UW-Madison) and colleagues at the Joint BioEnergy Institute (Emeryville, CA). Crystal structures of Cα-dehydrogenases LigD and LigL, as well as β-etherases LigE and LigF were obtained and, from their analyses, we formed hypotheses concerning the nature of Lig enzymes' catalytic reaction mechanisms and tested these by conducting biochemical kinetic assays for wild type and variant enzymes. For β-etherases LigE and LigF, we hypothesized that a serine in each was responsible for GSH deprotonation and activation of the thiol for attack. In each case, variation of that serine to an alanine resulted in reduced enzymatic activity, suggesting that each serine residue has either a structural or catalytic role in the reaction. I also synthesized a β-fluorinated analog of α-keto-guaiacyl-β-guaiacyl ether (Figure 1-3A) and demonstrated that it undergoes cleavage with LigE as a biocatalyst. Given these observations, and my previous findings that β-etherases cause inversion of β-chirality, we surmise that the enzymatic reaction mechanisms of LigE and LigF are governed by an S_N2-style nucleophilic attack by the GSH thiolate resulting in displacement of the β-ether linkage (Chapter 4).

Preliminary experiments testing the ability of *Rhodopseudomonas palustris* to grow on the expected monoaromatic products of β-ether cleavage in lignin (HPV and HPS, Figure 1-3E) revealed

that the aliphatic moiety of HPV supports *R. palustris* growth (Figure 1-5C), yet the resulting aromatic acid, vanillate (Figure 1-5B), is not degraded. Although neither vanillate nor syringate (the expected aromatic intermediate of HPS metabolism, Figure 1-5D) supported growth as sole sources of reducing power, I hypothesized that *R. palustris* may metabolize these, or other *m*-substituted aromatic carboxylates under certain growth conditions. Because protocatechuate and *m*-OH-benzoate contain *m*-hydroxyls and are thus structurally similar to vanillate and syringate (*e.g.*, *m*-O-demethylation of vanillate yields protocatechuate (Nishikawa *et al.* 1998; Sonoki *et al.* 2000)), I investigated whether *R. palustris* metabolizes either protocatechuate or *m*-OH-benzoate under photoheterotrophic conditions. Although neither protocatechuate nor *m*-OH-benzoate supported growth as sole sources of reducing power, I found that, in the presence of other aromatic carboxylates that are known to induce expression of the benzoyl-CoA pathway enzymes (*e.g.*, benzoate), *R. palustris* used both protocatechuate and *m*-OH-benzoate to stimulate photoheterotrophic growth. Subsequently, I carried out growth assays with mutant strains of *R. palustris* and these experiments demonstrated that, like benzoate, protocatechuate and *m*-OH-benzoate are metabolized via the benzoyl-CoA pathway (Chapter 5).

In the closing chapter (Chapter 6), I summarize my findings and examine future endeavors that may harness the implications of my work for using metabolic pathways to transform lignin into valued societal commodities. In sum, many of the goals that were envisioned at the onset of my project were accomplished, and as my training and experience progressed in topics outside of my academic major, I feel I was able to apply new ambitions to my research and successfully answer questions that were previously beyond my capacities. I believe that my work provides the scientific community with novel insight into strategies used by microorganisms for metabolism of recalcitrant organic compounds in the environment and hope that, by approaching my research with a multi-disciplinary approach, my data and analyses will be intriguing to engineers, chemists, microbiologists, and enzymologists alike.

References

Adler E. (1957). Structural elements of lignin. *Industrial & Engineering Chemistry* 49(9):1377-1383.

Adler E. (1977). Lignin chemistry - past, present and future. *Wood Science and Technology* 11(3):169-218.

Adler E., Eriksoo E. (1955). Guaiacylglycerol and its β-guaiacyl ether. *Acta chemica Scandinavica* 9341-342.

Akiyama T., Sugimoto T., Matsumoto Y., Meshitsuka G. (2002). *Erythro/threo* ratio of β-O-4 structures as an important structural characteristic of lignin. I: Improvement of ozonation method for the quantitative analysis of lignin side-chain structure. *Journal of Wood Science* 48(3):210-215.

Allocati N., Federici L., Masulli M., Di Ilio C. (2009). Glutathione transferases in bacteria. *Febs Journal* 276(1):58-75.

Argyropoulos D.S., Menachem S.B. (1998) Biopolymers from renewable resources. Springer, Berlin

Armstrong R.N. (1997). Structure, catalytic mechanism, and evolution of the glutathione transferases. *Chemical Research in Toxicology* 10(1):2-18.

Bhuiyan N.H., Selvaraj G., Wei Y., King J. (2009). Role of lignification in plant defense. *Plant signaling & behavior* 4(2):158-159.

Boerjan W., Ralph J., Baucher M. (2003). Lignin biosynthesis. *Annual Review of Plant Biology* 54519-546.

Boyce C.K., Zwieniecki M.A., Cody G.D., Jacobsen C., Wirick S., Knoll A.H., Holbrook N.M. (2004). Evolution of xylem lignification and hydrogel transport regulation. *Proceedings of the National Academy of Sciences of the United States of America* 101(50):17555-17558.

Chen F., Dixon R.A. (2007). Lignin modification improves fermentable sugar yields for biofuel production. *Nature Biotechnology* 25(7):759-761.

D'Argenio V., Petrillo M., Cantiello P., Naso B., Cozzuto L., Notomista E., Paolella G., Di Donato A., Salvatore F. (2011). *De novo* sequencing and assembly of the whole genome of *Novosphingobium* sp. strain PP1Y. *Journal of Bacteriology* 193(16):4296-4296.

Dale B.E., Latimer V.M., Leong C.K., Pham T.K., Esquivel V.M. (1993). Fermentable sugar yields from AFEX-treated corn fiber and switchgrass at low enzyme levels. *Abstracts of Papers of the American Chemical Society* 20534-BTEC.

Egland P.G., Pelletier D.A., Dispensa M., Gibson J., Harwood C.S. (1997). A cluster of bacterial genes for anaerobic benzene ring biodegradation. *Proceedings of the National Academy of Sciences of the United States of America* 94(12):6484-6489.

Freudenberg K. (1959). Biosynthesis and constitution of lignin. *Nature* 183(4669):1152-1155.

Fukuzumi T., Katayama Y. (1977). Bacterial degradation of dimer relating to structure of lignin I. β-hydroxypropiovanillone and coniferylalcohol as initial degradation products from guaiacylglycerol-β-coniferylether by *Pseudomonas putida*. *Mokuzai Gakkaishi* 23(4):214-215.

Gibson J., Dispensa M., Harwood C.S. (1997). 4-Hydroxybenzoyl coenzyme A reductase (dehydroxylating) is required for anaerobic degradation of 4-hydroxybenzoate by

Rhodopseudomonas palustris and shares features with molybdenum-containing hydroxylases. *Journal of Bacteriology* 179(3):634-642.

Gibson K.J., Gibson J. (1992). Potential early intermediates in anaerobic benzoate degradation by *Rhodopseudomonas palustris*. *Applied and Environmental Microbiology* 58(2):696-698.

Grabber J.H., Quideau S., Ralph J. (1996). *p*-Coumaroylated syringyl units in maize lignin: Implications for β-ether cleavage by thioacidolysis. *Phytochemistry* 43(6):1189-1194.

Hahlbrock K., Scheel D. (1989). Physiology and molecular biology of phenylpropanoid metabolism. *Annual Review of Plant Physiology and Plant Molecular Biology* 40347-369.

Harrison F.H., Harwood C.S. (2005). The *pimFABCDE* operon from *Rhodopseudomonas palustris* mediates dicarboxylic acid degradation and participates in anaerobic benzoate degradation. *Microbiology-Sgm* 151727-736.

Harwood C.S., Gibson J. (1988). Anaerobic and aerobic metabolism of diverse aromatic compounds by the photosynthetic bacterium *Rhodopseudomonas palustris*. *Applied and Environmental Microbiology* 54(3):712-717.

Hayes J.D., Flanagan J.U., Jowsey I.R. (2005). Glutathione transferases. *Annual Review of Pharmacology and Toxicology* 4551-88.

Hejnowicz Z. (2005). Unusual metaxylem tracheids in petioles of *Amorphophallus* (Araceae) giant leaves. *Annals of Botany* 96(3):407-412.

Higuchi T. (1980) Lignin structure and morphological distribution in plant cell walls. In: Kirk T.K., Higuchi T., Chang H. (eds) Lignin biodegradation: microbiology, chemistry and potential applications. CRC Press, Boca Raton, Florida, pp 1-20

Hirakawa H., Schaefer A.L., Greenberg E.P., Harwood C.S. (2012). Anaerobic *p*-coumarate degradation by *Rhodopseudomonas palustris* and identification of CouR, a MarR repressor protein that binds *p*-coumaroyl coenzyme A. *Journal of Bacteriology* 194(8):1960-1967.

Iiyama K., Lam T.B.T., Stone B.A. (1994). Covalent cross-links in the cell-wall. *Plant Physiology* 104(2):315-320.

Katayama Y., Fukuzumi T. (1978). Bacterial degradation of dimers structurally related to lignin II. Initial intermediate products from deydrodiconiferyl alcohol by *Pseudomonas putida*. *Mokuzai Gakkaishi* 24(9):643-649.

Katayama Y., Fukuzumi T. (1979a). Bacterial degradation of dimers structurally related to lignin. III. Metabolism of *α*-veratryl-*β*-guaiacylpropionic acid and D,L-pinoresinol by *Pseudomonas putida*. *Mokuzai Gakkaishi* 25(1):67-76.

Katayama Y., Fukuzumi T. (1979b). Bacterial degradation of dimers structurally related to lignin. IV. Metabolism of guaiacylglycerol-*β*-coniferyl ether by *Pseudomonas putida*. *Mokuzai Gakkaishi* 25(5):367-373.

Larimer F.W., Chain P., Hauser L., Lamerdin J., Malfatti S., Do L., Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R., Gibson J.L., Hanson T.E., Bobst C., Torres J., Peres C., Harrison F.H., Gibson J., Harwood C.S. (2004). Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris*. *Nature Biotechnology* 22(1):55-61.

Lewis N.G., Yamamoto E. (1990). Lignin - occurrence, biogenesis and biodegradation. *Annual Review of Plant Physiology and Plant Molecular Biology* 41455-496.

Lloyd T.A., Wyman C.E. (2005). Combined sugar yields for dilute sulfuric acid pretreatment of corn stover followed by enzymatic hydrolysis of the remaining solids. *Bioresource Technology* 96(18):1967-1977.

Lu F.C., Ralph J. (1997a). Derivatization followed by reductive cleavage (DFRC method), a new method for lignin analysis: Protocol for analysis of DFRC monomers. *Journal of Agricultural and Food Chemistry* 45(7):2590-2592.

Lu F.C., Ralph J. (1997b). DFRC method for lignin analysys. 1. New method for β-aryl ether cleavage: Lignin model studies. *Journal of Agricultural and Food Chemistry* 45(12):4655-4660.

Luo Y.R., Kang S.G., Kim S.-J., Kim M.-R., Li N., Lee J.-L., Kwon K.K. (2012). Genome Sequence of Benzo(a)pyrene-Degrading Bacterium Novosphingobium pentaromativorans US6-1. *Journal of Bacteriology* 194(4):907-907.

Lynd L.R., Cushman J.H., Nichols R.J., Wyman C.E. (1991). Fuel ethanol from cellulosic biomass. *Science* 251(4999):1318-1323.

Lynd L.R., Weimer P.J., van Zyl W.H., Pretorius I.S. (2002). Microbial cellulose utilization: Fundamentals and biotechnology (vol 66, pg 506, 2002). *Microbiology and Molecular Biology Reviews* 66(4):739-739.

Martinez A.T., Speranza M., Ruiz-Duenas F.J., Ferreira P., Camarero S., Guillen F., Martinez M.J., Gutierrez A., del Rio J.C. (2005). Biodegradation of lignocellulosics: microbial chemical, and enzymatic aspects of the fungal attack of lignin. *International Microbiology* 8(3):195-204.

Masai E., Ichimura A., Sato Y., Miyauchi K., Katayama Y., Fukuda M. (2003). Roles of the enantioselective glutathione S-transferases in cleavage of β-aryl ether. *Journal of Bacteriology* 185(6):1768-1775.

Masai E., Kamimura N., Kasai D., Oguchi A., Ankai A., Fukui S., Takahashi M., Yashiro I., Sasaki H., Harada T., Nakamura S., Katano Y., Narita-Yamada S., Nakazawa H., Hara H., Katayama Y., Fukuda M., Yamazaki S., Fujita N. (2012). Complete genome sequence of *Sphingobium* sp, strain SYK-6, a degrader of lignin-derived biaryls and monoaryls. *Journal of Bacteriology* 194(2):534-535.

Masai E., Katayama Y., Fukuda M. (2007). Genetic and biochemical investigations on bacterial catabolic pathways for lignin-derived aromatic compounds. *Bioscience Biotechnology and Biochemistry* 71(1):1-15.

Masai E., Katayama Y., Kawai S., Nishikawa S., Yamasaki M., Morohoshi N. (1991). Cloning and sequencing of the gene a *Pseudomonas-paucimobilis* enzyme that cleaves β-aryl ether. *Journal of Bacteriology* 173(24):7950-7955.

Masai E., Katayama Y., Nishikawa S., Yamasaki M., Morohoshi N., Haraguchi T. (1989). Detection and localization of a new enzyme catalyzing the β-aryl ether cleavage in the soil bacterium (*Pseudomonas-paucimobilis* SYK-6). *Febs Letters* 249(2):348-352.

Murnen H.K., Balan V., Chundawat S.P.S., Bals B., Sousa L.d.C., Dale B.E. (2007). Optimization of ammonia fiber expansion (AFEX) pretreatment and enzymatic hydrolysis of *Miscanthus x giganteus* to fermentable sugars. *Biotechnology Progress* 23(4):846-850.

Nishikawa S., Sonoki T., Kasahara T., Obi T., Kubota S., Kawai S., Morohoshi N., Katayama Y. (1998). Cloning and sequencing of the *Sphingomonas* (*Pseudomonas*) *paucimobilis* gene essential for the O-demethylation of vanillate and syringate. *Applied and Environmental Microbiology* 64(3):836-842.

Oakley A.J. (2005). Glutathione transferases: new functions. *Current Opinion in Structural Biology* 15(6):716-723.

Otsuka Y., Sonoki T., Ikeda S., Kajita S., Nakamura M., Katayama Y. (2003). Detection and characterization of a novel extracellular fungal enzyme that catalyzes the specific and hydrolytic cleavage of lignin guaiacylglycerol-β-aryl ether linkages. *European Journal of Biochemistry* 270(11):2353-2362.

Palmqvist E., Hahn-Hägerdal B. (2000a). Fermentation of lignocellulosic hydrolysates. I: inhibition and detoxification. *Bioresource Technology* 74(1):17-24.

Palmqvist E., Hahn-Hägerdal B. (2000b). Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. *Bioresource Technology* 74(1):25-33.

Pan C., Oda Y., Lankford P.K., Zhang B., Samatova N.F., Pelletier D.A., Harwood C.S., Hettich R.L. (2008). Characterization of anaerobic catabolism of *p*-coumarate in *Rhodopseudomonas palustris* by integrating transcriptomics and quantitative proteomics. *Molecular & Cellular Proteomics* 7(5):938-948.

Pelletier D.A., Harwood C.S. (2000). 2-Hydroxycyclohexanecarboxyl coenzyme A dehydrogenase, an enzyme characteristic of the anaerobic benzoate degradation pathway used by *Rhodopseudomonas palustris*. *Journal of Bacteriology* 182(10):2753-2760.

Ragauskas A.J., Beckham G.T., Biddy M.J., Chandra R., Chen F., Davis M.F., Davison B.H., Dixon R.A., Gilna P., Keller M., Langan P., Naskar A.K., Saddler J.N., Tschaplinski T.J., Tuskan G.A., Wyman C.E. (2014). Lignin valorization: Improving lignin processing in the biorefinery. *Science (New York, N.Y.)* 344(6185):1246843-1246843.

Ralph J. (2010). Hydroxycinnamates in lignification. *Phytochemistry Reviews* 9(1):65-83.

Ralph J., Bunzel M., Marita J.M., Hatfield R.D., Lu F., Kim H., Schatz P.F., Grabber J.H., Steinhart H. (2004a). Peroxidase-dependent cross-linking reactions of *p*-hydroxycinnamates in plant cell walls. *Phytochemistry Reviews* 3(1-2):79-96.

Ralph J., Lundquist K., Brunow G., Lu F., Kim H., Schatz P.F., Marita J.M., Hatfield R.D., Ralph S.A., Christensen J.H., Boerjan W. (2004b). Lignins: natural polymers from oxidative coupling of 4-hydroxyphenyl-propanoids. *Phytochem Rev* 329-60.

Ralph J., Peng J.P., Lu F.C., Hatfield R.D., Helm R.F. (1999). Are lignins optically active? *Journal of Agricultural and Food Chemistry* 47(8):2991-2996.

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. *Applied and Environmental Microbiology* 75(16):5195-5201.

Sheehan D., Meade G., Foley V.M., Dowd C.A. (2001). Structure, function and evolution of glutathione transferases: Implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochemical Journal* 3601-16.

Simmons B.A., Logue D., Ralph J. (2010). Advances in modifying lignin for enhanced biofuel production. *Current Opinion in Plant Biology* 13(3):313-320.

Sonoki T., Iimura Y., Masai E., Kajita S., Katayama Y. (2002). Specific degradation of β-aryl ether linkage in synthetic lignin (dehydrogenative polymerizate) by bacterial enzymes of *Sphingomonas paucimobilis* SYK-6 produced in recombinant *Escherichia coli*. *Journal of Wood Science* 48(5):429-433.

Sonoki T., Obi T., Kubota S., Higashi M., Masai E., Katayama Y. (2000). Coexistence of two different O demethylation systems in lignin metabolism by *Sphingomonas paucimobilis* SYK-6: Cloning and sequencing of the lignin biphenyl-specific O-demethylase (LigX) gene. *Applied and Environmental Microbiology* 66(5):2125-2132.

Sugimoto T., Akiyama T., Matsumoto Y., Meshitsuka G. (2002). The *erythro/threo* ratio of β-O-4 structures as an important structural characteristic of lignin - Part 2. Changes in *erythro/threo* (E/T) ratio of β-O-4 structures during delignification reactions. *Holzforschung* 56(4):416-421.

Tanamura K., Kasai D., Nakamura M., Katayama Y., Fukuda M., Masai E. (2010). Identification of the third glutathione S-transferase gene involved in the stereospecific cleavage of β-aryl ether in *Sphingobium* sp. strain SYK-6. *Journal of Biotechnology* 150S235-S235.

Teymouri F., Laureano-Perez L., Alizadeh H., Dale B.E. (2004). Ammonia fiber explosion treatment of corn stover. *Applied Biochemistry and Biotechnology* 113951-963.

Tuck C.O., Perez E., Horvath I.T., Sheldon R.A., Poliakoff M. (2012). Valorization of biomass: Deriving more value from waste. *Science* 337(6095):695-699.

Vuilleumier S. (1997). Bacterial glutathione S-transferases: What are they good for? *Journal of Bacteriology* 179(5):1431-1441.

Yang B., Wyman C.E. (2004). Effect of xylan and lignin removal by batch and flowthrough pretreatment on the enzymatic digestibility of corn stover cellulose. *Biotechnology and Bioengineering* 86(1):88-95.

Yang W.R., Sen A. (2010). One-step catalytic transformation of carbohydrates and cellulosic biomass to 2,5-dimethyltetrahydrofuran for liquid fuels. *Chemsuschem* 3(5):597-603.

CHAPTER 2: Stereochemical features of glutathione-dependent enzymes in the *Sphingobium* sp. strain SYK-6 β-Aryl Etherase Pathway

This chapter is published under the same title:

Gall D.L., Kim H., Lu F., Donohue T.J., Noguera D.R., Ralph J. (2014a). Stereochemical features of glutathione-dependent enzymes in the *Sphingobium* sp. strain SYK-6 β-aryl etherase pathway. *J Biol Chem* 289(12):8656-8667.

doi: 10.1074/jbc.M113.536250

Supplementary Information associated with this publication:
http://www.jbc.org/content/suppl/2014/02/07/M113.536250.DC1/jbc.M113.536250-1.pdf

Daniel L. Gall performed all of the experiments in this chapter with guidance pertaining to organic syntheses and structural analyses from Hoon Kim and Fachuang Lu.

Capsule

Background: A bacterial β-etherase pathway uses glutathione-dependent enzymes for catabolism of β-ether-linked substructures found in lignin.

Results: Racemic β-ether-linked substructures are stereoselectively converted to thioether-linked mono-aromatic compounds by β-etherase pathway enzymes.

Conclusion: Multiple enzymes with complementary stereochemical features are needed to catabolize racemic lignin substructures.

Significance: Understanding β-etherase pathway metabolism of lignin enhances our potential to use lignin for agricultural, industrial, and biotechnological purposes.

Abstract

Glutathione-dependent enzymes play important protective, repair, or metabolic roles in cells. In particular, enzymes in the glutathione-S-transferase (GST) superfamily function in stress responses, defense systems, or xenobiotic detoxification. Here, we identify novel features of bacterial GSTs that cleave β-aryl ether bonds typically found in plant lignin. Our data reveal several original features of the reaction cycle of these GSTs, including stereospecific substrate recognition and stereoselective formation of β-S-thioether linkages. Products of recombinant GSTs (LigE, LigP, and LigF) are β-S-glutathionyl-α-keto-β-aryl ethers that are degraded by a β-S-thioetherase (LigG). All three Lig GSTs produced the ketone product (β-S-glutathionyl-α-veratrylethanone) from an achiral sidechain-truncated model substrate (β-guaiacyl-α-veratrylethanone). However, when β-etherase assays were conducted with a racemic model substrate, β-guaiacyl-α-veratrylglycerone, LigE- or LigP-catalyzed reactions yielded only one of two potential product (β-S-glutathionyl-α-veratrylglycerone) epimers, whereas the other diastereomer (differing in configuration at the β-position, *i.e.*, its β-epimer) was produced only in the LigF-catalyzed reaction. Thus, β-etherase catalysis causes stereochemical inversion of the chiral center, converting a β(*R*)-substrate to a β(*S*)-product (LigE and LigP), and a β(*S*)-substrate to a β(*R*)-product (LigF). Further, LigG catalyzed glutathione-dependent β-S-thioether cleavage with β-S-glutathionyl-α-veratrylethanone and with β(*R*)-configured β-S-glutathionyl-α-veratrylglycerone, but exhibited no, or significantly reduced β-S-thioether-cleaving activity with the β(*S*)-epimer, demonstrating that LigG is a stereospecific β-thioetherase. We therefore propose that multiple Lig enzymes are needed in this β-aryl etherase pathway in order to cleave the racemic β-ether linkages that are present in the backbone of the lignin polymer.

Introduction

Glutathione (GSH) and GSH-dependent enzymes play crucial roles in cellular stress responses, protection against cell damage, or in detoxification of xenobiotic compounds (Vuilleumier 1997; Sheehan *et al.* 2001; Hayes *et al.* 2005; Oakley 2005). In well-studied members of the glutathione-S-transferase (GST) family, the GSH thiol initiates a nucleophilic attack on the substrate during catalysis. Here we report new and novel aspects of GSTs that are implicated in lignin degradation.

Lignin is a major component of plant cell walls that is composed of polymerized aromatic subunits (Freudenberg 1959; Higuchi 1980; Lewis and Yamamoto 1990). This polymer provides both structural rigidity and protection from pathogens (Sarkanen and Ludwig 1971; Dixon and Paiva 1995). Lignin could be of economic value if efficient depolymerization strategies can be developed (Simmons *et al.* 2010). β-O-4′-Aryl ether (henceforth termed β-ether) bonds comprise the majority of the linkages between monomer-derived aromatic units in lignin (Adler 1977), so cleavage of these bonds could yield valuable products (Zakzeski *et al.* 2010; Bugg *et al.* 2011). Here, we report on activities of GST family members that are proposed to cleave β-ether bonds typical of those found in the lignin backbone.

The GST proteins that we are studying are part of a bacterial pathway (Figure 2-1A–B) that allows growth on complex aromatic substrates (Masai *et al.* 1989; Masai *et al.* 1991; Masai *et al.* 1993a; Masai *et al.* 1993b; Hara *et al.* 2000; Masai *et al.* 2003; Masai *et al.* 2007). Although many GST family members are well studied, the properties of these enzymes are not well known. Plant lignin is synthesized via radical-coupling of aromatic monomers, the hydroxycinnamyl alcohols or monolignols, mainly by their chemical condensation with the growing polymer. As monolignols invariably couple at their β-positions, chiral centers are generated in each chain-extension step. Each dimeric unit in the polymer contains two chiral centers but the polymer is racemic (Ralph *et*

al. 1999; Akiyama *et al.* 2000). The mechanisms by which the enzymes in this degradative Lig pathway process the racemic β-ether units found in lignin is an enduring knowledge gap.

In the best studied β-ether degradation pathway from the α-Proteobacterium *Sphingobium* sp. strain SYK-6, the benzylic ketones proposed to be β-etherase substrates (Masai *et al.* 1989; Masai *et al.* 1991; Masai *et al.* 1993b) are produced by oxidation of their corresponding benzylic alcohols (Sato *et al.* 2009). Based on the existence of multiple α-dehydrogenases that oxidize these benzylic alcohols, it is proposed that each enzyme separately oxidizes substrates with different stereochemical configurations at the α-position (Akiyama *et al.* 2000; Akiyama *et al.* 2002; Masai *et al.* 2007; Ralph and Landucci 2010). For example, oxidation of a model dimeric Lig substrate β-guaiacyl-α-veratrylglycerol yields two enantiomeric products, β(*R*)- and β(*S*)-enantiomers (Figure 2-1A–B) of β-guaiacyl-α-veratrylglycerone.

In the next step of this pathway, three GSH-dependent GST enzymes (LigE, LigP, and LigF) contribute to degradation of substrates containing β-aryl ether linkages like β-guaiacyl-α-veratryl-glycerone, hereafter abbreviated as βGVG (Masai *et al.* 2003; Tanamura *et al.* 2010; Hishiyama *et al.* 2012). In *in vitro* assays with racemic substrates, cell extracts containing LigE and LigP separately degrade the β(*R*)GVG enantiomer while lacking detectable activity with the β(*S*)GVG substrate under identical assay conditions (Figure 2-1A) (Masai *et al.* 2003)**.** In contrast**,** LigF is active with the β(*S*)GVG enantiomer but has no detectable activity with β(*R*)GVG (Figure 2-1B). The LigE, LigP, and LigF reaction products include guaiacol and a previously unidentified compound. The properties of the uncharacterized compound obtained using partially purified LigF were consistent with its being a GSH-conjugated aromatic monomer. However, from available data (Masai *et al.* 2003), the type and location of the bond that links GSH and the aromatic moiety are unknown, leaving open the chemical identity of the intermediates produced by these enzymes (Figure 2-1A–B

shows a putative GSH conjugate to veratrylglycerone, namely β-S-glutathionyl-α-veratrylglycerone, hereafter referred to as GS-βVG).

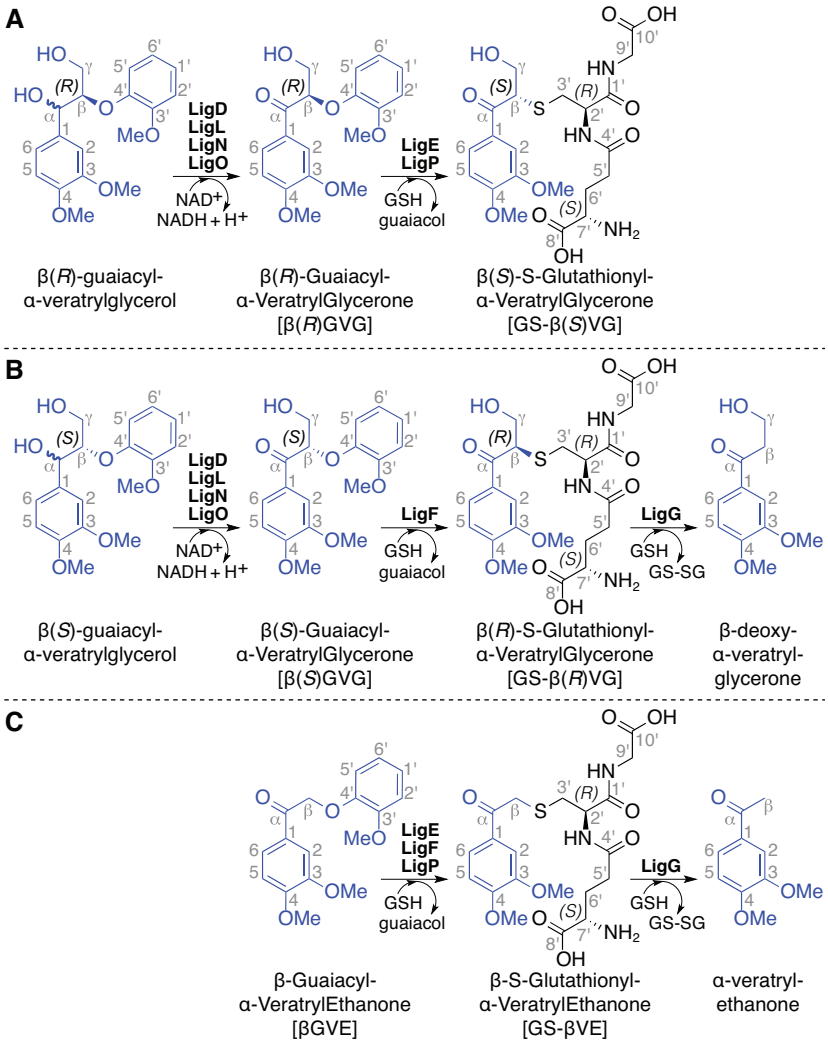

Figure 2-1. The proposed β-etherase pathway in *Sphingobium* sp. strain SYK-6. Panels **(A)** and **(B)** show pathways for metabolism of β(*R*)- and β(*S*)-configured β-aryl ether model compounds, with Lig dehydrogenases (LigD, LigL, LigN, and LigO) catalyzing oxidation of the benzylic alcohol in model substrates β(*R*)- and β(*S*)-guaiacyl-α-veratrylglycerol, producing β(*R*)- and β(*S*)-guaiacyl-α-veratrylglycerone [β(*R*)GVG and β(*S*)GVG]. In the presence of GSH, βGVG enantiomers are cleaved to guaiacol and one of two GS-conjugated β-epimers, β(*S*)- or β(*R*)-S-glutathionyl-α-veratrylglycerone [GS-β(*S*)VG or GS-β(*R*)VG], by either LigE, LigP, or LigF. Subsequent GSH-dependent cleavage of GS-β(*R*)VG by LigG generates β-deoxy-α-veratrylglycerone and glutathione-disulfide (GS-SG). Panel **(C)** shows the reaction products of Lig enzymes when the achiral model compound β-guaiacyl-α-veratrylethanone (βGVE) is used as a substrate. In this case, the product of β-etherase activity is β-S-glutathionyl-α-veratrylethanone (GS-βVE), and the product of GS-βVE thioether cleavage by LigG is α-veratrylethanone.

Evidence supporting the existence of a GSH conjugate in this pathway includes (a) cleavage of β(*S*)GVG and concomitant production of β-deoxy-α-veratrylglycerone (Figure 2-1B) in coupled assays containing racemic βGVG, GSH, LigF, and LigG (a subsequent pathway enzyme believed to have β-thioetherase activity) and (b) the finding that a coupled assay containing LigE, LigG, GSH and racemic βGVG resulted in β(*R*)GVG degradation but did not produce β-deoxy-α-veratrylglycerone (Figure 2-1A) (Masai *et al.* 2003). Based on these facts and the hypothesis that LigF is a stereoselective GST, LigG is proposed to stereospecifically cleave a putative β-S-glutathionyl thioether in the presence of GSH producing glutathione disulfide (GS-SG) and β-deoxy-α-veratryl-glycerone (Figure 2-1B) (Fukuzumi and Katayama 1977; Masai *et al.* 2003; Masai *et al.* 2007). Nevertheless, direct experimental evidence on the substrates of LigG activity is lacking.

Here, we determine the stereochemical properties of GSH-dependent enzymes in the Lig pathway. We show that the GST family members LigE, LigP, and LigF are each β-etherases, that LigG is a β-thioetherase, and we identify previously uncharacterized pathway intermediates. We demonstrate that the GST family members LigE, LigF and LigP, each catalyze stereospecific cleavage of α-keto-β-aryl ethers, β(*R*)GVG (Figure 2-1A) or β(*S*)GVG (Figure 2-1B). We further show that each GST β-etherase is stereoselective, yielding guaiacol and one of two potential β-epimers of GS-βVG. We also demonstrate that LigG is a β-S-thioetherase catalyzing the GSH-dependent cleavage of GS-βVG to β-deoxy-α-veratrylglycerone (Masai *et al.* 2003). To rationalize the observed stereochemical properties of these proteins, we present a model that links the observed substrate specificity of LigE, LigF, LigP and LigG to the racemic nature of the lignin polymer.

Experimental Procedures

Gene Cloning

General – Manipulation of DNA and preparation of *E. coli* transformant cells were carried out according to standard methods (Moore 2003). All DNA primers were obtained commercially from Integrated DNA Technologies (IDT). All restriction and DNA ligation enzymes were purchased from New England Biolabs. Plasmid pMHTΔ238 was obtained from the DNASU Plasmid Repository. All plasmids with the pM (first two letters) naming convention (*i.e.*, all plasmids encoding *Sphingobium* sp. strain SYK-6 Lig enzymes and *Vibrio cholarae* RtxA protease) were obtained commercially from GeneArt® (Life Technologies) and were codon-optimized for expression in *E. coli.*

Preparation of vector pVP102K – Preparation of expression vector pVP102K required the creation of intermediate plasmids pVP100K and pVP101K as follows: Plasmid pVP100K was first constructed by excising a 1,110-bp fragment from pMK1157550 with restriction enzymes BseRI and HindIII and then ligating into the compatible ends at the AsiSI and HindIII sites of pVP80K (obtained from DNASU Plasmid Repository). Intermediate plasmid pVP101K was then constructed by excising a 852-bp fragment from the 6,123-bp plasmid pVP100K using AsiSI and ligating the free ends together to form 5,271-bp plasmid pVP101K. Expression vector pVP102K was then constructed by amplifying a 4,169-bp fragment from plasmid pVP101K, restricting the resulting amplicon with SacI, and ligating the free ends together, resulting in removal of pVP101K's maltose binding protein-encoding region and formation of the 4,153-bp expression vector pVP102K, which was T5-promoter inducible and contained an in-frame N-terminal octa-histidine ($NHis_8$) affinity tag and Tobacco Etch Virus protease cleavage site upstream of the translation start codon.

Preparation of vector pVP202K – Intermediate 4,078-bp plasmid pVP201K was constructed via excision of a 75-bp fragment encoding the $NHis_8$ region of pVP102K by NcoI restriction and subsequent ligation. A 701-bp cysteine protease domain-encoding region of the *Vibrio cholarae*

RtxA protease was then amplified from pMA1157486, both the amplicon and vector pVP201K were restricted with SacII and PmeI, and the 681-bp and 4,037-bp fragments were ligated together to form the 4,718-bp expression vector pVP202K, which contained the C-terminal inducible RtxA protease fused to an $NHis_8$ affinity tag upstream of the translation stop codon.

Preparation of the pVP202KSSLigE expression plasmid – The ORF encoding *Sphingobium* sp. strain SYK-6 LigE on plasmid pMKT1025977 was cloned into expression vector pVP202K by the PCR overlap method (Horton 1993; Shevchuk *et al.* 2004; Bryksin and Matsumura 2010; Horton *et al.* 2013). The first round of PCR yielded a 906-bp amplicon, which was used to prime plasmid pVP202K in the second round of PCR, affording the 5,470-bp plasmid pVP202KSSLigE, which was used for expression of LigE.

Preparation of the pVP202KSSLigP expression plasmid – A 902-bp fragment was amplified from plasmid pJBEI_E2/P (obtained from collaborators at the Joint BioEnergy Institute, Berkeley, CA) and was cloned into vector pVP202K by the PCR overlap method (Horton 1993; Shevchuk *et al.* 2004; Bryksin and Matsumura 2010; Horton *et al.* 2013), affording plasmid pVP202KSSLigP, which was used for the expression of LigP.

Preparation of the pVP102KSSLigF expression plasmid – A 822-bp fragment containing the LigF-encoding ORF was amplified from plasmid pMKT1025979, the amplicon was restricted with AsiSI and SacII, and 809-bp fragment was inserted into the AsiI-SacII region of pVP102K, plasmid pVP102KSSLigF, which was used for the expression of LigF.

Preparation of the pVP102KSSLigG expression plasmid – A 849-bp fragment containing the LigG-encoding ORF from plasmid pMK1118986 was amplified, both the amplicon and vector pVP102K were restricted with AsiSI and SacII, and the 836-bp and 4,067-bp fragments were ligated together, forming 4,093-bp plasmid pVP102KSSLigG, which was used for the expression of LigG.

Enzyme Purification

General – Tranformant cultures of *E. coli* strain B834 were grown aerobically with shaking at 37 °C in Luria-Bertani (LB) medium (Bertani 1951; Bertani 2004) overnight and subcultured (1:100) into 1 L of auto-induction ZYM-5052 medium (Studier 2005) supplemented with 100 μg mL^{-1} Kanamycin. Auto-inducing cultures were grown overnight at 25 °C and then harvested by centrifugation at 10,500×g with a Sorvall RC-5B Superspeed Centrifuge. Pelleted cells were resuspended in ~30 mL (or 2.5 mL buffer g_{cells}^{-1}) Buffer A (50 mM NaH_2PO_4, 100 mM NaCl, 10 mM imidazole, 10% glycerol, 1.0% Triton X-100, 0.5 mM *tris*-(2-carboxyethyl)-phosphine hydrochloride [TCEP], pH 8.0) supplemented with 1 mg mL^{-1} lysozyme. Cells were incubated with lysozyme for 30 min at 4 °C, after which time, the NaCl concentration was brought up to 300 mM using a 5 M NaCl solution, and the slurry was augmented with 10 μg mL^{-1} RNAseA and 5 μg mL^{-1} DNAseI and allowed to sit for 15 min. Cells were then subjected twice to compression through a Spectronic Instruments French pressure cell at 137 MPa and cell debris was removed by centrifugation at 30,000×g. Histidine-tagged proteins were then purified from cell lysates through a series of chromatographic and size-exclusion chromatographic separations using a GE Healthcare ÄKTA Prime Plus FPLC system. First, cell lysates were loaded at 2 mL min^{-1} onto a Pharmacia Biotech XK-16 column packed with 25 mL of QIAGEN nickel-nitrilotriacetic acid (Ni-NTA) resin. The column was then washed with 100 mL of Buffer B (50 mM NaH_2PO_4, 300 mM NaCl, 25 mM imidazole, 0.5 mM TCEP, pH 8.0) and His-tagged proteins were then eluted with a 30-mL linear gradient to 100% Buffer C (50 mM NaH_2PO_4, 300 mM NaCl, 500 mM imidazole, 0.5 mM TCEP, pH 8.0). Collected fractions were then dialyzed for ~12 h at 4 °C using a 10 kDa MWCO Thermo Scientific Slide-a-Lyzer Dialysis Cassette into Buffer D (50 mM NaH_2PO_4, 50 mM NaCl, 10 mM imidazole, 0.5 mM TCEP, pH 8.0). Enzyme-specific methods (see below) were then carried out for the removal of recombinant affinity tags. After proteolysis, affinity tags were removed from the preparation by passage over a second Pharmacia Biotech XK-16 column packed with 30 mL of QIAGEN Ni-NTA resin. The preparation was

then subjected to size-exclusion chromatography in Buffer E (10 mM HEPES, 100 mM NaCl, pH 7.7) at a flow rate of 1 mL min^{-1} using a GE Healthcare HiLoad 16/600 Superdex™ 200 pg column. Finally, each enzyme was concentrated in Buffer E to ≥ 10 mg mL^{-1} by centrifugation at 3,500×g in a 10 kDa MWCO Pierce Concentrator before drop freezing in liquid N_2. Protein concentrations were determined by the Bradford method (Bradford 1976).

Purification of Tev Protease – Tev protease was expressed as a 70.8 kDa fusion to maltose binding protein (Blommel and Fox 2007). To ensure full removal of the 42.6 kDa maltose binding domain by autoproteolysis, cell lysates were allowed to sit for ~2 h between disruption by French press and Ni-NTA affinity chromatography. Tev protease was then subjected to size-exclusion chromatography where it eluted as a monomeric 28.2 kDa enzyme.

Purification of LigE and LigP – LigE and LigP were expressed as either a 56.1 kDa or 55.0 kDa fusions to RtxA and were dialyzed into Buffer D after Ni-NTA affinity chromatography. The RtxA and $CHis_8$ tag was then cleaved through 100 μM inositol hexaphosphate induction of RtxA (Sheahan *et al.* 2007; Prochazkova *et al.* 2009; Shen *et al.* 2009a; Shen *et al.* 2009b). The sample was subjected to a second round of Ni-NTA chromatography for removal of the 24.0 kDa RtxA-$CHis_8$ domain. SDS-12% PAGE electrophoresis of denatured 32.1 kDa LigE and 31.0 kDa LigP enzymes showed that each preparation was homogeneous (data not shown).

Purification of LigF and LigG – $NHis_8$-fused recombinant enzymes LigF (33.4 kDa) and LigG (34.1 kDa) were expressed with the amino acid motif for the consensus cut-site of Tev protease ENLYFQ (where cleavage occurs after Q) encoded immediately after the N-terminal His-tag and prior to the encoding sequence of the respective enzyme (Carrington *et al.* 1988; Dougherty *et al.* 1988; Dougherty *et al.* 1989; Dougherty and Parks 1991). After the first round of Ni-NTA chromatography, the $NHis_8$-tagged enzymes were dialyzed into Buffer D for cleavage of the affinity tags using Tev protease (2 mg TevΔ238 per 100 mg of recombinant enzyme). Subsequently, the

samples were subjected to a second round of Ni-NTA affinity, as well as size exclusion chromatography. Sodium dodecyl sulfate (SDS)-12% polyacrylamide gel electrophoresis (PAGE) of denatured 30.0 kDa LigF and 30.6 kDa LigG preparations showed that each was homogeneous (data not shown).

NMR spectroscopy

^{1}H and ^{13}C NMR spectra were recorded on a Bruker Biospin (Billerica, MA) AVANCE 700 MHz spectrometer fitted with a cryogenically-cooled 5-mm TXI gradient probe with inverse geometry (proton coils closest to the sample). All compounds were characterized and assigned using the usual array of homonuclear and heteronuclear 2D correlation (primarily COSY, HSQC, and HMBC) experiments.

Synthesis of β-ether-linked model compounds

Synthesis of βGVE – The synthesis of βGVE (Figure 2-2) began with bromination of commercially available α-veratrylethanone to produce crystalline β-bromo-α-veratrylethanone. β-Ether formation was via S_N2 displacement of the bromide by the phenolate ion of guaiacol, affording crystalline βGVE (Adler and Eriksoo 1955).

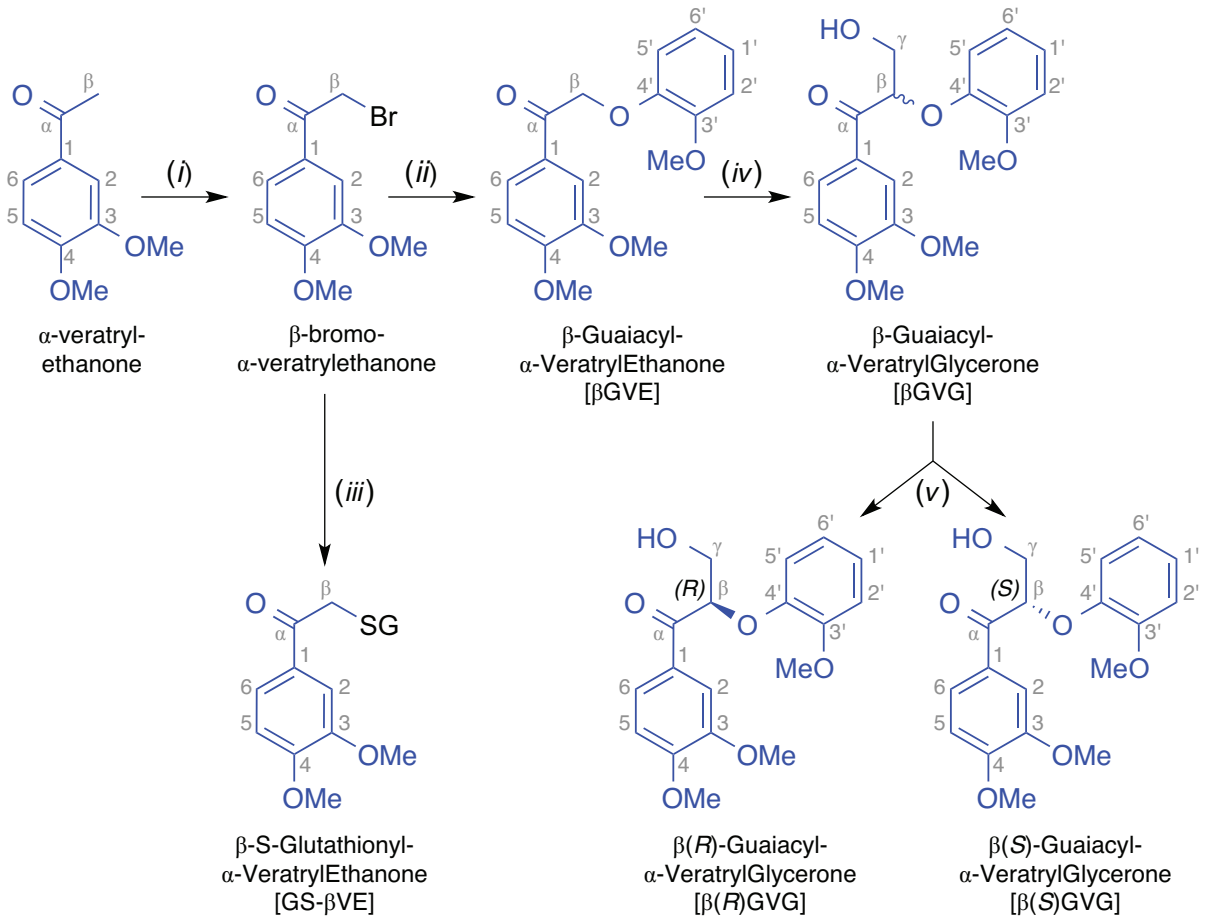

Figure 2-2. Scheme for the synthesis of β-guaiacyl-α-veratrylethanone (βGVE), β-S-glutathionyl-α-veratrylethanone (GS-βVE), and β-guaiacyl-α-veratrylglycerone [β(*R*)GVG, and β(*S*)GVG]. Reagents and conditions: (*i*) pyridinium tribromide, EtOAc, 30 min, 59%; (*ii*) guaiacol, K_2CO_3, acetone, 4 h, 82%; (*iii*) glutathione, $NaHCO_3$, 1:1 acetone/water, 14 h, 100%; (*iv*) formaldehyde, K_2CO_3, 1,4-dioxane, 3 h, 88% (*v*) chiral chromatography, ethanol/hexane. See Supplementary Information for details.

Synthesis of βGVG – βGVE was condensed with formaldehyde to yield racemic βGVG from which enantiopure β(*R*)GVG and β(*S*)GVG were isolated by preparative chiral chromatography. Additional synthetic details can be found in the Supplementary Information.

Synthesis of β-thioether-linked model compounds

Synthesis of GS-βVE – The synthesis of GS-βVE (Figure 2-2) commenced with chemically synthesized β-bromo-α-veratrylethanone and guaiacol as starting materials. β-Thioether formation was via S_N2 displacement of the bromide by the thiolate ion of GSH, yielding GS-βVE, which was then purified by preparative C_{18}-reversed phase chromatography.

Synthesis of GS-βVG – Condensation of commercially available α-veratrylethanone with diethyl carbonate yielded ethyl α-keto-veratrylpropionate (Figure 2-3). Its reduction produced β-deoxy-α-veratrylglycerol, which was converted to β-deoxy-α-veratrylglycerone through benzylic oxidation. Next, bromination yielded β-bromo-α-veratrylglycerone. Thioether formation with GSH then afforded the desired β-epimers of the GSH conjugate GS-βVG, which was then purified by preparative C_{18}-reversed phase chromatography. Additional synthetic details can be found in the Supplementary Information.

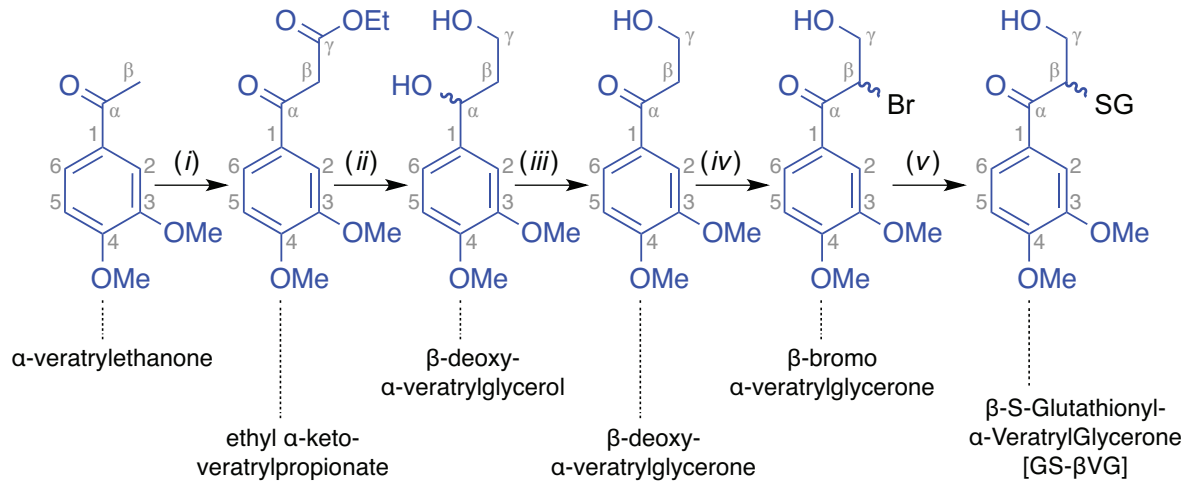

Figure 2-3. Scheme for the synthesis of β-S-glutathionyl-α-veratrylglycerone (GS-βVG) and β-deoxy-α-veratrylglycerone. Reagents and conditions: (*i*) diethyl carbonate, NaH, THF, reflux, 2 h, 88%; (*ii*) DIBAL-H, THF, 2 h, 91%; (*iii*) DDQ, 1,4-dioxane, 30 min, flash chromatography, 74%; (*iv*) pyridinium tribromide, EtOAc, 30 min, flash chromatography, 27%; (*v*) glutathione, $NaHCO_3$, 1:1 acetone/water, 14 h, 100%. See Supplementary Information for details.

Synthesis of CS-β(S)VP and CS-β(R)VP

Synthesis of 3,4-dimethoxy-cinnamaldehyde dimethyl acetal – Synthesis of the dimethyl acetal (Figure 2-4A) commenced with the condensation of triethyl phosphonoacetate with 3,4-dimethoxy-benzaldehyde, yielding ethyl 3,4-dimethoxy-cinnamate. Reduction of the ester yielded 3,4-dimethoxy-cinnamyl alcohol. Oxidation of the ring-conjugated alcohol afforded 3,4-dimethoxy-cinnamaldehyde. Protection of the carbonyl group then yielded 3,4-dimethoxy-cinnamaldehyde dimethyl acetal.

Synthesis of CS-β(S)VP and CS-β(R)VP – Syntheses of β(*S*)- and β(*R*)-isomers of γ,γ-dimethoxy-β-S-(methyl N-acetyl cysteinyl)-α-veratrylpropanone [CS-β(*S*)VP and CS-β(*R*)VP] were carried out in parallel (Figure 2-4B–C), and each thioether was derived from 3,4-dimethoxy-cinnamaldehyde dimethyl acetal. Then, in separate parallel reactions, the alkene was stereoselectively oxidized to diols γ,γ-dimethoxy-α-veratryl-α(*R*),β(*S*)-propanediol and γ,γ-dimethoxy-α-veratryl-α(*S*),β(*R*)-propanediol. Benzylic oxidation of the diols then yielded ketones γ,γ-dimethoxy-β(*R*)-hydroxy-α-veratrylpropanone and γ,γ-dimethoxy-β(*S*)-hydroxy-α-veratrylpropanone. Triflation of the β-hydroxyls afforded triflate esters γ,γ-dimethoxy-β(*R*)-O-triflyl-α-veratrylpropanone and γ,γ-dimethoxy-β(*S*)-O-triflyl-α-veratrylpropanone. Lastly, S_N2 displacement of the triflate groups by the thiolate ion of methyl N-acetyl cysteinate yielded β(*S*)- and β(*R*)-isomers of γ,γ-dimethoxy-β-S-(methyl N-acetyl cysteinyl)-α-veratrylpropanone [CS-β(*S*)VP and CS-β(*R*)VP] with high diastereometric purity. Because S_N1 displacement of the β-O-triflate synthetic intermediates would have yielded a mixture of both CS-βVP β-epimers in the synthesis schemes of both CS-β(*S*)VP (Figure 2-4B) and CS-β(*R*)VP (Figure 2-4C), we conclude that formation of the β-S-thioether linkages in CS-β(*S*)VP and CS-β(*R*)VP was governed by S_N2 inversion of the chiral center at β, affording highly diastereometrically pure products.

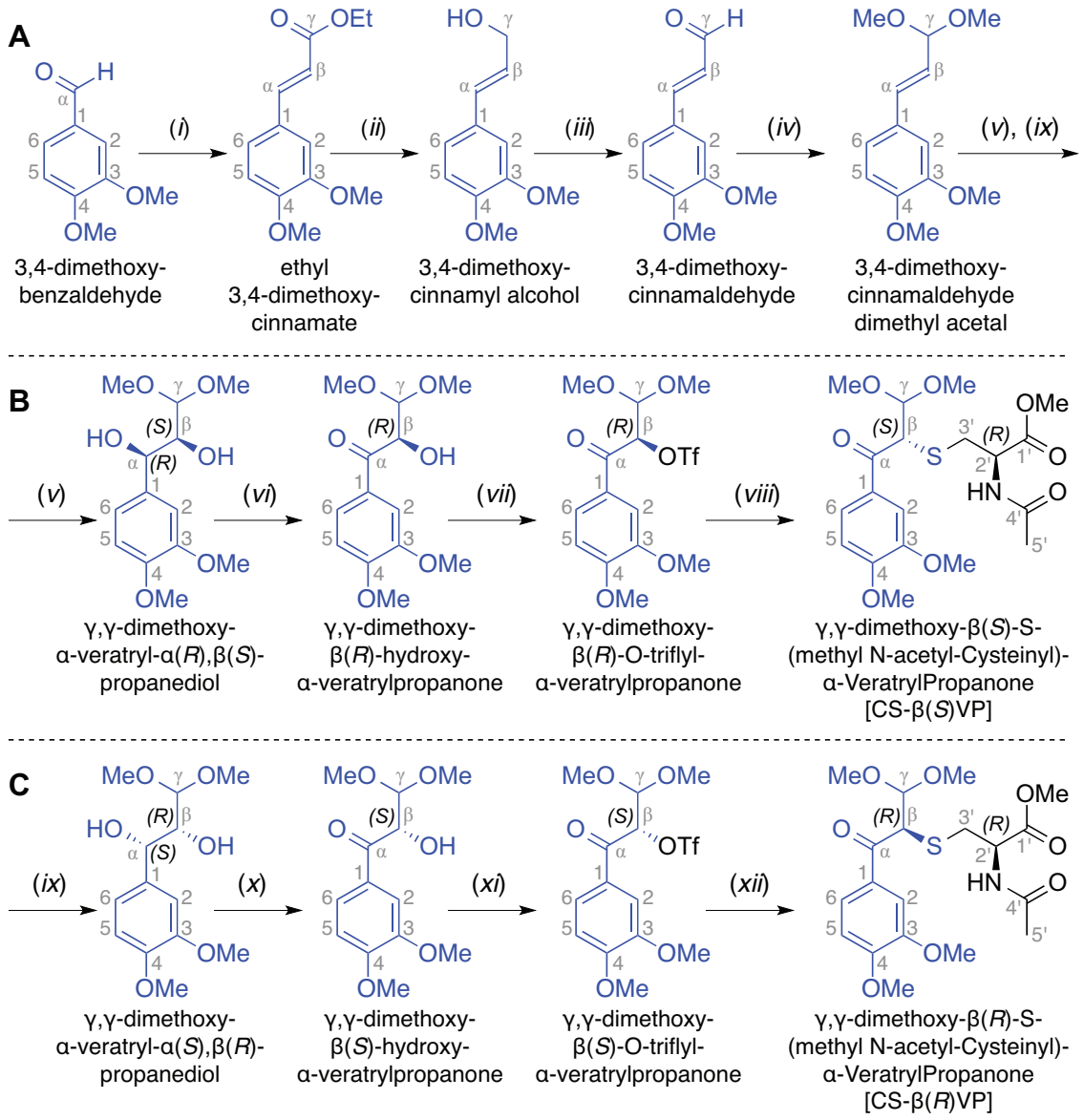

Figure 2-4. Scheme for the synthesis of **(A)** 3,4-dimethoxy-cinnamaldehyde dimethyl acetal, **(B)** γ,γ-dimethoxy-β(*S*)-S-(methyl N-acetyl-cysteinyl)-α-veratrylpropanone [CS-β(*S*)VP], and **(C)** γ,γ-dimethoxy-β(*R*)-S-(methyl N-acetyl-cysteinyl)-α-veratrylpropanone [CS-β(*R*)VP]. Reagents and conditions: (*i*) triethyl phosphonoacetate, NaH, THF, 2 h, 91%; (*ii*) DIBAL-H, THF, 2 h, 89%; (*iii*) DDQ, 1,4-dioxane, 30 min, flash chromatography, 72%; (*iv*) *p*-toluenesulfonic acid, trimethyl orthoformate, MeOH, 2 h, 96%; (*v*) AD-mix β, methanesulfonamide, 1:1 *t*-butanol/water, 4 °C, 18 h, 72%; (*vi*) DDQ, 1,4-dioxane, 30 min, flash chromatography, 73%; (*vii*) trifluoromethanesulfonic anhydride, 2,6-lutidine, CH_2Cl_2, 2 h, flash chromatography, 75%; (*viii*) methyl N-acetyl-(*R*)-cysteinate, K_2CO_3, dimethyl formamide, 2 h, 53%; (*ix*) AD-mix α, methanesulfonamide, 1:1 *t*-butanol/water, 4 °C, 18 h, 83%; (*x*) DDQ, 1,4-dioxane, 30 min, flash chromatography, 81%; (*xi*) trifluoromethanesulfonic anhydride, 2,6-lutidine, CH_2Cl_2, 2 h, flash chromatography, 59%; (*xii*) methyl N-acetyl-(*R*)-cysteinate, K_2CO_3, dimethyl formamide, 2 h, 73%. See Supplementary Information for details.

Chromatographic Techniques

General – All chromatographic separations were carried out using a Beckman 125NM solvent delivery module equipped with a Beckman 168 UV detector.

Preparative C_{18}-reversed phase chromatography – A pre-packed Biotage KP-C_{18} (100 g) reversed phase column was used for the purification of all enzymatically synthesized β-S-thioether compounds (GS-βVG and GS-βVE). A mixture of water and methanol was used for the mobile phase at a flow rate of 10 mL min^{-1}. The proportions of the total flow made up by each buffer were adjusted over a gradient: 0–15 min, 0% methanol; 15–20 min, gradient from 0-100% methanol; 20–35 min, 100% methanol; 35–40 min, gradient from 100-0% methanol; 40–50 min, 0% methanol. Fractions with UV absorption at 280 nm were collected, pooled, dried under a stream of nitrogen gas, and analyzed by ^{1}H, ^{13}C, COSY, HSQC, and HMBC NMR spectroscopy.

Analytical C_{18}-reversed phase chromatography – Parallel β-etherase reactions were carried out using either LigE, LigP, or LigF (1.0 mg mL^{-1}) with glutathione (2.0 mM) and βGVE (1.5 mM) as cosubstrates. Prior to addition of enzyme, 0.3 mL of the reaction mixture containing βGVE was collected for analysis by C_{18}-reversed phase chromatography. After 1 h of incubation with one of the β-etherases, 0.3 mL of the reaction mixture was also collected for HPLC analysis. The 0.3-mL samples from 0 h and 1 h reactions were then injected into an Ultra Aqueous (Restek Corporation, Bellefonte, PA) C_{18}-reversed stationary phase column (4 by 120 mm) for separation of GS-βVE, guaiacol, and βGVE, (Figure 2-5A). A mixture of water and methanol was used as the mobile phase at a flow rate of 1.0 mL min^{-1}. The methanol fraction of the total flow (with water as the remainder) was adjusted over a gradient as follows: 0–15 min, 0% methanol; 15–40 min, gradient from 0-100% methanol; 40–55 min, 100% methanol; 55–60 min, gradient from 100-0% methanol; 60–70 min,

0% methanol. GS-βVE, guaiacol, and substrate βGVE eluted from the C_{18} column after retention times (t_R) = 30.5, 39.6, and 46.3 min, respectively.

Preparative chiral chromatography – To separate chiral enantiomers β(*R*)GVG and β(*S*)GVG, crystalline racemic βGVG (2 mg, 6.0 nmol) was dissolved in ethanol (5 mL) and injected into a CHIRALPAK AY-H column (10 by 250 mm). A mixture of ethanol and hexane was used as the mobile phase at a flow rate of 2.0 mL $min.^{-1}$ The ethanol fraction of the total flow (with hexane as the remainder) was adjusted over a gradient as follows: 0–12 min, 50% ethanol; 12–20 min, gradient from 50-100% ethanol; 20–90 min, 100% ethanol; 90–95 min, gradient from 100-50% ethanol; 95–110 min, 0% ethanol. Fractions containing βGVG enantiomers were collected, pooled, and solvents were dried *in vacuo*. β(*S*)GVG or β(*R*)GVG eluted from the column with t_R = 21.8 and 80.6 min. The aforementioned procedure was repeated four additional times in order to collect approximately 5 mg of each enantiomer, to be used for enzymatic syntheses and isolation of GS-βVE, GS-β(*S*)VG and GS-β(*R*)VG.

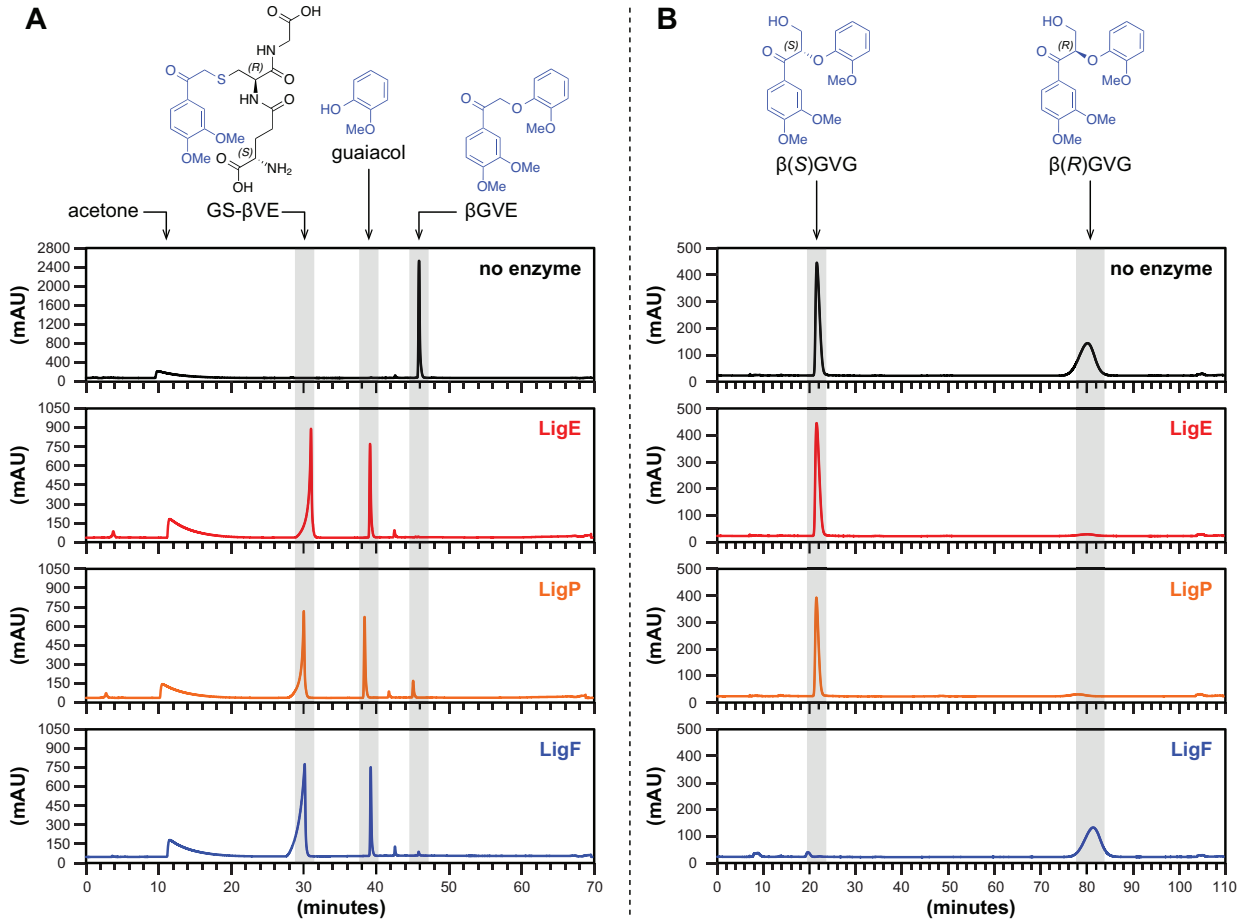

Figure 2-5. HPLC chromatogram traces of pre-enzyme addition (black), LigE (red), LigP (orange), and LigF (blue) enzyme assay samples. **(A)** C_{18}-reversed phase chromatography of βGVE prior to addition of and after 1 h incubation with GSH and either LigE, LigP, or LigF. **(B)** Chiral chromatography of *racem*-βGVG prior to addition of and after 1 h incubation with GSH and either LigE, LigP, or LigF.

Analytical chiral chromatography – Parallel β-etherase reactions were carried out with either LigE, LigP, or LigF (1.0 mg mL^{-1}) with glutathione (2.0 mM) and racemic βGVG (1.5 mM) as cosubstrates. Aliquots (0.5-mL) were collected prior to the addition of enzyme (0-h sample), and again after 1 h of incubation with either LigE, LigP, or LigF. The 0.5-mL samples from 0 h and 1 h were then each extracted six times with ethyl acetate and the solvent was subsequently dried *in vacuo*. Residues from the 0-h and 1-h samples were then dissolved in 0.1 mL ethanol and injected

into a Diacel Chemical Industries CHIRALPAK AY-H column (10 by 250 mm) for separation of guaiacol and chiral enantiomers β(*R*)GVG and β(*S*)GVG (Figure 2-5B). A mixture of ethanol and hexane was used as the mobile phase at a flow rate of 2.0 mL $min.^{-1}$ The ethanol fraction of the total flow (remainder was hexane) was adjusted over a gradient as follows: 0–12 min, 50% ethanol; 12–20 min, gradient from 50-100% ethanol; 20–90 min, 100% ethanol; 90–95 min, gradient from 100-50% ethanol; 95–110 min, 0% ethanol. Guaiacol, β(*S*)GVG, and β(*R*)GVG eluted from the chiral column with t_R = 8.0, 21.8, and 80.6 min, respectively.

Isolation of enzymatic reaction products

General – In vitro reaction assays for enzymatic synthesis of β-S-thioethers were conducted in an aqueous assay buffer [10 mM HEPES, 60 mM NaCl, 100 μM TCEP, 5% acetone, 2 mM GSH, pH 7.5]. GSH was added just prior to each assay to avoid disulfide formation, and NaOH was used to re-adjust the buffer to pH 7.5. All aqueous phase reaction products were desalted and purified by preparative C_{18}-reversed phase HPLC and fractions containing the reaction products were collected and dried over a stream of nitrogen gas. The ^{1}H, ^{13}C, COSY, HSQC, and HMBC NMR spectra of were analyzed for each of the isolated reaction products.

In vitro enzymatic synthesis of GS-β(S)VG from LigE or LigP – To 10 mL of the enzyme assay buffer, β(*R*)GVG (3.3 mg, 1.0 mM) was added and the buffer was separated into two 5-mL aliquots. To one 5-mL aliquot, LigE (1.0 mg, 6.2 μM) was added. LigP (1.0 mg, 6.5 μM) was added to the second 5-mL aliquot. Both reactions were incubated at room temperature for a period of 1 h. Guaiacol and trace amounts of β(*R*)GVG were removed from the reaction mixture by six successive ethyl acetate extractions. The LigE and LigP reaction products eluted from the Biotage KP-C18 column with t_R = 30.9 and t_R = 31.0 min, respectively.

In vitro enzymatic synthesis of GS-β(R)VG from LigF – Aside from the use of β(*S*)GVG as the LigF substrate, identical reaction conditions to those used for LigE and LigP were applied for LigF-

catalyzed synthesis of GS-β(*R*)VG from β(*S*)GVG. GS-β(*R*)VG eluted from the Biotage KP-C18 column with t_R = 31.1 min.

In vitro enzymatic synthesis of GS-βVE from LigE, LigP, and LigF – To 15 mL of the enzyme assay buffer, βGVE (4.5 mg, 1.0 mM) was added and the buffer was separated into three 5-mL aliquots. To each, either LigE (1.0 mg, 6.2 μM), LigP (1.0 mg, 6.5 μM), or LigF (1.0 mg, 6.7 μM) was added. The three reactions were incubated at room temperature for 1 h. After ethyl acetate extraction, the aqueous reaction products of LigE, LigP, and LigF eluted from the Biotage KP-C18 column after t_R = 31.0, 31.1, and 31.0 min, respectively.

In vitro enzymatic synthesis of β-deoxy-α-veratrylglycerone and α-veratrylethanone from LigG – To separate 5-mL aliquots of enzyme assay buffer, LigG (1.3 mg, 8.5 μM) and either chemically synthesized GS-βVG (5.0 mg, 1.0 mM each) or GS-βVE (2.4 mg, 1.0 mM) were added and the buffer was readjusted to pH 7.5 using NaOH. Reactions were incubated at room temperature for 1 h. Six ethyl acetate extractions were carried out and the organic fractions were pooled. Solvent was then removed by evaporation *in vacuo* and products (either β-deoxy-α-veratrylglycerone or α-veratrylethanone) were analyzed by NMR spectroscopy (^{1}H and ^{13}C NMR spectra can be found in the Supplementary Information). When GS-βVG was used as the substrate, the residual GS-β(*S*)VG in the aqueous fraction was purified further by preparative C_{18}-reversed phase HPLC, where GS-β(*S*)VG eluted from the Biotage KP-C18 column with t_R = 31.0 min.

Results

Characterization of the etherase reaction products using an achiral substrate

The GST family member LigE from strain SYK-6 liberates 4′-methyl-umbelliferone from the achiral substrate [β-O-7′-(4′-methyl)-umbelliferyl-α-guaiacylethanone] (Masai *et al.* 1989; Masai *et al.* 1991). This prompted us to use a synthetic achiral substrate (β-guaiacyl-α-veratrylethanone,

abbreviated βGVE, Figure 2-1C) in assays with the three proposed β-etherases from this strain. We also synthesized (Figure 2-2) a β-S-glutathione-conjugated model compound β-S-glutathionyl-α-veratrylethanone (abbreviated as GS-βVE) that is a potential βGVE cleavage product.

When we incubated βGVE with pure recombinant LigE, LigF or LigP, we found that, as expected from members of this GST superfamily, the substrate's degradation required the presence of GSH (Masai *et al.* 1993a). In the presence of GSH, degradation of βGVE by each enzyme (Figure 2-5A) was accompanied by production of guaiacol and an unidentified compound (later identified as GS-βVE). Thus, we conclude that LigE, LigP, and LigF were each active on the achiral βGVE as predicted (Figure 2-1C).

The unidentified product of βGVE degradation from each Lig enzyme was purified for analysis by NMR spectroscopy (Supplementary Information). The ^{1}H NMR spectrum of the product isolated from either LigE, LigP, or LigF reactions using βGVE and GSH (Figure 2-6B–D) was identical to that of chemically synthesized GS-βVE (Figure 2-6A). In addition, long-range 2D ^{1}H–^{13}C NMR (HMBC) correlations (a) between the β-protons and the 3′-carbon as well as (b) between carbon-β and each $3'H_a$ and $3'H_b$ proton in synthetic GS-βVE (Figure 2-7A) confirm that the GSH and aromatic moieties are linked via a β-S-3′-thioether. The task of identifying carbon-β and the β-protons was complicated by the β-protons' exchange with solvent (D_2O) deuterons. By analyzing the spectra of chemically synthesized GS-βVE in 90/10% H_2O/D_2O as the NMR solvent (with suppression of H_2O proton signals), we were able to unambiguously assign the β-carbon in the ^{13}C spectrum and the β-protons in the ^{1}H spectrum (See Supplementary Information). The HMBC spectrum in this solvent also revealed the above correlations between the β and 3′ nuclei, further establishing the existence of the β-S-3′-thioether linkage. Furthermore, both methoxyl moieties (3-OMe and 4-OMe) in GS-βVE could be unambiguously assigned (Figure 2-6), along with all aromatics, allowing GβVE β-ether bond cleavage by each GST family member to be fully elucidated. From these results we conclude

that LigE, LigF and LigP each catalyze GSH-dependent β-etherase activity with achiral substrate GβVE, and that β-etherase-catalyzed conversion of βGVE to GS-βVE (Figure 2-1C) links the β-carbon of the aryl moiety to GSH.

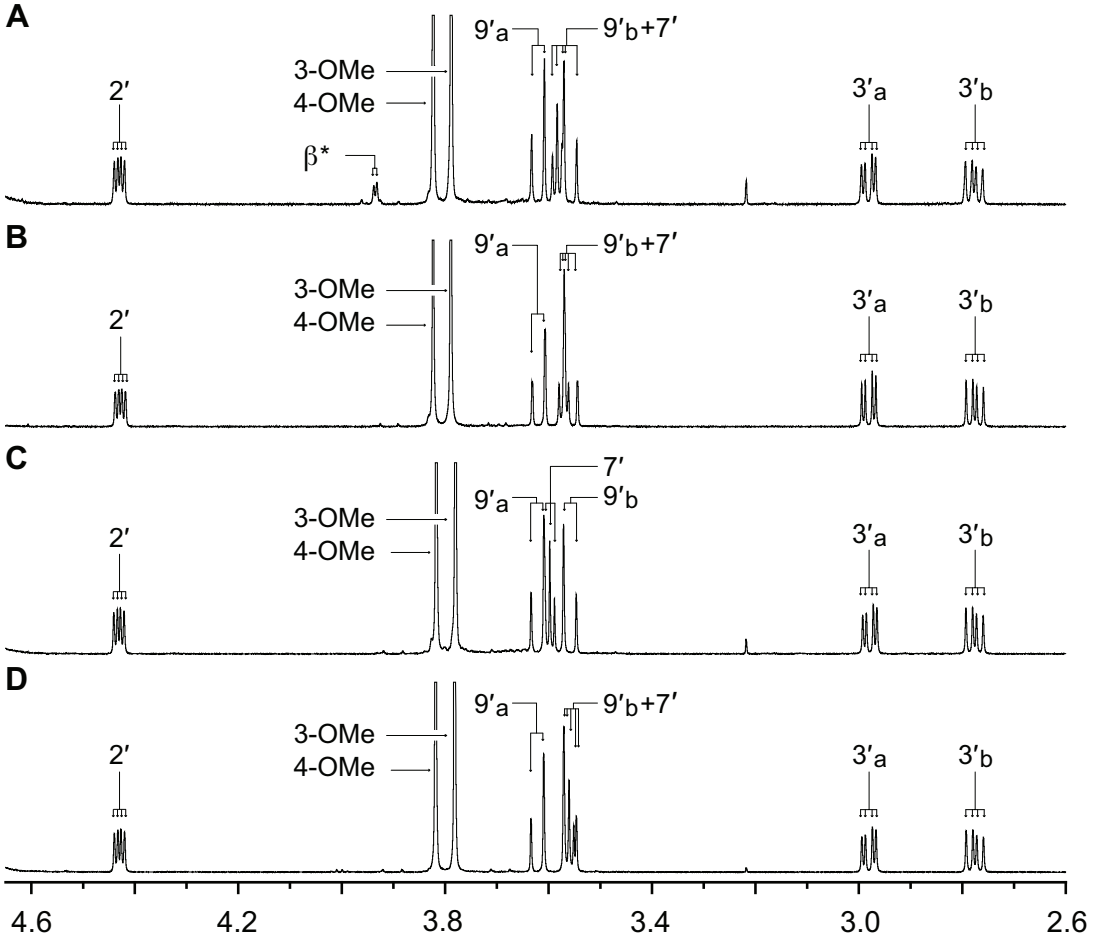

Figure 2-6. Aligned ^{1}H NMR partial spectra (2.60–4.65 ppm) of the GS-conjugate, β-S-glutathionyl-α-veratrylethanone (GS-βVE) in D_2O. Proton assignment labels correspond with the carbon to which the proton is bound. Alphabetical subscripts differentiate two non-identical geminal protons. Proton peaks that did not integrate as expected are denoted with an asterisk (*) – see text. **(A)** Compound GS-βVE obtained via chemical synthesis. Reaction product GS-βVE isolated from **(B)** LigE, **(C)** LigP, or **(D)** LigF activities using the achiral model compound GβVE as a substrate.

β-S-thioetherase activity of LigG with the product of β-etherase activity

The identification of GS-βVE as the product of β-etherase activity by the GST family members LigE, LigF and LigP, allowed us to test if this compound is a substrate for the putative GSH-dependent β-S-thioetherase, LigG. We incubated LigG from strain SYK-6 with GSH and chemically synthesized GS-βVE and observed GSH-dependent cleavage of GS-βVE and formation of the expected aromatic reaction product α-veratrylethanone (Masai *et al.* 2003) (Supplementary Information). We therefore conclude that LigG is a GSH-dependent β-S-thioetherase that cleaves the GSH-conjugate GS-βVE, the intermediate derived from β-etherase activity with achiral substrate βGVE (Figure 2-1C).

The use of the achiral substrate βGVE documents the combined activities of the GSH-dependent enzymes LigE, LigF, LigP, and LigG in a β-etherase pathway that releases two aromatic monomers from a dimeric model substrate (Figure 2-1C). In addition, we provide direct experimental evidence that LigG catalyzes thioether-cleavage at the β-S-glutathionyl linkage in the presence of GSH.

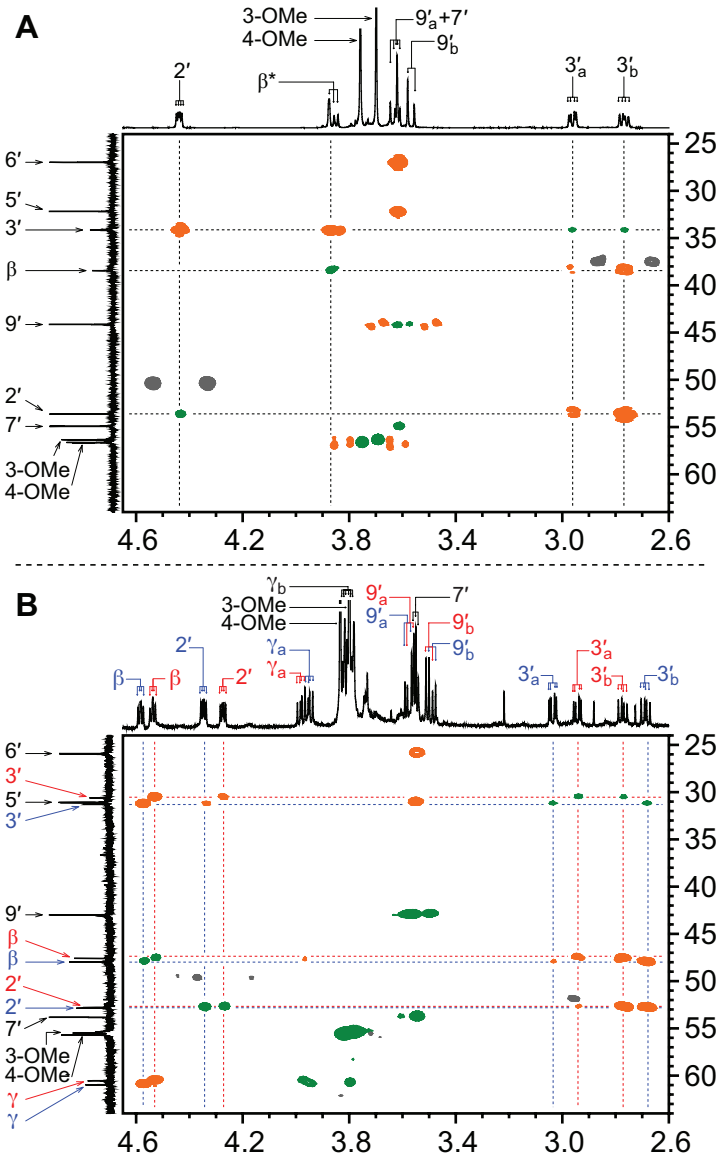

Figure 2-7. Partial ^{1}H-^{13}C 2-D HSQC (green) and HMBC (orange) NMR spectra of β-S-thioether-linked compounds in D_2O, where ^{1}H chemical shifts are plotted on the x-axis (2.60–4.65 ppm), ^{13}C chemical shifts are plotted on the y-axis (24.0–64.0 ppm), and non-^{1}H-^{13}C-correlating HMBC spectral regions are indicated (grey). Proton assignment labels correspond with the carbon to which the proton is bound. Alphabetical subscripts differentiate two non-identical geminal protons. **(A)** Chemically synthesized GS-βVE. Proton peaks that did not integrate as expected are denoted with an asterisk (*) – see text. **(B)** Chemically synthesized mixture of GS-β(*S*)VG (red labels) and GS-β(*R*)VG (blue labels). Overlapping GS-β(*S*)VG and GS-β(*R*)VG ^{1}H and ^{13}C spectral regions are indicated (black labels).

Stereoselectivity of the LigE, LigP, and LigF β-etherases

Given the racemic nature of β-ether units in the lignin backbone (Ralph *et al.* 1999; Akiyama *et al.* 2000), we analyzed activities of the β-etherases with chemically synthesized βGVG, a substrate, like its lignin counterpart, with a chiral center at the β-position and two enantiomeric configurations, β(*R*)GVG (Figure 2-1A) and β(*S*)GVG (Figure 2-1B). To aid these studies, we also synthesized (Figure 2-3) the GSH-conjugate, β-S-glutathionyl-α-veratrylglycerone (GS-βVG) that is the predicted product of β-etherase activity with βGVG. The existence of two β-epimers [GS-β(*S*)VG and GS-β(*R*)VG] in chemically synthesized GS-βVG was demonstrated by ^{1}H–^{1}H COSY, HSQC, and ^{1}H NMR spectral analyses (Figure 2-8A); the HMBC spectrum confirmed the existence of the β-S-3′-thioether linkage (Figure 2-7B).

To test for β-etherase activity, the racemic substrate βGVG was incubated with GSH and either LigE, LigP, or LigF. At the end of the assay, samples were extracted with ethyl acetate, partitioning guaiacol and residual βGVG enantiomers to the organic layer and β-epimers of GS-βVG to the aqueous layer. After evaporation of ethyl acetate and guaiacol *in vacuo*, residual organics were dissolved in ethanol and analyzed by chiral chromatography to determine if the enzymes cleaved one or both of the substrate enantiomers (Figure 2-5B). We found that LigF cleaved only the β(*S*)GVG enantiomer whereas LigE and LigP each cleaved only the β(*R*)GVG isomer. This confirmed that each of these enzymes is stereospecific for a single enantiomer (Masai *et al.* 2003; Hishiyama *et al.* 2012).

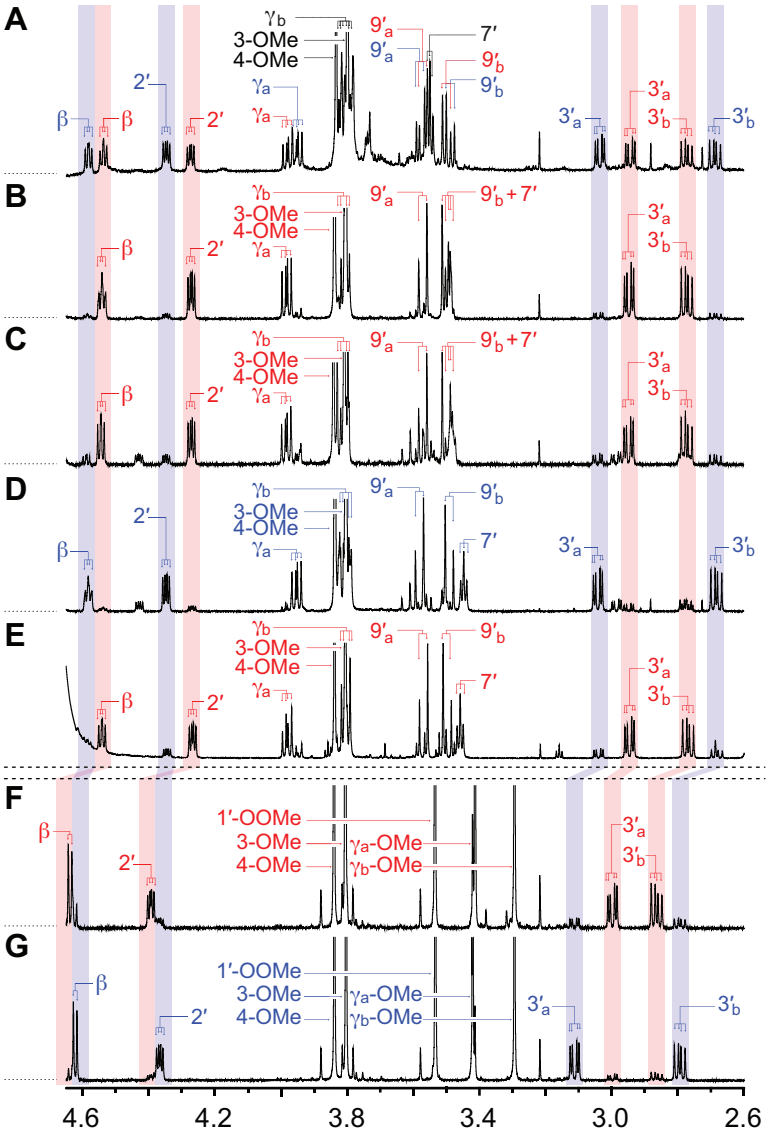

Figure 2-8. Aligned ^{1}H NMR partial spectra (2.60–4.65 ppm) of both β-epimers of β-S-glutathionyl-α-veratrylglycerone [GS-β(*S*)VG and GS-β(*R*)VG] and both β-epimers of γ,γ-dimethoxy-β-S-(methyl N-acetyl-cysteinyl)-α-veratrylpropanone [CS-β(*S*)VP and CS-β(*R*)VP] in D_2O. Proton assignment labels correspond with the carbon to which the proton is bound. Alphabetical subscripts differentiate two non-identical geminal protons or methoxyls. Red labels and shading are used for the two β(*S*)-configured compounds, GS-β(*S*)VG and CS-β(*S*)VP; blue labels and shading are used for the two β(*R*)-configured compounds, GS-β(*R*)VG and CS-β(*R*)VP; black labels are used for overlapped regions. Regions that differentiate β(*S*)- and β(*R*)-configurations (*i.e.*, protons at β, 2′, and 3′) are shaded across panels. **(A)** Chemically synthesized mixture (1:1) of GS-β(*S*)VG and GS-β(*R*)VG. **(B)** GS-β(*S*)VG from LigE. **(C)** GS-β(*S*)VG from LigP. **(D)** GS-β(*R*)VG from LigF. **(E)** residual GS-β(*S*)VG not degraded by LigG. **(F)** Chemically synthesized CS-β(*S*)VP. **(G)** Chemically synthesized CS-β(*R*)VP.

To further characterize the products of β-etherase activity, we analyzed their activities when either enantiopure β(*R*)GVG or β(*S*)GVG (purified by preparative chiral chromatography) was incubated with GSH and either LigE, LigP, or LigF. We found that LigE converted β(*R*)GVG to GS-β(*S*)VG with a 13:1 molar excess over the other diastereomer [GS-β(*R*)VG], as calculated from integration of the ^{1}H NMR spectral peaks (Figure 2-8B). LigP exhibited similar stereoselectivity as the product of β(*R*)GVG cleavage was a ~4:1:1 mixture of GS-β(*S*)VG, GS-β(*R*)VG, and an impurity likely to be a β,γ-unsaturated alkene (Figure 2-8C). In contrast, the product of LigF-catalyzed cleavage of β(*S*)GVG was a 1:7:2 mixture of the two β-epimers of GS-βVG [i.e., primarily GS-β(*R*)VG] and the above-mentioned alkene impurity (Figure 2-8D). We found that the abundance of the impurities and the level of the minor isomers in the β-etherase reaction products increased with time after isolation from enzymatic reactions, suggesting that non-enzymatic enolization and dehydration of GS-βVG are responsible for epimerization and side-product formation. Thus, we conclude that LigE and LigP each stereoselectively produces GS-β(*S*)VG from β(*R*)GVG whereas LigF stereoselectively yields GS-β(*R*)VG from β(*S*)GVG.

β-S-Thioetherase activity of LigG with GS-βVG

We also tested for the ability of LigG to catalyze GSH-dependent β-S-thioetherase activity in the presence of chemically synthesized GS-βVG [1:1 mixture of GS-β(*S*)VG and GS-β(*R*)VG]. We found that the β-S-thioether cleavage activity of LigG was insufficient for complete degradation of the synthetic GS-βVG. After chromatographic separation of the LigG assay substrates and reaction products, the ^{1}H NMR spectra of the residual GS-βVG showed that LigG had catalyzed β-S-thioether cleavage with high stereospecificity towards GS-β(*R*)VG compared to GS-β(*S*)VG. Integration of the ^{1}H NMR spectral peaks in the residual GS-βVG (Figure 2-8E) revealed that GS-β(*S*)VG was at a 5:1 molar excess over the other diastereomer, GS-β(*R*)VG. We also found that LigG-catalyzed degradation of the substrate resulted in the production of β-deoxy-α-veratrylglycerone

(Supplementary Information), the expected mono-aromatic product of GS-β(*R*)VG degradation (Figure 2-1B). These observations confirm that LigG is a stereospecific β-thioetherase that catalyzes GSH-dependent cleavage of GS-β(*R*)VG, which arises as an intermediate in the β-etherase pathway.

β-etherases cause inversion chiral carbon β

Although our data indicated that both GS-β(*S*)VG and GS-β(*R*)VG are intermediates in the β-etherase pathway (Figure 2-8A–E), the absolute configurations of the two GS-βVG β-epimers remained unknown. Thus, to assign absolute orientation of the chiral centers at position β in GS-β(*S*)VG and GS-β(*R*)VG, two closely-related models with known configurations at position β were synthesized (Figure 2-4). The parallel schemes used for β(*S*)- (Figure 2-4B) and β(*R*)-configured (Figure 2-4C) γ,γ-dimethoxy-β-S-(methyl N-acetyl cysteinyl)-α-veratrylpropanone [CS-β(*S*)VP and CS-β(*R*)VP] syntheses yielded high purity products with only trace amounts of the undesired β-epimer in each case, as indicated by the ^{1}H NMR spectra of CS-β(*S*)VP (Figure 2-8F) and CS-β(*R*)VP (Figure 2-8G). Analysis of the ^{1}H NMR spectra also revealed that the chemical shifts of protons at β, 2′, and 3′ were affected by the chiral configuration at carbon β, with a pronounced effect observed for the splitting of the two protons at carbon 3′.

The features of these synthetic compounds allowed us to compare them to the products of Lig etherase activity. The shifts of the 3′H_a and 3′H_b regions in CS-β(*S*)VP (Figure 2-8F) and CS-β(*R*)VP (Figure 2-8G) each aligned exclusively with one of the two β-epimers of GS-βVG (Figure 2-8A–E). Thus, we propose that the alignment of the CS-β(*S*)VP (Figure 2-8F) and GS-β(*S*)VG (Figure 2-8B–C, E) 3′H_a and 3′H_b spectral regions [and likewise of the CS-β(*R*)VP (Figure 2-8G) and GS-β(*R*)VG (Figure 2-8D) alignment] is attributable to the chiral configuration at carbon β. From these results, we conclude that β-etherase catalysis by LigE and LigP causes stereochemical inversion of chiral carbon β from a β(*R*)-substrate to product GS-β(*S*)VG (Figure 2-1A) whereas LigF carries out inversion of the β(*S*)-substrate chirality in forming product GS-β(*R*)VG (Figure 2-1B).

Discussion

The properties of GST family members and many other GSH-dependent enzymes have been well-studied due to their important roles in crucial cellular processes (Vuilleumier 1997; Sheehan *et al.* 2001; Hayes *et al.* 2005; Oakley 2005). However, much less in known about the role of GSH and the large number of bacterial GST proteins implicated in catabolic pathways. Our work provides several new insights into the properties of GST family members (LigE, LigF, and LigP) and a GSH-dependent thioetherase (LigG) in the degradation of oligomeric aromatic compounds. The results also provide direct support for the notion that the organism has evolved enzymes to independently deal with both the *R*- and *S*-configured centers in the racemic natural plant lignins (Ralph *et al.* 1999; Akiyama *et al.* 2000).

Proteins in the GST family typically use the GSH thiol to initiate a nucleophilic attack on the substrate. We showed that the GST family members, LigE, LigF, and LigP each produce a product in which the β-carbon of the substrate is covalently linked to the GSH thiol. We also provided the first experimental evidence that each of these three GST family members have stereospecific and stereo-selective β-etherase activity. Our data demonstrate that nucleophilic attack by the GSH thiol on the β-carbon of the substrate is responsible for the β-ether bond cleavage in βGVG and the release of an aromatic monomer (in our case guaiacol) and a second thiol-linked GSH-conjugated mono-aromatic product.

Although LigE, LigF and LigP are each active with β-ether-linked substrates, they are somewhat unusual GST family members as they are stereospecific for the configuration at the β-position of the substrate. We showed that with racemic β-aryl ether-linked model substrates, LigE and LigP each cleave only the β(*R*)-enantiomer and LigF cleaves only the β(*S*)-stereoisomer. We also found that LigE- and LigP-catalyzed β-etherase reactions exhibit stereoselectivity for the β(*S*)-configured GSH conjugate, whereas LigF yields the β(*R*)-diastereomer. Although the β-etherase catalytic reaction

mechanism remains unknown, our findings reveal that β-etherase catalysis causes S_N2-like inversion of the chiral configuration at carbon β, where β-ether cleavage and β-thioether formation are carried out on opposite faces of the molecule. In sum, our data indicate that these GST family members are both substrate-stereospecific and product-selective.

We also showed that LigG cleaves the GSH conjugates that are produced by β-etherase activity. LigG-mediated cleavage of these GSH conjugates requires addition of GSH, suggesting that thiol-mediated substrate reduction is needed for this β-S-thioetherase activity. Our data therefore provide direct experimental support for the contention that additional activities are not needed to release the other aromatic product of the β-etherase pathway. They also predict that the overall pathway uses GSH to derivatize and release one aromatic product (in a reaction catalyzed by LigE, LigF, or LigP) prior to reductive cleavage and release of the second aromatic (by LigG). Overall, this strategy is reminiscent of the chemical 'derivatization followed by reductive cleavage' (DFRC) method used to release aromatic monomers from lignin (Lu and Ralph 1997b; Lu and Ralph 1997a).

Our data show that β-etherase catalysis by the GST family members LigE, LigP, and LigF is achieved through stereospecific β-ether-bond cleavage, inversion of the chiral (*R*/*S*)-configuration at carbon β, and concomitant β-S-thioether bond formation. We have also unambiguously identified GSH conjugates as the previously uncharacterized glutathionyl-S-thioether-linked intermediates of the β-etherase pathway in *Sphingobium* sp. strain SYK-6. Further, given that (a) LigE/LigP and LigF stereoselectively produce GS-β(*S*)VG and GS-β(*R*)VG (respectively), (b) LigG converts the LigF β-etherase product GS-β(*R*)VG to β-deoxy-α-veratrylglycerone, and (c) LigG has little or no activity as a β-thioetherase with the LigE/LigP-produced GS-β(*S*)VG, we conclude that LigG is a stereospecific β-S-thioetherase that plays a role in the derivation of mono-aromatic compounds in the β-etherase pathway. The fate of the LigE/LigP-produced GSH-conjugated β-etherase pathway intermediate, GS-

β(*S*)VG, currently remains unknown. Assumedly, cells either use racemase-like enzymes for the conversion of the inactive β-epimer to the GSH-conjugate that is cleaved by LigG, possess a second β-S-thioetherase with the required stereospecificity for GSH-dependent cleavage of the β(*S*)-isomer, or employ other metabolic activities that enable the bacterium to utilize GS-β(*S*)VG as a growth substrate.

In sum, we have provided new insights into the enzymes, substrates and products of a novel catabolic β-etherase pathway (Masai *et al.* 1989; Masai *et al.* 1991; Masai *et al.* 1993a; Masai *et al.* 1993b; Hara *et al.* 2000; Masai *et al.* 2003; Masai *et al.* 2007) that is garnering considerable recent attention. We found a remarkably high degree of stereospecificity and stereoselectivity for these GSH-dependent enzymes. We propose that the existence of multiple GST family member β-etherases with complementary stereochemical features can be rationalized by the combinatorial radical chemistry that is used to synthesize plant lignins and the resulting racemic nature of the β-ether (and other) subunits of this polymer. The existence of β(*R*)- and β(*S*)-configurations in native plant lignins necessitates both β(*R*)- and β(*S*)-stereospecific β-etherases and β-thioetherases. Although these enzymes could cleave other structurally related compounds, the stereochemical features of the enzymes described support a hypothesis that they normally function in a pathway that processes racemic substrates similar to those found in nature. These defined properties of the GSH-dependent enzymes could help in producing valuable chiral products from individual stereoisomers of β-aryl substrates that might be derived from lignin degradation.

Abbreviations

GSH, glutathione; GS-SG, glutathione disulfide; βGVG, β-guaiacyl-α-veratrylglycerone; GS-βVG, β-S-glutathionyl-α-veratrylglycerone; βGVE, β-guaiacyl-α-veratrylethanone; GS-βVE, β-S-glutathionyl-α-veratrylethanone; CS-βVP, γ,γ-dimethoxy-β-S-(methyl N-acetyl cysteinyl)-α-veratrylpropanone; $NHis_8$, N-terminal octa-histidine; Ni-NTA, nickel-nitrilotriacetic acid; SDS, sodium

dodecyl sulfate; PAGE, polyacrylamide electrophoresis; TCEP, *tris*-(2-carboxyethyl)-phosphine hydrochloride; t_R, retention time; HSQC, (^{1}H–^{13}C) heteronuclear single-quantum coherence (NMR spectroscopy); HMBC, (^{1}H–^{13}C) heteronuclear multiple-bond correlation (NMR spectroscopy); COSY, (^{1}H–^{1}H) correlation spectroscopy.

Acknowledgments

This work was supported by the Department of Energy Office of Science's Great Lakes Bioenergy Research Center, Grant DE-FG02-07ER64495. Daniel Gall was supported by a traineeship from the NIGMS Biotechnology Training grant (Grant T32 GM08349). We thank Sally Ralph (US Forest Products Laboratory), and all members of the Ralph laboratory for aiding in synthesis, analysis, and characterization of model compounds, and Shannon Stahl for nomenclature advice. We also thank collaborators Blake Simmons and Richard Heins (the Joint BioEnergy Institute) for providing plasmid pJBEI_LigP (Supporting Information) containing *ligP* from *Sphingobium* sp. strain SYK-6, and for valuable discussion.

References

Adler E. (1977). Lignin chemistry - past, present and future. *Wood Science and Technology* 11(3):169-218.

Adler E., Eriksoo E. (1955). Guaiacylglycerol and its β-guaiacyl ether. *Acta chemica Scandinavica* 9341-342.

Akiyama T., Magara K., Matsumoto Y., Meshitsuka G., Ishizu A., Lundquist K. (2000). Proof of the presence of racemic forms of arylglycerol-β-aryl ether structure in lignin: studies on the stereo structure of lignin by ozonation. *Journal of Wood Science* 46(5):414-415.

Akiyama T., Sugimoto T., Matsumoto Y., Meshitsuka G. (2002). *Erythro*/*threo* ratio of β-O-4 structures as an important structural characteristic of lignin. I: Improvement of ozonation method for the quantitative analysis of lignin side-chain structure. *Journal of Wood Science* 48(3):210-215.

Bertani G. (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *Journal of bacteriology* 62(3):293-300.

Bertani G. (2004). Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. *Journal of Bacteriology* 186(3):595-600.

Blommel P.G., Fox B.G. (2007). A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expression and Purification* 55(1):53-68.

Bradford M.M. (1976). Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *Analytical Biochemistry* 72(1-2):248-254.

Bryksin A.V., Matsumura I. (2010). Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *Biotechniques* 48(6):463-465.

Bugg T.D.H., Ahmad M., Hardiman E.M., Singh R. (2011). The emerging role for bacteria in lignin degradation and bio-product formation. *Current Opinion in Biotechnology* 22(3):394-400.

Carrington J.C., Cary S.M., Dougherty W.G. (1988). Mutational analysis of tobacco etch virus polyprotein processing - *cis* and *trans* proteolytic activities of polyproteins containing the 49-kilodalton proteinase. *Journal of Virology* 62(7):2313-2320.

Dixon R.A., Paiva N.L. (1995). Stress-induced phenylpropanoid metabolism. *Plant Cell* 7(7):1085-1097.

Dougherty W.G., Carrington J.C., Cary S.M., Parks T.D. (1988). Biochemical and mutational analysis of a plant virus polyprotein cleavage site. *Embo Journal* 7(5):1281-1287.

Dougherty W.G., Cary S.M., Parks T.D. (1989). Molecular genetic-analysis of a plant-virus polyprotein cleavage site - a model. *Virology* 171(2):356-364.

Dougherty W.G., Parks T.D. (1991). Posttranslational processing of the tobacco etch virus 49-kda small nuclear inclusion polyprotein - identification of an internal cleavage site and delimitation of VPg and proteinase domains. *Virology* 183(2):449-456.

Freudenberg K. (1959). Biosynthesis and constitution of lignin. *Nature* 183(4669):1152-1155.

Fukuzumi T., Katayama Y. (1977). Bacterial degradation of dimer relating to structure of lignin I. β-hydroxypropiovanillone and coniferylalcohol as initial degradation products from

guaiacylglycerol-β-coniferylether by *Pseudomonas putida*. *Mokuzai Gakkaishi* 23(4):214-215.

Hara H., Masai E., Katayama Y., Fukuda M. (2000). The 4-oxalomesaconate hydratase gene, involved in the protocatechuate 4,5-cleavage pathway, is essential to vanillate and syringate degradation in *Sphingomonas paucimobilis* SYK-6. *Journal of Bacteriology* 182(24):6950-6957.

Hayes J.D., Flanagan J.U., Jowsey I.R. (2005). Glutathione transferases. *Annual Review of Pharmacology and Toxicology* 4551-88.

Higuchi T. (1980) Lignin structure and morphological distribution in plant cell walls. In: Kirk T.K., Higuchi T., Chang H. (eds) Lignin biodegradation: microbiology, chemistry and potential applications. CRC Press, Boca Raton, Florida, pp 1-20

Hishiyama S., Otsuka Y., Nakamura M., Ohara S., Kajita S., Masai E., Katayama Y. (2012). Convenient synthesis of chiral lignin model compounds via optical resolution: four stereoisomers of guaiacylglycerol-β-guaiacyl ether and both enantiomers of 3-hydroxy-1-(4-hydroxy-3-methoxyphenyl)-2-(2-methoxy-phenoxy)-propan-1-one (erone). *Tetrahedron Letters* 53842-845.

Horton R.M. (1993). *In vitro* recombination and mutagenesis of DNA : SOEing together tailor-made genes. *Methods in molecular biology (Clifton, N.J.)* 15251-261.

Horton R.M., Cai Z., Ho S.N., Pease L.R. (2013). Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *Biotechniques* 54(3):129-133.

Lewis N.G., Yamamoto E. (1990). Lignin - occurrence, biogenesis and biodegradation. *Annual Review of Plant Physiology and Plant Molecular Biology* 41455-496.

Lu F.C., Ralph J. (1997a). Derivatization followed by reductive cleavage (DFRC method), a new method for lignin analysis: Protocol for analysis of DFRC monomers. *Journal of Agricultural and Food Chemistry* 45(7):2590-2592.

Lu F.C., Ralph J. (1997b). DFRC method for lignin analysys. 1. New method for β-aryl ether cleavage: Lignin model studies. *Journal of Agricultural and Food Chemistry* 45(12):4655-4660.

Masai E., Ichimura A., Sato Y., Miyauchi K., Katayama Y., Fukuda M. (2003). Roles of the enantioselective glutathione S-transferases in cleavage of β-aryl ether. *Journal of Bacteriology* 185(6):1768-1775.

Masai E., Katayama Y., Fukuda M. (2007). Genetic and biochemical investigations on bacterial catabolic pathways for lignin-derived aromatic compounds. *Bioscience Biotechnology and Biochemistry* 71(1):1-15.

Masai E., Katayama Y., Kawai S., Nishikawa S., Yamasaki M., Morohoshi N. (1991). Cloning and sequencing of the gene a *Pseudomonas-paucimobilis* enzyme that cleaves β-aryl ether. *Journal of Bacteriology* 173(24):7950-7955.

Masai E., Katayama Y., Kubota S., Kawai S., Yamasaki M., Morohoshi N. (1993a). A bacterial enzyme degrading the model lignin compound β-etherase is a member of the glutathione-S-transferase superfamily. *Febs Letters* 323(1-2):135-140.

Masai E., Katayama Y., Nishikawa S., Yamasaki M., Morohoshi N., Haraguchi T. (1989). Detection and localization of a new enzyme catalyzing the β-aryl ether cleavage in the soil bacterium (*Pseudomonas-paucimobilis* SYK-6). *Febs Letters* 249(2):348-352.

Masai E., Kubota S., Katayama Y., Kawai S., Yamasaki M., Morohoshi N. (1993b). Characterization of the Cα-dehydrogenase gene involved in the cleavage of β-aryl ether by *Pseudomonas paucimobilis*. *Bioscience Biotechnology and Biochemistry* 57(10):1655-1659.

Moore D.D. (2003) Current protocols in molecular biology. In: Ausubel F.M., Brent R., Kingston R.E., Moore D.D., Seidman J.G., Smith J.A., Struhl K. (eds). John Wiley & Sons, New York, NY

Oakley A.J. (2005). Glutathione transferases: new functions. *Current Opinion in Structural Biology* 15(6):716-723.

Prochazkova K., Shuvalova L.A., Minasov G., Voburka Z., Anderson W.F., Satchell K.J.F. (2009). Structural and molecular mechanism for autoprocessing of MARTX toxin of *Vibrio cholerae* at multiple sites. *Journal of Biological Chemistry* 284(39):26557-26568.

Ralph J., Landucci L.L. (2010) Lignin and Lignans; Advances in Chemistry. In: Heitner C., Dimmel D.R., Schmidt J.A. (eds). CRC Press (Taylor & Francis Group), Boca Raton, FL, pp 137-234

Ralph J., Peng J.P., Lu F.C., Hatfield R.D., Helm R.F. (1999). Are lignins optically active? *Journal of Agricultural and Food Chemistry* 47(8):2991-2996.

Sarkanen K.V., Ludwig C.H. (1971) Lignins: Occurrence, formation, structure and reactions In: Sarkanen K.V., Ludwig C.H. (eds). John Wiley & Sons, Inc., New York, pp 1-916

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. *Applied and Environmental Microbiology* 75(16):5195-5201.

Sheahan K.-L., Cordero C.L., Satchell K.J.F. (2007). Autoprocessing of the *Vibrio cholerae* RTX toxin by the cysteine protease domain. *Embo Journal* 26(10):2552-2561.

Sheehan D., Meade G., Foley V.M., Dowd C.A. (2001). Structure, function and evolution of glutathione transferases: Implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochemical Journal* 3601-16.

Shen A., Lupardus P.J., Albrow V.E., Guzzetta A., Powers J.C., Garcia K.C., Bogyo M. (2009a). Mechanistic and structural insights into the proteolytic activation of *Vibrio cholerae* MARTX toxin. *Nature Chemical Biology* 5(7):469-478.

Shen A., Lupardus P.J., Morell M., Ponder E.L., Sadaghiani A.M., Garcia K.C., Bogyo M. (2009b). Simplified, enhanced protein purification using an inducible, autoprocessing enzyme tag. *Plos One* 4(12).

Shevchuk N.A., Bryksin A.V., Nusinovich Y.A., Cabello F.C., Sutherland M., Ladisch S. (2004). Construction of long DNA molecules using long PCR-based fusion of several fragments simultaneously. *Nucleic Acids Research* 32(2).

Simmons B.A., Logue D., Ralph J. (2010). Advances in modifying lignin for enhanced biofuel production. *Current Opinion in Plant Biology* 13(3):313-320.

Studier F.W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* 41(1):207-234.

Tanamura K., Kasai D., Nakamura M., Katayama Y., Fukuda M., Masai E. (2010). Identification of the third glutathione S-transferase gene involved in the stereospecific cleavage of β-aryl ether in *Sphingobium* sp. strain SYK-6. *Journal of Biotechnology* 150S235-S235.

Vuilleumier S. (1997). Bacterial glutathione S-transferases: What are they good for? *Journal of Bacteriology* 179(5):1431-1441.

Zakzeski J., Bruijnincx P.C.A., Jongerius A.L., Weckhuysen B.M. (2010). The catalytic valorization of lignin for the production of renewable chemicals. *Chemical Reviews* 110(6):3552-3599.

CHAPTER 3: A group of sequence-related sphingomonad enzymes catalyzes cleavage of β-aryl ether linkages in lignin β-guaiacyl and β-syringyl ether dimers

This chapter is published under the same title:

Gall D.L., Ralph J., Donohue T.J., Noguera D.R. (2014b). A group of sequence-related sphingomonad enzymes catalyzes cleavage of β-aryl ether linkages in lignin β-guaiacyl and β-syringyl ether dimers. *Environmental Science & Technology* 48(20):12454-12463.

doi: 10.1021/es503886d

Supporting Information associated with this publication:
http://pubs.acs.org/doi/suppl/10.1021/es503886d/suppl_file/es503886d_si_001.pdf

Daniel L. Gall performed all of the experiments and analyses in this chapter.

Abstract

Lignin biosynthesis occurs via radical coupling of guaiacyl and syringyl hydroxycinnamyl alcohol monomers (*i.e.*, "monolignols") through chemical condensation with the growing lignin polymer. With each chain-extension step, monolignols invariably couple at their β-positions, generating chiral centers. Here, we report on activities of bacterial glutathione-S-transferase (GST) enzymes that cleave β-aryl ether bonds in lignin dimers that are composed of different monomeric units. Our data reveal that these sequence-related enzymes from *Novosphingobium* sp. strain PP1Y, *Novosphingobium aromaticivorans* strain DSM12444, and *Sphingobium* sp. strain SYK-6 have conserved functions as β-etherases, catalyzing cleavage of each of the four dimeric α-keto-β-aryl ether-linked substrates (*i.e.*, guaiacyl-β-guaiacyl, guaiacyl-β-syringyl, syringyl-β-guaiacyl, and syringyl-β-syringyl). Although each β-etherase cleaves β-guaiacyl and β-syringyl substrates, we have found that each is stereospecific for a given β-enantiomer in a racemic substrate; LigE and LigP β-etherase homologs exhibited stereospecificity towards β(*R*)-enantiomers whereas LigF and its homologs exhibited β(*S*)-stereospecificity. Given the diversity of lignin's monomeric units and the racemic nature of lignin polymers, we propose that bacterial catabolic pathways have overcome

the existence of diverse lignin-derived substrates in nature by evolving multiple enzymes with broad substrate specificities. Thus, each bacterial β-etherase is able to cleave β-guaiacyl and β-syringyl ether-linked compounds while retaining either β(*R*)- or β(*S*)-stereospecificity.

Introduction

Lignin, a major component of plant cell walls, is a recalcitrant polymer composed of monomeric units (*i.e.*, components derived from guaiacyl and syringyl monomers) (Freudenberg 1959; Higuchi 1980; Lewis and Yamamoto 1990), providing plants with both pathogenic resistance and structural integrity (Sarkanen and Ludwig 1971; Dixon and Paiva 1995). The β-O-4′-ether (hereafter termed β-ether) is the most prevalent type of intermolecular bond through which the guaiacyl (monomethoxylated) and syringyl (dimethoxylated) aromatic units are linked (Adler 1977). Thus, the development of methodologies for β-ether cleavage and depolymerization of the lignin backbone may reveal novel aspects of catalysis and lead to lignin-derived products of high economic value (Chen and Dixon 2007; Simmons *et al.* 2010; Zakzeski *et al.* 2010; Bugg *et al.* 2011).

The formation of lignin polymers by radical coupling of monomeric units generates a racemic product containing both β(*R*)- and β(*S*)-ether bonds. Here, we report on enzyme activity with a set of newly analyzed substrates for a group of sequence-related bacterial β-etherases that are glutathione-S-transferase (GST) superfamily member enzymes, each of which catalyzes cleavage of β-ether bonds that are characteristically found in lignin polymers. Specifically, we reveal that each of these enzymes has activity with guaiacyl- and syringyl-containing substrates and that each enzyme exhibits stereospecifity for cleavage of either β(*R*)- or β(*S*)-ether-linked enantiomers.

The bacterium *Sphingobium* sp. strain SYK-6 possesses several metabolic enzymes that mediate metabolism of lignin-derived compounds (Masai *et al.* 2007). “Lig enzymes” that act in the proposed β-etherase pathway enable this organism to derive monoaromatic growth substrates from β-ether-linked α-keto diguaiacyl compounds such as α-(4-O-Me)-guaiacylglycerone-β-(1′-

formyl)-guaiacyl ether (GβG). The β(*R*)- and β(*S*)-enantiomers of GβG (Gβ(*R*)G and Gβ(*S*)G, Figure 3-1A) arise as β-etherase pathway intermediates from the activities of nicotinamide adenine dinucleotide (NAD)-dependent Lig dehydrogenases, which oxidize the corresponding benzylic alcohols to α-ketones (Masai *et al.* 1993b; Sato *et al.* 2009). It has been shown that, using glutathione (GSH) and GβG as cosubstrates, the β-etherases (LigE, LigP, and LigF1) cleave this aromatic dimer (Masai *et al.* 1989; Masai *et al.* 1991; Masai *et al.* 1993a; Masai *et al.* 2003), producing vanillin and a GSH-conjugated guaiacyl monomer (Gβ-SG) as reaction products (Gall *et al.* 2014). Gβ-SG is further degraded by LigG (and other enzymes that have not yet been identified), yielding glutathione disulfide (GSSG) and the mono-aromatic compound β-deoxy-α-(4-O-Me)-guaiacylglycerone (Figure 3-1A) (Masai *et al.* 2003).

The racemic nature of the lignin backbone (Ralph *et al.* 1999; Akiyama *et al.* 2000; Akiyama *et al.* 2002; Sugimoto *et al.* 2002) and the existence of both β(*R*)- and β(*S*)- configurations in lignin necessitate the ability to degrade both Gβ(*R*)G and Gβ(*S*)G enantiomers (Figure 3-1A). In *Sphingobium* sp. strain SYK-6, this is accomplished via the activities of multiple β-etherases with complementary stereochemical properties (Masai *et al.* 2003; Sato *et al.* 2009). *Sphingobium* sp. strain SYK-6 LigE and LigP catalyze stereospecific cleavage of Gβ(*R*)G and LigF1 exhibits stereospecificity for the Gβ(*S*)G enantiomer. In this organism, β-ether cleavage is coupled to GSH-conjugation, inversion of β-chirality, and stereoselective formation of Gβ(*S*)-SG (LigE and LigP) and Gβ(*R*)-SG (LigF1) (Gall *et al.* 2014).

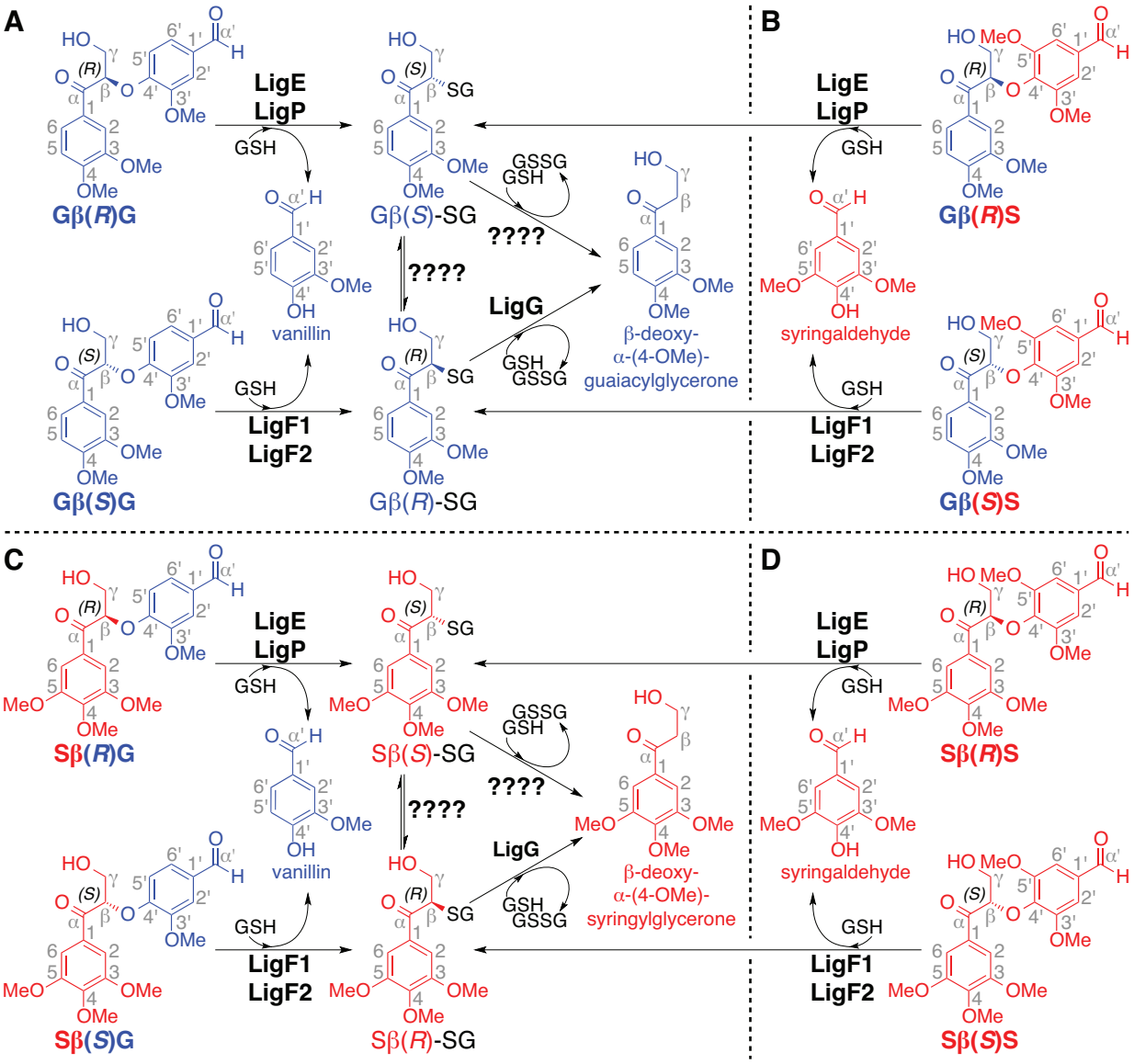

Figure 3-1. β-Etherase pathway-mediated conversion of β-enantiomers of substrates GβG, GβS, SβG, and SβS, in which vanillin and syringaldehyde are formed from cleavage of β-guaiacyl (in panels **A** and **C**) and β-syringyl (in panels **B** and **D**) ether-linked compounds. Compound names are displayed below each structure and 3-methoxylated (*i.e.*, guaiacyl) and 3,5-dimethoxylated (*i.e.*, syringyl) units are shown in blue and red. Catabolism of **(A)** Gβ(*R*)G and Gβ(*S*)G, as well as **(B)** Gβ(*R*)S and Gβ(*S*)S, yields aromatic monomers Gβ(*S*)-SG, Gβ(*R*)-SG, and β-deoxy-α-(4-O-Me)-guaiacylglycerone as metabolic intermediates. Catabolism of **(C)** Sβ(*R*)G and Sβ(*S*)G, as well as **(D)** Sβ(*R*)S and Sβ(*S*)S, yields aromatic monomers Sβ(*S*)-SG, Sβ(*R*)-SG, and β-deoxy-α-(4-O-Me)-syringylglycerone as metabolic intermediates.

The existence of guaiacyl and syringyl units in the lignin polymers of all land plants other than softwoods also necessitates the existence of enzymes that will cleave β-ether linked units of different subunit composition (*i.e.*, guaiacyl-β-guaiacyl (GβG), guaiacyl-β-syringyl (GβS), syringyl-β-guaiacyl (SβG), and syringyl-β-syringyl (SβS), Figure 3-1). Although the activities of *Sphingobium* sp. strain SYK-6 β-etherases have been shown to contribute to the stereospecific and stereoselective degradation of model compounds containing guaiacyl units, such as Gβ(*R*)G and Gβ(*S*)G (Sonoki *et al.* 2002; Masai *et al.* 2003; Gall *et al.* 2014; Picart *et al.* 2014), the role served by Lig enzymes in the catabolism of native lignin-derived compounds is largely unknown because (a) investigation of enzymes that might be involved in this pathway has been limited to those encoded in the genome of *Sphingobium* sp. strain SYK-6, and (b) the activities of β-etherase pathway enzymes have not been tested with the range of β-ether-containing oligomers composed of guaiacyl and syringyl subunits that are typically found in lignin (Figure 3-1B–D).

In this work, we reveal the ability of β-etherases from *Sphingobium* sp. strain SYK-6 (SsLigE, SsLigP, and SsLigF1) to cleave model dimeric lignin compounds containing GβS, SβG and SβS β-ether linkages, in addition to the previously reported GβG substrate (Masai *et al.* 2003; Gall *et al.* 2014). Further, we identify several additional sequence-related proteins with β-etherase activity from *Novosphingobium aromaticivorans* DSM12444 (NaLigE, NaLigF1, and NaLigF2) and *Novosphingobium* sp. strain PP1Y (NsLigE). We demonstrate that each enzyme catalyzes cleavage of all four combinations of β-ether-linked substrates, GβG (Figure 3-1A), GβS (Figure 3-1B), SβG (Figure 3-1C), and SβS (Figure 3-1D), where each LigE/LigP β-etherase homolog has the conserved function of degrading β(*R*)-enantiomers whereas each LigF1/LigF2 β-etherase homolog exhibits stereospecificity for the β(*S*)-enantiomers. Thus, we show that several bacteria possess β-etherases that have a previously unreported ability to cleave lignin dimers containing GβG, GβS, SβG and SβS β-ether linkages. Our results also reveal that each of these enzymes exhibits similar stereospecifity to that previously described for the enzymes from *Sphingobium* sp. strain SYK-6 (Masai *et al.* 2003;

Tanamura *et al.* 2010; Gall *et al.* 2014). These observations reveal important features of a conserved class of bacterial enzymes that have utility in the conversion of lignin during either plant biomass processing or the potential production of valuable compounds from this abundant polymer.

Experimental

Gene cloning and enzyme purification

DNA manipulation and transformation into *Escherichia coli* were carried out according to standard methods(Moore 2003) and as previously described (see Supporting Information for details) (Gall *et al.* 2014). DNA primers and restriction enzymes were obtained from Integrated DNA Technologies (Coralville, IA) and New England Biolabs (Ipswich, MA). Plasmids containing genes encoding SsLigE (locus tag SLG_08660), SsLigP (SLG_32600), and SsLigF1 (SLG_08650) from *Sphingobium* sp. strain SYK-6, the gene encoding potential Lig enzyme NsLigE (PP1Y_AT11664) from *Novosphingobium* sp. strain PP1Y, and the *Vibrio cholarae* rtxA gene (Vch1786_I0951) were obtained from Invitrogen (Carlsbad, CA) and were codon-optimized for expression in *E. coli.* Genes encoding potential Lig enzymes NaLigE (Saro_2405), NaLigF1 (Saro_2091), NaLigF2 (Saro_2865), and RpHypGST (RPA4340) were cloned from genomic DNA from *Novosphingobium aromaticivorans* strain DSM12444 or *Rhodopseudomonas palustris* strain CGA009, respectively.

Protein expression and purification – Each N-terminal (encoded on vector pVP302K) and C-terminal (encoded on vector pVP202K) octa-histidine affinity tagged ($NHis_8$ and $CHis_8$, respectively) enzyme was purified using nickel-nitrilotriacetic acid resin (Ni-NTA) affinity chromatography. $NHis_8$ tags were cleaved using Tev protease(Blommel and Fox 2007) and $CHis_8$ tags were cleaved by induction of the fused *Vibrio cholarae* RtxA protease (Sheahan *et al.* 2007; Prochazkova *et al.* 2009; Shen *et al.* 2009). A second round of Ni-NTA affinity chromatography removed cleaved tags from enzyme preparations that were subsequently purified by size-exclusion chromatography (Gall *et al.* 2014), and evaluated by SDS-PAGE (Figure 3-2).

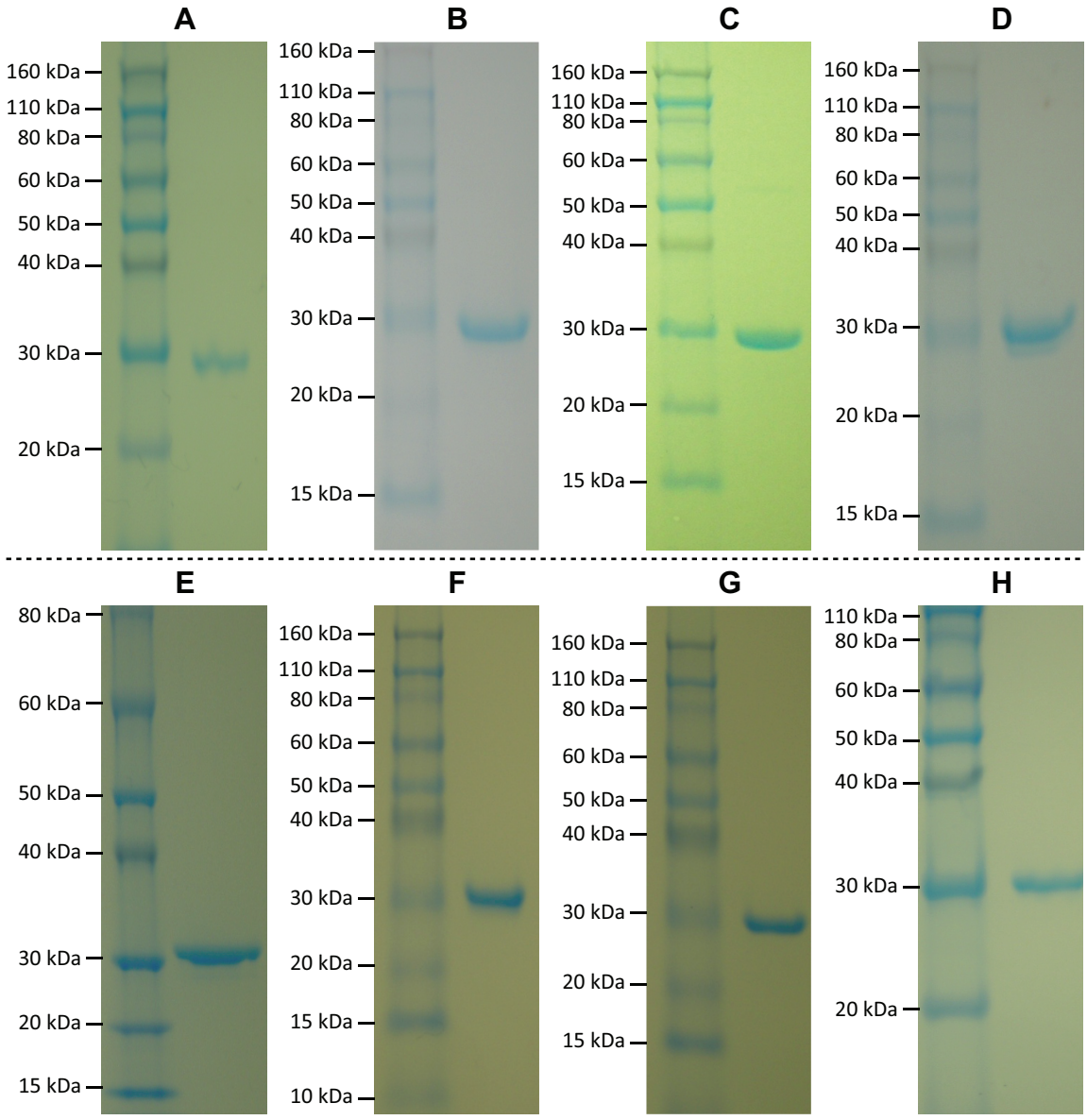

Figure 3-2. Images of SDS-12% PAGE gels loaded with enzyme preparations **(A)** 28.6-kDa RpHypGST, **(B)** 31.1-kDa NaLigE, **(C)** 30.8-kDa NsLigE, **(D)** 32.1-kDa SsLigE, **(E)** 31.0-kDa SsLigP, **(F)** 28.9-kDa NaLigF1, **(G)** 29.3-kDa NaLigF2, and **(H)** 30.0-kDa SsLigF.

NMR spectroscopy

^{1}H and ^{13}C NMR spectra were recorded on a Bruker Biospin (Billerica, MA) AVANCE 700 MHz spectrometer fitted with a cryogenically-cooled 5-mm TXI gradient probe with inverse geometry (proton coils closest to the sample). See Supporting Information for additional details.

Synthesis of β-ether-linked dimeric model compounds

Syntheses of β-brominated intermediates – β-bromination of commercially available α-(4-O-Me)-guaiacylethanone produced crystalline β-bromo-α-(4-O-Me)-guaiacylethanone (for additional details, see Supporting Information and Figure 3-3). Similarly, commercially available α-(4-O-Me)-syringylethanone was brominated, yielding crystalline β-bromo-α-(4-O-Me)-syringylethanone.

Syntheses of achiral β-ether-linked intermediates – Four achiral β-ether-linked compounds were synthesized using the above β-bromides as starting materials. The phenolate ion of vanillin was used for S_N2 displacement of the β-bromo-α-(4-O-Me)-guaiacylethanone bromide, yielding α-(4-O-Me)-guaiacylethanone-β-(1′-formyl)-guaiacyl ether (Figure 3-3A). Similarly, β-bromo-α-(4-O-Me)-guaiacylethanone and syringaldehyde were used to synthesize α-(4-O-Me)-guaiacylethanone-β-(1′-formyl)-syringyl ether (Figure 3-3B). The vanillin phenolate ion was used to displace the β-bromo-α-(4-O-Me)-syringylethanone bromide, yielding α-(4-O-Me)-syringylethanone-β-(1′-formyl)-guaiacyl ether (Figure 3-3C). Using β-bromo-α-(4-O-Me)-syringylethanone and syringaldehyde as starting materials under similar reaction conditions yielded α-(4-O-Me)-syringylethanone-β-(1′-formyl)-syringyl ether (Figure 3-3D).

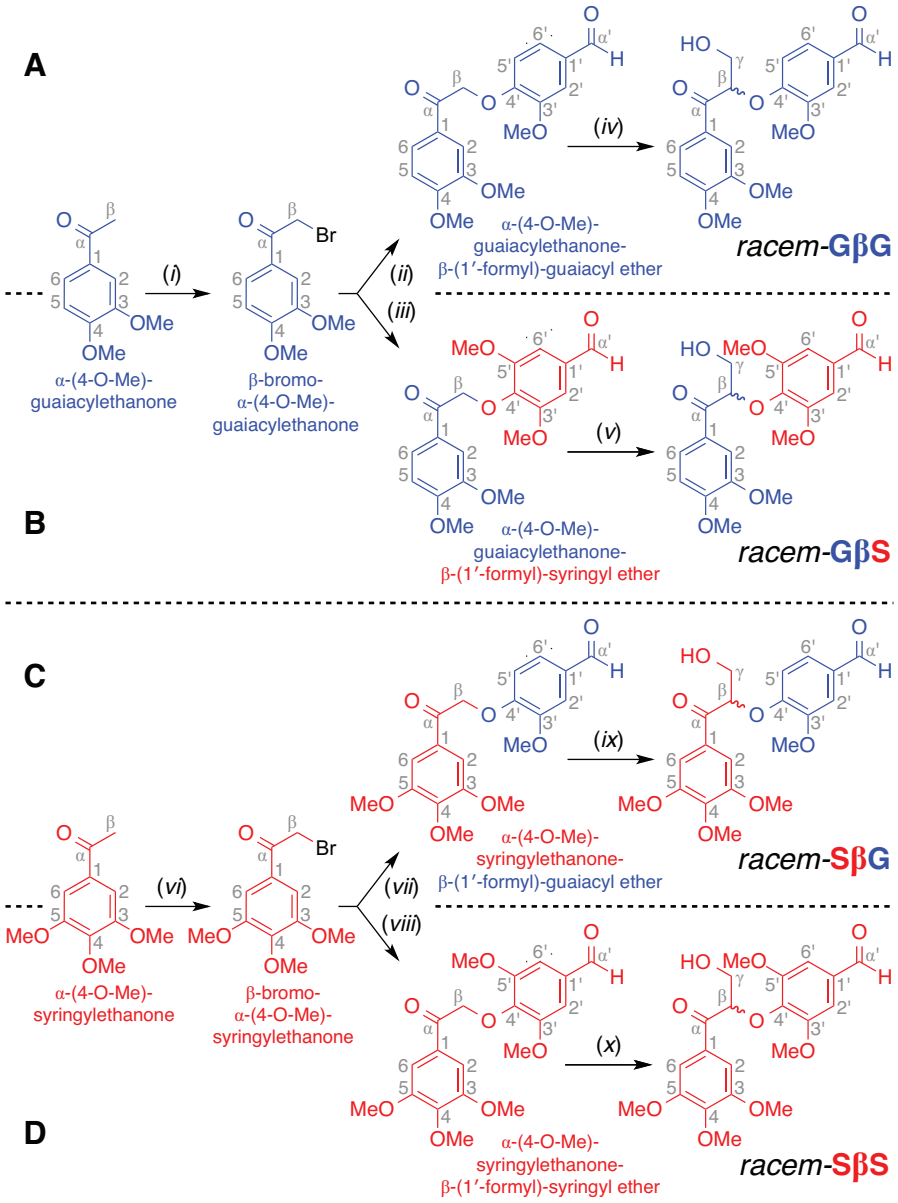

Figure 3-3. Synthetic schemes for the preparation of β-etherase substrates **(A)** GβG, **(B)** GβS, **(C)** SβG, and **(D)** SβS. Reagents and conditions: (*i*) pyridinium tribromide, EtOAc, 30 min, 59%; (*ii*) vanillin, K_2CO_3, acetone, 1 h, 87%; (*iii*) syringaldehyde, K_2CO_3, acetone, 1 h, 84%; (*iv*) formaldehyde, K_2CO_3, 1,4-dioxane, 3 h, 79%; (*v*) formaldehyde, K_2CO_3, 1,4-dioxane, 3 h, 88%; (*vi*) pyridinium tribromide, EtOAc, 30 min, 66%; (*vii*) vanillin, K_2CO_3, acetone, 1 h, 92%; (*viii*) syringaldehyde, K_2CO_3, acetone, 1 h, 90%; (*ix*) formaldehyde, K_2CO_3, 1,4-dioxane, 3 h, 82%; (*x*) formaldehyde, K_2CO_3, 1,4-dioxane, 3 h, 78%.

Syntheses of racemic β-etherase substrates – In parallel, each of the four racemic β-ether-linked intermediates were condensed with formaldehyde (Adler and Eriksoo 1955; Landucci *et al.* 1981), yielding racemic β-aryl ether-linked dimeric model compounds that served as substrates for β-etherase enzyme assays. Accordingly, α-(4-O-Me)-guaiacylglycerone-β-(1′-formyl)-guaiacyl ether (GβG) was derived from α-(4-O-Me)-guaiacylethanone-β-(1′-formyl)-guaiacyl ether (Figure 3-3A), α-(4-O-Me)-guaiacylglycerone-β-(1′-formyl)-syringyl ether (GβS) was derived from α-(4-O-Me)-guaiacylethanone-β-(1′-formyl)-syringyl ether (Figure 3-3B), α-(4-O-Me)-syringylglycerone-β-(1′-formyl)-guaiacyl ether (SβG) was derived from α-(4-O-Me)-syringylethanone-β-(1′-formyl)-guaiacyl ether (Figure 3-3C), and α-(4-O-Me)-syringylglycerone-β-(1′-formyl)-syringyl ether (SβG) was derived from α-(4-O-Me)-syringylethanone-β-(1′-formyl)-syringyl ether (Figure 3-3D). Additional details on the syntheses can be found in the Supporting Information.

β-Etherase enzyme assays

Parallel 5-mL β-etherase reactions were conducted (assay buffer: 10 mM HEPES, 60 mM NaCl, 100 μM TCEP, 5% acetone, 2 mM GSH, pH 7.5) in which individual proteins NaLigE, NsLigE, SsLigE, SsLigP, NaLigF1, NaLigF2, or SsLigF1 (0.25 mg mL^{-1}) were individually incubated with GSH and one of the racemic β-ethers GβG, GβS, SβG, or SβS (1.0 mM) as co-substrates. Aliquots (2.5-mL) were collected prior to protein addition (0-h sample), and again after 1 h of incubation with each of the putative β-etherases. Each 2.5-mL sample was extracted six times with ethyl acetate, partitioning residual β-ether-linked enantiomers and aromatic aldehydes to the organic phase and glutathione-conjugated products, Gβ-SG and Sβ-SG, to the aqueous layer. Ethyl acetate was then dried *in vacuo*, yielding residues containing the hydrophobic reaction products and residual substrate enantiomers. Residues from each sample were dissolved in 0.1 mL ethanol and analyzed by chiral chromatography.

Chiral chromatography

Analytical separation of GβG and GβS enantiomers – Analyses of GβG- and GβS-derived β-etherase reaction products and residual substrates were conducted via chiral chromatographic separation using a Diacel Chemical Industries CHIRALPAK AD-H column (4.6 by 250 mm). A mobile phase of 3/2 hexane/ethanol was used at a flow rate of 1.0 mL $min.^{-1}$ Vanillin, Gβ(*S*)G, and Gβ(*R*)G were detected in enzymatic reaction samples when *racem*-GβG was used as the substrate, with each eluting after t_R = 4.8, 16.6, and 20.2 min, respectively. Absolute configurations of GβG enantiomers were determined previously (Hishiyama *et al.* 2012). Syringaldehyde, Gβ(*S*)S, and Gβ(*R*)S were detected in reaction samples when *racem*-GβS was used as the substrate, eluting after t_R = 6.3, 16.0, and 18.1 min.

Analytical separation of SβG and SβS enantiomers – Analyses of SβG- and SβS-derived β-etherase reaction products and residual substrate enantiomers were conducted via chiral chromatographic separation using a Diacel Chemical Industries CHIRALPAK AY-H column (10 by 250 mm). A mobile phase of 1/1 hexane/ethanol was used at a flow rate of 2.5 mL $min.^{-1}$ Vanillin, Sβ(*S*)G, and Sβ(*R*)G were detected in reaction samples when *racem*-SβG was used as the substrate, eluting after t_R = 6.9, 16.7, and 19.5 min, respectively. Syringaldehyde, Sβ(*R*)S, and Sβ(*S*)S were detected in reaction samples when *racem*-SβS was used as the substrate, eluting after t_R = 8.0, 18.4, and 24.2 min, respectively.

Results

Identification of a conserved class of putative β-etherases

Given what is known about the β-etherase pathway in *Sphingobium* sp. strain SYK-6 (Masai *et al.* 2003; Sato *et al.* 2009; Gall *et al.* 2014), we sought to investigate whether or not this pathway could be utilized for β-ether catabolism by other sphingomonads (bacteria from genera:

Novosphingobium, *Sphingobium*, *Sphingomonas*, and *Sphingopyxis*) (Takeuchi *et al.* 2001), organisms that are often associated with the biodegradation of aromatic compounds in the environment (Xia *et al.* 2005; LaRoe *et al.* 2010; Notomista *et al.* 2011). At the onset of this study, BLASTP searches (Altschul *et al.* 1990) querying the amino acid sequences of Lig enzymes from *Sphingobium* sp. strain SYK-6 (Masai *et al.* 2012) revealed the existence of genes for putative LigE and LigF enzymes in two additional organisms for which full genome sequences were available: *Novosphingobium* sp. strain PP1Y and *N. aromaticivorans* strain DSM12444 (D'Argenio *et al.* 2011). Further, two homologs of each enzyme (LigE/LigP and LigF1/LigF2, respectively) were identified in strain SYK-6. We also found that both *Novosphingobium* strains encoded homologs of the NAD-dependent dehydrogenases that are essential for forming the α-ketones that undergo β-ether cleavage in strain SYK-6. In addition to the characterized β-etherase Lig enzymes (SsLigE, SsLigP, and SsLigF1), we expressed and purified homologous proteins encoded in the genomes of *Novosphingobium* sp. strain PP1Y (NsLigE) and *N. aromaticivorans* strain DSM12444 (NaLigE, NaLigF1, and NaLigF2). In sum, seven Lig homologs were tested for β-etherase activity with substrates GβG, GβS, SβG, and SβS; amino acid similarity to SsLigE or SsLigF1 is given in parentheses: SsLigE (100%), SsLigP (62%), NsLigE (78%), NaLigE (61%), SsLigF1 (100%), NaLigF1 (60%), and NaLigF2 (40%).

Cleavage of GβG aromatic β-ethers

Previously, it was reported that SsLigE, SsLigP, and SsLigF1 catalyzed stereospecific cleavage of a racemic diguaiacyl β-ether-linked substrate having a similar structure to that of GβG (Masai *et al.* 2003), but which contained an α′-H (rather than an α′-aldehyde) and a 4-OH (rather than a 4-OMe) (Figure 3-1A). We hypothesized that the Lig β-etherase homologs from each strain would catalyze the same reactions. In addition, we thought that it was likely, based on previously published data with diguaiacyl substrates (Masai *et al.* 2003; Gall *et al.* 2014), that neither the α′-aldehyde, nor the

4-O-Me moieties of GβG would inhibit β-etherase activity. Also, if each LigE/LigP enzyme had activity similar to that of the enzymes from *Sphingobium* sp. strain SYK-6, then they would each catalyze stereospecific degradation of Gβ(*R*)G whereas the Gβ(*S*)G enantiomer would be cleaved stereospecifically by the LigF1/LigF2 homologues. To test these predictions, we synthesized *racem*-GβG (Figure 3-3A) and performed β-etherase assays with each of the seven recombinant putative Lig enzymes using GSH and *racem*-GβG as cosubstrates. In comparing the chiral chromatogram of a sample containing substrates without protein (Figure 3-4A) with those representing materials from enzymatic assays that had been incubated with a homolog of LigE (Figure 3-4B–E), we found that, in each case, the LigE homologs released the expected product vanillin (t_R = 4.8 min) and degraded the high-t_R enantiomer (20.2 min) of GβG. Conversely, we found that each LigF homolog (Figure 3-4F–H) yielded vanillin as a reaction product while degrading the low-t_R enantiomer (16.6 min). Given that previous work has shown that SsLigE and SsLigP, when incubated with substrate analogs of GβG, show β(*R*)-stereospecificity whereas SsLigF1 exhibits β(*S*)-stereospecificity (Masai *et al.* 2003; Hishiyama *et al.* 2012), we propose that the high-t_R compound degraded by the LigE homologs (Figure 3-4B–E) was Gβ(*R*)G and the low-t_R compound degraded by the LigF homologs (Figure 3-4F–H) was Gβ(*S*)G. Further, analysis of the aqueous layers from these reactions (Gall *et al.* 2014) confirmed that SsLigE and SsLigP catalyzed formation of Gβ(*S*)-SG whereas Gβ(*R*)-SG was formed as a product of SsLigF-catalyzed reactions, demonstrating that β-etherase catalysis involves formation of β-thioether compounds (Figure 3-1A). We therefore conclude that LigE homologs have the conserved function of β(*R*)-etherase activity with substrate GβG, and similarly, that LigF homologs in the sphingomonads each have the conserved function of catalyzing β(*S*)-ether cleavage.

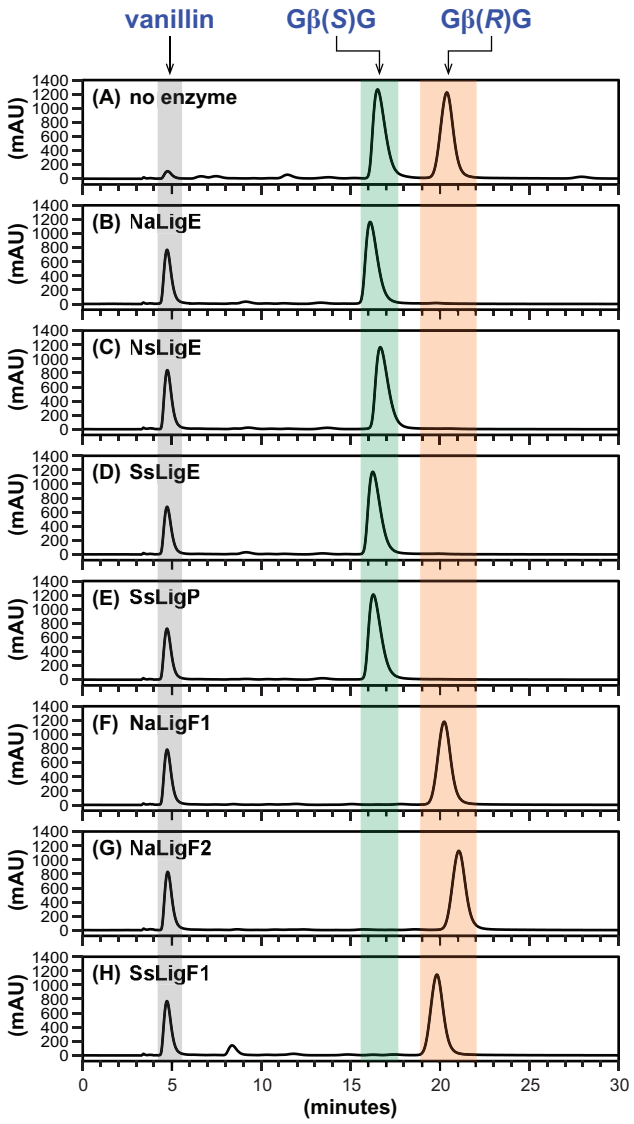

Figure 3-4. HPLC chromatographic traces (CHIRALPAK AD-H column, λ = 280 nm) of β-etherase enzyme assay samples from cosubstrates *racem*-GβG and glutathione. Chromatographic regions for vanillin (grey), Gβ(*S*)G (green), and Gβ(*R*)G (orange) peak elution times are highlighted by shading. **(A)** No enzyme added, 0 h sample, where the ratio of peak area integrals of Gβ(*S*)G to Gβ(*R*)G was ~1:1. After 1 h incubation with either enzymatic catalyst: **(B)** NaLigE, **(C)** NsLigE, **(D)** SsLigE, **(E)** SsLigP, **(F)** NaLigF1, **(G)** NaLigF2, or **(H)** SsLigF1. Structures of vanillin, Gβ(*S*)G, and Gβ(*R*)G are shown in Figure 3-1A. Abbreviations: Na, *Novosphingobium aromaticivorans* strain DSM12444; Ns, *Novosphingobium* sp. strain PP1Y; Ss, *Sphingobium* sp. strain SYK-6. See Experimental Procedures for details.

Cleavage of GβS aromatic β-ethers

To date, diguaiacyl compounds have been the only type of β-ether-linked lignin compound tested as a substrate of the *Sphingobium* sp. SYK-6 β-etherase pathway enzymes (Masai *et al.* 2003; Tanamura *et al.* 2010; Hishiyama *et al.* 2012). We hypothesized that β-ether cleavage would occur with substrates containing additional methoxy groups on the aromatic rings, *i.e.*, syringyl units. To test this hypothesis, we conducted additional β-etherase assays with each putative Lig enzyme using model β-ether compounds that contained either one or two syringyl units (Figure 3-1B–D) as substrates. The resulting data from assays in which GSH and *racem*-GβS (Figure 3-1B) were used as cosubstrates revealed that each β-etherase homolog catalyzed cleavage. Chiral chromatography of the reaction samples (Figure 3-5) indicated that each of the seven putative Lig enzymes produced the expected product (Figure 3-1B) syringaldehyde (t_R = 6.3 min), with each LigE homolog (Figure 3-5B–E) degrading only the high-t_R enantiomer (18.1 min) and each LigF homolog (Figure 3-5F–H) degrading the low-t_R enantiomer (16.0 min). We propose that the GβG cleavage stereospecificity exhibited by each enzyme is also observed with the degradation of the GβS enantiomers. From this, we conclude that the LigE/LigP homologs exhibited β(*R*)-etherase activity, degrading Gβ(*R*)S (t_R = 18.1 min), whereas each LigF catalyzed β(*S*)-ether cleavage of Gβ(*S*)S (t_R = 16.0 min).

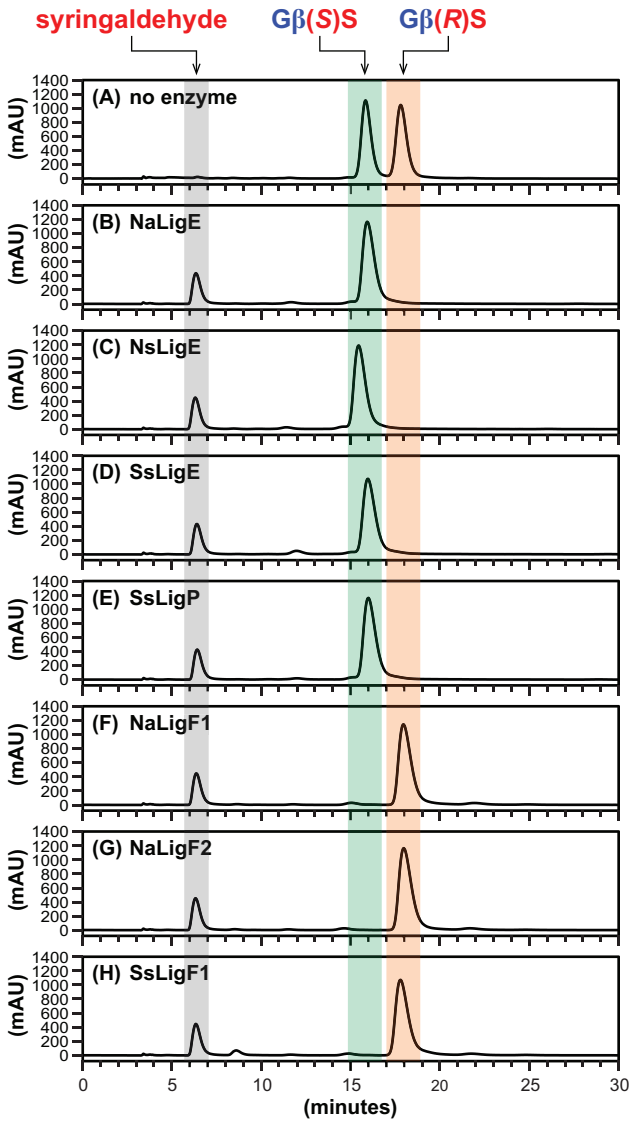

Figure 3-5. HPLC chromatographic traces (CHIRALPAK AD-H column, λ = 280 nm) of β-etherase enzyme assay samples from cosubstrates *racem*-GβS and glutathione. Chromatographic regions for syringaldehyde (grey), Gβ(*S*)S (green), and Gβ(*R*)S (orange) peak elution times are highlighted by shading. **(A)** No enzyme added, 0 h sample, where the ratio of peak area integrals of Gβ(*S*)S to Gβ(*R*)S was ~1:1. After 1 h incubation with either enzymatic catalyst: **(B)** NaLigE, **(C)** NsLigE, **(D)** SsLigE, **(E)** SsLigP, **(F)** NaLigF1, **(G)** NaLigF2, or **(H)** SsLigF1. Structures of syringaldehyde, Gβ(*S*)S, and Gβ(*R*)S are shown in Figure 3-1B. Abbreviations: Na, *Novosphingobium aromaticivorans* strain DSM12444; Ns, *Novosphingobium* sp. strain PP1Y; Ss, *Sphingobium* sp. strain SYK-6. See Experimental Procedures for details.

Cleavage of SβG aromatic β-ethers

To test whether Lig β-etherases catalyze cleavage of the geometric isomer containing its syringyl and guaiacyl units in the opposite bonding orientation of GβS, we assayed β-etherase activity of each putative Lig enzyme using GSH and *racem*-SβG as cosubstrates and found that each catalyzed cleavage. An alignment of the chiral chromatograms (Figure 3-6) reveals that each of the seven enzymes cleaved a single SβG enantiomer, yielding the expected product (Figure 3-1C), vanillin (t_R = 6.9 min). We also found that each putative LigE/LigP enzyme (Figure 3-6B–E) catalyzed stereospecific cleavage of Sβ(*R*)G (t_R = 19.5 min), whereas the LigF homologs (Figure 3-6F–H) exhibited stereospecificity towards Sβ(*S*)G (t_R = 16.7 min).

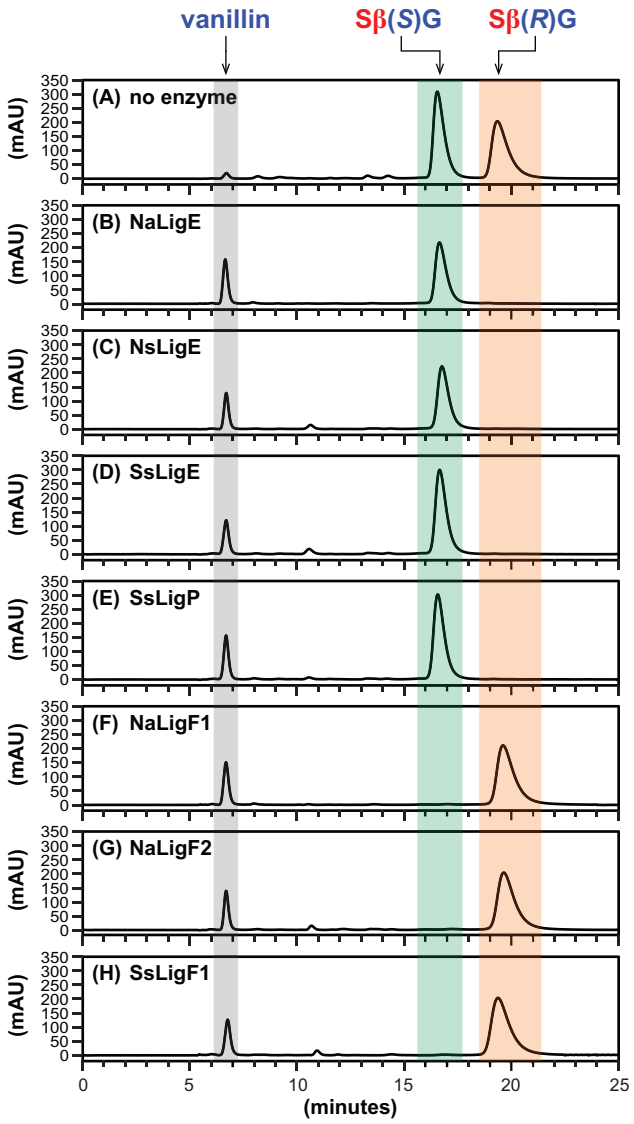

Figure 3-6. HPLC chromatographic traces (CHIRALPAK AY-H column, λ = 280 nm) of β-etherase enzyme assay samples from cosubstrates *racem*-SβG and glutathione. Chromatographic regions for vanillin (grey), Sβ(*S*)G (green), and Sβ(*R*)G (orange) peak elution times are highlighted by shading. **(A)** No enzyme added, 0 h sample, where the ratio of peak area integrals of Sβ(*S*)G to Sβ(*R*)G was ~1:1. After 1 h incubation with either enzymatic catalyst: **(B)** NaLigE, **(C)** NsLigE, **(D)** SsLigE, **(E)** SsLigP, **(F)** NaLigF1, **(G)** NaLigF2, or **(H)** SsLigF1. Structures of vanillin, Sβ(*S*)G, and Sβ(*R*)G are shown in Figure 3-1C. Abbreviations: Na, *Novosphingobium aromaticivorans* strain DSM12444; Ns, *Novosphingobium* sp. strain PP1Y; Ss, *Sphingobium* sp. strain SYK-6. See Experimental Procedures for details.

Cleavage of SβS aromatic β-ethers

To test for activity with a lignin compound composed of two syringyl units, we assayed for β-etherase activity with each putative Lig enzyme using GSH and *racem*-SβS as cosubstrates. Chiral chromatography (Figure 3-7) revealed that each enzyme degraded a single SβS enantiomer, resulting in the release of the expected product (Figure 3-1D), syringaldehyde (t_R = 8.0 min). In contrast with chromatogram alignments from GβG (Figure 3-4), GβS (Figure 3-5), and SβG assay samples (Figure 3-6), where LigE/LigP homologs degraded the high-t_R isomer and LigF homologs cleaved the low-t_R enantiomer, we found that the putative LigE/LigP enzymes (Figure 3-7B–E) catalyzed degradation of the low-t_R SβS isomer (t_R = 18.4 min) and the LigF homologs (Figure 3-7F–H) cleaved the high-t_R SβS isomer (t_R = 24.2 min). Because this result was in contrast to our findings with substrates GβG (Figure 3-4), GβS (Figure 3-5), and SβG (Figure 3-6), preparative chiral chromatography was used for the isolation of each isomer (for additional details, see Supporting Information and Figure 3-8) and the resulting enantiopure compounds were used to derive MTPA(*R*) esters that aided in the assignment of absolute configurations to the low-t_R (Sβ(*R*)S) and high-t_R (Sβ(*S*)S) isomers by ^{1}H NMR spectroscopy (for additional details, see Supporting Information and Figure 3-9). As was the case with racemic substrates GβG, GβS, and SβG, we again conclude that each LigE/LigP homolog exhibits β(*R*)-stereospecificity whereas each LigF homolog catalyzes β(*S*)-ether cleavage of SβS enantiomers, and that the isomers simply elute in reverse order in this case.

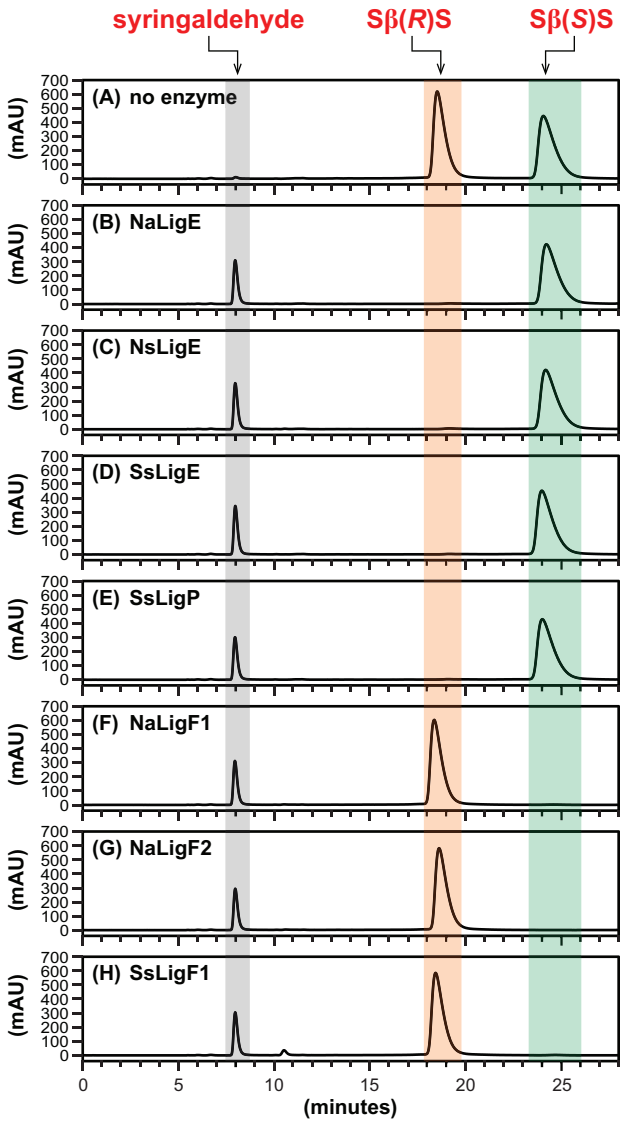

Figure 3-7. HPLC chromatographic traces (CHIRALPAK AY-H column, λ = 280 nm) of β-etherase enzyme assay samples from cosubstrates *racem*-SβS and glutathione. Chromatographic regions for syringaldehyde (grey), Sβ(*R*)S (orange), and Sβ(*S*)S (green). peak elution times are highlighted by shading. **(A)** No enzyme added, 0 h sample, where the ratio of peak area integrals of Sβ(*R*)S to Sβ(*S*)S was ~1:1. After 1 h incubation with either enzymatic catalyst: **(B)** NaLigE, **(C)** NsLigE, **(D)** SsLigE, **(E)** SsLigP, **(F)** NaLigF1, **(G)** NaLigF2, or **(H)** SsLigF1. Structures of syringaldehyde, Sβ(*R*)S, and Sβ(*S*)S are shown in Figure 3-1D. Abbreviations: Na, *Novosphingobium aromaticivorans* strain DSM12444; Ns, *Novosphingobium* sp. strain PP1Y; Ss, *Sphingobium* sp. strain SYK-6. See Experimental Procedures for details.

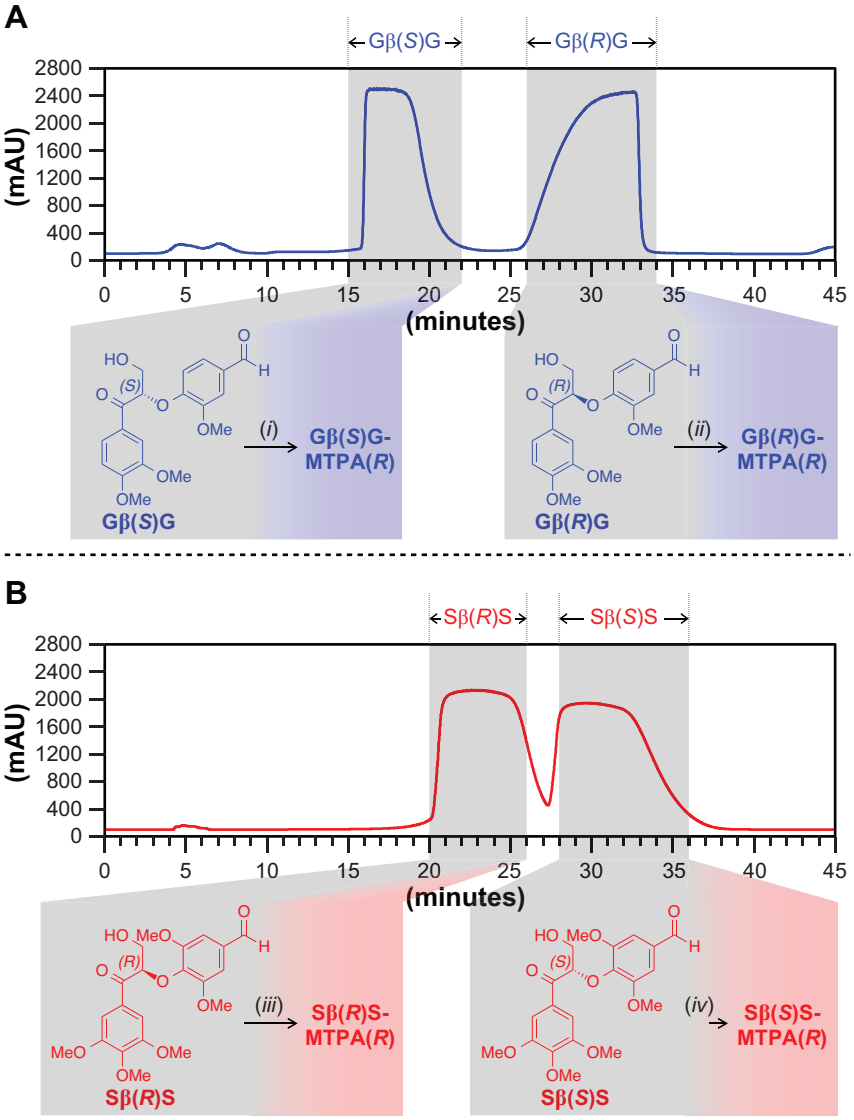

Figure 3-8. Preparative chiral HPLC chromatographic separations (CHIRALPAK AY-H column, λ = 280 nm) of **(A)** *racem*-GβG starting material, yielding Gβ(*S*)G (15-22 min) and Gβ(*R*)G (26-34 min) and **(B)** *racem*-SβS starting material, yielding Sβ(*R*)S (20-26 min) and Sβ(*R*)S (28-36 min). Reagents and conditions: (*i-iv*) diisopropylethylamine, dimethylaminopyridine, MTPACl(*S*), dichloromethane, 5 min, flash chromatography. Reaction yields: (*i*) 38%, (*ii*) 44%, (*iii*) 65%, (*iv*) 59%. Product MTPA(*R*) esters: (*i*) Gβ(*S*)G-MTPA(*R*); (*ii*) Gβ(*R*)G-MTPA(*R*); (*iii*) Sβ(*R*)S-MTPA(*R*); and (*iv*) Sβ(*S*)S-MTPA(*R*). Chemical structures of MTPA(*R*) esters are shown in Figure 3-9.

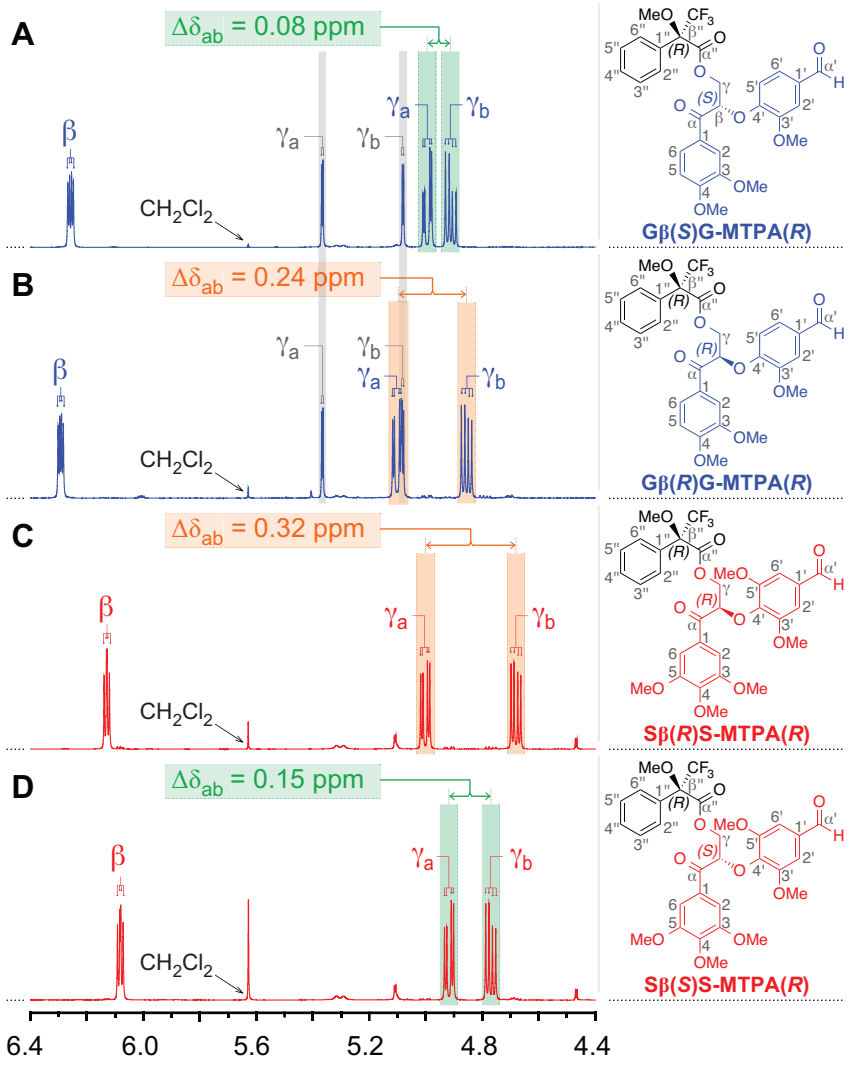

Figure 3-9. Aligned ^{1}H NMR spectra (4.40–6.40 ppm) of β-(1′-formyl)-guaiacyl-α-(4-O-Me)-guaiacylglyceryl α(*R*)-methoxy-trifluoromethyl-phenyl-acetate (GβG-MTPA(*R*) esters; shown in blue) and β-(1′-formyl)-syringyl-α-(4-O-Me)-syringylglyceryl α(*R*)-methoxy-trifluoromethyl-phenyl-acetate (SβS-MTPA(*R*) esters; shown in red). Proton assignment labels correspond with the carbon to which the proton is bound. Alphabetical subscripts differentiate two non-identical geminal protons. γH$_a$ and γH$_b$ proton spectral regions are highlighted by shading for β(*R*)-configured (orange), β(*S*)-configured (green), and achiral GβG-propenone (grey). **(A)** Gβ(*S*)G-MTPA(*R*), **(B)** Gβ(*R*)G-MTPA(*R*), **(C)** Sβ(*R*)S-MTPA(*R*), and **(D)** Sβ(*S*)S-MTPA(*R*).

β-Etherase assays with RpHypGST

In seeking to identify other potential β-etherases with the ability to cleave lignin model substrates, we constructed a phylogenetic tree from an alignment of closely related LigE/LigP and LigF homologs (Figure 3-10). *Rhodopseudomonas palustris* strain CGA009 is a bacterium previously shown to metabolize aromatic monomers likely to be derived from native lignin (Harwood and Gibson 1988). Thus, we cloned an *R. palustris* gene that encodes a hypothetical Lig β-etherase (RpHypGST, having 36% amino acid sequence similarity to SsLigE) and purified recombinant protein to be tested for activity in the same β-etherase assays. Recombinant RpHypGST was expressed and purified as either N-terminally tagged (affording N-RpHypGST) and C-terminally tagged (affording C-RpHypGST) His_8 fusions (see Experimental Procedures). N-RpHypGST and C-RpHypGST were each assayed using GSH and *racem*-GβG as cosubstrates. Under conditions identical to those where the sphingomonad Lig β-etherases exhibited β-etherase activity, neither substrate degradation nor release of the expected product (vanillin, Figure 3-1A) was detected (data not shown). We conclude that neither recombinant RpHypGST protein is a catalyst of β-etherase activity with these substrates.

Although the putative GSH binding domain was conserved across all sequences in both the LigE cluster and the HypGST cluster A (Figure 3-10), sequence analysis revealed several dissimilarities between the two clades. We found that each of the amino acid sequences in the HypGST cluster A, including RpHypGST, were significantly shorter (230 amino acids) than those in the LigE cluster (264-280 amino acids), which includes the four LigE homologs with confirmed β-etherase activity. Further, the putative GST superfamily substrate binding domains (residues 95-132) were conserved within a clade but dissimilar across the two clusters (Ji *et al.* 1994; Chen *et al.* 2003), suggesting that they bind different substrates. While this work was being reviewed, another member of HypGST cluster A encoded in *Sorangium cellulosum* (Figure 3-10) was also reported to

be inactive with β-ether-linked substrates (Picart *et al.* 2014). Given these findings, we propose that the shorter sequences in HypGST cluster A from non-sphingomonad strains have an alternative glutathione-dependent function, and do not encode active β-etherase enzymes.

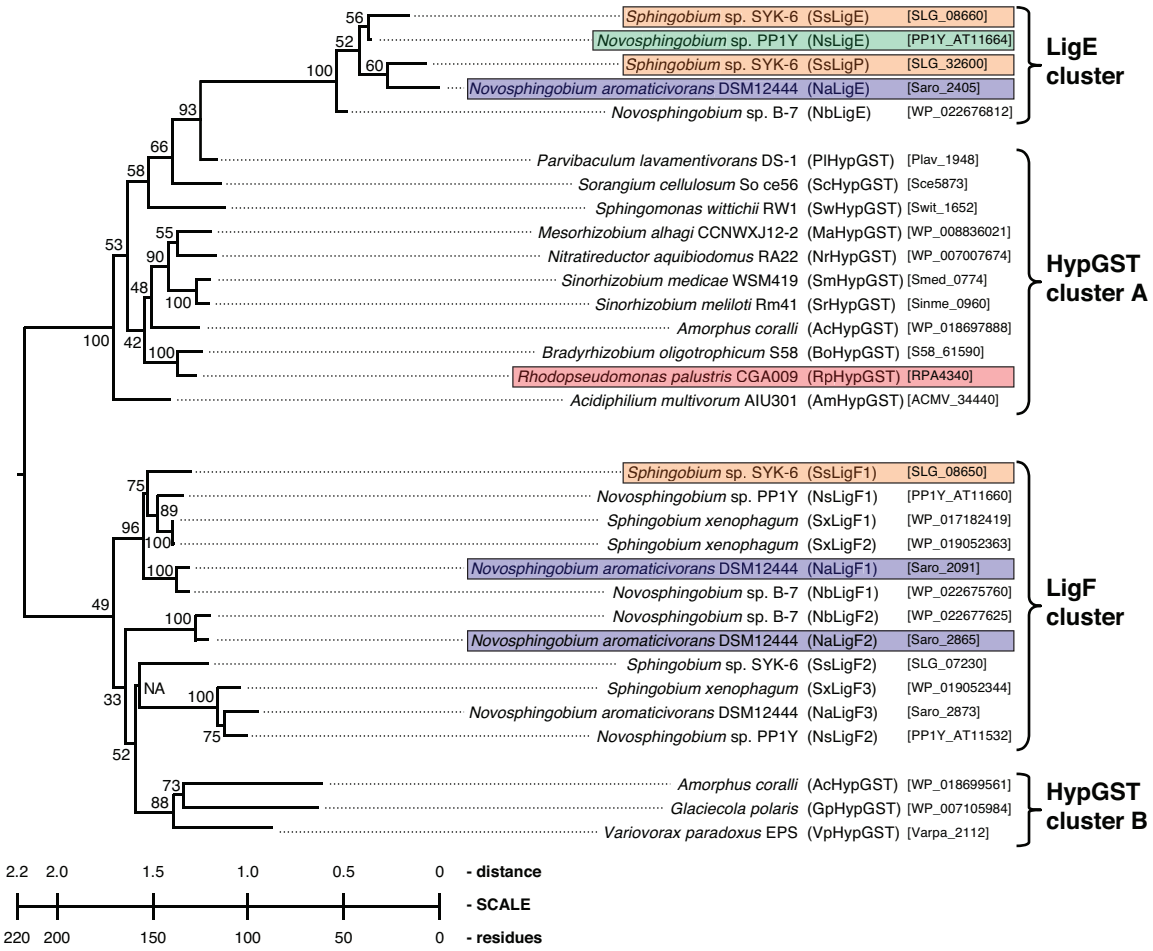

Figure 3-10. Phylogenetic tree of aligned β-etherase amino acid sequences (10,000 bootstrap trials, 111 seeds). The 31 aligned sequences depicted were from the 15 most similar sequences to each SsLigE and SsLigF1 found in the BLASTP database, in addition to the sequence encoding RpHypGST. Gene symbols and locus tags are shown in parentheses and brackets. The LigE cluster shows the five closely related LigE-like sequences. HypGST cluster A shows the eleven divergent sequences from the SsLigE BLASTP search. The LigF cluster shows the twelve closely related LigF sequences (gene symbol numerals indicate relatedness to SsLigF1, where "1" indicates most similar). HypGST cluster B shows the three divergent sequences from the SsLigF1 BLASTP search. The LigE and LigF enzymes from selected strains that were tested for β-etherase activity in this study are highlighted by color: *Sphingobium* sp. strain SYK-6 (orange), *Novosphingobium* sp. strain PP1Y (green), *Novosphingobium aromaticivorans* strain DSM12444 (blue), and *Rhodopseudomonas palustris* CGA009 (red).

Discussion

Recently, it has been shown that GST superfamily enzymes from *Sphingobium* sp. strain SYK-6 have the ability to act as stereospecific β-etherases using lignin model compounds as substrates (Masai *et al.* 2003; Gall *et al.* 2014). These so-called Lig β-etherases have been shown to cleave lignin dimers composed of guaiacyl monomers. In this study, we investigated whether Lig β-etherases from *Sphingobium* sp. strain SYK-6 also exhibit enzyme activity with substrates that contain syringyl units, the other major monomeric constituent of lignin. Further, we investigated whether other bacteria possess sequence-related proteins with similar or different substrate or stereospecificities as those reported for the *Sphingobium* sp. strain SYK-6 enzymes.

This study reveals for the first time that (a) several species of sphingomonads encode glutathione-dependent enzymes that catalyze cleavage of β-ether linkages that are found in lignin, (b) each Lig homolog cleaves guaiacyl-β-guaiacyl, guaiacyl-β-syringyl, syringyl-β-guaiacyl, and syringyl-β-syringyl β-ether-linked substrates, and (c) with each substrate, LigE/LigP and their homologs exhibit β(*R*)-stereospecificity whereas LigF and its homologs have β(*S*)-ether stereospecificity. These results show that methoxy group ring substitutions on the aromatic monomeric units are not inhibitory to the function of these β-etherase enzymes. Rather, sphingomonads use enzymes with active sites that are receptive to variably methoxylated rings. Also, these findings give insight into how a set of β-etherase pathway enzymes from different species accommodate substrates containing the multiple chiral centers (*i.e.*, at carbons α and β) that exist in the β-ether-linked structures found in lignin (Ralph *et al.* 1999; Akiyama *et al.* 2000). The NAD-dependent dehydrogenases oxidize and eliminate the chiral center at carbon α, forming α-keto-β(*R*)- and α-keto-β(*S*)-enantiomers. Further, the existence of both β(*R*)- and β(*S*)-ether enantiomers

in nature is overcome by the evolution of separate glutathione-dependent enzymes with either β(*R*)- or β(*S*)-ether-cleaving reaction mechanisms.

Our results predict that a single organism may contain multiple β(*R*)-etherases (*e.g.*, SsLigE and SsLigP) or numerous β(*S*)-etherases (*e.g.*, NsLigF1 and NsLigF2), each of which is capable of catalyzing cleavage of GβG, GβS, SβG, and SβS enantiomers. In sphingomonads *Sphingobium* sp. SYK-6, *Novosphingobium* sp. strain PP1Y, *N. aromaticivorans* strain DSM12444, and another *Novosphingobium* strain with sequence-related homologs to Lig enzymes, strain B-7 (Figure 3-10), it appears that metabolism of α-keto-β-ether-linked compounds is achieved via catalysis by multiple Lig β-etherases with overlapping function. However, it is possible that variations of the pathway may exist in closely related bacteria. For example, a phylogenetic tree constructed from an alignment of LigE/LigP and LigF homologs (Figure 3-10) reveals that five LigE/LigP homologs belonging to four sphingomonad strains (Figure 3-10, LigE cluster) were more closely related to each other than the next eleven sequences identified in the SsLigE BLASTP search (Figure 3-10, HypGST cluster A). The genome of each sphingomonad encodes multiple LigF homologs. Five such sphingomonad strains encode closely related putative LigF enzymes (Figure 3-10, LigF cluster) that exhibited sequence dissimilarity with the three non-sphingomonad LigF homologs (Figure 3-10, HypGST cluster B), perhaps because the HypGST sequences encode different functions.

Overall, BLASTP analysis predicts that six sphingomonads (α-Proteobacteria of the Order *Sphingomonadales*) encoded Lig homologs that were aligned in the phylogenetic tree (Figure 3-10). Additional BLASTP searches within the genomes of *Sphingobium* sp. strain SYK-6 and *Novo-sphingobium* strains B-7, PP1Y, and DSM12444, each of which had multiple sequences in the phylogenetic tree, revealed that each organism encoded both the LigE homolog needed for β(*R*)-enantiomer degradation, and the LigF homolog required for catabolism of β(*S*)-enantiomers. Of these, only *Sphingobium* sp. strain SYK-6 encoded multiple β(*R*)-specific (SsLigE and SsLigP) and

multiple β(*S*)-specific enzymes (SsLigF1 and SsLigF2). However, *Novosphingobium* sp. strain PP1Y, *Novosphingobium* sp. strain B-7, and *N. aromaticivorans* strain DSM12444 were each found to encode a single LigE homolog and multiple sequences with LigF homology. Also, all four sphingomonad strains additionally encode multiple NAD-dependent dehydrogenases that catalyze the formation of the α-ketones required for β-ether cleavage activity.

The fifth sphingomonad that encodes putative β-etherases, *Sphingomonas wittichii* strain RW1, had a single LigE homolog (Figure 3-10, HypGST cluster A) that, based on sequence analysis, is more similar to RpHypGST (which had a shorter sequence and did not exhibit β-etherase activity) than to the confirmed β-etherases in the LigE cluster. Further, the *S. wittichii* genome did not encode a protein related to those that have β(*S*)-etherase activity or putative NAD-dependent Lig dehydrogenase activity, suggesting that the LigE homolog in *S. wittichii* does not encode a function related to β-ether catabolism. Another sphingomonad, *Sphingobium xenophagum*, encoded three homologs with potential β(*S*)-specific activity (SxLigF1, SxLigF2, and SxLigF3), but did not encode a LigE homolog. Given the high sequence similarity to enzymes with demonstrated β(*S*)-etherase activity, it is possible that *S. xenophagum* carries out β(*S*)-enantiomer catabolism with its various LigF homologs but uses alternative metabolic pathways for the degradation of β(*R*)-enantiomers.

Thirteen of the thirty-one sequences in the phylogenetic tree (Figure 3-10) are derived from non-sphingomonads, one from each of α- (of the Order *Rhodospirillales*), β-, γ-, and δ-Proteobacteria, and nine from α-Proteobacteria (of the Order *Rhizobiales*). *Amorphus coralii* was the only non-sphingomonad that encoded both a LigE- and a LigF-like protein. However, unlike in *Novosphingobium* sp. strains B-7 and PP1Y, *N. aromaticivorans* strain DSM12444, and *Sphingobium* sp. strain SYK-6, the *A. coralii* genome encoded no sequences with homology to the NAD-dependent Lig dehydrogenases, suggesting that homologs from *A. coralii* are HypGSTs with alternative functions to those of the Lig β-etherases. Further, the *A. coralii* LigE homolog clustered with the

other homologs with shorter sequences that we predict not to have β-etherase activity (Figure 3-10, HypGST cluster A). The genomes of *Glaciecola polaris* and *Variovorax paradoxus* EPS, each encode a single LigF-like sequence (Figure 3-10, HypGST cluster B), but did not encode homologs of any of the other essential β-etherase pathway enzymes. We therefore propose that the HypGST proteins in clusters A and B do not have activity as β-etherases with the lignin compounds used in this study.

Given that each of the LigE/LigP enzymes that we tested catalyzed β(*R*)-ether cleavage, whereas each LigF enzyme exhibited β(*S*)-stereospecificity, we propose that the β-etherase pathway functions similarly in *Novosphingobium* sp. strains B-7 and PP1Y, *N. aromaticivorans* strain DSM12444, and *Sphingobium* sp. strain SYK-6. These organisms appear to have adapted to the racemic nature of lignin by evolving multiple glutathione-dependent enzymes with complementary β-etherase stereospecificities. It will be intriguing to learn if the functions of the β-etherase pathway are unique to the sphingomonads as the availability of additional genome sequences pave the way for future studies of lignin catabolism in other bacteria.

Abbreviations and Nomenclature

NAD, nicotinamide adenine dinucleotide; GSH, glutathione; GSSG, glutathione disulfide; OMe, methoxyl; GβG, α-(4-O-Me)-guaiacylglycerone-β-(1′-formyl)-guaiacyl ether; GβS, α-(4-O-Me)-guaiacylglycerone-β-(1′-formyl)-syringyl ether; SβG, α-(4-O-Me)-syringylglycerone-β-(1′-formyl)-guaiacyl ether; SβS, α-(4-O-Me)-syringylglycerone-β-(1′-formyl)-syringyl ether; Gβ-SG, β-glutathionyl-α-(4-O-Me)-guaiacylglycerone; Sβ-SG, β-glutathionyl-α-(4-O-Me)-syringylglycerone; TCEP, *tris*-(2-carboxyethyl)-phosphine hydrochloride; GST, glutathione-S-transferase; Na, *Novosphingobium aromaticivorans* strain DSM12444; Ns, *Novosphingobium* sp. strain PP1Y; Rp, *Rhodopseudomonas palustris* strain CGA009; Ss, *Sphingobium* sp. strain SYK-6; HypGST, hypothetical glutathione-S-transferase; His_8, octa-histidine affinity tag; Ni-NTA, nickel-nitrilotriacetic acid resin;

HSQC, (^{1}H–^{13}C) heteronuclear single quantum coherence (NMR spectroscopy); HMBC, (^{1}H–^{13}C) heteronuclear multiple-bond correlation (NMR spectroscopy); COSY, (^{1}H–^{1}H) correlation spectroscopy; t_R, retention time; MTPA(*R*), α(*R*)-methoxy-trifluoromethyl-phenylacetate; MTPACl(*S*), α(*S*)-methoxy-trifluoromethyl-phenylacetyl chloride; GβG-MTPA(*R*), β-(1′-formyl)-guaiacyl-α-(4-O-Me)-guaiacylglyceryl α(*R*)-methoxy-trifluoromethyl-phenyl-acetate; SβS-MTPA(*R*), β-(1′-formyl)-syringyl-α-(4-O-Me)-syringylglyceryl α(*R*)-methoxy-trifluoromethyl-phenyl-acetate; GβG-propenone, α-(4-O-Me)-guaiacyl-β,γ-propenone-β-(1′-formyl)-guaiacyl ether.

Acknowledgments

This work was supported by the Department of Energy Office of Science's Great Lakes Bioenergy Research Center, Grant DE-FC02-07ER64494. Daniel Gall was supported by a NIGMS Biotechnology Training grant (Grant T32 GM08349). We thank Sally Ralph at the US Forest Product Laboratory, and members of the Ralph laboratory for aiding in synthesis, analysis, and characterization of model compounds.

References

Adler E. (1977). Lignin chemistry - past, present and future. *Wood Science and Technology* 11(3):169-218.

Adler E., Eriksoo E. (1955). Guaiacylglycerol and its β-guaiacyl ether. *Acta chemica Scandinavica* 9341-342.

Akiyama T., Magara K., Matsumoto Y., Meshitsuka G., Ishizu A., Lundquist K. (2000). Proof of the presence of racemic forms of arylglycerol-β-aryl ether structure in lignin: studies on the stereo structure of lignin by ozonation. *Journal of Wood Science* 46(5):414-415.

Akiyama T., Sugimoto T., Matsumoto Y., Meshitsuka G. (2002). *Erythro*/*threo* ratio of β-O-4 structures as an important structural characteristic of lignin. I: Improvement of ozonation method for the quantitative analysis of lignin side-chain structure. *Journal of Wood Science* 48(3):210-215.

Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* 215(3):403-410.

Blommel P.G., Fox B.G. (2007). A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expression and Purification* 55(1):53-68.

Bugg T.D.H., Ahmad M., Hardiman E.M., Singh R. (2011). The emerging role for bacteria in lignin degradation and bio-product formation. *Current Opinion in Biotechnology* 22(3):394-400.

Chen F., Dixon R.A. (2007). Lignin modification improves fermentable sugar yields for biofuel production. *Nature Biotechnology* 25(7):759-761.

Chen L.Q., Hall P.R., Zhou X.Y.E., Ranson H., Hemingway J., Meehan E.J. (2003). Structure of an insect δ-class glutathione S-transferase from a DDT-resistant strain of the malaria vector *Anopheles gambiae*. *Acta Crystallographica Section D-Biological Crystallography* 592211-2217.

D'Argenio V., Petrillo M., Cantiello P., Naso B., Cozzuto L., Notomista E., Paolella G., Di Donato A., Salvatore F. (2011). *De novo* sequencing and assembly of the whole genome of *Novosphingobium* sp. strain PP1Y. *Journal of Bacteriology* 193(16):4296-4296.

Dixon R.A., Paiva N.L. (1995). Stress-induced phenylpropanoid metabolism. *Plant Cell* 7(7):1085-1097.

Freudenberg K. (1959). Biosynthesis and constitution of lignin. *Nature* 183(4669):1152-1155.

Gall D.L., Kim H., Lu F., Donohue T.J., Noguera D.R., Ralph J. (2014). Stereochemical features of glutathione-dependent enzymes in the *Sphingobium* sp. strain SYK-6 β-aryl etherase pathway. *J Biol Chem* 289(12):8656-8667.

Harwood C.S., Gibson J. (1988). Anaerobic and aerobic metabolism of diverse aromatic compounds by the photosynthetic bacterium *Rhodopseudomonas palustris*. *Applied and Environmental Microbiology* 54(3):712-717.

Higuchi T. (1980) Lignin structure and morphological distribution in plant cell walls. In: Kirk T.K., Higuchi T., Chang H. (eds) Lignin biodegradation: microbiology, chemistry and potential applications. CRC Press, Boca Raton, Florida, pp 1-20

Hishiyama S., Otsuka Y., Nakamura M., Ohara S., Kajita S., Masai E., Katayama Y. (2012). Convenient synthesis of chiral lignin model compounds via optical resolution: four stereoisomers of

guaiacylglycerol-β-guaiacyl ether and both enantiomers of 3-hydroxy-1-(4-hydroxy-3-methoxyphenyl)-2-(2-methoxy-phenoxy)-propan-1-one (erone). *Tetrahedron Letters* 53842-845.

Ji X.H., Johnson W.W., Sesay M.A., Dickert L., Prasad S.M., Ammon H.L., Armstrong R.N., Gilliland G.L. (1994). Structure and function of the xenobiotic substrate-binding site of a glutathione-S-transferase as revealed by x-ray crystallographic analysis of product complexes with the diastereomers of 9-(*S*-glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene. *Biochemistry* 33(5):1043-1052.

Landucci L.L., Geddes S.A., Kirk T.K. (1981). Synthesis of C-14-labeled 3-methoxy-4-hydroxy-α-(2-methoxyphenoxy)-β-hydroxypropiophenone, a lignin model-compound. *Holzforschung* 35(2):67-70.

LaRoe S.L., Wang B., Han J.-I. (2010). Isolation and characterization of a novel polycyclic aromatic hydrocarbon-degrading bacterium, *Sphingopyxis* sp. strain M2R2, capable of passive spreading motility through soil. *Environmental Engineering Science* 27(6):505-512.

Lewis N.G., Yamamoto E. (1990). Lignin - occurrence, biogenesis and biodegradation. *Annual Review of Plant Physiology and Plant Molecular Biology* 41455-496.

Masai E., Ichimura A., Sato Y., Miyauchi K., Katayama Y., Fukuda M. (2003). Roles of the enantioselective glutathione S-transferases in cleavage of β-aryl ether. *Journal of Bacteriology* 185(6):1768-1775.

Masai E., Kamimura N., Kasai D., Oguchi A., Ankai A., Fukui S., Takahashi M., Yashiro I., Sasaki H., Harada T., Nakamura S., Katano Y., Narita-Yamada S., Nakazawa H., Hara H., Katayama Y., Fukuda M., Yamazaki S., Fujita N. (2012). Complete genome sequence of *Sphingobium* sp, strain SYK-6, a degrader of lignin-derived biaryls and monoaryls. *Journal of Bacteriology* 194(2):534-535.

Masai E., Katayama Y., Fukuda M. (2007). Genetic and biochemical investigations on bacterial catabolic pathways for lignin-derived aromatic compounds. *Bioscience Biotechnology and Biochemistry* 71(1):1-15.

Masai E., Katayama Y., Kawai S., Nishikawa S., Yamasaki M., Morohoshi N. (1991). Cloning and sequencing of the gene a *Pseudomonas-paucimobilis* enzyme that cleaves β-aryl ether. *Journal of Bacteriology* 173(24):7950-7955.

Masai E., Katayama Y., Kubota S., Kawai S., Yamasaki M., Morohoshi N. (1993a). A bacterial enzyme degrading the model lignin compound β-etherase is a member of the glutathione-S-transferase superfamily. *Febs Letters* 323(1-2):135-140.

Masai E., Katayama Y., Nishikawa S., Yamasaki M., Morohoshi N., Haraguchi T. (1989). Detection and localization of a new enzyme catalyzing the β-aryl ether cleavage in the soil bacterium (*Pseudomonas-paucimobilis* SYK-6). *Febs Letters* 249(2):348-352.

Masai E., Kubota S., Katayama Y., Kawai S., Yamasaki M., Morohoshi N. (1993b). Characterization of the Cα-dehydrogenase gene involved in the cleavage of β-aryl ether by *Pseudomonas paucimobilis*. *Bioscience Biotechnology and Biochemistry* 57(10):1655-1659.

Moore D.D. (2003) Current protocols in molecular biology. In: Ausubel F.M., Brent R., Kingston R.E., Moore D.D., Seidman J.G., Smith J.A., Struhl K. (eds). John Wiley & Sons, New York, NY

Notomista E., Pennacchio F., Cafaro V., Smaldone G., Izzo V., Troncone L., Varcamonti M., Di Donato A. (2011). The marine isolate *Novosphingobium* sp. PP1Y shows specific adaptation to use

the aromatic fraction of fuels as the sole carbon and energy source. *Microbial Ecology* 61(3):582-594.

Picart P., Müller C., Mottweiler J., Wiermans L., Bolm C., Domínguez de María P., Schallmey A. (2014). From gene towards selective biomass valorization: bacterial β-etherases with catalytic activity on lignin-like polymers. *ChemSusChem*.

Prochazkova K., Shuvalova L.A., Minasov G., Voburka Z., Anderson W.F., Satchell K.J.F. (2009). Structural and molecular mechanism for autoprocessing of MARTX toxin of *Vibrio cholerae* at multiple sites. *Journal of Biological Chemistry* 284(39):26557-26568.

Ralph J., Peng J.P., Lu F.C., Hatfield R.D., Helm R.F. (1999). Are lignins optically active? *Journal of Agricultural and Food Chemistry* 47(8):2991-2996.

Sarkanen K.V., Ludwig C.H. (1971) Lignins: Occurrence, formation, structure and reactions In: Sarkanen K.V., Ludwig C.H. (eds). John Wiley & Sons, Inc., New York, pp 1-916

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. *Applied and Environmental Microbiology* 75(16):5195-5201.

Sheahan K.-L., Cordero C.L., Satchell K.J.F. (2007). Autoprocessing of the *Vibrio cholerae* RTX toxin by the cysteine protease domain. *Embo Journal* 26(10):2552-2561.

Shen A., Lupardus P.J., Albrow V.E., Guzzetta A., Powers J.C., Garcia K.C., Bogyo M. (2009). Mechanistic and structural insights into the proteolytic activation of *Vibrio cholerae* MARTX toxin. *Nature Chemical Biology* 5(7):469-478.

Simmons B.A., Logue D., Ralph J. (2010). Advances in modifying lignin for enhanced biofuel production. *Current Opinion in Plant Biology* 13(3):313-320.

Sonoki T., Iimura Y., Masai E., Kajita S., Katayama Y. (2002). Specific degradation of β-aryl ether linkage in synthetic lignin (dehydrogenative polymerizate) by bacterial enzymes of *Sphingomonas paucimobilis* SYK-6 produced in recombinant *Escherichia coli*. *Journal of Wood Science* 48(5):429-433.

Sugimoto T., Akiyama T., Matsumoto Y., Meshitsuka G. (2002). The *erythro/threo* ratio of β-O-4 structures as an important structural characteristic of lignin - Part 2. Changes in *erythro/threo* (E/T) ratio of β-O-4 structures during delignification reactions. *Holzforschung* 56(4):416-421.

Takeuchi M., Hamana K., Hiraishi A. (2001). Proposal of the genus *Sphingomonas sensu stricto* and three new genera, *Sphingobium*, *Novosphingobium* and *Sphingopyxis*, on the basis of phylogenetic and chemotaxonomic analyses. *International Journal of Systematic and Evolutionary Microbiology* 511405-1417.

Tanamura K., Kasai D., Nakamura M., Katayama Y., Fukuda M., Masai E. (2010). Identification of the third glutathione S-transferase gene involved in the stereospecific cleavage of β-aryl ether in *Sphingobium* sp. strain SYK-6. *Journal of Biotechnology* 150S235-S235.

Xia Y., Min H., Rao G., Lv Z.M., Liu J., Ye Y.F., Duan X.J. (2005). Isolation and characterization of phenanthrene-degrading *Sphingomonas paucimobilis* strain ZX4. *Biodegradation* 16(5):393-402.

Zakzeski J., Bruijnincx P.C.A., Jongerius A.L., Weckhuysen B.M. (2010). The catalytic valorization of lignin for the production of renewable chemicals. *Chemical Reviews* 110(6):3552-3599.

CHAPTER 4: Structural basis of bacterial β-aryl ether lignin degradation

This chapter is formatted as a manuscript and has been submitted for publication:

Pereira J.H., Helmich K.E., Heins R.A., Gall D.L., McAndrew R.P., Bingman C., Deng K., Holland K.C., Donohue T.J., Noguera D.R., Simmons B.A., Sale K.L., Phillips G.N., Ralph J., Adams P.D. Structural basis of bacterial β-aryl ether lignin degradation.

Daniel L. Gall and Richard A. Heins carried out the biochemical assays described in this chapter. Daniel L. Gall and Kai Deng conducted organic syntheses and structural analyses. Jose H. Pereira, Kate E. Helmich, Ryan P. McAndrew, Craig Bingman, and Keefe C. Holland carried out crystallographic techniques and analyses.

Abstract

Lignin is a combinatorial polymer comprised of monoaromatic units that are linked together via strong chemical bonds. Although lignin is a potential source of valuable aromatic chemicals, its recalcitrance presents major obstacles to both the extraction of sugars from lignocellulosic biomass for production of second-generation biofuels and the generation of valuable coproducts from degradation of this polymer. Degradation of lignin has been relatively well characterized in fungi, but is less well characterized in bacteria. A catabolic pathway for the enzymatic breakdown of β-aryl ether-linked units in lignin has been reported in the bacterium *Sphingobium* sp. SYK-6. As β-aryl ether bonds account for 50–70% of all inter-unit linkages in lignin, understanding the mechanism of β-aryl ether cleavage has great potential for the effective breakdown of lignin. Here we present a structural and biochemical characterization of enzymes in this pathway, LigD, LigO, LigL, LigE, LigF and LigG, including cofactor binding sites, substrate binding sites, and catalytic mechanisms. This new information can be used to enhance the degradation of lignin and its use as a source for valuable aromatics and other renewable chemicals, including second-generation biofuels.

Introduction

The production of renewable chemicals and advanced biofuels from lignocellulosic biomass is a potentially sustainable alternative route to support the growing demand for energy. The primary obstacle in the production of lignocellulosic biofuels is the release of sugars in high quantities at low cost from recalcitrant lignocellulosic biomass feedstocks (Simmons *et al.* 2008). As lignin is the prime source of this recalcitrance, there has been renewed interest in the microbial enzymes capable of lignin degradation (Chang and Holtzapple 2000; Bugg *et al.* 2011b). Additionally, the lignin by-product of biomass processing is a potential source of aromatic compounds that are widely used in the chemical industry, as lignin is the most abundant aromatic polymer in nature (Masai *et al.* 2007; Reiter *et al.* 2013). Generally, white rot and brown rot fungi secrete lignin peroxidases, manganese peroxidases and laccases that are involved in initial degradation of lignin (Leonowicz *et al.* 2001; Martinez *et al.* 2005), whereas bacteria are thought to play a role in degradation of lignin-derived lower molecular weight compounds (Masai *et al.* 2007).

Sphingobium sp. strain SYK-6, one of the most well-studied bacteria implicated in lignin degradation, has the ability to derive nutrients needed for growth from a wide variety of dimeric aromatic compounds representing the various units, with their characteristic inter-unit linkages, present in plant lignin (Masai *et al.* 2007; Sato *et al.* 2009). The cleavage of β-aryl ether (termed β-ether hereafter) linkages is an essential step in the lignin degradation and catabolic process, as this bond type accounts for 50–70% of all inter-unit linkages in lignin polymers (Adler 1957). Using a β-ether-linked lignin model substrate, guaiacylglycerol-β-guaiacyl ether (GGE, Figure 4-1), three enzymatic reactions comprising the β-ether degradation pathway were identified in *Sphingobium* sp. strain SYK-6 (Masai *et al.* 2003; Sato *et al.* 2009; Gall *et al.* 2014a):

1. Stereospecific oxidation of the α-hydroxyl group of GGE to β-(3′-methoxyphenoxy)-γ-hydroxypropiovanillone (MPHPV) is catalyzed by the nicotinamide adenine dinucleotide

(NAD^+)-dependent Cα-dehydrogenases LigD, LigL, LigN and LigO. The LigD and LigO enzymes catalyze oxidation of the α(*R*)-substrates (α*R*,β*R*)-GGE and (α*R*,β*S*)-GGE whereas LigL and LigN exhibit stereospecificity towards α(*S*)-configured substrates (α*S*,β*R*)-GGE and (α*S*,β*S*)-GGE.

2. Members of the glutathione-S-transferase (GST) superfamily of enzymes, β-etherases LigE and LigF, stereospecifically catalyze glutathione (GSH)-dependent cleavage of the β-ether linkage in MPHPV, forming β-glutathionyl-γ-hydroxypropiovanillone (GS-HPV) and guaiacol. LigE catalyzes cleavage of (β*R*)-MPHPV and results in stereoselective formation of (β*S*)-GS-HPV, whereas LigF converts (β*S*)-MPHPV to (β*R*)-GS-HPV.

3. The GSH-dependent lyase LigG catalyzes elimination of the GSH thioether linkage in (β*S*)-GS-HPV, generating glutathione disulfide (GSSG) and the achiral derivative γ-hydroxy-propiovanillone (HPV) that ultimately serves as the growth substrate for strain SYK-6 (Masai *et al.* 2007; Sato *et al.* 2009) (Figure 4-1). Although the metabolic fate of (β*R*)-GS-HPV is uncertain, it is likely that either (β*R*)-GS-HPV is converted to (β*S*)-GS-HPV via racematization or an unknown enzyme, a stereochemical complement to LigG, cleaves (β*R*)-GS-HPV to HPV.

Because plant lignin is biosynthesized via radical coupling reactions that give rise to the formation of chiral centers as the polymer assembles, complementary stereospecificities of the multiple enzymes in the β-ether degradation pathway are required to oxidize and cleave the various stereoisomers that are present in lignin polymers (Ralph *et al.* 1999; Akiyama *et al.* 2000; Sugimoto *et al.* 2002).

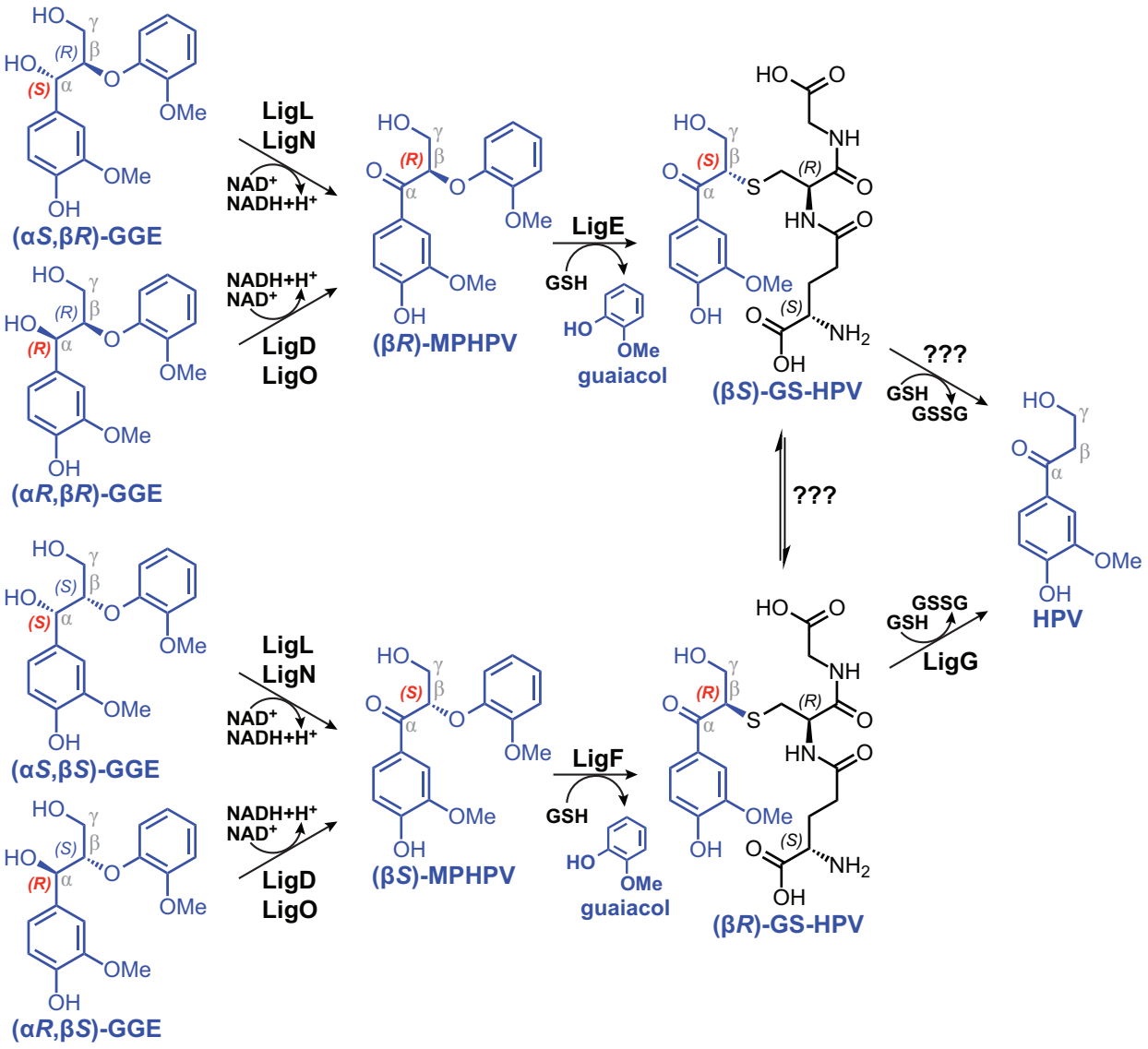

Figure 4-1. The *Sphingobium* sp. strain SYK-6 β-etherase pathway. Enantiomeric configuration labels for chiral carbons at which stereospecific reactions occur are shown in red. Stereospecific reactions for (α*S*,β*R*)-GGE and (α*S*,β*S*)-GGE oxidation (by LigL and LigN), (α*R*,β*R*)-GGE and (α*R*,β*S*)-GGE oxidation (by LigD and LigO), the GSH-dependent stereospecific cleavage reactions of (β*R*)-MPHPV (by LigE) and (β*S*)-MPHPV (by LigF), as well as the stereospecific lyase reaction of LigG with (β*S*)-GS-HPV, are shown.

Here, we describe 12 protein crystal structures and the corresponding biochemical data for native and variant enzymes involved in all the enzymatic steps of the *Sphingobium* sp. strain SYK-6 β-ether degradation pathway. The crystal structures of the Cα-dehydrogenases LigD, LigL and LigO

were solved as apoenzymes, with cosubstrate NADH (the reduced form of NAD^+) bound, or in the ternary complex of protein-NADH-GGE, providing a complete structural illustration of each of the different enzymatic states in the catalytic cycle. Structures of the β-etherase enzymes LigE and LigF with GSH provide insight into the potential enzymatic reaction mechanism for β-ether cleavage. Insight into the last reaction of the β-ether cleavage pathway, which is catalyzed by a glutathione lyase, is provided by the crystal structures of apo-LigG and the LigG-GS-AV (β-glutathionyl-aceto-veratrone) substrate analog complex (Figure 4-2A). The structural and biochemical information presented provides new insights into the structure-function relationships and biochemistry of these enzymes that expands the knowledge of bacterial catabolism of lignin breakdown products. We describe how this information will enable future development of efficient pathways for lignin conversion into renewable aromatics with applications in advanced biofuels and chemicals.

Results

NAD^+-dependent Cα-dehydrogenases (LigD, LigO and LigL)

Structural Analysis – We have solved the crystal structures of LigD, LigO, and LigL, which belong to the SDR superfamily of enzymes (Reid and Fewson 1994) that catalyze the oxidation of the α-hydroxyl group of GGE to form MPHPV (Figure 4-1). LigN, another Cα-dehydrogenase enzyme in the pathway proved recalcitrant to crystallization.

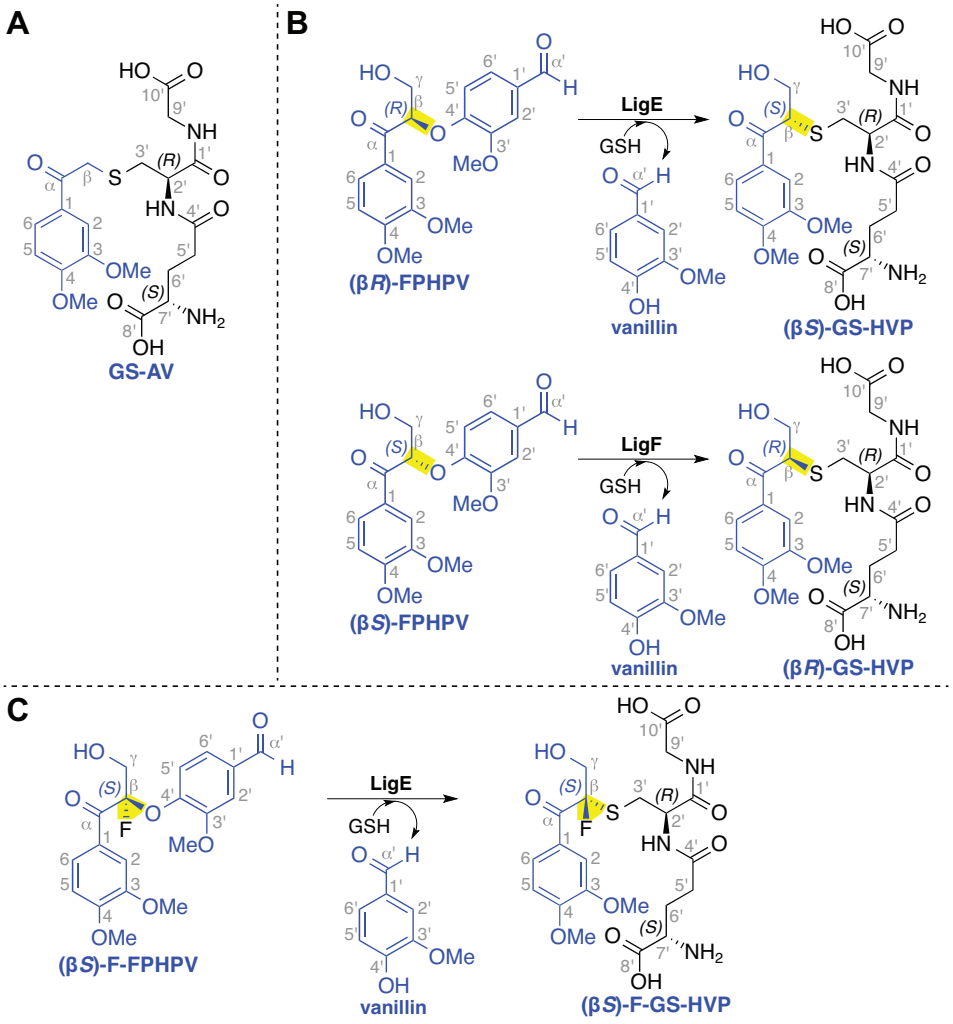

Figure 4-2. (A) Structure of the GS-HPV analog substrate, GS-AV, that was used in the crystallization of LigG. **(B)** Structure of an MPHPV analog substrate, FPHPV, that was used in the LigE- and LigF-catalyzed reactions, converting FPHPV to vanillin and GS-HVP. **(C)** LigE-catalyzed β-ether elimination reaction with fluorinated model substrate (β*S*)-F-FPHPV, resulting in formation of vanillin and (β*S*)-F-GS-HVP.

The SDR family is characterized as a large group of NADPH (2′-phosphorylated NADH)-dependent enzymes displaying a α/β folding pattern containing a Rossman-fold (Oppermann *et al.* 2003); this overall organization is seen in the LigD, LigO, and LigL crystal structures (Figure 4-3). LigD, LigO and LigL are classified as classical SDR members, which have a core structure of approximately 300 residues, and share the $_{11}$TGXXXGXG/A$_{18}$ sequence motif at the cofactor binding

site and the catalytic tetrad N_{115}-S_{144}-Y_{158}-K_{162} (LigL numbering) (Persson *et al.* 2009). LigD and LigO catalyze the oxidation of the α(*R*)-substrates whereas LigL and LigN exhibit stereospecificity towards α(*S*)-configured substrates (Figure 4-1). A sequence alignment of LigD, LigO and LigL shows an identity between LigD-LigO of 40%, LigD-LigL of 38%, and LigO-LigL of 37%. In these alignments, the most sequence-divergent region is in the predicted substrate binding loop (Figure 4-4). Interestingly this proposed substrate binding region is disordered in all of the crystal structures solved in the apo-form (LigD, LigO and LigL) and in the structures of the LigD-NADH and LigO-NADH complexes. However, in the structures of the binary complex of LigL-NADH and the ternary complex structure of LigL-NADH-(α*S*,β*R*)-GGE, this region is very well-ordered, suggesting that this is a flexible loop that undergoes a major conformational change upon cosubstrate binding to the LigL enzyme (Figure 4-5A).

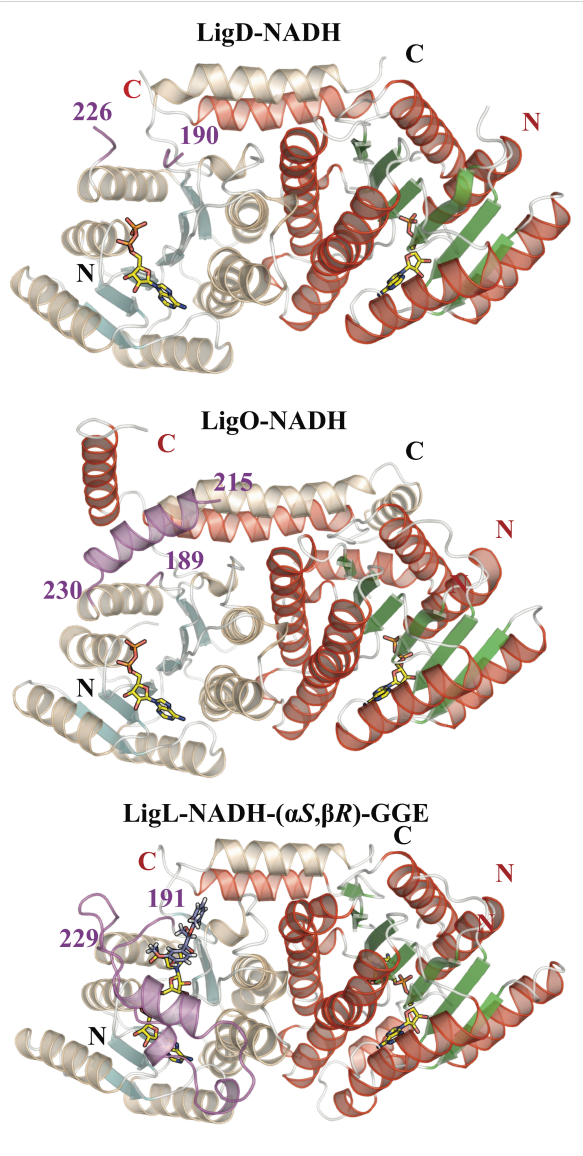

Figure 4-3. Cartoon representations of the biological dimers of LigD-NADH, LigO-NADH and LigL-NADH-(α*S*,β*R*)-GGE, showing the overall SDR family fold composed of a central Rossmann fold. The most sequence divergent region of SDR family members is the substrate binding loop represented in magenta. This region is disordered in all the crystal structures solved in the apo-forms (LigD, LigO and LigL) and in the structures of the LigD-NADH and LigO-NADH complexes. The apo-LigO and LigO-NADH structures showed a partially ordered region with an α-helix at the N-terminus of the substrate binding loop. This loop is ordered and modeled in the binary complex of LigL-NADH and ternary complex structure of LigL-NADH-(α*S*,β*R*)-GGE indicating a conformational change of this loop upon cosubstrate binding.

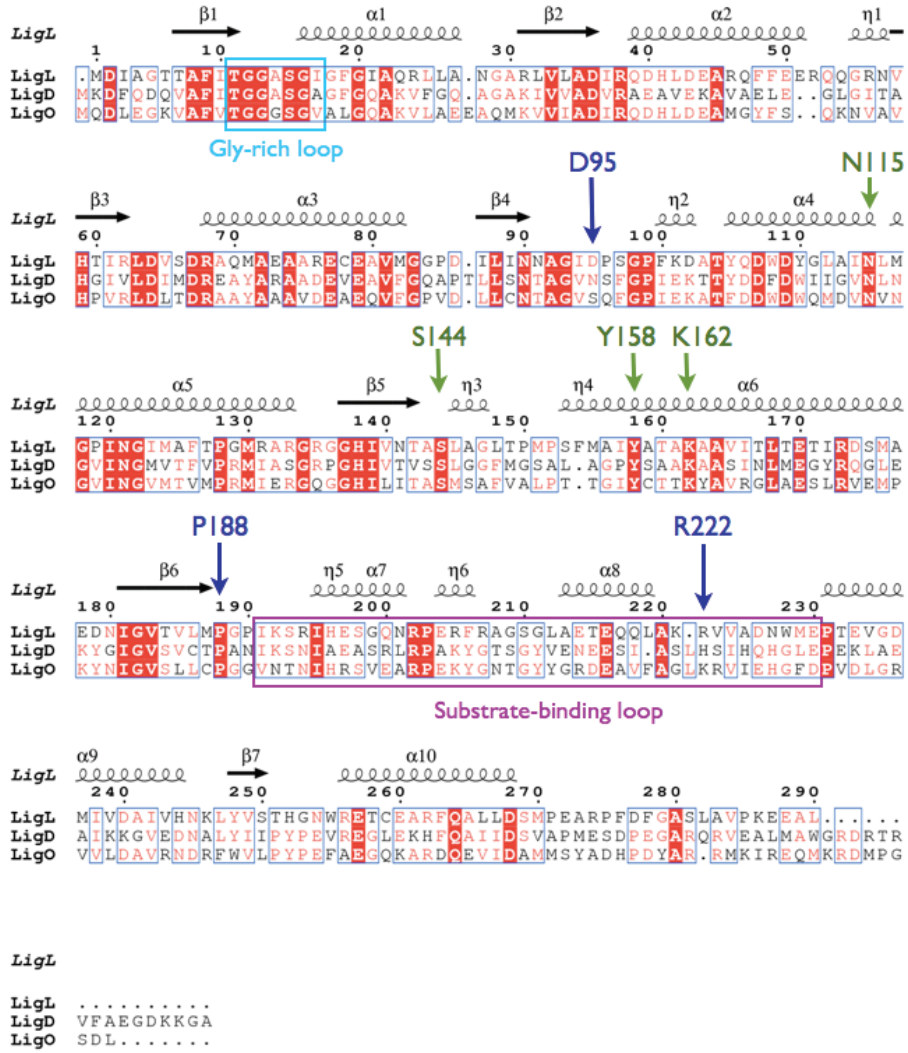

Figure 4-4. The sequence alignment between the Cα dehydrogenase LigD, LigO and LigL shows an identity between LigD-LigO, LigD-LigL and LigO-LigL of 40%, 38% and 37% respectively. The Glycine-rich loop consensus sequences (cyan box) located at N-terminal region $_{11}$TGXXXGXG/A$_{18}$ observed in all SDR family members are in contact with pyrophosphate group of NADH cosubstrate. Green arrows indicate the catalytic tetrad of SDR members N115-S144-Y158-K162. Blue arrows indicate the residues D-95, P-188 and R-222 of LigL that make direct contacts with GGE substrate. The most sequence-divergent region is the substrate-binding loop shown inside the magenta box.

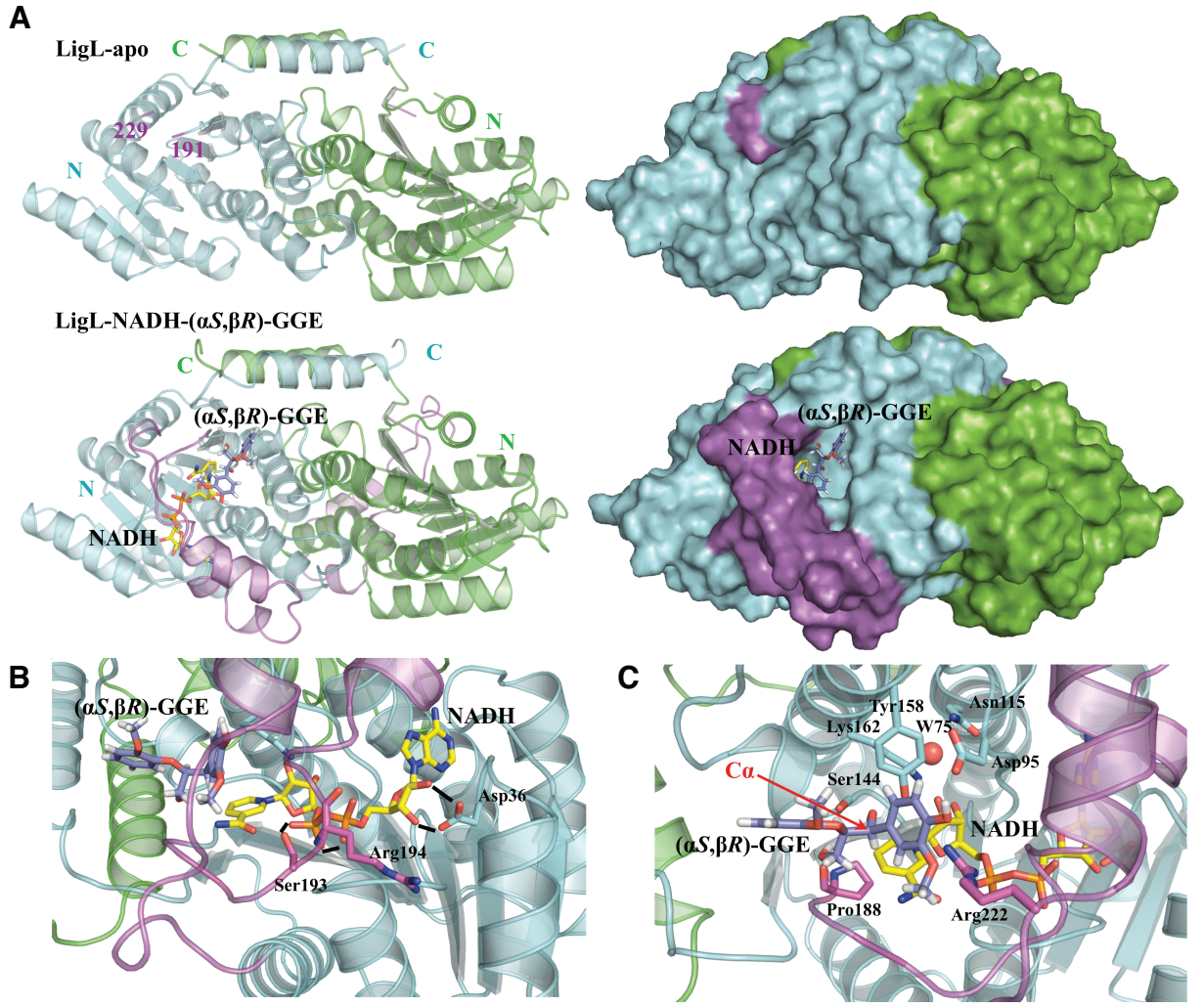

Figure 4-5. (A) Cartoon and molecular surface representations of apo-LigL and the LigL-NADH-(α*S*,β*R*)-GGE complex. The substrate binding loop (residues 191 to 229) is completely disordered in the apo-LigL structure. In the LigL-NADH-(α*S*,β*R*)-GGE complex structure the substrate binding loop region (magenta) works as a lid above the NADH and GGE binding sites. **(B)** Active site of LigL in complex with NADH and GGE displaying the interactions involving the cosubstrate NADH and the residues Ser193 and Arg194 located at the N-terminus of the substrate binding loop. Asp36 interacts with the 2′- and 3′- hydroxyl groups of the adenosine ribose sugar. **(C)** Active site of LigL-NADH-(α*S*,β*R*)-GGE showing the Cα position of GGE, the catalytic tetrad N_{115}-S_{144}-Y_{158}-K_{162}, and a water molecule (W75) involved in the extended proton relay system described for the SDR family (Filling *et al.* 2002). The residues Asp95, Ser144, Pro188 and Arg222 interact directly with the GGE substrate.

Despite the high quality of diffraction data for LigD-NADH (2.0 Å resolution) and LigO-NADH (1.7 Å resolution), disorder was observed for the nicotinamide moiety of NADH and only the

adenosine diphosphate of the cosubstrate was visible in the electron density maps (Figure 4-6A). Similar results were observed in crystals of other SDR members including a stereospecific short-chain alcohol dehydrogenase solved at 1.0 Å resolution, suggesting this is an intrinsic feature related to the flexibility of the nicotinamide region for some members of this family (Schlieben *et al.* 2005). However, the ternary complex structure of LigL-NADH-(α*S*,β*R*)-GGE reveals clear electron density for both NADH and the (α*S*,β*R*)-GGE substrate (Figure 4-6B). Interactions with the NADPH cosubstrate have been described previously for several short-chain dehydrogenase/reductase (SDR) family members (Filling *et al.* 2002; Oppermann *et al.* 2003; Schlieben *et al.* 2005; Javidpour *et al.* 2014). LigL interacts with adenosine region of NADH via residues located in the loop between strand β2 and helix α2; Asp36 contacts the 2′- and 3′- hydroxyl groups of the adenosine ribose sugar, the catalytic residues Tyr158 and Lys162 contact the nicotinamide ribose sugar, and finally the nicotinamide interacts with the side-chain atoms from Ser193 and Arg194. These residues are located at the start of the observed substrate binding loop, indicating that the hydrogen bonds involving the residues Ser193, Arg194 and the phosphate oxygen atoms of NADH seem to be important in stabilizing the substrate binding loop "closed" conformation (Figure 4-5B). The presence of a negatively charged aspartate (Asp36, Asp37 and Asp38 in LigL, LigD and LigO respectively) interacting with the hydroxyl groups of the adenosine ribose is likely responsible for favoring a NADH rather than NADPH cosubstrate. The binding of NADPH is likely unfavorable as a result of electrostatic repulsion between the aspartate residue and the NADPH 2′-phosphate group (Fan *et al.* 1991; Schlieben *et al.* 2005; Javidpour *et al.* 2014).

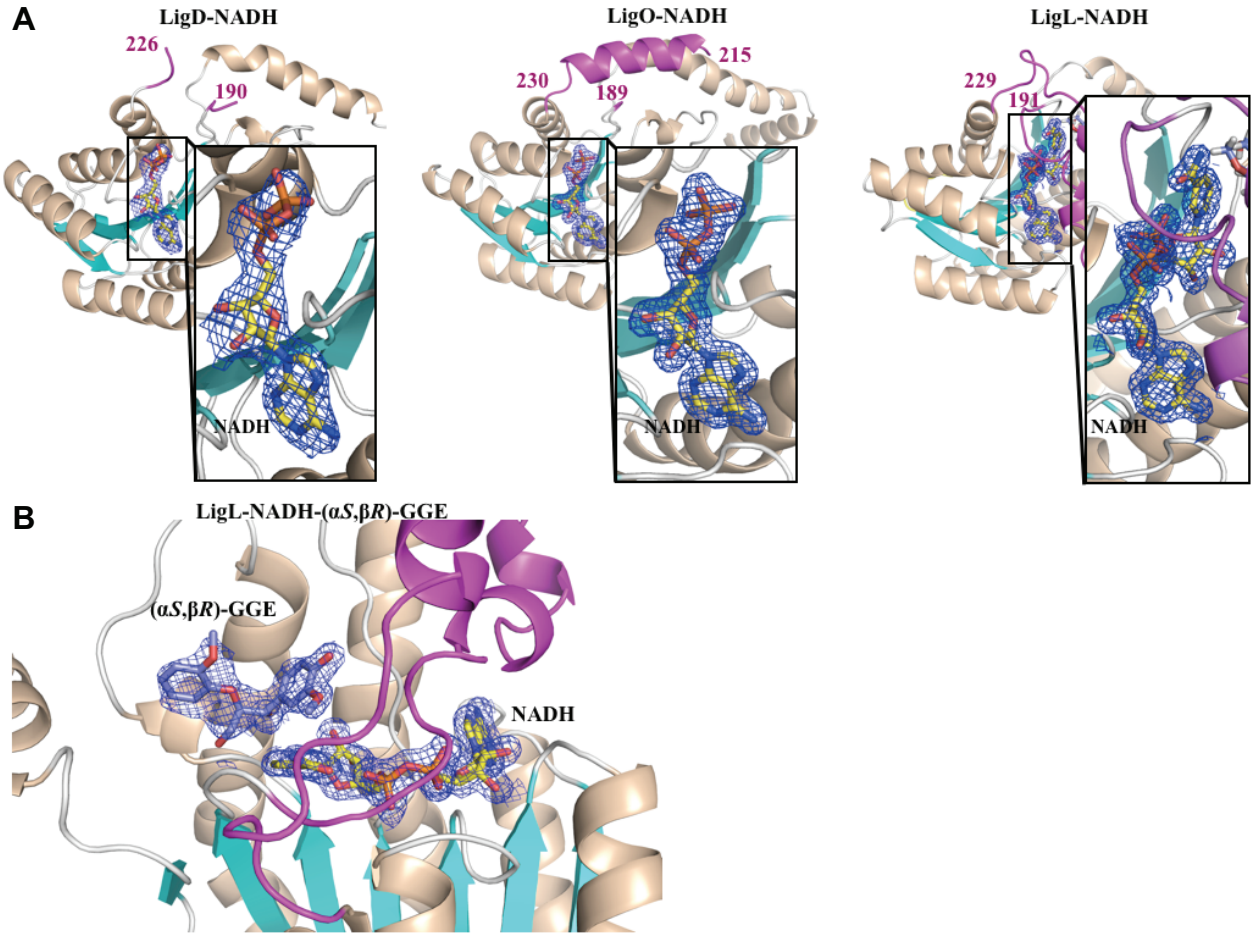

Figure 4-6. (A) A $2mF_o$-DF_c electron density map around the NADH nucleotide contoured at 1.0 σ is shown in blue. The adenine part of cosubstrate was defined on the electron density maps whereas the nicotinamide portion is missing from LigD-NADH and LigO-NADH structures. **(B)** The ternary complex structure of LigL-NADH-(α*S*,β*R*)-GGE show a clear A $2mF_o$-DF_c electron density map contoured at 1.0 σ is shown in blue around the NADH and GGE ligands.

The active site in the LigL-NADH-(α*S*,β*R*)-GGE ternary complex reveals that (α*S*,β*R*)-GGE makes direct interactions, via hydrogen bonds, with Asp95, Ser144 (a catalytic residue), the main chain carbonyl group of Pro188, and Arg222 located at the C-terminal of the substrate binding loop region (Figure 4-5C). The catalytic "extended proton relay system" mechanism for this class of enzymes, as proposed by Filling *et al.*, describes the role of the N_{115}-S_{144}-Y_{158}-K_{162} tetrad, the 2′-OH group of NAD^+ and a water molecule (Wat75 in the Figure 4-5C) in the transfer of a proton from the

active site to the bulk solvent (Filling *et al.* 2002). By analogy, the LigL tyrosine residue (Y158) functions as the catalytic base, the serine (S144) stabilizes the substrate via a hydrogen bond to the Cα-OH-group of (α*S*,β*R*)-GGE, and lysine (K162) interacts with the nicotinamide ribose sugar and is proposed to lower the pK_a of the tyrosine. Finally, the asparagine residue (N115) stabilizes the water molecule involved in the extended proton relay system (Filling *et al.* 2002). The conversion from GGE to MPHPV during the Cα-dehydrogenase reaction is achieved by the loss of a proton and a hydride ion from the substrate, creating NADH from NAD^+, and transfer of the proton to the bulk solvent (Figure 4-1). The ternary complex LigL-NADH-(α*S*,β*R*)-GGE reveals that the hydrogen bond to the Cα position of (α*S*,β*R*)-GGE is directed toward the NADH cofactor whereas the hydrogen atom from the Cα-OH-group of (α*S*,β*R*)-GGE is stabilized by the catalytic Ser144 (Figure 4-5C).

Enzymatic Analysis and Mutagenesis

We observed that the dehydrogenation of GGE to MPHPV is governed by solution pH for all four enzymes, with equilibrium conversion reaching approximately 90% at pH 9 but only 25% at pH 7 (Figure 4-7). Enzyme kinetics for LigL were therefore measured at pH 9 to reduce the influence of equilibrium effects on kinetic parameters. LigL exhibited Michaelis-Menton kinetics using both (α*S*,β*R*)-GGE and (α*S*,β*S*)-GGE stereoisomers (Table 4-1) with a higher turnover number (k_{cat}) and a lower Michaelis constant (K_M) toward the latter substrate. As expected from a prior report, LigL did not dehydrogenate (α*R*,β*S*)-GGE or (α*R*,β*R*)-GGE substrates (Sato *et al.* 2009).

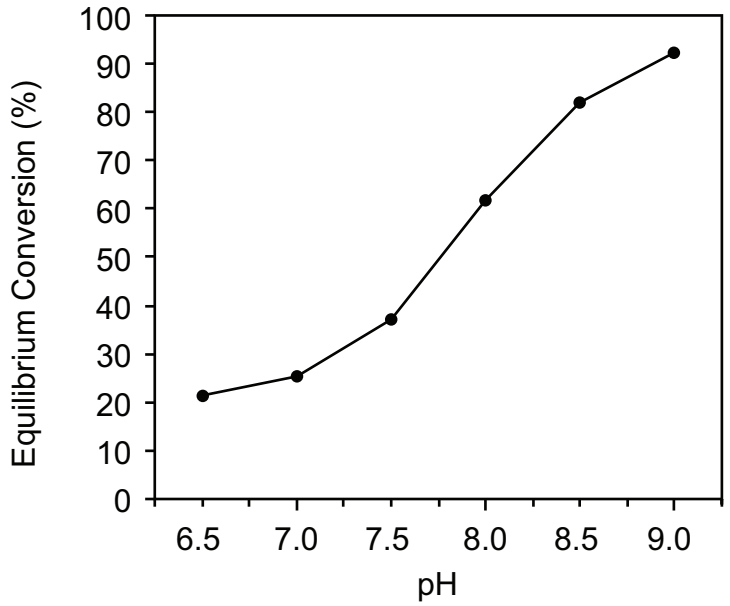

Figure 4-7. The dehydrogenation of GGE to MPHPV by LigL is strongly governed by solution pH, with equilibrium conversion reaching 90% at pH 9 but only 25% at pH 7. A phosphate buffer was used for pH 6.5 through 7.5, and a Tris buffer was used for pH 8 through 9.

Table 4-1. Kinetic parameters, determined from Michaelis-Menton curves for NAD^+-dependent Cα-dehydrogenase LigL and its variants with substrates (α*S*,β*R*)-GGE and (α*S*,β*S*)-GGE at pH 9.0.

Enzyme	Substrate	V_{MAX} (U mg^{-1})[a]	% WT activity with (α*S*,β*R*)-GGE[b]	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($mM^{-1}\ s^{-1}$)
LigL	(α*S*,β*S*)-GGE	33.7	154[b]	7.49	10.90	687.56
LigL	(α*S*,β*R*)-GGE	21.8	100	4.86	20.29	239.57
LigL-R222A	(α*S*,β*R*)-GGE	NDA	0	-	-	-
LigL-D95A	(α*S*,β*R*)-GGE	52.7	241	-	-	-
LigL-P188A	(α*S*,β*R*)-GGE	36.3	166	-	-	-
LigL-P188A-R222A	(α*S*,β*R*)-GGE	NDA	0	-	-	-

[a] Abbreviation: NDA, no detectable activity

[b] Where noted, alternative substrate (α*S*,β*S*)-GGE was used in biochemical assays

The crystal structure of the LigL-NADH-(α*S*,β*R*)-GGE complex revealed contacts between the substrate and Asp95, Ser144, Pro188 and Arg222 (Figure 4-5C). Prior analysis of SDR enzymes indicated that Ser144 was an essential catalytic residue being part of the catalytic tetrad N_{115}-S_{144}-Y_{158}-K_{162}. The remaining three residues (Asp95, Pro188 and Arg222) were therefore individually mutated to alanine to explore their contribution to enzyme catalysis (Table 4-1). A higher enzymatic activity was observed for the mutant LigL-Asp95Ala showing that Asp95 is not required for the binding of the (α*S*,β*R*)-GGE substrate. The mutant LigL-Asp95Ala showed 2.5X higher activity compared to wild-type LigL. The interaction between Pro188 and (α*S*,β*R*)-GGE substrate occurs via a hydrogen bond to the main chain carbonyl group of Pro188. Although the mutant LigL-Pro188Ala cannot eliminate this interaction it may inform the influence Pro188 has on substrate binding loop flexibility and consequently to the enzyme activity. The LigL-Pro188Ala mutant showed an approximately 60% increase in enzymatic activity compared to wild-type LigL. Finally, the LigL-Arg222Ala mutant completely abolished enzyme activity indicating this is a key residue for the substrate binding mechanism of LigL. The residue Arg222 is the only residue located at the substrate binding loop region that interacts direct with the (α*S*,β*R*)-GGE substrate. Sequence alignment against LigD and LigO showed that the Arg222 residue is not conserved (Figure 4-4). As LigL catalyzes the α(*S*)-configured substrates (α*S*,β*R*)-GGE and (α*S*,β*S*)-GGE and LigD and LigO catalyzes the α(*R*)-substrates (α*R*,β*R*)-GGE and (α*R*,β*S*)-GGE the difference on this particular position could be related to substrate stereoisomer recognition in this class of enzymes.

GSH-dependent β-etherases (LigE and LigF)

Structural Analysis – Attempts to solve the structure of full-length wild-type LigE (282 residues) and LigF (254 residues) were unsuccessful, but C-terminal truncation constructs of both proteins were successfully crystallized and used for structural analysis. Truncations of LigE and LigF were designed based on homology models generated by I-TASSER Online and disorder

predictions generated using PONDR (Predictor of Naturally Disordered Regions, Molecular Kinetics Inc., Indianapolis, IN) (Roy *et al.* 2010). LigEΔ255 and the LigEΔ255-GSH complex crystallized in the space group C2 with four molecules in the asymmetric unit with well-defined electron density for the bound GSH molecule. LigFΔ242-GSH crystallized in the space group $P6_322$ with one molecule in the asymmetric unit. Well-defined electron density corresponding to the GSH molecule is also visible in the structure.

GST superfamily members are multifunctional enzymes often involved in cellular detoxification processes via GSH-conjugation (Mathieu *et al.* 2012). However, some bacterial GSTs, are implicated in basal metabolism and supply bacterial cells with carbon sources (Allocati *et al.* 2009). Consistent with their predicted role in metabolism of β-ether-linked aromatic compounds, LigE and LigF each adopt the canonical GSTs domain fold with an N-terminal thioredoxin domain (residues 1-82 and 1-76 respectively) and a C-terminal α-helical domain (residues 93-255 and 93-242 respectively) connected by a short linker (residues 83-92 and 77-92 respectively) (Figure 4-8). In both LigE and LigF the thioredoxin domain is composed of four β-strands and three α-helices following the topology β1α1β2α2β3β4α3. The loop between β1 and α1 is longer in LigE than in LigF and occupies the space between the thioredoxin domain and the α-helical domain whereas in LigF this loop is moved away from the domain interface toward the surface of the thioredoxin domain. The loop between β2 and α2 is longer in LigF than in LigE, but both interact with the α-helical domain on the protein face opposite the linker (Figure 4-8). The C-terminal domains of both LigE and LigF are composed of 6 and 8 α-helices respectively. The RMSD between the monomers of LigE and LigF is 4.42Å indicating that, even though they catalyze very similar reactions, the enzymes display significant structural differences.

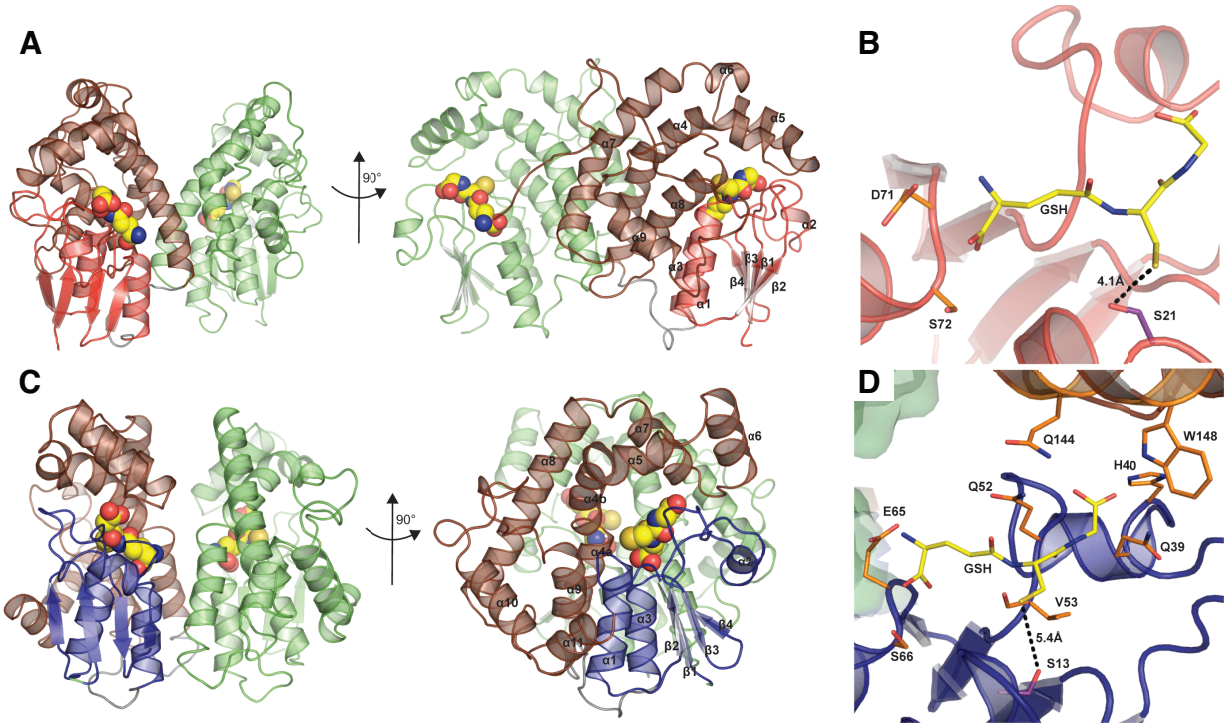

Figure 4-8. (A) Cartoon representation of the dimer of LigE, including the N-terminal thioredoxin domain (red), the C-terminal α-helical domain (brown) and the short linker (grey). Bound GSH is shown as yellow spheres. **(B)** Active site of LigE showing bound GSH (yellow sticks) and its interactions with residues D71 and S72 (orange sticks). Distance between the GSH sulfur and the catalytic serine 21 (purple sticks) is 4.1 Å. **(C)** Cartoon representation of the LigF dimer, including the N-terminal thioredoxin domain (blue), the C-terminal α-helical domain (brown) and the short linker (grey). Bound GSH is shown as yellow spheres. **(D)** Active site of LigF showing residues (in orange sticks) interacting with the γ-glutamyl (E65 and S66), cysteinyl (Q52 and V53) and glycine (Q144, H40, W148, and Q39) residues of bound GSH (yellow sticks). Distance between the GSH sulfur and the catalytic serine 13 (purple sticks) is 5.4 Å.

Biochemical data suggest that both LigE and LigF exist as dimers in solution and these dimers, related by two-fold symmetry, can be seen in the respective crystal structures. The dimer interface accounts for 1066Å^2 and 1092Å^2 of buried surface area in LigE and LigF, respectively (PISA European Bioinformatics Institute) (Krissinel and Henrick 2007). The LigE dimer forms via interactions between helices 4 and 7 of each monomer forming a pseudo four-helix bundle. The entire dimer interface of LigE is contained within the α-helical domain. In contrast, the LigF dimer forms via interactions between helices 3 and 4 of each monomer forming a four-helix bundle. The

dimer interface of LigF spans both the thioredoxin domain and the α-helical domain. The overall dimeric shape of both LigF and LigE were confirmed using small-angle X-ray solution scattering on both the truncated and full-length proteins. The protein envelopes determined by *ab initio* modeling align well with the crystal structures of both proteins, confirming the more elongated and extended dimer form of LigE as compared to LigF (Figure 4-9).

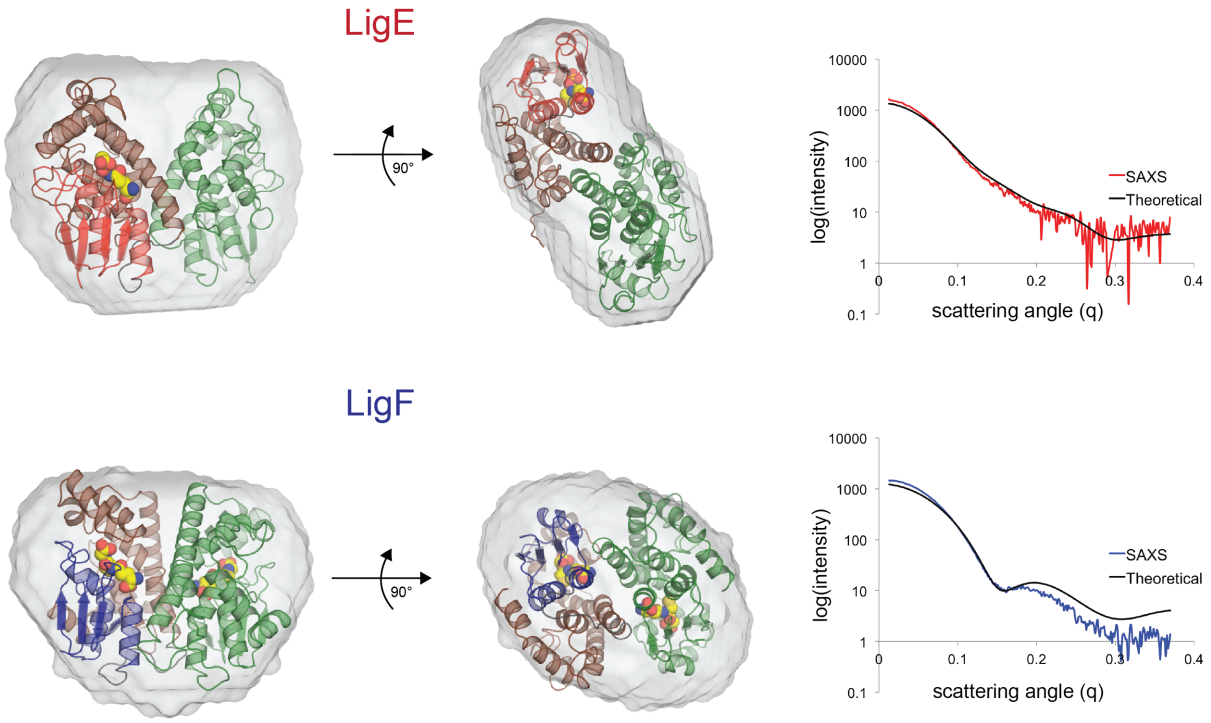

Figure 4-9. Cartoon representations of LigE (top) and LigF (bottom) with *ab initio* protein envelopes from SAXS data. One monomer of each dimer is colored to highlight the α-helical domain (brown), linker (gray), and thioredoxin domain (LigE-red, LigF-blue). The second molecule of the dimer is colored green. The protein envelopes shown were determined using DAMMIF *ab initio* modeling of SAXS data. The scattering angle (q) vs the log of the intensity of the scattering plots show the experimentally observed data (LigE-red, LigF-blue) and the theoretical scattering determined using CRYSOL from the X-ray structures of the dimers.

The active site of these GST family members is often located in a cleft between the thioredoxin domain and the α-helical domain. Both the LigE and LigF enzymes contain the ββα motif required for anchoring GSH in the active site (Armstrong 1997). In LigE Asp71 and Ser72, both located in the turn between β4 and α3, hydrogen bond with the amino and carboxylate groups, respectively, of the γ-glutamyl residue of the GSH molecule (Figure 4-8B). In LigF, Glu65 and Ser66 located in the turn connecting β4 and α3, recognize the γ-glutamyl moiety of GSH as part of the ββα motif (Figure 4-8D). Additionally, Gln52 and the backbone of Val53 interact with the cysteinyl moiety while Gln144, His40, Tyr148 and Gln39 anchor the glycine residue of the active site GSH molecule.

From analysis of a cross-validated σ_A weighted difference electron density map (mF_o-DF_c) contoured at 3.0 σ, the apo-LigE structure shows electron density located at the substrate-binding site. This can most likely be accounted for by a xenobiotic *E. coli* metabolite remaining from expression and purification. Based on these observations we propose a potential location for the native substrate-binding site at the highly hydrophobic region consisting of residues Tyr23, Phe45, Trp107, Phe115, Phe142 and Trp197. The aromatic rings of these hydrophobic residues are likely important in stacking interactions with the aromatic compounds from low molecular weight lignin derivative compounds. The LigEΔ255-GSH and LigFΔ242-GSH structures also revealed LigE-Ser21 (Figure 4-8B) and LigF-Ser13 (Figure 4-8D) as potential residues capable of deprotonating GSH. The proximities of the LigE-Ser21 and LigF-Ser13 hydroxyls with the GSH thiol (4.1 Å and 5.4 Å, respectively) support a hypothesis that the role of these serine residues in β-etherase catalysis is to deprotonate GSH, thereby activating the thiol for nucleophilic attack on the substrate's β-carbon. To further investigate the roles of LigE-Ser21 and LigF-Ser13 in β-etherase catalysis, variants LigE-S21A and LigFΔ242-S13A, in which serine residues were replaced with alanine, were expressed, purified, and tested for activity in the β-etherase assays (see below).

GSTs with greater than 40% sequence identity are traditionally considered to be in the same class, whereas proteins of different classes have typically less than 25% protein sequence identity (Sheehan *et al.* 2001). However, this classification is also based on a number of other considerations including structure, function and biochemical properties (Sheehan *et al.* 2001). Although there are seven classes of GSTs in mammals (alpha, mu, pi, sigma, theta, omega, and zeta), there is an ever-increasing number of non-mammalian classes including beta, chi, delta, epsilon, lambda, phi, and tau, as well as a number of more recently defined novel classes (Sheehan *et al.* 2001; Hayes *et al.* 2005; Wiktelius and Stenberg 2007). Based on structural properties, LigE is most similar to the fungal GSTFuA class (Mathieu *et al.* 2012), suggesting that the enzymes in this class are present in both prokaryotes and fungi. Other representatives in this class are from saprotrophic fungi suggesting a functional connection among the members of the class (Mathieu *et al.* 2013). While it has been suggested that LigF also belongs to the GSTFuA class (Mathieu *et al.* 2012), the dimer-interface present in the structure is inconsistent with others members of this class. Based on our data, LigF is best placed in a new structural class closely related to GSTFuAs or as a fungal Ure2p-like GST based on structural similarities and function in saprotrophic organisms, although it does not strictly fit the class (Thuillier *et al.* 2011).

Enzymatic Analysis and Mutagenesis

To further analyze the enzymatic activities of the GSH-dependent β-etherase enzymes, β-(1′-formyl-3′-methoxyphenoxy)-γ-hydroxypropioveratrone (FPHPV) degradation rates were measured by the accumulation of vanillin, a monoaromatic product of FPHPV cleavage (Figure 4-2C). Whereas β-etherase catalysis with MPHPV results in the release of guaiacol (Figure 4-1), vanillin is more easily detected by UV absorption, thus improving the sensitivity of the assays. In addition to LigE and LigF, we tested the rates of β-etherase catalysis for LigE variant LigE-S21A and two LigF variants, LigFΔ242 and LigFΔ242-S13A.

Consistent with previous reports that LigE and LigF catalyzed cleavage of (β*R*)-MPHPV and (β*S*)-MPHPV, respectively (Masai *et al.* 2003; Gall *et al.* 2014a), we found that LigE catalysis resulted in stereospecific (β*R*)-FPHPV cleavage (Gall *et al.* 2014b) whereas LigF and its variants selectively degraded the (β*S*)-FPHPV enantiomer. The effect of pH on β-etherase activities was determined for each enzyme, revealing that LigE, LigF, LigFΔ242, and LigFΔ242-S13A have pH optima at pH 8.0 (Figure 4-10). The activity of LigE was relatively unaffected by pH whereas, in contrast to previous reports of pH optimum 10.0 (Reiter *et al.* 2013), we observed that LigF, LigFΔ242, and LigFΔ242-S13A activity was significantly reduced above pH 8.0. Interestingly, LigFΔ242 exhibited higher rates of catalysis than the full-length LigF enzyme at all pH values, indicating that the predicted disordered region in the C-terminus may actually be inhibitory to β-etherase catalysis. We also found that β-etherase activities of LigE and LigF were reduced significantly upon mutation of residues LigE-Ser21 and LigF-Ser13, consistent with our structure-based predictions. The specific activity of LigE-S21A was 14% (Table 4-2) of the wild type LigE (*i.e.*, greater than a 7-fold reduction in activity) and the specific activity of LigFΔ242-S13A was less than 5% of LigFΔ242 (Figure 4-10, Table 4-2). These results support our hypothesis that LigE-Ser21 and LigF-Ser13 play critical roles in β-etherase catalysis and, based on structural observations, they are responsible for deprotonation of the GSH thiol.

Under optimal pH conditions, we found that LigF exhibited a higher affinity (*i.e.*, lower K_M) for its cognate substrate enantiomer ((β*S*)-FPHPV) than did LigE for (β*R*)-FPHPV (Table 4-2). Although the LigE and LigF k_{cat} were comparable at pH 8.0 (~31.9 s^{-1}), the higher affinity exhibited by LigF helps to explain why previous studies, in which activities were measured from initial substrate concentrations of less than 100 μM, reported a significantly higher specific activity for LigF than for LigE (Masai *et al.* 2003).

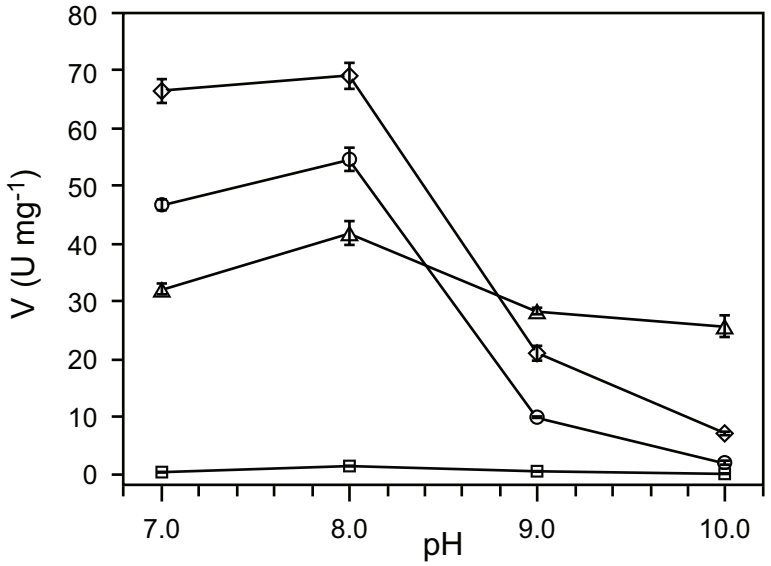

Figure 4-10. The effect of pH on β-etherase activities was determined for each enzyme, revealing that LigE, LigF, LigFΔ242, and LigFΔ242-S13A have pH optima at pH 8.0. Plotted as a function of pH (x-axis) are the specific enzymatic activities (y-axis) of β-etherases with either (β*R*)-FPHPV (LigE) or (β*S*)-FPHPV (LigF, LigFΔ242, and LigFΔ242-S13A) as the assay substrate (1.5 mM initial concentration). Error bars indicate the standard deviation of triplicate measurements. Legend: LigE (Δ), LigF (○), LigFΔ242 (◊), and LigFΔ242-S13A (□).

Because LigF-Ser13 was the only potential acid-base catalyst revealed in the active site of the LigFΔ242-GSH structure (Figure 4-8D), we hypothesized that an S_N2-type nucleophilic attack, which requires only a single basic residue for deprotonation of GSH, is responsible for β-etherase catalysis in LigF. To test this, SwissDock (Grosdidier *et al.* 2011a; Grosdidier *et al.* 2011b) was used to generate a LigFΔ242-GSH-(β*S*)-MPHPV complex (Figure 4-11) from the LigFΔ242-GSH structure, and a molecular model of (β*S*)-MPHPV. The LigFΔ242-GSH-(β*S*)-MPHPV complex revealed that, if deprotonated by Ser13 as expected, the resulting thiolate ion's position relative to the substrate β-carbon is in the appropriate orientation for an S_N2 attack.

Table 4-2. Kinetic parameters, determined from Michaelis-Menton curves for GSH-dependent β-etherases LigE, LigF, and their variants with substrates (β*R*)-MPHPV, (β*R*)-FPHPV, (β*S*)-F-FPHPV, and (β*S*)-FPHPV at pH 8.0.

Enzyme	Substrate	V_{MAX} (U mg^{-1})[a,b]	% WT activity with (β*S*)-FPHPV[c]	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($mM^{-1}\,s^{-1}$)
LigE	(β*S*)-FPHPV	59.73	-	31.91	554.1	57.59
LigE-S21A	(β*S*)-MPHPV	-	13.5[c]	-	-	-
LigF	(β*S*)-FPHPV	63.77	100	31.85	268.90	118.42
LigFΔ242	(β*S*)-FPHPV	69.34[b]	108.7	-	-	-
LigFΔ242-S13A	(β*S*)-FPHPV	1.45[b]	2.3	-	-	-
LigE	(β*S*)-F-FPHPV	0.02[b]	-	-	-	-
LigF	(β*R*)-F-FPHPV	NDA	-	-	-	-

[a] Abbreviation: NDA, no detectable activity

[b] Where noted (*i.e.*, in the absence of Michaelis-Menton curves), activity is reported as the velocity from biochemical assays in which the initial substrate concentration was 1.5 mM.

[c] Where noted, alternative substrate (β*R*)-MPHPV was used in biochemical assays with either LigE or LigE-S21A as catalysts

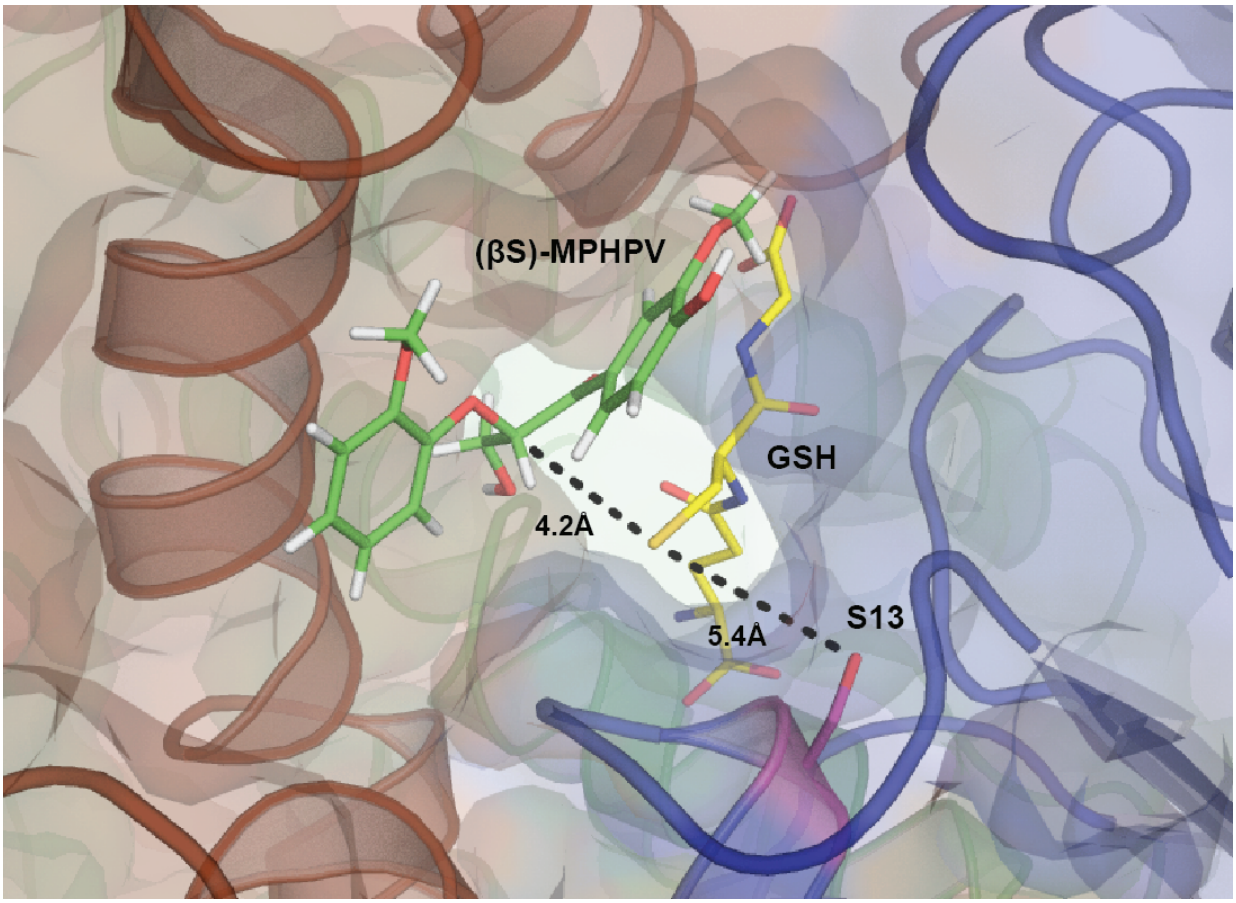

Figure 4-11. Model of ternary complex LigFΔ242-GSH-(β*S*)-MPHPV. Cartoon representations are shown of the N-terminal thioredoxin domain (blue) and the C-terminal α-helical domain (brown) with a transparent surface rendering. The bound glutathione (yellow) and docked (β*S*)-MPHPV (green) are shown as sticks, with interatomic distances between the catalytic serine 13 (purple), the GSH sulfur, and the β-carbon of MPHPV labelled in black.

Conversely, the LigEΔ255-GSH structure (Figure 4-8B) revealed several potential catalytic residues in the active site, leaving open the possibility that the LigE β-etherase mechanism involves additional acid-base reactions. Thus, in support of our hypothesis that LigE also catalyzes β-ether cleavage via S_N2 elimination, we sought to test whether or not the potential catalytic residues in the LigE active site requisitely carry out deprotonation of the proton bound to carbon β, the lone acidic proton of the LigE substrate, (β*R*)-FPHPV (Figure 4-2B). To test this possibility, we synthesized a model compound, (β*S*)-fluoro-(1′-formyl-3′-methoxyphenoxy)-γ-hydroxypropioveratrone ((β*S*)-F-FPHPV, Figure 4-2C) as an analog of the LigE substrate (β*R*)-FPHPV. (β*S*)-F-FPHPV and (β*R*)-FPHPV (despite their Cahn-Ingold-Prelog-derived *R*/*S* notations (Cahn *et al.* 1966)) have the same enantiomeric configuration with respect the orientation of their β-ether bonds and differ only in replacement of the hydrogen at carbon β in (β*R*)-FPHPV with a fluorine in (β*S*)-F-FPHPV, and this fluorine is predicted to prohibit deprotonation. We found that LigE catalyzed conversion of (β*S*)-F-FPHPV to vanillin and a GS-conjugated coproduct, albeit at a much lower velocity compared to cleavage of (β*R*)-FPHPV (Table 4-2)**,** exactly as predicted based on the hypothesis that an S_N2 catalytic mechanism would not involve deprotonation of the β-proton. Because of the low specific activity of the LigE enzyme with (β*S*)-F-FPHPV, we purified and analyzed the structural properties of the GS-conjugated coproduct by NMR spectroscopy to test if its spectrum was consistent with the predicted β(*S*)-fluoro-glutathionyl-γ-hydroxy-α-veratrylpropanone [(β*S*)-F-GS-HVP, Figure 4-2C], with the following results.

Previously, we characterized by NMR spectroscopy the two β-epimers of β-glutathionyl-γ-hydroxy-α-veratrylpropanone ((β*S*)-GS-HVP and (β*R*)-GS-HVP, (Figure 4-2B) that are produced from cleavage of (β*R*)-FPHPV and (β*S*)-FPHPV (respectively) and demonstrated that both LigE and LigF invert the chiral center at carbon β during catalysis, which is consistent with an S_N2 reaction

mechanism (Gall *et al.* 2014a). Using the same approach, we analyzed and compared the spectra of LigE-derived (β*S*)-F-GS-HVP to those of LigE-derived (β*S*)-GS-HVP, analyzing the 2-D ^{1}H-^{1}H COSY, ^{1}H-^{13}C HSQC, and ^{1}H-^{13}C HMBC NMR spectra of each compound (Figure 4-12). The ^{1}H-^{1}H COSY spectrum of (β*S*)-GS-HVP (Figure 4-12A) demonstrates the predicted Hβ-Hγ_a, Hβ-Hγ_b, and Hγ_a-Hγ_b correlations whereas, in the ^{1}H-^{1}H COSY spectrum of (β*S*)-F-GS-HVP (Figure 4-12B), only the Hγ_a-Hγ_b correlation is observed because (β*S*)-F-GS-HVP lacks a β-proton. Further, in the corresponding 1-D ^{1}H spectra, Hγ_a and Hγ_b each exhibits a doublet of doublet splitting pattern that is indicative of Hβ coupling for (β*S*)-GS-HVP (Figure 4-12A) or Fβ coupling for (β*S*)-F-GS-HVP (Figure 4-12B). The predicted Hβ-Cγ, Hγ_a-Cβ, Hβ-C3′ and H3′$_a$-Cβ correlations were observed in the ^{1}H-^{13}C HMBC spectrum of (β*S*)-GS-HVP (Figure 4-12C). In contrast, the ^{1}H-^{13}C HMBC spectrum of (β*S*)-F-GS-HVP (Figure 4-12D) revealed only Hγ_b-Cβ and H3′$_{a/b}$-Cβ correlations, demonstrating that the lack of a β-proton had abolished correlations to Cβ and C3′. Further, a comparison of the 1-D ^{13}C spectrum of (β*S*)-GS-HVP (Figure 4-12C) and the HMBC-projected ^{13}C spectrum of (β*S*)-F-GS-HVP (Figure 4-12D) shows the anticipated downfield shift caused by Cβ-fluorination as the chemical shifts of Cβ were 48.6 and ~107.7 ppm for (β*S*)-GS-HVP and (β*S*)-F-GS-HVP, respectively. Thus, we conclude that the LigE catalyzed β-ether cleavage of (β*S*)-F-FPHPV resulted in formation of the expected GS-conjugated product, (β*S*)-F-GS-HVP. Although it is unclear why this reaction was ~3,000 times slower than LigE-catalyzed cleavage of (β*R*)-FPHPV (Table 4-2), we hypothesize that the fluorine atom affects the β-ether bond angle and inhibits the approach of the thiolate ion for S_N2 elimination. It is possible that these effects were more pronounced in the active site of LigF as LigF showed no detectable activity with the (β*R*)-F-FPHPV enantiomer.

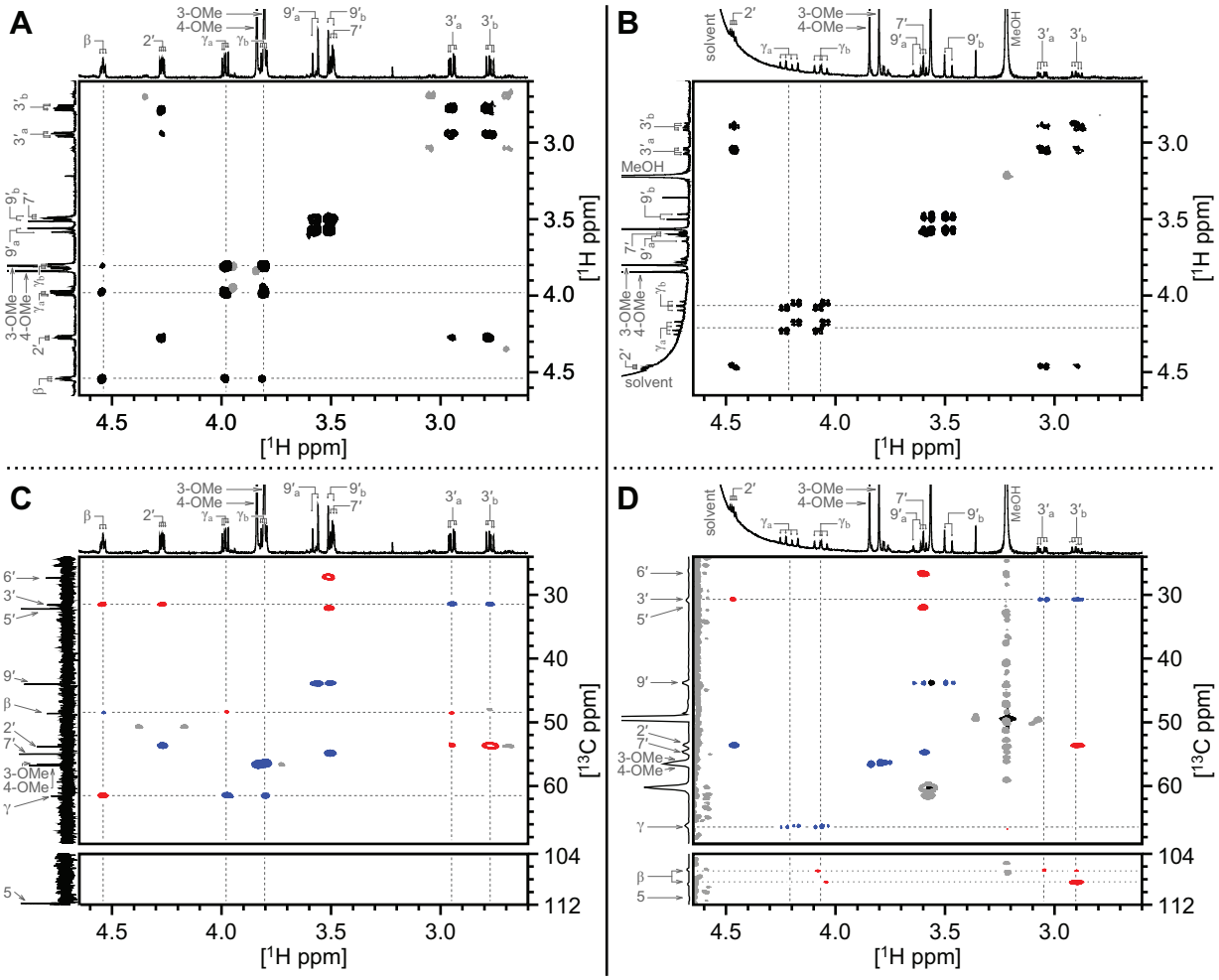

Figure 4-12. Partial 2D NMR spectra of β(*S*)-GS-HPV (panels **A** and **C**) and fluorinated analog β(*S*)-F-GS-HVP (panels **B** and **D**) in NMR solvent D_2O. Panels **(A)** and **(B)**, β(*S*)-GS-HVP and β(*S*)-F-GS-HVP (resp.), show the 2D ^{1}H-^{1}H COSY spectra (in black), where ^{1}H chemical shifts are plotted on the x- and y-axes (2.60–4.65 ppm); non-^{1}H-^{1}H-correlating COSY spectral regions are indicated (grey); dashed lines highlight the correlations between Hβ, Hγ_a, and Hγ_b. Panels **(C)** and **(D)**, β(*S*)-GS-HVP and β(*S*)-F-GS-HVP (resp.), show the 2D ^{1}H-^{13}C HSQC (in blue) and the 2D ^{1}H-^{13}C HMBC (in red) spectra, where ^{1}H chemical shifts are plotted on the x-axis (2.60–4.65 ppm) and ^{13}C chemical shifts are plotted on the y-axis (24.0–69.0 ppm); non-^{1}H-^{13}C-correlating HSQC spectral regions are indicated (in black); non-^{1}H-^{13}C-correlating HMBC spectral regions are indicated (in grey); dashed lines highlight the correlations between Hβ, Hγ_a, Hγ_b, H3′$_a$, H3′$_b$, Cβ, Cγ, and C3′. Proton assignment labels correspond with the carbon to which the proton is bound. Alphabetical subscripts differentiate two non-identical geminal protons. The chemical structures of β(*S*)-GS-HVP and β(*S*)-F-GS-HVP are shown in Figure 4-1 and Figure 4-2C, respectively.

GSH-dependent glutathione lyase (LigG)

Structural Analysis – We obtained apo-LigG and LigG-GS-AV crystals that diffracted to 1.1 Å and 1.4 Å resolution respectively. The crystal structure of LigG was solved using selenomethionine- (Se-Met)-labeled protein by single-wavelength anomalous dispersion (SAD) methods (Hendrickson 1991). LigG belongs to the omega class of GSTs that have a catalytic cysteine residue, as confirmed by loss of activity in the LigG-C15S variant (Meux *et al.* 2012). The LigG enzyme catalyzes cleavage of the (β*R*)-GS-HPV β-thioether bond and formation of GSSG (Figure 4-1). The structure of the LigG-GS-AV complex supports a reaction mechanism in which a disulfide bond is formed between sulfur atoms of the GS-AV substrate analog and the catalytic residue Cys15, releasing the aromatic portion of the substrate. GSH enters the active site and a disulfide bond exchange takes place with formation and release of GSSG, restoring the enzyme to a substrate-accepting state (Tocheva *et al.* 2006; Allocati *et al.* 2009).

The LigG structure possesses the canonical GST domain fold with an N-terminal thioredoxin domain (β1α1β2α2β3β4α3) and a C-terminal α-helical domain composed of 6 α-helices (Figure 4-13A). A comparison between the apo-LigG form and that of a LigG-GSH complex (Meux *et al.* 2012) revealed a conformational change upon GSH binding at the site located in the loops between β1/α1, β2/α2 and α2/β3 of the N-terminal domain (Figure 4-13B). In order to obtain the LigG structure in a complex with the substrate analog GS-AV (Figure 4-13C), ~1-year old crystals of apo-LigG were used for soaking experiments. These older crystals show oxidation of the catalytic Cys15 that rendered the enzyme inactive. The active site, located between the two domains, shows the glutathionyl moiety of the GS-AV substrate sitting on the top of the four β-strands of the N-terminal thioredoxin domain and the aromatic ring of the substrate contacting residues of α4, α8 and a cap loop region $_{220}$GGGNG$_{224}$ from the C-terminal α-helical domain (Figure 4-13D). The sulfur atom of the catalytic residue Cys15, located in the loop between β1/α1, was observed in two distinct

positions with distances of 3.6 Å and 4.7 Å from the sulfur atom of GS-AV. The glutathionyl moiety of the substrate contacts the loops of the thioredoxin-like domain β1/α1 (residues Ile12, Cys15 and Phe17), α2/β3 (residues Thr55, Ala56, Leu57 and Pro58) and β4/α3 (residues Glu70 and Ser71). The aromatic portion of the substrate contacts the C-terminal α-helical domain via residues Ser109, Tyr113 and Leu117 located in α4 and the residues Tyr217 and Asn223 located in α8 and cap loop respectively. Interestingly, the residue Tyr113 in α4 interacts via π–π stacking with the aromatic ring of the substrate (Figure 4-13D).

Enzymatic Analysis and Mutagenesis

GSH-dependent glutathione lyase (LigG) – The thiol transferase activity of LigG was probed with two substrates, GS-AV and (βR)-GS-HPV at pH 9.0 (Figure 4-1). As reported previously, LigG was unreactive towards (βS)-GS-HPV and the mutation of Cys15 to serine abolished enzyme activity with (βR)-GS-HPV (Meux et al. 2012). Interestingly, the KM measured for GS-AV was ~12-fold higher than for (βR)-GS-HPV, while the kcat was unchanged (Table 4-3). These results suggest that the active site of LigG has evolved to tightly bind the GS-conjugated product from LigF, (βR)-GS-HPV.

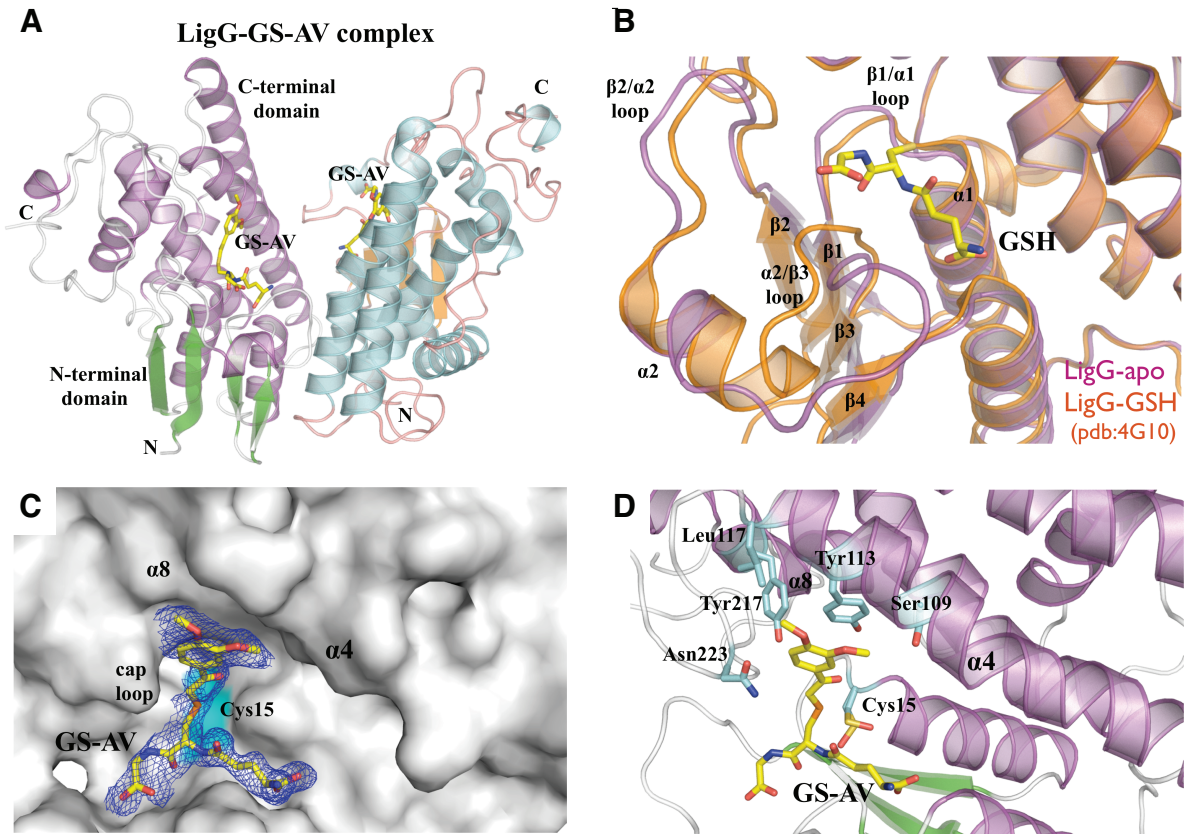

Figure 4-13. (A) Overall cartoon representation of the LigG-GS-AV complex dimer. **(B)** Superposition of the GSH binding site of apo-LigG (magenta) and LigG-GSH (PDB ID 4G10) (Meux *et al.* 2012) (orange) structures. Significant conformational changes of the GSH binding-site were observed on the loop regions at N-terminal domain connecting the β1/α1, β2/α2 and α2/β3 structural elements. **(C)** Molecular surface representation of the LigG monomer in complex with the GS-AV substrate analog. A $2mF_o$-DF_c electron density map around the substrate analog GS-AV contoured at 0.8 σ is shown in blue. **(D)** Active site of the LigG-GS-AV complex. The aromatic region of the substrate contacts C-terminal α-helical domain via residues Ser109, Tyr113 and Leu117 on α4 and residues Tyr217 and Asn223 on α8 and the cap loop, respectively.

Table 4-3. Kinetic parameters, determined from Michaelis-Menton curves for GSH-dependent glutathione lyase LigG with substrates GS-AV and (β*R*)-GS-HPV at pH 9.0.

Enzyme	Substrate	V_{MAX} (U mg^{-1})[a]	% WT activity with (β*R*)-GS-HPV[b]	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($mM^{-1}\ s^{-1}$)
LigG	GS-AV	83.7	105[b]	29.1	204.6	142.38
LigG	(β*R*)-GS-HPV	79.5	100	27.7	16.2	1709.88
LigG-I12A	(β*R*)-GS-HPV	30.3	38	-	-	-
LigG-I12V	(β*R*)-GS-HPV	58.8	74	-	-	-
LigG-N223A	(β*R*)-GS-HPV	17.8	22	-	-	-
LigG-N223G	(β*R*)-GS-HPV	31.6	40	-	-	-
LigG-N223V	(β*R*)-GS-HPV	2.3	3	-	-	-
LigG-Y217F	(β*R*)-GS-HPV	0.1	< 1	-	-	-
LigG-Y113F	(β*R*)-GS-HPV	0.3	< 1	-	-	-
LigG-Y214F	(β*R*)-GS-HPV	0.1	< 1	-	-	-
LigG-C15S	(β*R*)-GS-HPV	NDA	-	-	-	-

[a] Abbreviation: NDA, no detectable activity

[b] Where noted, alternative substrate GS-AV was used in biochemical assays

From inspection of the LigG-GS-AV crystal structure, Ile12, Tyr113, Tyr214, Tyr217 and Asn223 appear to interact with the substrate in either the first or second coordinating residue shell. We investigated the contribution of these residues to LigG activity and stereoselectivity via a series of mutations (Table 4-3). None of the variants displayed activity towards (β*S*)-GS-HPV, and all had compromised activity relative to the wild-type enzyme. Mutation of Tyr113, Tyr214 or Tyr217 to phenylalanine abolished enzyme activity suggesting that the hydroxyl groups on these three residues play an important role in substrate binding and/or active site organization.

Discussion

The biocatalytic breakdown of lignin compounds represents a major potential source of aromatic products that would be valuable for the chemical, food and pharmaceutical industry (Bugg *et al.* 2011a). In contrast to known fungal systems, the bacterium *Sphingobium sp.* strain SYK-6 possesses an enzymatic route to the breakdown of lignin components that is independent of chemical mediators and requires common cellular cofactors such as pyridine nucleotides, and glutathione. Our combined structural and enzymological studies of the β-aryl ether cleavage pathway enzymes provide insights into the features important for substrate and cofactor binding and catalysis. This opens up the possibility of directing these catalysts for the production of valuable commodities from lignin by changing substrate specificities, or for the enhancement of lignin degradation and release of sugars from lignocellulosic feedstocks.

The Cα-dehydrogenase crystal structures show that the catalytic mechanisms of LigD, LigO, and LigL are likely similar to those of other SDR superfamily member enzymes, where Cα-oxidation of GGE commences with base-catalyzed deprotonation of the Cα-hydroxyl group, followed by ketone formation coupled with hydride transfer from Cα to NAD^+, resulting in the coproducts NADH and MPHPV. Similar to other SDRs, basic conditions were optimal for the Cα-dehydrogenation reactions indicating that Cα-hydroxyl deprotonation is a rate-limiting step during catalysis.

Further, our observations with LigG are consistent with other omega-class GST member enzymes. The loss of lyase activity in the LigG-C15S variant is consistent with the role of Cys15 in the LigG catalytic cycle, where a disulfide bond is formed between the Cys15 thiol and the glutathionyl moiety of GS-HPV thereby releasing HPV, and subsequently, a deprotonated GSH molecule is used to cleave the GS-enzyme disulfide linkage, yielding GSSG and restored LigG.

Previously, it was reported that the β-ether-cleaving reactions of LigE and LigF caused inversion of the chiral center at carbon β (*i.e.*, β(*R*)- and β(*S*)-substrates are converted to β(*S*)- and

β(*R*)-products, respectively) (Gall *et al.* 2014a). However, we were unable to link the reactions facilitated by LigE and LigF to those of other known GST family member enzymes. For LigF, we hypothesized an S_N2-type reaction mechanism based on our observations in the LigFΔ242-GSH crystal structure, where LigF-Ser13 was identified as the only potentially basic residue in the active site. This hypothesis was supported by our finding that the LigFΔ242-S13A variant had reduced β-etherase activity and its level of inactivation was more severe at lower pH, suggesting that the role of LigF-Ser13 is to deprotonate the GSH thiol, thereby activating it for S_N2 nucleophilic attack on the β-carbon of the substrate. Given that inversion of the chiral center undergoing attack is requisite of an S_N2 reaction we surmise that the likely reaction mechanism of LigF is S_N2. Although the existence of several potentially catalytic residues in the active site of the LigEΔ255-GSH structure suggest that the LigE mechanism may be different from that of LigF, we find that, of the potential LigE mechanisms proposed, only S_N2-catalysis is consistent with our observations. This is supported by our observation that the mutation of LigE-Ser21 resulted in significantly reduced enzymatic activity. Further, because no nucleophilic residues could be identified in the predicted active site of the LigEΔ255-GSH structure near the proposed hydrophobic binding pocket, and because LigE catalyzed the conversion of (β*S*)-F-FPHPV (Figure 4-2C) to (β*S*)-F-GS-HVP (Figure 4-12), we conclude that the LigE mechanism is unlikely to involve formation of an enzyme-substrate adduct and does not involve Cβ-deprotonation or substrate enolization. From these observations, we surmise that S_N2-catalysis, or possibly a yet-to-be-identified reaction mechanism, is responsible for β-ether cleavage by LigE.

The structures of the LigE and LigF enzymes also highlight the nature of stereospecific control that is key to this pathway. These enzymes possess dramatically different structural arrangements within the monomers, reflected in very different dimer shapes. As a result the substrate binding surfaces of the two enzymes are on opposite faces of the thioredoxin domain and glutathione binding. This means that if a substrate with the wrong stereochemistry were to bind, it would not

be within reach of the catalytic residues, hence controlling stereospecificity. Additionally, there is no simple set of mutations within the respective active sites that would switch substrate specificity.

The detailed structural and biochemical characterization of the members of the β-aryl etherase pathway in this study reveals important new aspects of the enzyme mechanism and the determinants of substrate stereospecificity. Future enzyme engineering studies enabled by these results may focus on optimizing the pathway for specific lignin-derived subunits available from the processing of biomass into suitable products for use as advanced biofuels and/or renewable chemicals, and could lead to the development of enzymatic pathways for the creation of specific aromatic units of defined composition for further biochemical or chemical conversion.

Materials and Methods

Cloning, expression and purification of LigD, LigO, LigL, LigE, LigF and LigG

Gene Cloning – LigD, LigO, LigL, LigE and LigG were synthesized and cloned into a custom vector (pCPD) assembled by Genscript (Piscataway, NJ). This vector combined the pVP16 backbone (provided by the Center for Eukaryotic Structural Genomics, Madison, WI) with the gene of interest and a C-terminal fusion protein tag comprised of the Vibro cholera MARTX toxin cysteine protease domain (CPD) (Shen *et al.* 2009); see Supplemental Material for details of the constructs. During protein purification, the CPD tag can be activated by the addition of inositol hexaphosphate ($InsP_6$), cleaving at a leucine positioned in-between the protein of interest and CPD. The pVP80K_LigFΔ242 vector was prepared using polymerase incomplete primer extension as previously described using Phusion High-Fidelity PCR master mix with HF buffer (New England Biolabs Inc, Ipswich, MA), and primers from Integrated DNA Technologies (Coralville, IA) (Klock *et al.* 2008). The pVP80K vector was provided by the Center for Eukaryotic Structural Genomics (Madison, WI) and the pVP102KSSLigF vector containing full-length wild type LigF was prepared as previously described (Gall *et al.* 2014a). Insert and vector backbone PCR products were mixed 1:1 and immediately

transformed into *Escherichia coli* One Shot® TOP10 cells (Invitrogen, Carlsbad, CA). The pVP80K_LigFΔ242 vector was purified from *E. coli* (One Shot® TOP10, 10 mL LB with kanamycin, 18 h at 37 °C) using the QIAperep® Spin Miniprep Kit (Qiagen, Germantown, MD) and transformed into the lab strain *E. coli* B834(DE3) Z-competent cells (Zymo Research, Orange, CA).

Enzyme Purification – NEB Express protein expression cells (New England Biolabs Inc, Ipswich, MA) containing either pCPD-LigD, -LigO, -LigL, -LigE, and -LigG were grown in auto-inducing selenomethionine media as previously described (Sreenath *et al.* 2005) and harvested via centrifugation. Harvested cells were resuspended in 30 mL lysis buffer (50 mM HEPES buffer, pH 7.4, 150 mM NaCl and 40 mM imidazole) and lysed by an Avestin EmulsiFlex-C3 homogenizer. The C-terminally His-tagged proteins were purified from the clarified supernatant using pre-charged nickel-IMAC resin (GE Healthcare, Piscataway, NJ). After protein binding and washing twice with lysis buffer, $InsP_6$ was added to a final concentration of 200 μM. Note that the $InsP_6$ was first diluted to 10 mM in lysis buffer to neutralize the acidic pH of the stock solution. After 1 h of incubation, the resin was washed with 1 mL of lysis buffer to elute the cleaved protein. Following buffer exchange into 20 mM Tris pH 8, proteins were further purified using a HiTrap Q HP anion exchange column. Pooled fractions containing the protein of interest, as confirmed by SDS-PAGE were pooled and concentrated. Final protein clean up was done using gel filtration on a Superdex 200 10/300 GL column (GE Healthcare, Piscataway, NJ).

Lab strain *E. coli* B834(DE3) Z-competent cells (Zymo Research, Orange, CA) containing the pVP80K_LigFΔ242 plasmid were grown in auto-inducing selenomethionine media as previously described (Sreenath *et al.* 2005) and harvested via centrifugation. Harvested cells were resuspended in 20 mL lysis buffer (20 mM sodium phosphate buffer, pH 7.5, 500 mM sodium chloride, 20% ethylene glycol) and lysed by sonication. The N-terminally His-tagged LigFΔ242 fusion protein was purified from the supernatant by immobilized nickel affinity chromatography

(IMAC) using a HiTrap Q HP anion exchange column on an ÄKTA FLPRC (GE Healthcare, Piscataway, NJ). Fractions containing LigFΔ242, as determined by SDS-PAGE, were combined and dialyzed overnight at 4 °C. LigFΔ242 was cleaved from the fusion protein using TEV protease (1 mg/100 mg protein, provided by the Center for Eukaryotic Structural Genomics). Following cleavage LigFΔ242 and the poly-histidine tag were separated using a HiTrap Q HP anion exchange column. Pooled fractions containing LigFΔ242, as confirmed by SDS-PAGE were pooled and concentrated to 3 mL. Final protein clean up was done using gel filtration on a HiLoad™ 26/60 Supradex™ 200 prep grade column.

Enzyme kinetic assays

NAD^+-dependent dehydrogenation assays – *In vitro* NAD-dependent dehydrogenase assays with LigL, were preformed in an aqueous assay buffer (25 mM phosphate (pH 6.5–7.5) or Tris (pH 7.5–9.0) and 5 mM NAD) at 30 °C with substrate concentrations ranging from 6.25 mM to 100 mM and an enzyme concentration of 3.125 nM. Enantiopure preparations of (α*S*, β*S*)-GGE and (α*S*,β*R*)-GGE were synthesized as described previously (Hishiyama *et al.* 2012). After incubating the enzyme with substrate for 10 min, the reaction was quenched by the addition of 1 volume of 5% formic acid in water. Each sample was then subjected to reverse-phase HPLC using a Kinetex™ 5 µm Phenyl-Hexyl 100 Å, LC Column 250 x 4.6 mm attached to an Agilent 1100 HPLC. Samples and external standards were quantified by UV absorption at 280 and 300 nm. The HPLC mobile phase was a mixture of aqueous (0.1% formic acid in water) and acetonitrile (0.1% formic acid) at a flow rate of 1 mL min^{-1}. MPHPV and GGE concentrations were quantified in order to calculate the specific activity of each reaction. The average of the triplicate assays were reported.

GSH-dependent β-etherase assays – *In vitro* β-etherase assays with LigE, LigF, LigFΔ242, and LigFΔ242-S13A were conducted in an aqueous assay buffer (25 mM Tris, 2.5% DMSO, 5 mM GSH, pH 7.0–10.0) at 30 °C with an initial substrate concentration of 1.5 mM and enzyme concentrations

of either 160 nM (LigE), 170 nM (LigF), 180 nM (LigFΔ242), 3.9 μM (LigFΔ242-S13A), 11.2 μM (LigE with (β*S*)-F-FPHPV), or 12.0 μM (LigF with (β*R*)-F-FPHPV). Enantiopure preparations of (β*R*)-FPHPV and (β*S*)-FPHPV were obtained from chiral chromatographic separation of the parent racemate as previously described (Gall *et al.* 2014b). Racemic F-FPHPV was prepared via organic synthesis (Figure 4-14), chiral chromatography was used for the separation of (β*S*)-F-FPHPV and (β*R*)-F-FPHPV (Supplementary Information), and (β*S*)-F-FPHPV was used as a substrate in the LigE assays. Synthetic details for enzymatic substrates FPHPV and F-FPHPV are described in the Supplementary Information.

Michaelis-Menton curves were generated by measuring the enzymatic specific activities over a range of initial substrate concentrations (1.50, 1.25, 1.00, 0.75, 0.50, and 0.25 mM) obtained from serial dilution of a 1.5 mM substrate buffer made immediately prior to conducting the assays. The 1 mL assays were conducted in triplicate and were managed as follows: (1) the substrate was dissolved in DMSO at 60 mM and 25 μL were added to a 2 mL vial, (2) 875 μL of 25.7 mM Tris pH X is added (where X is higher than the intended pH of the assay to account for the acidic effect of GSH – *e.g.*, pH X = 11.5 drops to pH 8.0 after addition of 5 mM GSH), (3) 50 μL of 100 mM GSH is added (100 mM GSH stock solution is prepared by adding GSH to 25 mM Tris (pH X)), (4) 50 μL of 20-times concentrated enzyme is added, (5) 150 μL samples were collected after 0, 6, 12, 18, 24, and 30 s of incubation and enzymatic activity was abolished by pipetting each sample into 5 μL of 5 M phosphoric acid, and (6) the remaining reaction volume was used to measure the pH of the mixture with pH paper.

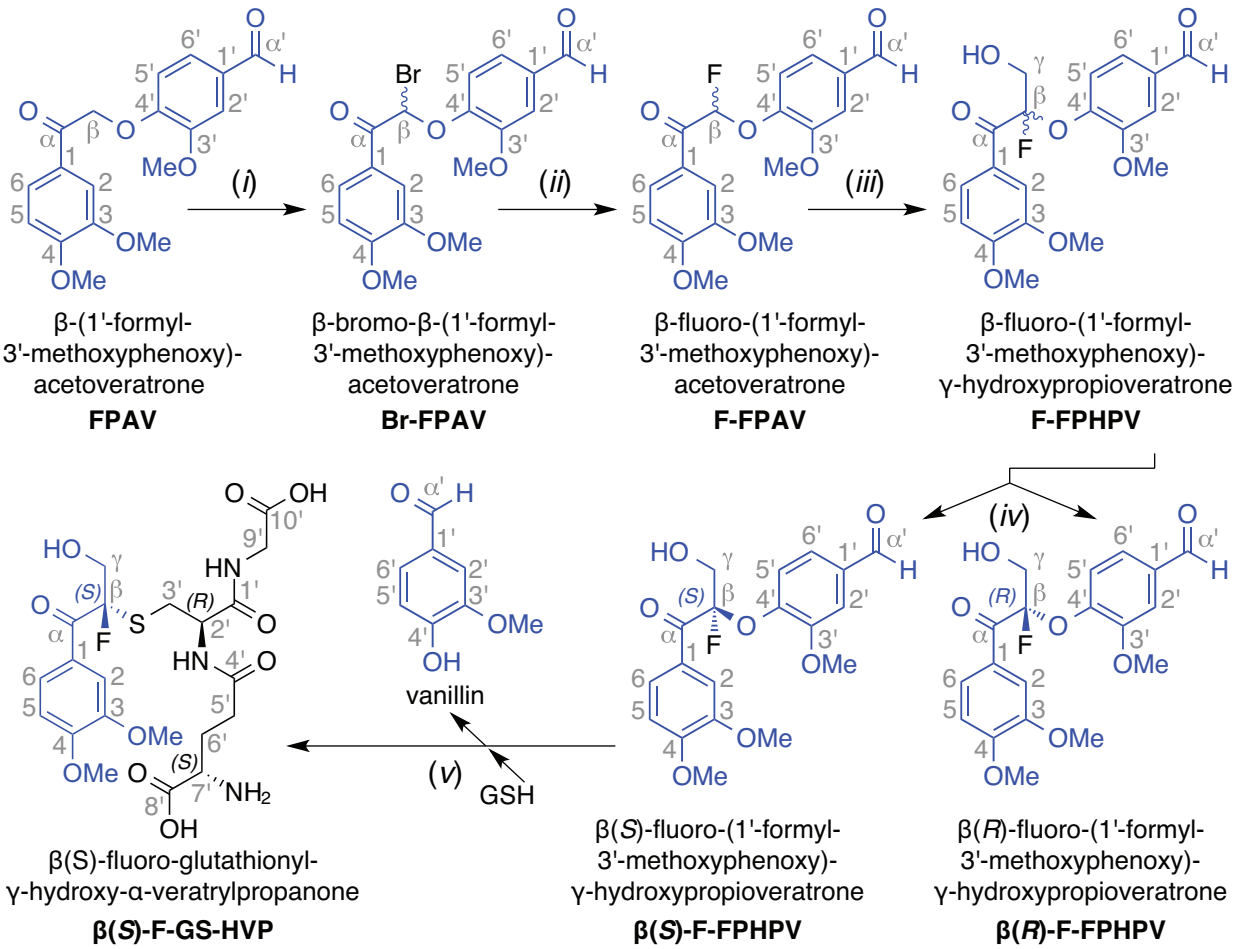

Figure 4-14. Scheme for organic synthesis of (*i*) racemic β-bromo-(1'-formyl-3'-methoxyphenoxy)-acetoveratrone (Br-FPAV), (*ii*) racemic β-fluoro-(1'-formyl-3'-methoxyphenoxy)-acetoveratrone (F-FPAV), (*iii*) racemic β-fluoro-(1'-formyl-3'-methoxyphenoxy)-γ-hydroxypropioveratrone (F-FPHPV), (*iv*) chiral chromatographic separation of enantiomers β(*S*)-F-FPHPV and β(*R*)-F-FPHPV, and (*v*) enzymatic synthesis of β(*S*)-fluoro-glutathionyl-γ-hydroxy-α-veratrylpropanone (β(*S*)-F-GS-HVP). Reagents and conditions: (*i*) pyridinium tribromide, EtOAc, 90 min, 68%; (*ii*) Ag(I)F, acetonitrile, 18 h, 90%; (*iii*) formaldehyde, K_2CO_3, 1,4-dioxane, 18 h, 23%; (*iv*) chiral chromatography, ethanol/hexane; (*v*) LigE, glutathione, 25 mM Tris in H_2O, pH 8.0, 18 h, C_{18} chromatography, H_2O/methanol.

Each sample was then subjected to C_{18}-reversed phase HPLC using a Beckman 125NM solvent delivery module equipped with a Beckman 168 UV detector. Samples and external standards were quantified by UV absorption at 280 nm. The HPLC mobile phase was a mixture of aqueous (5 mM formic acid in 95/5 water/acetonitrile) and methanol at a flow rate of 1.0 mL min^{-1}. The ratio of

buffers was adjusted as follows: 0–6 min, 30% methanol; 6–15 min, gradient from 30–80% methanol; 15–25 min, 80% methanol; 25–26 min, gradient from 80–30% methanol; 26–33 min, 30% methanol. Vanillin concentrations were quantified for each time point and a linear regression was generated over the 30 s assay period in order to calculate the specific activity of each reaction. The average of the triplicate assays were reported.

GSH-dependent glutathione lyase assays – *In vitro* GSH-dependent glutathione lyase assays with LigG, were preformed in an aqueous assay buffer (25 mM Tris pH 9 and 5 mM GSH) at 30 °C with substrate concentrations ranging from 3.125 mM to 50 mM and an enzyme concentration of 1 nM. The GS-AV substrate was synthesized as described previously (Gall *et al.* 2014a). Enantiopure preparation of (β*R*)-GS-HPV was synthesized by reacting a known quantity of enantiopure (β*S*)-MPHPV with LigF until the reaction reached completion (verified by the disappearance of (β*S*)-MPHPV via HPLC) after which LigF was heat inactivated (verified by the absence of catalytic activity upon addition of fresh (β*S*)-MPHPV). After incubating the enzyme with substrate for 10 min, the reaction was quenched by the addition of 1 volume of 5% formic acid in water. Samples were then measured by HPLC as described above.

Crystallization of LigD, LigO, LigL, LigE, LigF and LigG

LigD, LigO, LigL, LigE and LigG were concentrated to 10 to 15 mg mL^{-1} and dialyzed against 20 mM HEPES pH 7.4 and 50 mM NaCl. The LigF was dialyzed in 10mM HEPES buffer, pH 7.5 containing 50 mM sodium chloride, 0.5 mM TCEP and 1 mM GSH, and concentrated to 18.5 mg mL^{-1}. The LigD, LigO, LigL, LigE, LigF and LigG proteins were screened using the sparse matrix method (Jancarik and Kim 1991) with a Phoenix Robot (Art Robbins Instruments, Sunnyvale, CA) and a Mosquito dispenser (TTP LabTech, Melbourn, United Kingdom) utilizing the following crystallization screens: Berkeley Screen (Lawrence Berkeley National Laboratory), Crystal Screen, SaltRx, PEG/Ion, Index and PEGRx (Hampton Research, Aliso Viejo, CA), JSCG-plus HT-96 and PACT

premier HT-96 (Molecular Dimensions, Altamonte Springs, FL). The optimum conditions for crystallization of the different pathway proteins were found as follows: LigD, 0.1 M Hepes pH 7.5 and 1.5 M lithium sulfate; LigO, 0.1 M ammonium citrate, 0.1 M MES pH 5.5, 20% PEG 3,350 and 5% isopropanol; LigL, 0.2 M magnesium chloride, 0.1 M Hepes pH 7.5 and 25% PEG 3,350; LigE, 0.1 M ammonium citrate, 0.1 M MES pH 5.5, 20% PEG 3,350 and 5% isopropanol; LigF, 25% polyethylene glycol monomethyl ether 2000, 0.25 M TMAO, and 0.1 M Tris, pH 8.5; and LigG, 0.1 M Bis-Tris propane pH 7.0 and 1.5 M ammonium sulfate.; and LigG, 0.1 M Bis-Tris propane pH 7.0 and 1.5 M ammonium sulfate. LigO, LigL, LigE, and LigG crystals were obtained after 2 to 7 days by the sitting-drop vapor-diffusion method with the drops consisting of a mixture of 0.2 μL of protein solution and 0.2 μL of reservoir solution. LigD crystals were obtained after 45 days. LigF crystals were obtained in less than 24 h with drops containing a mixture of 1 μL of protein solution, 0.8 μL of reservoir solution and 0.2 μL of seed crystals (pulverized LigFΔ242 crystals in 0.2 M magnesium formate, 30% polyethylene glycol 3350, and 1 mM GSH).

X-ray data collection and structure determination

The crystals of LigD, LigO, LigL, LigE, and LigG was placed in a reservoir solution containing 10 to 20% (*v/v*) glycerol**,** and then flash-cooled in liquid nitrogen. The X-ray data sets for LigD, LigO, LigL, LigE and LigG were collected at the Berkeley Center for Structural Biology beamLines 8.2.1 and 8.2.2 of the Advanced Light Source at Lawrence Berkeley National Laboratory (LBNL). LigF crystals were cryoprotected with a reservoir solution containing 30% polyethylene glycol monomethyl ether 2000 and 1 mM GSH. X-ray diffraction data were collected at the Life Sciences Collaborative Access Team (LS-CAT) with X-ray wavelength 0.9793 at the Advanced Photon Source at Argonne National Laboratory, USA. Data sets were indexed and scaled using HKL2000 (Otwinowski and Minor 1997). LigD, LigO, LigL, and LigF crystal structure were determined by the molecular-replacement method with the program *PHASER* (McCoy *et al.* 2007) within the *Phenix*

suite (Adams *et al.* 2010). LigD was solved using as a search model the PDB ID coordinates of 3IOY, which has 35% sequence identity with the target. LigO and LigL crystal structures were solved using the LigD coordinates as the search model. The structure of LigF solved was using the coordinates of a LigF homologue (Lig37), whose sequence was identified from a metagenomic analysis of a rice-straw enriched compost microbial community from Berkeley, CA (Reddy *et al.* 2013; Simmons *et al.* 2014). The atomic positions obtained from molecular replacement and the resulting electron density maps were used to build the LigD, LigO, LigL, and LigF structures and initiate crystallographic refinement and model rebuilding. The crystal structure of LigE and LigG were solved using selenomethionine (Se-Met) labeled protein by single-wavelength anomalous dispersion (SAD) methods (Hendrickson 1991) with the *phenix.autosol* (Terwilliger *et al.* 2009) and *phenix.autobuild* (Terwilliger *et al.* 2008) programs. Structure refinement was performed using the *phenix.refine* program (Afonine *et al.* 2009). Manual rebuilding using *COOT* (Emsley and Cowtan 2004) and the addition of water molecules allowed construction of the final models. Root-mean-square deviation differences from ideal geometries for bond lengths, angles and dihedrals were calculated with *Phenix* (Adams *et al.* 2010). The overall stereochemical quality of all final models was assessed using the program *MOLPROBITY* (Davis *et al.* 2007).

Small-Angle X-Ray Scattering

LigE and LigF were dialyzed for 15 h at 4 °C into buffer containing 10 mM HEPES pH 7.5, 50 mM sodium chloride, 1 mM GSH, and 0.5 mM TCEP. Prior to data collection samples were filtered through a 0.2 μm syringe filter and diluted to the working concentrations. After dilution, samples were clarified via centrifugation. The buffer blank was also syringe-filtered and clarified by centrifugation. Small-angle scattering data was collected on a Bruker NANOSTAR X-ray generator located at the National Magnetic Resonance Facility at the University of Wisconsin–Madison. Three data collections of 1 h each were taken for each sample and buffer. Data was merged and indexed

using the Bruker NANOSTAR Small-Angle X-ray Scattering System software (Bruker AXS, Madison, WI). The scattering intensity was obtained by subtracting the scattering of the buffer blank from the sample scattering using the PRIMUS software (Konarev *et al.* 2003). All SAXS data was processed using GNOM, integrated in the PRIMUS software, to obtain the pair distance distribution function (PDDF) (Svergun 1992). The GNOM output was used with DAMMIF to calculate 10 *ab initio* dummy atom models (Franke and Svergun 2009). Models were averaged using DAMAVER and aligned to X-ray crystal structures using SUPCOMB (Kozin and Svergun 2001; Volkov and Svergun 2003). Theoretical scattering curves for the X-ray crystal structure of LigE and a model of the dimer of LigF were calculated using CRYSOL (Svergun *et al.* 1995).

Molecular docking

Docking of MPHPV to the LigFΔ242-GSH structure was performed using the SwissDock server (http://www.swissdock.ch) (Grosdidier *et al.* 2011a; Grosdidier *et al.* 2011b). Docking was performed using the 'Accurate' parameter and otherwise default parameters, with the search space limited to a $10 \times 10 \times 10$ Å^3 region around the GSH binding. Both the protein and the MPHPV ligand were rigid during docking. The structure of MPHPV was built in ChemDraw (Mills 2006), converted to 3D coordinates using OpenBabel (O'Boyle *et al.* 2011). Docking results were visualized and screened using the UCSF Chimera molecular modeling system (www.cgl.ucsf.edu/chimera) (Pettersen *et al.* 2004).

Acknowledgments

This work was performed through collaboration between the Joint BioEnergy Institute (http://www.jbei.org), and the Great Lakes Bioenergy Research Center (http://www.glbrc.org). The Joint BioEnergy Institute (http://www.jbei.org) is supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-

05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. The Great Lakes Bioenergy Research Center is supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through grant DE-FG02-07ER64495. We are grateful to the staff of the Berkeley Center for Structural Biology at the Advanced Light Source of Lawrence Berkeley National Laboratory. The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health, National Institute of General Medical Sciences. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We thank Dr. Brian G. Fox and Dr. Christopher M. Bianchetti for helpful discussion, and the NIH-funded Center for Eukaryotic Structural Genomics for general access to computers and equipment for the structural work. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817). Use of GM/CA@APS Sector 23 was supported by Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). This study made use of the National Magnetic Resonance Facility at Madison, which is supported by NIH grant P41GM103399 (NIGMS). SAXS studies were supported by funds from NIH shared instrumentation grant S10RR027000 and the University of Wisconsin-Madison.

Notes

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.rscb.org. apo-LigD (4Y98), LigD-NADH (4Y9D), apo-LigO (4YA6), LigO-NADH (4YAC), apo-LigL (4YAE), LigL-NADH (4YAG), LigL-NADH-(α*S*,β*R*)-GGE (4YAI), apo-LigE (4YAM), LigE-GSH (4YAN), LigF-GSH (4XT0), apo-LigG (4YAP) and LigG-GS-AV (4YAV).

Abbreviations

GGE, guaiacylglycerol-β-guaiacyl ether; NAD^+, nicotinamide adenine dinucleotide; NADH, reduced nicotinamide adenine dinucleotide; NADPH, reduced nicotinamide adenine dinucleotide phosphate; MPHPV β-(3′-methoxyphenoxy)-γ-hydroxypropiovanillone; FPHPV, β-(1′-formyl-3′-methoxyphenoxy)-γ-hydroxypropioveratrone; F-FPHPV, β-fluoro-(1′-formyl-3′-methoxyphenoxy)-γ-hydroxypropioveratrone; GSH, glutathione; GS-HPV, β-glutathionyl-γ-hydroxypropiovanillone; GS-HVP, β-glutathionyl-γ-hydroxy-α-veratrylpropanone; F-GS-HVP, β-fluoro-glutathionyl-γ-hydroxy-α-veratrylpropanone; GS-AV, β-glutathionyl-acetoveratrone; SDR, short-chain dehydrogenase/reductase; GST, glutathione-S-transferase; GSSG, glutathione disulfide; k_{cat}, enzymatic turnover number; K_M, Michaelis (half-saturation) constant; V_{MAX}, maximum enzymatic specific activity; HSQC, (^1H–^{13}C) heteronuclear single-quantum coherence (NMR spectroscopy); HMBC, (^1H–^{13}C) heteronuclear multiple-bond correlation (NMR spectroscopy); COSY, (^1H–^1H) correlation spectroscopy; $InsP_6$, inositol hexaphosphate.

References

Adams P.D., Afonine P.V., Bunkoczi G., Chen V.B., Davis I.W., Echols N., Headd J.J., Hung L.-W., Kapral G.J., Grosse-Kunstleve R.W., McCoy A.J., Moriarty N.W., Oeffner R., Read R.J., Richardson D.C., Richardson J.S., Terwilliger T.C., Zwart P.H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D-Biological Crystallography* 66213-221.

Adler E. (1957). Structural elements of lignin. *Industrial & Engineering Chemistry* 49(9):1377-1383.

Afonine P.V., Grosse-Kunstleve R.W., Urzhumtsev A., Adams P.D. (2009). Automatic multiple-zone rigid-body refinement with a large convergence radius. *Journal of Applied Crystallography* 42607-615.

Akiyama T., Magara K., Matsumoto Y., Meshitsuka G., Ishizu A., Lundquist K. (2000). Proof of the presence of racemic forms of arylglycerol-β-aryl ether structure in lignin: studies on the stereo structure of lignin by ozonation. *Journal of Wood Science* 46(5):414-415.

Allocati N., Federici L., Masulli M., Di Ilio C. (2009). Glutathione transferases in bacteria. *Febs Journal* 276(1):58-75.

Armstrong R.N. (1997). Structure, catalytic mechanism, and evolution of the glutathione transferases. *Chemical Research in Toxicology* 10(1):2-18.

Bugg T.D.H., Ahmad M., Hardiman E.M., Rahmanpour R. (2011a). Pathways for degradation of lignin in bacteria and fungi. *Natural Product Reports* 28(12):1883-1896.

Bugg T.D.H., Ahmad M., Hardiman E.M., Singh R. (2011b). The emerging role for bacteria in lignin degradation and bio-product formation. *Current Opinion in Biotechnology* 22(3):394-400.

Cahn R.S., Ingold C., Prelog V. (1966). Specification of molecular chirality. *Angewandte Chemie International Edition in English* 5(4):385-415.

Chang V.S., Holtzapple M.T. (2000). Fundamental factors affecting biomass enzymatic reactivity. *Applied Biochemistry and Biotechnology* 84-65-37.

Davis I.W., Leaver-Fay A., Chen V.B., Block J.N., Kapral G.J., Wang X., Murray L.W., Arendall W.B., III, Snoeyink J., Richardson J.S., Richardson D.C. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research* 35W375-W383.

Emsley P., Cowtan K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D-Biological Crystallography* 602126-2132.

Fan F., Lorenzen J.A., Plapp B.V. (1991). An aspartate residue in yeast alcohol dehydrogenase I determines the specificity for coenzyme. *Biochemistry* 30(26):6397-6401.

Filling C., Berndt K.D., Benach J., Knapp S., Prozorovski T., Nordling E., Ladenstein R., Jornvall H., Oppermann U. (2002). Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *Journal of Biological Chemistry* 277(28):25677-25684.

Franke D., Svergun D.I. (2009). *DAMMIF*, a program for rapid *ab-initio* shape determination in small-angle scattering. *Journal of Applied Crystallography* 42342-346.

Gall D.L., Kim H., Lu F., Donohue T.J., Noguera D.R., Ralph J. (2014a). Stereochemical features of glutathione-dependent enzymes in the *Sphingobium* sp. strain SYK-6 β-aryl etherase pathway. *J Biol Chem* 289(12):8656-8667.

Gall D.L., Ralph J., Donohue T.J., Noguera D.R. (2014b). A group of sequence-related sphingomonad enzymes catalyzes cleavage of β-aryl ether linkages in lignin β-guaiacyl and β-syringyl ether dimers. *Environmental Science & Technology* 48(20):12454-12463.

Grosdidier A., Zoete V., Michielin O. (2011a). Fast Docking Using the CHARMM Force Field with EADock DSS. *Journal of Computational Chemistry* 32(10):2149-2159.

Grosdidier A., Zoete V., Michielin O. (2011b). SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Research* 39W270-W277.

Hayes J.D., Flanagan J.U., Jowsey I.R. (2005). Glutathione transferases. *Annual Review of Pharmacology and Toxicology* 4551-88.

Hendrickson W.A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* 254(5028):51-58.

Hishiyama S., Otsuka Y., Nakamura M., Ohara S., Kajita S., Masai E., Katayama Y. (2012). Convenient synthesis of chiral lignin model compounds via optical resolution: four stereoisomers of guaiacylglycerol-β-guaiacyl ether and both enantiomers of 3-hydroxy-1-(4-hydroxy-3-methoxyphenyl)-2-(2-methoxy-phenoxy)-propan-1-one (erone). *Tetrahedron Letters* 53842-845.

Jancarik J., Kim S.H. (1991). Sparse-matrix sampling - a screening method for crystallization of proteins. *Journal of Applied Crystallography* 24409-411.

Javidpour P., Pereira J.H., Goh E.-B., McAndrew R.P., Ma S.M., Friedland G.D., Keasling J.D., Chhabra S.R., Adams P.D., Beller H.R. (2014). Biochemical and structural studies of NADH-dependent FabG used to increase the bacterial production of fatty acids under anaerobic conditions. *Applied and Environmental Microbiology* 80(2):497-505.

Klock H.E., Koesema E.J., Knuth M.W., Lesley S.A. (2008). Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts. *Proteins-Structure Function and Bioinformatics* 71(2):982-994.

Konarev P.V., Volkov V.V., Sokolova A.V., Koch M.H.J., Svergun D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography* 361277-1282.

Kozin M.B., Svergun D.I. (2001). Automated matching of high- and low-resolution structural models. *Journal of Applied Crystallography* 3433-41.

Krissinel E., Henrick K. (2007). Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* 372(3):774-797.

Leonowicz A., Cho N.S., Luterek J., Wilkolazka A., Wojtas-Wasilewska M., Matuszewska A., Hofrichter M., Wesenberg D., Rogalski J. (2001). Fungal laccase: properties and activity on lignin. *Journal of Basic Microbiology* 41(3-4):185-227.

Martinez A.T., Speranza M., Ruiz-Duenas F.J., Ferreira P., Camarero S., Guillen F., Martinez M.J., Gutierrez A., del Rio J.C. (2005). Biodegradation of lignocellulosics: microbial chemical, and enzymatic aspects of the fungal attack of lignin. *International Microbiology* 8(3):195-204.

Masai E., Ichimura A., Sato Y., Miyauchi K., Katayama Y., Fukuda M. (2003). Roles of the enantioselective glutathione S-transferases in cleavage of β-aryl ether. *Journal of Bacteriology* 185(6):1768-1775.

Masai E., Katayama Y., Fukuda M. (2007). Genetic and biochemical investigations on bacterial catabolic pathways for lignin-derived aromatic compounds. *Bioscience Biotechnology and Biochemistry* 71(1):1-15.

Mathieu Y., Prosper P., Buee M., Dumarcay S., Favier F., Gelhaye E., Gerardin P., Harvengt L., Jacquot J.-P., Lamant T., Meux E., Mathiot S., Didierjean C., Morel M. (2012). Characterization of a *Phanerochaete chrysosporium* glutathione transferase reveals a novel structural and functional class with ligandin properties. *Journal of Biological Chemistry* 287(46):39001-39011.

Mathieu Y., Prosper P., Favier F., Harvengt L., Didierjean C., Jacquot J.-P., Morel-Rouhier M., Gelhaye E. (2013). Diversification of fungal specific class a glutathione transferases in saprotrophic fungi. *Plos One* 8(11).

McCoy A.J., Grosse-Kunstleve R.W., Adams P.D., Winn M.D., Storoni L.C., Read R.J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography* 40658-674.

Meux E., Prosper P., Masai E., Mulliert G., Dumarcay S., Morel M., Didierjean C., Gelhaye E., Favier F. (2012). *Sphingobium* sp SYK-6 LigG involved in lignin degradation is structurally and biochemically related to the glutathione transferase omega class. *Febs Letters* 586(22):3944-3950.

Mills N. (2006). ChemDraw ultra 10.0. *Journal of the American Chemical Society* 128(41):13649-13650.

O'Boyle N.M., Banck M., James C.A., Morley C., Vandermeersch T., Hutchison G.R. (2011). Open Babel: An open chemical toolbox. *Journal of Cheminformatics* 3.

Oppermann U., Filling C., Hult M., Shafqat N., Wu X.Q., Lindh M., Shafqat J., Nordling E., Kallberg Y., Persson B., Jornvall H. (2003). Short-chain dehydrogenases/reductases (SDR): the 2002 update. *Chemico-Biological Interactions* 143247-253.

Otwinowski Z., Minor W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A* 276307-326.

Persson B., Kallberg Y., Bray J.E., Bruford E., Dellaporta S.L., Favia A.D., Gonzalez Duarte R., Jornvall H., Kavanagh K.L., Kedishvili N., Kisiela M., Maserk E., Mindnich R., Orchard S., Penning T.M., Thornton J.M., Adamski J., Oppermann U. (2009). The SDR (short-chain dehydrogenase/reductase and related enzymes) nomenclature initiative. *Chemico-Biological Interactions* 178(1-3):94-98.

Pettersen E.F., Goddard T.D., Huang C.C., Couch G.S., Greenblatt D.M., Meng E.C., Ferrin T.E. (2004). UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25(13):1605-1612.

Ralph J., Peng J.P., Lu F.C., Hatfield R.D., Helm R.F. (1999). Are lignins optically active? *Journal of Agricultural and Food Chemistry* 47(8):2991-2996.

Reddy A.P., Simmons C.W., D'Haeseleer P., Khudyakov J., Burd H., Hadi M., Simmons B.A., Singer S.W., Thelen M.P., VanderGheynst J.S. (2013). Discovery of microorganisms and enzymes involved in high-solids decomposition of rice straw using metagenomic analyses. *Plos One* 8(10).

Reid M.F., Fewson C.A. (1994). Molecular characterization of microbial alcohol dehydrogenases. *Critical Reviews in Microbiology* 20(1):13-56.

Reiter J., Strittmatter H., Wiemann L.O., Schieder D., Sieber V. (2013). Enzymatic cleavage of lignin β-O-4 aryl ether bonds via net internal hydrogen transfer. *Green Chemistry* 15(5):1373-1381.

Roy A., Kucukural A., Zhang Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* 5(4):725-738.

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. *Applied and Environmental Microbiology* 75(16):5195-5201.

Schlieben N.H., Niefind K., Muller J., Riebel B., Hummel W., Schomburg D. (2005). Atomic resolution structures of *R*-specific alcohol dehydrogenase from Lactobacillus brevis provide the structural bases of its substrate and cosubstrate specificity. *Journal of Molecular Biology* 349(4):801-813.

Sheehan D., Meade G., Foley V.M., Dowd C.A. (2001). Structure, function and evolution of glutathione transferases: Implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochemical Journal* 3601-16.

Shen A., Lupardus P.J., Morell M., Ponder E.L., Sadaghiani A.M., Garcia K.C., Bogyo M. (2009). Simplified, enhanced protein purification using an inducible, autoprocessing enzyme tag. *Plos One* 4(12).

Simmons B.A., Loque D., Blanch H.W. (2008). Next-generation biomass feedstocks for biofuel production. *Genome Biology* 9(12).

Simmons C.W., Reddy A.P., D'Haeseleer P., Khudyakov J., Billis K., Pati A., Simmons B.A., Singer S.W., Thelen M.P., VanderGheynst J.S. (2014). Metatranscriptomic analysis of lignocellulolytic microbial communities involved in high-solids decomposition of rice straw. *Biotechnology for Biofuels* 7(495):(31 December 2014)-(2031 December 2014).

Sreenath H.K., Bingman C.A., Buchan B.W., Seder K.D., Burns B.T., Geetha H.V., Jeon W.B., Vojtik F.C., Aceti D.J., Frederick R.O., Phillips G.N., Fox B.G. (2005). Protocols for production of selenomethionine-labeled proteins in 2-L polyethylene terephthalate bottles using auto-induction medium. *Protein Expression and Purification* 40(2):256-267.

Sugimoto T., Akiyama T., Matsumoto Y., Meshitsuka G. (2002). The *erythro/threo* ratio of β-O-4 structures as an important structural characteristic of lignin - Part 2. Changes in *erythro/threo* (E/T) ratio of β-O-4 structures during delignification reactions. *Holzforschung* 56(4):416-421.

Svergun D., Barberato C., Koch M.H.J. (1995). CRYSOL-a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography* 28768-773.

Svergun D.I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography* 25495-503.

Terwilliger T.C., Adams P.D., Read R.J., McCoy A.J., Moriarty N.W., Grosse-Kunstleve R.W., Afonine P.V., Zwart P.H., Hung L.-W. (2009). Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallographica Section D-Biological Crystallography* 65582-601.

Terwilliger T.C., Grosse-Kunstleve R.W., Afonine P.V., Moriarty N.W., Zwart P.H., Hung L.-W., Read R.J., Adams P.D. (2008). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallographica Section D-Biological Crystallography* 6461-69.

Thuillier A., Ngadin A.A., Thion C., Billard P., Jacquot J.-P., Gelhaye E., Morel M. (2011). Functional diversification of fungal glutathione transferases from the ure2p class. *International journal of evolutionary biology* 2011938308-938308.

Tocheva E.I., Fortin P.D., Eltis L.D., Murphy M.E.P. (2006). Structures of ternary complexes of BphK, a bacterial glutathione-S-transferase that reductively dechlorinates polychlorinated biphenyl metabolites. *Journal of Biological Chemistry* 281(41):30933-30940.

Volkov V.V., Svergun D.I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography* 36860-864.

Wiktelius E., Stenberg G. (2007). Novel class of glutathione transferases from cyanobacteria exhibit high catalytic activities towards naturally occurring isothiocyanates. *Biochemical Journal* 406115-123.

CHAPTER 5: Benzoyl coenzyme A pathway-mediated metabolism of *meta*-hydroxy-aromatic acids in *Rhodopseudomonas palustris*

This chapter is published under the same title:

Gall D.L., Ralph J., Donohue T.J., Noguera D.R. (2013). Benzoyl coenzyme A pathway-mediated metabolism of *meta*-hydroxy-aromatic acids in *Rhodopseudomonas palustris*. *Journal of Bacteriology* 195(18):4112-4120.

doi: 10.1128/jb.00634-13

Daniel L. Gall performed all of the experiments and analyses in this chapter.

Abstract

Photoheterotrophic metabolism of two *meta*-hydroxy-aromatic acids, *meta*-, *para*-dihydroxy-benzoate (protocatechuate) and *meta*-hydroxybenzoate, was investigated in *Rhodopseudomonas palustris*. When protocatechuate was the sole organic carbon source, photoheterotrophic growth in *R. palustris* was slow relative to cells using compounds known to be metabolized by the benzoyl coenzyme A (benzoyl-CoA) pathway. *R. palustris* was unable to grow when *meta*-hydroxybenzoate was provided as a sole source of organic carbon under photoheterotrophic growth conditions. However, in cultures supplemented with known benzoyl-CoA pathway inducers (*para*-hydroxy-benzoate, benzoate, or cyclohexanoate), protocatechuate and *meta*-hydroxybenzoate were taken up from the culture medium. Further, protocatechuate and *meta*-hydroxybenzoate were each removed from cultures containing both *meta*-hydroxy-aromatic acids at equimolar concentrations in the absence of other organic compounds. Analysis of changes in culture optical density and in the concentration of soluble organic compounds indicated that the loss of these *meta*-hydroxy-aromatic acids was accompanied by biomass production. Additional experiments with defined mutants demonstrated that enzymes known to participate in the dehydroxylation of *para*-hydroxybenzoyl-CoA (HbaBCD) and reductive dearomatization of benzoyl-CoA (BadDEFG) were required for

metabolism of protocatechuate and *meta*-hydroxybenzoate. These findings indicate that, under photoheterotrophic growth conditions, *R. palustris* can degrade *meta*-hydroxy-aromatic acids via the benzoyl-CoA pathway, apparently due to the promiscuity of the enzymes involved.

Introduction

A combination of human and plant activities give rise to a variety of aromatic compounds in the environment. Among these compounds are aromatic carboxylic acids that contain hydroxyl groups disposed *meta*, *ortho*, or *para* (*m*, *o*, or *p*) to the carboxyl. Aromatic acids such as benzoate, *para*-hydroxybenzoate (*p*HB), *meta*-hydroxybenzoate (*m*HB) and *meta*-,*para*-dihydroxybenzoate (protocatechuate) can serve as carbon sources for some bacteria. In aerobic pathways, molecular oxygen (O_2) is used as an electrophilic cosubstrate for α-electron destabilization and aromatic ring fission (Harwood and Parales 1996). In contrast, strict and facultative anaerobes use O_2-independent mechanisms for aromatic ring cleavage (Carmona *et al.* 2009). In these O_2-independent pathways, benzoyl coenzyme A (benzoyl-CoA) thioesters are formed and reductively transformed into essential precursors of central metabolism (Breese *et al.* 1998; Harrison and Harwood 2005). This work sought to gain additional insight into anaerobic metabolism of hydroxylated aromatic compounds by the facultative photoheterotroph *Rhodopseudomonas palustris*.

Anaerobic metabolism of aromatic compounds has been most extensively studied in *R. palustris* and the denitrifying bacterium *Thauera aromatica* (Harwood *et al.* 1998). In *R. palustris* and *T. aromatica*, catabolism of *p*HB (Figure 5-1A) proceeds via benzoyl-CoA, an intermediate that is also used for metabolism of benzoate (Harwood and Gibson 1986; Dangel *et al.* 1991; Egland *et al.* 1997). In both organisms, specific CoA-dependent ligases catalyze the ATP-dependent synthesis of *p*HB-CoA (HbaA, EC 6.2.1.27) (Merkel *et al.* 1989; Biegert *et al.* 1993; Gibson *et al.* 1994) and benzoyl-CoA (BadA, EC 6.2.1.25) (Geissler *et al.* 1988; Egland *et al.* 1995; Schuhle *et al.* 2003). The

*p*HB-CoA thioester is subsequently dehydroxylated by *p*HB-CoA reductase (HbaBCD, EC 1.3.7.9) to produce benzoyl-CoA (Brackmann and Fuchs 1993; Gibson *et al.* 1997; Breese and Fuchs 1998). Benzoyl-CoA dearomatization is catalyzed by benzoyl-CoA reductase (BadDEFG, EC 1.3.7.8) forming aliphatic intermediates (1,5-diene-cyclohexanoyl-CoA in *T. aromatica* (Boll and Fuchs 1995; Boll and Fuchs 1998) and 1-ene-cyclohexanoyl-CoA in *R. palustris* (Gibson and Gibson 1992)) that serve as carbon sources and reducing power in catabolic pathways that ultimately yield acetyl-CoA, CO_2, and NADH (Perrotta and Harwood 1994; Laempe *et al.* 1998; Pelletier and Harwood 1998; Laempe *et al.* 1999).

Despite the striking similarities between the metabolic routes through which *p*HB and benzoate are catabolized in *T. aromatica* and *R. palustris*, there are significant differences in the ways these two organisms metabolize *meta*-hydroxy-aromatic acids. *R. palustris* is not reported to grow photoheterotrophically with *m*HB as a sole organic carbon source, whereas protocatechuate serves only as a poor organic carbon source relative to the known benzoyl-CoA pathway-inducing substrates, benzoate, *p*HB, and cyclohexanoate (Harwood and Gibson 1988; Kim and Harwood 1991; Heider *et al.* 1998; Ding *et al.* 2008). In contrast, *T. aromatica* is known to activate both protocatechuate and *m*HB via enzymatic activity with *m*HB-CoA ligase (EC 6.2.1.37) to form the cognate benzylic thioesters (Laempe *et al.* 2001; Philipp *et al.* 2002). In *T. aromatica*, metabolism of *meta*-hydroxy-aromatic acids proceeds through intermediate *m*HB-CoA (not benzoyl-CoA). Derived as a product of benzoyl-CoA reductase activity on *m*HB-CoA, the aliphatic thioester 1,5-diene-3-hydroxycyclohexanoyl-CoA (Figure 5-1B) is further transformed to ultimately yield metabolic intermediates and reducing equivalents necessary for growth (Laempe *et al.* 1998; Laempe *et al.* 2001; Philipp *et al.* 2002).

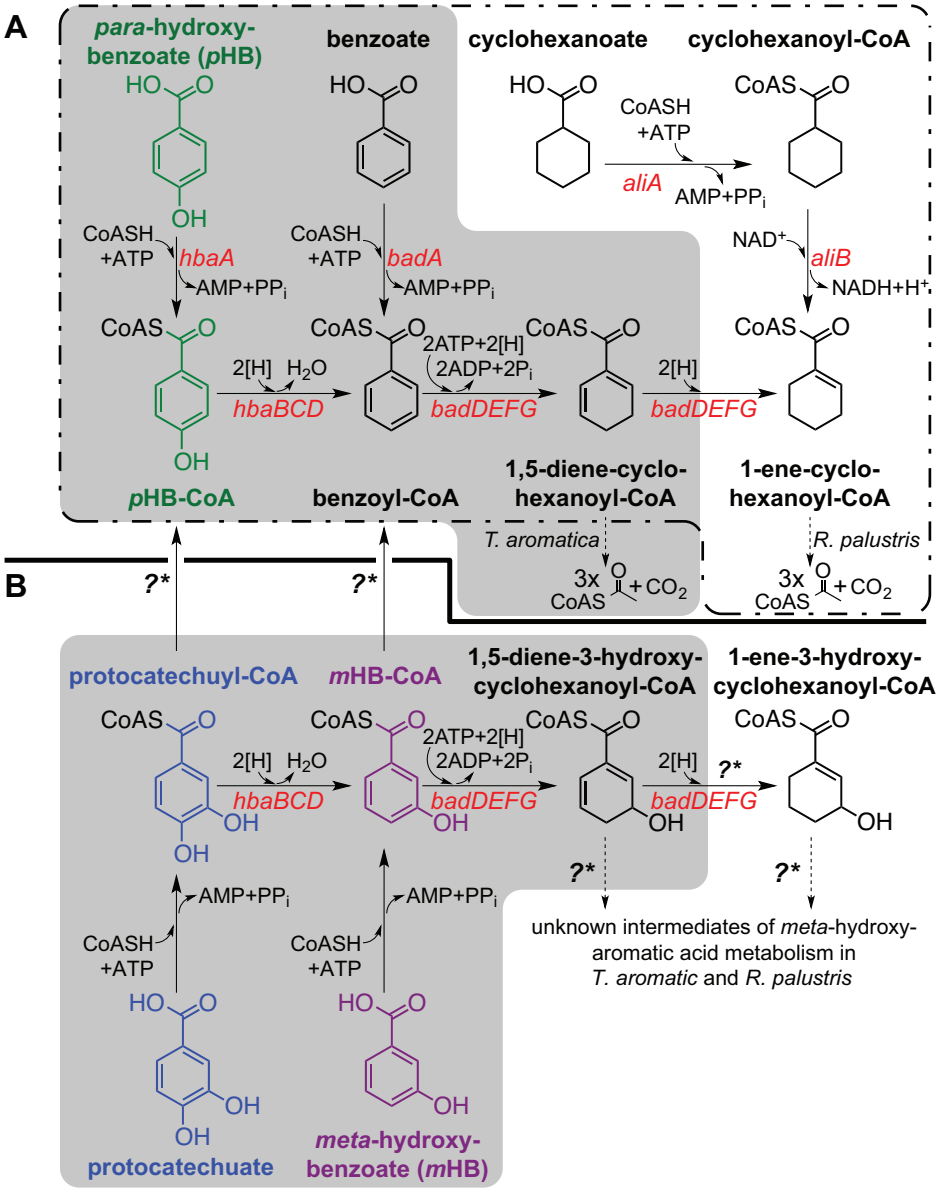

Figure 5-1. Benzoyl-CoA pathway-mediated transformation of aromatic acids. Genes are shown (red) for previously characterized reactions and correspond to annotations in the *R. palustris* genome sequence (Egland *et al.* 1997; Larimer *et al.* 2004). Reactions and intermediates that have been previously characterized in *R. palustris* are shown within the dashed outline (Gibson and Gibson 1992; Egland *et al.* 1995; Egland *et al.* 1997; Gibson *et al.* 1997). Reactions and intermediates that have been previously characterized in *T. aromatica* are indicated by shading (Breese *et al.* 1998; Laempe *et al.* 2001; Philipp *et al.* 2002). Reactions that have yet to be demonstrated in vivo or in vitro are denoted with an asterisk and question mark (*?**). **(A)** Benzoyl-CoA pathway functions used in metabolism of *para*-hydroxybenzoate, benzoate and cyclo-hexanoate; **(B)** proposed benzoyl-CoA pathway functions used in metabolism of protocatechuate and *meta*-hydroxybenzoate.

We analyzed here the ability of *R. palustris* to grow on *meta*-hydroxy-aromatic acids. Our results provide evidence for benzoyl-CoA pathway-mediated metabolism of *meta*-hydroxy-aromatic acids when cells are grown under conditions known to induce expression of benzoyl-CoA pathway enzymes. We provide genetic evidence for the requirement of the benzoyl-CoA pathway for the anaerobic metabolism of *meta*-hydroxy-aromatic acids. Based on this, we present a model to explain why metabolism of *meta*-hydroxy-aromatic acids requires or is enhanced by conditions that induce expression of the benzoyl-CoA pathway enzymes.

Materials and Methods

Bacterial strains and growth conditions

The bacterial strains and plasmids used in the present study are described in Table 5-1. *R. palustris* strains used in this study. *R. palustris* strain CGA009 and its derivatives were grown in pre-sterilized photosynthetic media (PM) supplemented with carbon sources at selected concentrations (Kim and Harwood 1991). As needed, 100 μg of kanamycin mL^{-1} was added to media for the maintenance of *R. palustris* mutant strains. Photoheterotrophic cultures were grown at 30°C and illuminated with incandescent light bulbs at ~10 W m^{-2}. *R. palustris* growth was monitored using a Klett-Summerson photoelectric colorimeter (Klett MFG Co., New York, NY).

Table 5-1. *R. palustris* strains used in this study

Strain	Genotype/Relevant characteristic(s)[a]	Reference
CGA009	Wild-type; spontaneous Cm^r derivative of CGA001	(Kim and Harwood 1991)
CGA506	*hbaB*::*'lacZ*-Km^r	(Gibson *et al.* 1997)
CGA606	*badE*::*'lacZ*-Km^r	(Egland *et al.* 1997)

[a] Cm, chloramphenicol; Km, kanamycin; *'lacZ*, promoterless *lacZ* gene

For aromatic acid- and cyclohexanoate-supported growth experiments, the initial medium pH was set to 7.0, except when protocatechuate was used as an organic carbon source, where the initial pH was adjusted to pH 6.5 using phosphoric acid to prevent photochemical degradation of protocatechuate. At the time of inoculation, the culture medium was sparged with argon gas to ensure anaerobic conditions at the onset of each experiment. Replicate (two–four) 21-mL *R. palustris* cultures were inoculated with ~250 µl of succinate-grown cells and supplemented with 30 mM $NaHCO_3$, as well as 4.4 to 4.6 mM total aromatic acid, unless specified otherwise. Final aromatic substrate concentrations were chosen to maintain equal total reducing equivalents for all cultures tested at 1 g of chemical oxygen demand (COD) per liter of medium.

Analytical methods

Samples were aseptically removed from cultures by removing 200-µl aliquots while adding argon gas to maintain headspace atmospheric pressure. Samples were filtered through 0.22-µm-pore-size hydrophilic Durapore polyvinylidene difluoride membranes (Merck KGaA, Darmstadt, Germany) and acidified with phosphoric acid (0.5% [vol/vol]) before analysis by gas chromatography (GC) or high-pressure liquid chromatography (HPLC).

GC analyses – Acetate, butyrate, and cyclohexanoate concentrations were quantified using a GC-2010 gas chromatograph (Shimadzu, Kyoto, Japan) equipped with a flame ionization detector. Compounds were separated using a Stabilwax-DA capillary free fatty acid-phase column (0.32 mm [inner diameter] by 30 m; Agilent Technologies, Wilmington, DE) using helium as the mobile phase. Ramping of the column temperature began 2 min after the time of injection, increasing from 60°C at a rate of 12°C min^{-1} and held at 240°C for 2 min. The injector and detector temperatures were 250°C.

HPLC analyses – Aromatic acid concentrations were quantified using an LC-10AT$_{VP}$ solvent delivery module HPLC system (Shimadzu, Kyoto, Japan) equipped with an SPD-M10A$_{VP}$ diode array detector (Shimadzu, Kyoto, Japan). Aromatic compounds were separated using an Ultra Aqueous (Restek Corp., Bellefonte, PA) C_{18}-reversed stationary phase column (5 μm particle size; 120 mm × 4 mm I.D.) and an isocratic aqueous mobile phase of methanol (30% [wt/vol]), acetonitrile (6% [wt/vol]), and 5 mM formic acid in water (64% [wt/vol]) at a flow rate of 0.8 mL min^{-1} (Phelps and Young 1997). Aromatic acids were detected by UV absorption at 280 nm. Concentrations of aromatic compounds were calculated from linear regressions created from standards of known concentration.

COD analyses – Culture supernatants were filtered through 0.22-μm-pore-size membranes prior to measuring the concentrations of soluble organic compounds using a HACH High-Range (0 to 1,500 mg L^{-1}) COD kit (HACH Company, Loveland, CO). COD is a standard test specific for organic substrates (APHA/AWWA/WEF 2005) and measures the amount of oxygen required to fully oxidize and organic substrate to CO_2. As such, it has been used to understand the fate of reducing equivalents in photoheterotrophic cultures (Yilmaz *et al.* 2010). The theoretical COD values for various carbon sources used in the present study are as follows (in mg of COD mmol^{-1} of substrate): acetate, 64; benzoate, 240; butyrate, 160; caffeate, 288; cyclohexanoate, 288; *m*HB, 224; *p*HB, 224; protocatechuate, 208; bicarbonate, 0; and succinate, 112.

Results

Conditions to induce meta-hydroxy-aromatic acid metabolism in wild-type R. palustris

The goal of the present study was to identify culture conditions that allowed anaerobic metabolism of *meta*-hydroxy-aromatic compounds by wild-type *R. palustris* strain CGA009. Consistent with previous findings (Harwood and Gibson 1988; Gibson *et al.* 1997), photo-heterotrophic cultures of strain CGA009 incubated with either *m*HB or protocatechuate as a sole

organic carbon source were incapable of doubling at rates similar to cells grown with benzoate, *p*HB, or cyclohexanoate (0.4-, 0.6-, and 0.4-day doubling times, respectively) (Table 5-2). The presence of *m*HB (2.2 mM) as a sole organic carbon source did not support detectable growth of wild-type cultures even after an extended incubation period under photoheterotrophic conditions. However, protocatechuate (2.3 mM) supported very slow photoheterotrophic growth of *R. palustris*. For all experiments during which growth was observed, measurements of medium COD indicated >82% removal of the organic substrates as the biomass increased, demonstrating that growth was associated with the degradation of the substrates, even in the protocatechuate-fed cultures. By-products from the degradation of the different substrates were not detected by either GC or HPLC analyses. However, a small accumulation of undetected by-products in these growth conditions cannot be ruled out, given the incomplete removal of COD from the culture medium.

Table 5-2. Doubling times of wild-type *R. palustris* strain CGA009 and COD removal efficiency during photoheterotrophic growth in cultures containing a single source of organic carbon

		Avg (SD)	
Growth Substrate	Concentration (mM)	T_d[a]	%COD[b]
Benzoate	2.2	0.4 (0.0)	88.4 (1.0)
*p*HB	2.2	0.6 (0.0)	86.6 (3.6)
Cyclohexanoate	2.0	0.4 (0.0)	89.1 (1.5)
*m*HB	2.2	NDG	-
Protocatechuate	2.3	> 60	82.9 (5.1)

[a] The doubling time (T_d) is reported as the average of replicate (two or more) cultures incubated with the indicated concentrations of carbon sources. NDG, no detectable growth.

[b] The COD removal efficiency is reported as the average of at least two cultures from which the soluble COD was measured before inoculation and after visible growth of the culture had stopped. –, Condition not texted.

We also found that strain CGA009 was able to metabolize *m*HB or protocatechuate when either benzoate, *p*HB, or cyclohexanoate (for *m*HB only) were also supplied as an additional cosubstrate (Table 5-3). Under these conditions, cultures exhibited a biphasic mode of growth, where benzoate, *p*HB, or cyclohexanoate were degraded first (defined here as primary substrates) with apparent doubling times similar to those observed when these substrates were sole growth substrates (Table 5-2). After the primary substrate was consumed, *meta*-hydroxy-aromatic acid cosubstrates, either *m*HB or protocatechuate (defined here as secondary substrates), were subsequently removed from the media, and the cultures exhibited longer doubling times (average 1.6 days and 19 days, respectively). Notably, in comparison to when benzoate or *p*HB was the sole organic carbon source (Table 5-2), growth with benzoate and *p*HB as primary substrates was slower when protocatechuate was the secondary substrate (Table 5-3), perhaps as an effect of competition between protocatechuate and the primary substrate for required enzymes. Although it was not an enhancer of protocatechuate metabolism, cyclohexanoate was utilized as a primary substrate when *m*HB was degraded as the secondary substrate, but growth during *m*HB uptake was slower (8.7-day doubling time). The inability of strain CGA009 to degrade protocatechuate as the secondary substrate when cyclohexanoate was the cosubstrate may be due to the additional enzymes (*e.g.*, HbaBCD) that are expected to be needed for protocatechuate degradation compared to *m*HB metabolism (Figure 5-1B). In these two-substrate experiments, the primary and secondary substrates were added in equimolar concentrations, and each corresponded to *ca.* 50% of the initial COD in the media, so we found approximately equal increases in biomass during growth on the primary and secondary substrate (between 100 and 120 Klett units per substrate). The COD removal efficiency when *m*HB was the secondary substrate was between 75.0 and 81.8% (Table 5-3), demonstrating that *m*HB was utilized for growth. When protocatechuate was the secondary substrate, the COD removal efficiencies were 86.7 and 76.9% when biphasic growth was observed, also demonstrating that

protocatechuate was utilized as a carbon source for growth. In contrast, when cyclohexanoate was used as the primary substrate and neither biphasic growth nor protocatechuate degradation was subsequently observed, the COD removal efficiency was only 47.9%, in agreement with protocatechuate not being used for growth under this condition. Similar to the experiments with single substrates, by-products of these transformations were not detected but small accumulations of metabolites cannot be ruled out since the COD removal was incomplete in these cultures.

Table 5–3. Doubling times of *R. palustris* strains and COD removal efficiencies during photoheterotrophic growth in cultures containing equimolar combinations of primary substrates and secondary substrates

Equimolar cosubstrate combination (concn [mM])[a]		Avg T_d (days) and %COD values in three *R. palustris* strains[b]								
		CGA009			CGA506			CGA606		
PS	SS	T_d (PS)	T_d (SS)	%COD	T_d (PS)	T_d (SS)	%COD	T_d (PS)	T_d (SS)[c]	%COD
Benzoate (2.2)	*m*HB (2.2)	0.4 (0.0)	1.6 (0.1)	75.0 (0.5)	-	-	-	NDG	NDG	-
*p*HB (2.2)	*m*HB (2.2)	0.7 (0.2)	1.6 (0.5)	77.1 (0.5)	-	-	-	NDG	NDG	-
Cyclohexanoate (2.0)	*m*HB (2.0)	0.4 (0.0)	8.7 (2.0)	81.8 (2.9)	-	-	-	0.3 (0.1)	NDG	54.4 (1.6)
Benzoate (2.2)	Protocatechuate (2.2)	0.7 (0.0)	16.2 (0.9)	86.7 (0.7)	0.7 (0.1)	>60	68.5 (2.5)	NDG	>0*	-
*p*HB (2.2)	Protocatechuate (2.2)	1.1 (0.0)	21.7 (1.8)	76.9 (0.1)	NDG	NDG	-	NDG	>60*	-
Cyclohexanoate (2.0)	Protocatechuate (2.0)	0.4 (0.0)	NDG	47.9 (4.5)	-	-	-	-	-	-
*m*HB (2.3)	Protocatechuate (2.3)	2.1 (0.3)	32.7 (1.0)	67.8 (2.7)	2.1 (0.1)	>60	57.6 (5.3)	NDG	>60*	-

[a] PS, primary substrate; SS, secondary substrate.

[b] Doubling times (T_d) are reported as the average of replicate (two or more) cultures incubated with the indicated concentrations of carbon sources. HPLC analysis of media (as in Fig. 1 to 5) showed that primary substrates were consumed prior to significant utilization of secondary substrates during each experiment. The percentages of soluble COD removed (%COD) are averages of at least two cultures from which the soluble COD concentration was measured before inoculation and after visible growth of the culture had stopped. Standard deviations are indicated in parentheses. NDG, no detectable growth; –, condition not tested.

[c] *, Observed growth was attributed to utilization of protocatechuate because the primary substrate concentrations did not change during the course of the experiments.

To further investigate the utilization of *m*HB and protocatechuate, we performed another set of growth experiments with a molar excess of the secondary substrates. For example, when strain CGA009 was incubated in media with an approximately 5-to-1 molar excess of *m*HB (3.7 mM) to *p*HB (0.7 mM), *p*HB was depleted first, followed by *m*HB (Figure 5-2A). Under these conditions, the contribution of *p*HB to the total COD in the medium was 16.6%, whereas the measured COD removal in the experiment was 71.5%. This confirms that *m*HB is being metabolized by the culture. Furthermore, there is an ~4-fold increase in optical density as the *m*HB is degraded compared to that observed when *p*HB was degraded (178 Klett units compared to 50 Klett units), a finding consistent with the excess *m*HB providing most of the organic carbon for growth. Similarly, in experiments where protocatechuate (3.9 mM) was present at a 5-to-1 molar ratio over benzoate (0.8 mM) in the medium (Figure 5-2B), benzoate degradation and concomitant growth preceded subsequent consumption of protocatechuate and slower cell growth, with an ~4-fold increase in optical density during the period of protocatechuate degradation. In this case, medium COD concentrations were reduced by an average of 66.5%, whereas benzoate represented only 18.7% of the initial COD. Therefore, the results of these experiments also confirmed that protocatechuate was used as a source of carbon and reducing equivalents to support cell growth in these cultures, without excluding the possibility of some undetected metabolites accumulated in the media.

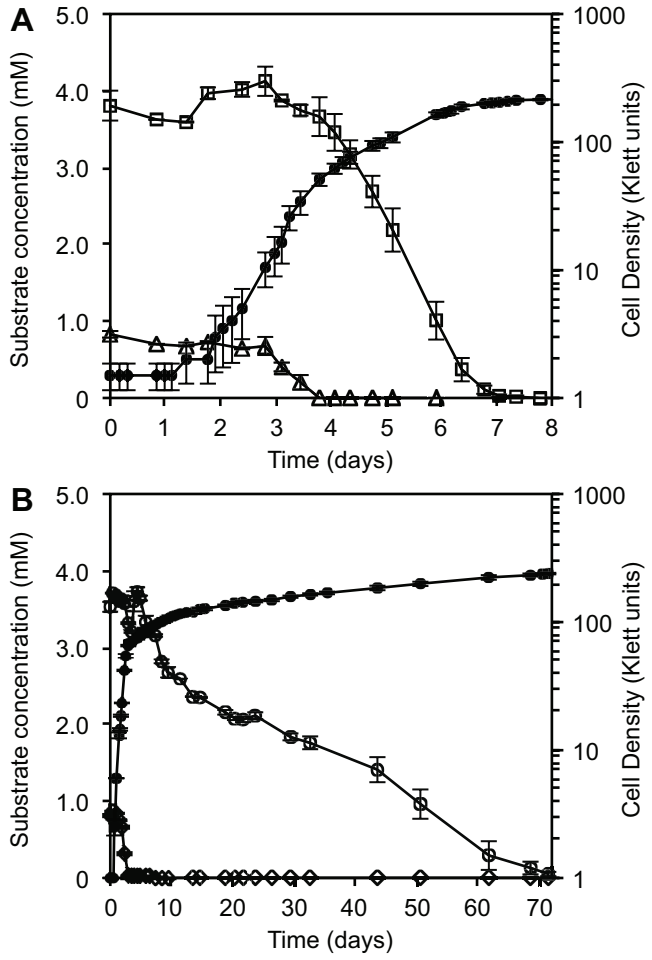

Figure 5-2. Photoheterotrophic growth of the *R. palustris* wild-type strain CGA009 supplemented with excess of *meta*-hydroxy-aromatic acid to benzoyl-CoA pathway inducing substrate (5-to-1 molar ratio). Cell density (•) is reported in Klett units (right y-axis). Medium benzoate (◊), *para*-hydroxybenzoate (Δ), *meta*-hydroxybenzoate (□), and protocatechuate (○) concentrations (left y-axis) are indicated by the corresponding symbols. Error bars represent the standard deviations of replicate cultures. **(A)** *para*-Hydroxybenzoate (0.7 mM) plus *meta*-hydroxybenzoate (3.7 mM)-supported photoheterotrophic growth of strain CGA009; **(B)** benzoate (0.8 mM) + protocatechuate (3.9 mM)-supported photoheterotrophic growth of strain CGA009.

Surprisingly, we found that strain CGA009 also grew when provided with *m*HB and protocatechuate as cosubstrates at equimolar concentrations, with cells utilizing *m*HB as the primary substrate at short doubling times and protocatechuate as the slowly degraded secondary substrate (Figure 5-3A), supporting growth at relatively longer doubling times (Table 5-3). A COD removal

efficiency of 67.8% and a nearly 2-fold increase in optical density after *m*HB had been depleted (Figure 5-3A) confirmed the utilization of both substrates for growth, although the accumulation of metabolites cannot be ruled out given the low COD removal efficiency.

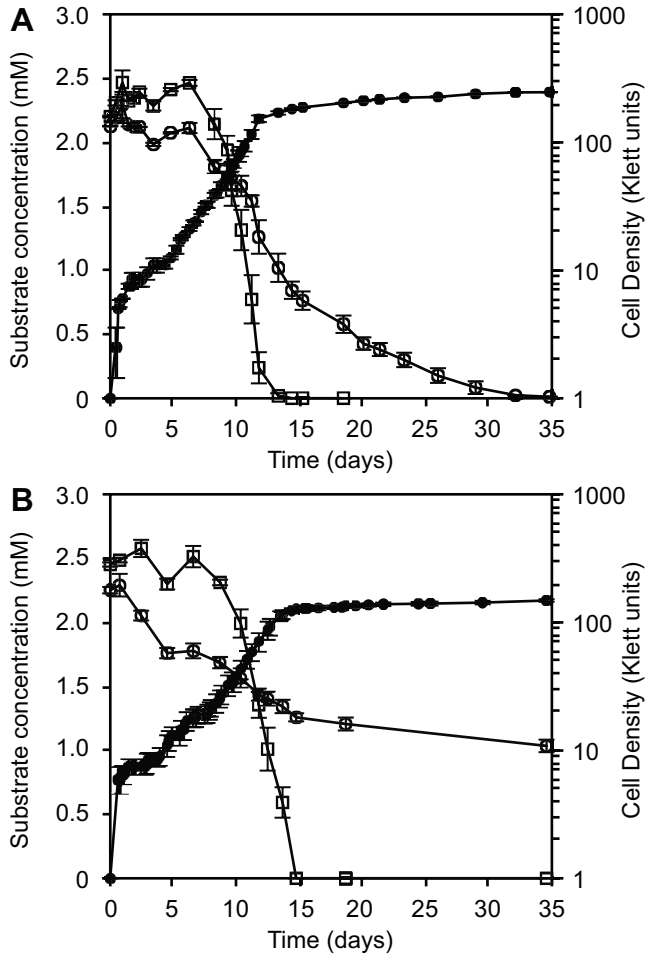

Figure 5-3. *meta*-Hydroxybenzoate (2.3 mM) plus protocatechuate (2.3 mM)-supported photoheterotrophic growth of *R. palustris* strains. The cell density (•) is reported in Klett units (right y-axis). Medium *meta*-hydroxybenzoate (□) and protocatechuate (○) concentrations (left y-axis) are represented by the corresponding symbols. Error bars represent the standard deviations of replicate cultures. **(A)** *meta*-Hydroxybenzoate- and protocatechuate-grown cells of wild-type strain CGA009; **(B)** *meta*-hydroxybenzoate-grown cells of strain CGA506, deficient in *p*HB-CoA reductase activity. Protocatechuate decrease in this culture is due to photochemical decay.

We also found that neither *m*HB nor protocatechuate (2.2 mM) were removed from cultures of wild-type cells when either acetate (7.8 mM) or butyrate (3.1 mM) was added as a cosubstrate

(data not shown). This indicates that catabolism of the *meta*-hydroxy-aromatic acids was not supported by other organic carbon sources and suggests that *meta*-hydroxy-aromatic acid metabolism requires the presence of one or more benzoyl-CoA pathway enzymes, the expression of which is induced by the presence of benzoate, *p*HB, or cyclohexanoate (Kim and Harwood 1991; Egland and Harwood 1999; Egland and Harwood 2000; Pan *et al.* 2008). Consistent with previous findings (Harwood and Gibson 1988), in cultures where caffeate (an analog of protocatechuate with a 3-carbon aliphatic side chain) was the sole organic carbon source (2.2 mM), this compound was taken up from the medium concomitantly with an increase in the medium's protocatechuate concentration. In this case, the protocatechuate concentration rose to ~2.2 mM (the same as the initial caffeate concentration; data not shown), suggesting that the aliphatic side chain of caffeate was being used to support growth without significant metabolism of the *meta*-hydroxy-aromatic acid moiety. Combined, these results suggest that utilization of the *meta*-hydroxy-aromatic compounds requires the activity of benzoyl-CoA pathway enzymes.

Role of benzoyl-CoA pathway enzymes in meta-hydroxy-aromatic acid metabolism

There are reported differences in the pathways used for the anaerobic metabolism of *meta*-hydroxy-aromatic acids by *R. palustris* and *T. aromatica*. Thus, we tested whether previously characterized genes that encode benzoyl-CoA pathway enzymes are needed for the observed metabolism of *meta*-hydroxy-aromatic acids in *R. palustris*. It was previously reported that *R. palustris* cells lacking pHB-CoA reductase (HbaBCD) activity (CGA506; Table 5-1) grow at wild-type rates while using benzoate but exhibit a growth defect when *p*HB is the organic carbon source under photoheterotrophic conditions (Gibson *et al.* 1997). In media containing equimolar benzoate and protocatechuate (2.2 mM), strain CGA506 grew with wild-type photoheterotrophic generation times while benzoate was present in the media (Figure 5-4B and Table 5-3). However, once benzoate was consumed, the growth of strain CGA506 was impaired, and protocatechuate in the

media decayed slowly (Figure 5-4B). Further, compared to wild-type cultures incubated under identical conditions (Figure 5-4A), the COD removal efficiency was lower in the CGA506 culture at the end of the experiment (Table 5-3). This, plus the lower overall cell yield of CGA506 under these conditions, supports the hypothesis that *p*HB-CoA reductase activity is required for protocatechuate metabolism in *R. palustris*. In further support of this notion, we found that when strain CGA506 is incubated with equimolar concentrations of *m*HB and protocatechuate (2.3 mM), *m*HB supported growth with wild-type generation times, but protocatechuate metabolism was inhibited (Figure 5-3B) compared to the wild-type strain (Figure 5-3A and Table 5-3). Taken together, our findings suggest that the protocatechuate thioester (protocatechuyl-CoA), or its potential *meta*-dehydroxylated intermediate (*p*HB-CoA), is subject to *para*-dehydroxylation by *p*HB-CoA reductase (Figure 5-1) and that this activity is required for protocatechuate-supported growth by *R. palustris*. In addition, *p*HB-CoA reductase is not required for the degradation of *m*HB.

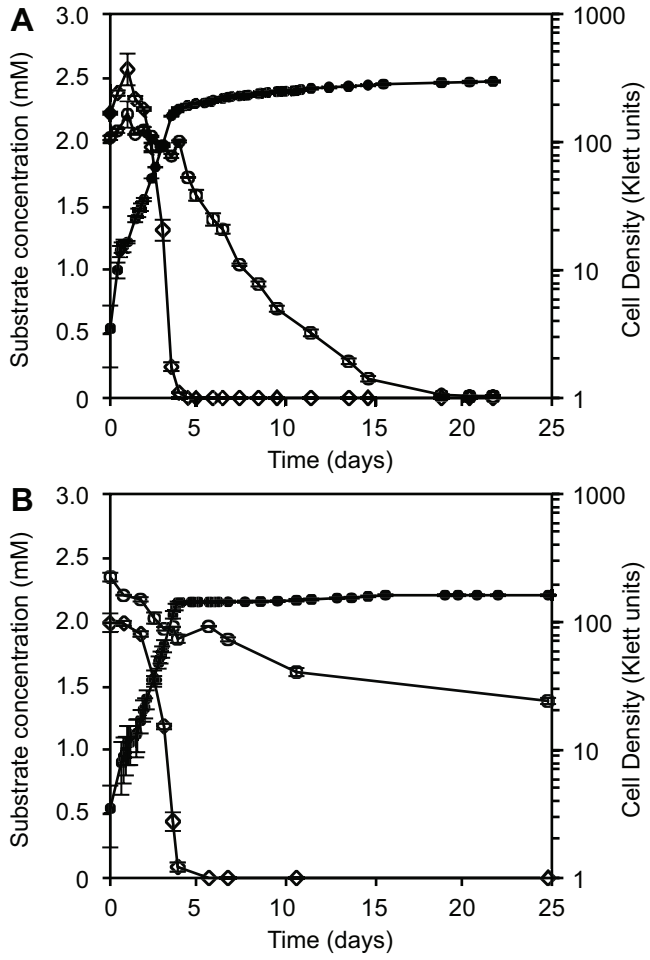

Figure 5-4. Benzoate (2.2 mM) plus protocatechuate (2.2 mM)-supported photoheterotrophic growth of *R. palustris* strains. The cell density (•) is reported in Klett units (right y-axis). Medium benzoate (◊) and protocatechuate (○) concentrations (left y-axis) are indicated by the corresponding symbols. Error bars represent the standard deviations of replicate cultures. **(A)** Benzoate- and protocatechuate-grown cells of wild-type strain CGA009; **(B)** benzoate-grown cells of strain CGA506, deficient in *p*HB-CoA reductase activity. Protocatechuate decrease in this culture is due to photochemical decay.

To test the role of benzoyl-CoA reductase (BadDEFG) in the metabolism of *meta*-hydroxy-aromatic acids, we used a mutant (CGA606) lacking benzoyl-CoA reductase activity (Egland *et al.* 1997). In this case, we analyzed growth of strain CGA606 on cyclohexanoate and *m*HB since cyclo-hexanoate enters the benzoyl-CoA pathway downstream of benzoyl-CoA reductase activity (Figure 5-1A). When incubated with equimolar concentrations of cyclohexanoate and *m*HB (2.0 mM), strain

CGA606 (Figure 5-5B) grew with wild-type doubling times (Figure 5-5A) as long as cyclohexanoate was present in the medium (Table 5-3). However, growth ceased after cyclohexanoate was depleted from the medium, and there was no net removal of *m*HB (Figure 5-5B), indicating that the loss of benzoyl-CoA reductase activity blocked the ability of cells to metabolize *m*HB. Further, the amount of COD removed by strain CGA606 was lower than in wild-type cells under this growth condition (Table 5-3), a finding consistent with the hypothesis that BadDEFG is required for *m*HB degradation. These data suggest that benzoyl-CoA reductase is likely facilitating the dearomatization of *m*HB-CoA or its potential *meta*-dehydroxylated intermediate (benzoyl-CoA), once it is formed in wild-type cells (Figure 5-1).

We also found that strain CGA606 was capable of only slow growth (>60-day generation times) with protocatechuate as the sole carbon source (2.2 mM), similar to the observations that had been made when testing the wild-type strain (Table 5-2). This observation may be due to both strains making use of protocatechuate photochemical degradation products as growth substrates, given our observation that protocatechuate degrades at approximately the same rate under light in abiotic tubes and in cultures where protocatechuate is the sole organic carbon source (data not shown). These data also leave open the possibility that a yet-to-be-characterized pathway for anaerobic aromatic metabolism that has been proposed in previous studies (Harwood and Gibson 1988; Merkel *et al.* 1989) may exist in strain CGA009 and contribute to protocatechuate metabolism under these conditions. However, the potential contribution of this pathway is apparently minimal relative to that of the benzoyl-CoA pathway when cells are grown in the presence of substrates that induce the activity of benzoyl-CoA pathway enzymes.

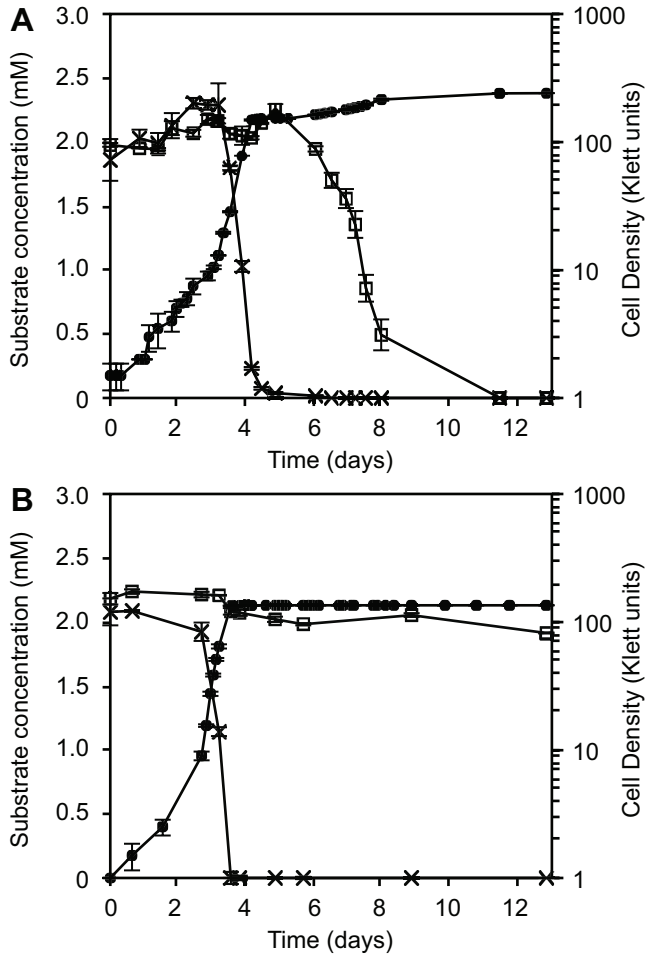

Figure 5-5. Cyclohexanoate (2.0 mM) plus *meta*-hydroxybenzoate (2.0 mM)-supported photo-heterotrophic growth of *R. palustris* strains. The cell density (•) is reported in Klett units (right y-axis). Medium cyclohexanoate (×) and *meta*-hydroxybenzoate (□) concentrations (left y-axis) are indicated by the corresponding symbols. Error bars represent the standard deviations of replicate cultures. **(A)** Cyclohexanoate and *meta*-hydroxybenzoate grown cells of wild-type strain CGA009; **(B)** cyclohexanoate-grown cells of strain CGA606, deficient in benzoyl-CoA reductase activity.

Discussion

The ability to utilize aromatic compounds with one or more ring substitutions is integral to the biodegradation of natural or xenobiotic aromatic compounds. For example, lignin is a significant portion of plant lignocellulose and is a polymer of aromatic subunits containing an aliphatic side chain, a *para*-substitution, and up to two substitutions in the *meta*-positions (relative to the

aliphatic side chain). In addition, there are many toxic products of human activity that contain functional group substitutions on the aromatic ring. Thus, there is considerable interest in identifying microbial pathways to either degrade ring-substituted aromatics or convert them into high-value products. The data in the present study predicts that HbaBCD and BadDEFG have previously unreported roles in the degradation of aromatic compounds containing *meta*-hydroxy-aromatic functional groups.

Benzoyl-CoA pathway-inducing substrates enhance meta-hydroxy-aromatic acid catabolism in R. palustris

Aromatic-supported photoheterotrophic growth by *R. palustris* was first described with benzoate as a sole organic carbon source (Geissler *et al.* 1988; Egland *et al.* 1997). It was subsequently shown that *R. palustris* reductively transforms *para*-hydroxyl ring substitutions via the benzoyl-CoA pathway (Figure 5-1A) using much of the same enzymology used during utilization of benzoate (Merkel *et al.* 1989; Dispensa *et al.* 1992; Gibson *et al.* 1994; Gibson *et al.* 1997) and that benzoate, *p*HB, and cyclohexanoate are inducers of this pathway (Harwood and Gibson 1988; Kim and Harwood 1991; Heider *et al.* 1998; Ding *et al.* 2008). Here, we demonstrate that under conditions known to induce expression of benzoyl-CoA pathway genes (Pan *et al.* 2008), *R. palustris* also utilizes the *meta*-hydroxy-aromatic acids protocatechuate and *m*HB during photoheterotrophic growth.

A comparison of growth characteristics of *R. palustris* in single-substrate- and two-substrate-fed cultures reveals new aspects of aromatic acid metabolism by the benzoyl-CoA pathway. *R. palustris* strain CGA009 was unable to utilize *m*HB for growth either when it was the sole photo-heterotrophic carbon source or when the media also contained rapidly metabolized nonaromatic organic acids as cosubstrates (butyrate or acetate). Protocatechuate was inefficiently utilized when supplied as a sole organic carbon source. However, degradation of both *meta*-hydroxy-aromatic

compounds was observed in cultures that were either (i) supplemented with one of the three known inducers of benzoyl-CoA pathway gene expression—cyclohexanoate (for *m*HB only), benzoate (for *m*HB or protocatechuate), or *p*HB (for *m*HB or protocatechuate)—or (ii) fed both *m*HB and protocatechuate as cosubstrates. The means by which (i) the three known benzoyl-CoA pathway inducers enhance metabolism of one or both *meta*-hydroxy-aromatic acids or (ii) the combination of *m*HB and protocatechuate as cosubstrates enhances the metabolism of each *meta*-hydroxy-aromatic acid remains to be determined since pathway expression is affected by both transcriptional (Dispensa *et al.* 1992; Egland and Harwood 1999; Egland and Harwood 2000) and posttranslational (Crosby *et al.* 2010; Crosby *et al.* 2012) mechanisms.

Benzoyl-CoA pathway enzymes mediate meta-hydroxy-aromatic acid metabolism in R. palustris

The benzoyl-CoA pathway enzymes in *T. aromatica* are known to be both expressed and catalytically active with *meta*-hydroxy-aromatic substrates (Dangel *et al.* 1991; Laempe *et al.* 2001; Philipp *et al.* 2002; Ding *et al.* 2008). However, it is not known whether *T. aromatica* requires the activity of benzoyl-CoA pathway enzymes for *meta*-hydroxy-aromatic acid degradation or if other pathways contribute to metabolism in this organism. In the present study, we show that photo-heterotrophic metabolism of protocatechuate or *m*HB is only observed in *R. palustris* under conditions that increase benzoyl-CoA pathway activity. Further, our data provide direct genetic evidence that the metabolism of *meta*-hydroxy-aromatic compounds requires the activity of benzoyl-CoA pathway enzymes in *R. palustris*.

For example, mutants lacking BadDEFG activity (Egland *et al.* 1997) are known to be incapable of utilizing benzoate as a sole organic carbon source under photoheterotrophic growth conditions. Whereas wild-type cells were capable of using *m*HB to support growth after cyclohexanoate was removed from culture medium (Figure 5-5A), cells lacking benzoyl-CoA reductase activity (strain CGA606) were unable to grow using mHB under identical conditions (Figure 5-5B). This is direct

evidence that BadDEFG activity is necessary for utilization of *m*HB as a growth substrate by *R. palustris* under photoheterotrophic conditions. Similarly, the properties of the *hbaB* mutant (strain CGA506; Table 5-3) indicates that HbaBCD activity is needed for metabolism of protocatechuate (Figure 5-3; Figure 5-4). Collectively, our data predict a potentially new role for *R. palustris* HbaBCD and BadDEFG enzyme activities in *para*-dehydroxylation and reductive dearomatization of *meta*-hydroxylated metabolic intermediates during photoheterotrophic growth with *meta*-hydroxy-aromatic acids as organic carbon sources. Although HbaBCD and BadDEFG may function directly in *para*-dehydroxylation of protocatechuyl-CoA and dearomatization of *m*HB-CoA, respectively (Figure 5-1B), the possibility remains that other uncharacterized enzymes function in reductive *meta*-dehydroxylation, producing alternative intermediates *p*HB-CoA and benzoyl-CoA as substrates for HbaBCD and BadDEFG (Figure 5-1A).

Inducers of benzoyl-CoA pathway gene expression enhance metabolism of additional substrates

Given the various conditions that enhanced *meta*-hydroxy-aromatic acid metabolism, we present a model for benzoyl-CoA pathway expression that enables *R. palustris* to utilize non-inducing compounds as growth substrates. In experiments in which one of the inducers of benzoyl-CoA pathway expression (*p*HB, benzoate, or cyclohexanoate) was used to enhance the metabolism of one of the *meta*-hydroxy-aromatic acids (*m*HB or protocatechuate), the ability to metabolize the new substrate may be a result of differing substrate specificities of the proteins in the pathway. Whereas the known inducers are allosteric effectors of the benzoyl-CoA transcription factors (Egland and Harwood 1999; Egland and Harwood 2000; Peres and Harwood 2006), leading to the upregulated expression of pathway genes, it is likely that the *meta*-hydroxy-aromatic acids either serve as very weak inducers or do not induce any expression of the benzoyl-CoA pathway. However, we hypothesize that, as is the case in *T. aromatica*, CoA ligases, HbaBCD, and BadDEFG exhibit broad substrate specificities (Heider *et al.* 1998; Laempe *et al.* 2001) and, once expressed,

facilitate catabolism of substrates that do not induce their expression and lead to growth in *R. palustris*. A slight variation of this model can help describe the results observed in cultures incubated with equimolar concentrations of protocatechuate and *m*HB as cosubstrates (Figure 5-3). Although we did not expect CGA009 to grow under this condition, a review of previous findings provides some explanation of these results. Earlier observations (Kim and Harwood 1991) showed that, although protocatechuate (and its metabolic derivatives) does not induce expression of the entire benzoyl-CoA pathway needed for exponential growth, it does induce expression of CoA-ligase enzymes required for thioesterification of aromatic acids at the onset of the pathway (Figure 5-1). Thus, the CoA ligase expressed in the presence of protocatechuate may have a sufficiently high affinity for *m*HB (versus protocatechuate) to allow the formation of *m*HB-CoA as a potential inducer for the expression of other benzoyl-CoA pathway genes (*e.g.*, *hbaBCD* and *badDEFG*). Thereafter, the products *m*HB catabolism may induce the expression of pathways needed for subsequent catabolism of protocatechuate. We hypothesize that each substrate (and its cognate intermediates) induces the expression of only a subset of genes needed for growth, and those genes not induced by it or its intermediates are compensated for by expression induced by the other meta-hydroxy-aromatic acid under this condition.

Different factors control benzoyl-CoA pathway metabolism in R. palustris and T. aromatica

Our results provide evidence of (i) similar roles played by benzoyl-CoA pathway enzymes in *T. aromatica* and *R. palustris* for the degradation of *meta*-hydroxy-aromatic acids and of (ii) the regulatory control of benzoyl-CoA pathway expression being restricted to a narrower set of allosteric effector molecules in *R. palustris* compared to *T. aromatica*. It appears that, although broad substrate specificity may be characteristic of the metabolic enzymes in both *T. aromatica* and *R. palustris*, the latter organism is lacking regulatory elements needed for protocatechuate- or *m*HB-supported photoheterotrophic growth. One unique feature of anaerobic aromatic metabolism in *R.*

palustris is that the transcription factors controlling the expression of benzoyl-CoA pathway genes (Egland and Harwood 1999; Egland and Harwood 2000) apparently respond to only a narrow set of allosteric effectors but trigger upregulation of benzoyl-CoA pathway enzymes, which exhibit catalytic activity toward a broader set of aromatic growth substrates. Here, we demonstrated that protocatechuate and *m*HB comprise a group of benzoyl-CoA substrates that are not able to support rapid growth unless other compounds metabolized by this pathway are present as cosubstrates. A third aromatic acid, *meta*-chlorobenzoate, may also be a member of this group as it has been reported to support photoheterotrophic growth of *R. palustris* in the presence of benzoate but does not support growth as a sole organic carbon source (Egland *et al.* 2001; Samanta and Harwood 2005). The findings presented here also provide insight into how selective pressures in nature may influence evolution of systems for metabolism of chemically related compounds. For example, it is possible that the observed behavior of *R. palustris* reflects the fact that *meta*-hydroxy-aromatic acids are typically found in nature in environments where a known inducer of the benzoyl-CoA pathway is also present. If this were the case, the common presence of both classes of aromatic acids in nature has apparently not required *R. palustris* to evolve systems to induce benzoyl-CoA pathway expression in the presence of only *meta*-hydroxy-aromatic acids.

In sum, we identified here conditions that enable the metabolism of *meta*-hydroxy-aromatic acids in *R. palustris*. We also provided genetic evidence that benzoyl-CoA pathway enzymes are needed for metabolism of protocatechuate and *m*HB in *R. palustris*. Our findings leave open the possibility of *R. palustris* metabolizing other aromatic compounds that are found in either plant lignin, the environment, or as a product of industrial activity in the presence of one or more benzoyl-CoA pathway inducers.

Acknowledgments

This study was supported by a Department of Energy Office of Science's Great Lakes Bioenergy Research Center grant (DE-FG02-07ER64495). D.L.G. was supported by a traineeship from the National Institute of General Medical Sciences (NIGMS) under biotechnology training grant T32 GM08349.

We extend our thanks to Carrie S. Harwood and members of the Harwood laboratory for advice and for providing wild-type and mutant *R. palustris* strains used in this study.

References

APHA/AWWA/WEF (2005) Standards methods for the examination of water and wastewater, 21 edn. American Public Health Association, Washington, DC, USA

Biegert T., Altenschmidt U., Eckerskorn C., Fuchs G. (1993). Enzymes of anaerobic metabolism of phenolic-compounds - 4-hydroxybenzoate-CoA ligase from a denitrifying *Pseudomonas* species. *European Journal of Biochemistry* 213(1):555-561.

Boll M., Fuchs G. (1995). Benzoyl-coenzyme A reductase (dearomatizing), a key enzyme of anaerobic aromatic metabolism - ATP dependence of the reaction, purification and some properties of the enzyme from *Thauera aromatica* strain K172. *European Journal of Biochemistry* 234(3):921-933.

Boll M., Fuchs G. (1998). Identification and characterization of the natural electron donor ferredoxin and of FAD as a possible prosthetic group of benzoyl-CoA reductase (dearomatizing), a key enzyme of anaerobic metabolism. *European Journal of Biochemistry* 251(3):946-954.

Brackmann R., Fuchs G. (1993). Enzymes of anaerobic metabolism of phenolic-compounds - 4-hydroxybenzoyl-CoA reductase (dehydroxylating) from a denitrifying *Pseudomonas* species. *European Journal of Biochemistry* 213(1):563-571.

Breese K., Boll M., Alt-Morbe J., Schagger H., Fuchs G. (1998). Genes coding for the benzoyl-CoA pathway of anaerobic aromatic metabolism in the bacterium *Thauera aromatica*. *European Journal of Biochemistry* 256(1):148-154.

Breese K., Fuchs G. (1998). 4-hydroxybenzoyl-CoA reductase (dehydroxylating) from the denitrifying bacterium *Thauera aromatica* - Prosthetic groups, electron donor, and genes of a member of the molybdenum-flavin-iron-sulfur proteins. *European Journal of Biochemistry* 251(3):916-923.

Carmona M., Teresa Zamarro M., Blazquez B., Durante-Rodriguez G., Juarez J.F., Valderrama J.A., Barragan M.J.L., Garcia J.L., Diaz E. (2009). Anaerobic catabolism of aromatic compounds: a genetic and genomic view. *Microbiology and Molecular Biology Reviews* 73(1):71-+.

Crosby H.A., Heiniger E.K., Harwood C.S., Escalante-Semerena J.C. (2010). Reversible N epsilon-lysine acetylation regulates the activity of acyl-CoA synthetases involved in anaerobic benzoate catabolism in Rhodopseudomonas palustris. *Molecular Microbiology* 76(4):874-888.

Crosby H.A., Pelletier D.A., Hurst G.B., Escalante-Semerena J.C. (2012). System-wide studies of N-lysine acetylation in *Rhodopseudomonas palustris* reveal substrate specificity of protein acetyltransferases. *Journal of Biological Chemistry* 287(19):15590-15601.

Dangel W., Brackmann R., Lack A., Mohamed M., Koch J., Oswald B., Seyfried B., Tschech A., Fuchs G. (1991). Differential expression of enzyme activities initiating anoxic metabolism of various aromatic compounds via benzoyl-CoA. *Archives of Microbiology* 155(3):256-262.

Ding B., Schmeling S., Fuchs G. (2008). Anaerobic metabolism of catechol by the denitrifying bacterium *Thauera aromatica* - a result of promiscuous enzymes and regulators? *Journal of Bacteriology* 190(5):1620-1630.

Dispensa M., Thomas C.T., Kim M.K., Perrotta J.A., Gibson J., Harwood C.S. (1992). Anaerobic growth of *Rhodopseudomonas palustris* on 4-hydroxybenzoate is dependent on AadR, a member of

the cyclic-AMP receptor protein family of transcriptional regulators. *Journal of Bacteriology* 174(18):5803-5813.

Egland P.G., Gibson J., Harwood C.S. (1995). Benzoate-coenzyme a ligase, encoded by *badA*, is one of 3 ligases able to catalyze benzoyl-coenzyme A formation during anaerobic growth of *Rhodopseudomonas palustris* on benzoate. *Journal of Bacteriology* 177(22):6545-6551.

Egland P.G., Gibson J., Harwood C.S. (2001). Reductive, coenzyme A-mediated pathway for 3-chlorobenzoate degradation in the phototrophic bacterium *Rhodopseudomonas palustris*. *Applied and Environmental Microbiology* 67(3):1396-1399.

Egland P.G., Harwood C.S. (1999). BadR, a new MarR family member, regulates anaerobic benzoate degradation by *Rhodopseudomonas palustris* in concert with AadR, an Fnr family member. *Journal of Bacteriology* 181(7):2102-2109.

Egland P.G., Harwood C.S. (2000). HbaR, a 4-hydroxybenzoate sensor and FNR-CRP superfamily member, regulates anaerobic 4-hydroxybenzoate degradation by *Rhodopseudomonas palustris*. *Journal of Bacteriology* 182(1):100-106.

Egland P.G., Pelletier D.A., Dispensa M., Gibson J., Harwood C.S. (1997). A cluster of bacterial genes for anaerobic benzene ring biodegradation. *Proceedings of the National Academy of Sciences of the United States of America* 94(12):6484-6489.

Geissler J.F., Harwood C.S., Gibson J. (1988). Purification and properties of benzoate-coenzyme A ligase, a *Rhodopseudomonas palustris* enzyme involved in the anaerobic degradation of benzoate. *Journal of Bacteriology* 170(4):1709-1714.

Gibson J., Dispensa M., Fogg G.C., Evans D.T., Harwood C.S. (1994). 4-Hydroxybenzoate-coenzyme A ligase from *Rhodopseudomonas palustris* - purification, gene sequence, and role in anaerobic degradation. *Journal of Bacteriology* 176(3):634-641.

Gibson J., Dispensa M., Harwood C.S. (1997). 4-Hydroxybenzoyl coenzyme A reductase (dehydroxylating) is required for anaerobic degradation of 4-hydroxybenzoate by *Rhodopseudomonas palustris* and shares features with molybdenum-containing hydroxylases. *Journal of Bacteriology* 179(3):634-642.

Gibson K.J., Gibson J. (1992). Potential early intermediates in anaerobic benzoate degradation by *Rhodopseudomonas palustris*. *Applied and Environmental Microbiology* 58(2):696-698.

Harrison F.H., Harwood C.S. (2005). The *pimFABCDE* operon from *Rhodopseudomonas palustris* mediates dicarboxylic acid degradation and participates in anaerobic benzoate degradation. *Microbiology-Sgm* 151727-736.

Harwood C.S., Burchhardt G., Herrmann H., Fuchs G. (1998). Anaerobic metabolism of aromatic compounds via the benzoyl-CoA pathway. *Fems Microbiology Reviews* 22(5):439-458.

Harwood C.S., Gibson J. (1986). Uptake of benzoate by *Rhodopseudomonas palustris* grown anaerobically in light. *Journal of Bacteriology* 165(2):504-509.

Harwood C.S., Gibson J. (1988). Anaerobic and aerobic metabolism of diverse aromatic compounds by the photosynthetic bacterium *Rhodopseudomonas palustris*. *Applied and Environmental Microbiology* 54(3):712-717.

Harwood C.S., Parales R.E. (1996). The beta-ketoadipate pathway and the biology of self-identity. *Annual Review of Microbiology* 50553-590.

Heider J., Boll M., Breese K., Breinig S., Ebenau-Jehle C., Feil U., Gad'on N., Laempe D., Leuthner B., Mohamed M.E.S., Schneider S., Burchhardt G., Fuchs G. (1998). Differential induction of enzymes involved in anaerobic metabolism of aromatic compounds in the denitrifying bacterium *Thauera aromatica*. *Archives of Microbiology* 170(2):120-131.

Kim M.K., Harwood C.S. (1991). Regulation of benzoate-CoA ligase in *Rhodopseudomonas palustris*. *Fems Microbiology Letters* 83(2):199-203.

Laempe D., Eisenreich W., Bacher A., Fuchs G. (1998). Cyclohexa-1,5-diene-1-carboxyl-CoA hydratase, an enzyme involved in anaerobic metabolism of benzoyl-CoA in the denitrifying bacterium *Thauera aromatica*. *European Journal of Biochemistry* 255(3):618-627.

Laempe D., Jahn M., Breese K., Schagger H., Fuchs G. (2001). Anaerobic metabolism of 3-hydroxybenzoate by the denitrifying bacterium *Thauera aromatica*. *Journal of Bacteriology* 183(3):968-979.

Laempe D., Jahn M., Fuchs G. (1999). 6-hydroxycyclohex-1-ene-1-carbonyl-CoA dehydrogenase and 6-oxocycsohex-1-ene-1-carbonyl-CoA hydrolase, enzymes of the benzoyl-CoA pathway of anaerobic aromatic metabolism in the denitrifying bacterium *Thauera aromatica*. *European Journal of Biochemistry* 263(2):420-429.

Larimer F.W., Chain P., Hauser L., Lamerdin J., Malfatti S., Do L., Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R., Gibson J.L., Hanson T.E., Bobst C., Torres J., Peres C., Harrison F.H., Gibson J., Harwood C.S. (2004). Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris*. *Nature Biotechnology* 22(1):55-61.

Merkel S.M., Eberhard A.E., Gibson J., Harwood C.S. (1989). Involvement of coenzyme A thioesters in anaerobic metabolism of 4-hydroxybenzoate by *Rhodopseudomonas palustris*. *Journal of Bacteriology* 171(1):1-7.

Pan C., Oda Y., Lankford P.K., Zhang B., Samatova N.F., Pelletier D.A., Harwood C.S., Hettich R.L. (2008). Characterization of anaerobic catabolism of *p*-coumarate in *Rhodopseudomonas palustris* by integrating transcriptomics and quantitative proteomics. *Molecular & Cellular Proteomics* 7(5):938-948.

Pelletier D.A., Harwood C.S. (1998). 2-Ketocyclohexanecarboxyl coenzyme A hydrolase, the ring cleavage enzyme required for anaerobic benzoate degradation by *Rhodopseudomonas palustris*. *Journal of Bacteriology* 180(9):2330-2336.

Peres C.M., Harwood C.S. (2006). BadM is a transcriptional repressor and one of three regulators that control benzoyl coenzyme A reductase gene expression in *Rhodopseudomonas palustris*. *Journal of Bacteriology* 188(24):8662-8665.

Perrotta J.A., Harwood C.S. (1994). Anaerobic metabolism of cyclohex-1-ene-1-carboxylate, a proposed intermediate of benzoate degradation by *Rhodopseudomonas palustris*. *Applied and Environmental Microbiology* 60(6):1775-1782.

Phelps C.D., Young L.Y. (1997). Microbial metabolism of the plant phenolic compounds ferulic and syringic acids under three anaerobic conditions. *Microbial Ecology* 33(3):206-215.

Philipp B., Kemmler D., Hellstern J., Gorny N., Caballero A., Schink B. (2002). Anaerobic degradation of protocatechuate (3,4-dihydroxybenzoate) by Thauera aromatica strain AR-1. *Fems Microbiology Letters* 212(1):139-143.

Samanta S.K., Harwood C.S. (2005). Use of the *Rhodopseudomonas palustris* genome sequence to identify a single amino acid that contributes to the activity of a coenzyme A ligase with chlorinated substrates. *Molecular Microbiology* 55(4):1151-1159.

Schuhle K., Gescher J., Feil U., Paul M., Jahn M., Schagger H., Fuchs G. (2003). Benzoate-coenzyme A ligase from *Thauera aromatica*: an enzyme acting in anaerobic and aerobic pathways. *Journal of Bacteriology* 185(16):4920-4929.

Yilmaz L.S., Kontur W.S., Sanders A.P., Sohmen U., Donohue T.J., Noguera D.R. (2010). Electron partitioning during light and nutrient-powered hydrogen production by *Rhodobacter sphaeroides*. *Bioenergy Research* 3(1):55-66.

CHAPTER 6: Harnessing pathway functions for biocatalytic transformation of lignin

Although identifying the substrates, intermediates, products, catalysts, and underlying mechanisms that govern the β-etherase and benzoyl-CoA pathway reactions (Chapters 2–5) has revealed much concerning the relevance of these pathways in the field of biotechnology, identifying a platform with which to integrate and implement their functions is imperative to the development of full-scale lignin biodegradation systems. As the use of cellulosic feedstocks derived from ligno-cellulose for industrial biomanufacturing becomes more prevalent (Palmqvist and Hahn-Hägerdal 2000a; Palmqvist and Hahn-Hägerdal 2000b; Alriksson *et al.* 2010; Yang and Sen 2010), so follows the need for strategies for deriving valued chemicals from lignin or, at the very least, detoxifying lignin derivatives that inhibit the cellulose-metabolizing microbial catalysts in downstream processes (Palmqvist and Hahn-Hägerdal 2000a).

For my dissertation, I conducted fundamental research into biodegradation reactions with *in vitro* investigation of β-etherase enzymes and organismal-level *in vivo* studies of aromatic metabolism via the benzoyl-CoA pathway in *R. palustris*. The conclusions from these studies have helped to delineate the route through which lignin can be metabolized and potentially converted to valuable intermediates on a chemical level (Figure 6-1). However, a combination of fundamental and systems-level approaches is required in order to harness the capabilities of these pathways in full-scale processes. In this chapter, I describe future research objectives for integrating the fundamental aspects of β-etherase and benzoyl-CoA pathway reactions and examine potential strategies for microbe-catalyzed conversion of lignin to valued commodities.

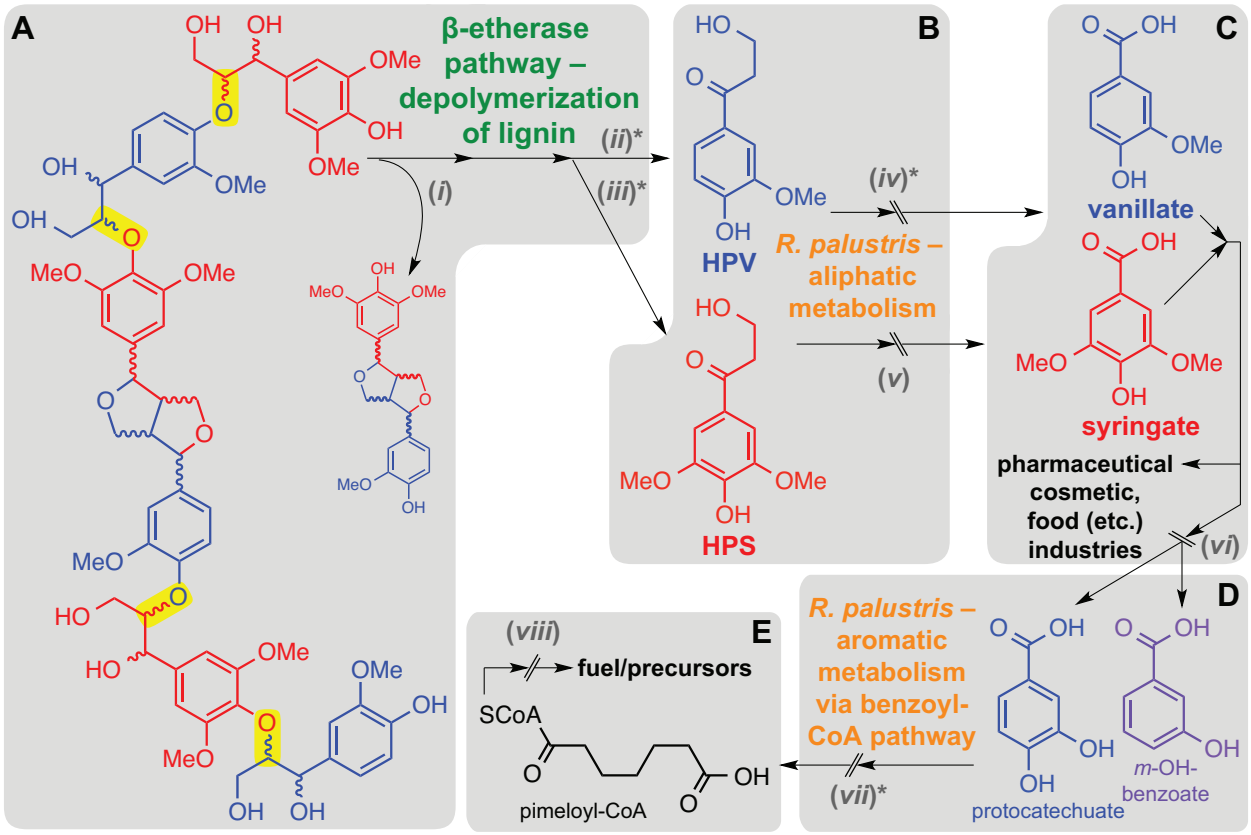

Figure 6-1. Overall scheme for lignin biodegradation via β-etherase and benzoyl-CoA pathway-mediated cleavage of lignin β-ether linkages and metabolism of monoaromatic derivatives. Scheme processes studied for my thesis are denoted with an asterisk (*). Steps (*i*)–(*viii*) are described in the text. **(A)** β-Etherase pathway catalyzing the conversion of lignin oligomers containing β-ether linkages (highlighted in yellow) to monoaromatic products and representative di-aryl byproducts. **(B)** Metabolism of the aliphatic moieties of monoaromatic derivatives in *R. palustris*, yielding mono-aromatic carboxylates. **(C)** Hypothetical conversion of lignin-derived aromatic carboxylates to *m*-hydroxy-aromatic carboxylates. **(D)** Metabolism of protocatechuate and *m*-OH-benzoate via the benzoyl-CoA pathway in *R. palustris*. **(E)** Hypothetical chemical reduction of pimeloyl-CoA to fuel precursors or valued commodities.

Future Research Objectives – Enzymology and Metabolism

In vitro β-ether cleavage in lignin polymers using β-etherase pathway enzymes

As described in the preceding chapters, the β-etherase pathway enzymes catalyze cleavage of β-ether linkages in synthetic model lignin compounds (Figure 6-1A) (Masai *et al.* 2003; Sato *et al.* 2009; Gall *et al.* 2014a; Gall *et al.* 2014b). The genome of *Sphingobium* sp. strain SYK-6, an isolate

from a paper-pulping mill, encodes several pathways for the degradation of aromatic compounds (Masai *et al.* 2007) and this organism is able to use synthetic β-ether-linked aromatic dimers as sources of carbon and reducing power for growth (Fukuzumi and Katayama 1977; Katayama and Fukuzumi 1978; Katayama and Fukuzumi 1979a; Katayama and Fukuzumi 1979b). Previous findings suggest that metabolism of model dimers occurs cytoplasmically, indicating that the organism is able to transport its substrates across the cellular membrane (Masai *et al.* 1989). To date, it remains unclear whether strain SYK-6 gains access to β-ether-linked dimers in nature by (a) scavenging them as degradation products of extracellular lignin oxidation (*i.e.*, from fungal export of hydroxyl radicals) (Leonowicz *et al.* 2001; Otsuka *et al.* 2003), (b) importing lignin polymers across the cell membrane to the cytoplasm, or (c) exporting Lig enzymes (and their required co-substrates – *e.g.*, NAD^+ and GSH) to enable extracellular β-ether cleavage in lignin polymers. Incorporation of (a) to aid strain SYK-6 is undesirable due to the fact that the fungal pathways for β-ether cleavage are highly oxidative and non-selective (Leonowicz *et al.* 2001; Otsuka *et al.* 2003). Further, options (b) and (c) seem unlikely because of lignin's low solubility and because of the cost that would be incurred by exportation of the enzymatic cosubstrates (respectively). Although growth experiments using higher molecular weight synthetic lignin oligomers (*e.g.*, β-ether-linked trimers or tetramers) may reveal if strain SYK-6 is capable of either transport mechanism, because this organism only possesses oxidative pathways for monoaromatic catabolism (Masai *et al.* 2007), it may be advantageous to express these genes heterologously in an organism capable of anaerobic metabolism (*e.g.*, *R. palustris*), or to use the enzymes *in vitro* to catalyze lignin depolymerization.

Although *in vitro* β-etherase pathway functions have been evaluated with model lignin dimers as substrates, the abilities of these enzymes to catalyze the cleavage in higher-order oligomers have not been tested. To achieve this, coupled biochemical assays (*i.e.*, with Cα-dehydrogenases, β-etherases, and GSH lyases) with β-ether-linked oligomeric substrates are necessary. Of interest in these experiments is whether or not Lig enzymes are structurally and functionally capable of

catalyzing cleavage in compounds containing multiple β-ether linkages and non-β-ether-type linkages that are typically found in lignin (Figure 6-1A(*i*)). If *in vitro* cleavage of the oligomeric model compounds is successful, it will be intriguing to see whether high yields of HPV (Figure 6-1A(*ii*)) and HPS (Figure 6-1A(*iii*)) can be obtained from extracted natural plant lignin. The α(*R*)- and α(*S*)-stereospecific dehydrogenases, LigD and LigN (respectively) were expressed and purified, along with the β-etherases (LigE, F, P) and GSH lyase (LigG) purified previously (Chapters 2–3) to support ongoing studies that are described further in the following section.

In vitro replenishment of β-etherase pathway cosubstrates NAD⁺ and GSH using glutathione reductase

Although many of the enzymes required for cleavage of β-ether bonds have been identified in strain SYK-6 (Figure 6-2), the identity of the presumed stereochemical complement of LigG glutathione lyase, and the route through which NADH and GSSG are restored to NAD^+ and GSH remain in question. Identification of the LigG complement via mutational or biochemical studies may be necessary in order to maximize the yield of HPV and HPS (Figure 6-1B), and thus, downstream targeted intermediates from this scheme, given that a 1:1 ratio of β(*R*)- and β(*S*)-GS-HPV is expected from cleavage of the racemic β-ether linkages in lignin. Further, implementing a strategy for transferring a hydride ion from NADH to GSSG, thereby restoring the pathway's co-substrates, NAD^+ and GSH, will be vital to reducing cosubstrate input costs in a potential full-scale system (Figure 6-2).

Typically, reduction of GSSG to GSH (2 eq.) is facilitated by glutathione reductase (EC 1.8.1.7) using NAD(P)H (2′-phosphorylated NADH) as the hydride donor (Bohme *et al.* 2000). The halo-tolerant bacterium *Allochromatium vinosum* strain DSM180 glutathione reductase (AvGR, Figure 6-2) has been characterized as having a higher affinity for NADH than for NAD(P)H (Chung and Hurlbert 1975; Reiter *et al.* 2014). After having the codon-optimized (for expression in *E. coli*) *avGR*

gene synthesized (Life Technologies – GeneArt®), I expressed and purified the AvGR enzyme from *E. coli* and have conducted biochemical assays in which NADH and GSSG were used as cosubstrates, yielding coproducts NAD^+ and GSH (2 eq.) (Figure 6-2). With Cα-dehydrogenases (LigD and LigN), β-etherases (LigE and LigF), GSH lyase (LigG), and NADH-dependent glutathione reductase (AvGR) in hand, it is now possible to test whether the activities of these enzymes can be coupled to catalyze cleavage of β-ether linkages using only catalytic quantities (*e.g.*, ~0.01 eq.) of the pathway's co-substrates, NAD^+ and GSH. For example, in a system using guaiacyl-β-guaiacyl ether (1 eq., Figure 6-2), NAD^+ (0.01 eq.), and GSH (0.02 eq.) as cosubstrates and Cα-dehydrogenases (LigD and LigN), β-etherases (LigE and LigF), and GSH lyases (LigG and its complement) as catalysts, guaiacol (1 eq.) and HPV (1 eq.) are the expected products since NADH and GSSG will be used for replenishing NAD^+ and GSH. Thus, by including AvGR in the *in vitro* depolymerization assays, a minimal supply of co-substrates may lead to complete β-ether cleavage and derivation of HPV and HPS from model lignin compounds and natural, plant-extracted lignins.

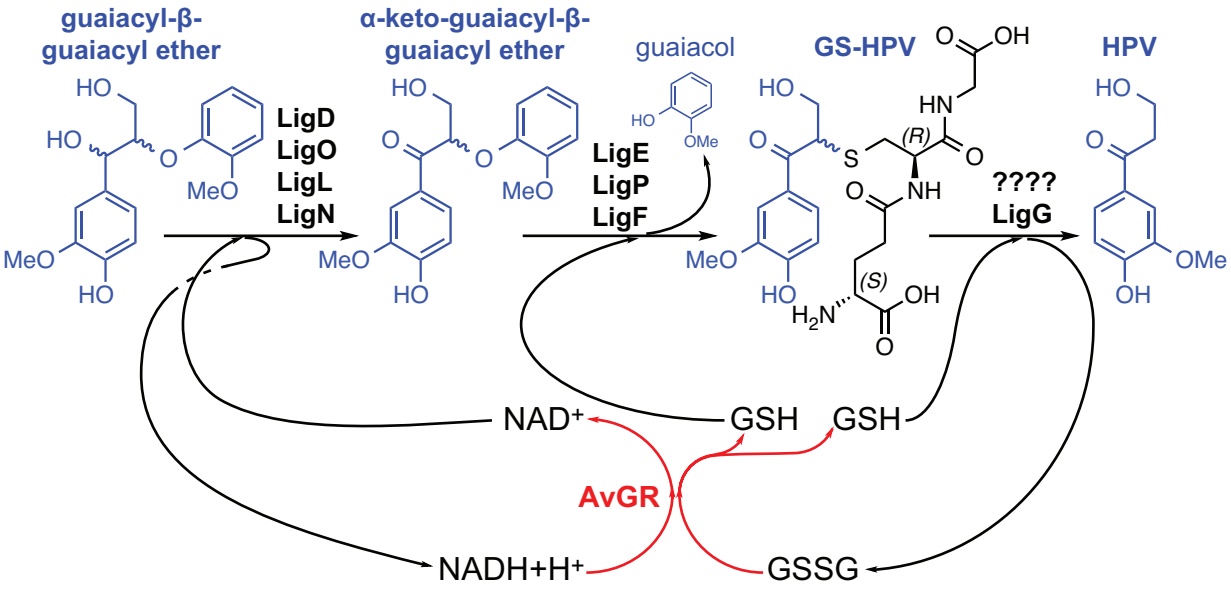

Figure 6-2. The β-etherase pathway-mediated conversion of guaiacyl-β-guaiacyl ether, NAD^+, and GSH (2 eq.) to guaiacol, HPV, NADH, and GSSG, in addition to the NADH-dependent reduction of GSSG (in red) by the *Allochromatium vinosum* strain DSM glutathione reductase (AvGR), yielding the original cosubstrates, NAD^+ and GSH (2 eq.).

In vivo metabolism of HPV and HPS aliphatic side-chains in R. palustris

Because neither HPV nor HPS are commercially available compounds, neither had been previously tested as growth substrates of *R. palustris*. Given that these are the expected mono-aromatic derivatives of β-ether cleavage in lignin however, I obtained HPV (synthesized by Fachuang Lu in the John Ralph laboratory) and tested the ability of *R. palustris* to use it as a sole source of reducing power in photoheterotrophic growth experiments. This experiment revealed that the wild-type strain CGA009 was able to use HPV for growth, first utilizing the aliphatic moiety of HPV (T_d = 4.6 d) with concomitant release of vanillate (Figure 6-1C) to the culture medium, followed by relatively slow growth during uptake of vanillate (T_d = 63.6 d) (Figure 6-3). Although vanillate-supported growth was slow, this finding indicates that HPV metabolism causes expression of genes that enable vanillate degradation and that expression of these genes is not induced by vanillate alone, since vanillate did not support growth as a sole source of reducing equivalents. As *R. palustris* was also unable to utilize syringate (Figure 6-1C) as a sole source of nutrients, it will be intriguing to see whether, in similar fashion with which *R. palustris* converts HPV to vanillate (Figure 6-1B(*iv*)), *R. palustris* may be used as a biocatalyst for the conversion of HPS to syringate (Figure 6-1B(*v*)), yielding two aromatic carboxylates of interest for industrial uses (Itoh *et al.* 2009).

Of equally desirable value to the food, pharmaceutical, and cosmetic industries are the cognate aromatic aldehydes of vanillate and syringate, vanillin and syringaldehyde (Sinha *et al.* 2008). Previously, it was reported that metabolism of ferulate (the cognate hydroxycinnamate of HPV) is metabolized through vanillin as a metabolic intermediate, which is then oxidized to vanillate by an unidentified dehydrogenase in *R. palustris* (Pan *et al.* 2008; Hirakawa *et al.* 2012) (Chapter 1, Figure 1-5B). Although it is unclear whether HPV metabolism proceeds through vanillin as an intermediate

or directly to vanillate (Figure 1-5C), further evaluation of this pathway and chromosomal deletion of the gene encoding for vanillate dehydrogenase activity may yield a mutant strain that can convert HPV and HPS to vanillin and syringaldehyde. Future investigations into this matter should also take advantage of the existence of *R. palustris* mutants (Chapter 5) that are deficient in benzoyl-CoA pathway activities. Although degradation of vanillate by the wild type strain was slow (Figure 6-3), HPV-supported growth of the *hbaBCD* (Figure 1-5A) *R. palustris* strain may lead to equimolar production of vanillate.

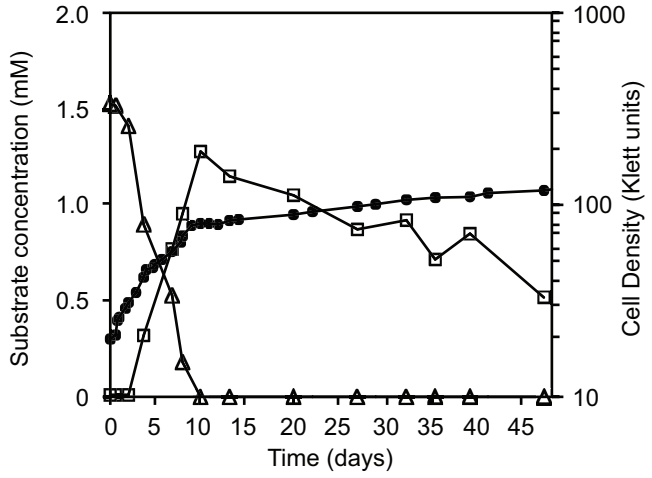

Figure 6-3. HPV (1.5 mM)-supported photoheterotrophic growth of *R. palustris* wild-type strain CGA009. Cell density (•) is reported in Klett units (right y-axis). Medium HPV (Δ) and vanillate (□) concentrations (left y-axis) are indicated by the corresponding symbols.

Microbial strategies for metabolism of vanillate and syringate

The known ring-shaped substrates of the benzoyl-CoA pathway, through which benzoyl-CoA is formed as a central intermediate, and that support growth of *R. palustris* as a sole source of reducing equivalents, are *p*-OH-benzoate, benzoate, cyclohexanoate, and *p*-coumarate (Merkel *et al.* 1989; Perrotta and Harwood 1994; Egland *et al.* 1997; Gibson *et al.* 1997; Egland *et al.* 2001; Harrison and Harwood 2005; Pan *et al.* 2008; Hirakawa *et al.* 2012). The expected aromatic

carboxylate derivatives from lignin via the combined activities of the β-etherase pathway enzymes and aliphatic metabolism in *R. palustris*, vanillate and syringate (Figure 6-1C), differ from *p*-OH-benzoate by either a single *m*-methoxyl or two *m,m*-methoxy groups, respectively. Although my investigations indicated that neither vanillate, syringate, nor any *m*-substituted aromatic carboxylate that were tested could support growth as a sole source of reducing equivalents, the presence of a known benzoyl-CoA pathway substrate (*e.g.*, benzoate) resulted in induction of benzoyl-CoA pathway expression, and thus, enabled *R. palustris* to degrade protocatechuate and *m*-OH-benzoate (Figure 6-1D) as described previously (Chapter 5). To date, no photoheterotrophic conditions have been found in which *R. palustris* definitively degraded syringate to support growth, and only slow degradation of vanillate (once derived from HPV) was observed (Figure 6-3). Thus, further biological transformation of lignin-derived vanillate and syringate to additional valued metabolic intermediates may require enzymatic activities that are not encoded in the *R. palustris* genome.

Chromosomal integration of exogenous DNA or heterologous expression may enable *R. palustris* to derive known benzoyl-CoA pathway substrates such as protocatechuate or *m*-OH-benzoate from vanillate and syringate (Figure 6-1C(*vi*)). Given that protocatechuate and *m*-OH-benzoate degradation via reductive transformation by "upper benzoyl-CoA pathway" enzymes (Harrison and Harwood 2005) was demonstrated (Figure 6-1D(*vii*)), the aliphatic product of ring cleavage likely to arise as an intermediate in the "lower benzoyl-CoA pathway" is pimeloyl-CoA (the product of *p*-OH-benzoate and benzoate metabolism, Figure 6-1E) or perhaps, a hydroxy-analog. Because it may be possible to biochemically upgrade fatty carboxylic thioesters or dicarboxylic fatty-CoA thioesters (*e.g.*, pimeloyl-CoA) to liquid transportation fuels such as fatty alcohols or alkanes (Akhtar *et al.* 2013), it appears that achieving conversion of vanillate and syringate to aromatic compounds that are metabolized via the benzoyl-CoA pathway is the missing link in a pipeline for deriving fuel precursors from lignin (Figure 6-1).

Although the reason for the slow growth rate exhibited while using HPV-derived vanillate (Figure 6-3) is unclear, introduction of the *Sphingobium* sp. strain SYK-6 *ligM* gene to *R. palustris* could drastically improve vanillate degradation rates, considering that LigM catalyzes demethylation of vanillate to protocatechuate (Figure 6-4A), a substrate that supported a four-fold higher doubling rate in *R. palustris* (Chapter 5) (Sonoki *et al.* 2000; Abe *et al.* 2005). Also, strain SYK-6 uses LigM for oxygen-independent di-demethylation of syringate to gallate (Figure 6-4B). LigM uses tetrahydrofolate (THF) as a cosubstrate, yielding *m*-hydroxy-aromatic carboxylates and coproduct 5-methyltetrahydrofolate (THF-CH_3) that undergoes further oxidation, supplying reducing equivalents for the organism and eventually restoring THF to its premethylated state. If *R. palustris* can tolerate the foreign *ligM* gene (there is no homolog in *R. palustris*), and complement the LigM enzyme via upregulation of THF biosynthesis, then the resulting strain is likely to derive protocatechuate from vanillate, and gallate from syringate.

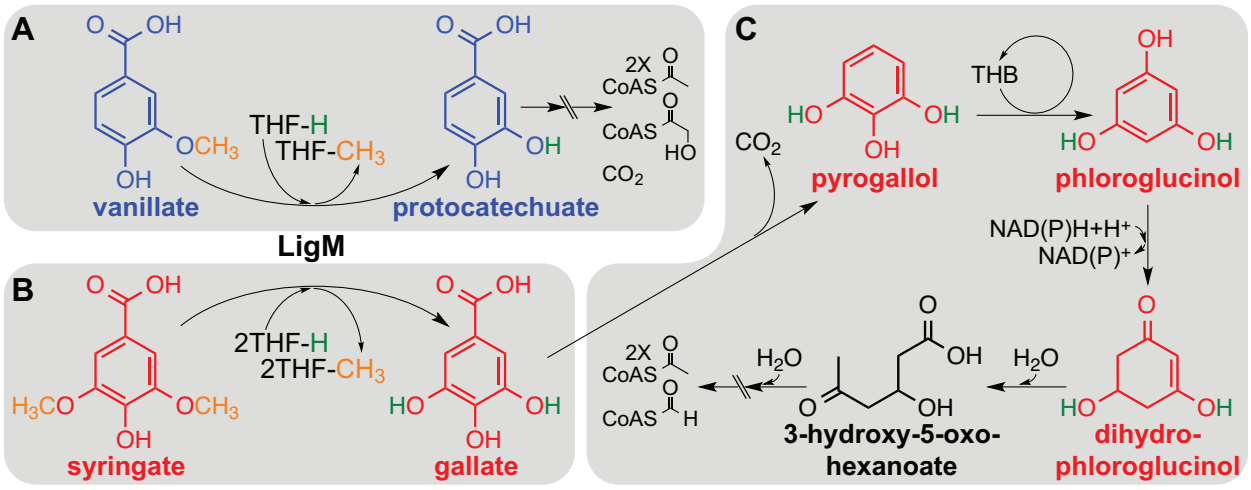

Figure 6-4. Metabolic routes for anaerobic degradation of **(A)** vanillate to protocatechuate and TCA cycle intermediates, **(B)** syringate to gallate, and **(C)** gallate to TCA cycle intermediates. THF-dependent LigM is encoded in the genome of *Sphingobium* sp. strain SYK-6 (Abe *et al.* 2005). Decarboxylation of gallate to pyrogallol is described in *Eubacterium oxidoreducens* (Haddock and Ferry 1993) and THB-dependent transhydroxylation of pyrogallol to phloroglucinol and subsequent steps are described in *Pelobacter acidigallici* (Reichenbecher *et al.* 1994; Reichenbecher and Schink 1999). Abbreviations: THF, tetrahydrofolate; THB, 1,2,3,5-tetrahydroxybenzene; $NADP^+$, 2′-phosphorylated nicotinamide adenine dinucleotide.

Unfortunately, the genes encoding for anaerobic degradation of gallate in microorganisms have not yet been identified. In *Eubacterium oxidoreducens*, gallate decarboxylase has been studied in crude cell extract experiments (Haddock and Ferry 1993) and catalyzes formation of pyrogallol (Figure 6-4C). Anaerobic degradation of pyrogallol has been described in *Pelobacter acidigallici* (Reichenbecher *et al.* 1994; Reichenbecher and Schink 1999), which uses a unique molybdoenzyme that apparently involves 1,2,3,5-tetrahydroxybenzene (THB) as both a cosubstrate and coproduct. In this reaction, phloroglucinol (*i.e.*, 1,3,5-trihydroxybenzene) is actually derived from cosubstrate THB as its 2-hydroxyl is reductively removed and transferred, oxidatively, to the 5-position of phloroglucinol (*i.e.*, transhydroxylated), yielding THB as a coproduct (Reichenbecher *et al.* 1994). From there, it is believed that typical β-oxidation mechanisms are utilized for deriving TCA cycle intermediates from phloroglucinol because the *m*,*m*,*m*-trihydroxylated compound is no longer truly aromatic (Reichenbecher and Schink 1999). Like the benzoyl-CoA pathway, with which six electrons are added to a *p*-hydroxy-aromatic prior to ring cleavage, the anaerobic gallate degradation pathway is reductive, adding two electrons to the ring and eventually yielding 3-hydroxy-5-oxo-hexanoate as an aliphatic intermediate. It will be necessary, however, to identify the genes encoding the aforementioned enzymes before these activities may be implemented in *R. palustris*.

In vivo conversion of aromatic carboxylates to biofuel precursors in R. palustris

Although it may be possible to integrate foreign physiological strategies with benzoyl-CoA pathway metabolism in *R. palustris* to create a strain that can convert vanillate and syringate to aliphatic biofuel precursors, several challenges to this goal exist. Given that we hope to derive metabolic intermediates that are reduced in comparison to their aromatic precursors, the viability of the strain in this process will rely on supplementation of an alternative electron donor. In the

best-case scenario, the strain could obtain electrons from lignocellulose-derived sugars, using a fraction of the reducing equivalents to support growth and the remainder for the reductive dearomatization of aromatics without catabolizing the aliphatic products of ring cleavage. Although accomplishing this would be extraordinary, extraction of the cellular intermediates, distillation or purification, as well as chemical modification are additional concerns that will need to be addressed if microbial catalysts are to be used for *in vivo* aromatic metabolism in which the desired products are not readily exported by the strain in use. Given that *R. palustris* has the inherent ability to convert HPV (and possibly HPS) to vanillate (and syringate), and export the aromatic carboxylate from the cytoplasm to the extracellular medium, I explore exploitation of this possibility on a larger scale in the following sections.

Future Research Objectives – Development of Lignin Biorefinery Processes

Large-scale production of thermostable AvGR and β-etherase pathway enzyme variants

Each enzyme used during the course of my dissertation research was obtained from a culture of *E. coli* heterologously expressing a single target enzyme-encoding gene. In other words, purification of LigD, LigN, LigE, LigF, LigG, and AvGR required the growth of six separate cultures and six protein purification procedures. The time and effort required to do this on a larger scale would render this *in vitro* approach inefficient. Thus, future studies aimed at obtaining high yields of these enzymes should take advantage of strategies for expressing multiple genes in *E. coli* and high-throughput protein extraction methods. To accomplish this, the *ligD*, *ligN*, *ligE*, *ligF*, *ligG*, and *avGR* genes may be assembled in a recombinant operon and expressed simultaneously in *E. coli* (Ninh *et al.* 2015). However, because this approach requires the use of high temperatures (70 °C) for *E. coli* lysis and enzyme extraction, and given that this temperature is known to deactivate some Lig enzymes

(Chapter 4), it will be necessary to first obtain genes that encode thermostable variants of each enzyme.

In the case of glutathione reductase (GR), this may be accomplished by selecting an alternative glutathione reductase from a thermophilic species. However, inspection of the available GR amino acid sequences from thermophilic organisms suggest that site-directed mutagenesis may be necessary for achieving high NADH affinity. It is suggested that differentiation between NADH and NAD(P)H is governed by two key amino acids in the enzyme: at position 198, a negatively charged residue (glutamate in AvGR) stabilizes the NADH 2′-hydroxyl or a positively charged residue (*e.g.*, lysine) creates hydrogen bonds with the NAD(P)H 2′-phosphate (Schierbeek *et al.* 1989); at position 178, it is suggested that glycine conveys NADH affinity whereas alanine and bulkier amino acids are found in NAD(P)H-specific enzymes (Scrutton *et al.* 1990). Although no thermophile-derived sequence meeting both criteria was found, the GR sequence encoded in the genome of the thermophilic cyanobacterium *Chroococcidiopsis thermalis* (Li *et al.* 2014) contains the necessary glycine-178, but requires variation of its lysine-198 to glutamate.

Unfortunately, no Lig homolog-encoding genes have been identified in the available thermophilic genome sequences in public databases. An alternative approach to obtaining thermostable enzyme variants is to introduce the mesophile-derived gene into a thermophilic organism and then, via directed evolution, place selective pressure on the thermophile (Liao *et al.* 1986). For example, vanillate-degrading *Bacillus* sp. have been isolated previously, and such a strain may be a reasonable candidate for hosting directed evolution (Peng *et al.* 2003). Optimally, the ability of the *Bacillus* strain to use HPV as a growth substrate should be tested, and at the very least, it should be confirmed that the organism does not have alternative ways of degrading the target substrate, α-keto-guaiacyl-β-guaiacyl ether (Figure 6-2), in this case. Given that the parent strain is able to grow using HPV as a substrate, but unable to grow using α-keto-guaiacyl-β-guaiacyl

ether, transformation of the *ligE*, *ligF*, and *ligG* genes (these genes are encoded in a single operon in strain SYK-6) to the parent strain (on an antibiotic resistant plasmid) and growing the transformant under thermophilic conditions using α-keto-guaiacyl-β-guaiacyl ether (and perhaps trace amounts of vanillate) as the growth substrate would select for growth of transformants that have created mutations in the *ligE*, *ligF*, and *ligG* genes that confer enzymatic thermostability. If these can be isolated, a similar procedure may be employed for isolation of thermostable LigD- and LigN-encoding genes using guaiacyl-β-guaiacyl ether (Figure 6-2) as the growth substrate.

Large-scale production of vanillate and syringate from lignin using biological catalysts

Given the industrial demands for vanillate and syringate (and cognate aldehydes), the use of *R. palustris* as biocatalyst of partial HPV and HPS degradation may be of value. Syringate has been implicated as a preventative agent for treatment of liver hepatotoxicity (Itoh *et al.* 2009) and vanillate is employed as an inhibitor of snake venom activity (Dhananjaya *et al.* 2009), carcinogenesis (Vetrano *et al.* 2005), apoptosis (Shang-Ming *et al.* 2008), and inflammation (Itoh *et al.* 2009), in addition to being used in fragrances. Thus, it may be possible to use lignin (post-separation from lignocellulose) as a feedstock for vanillate and syringate production in a multi-stage system in which thermophilic variants of AvGR and the β-etherase pathway enzymes are used for *in vitro* depolymerization of lignin and, in a downstream reactor, *R. palustris* catalyzes conversion of HPV and HPS to vanillate and syringate, or vanillin and syringaldehyde.

A simplified process flow schematic of such a system is shown in Figure 6-5. It involves an aerated basin supplied with nutrients to support growth of *E. coli* cells that are constitutively expressing the desired thermostable variants of β-etherase pathway and AvGR enzymes (Stage 1). Downstream, application of heat is necessary for extraction of the thermostable enzymes and inactivation of native *E. coli* proteins for prevention of undesired side-reactions (Stage 2). Insoluble *E. coli* cell debris may be removed from the system via filtration or flotation (Stage 3) and either

sent to waste or downstream to Stage 5 where it can be mineralized by *R. palustris*, a known degrader of long-chain fatty acids (Harrison and Harwood 2005). Combined with enzymes produced in Stages 1–3 and catalytic amounts of the β-etherase pathway cosubstrates NAD^+ and GSH is the raw lignin feedstock that will require the use of a water-miscible organic solvent (*e.g.*, DMSO) to introduce the substrate to aqueous enzymes (Stage 4). Hypothetically, HPV and HPS are derived from lignin in Stage 4 and sent downstream for partial metabolism by *R. palustris* (Stage 5), and, after gravity settling of biomass (Stage 6), the aqueous product stream contains a mixture of either vanillate and syringate, or vanillin and syringaldehyde, depending on the type of *R. palustris* strain used.

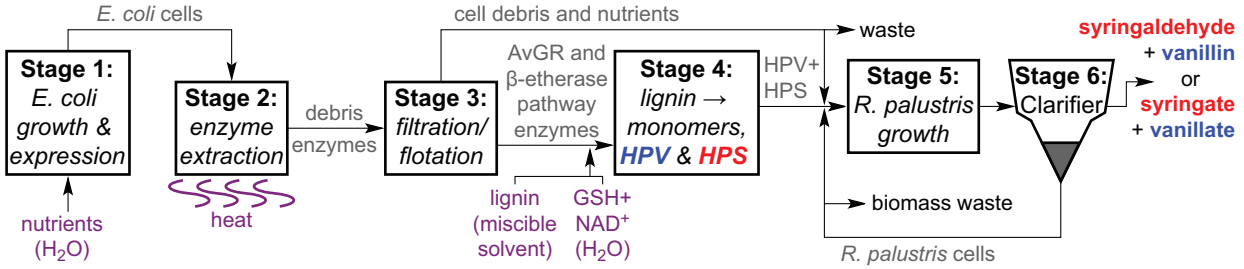

Figure 6-5. Hypothetical process flow scheme for conversion of lignin to vanillate and syringate or vanillin and syringaldehyde. Stage 1: Nutrients are supplemented in an aqueous system to support *E. coli* growth and heterologous expression of thermostable AvGR and β-etherase pathway enzyme variants. Stage 2: Heat is used to extract thermophilic Lig enzymes (as well as AvGR) and inactivate native enzymes of *E. coli*. Stage 3: Filtration or flotation for removal of cell debris and diversion to Stage 5. Stage 4: Lignin and cosubstrates NAD^+ and GSH are combined with enzymes from Stage 2, converting lignin to HPV and HPS. Stage 5: Growth of *R. palustris* and conversion of HPV and HPS to aromatic carboxylates or aldehydes. Stage 6: Clarifier for gravity separation of *R. palustris* biomass from monoaromatic products. System inputs and intermediates are shown in purple and grey, respectively. Solvents are shown in parentheses.

The development of biologically catalyzed lignin-to-products process flow schemes (*e.g.*, Figure 6-5) has not kept pace with their chemical counterparts (Rahimi *et al.* 2013; Rahimi *et al.* 2014) but, ultimately, the conversion of lignin to valued commodities may involve a combination of chemical and biological processes. Although the selectivity of the β-etherase pathway makes it an attractive

strategy for approaching lignin depolymerization, chemical strategies may accomplish this with greater efficiency because of lignin's low solubility in aqueous systems and due to the sensitivity of enzymes in organic solvents. Chemical catalysts of β-ether cleavage and derivation of mono-aromatic compounds may however be coupled with selective biological processes that can further refine the derivatives for production of commercially valuable aromatic derivatives from lignocellulosic lignin polymers.

References

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. *Journal of Bacteriology* 187(6):2030-2037.

Akhtar M.K., Turner N.J., Jones P.R. (2013). Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities. *Proceedings of the National Academy of Sciences of the United States of America* 110(1):87-92.

Alriksson B., Horvath I.S., Jonsson L.J. (2010). Overexpression of *Saccharomyces cerevisiae* transcription factor and multidrug resistance genes conveys enhanced resistance to lignocellulose-derived fermentation inhibitors. *Process Biochemistry* 45(2):264-271.

Bohme C.C., Arscott L.D., Becker K., Schirmer R.H., Williams C.H. (2000). Kinetic characterization of glutathione reductase from the malarial parasite *Plasmodium falciparum* - Comparison with the human enzyme. *Journal of Biological Chemistry* 275(48):37317-37323.

Chung Y.C., Hurlbert R.E. (1975). Purification and properties of glutathione reductase of *Chromatium vinosum*. *Journal of Bacteriology* 123(1):203-211.

Dhananjaya B.L., Nataraju A., Gowda C.D.R., Sharath B.K., D'Souza C.J.M. (2009). Vanillic acid as a novel specific inhibitor of snake venom 5'-nucleotidase: A pharmacological tool in evaluating the role of the enzyme in snake envenomation. *Biochemistry-Moscow* 74(12):1315-1319.

Egland P.G., Gibson J., Harwood C.S. (2001). Reductive, coenzyme A-mediated pathway for 3-chlorobenzoate degradation in the phototrophic bacterium *Rhodopseudomonas palustris*. *Applied and Environmental Microbiology* 67(3):1396-1399.

Egland P.G., Pelletier D.A., Dispensa M., Gibson J., Harwood C.S. (1997). A cluster of bacterial genes for anaerobic benzene ring biodegradation. *Proceedings of the National Academy of Sciences of the United States of America* 94(12):6484-6489.

Fukuzumi T., Katayama Y. (1977). Bacterial degradation of dimer relating to structure of lignin I. β-hydroxypropiovanillone and coniferylalcohol as initial degradation products from guaiacylglycerol-β-coniferylether by *Pseudomonas putida*. *Mokuzai Gakkaishi* 23(4):214-215.

Gall D.L., Kim H., Lu F., Donohue T.J., Noguera D.R., Ralph J. (2014a). Stereochemical features of glutathione-dependent enzymes in the *Sphingobium* sp. strain SYK-6 β-aryl etherase pathway. *J Biol Chem* 289(12):8656-8667.

Gall D.L., Ralph J., Donohue T.J., Noguera D.R. (2014b). A group of sequence-related sphingomonad enzymes catalyzes cleavage of β-aryl ether linkages in lignin β-guaiacyl and β-syringyl ether dimers. *Environmental Science & Technology* 48(20):12454-12463.

Gibson J., Dispensa M., Harwood C.S. (1997). 4-Hydroxybenzoyl coenzyme A reductase (dehydroxylating) is required for anaerobic degradation of 4-hydroxybenzoate by *Rhodopseudomonas palustris* and shares features with molybdenum-containing hydroxylases. *Journal of Bacteriology* 179(3):634-642.

Haddock J.D., Ferry J.G. (1993). Initial steps in the anaerobic degradation of 3,4,5-trihydroxybenzoate by *Eubacterium oxidoreducens* - characterization of mutants and role of 1,2,3,5-tetrahydroxybenzene. *Journal of Bacteriology* 175(3):669-673.

Harrison F.H., Harwood C.S. (2005). The *pimFABCDE* operon from *Rhodopseudomonas palustris* mediates dicarboxylic acid degradation and participates in anaerobic benzoate degradation. *Microbiology-Sgm* 151727-736.

Hirakawa H., Schaefer A.L., Greenberg E.P., Harwood C.S. (2012). Anaerobic *p*-coumarate degradation by *Rhodopseudomonas palustris* and identification of CouR, a MarR repressor protein that binds *p*-coumaroyl coenzyme A. *Journal of Bacteriology* 194(8):1960-1967.

Itoh A., Isoda K., Kondoh M., Kawase M., Kobayashi M., Tamesada M., Yagi K. (2009). Hepatoprotective effect of syringic acid and vanillic acid on concanavalin A-induced liver injury. *Biological & Pharmaceutical Bulletin* 32(7):1215-1219.

Katayama Y., Fukuzumi T. (1978). Bacterial degradation of dimers structurally related to lignin II. Initial intermediate products from deydrodiconiferyl alcohol by *Pseudomonas putida*. *Mokuzai Gakkaishi* 24(9):643-649.

Katayama Y., Fukuzumi T. (1979a). Bacterial degradation of dimers structurally related to lignin. III. Metabolism of *α*-veratryl-*β*-guaiacylpropionic acid and D,L-pinoresinol by *Pseudomonas putida*. *Mokuzai Gakkaishi* 25(1):67-76.

Katayama Y., Fukuzumi T. (1979b). Bacterial degradation of dimers structurally related to lignin. IV. Metabolism of guaiacylglycerol-*β*-coniferyl ether by *Pseudomonas putida*. *Mokuzai Gakkaishi* 25(5):367-373.

Leonowicz A., Cho N.S., Luterek J., Wilkolazka A., Wojtas-Wasilewska M., Matuszewska A., Hofrichter M., Wesenberg D., Rogalski J. (2001). Fungal laccase: properties and activity on lignin. *Journal of Basic Microbiology* 41(3-4):185-227.

Li M., Semchonok D.A., Boekema E.J., Bruce B.D. (2014). Characterization and evolution of tetrameric photosystem I from the thermophilic cyanobacterium *Chroococcidiopsis* sp TS-821. *Plant Cell* 26(3):1230-1245.

Liao H., McKenzie T., Hageman R. (1986). Isolation of a thermostable enzyme variant by cloning and selection in a thermophile. *Proceedings of the National Academy of Sciences of the United States of America* 83(3):576-580.

Masai E., Ichimura A., Sato Y., Miyauchi K., Katayama Y., Fukuda M. (2003). Roles of the enantioselective glutathione S-transferases in cleavage of β-aryl ether. *Journal of Bacteriology* 185(6):1768-1775.

Masai E., Katayama Y., Fukuda M. (2007). Genetic and biochemical investigations on bacterial catabolic pathways for lignin-derived aromatic compounds. *Bioscience Biotechnology and Biochemistry* 71(1):1-15.

Masai E., Katayama Y., Nishikawa S., Yamasaki M., Morohoshi N., Haraguchi T. (1989). Detection and localization of a new enzyme catalyzing the β-aryl ether cleavage in the soil bacterium (*Pseudomonas-paucimobilis* SYK-6). *Febs Letters* 249(2):348-352.

Merkel S.M., Eberhard A.E., Gibson J., Harwood C.S. (1989). Involvement of coenzyme A thioesters in anaerobic metabolism of 4-hydroxybenzoate by *Rhodopseudomonas palustris*. *Journal of Bacteriology* 171(1):1-7.

Ninh P.H., Honda K., Sakai T., Okano K., Ohtake H. (2015). Assembly and multiple gene expression of thermophilic enzymes in *Escherichia coli* for *in vitro* metabolic engineering. *Biotechnology and Bioengineering* 112(1):189-196.

Otsuka Y., Sonoki T., Ikeda S., Kajita S., Nakamura M., Katayama Y. (2003). Detection and characterization of a novel extracellular fungal enzyme that catalyzes the specific and hydrolytic cleavage of lignin guaiacylglycerol-β-aryl ether linkages. *European Journal of Biochemistry* 270(11):2353-2362.

Palmqvist E., Hahn-Hägerdal B. (2000a). Fermentation of lignocellulosic hydrolysates. I: inhibition and detoxification. *Bioresource Technology* 74(1):17-24.

Palmqvist E., Hahn-Hägerdal B. (2000b). Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. *Bioresource Technology* 74(1):25-33.

Pan C., Oda Y., Lankford P.K., Zhang B., Samatova N.F., Pelletier D.A., Harwood C.S., Hettich R.L. (2008). Characterization of anaerobic catabolism of *p*-coumarate in *Rhodopseudomonas palustris* by integrating transcriptomics and quantitative proteomics. *Molecular & Cellular Proteomics* 7(5):938-948.

Peng X., Misawa N., Harayama S. (2003). Isolation and characterization of thermophilic Bacilli degrading cinnamic, 4-coumaric, and ferulic acids. *Applied and Environmental Microbiology* 69(3):1417-1427.

Perrotta J.A., Harwood C.S. (1994). Anaerobic metabolism of cyclohex-1-ene-1-carboxylate, a proposed intermediate of benzoate degradation by *Rhodopseudomonas palustris*. *Applied and Environmental Microbiology* 60(6):1775-1782.

Rahimi A., Azarpira A., Kim H., Ralph J., Stahl S.S. (2013). Chemoselective Metal-Free Aerobic Alcohol Oxidation in Lignin. *Journal of the American Chemical Society* 135(17):6415-6418.

Rahimi A., Ulbrich A., Coon J.J., Stahl S.S. (2014). Formic-acid-induced depolymerization of oxidized lignin to aromatics. *Nature* 515(7526):249-252.

Reichenbecher W., Brune A., Schink B. (1994). Transhydroxylase of *Pelobacter acidigallici* - a molybdoenzyme catalyzing the conversion of pyrogallol to phloroglucinol. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* 1204(2):217-224.

Reichenbecher W., Schink B. (1999). Towards the reaction mechanism of pyrogallol-phloroglucinol transhydroxylase of *Pelobacter acidigallici*. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* 1430(2):245-253.

Reiter J., Pick A., Wiemann L.O., Schieder D., Sieber V. (2014). A novel natural NADH and NADPH dependent glutathione reductase as tool in biotechnological applications. *JSM Biotechnol Bioeng* 2(1):1028-1035.

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. *Applied and Environmental Microbiology* 75(16):5195-5201.

Schierbeek A.J., Swarte M.B.A., Dijkstra B.W., Vriend G., Read R.J., Hol W.G.J., Drenth J., Betzel C. (1989). X-ray structure of lipoamide dehydrogenase from *Azotobacter vinelandii* determined by a combination of molecular and isomorphous replacement techniques. *Journal of Molecular Biology* 206(2):365-379.

Scrutton N.S., Berry A., Perham R.N. (1990). Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature* 343(6253):38-43.

Shang-Ming H., Hong-Chih C., Chi-Hao W., Gow-Chin Y. (2008). Cytoprotective effects of phenolic acids on methylglyoxal-induced apoptosis in Neuro-2A cells. *Molecular Nutrition and Food Research* 52(8):940-949.

Sinha A.K., Sharma U.K., Sharma N. (2008). A comprehensive review on vanilla flavor: Extraction, isolation and quantification of vanillin and others constituents. *International Journal of Food Sciences and Nutrition* 59(4):299-326.

Sonoki T., Obi T., Kubota S., Higashi M., Masai E., Katayama Y. (2000). Coexistence of two different O demethylation systems in lignin metabolism by *Sphingomonas paucimobilis* SYK-6: Cloning and sequencing of the lignin biphenyl-specific O-demethylase (LigX) gene. *Applied and Environmental Microbiology* 66(5):2125-2132.

Vetrano A.M., Heck D.E., Mariano T.M., Mishin V., Laskin D.L., Laskin J.D. (2005). Characterization of the oxidase activity in mammalian catalase. *Journal of Biological Chemistry* 280(42):35372-35381.

Yang W.R., Sen A. (2010). One-step catalytic transformation of carbohydrates and cellulosic biomass to 2,5-dimethyltetrahydrofuran for liquid fuels. *Chemsuschem* 3(5):597-603.