Title
Exploring enzymatic diversity in the environment: Discovery and characterization of enzymes from lignolytic soil bacteria

Author
Brown, Margaret Elizabeth

Publication Date
2013

Peer reviewed|Thesis/dissertation

Exploring enzymatic diversity in the environment: Discovery and characterization of enzymes from lignolytic soil bacteria

by

Margaret Elizabeth Brown

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy in
Chemistry
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Michelle C. Y. Chang, Chair
Professor Christopher J. Chang
Professor Danielle Tullman Ercek

Fall 2013

Exploring enzymatic diversity in the environment: Discovery and characterization of enzymes from lignolytic soil bacteria

by Margaret Elizabeth Brown

Abstract

Exploring enzymatic diversity in the environment: Discovery and characterization of enzymes from lignolytic soil bacteria

by

Margaret Elizabeth Brown

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Michelle C. Y. Chang, Chair

In order to explore new chemical strategies for lignin degradation, we have initiated studies aimed at discovery and characterization of oxidative and accessory enzymes in lignin-reactive soil bacteria that exhibit particularly rich activity towards the depolymerization and utilization of biomass-derived carbon sources. Our interest lies in studying the biochemical logic underlying the transformation of complex substrates by living organisms with the overall goal of elucidating new molecular strategies for lignin degradation. Towards this goal, we began by creating a pipeline for rapid discovery and functional identification of new enzymes from unsequenced bacteria under lignin-reactive growth conditions, an approach that can be applied to any culturable microbe of interest. To do this, we assembled a *de novo* genome for *Amycolatopsis* sp. 75iv2 and used this genome sequence to identify genes potentially involved in the degradation system of lignocellulose as well as their downstream degradation products. With a genome sequence now available for proteomics analysis, we analyzed the secretome of *A.* sp. 75iv2 grown in the presence of lignocellulose by extracellular peroxidase activity of the culture; separation of the secretome by SDS-PAGE and heme-staining revealed the presence of heme-containing proteins which we identified as a catalase-peroxidase and a catalase by LC-MS/MS. Biochemical characterization of the catalase-peroxidase revealed an inability to oxidize non-phenolic ring motifs but showed that phenolic moieties, which encompass ~20% sites in lignin, could be oxidized.

Further analysis of the genome sequence allowed for identification of an interesting and relatively new family of enzymes, the dye decolorizing peroxidases (DyPs). One phylogenetically unique member of this family, DyP2, was heterologously expressed for biochemical and structural characterization. Comparison of the enzymatic activity of this protein indicated that it displayed not only high peroxidase activity against a suite of low- and high-potential DyP substrates, similar to that of fungal DyPs and LiPs, but also an unusually active manganese peroxidase activity akin to that of versatile peroxidases. Furthermore, studies showed DyP2 to have a Mn-dependent oxidase activity that expands its substrate scope to include the oxidation of substrates with non-phenolic aromatic rings. We also solved a crystal structure of DyP2 at 2.25A resolution which revealed the presence of a Mn binding pocket 15 Å from the heme active site.

Finally to begin studying the physiological response of *A.* sp. 75iv2 to lignin, we developed a set of minimal media conditions that demonstrated *A.* sp. 75iv2 to be indeed competent to utilize lignin as a sole carbon source. Interestingly, the presence of lignin induced a differential growth response compared to growth on sugar controls. Analysis of the *A.* sp. 75iv2 secretome further showed the existence of an induction of extracellular peroxidase activity and changes in the secreted heme protein profile. RNA sequencing and proteomics studies were then utilized to profile the global response of *A.* sp. 75iv2 to lignin and identify candidate proteins and pathways involved in its metabolism. These studies revealed that the dominant response involved a strong activation of Fe assimilation pathways as well as significant upregulation in the expression of universal stress response regulators. Both up- and down-regulation of aromatic degradation pathways were also observed. In order to continue to explore this response, we developed a synthesis for ^{13}C-labeled synthetic lignin as a tool for structural studies of lignin by NMR as well as a peroxidase fractionation method as an approach to begin identifying upregulated peroxidases in the *A.* sp. 75iv2 secretome.

Table of Contents

List of Figures, Schemes, and Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Abbreviations

LiP	lignin peroxidase
MnP	manganese peroxidase
VP	versatile peroxidase
DyP	dye-decolorizing peroxidase
CYP	cytochrome P450
Amyco1	catalase-peroxidase from *A.* sp. 75iv2
Amyco2	catalase from *A.* sp. 75iv2
dNTP	deoxynucleotide triphosphate
H_2O_2	hydrogen peroxide
DCP	2,4-dichlorophenol
DL	dioxane-extracted lignin
EL	ethanol-extracted lignin
OD	optical density
VA	veratryl alcohol
L-DOPA	L-3,4-dihydroxyphenylalanine
AAP	4-aminoantipyrine
MMA	4-methoxymandelic acid
TCEP	tris(2-carboxyethyl)phosphine
IPTG	isopropyl β-D-1-thiogalactopyranoside
TCA	trichloroacetic acid
DTT	diothiothreitol
PMSF	phenylmethanesulfonyl fluoride
TB	Terrific broth
EDTA	ethylenediaminetetraacetic acid
ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)
APPL	acid-precipitable polymeric lignin
ORFs	open reading frames
COGs	clusters of orthologous groups
SMM	supplemented minimal media
MM	minimal media
DHP	dehydropolymer, synthetic lignin
SDS-PAGE	sodium dodecyl sulfate- polyacrylamide gel electrophoresis
TEMED	N,N,N′,N′-Tetramethylethane-1,2-diamine

NMR	nuclear magnetic resonance
GPC	gel permeation chromatography
HSQC	heteronuclear single quantum correlation
ICP-OES	inductively coupled plasma-optical emission spectrometry
GC-MS	gas chromatography-mass spectrometry
FPLC	fast protein liquid chromatraphy
RP-HPLC	reversed phase-high performance liquid chromatography
LC-MS/MS	liquid chromatography-tandem mass spectrometry

Acknowledgments

First and foremost, I would like to thank Professor Michelle Chang. As she made a gamble by taking me into the lab, having absolutely no academic background in biochemistry and related fields, I appreciate the opportunity she gave me, even as it was extremely risky for her. With her zeal and precision for this field of science, she single-handedly, accurately, and rapidly taught me this world, a world for which I have grown a great passion. She allowed me to work on a project for which, every day, I could wake up with a purpose, questions to answer, and a desire to answer them. For even though it's almost silly, I've attempted to live by Horace Mann's 1859 commencement speech in which he stated "be ashamed to die until you have won some victory for humanity." So much can be accomplished through hard work, dedication, interesting insights, and passion, and Professor Chang embodies these characteristics. I aspire to live my life according to these. As members of my thesis committee, I thank Professor Christopher Chang for his invaluable suggestions, support, and help and I thank Professor Danielle Ercek for her hallway pep talks, scientific and life advice, and her general supportive demeanor.

A number of collaborations with invaluable scientific experts have allowed this research to flow smoothly. For their help with LC-MS/MS, I thank both Dr. Anthony Iavarone and Dr. Lori Kohlstaedt. For MUDPIT analysis of complex protein mixtures, I thank Dr. Mela Mulvihill and Professor Daniel Nomura. For their help with my extremely complicated lignin structure determination using NMR as well as general advice, I thank Dr. Hagit Sorek and Professor Dave Wemmer. For two reasons, I thank Dr. Tiago Barros: for his help with my DyP2 diffraction datasets using his decade of crystallographic experience; also, for his amazing friendship that I gained through this project. Similarly, to my labmate, Mark Walker, who helped me with my *de novo* genome assembly as well as RNAseq analysis. His assistance to my project is inestimable; almost more importantly, he has become a great friend, one whom I wish our busy scientific lives had allotted more time for further friendship development. Even with this, I trust him in a way I trust few and will miss him dearly as he leaves the Bay Area to continue his scientific endeavors. To the other two members leaving Professor Chang's lab as her first crew, Brooks Bond-Watts and Amy Weeks, I cannot verbalize my appreciation to them for their support, general friendship, and overall brilliance. The lab was made an amazing place by them. In fact, the whole M. Chang lab was a wonderful place to perform research, as it is a family built of beautiful and intensely amazing personalities. I will miss them as I leave.

Finally, as the world and life are so simultaneously ugly and beautiful, wonderful and painful, it is with the support and love of a number of people that I continue to function. For some, family is given, and for some, a family is made. Mine is somewhere in between. To my sister, Alison Brown (Ali B), she is my inspiration as she is so utterly beautiful inside and out that she has reached superhero status in my mind. With my new-found time, I will make her the cape she deserves. A number of friends have stayed long-distance with me over the last decade: Jared Collins and Megan Powers who have been with me through long nights of intense life experiences. They keep me honest; they force me to analyze my world and continually question my role; they make me a better person. For these things, I cannot impart their importance and beauty even as they live states away. The years of living in the Bay Area have given me the amazing friendships of a number of people. Jason and Carolyn Hudak understand the minutae that makes me appear crazy to the rest of the world; their scientific drive as well as their passion and duty toward life inspire me to be better and do more. Katherine Miller and many other amazing people have spent hundred of hours on a bicycle

with me, distance cyling out the stress through the beautiful bay area and its local mountains. Katherine is an invaluable friend whom I love dearly. Finally, to my partner in crime, Steve Hanna, a fellow survivor of UC Berkeley graduate school, I have to thank him for everything in the last 1.5 years. Steve makes my life brighter with his passion to do everything, and I cannot express my gratitude for his support and love. With him, I've had great experiences and adventures, and I know it only gets better from here.

Chapter 1: *Lignin—structure, function, and depolymerization*

1.1. Introduction

Living organisms have used the power of evolution to innovate new and remarkable chemical transformations that not only allow them to compete successfully with rivals for environmental resources but also address synthetic challenges of interest to human society. For example, the effective and efficient activation of small molecules such as carbon dioxide, oxygen, or nitrogen by organisms for downstream multistep reactions takes advantage of the high availability of gaseous substrates to fix C, N, and O for cell growth. As such, the enzymes responsible for these transformations have served as the longstanding inspiration for the development of homogeneous and heterogeneous chemical catalysts [*1-7*] and other enzyme mimicks [*8, 9*]. Similarly, organisms have evolved entire pathways for degrading environmental toxins such as chlorinated aliphatics and substituted aromatics [*10, 11*] or synthesizing complex molecules [*12-15*] that would be difficut to achieve through chemical approaches. These naturally-occurring multi-step reaction sequences provide the inspiration for synthetic biology approaches to introduce multi-gene synthetic pathways into tractable hosts, potentially enabling the large-scale production or degradation of a target compound in a single step [*16-21*].

One critical advance that biological systems could provide is new routes to the utilization of alternative carbon sources. Despite escalating interest in renewable resources, the large-scale production of fine and commodity chemicals that forms the basis for modern society in the

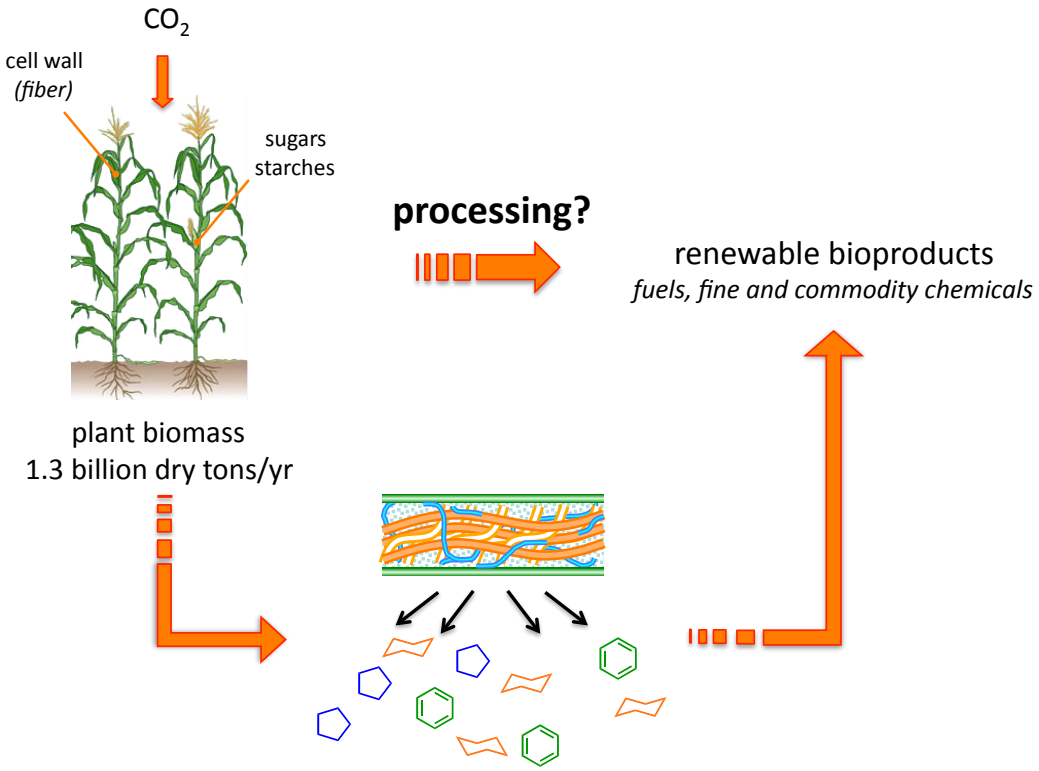


Figure 1.1. *Processing of complete lignocellulosic fraction of plant materials expands the scope of plant materials available for industrial usage. As industry is currently only using the expensive sugar and starch components of the plants, most of the carbon and, thus, energy is left behind as waste in the lignocellulosic fraction, or the plant cell wall. If processes can be developed to degrade the crop waste into useful sugar and aromatic monomers, a new stock of starting material will be available for the production of fine and commodity chemicals.*

United States relies almost completely on petroleum-derived feedstocks as a carbon source [*22-26*]. One of the few alternative natural carbon-based resources that is produced in high enough volume to displace a significant level of fossil fuels is plant biomass, which is both renewable and has the potential to provide carbon on the order of 10–50 billion tons per year [*24, 27*]. Although sugars, starches, and oils from plants have entered the industrial processing pipeline, the majority of plant material is made up of structural biopolymers that are resistant to breakdown and downstream utilization for chemical and biological processes [*23, 28*]. Thus, the majority of the carbon in plant material remains trapped in these structural biopolymers, and only easily accessible sugars and starches are currently utilized (*Figure 1.1*) [*28*]. This recalcitrant fraction, the plant cell wall, is comprised of cellulose, hemicellulose, and lignin (*Figure 1.2A*). Of these biopolymers, lignin is arguably the most chemically difficult target to approach for degradation because of its aromatic monomer units as well as diverse set of C–C and C–O crosslinks [*29*], and this high degree of complexity requires multiple families of enzymes to achieve its breakdown [*30, 31*].

The importance of lignin to biomass depolymerization has led to the development of a variety of engineering, chemical, and biochemical approaches to its degradation [*32, 33*]. Both chemical and thermochemical approaches are capable of breaking lignin down a wide range of products but are also quite energy intensive. In contrast, biological approaches provide an alternative strategy that could both potentially bypass this energy-intensive process and also produce a more specific array of products. Despite its recalcitrance, select microbes have evolved the ability to react with and depolymerize lignin. Fungal systems have been fairly well-characterized both *in vivo* and *in vitro* to be highly active for lignin depolymerization, using families of heme- and multicopper-dependent oxidative enzymes to carry out radical-mediated degradation [*30, 31*]. These fungal metalloenzymes have been used to develop the current model for fungal lignin breakdown: an oxidative combustion of lignin through various radical-mediated

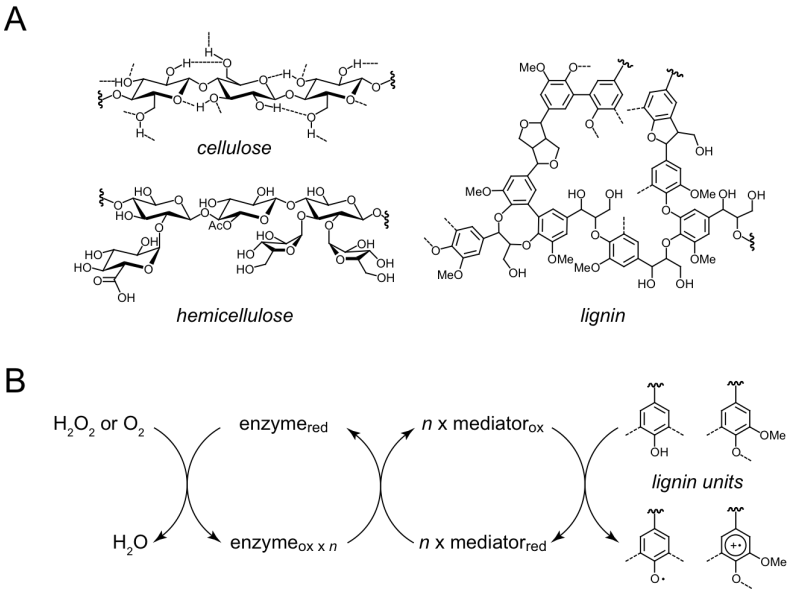


Figure 1.2. *(A) Representative biopolymers found in lignocellulose. (B) Oxidative reactions initiating the radical-mediated breakdown of lignin. Small molecule mediators, including veratryl alcohol, Mn(II), and oxalate, are oxidized directly by fungal peroxidases and laccases and transfer an oxidizing equivalent to the lignin substrate to generate either a phenoxyl or phenyl radical equivalent.*

paths that ultimately generates phenoxyl and phenyl radicals on the substrate leading to a cascade of bond scission reactions within the lignin substrate (*Figure 1.2B*) [*30, 31, 34, 35*]. Studies on these canonical lignin-degrading fungal proteins, primarily the lignin peroxidases (LiPs) and manganese peroxidases (MnPs), have revealed these proteins to be extremely effective at lignin depolymerization, producing a suite of monomeric aromatic compounds and even complete degradation to carbon dioxide and water [*30, 31*]. In comparison, bacterial systems are less well understood with regard to lignin reactivity and thus represent an interesting area for exploring the diversity of enzymatic strategies for biomass deconstruction [*36-38*]. While this overall oxidative strategy for lignin reactivity is most likely conserved within bacteria, it is also known that bacteria exhibit quite different reactivity towards lignin compared to fungi based on their altered product distribution. Many soil bacteria have been reported to modify lignin to produce a different product, acid-precipitable polymeric lignin (APPL), a water-soluble lignin product which precipitates in acidic conditions [*39*]. These observations suggest that the enzymatic approaches to lignin degradation are organism-dependent and that exploring more biodiversity could expand our understanding of the different roles played by bacteria and fungi in environmental lignin degradation as well as lead to the discovery of new enzymatic mechanisms for lignin degradation. These new approaches for lignin breakdown could enable our ability to utilize lignocellulosic biomass as a carbon and energy source as well as provide new routes for aromatic feedstocks.

1.2. Lignin structure and function

The plant cell wall is comprised of three primary biopolymers – cellulose, hemicellulose, and lignin. Cellulose, the β-1,4-linked glucose polymer, provides the basic foundation for the primary and secondary cell walls, making up about 35-50% of the cell wall mass [*40*]. Hemicellulose (20-35%), which includes a diverse set of sugar polymers such as xyloglucan, then links the cellulose microfibrils to form a network. Lignin (5-25%) is contained within this framework and fills the space to increase cell wall integrity, hydrophobicity, rigidity, and resistance to attack by pathogens [*41*]. The structure of lignin has proven an extremely beneficial adaptation for vascular plants, as it has been found to be highly resistant to both enzymatic and chemical degradation. Despite the challenges, the development of new methods for lignin degradation has been of high interest for a variety of biotechnological applications.

Monolignols. Lignin is formed within the plant cell wall by the radical polymerization of three primary monomers, *p*-coumaryl alcohol, and its methoxylated derivatives, coniferyl and sinapyl alcohol (*Figure 1.3*). These three monolignols are synthesized by the plant from phenylalanine, beginning with deamination of the amino acid and followed by a complex series of modifications by ten enzymes to produce the monolignols needed for lignin biosynthesis [*29, 41*]. The concentration of phenylpropanoid monomers found within each plant is different and genetically determined. Indeed, many isoforms of the phenylpropanoid pathway enzymes are found within plant genomes, each with different biochemical properties and distribution throughout the plant, further suggesting that lignin synthesis is complex and highly regulated [*41*].

Lignin polymerization. Lignin polymerization is initiated by enzyme-dependent formation of a phenoxy radical on the monolignol and progresses by radical coupling between the monomers. At this time, gene knockout studies support the involvement of laccases in this

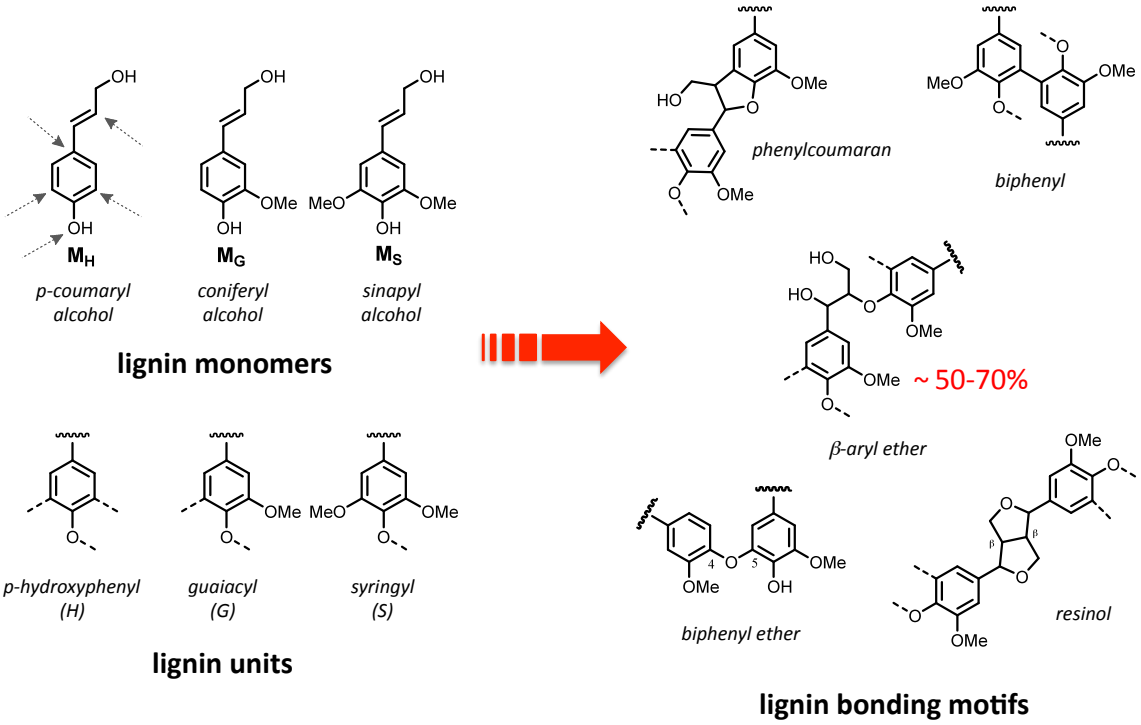


Figure 1.3. *Lignin is polymerized from 3 primary monomers: p-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol. As the polymerization process is radical-dependent, susceptible locations to radical attack are denoted by gray arrows on the p-coumaryl alcohol molecule. Coniferyl and sinapyl alcohol are methoxylated at the 3- and 5- positions, respectively, so the route for crosslinking at these positions is unavailable. Thus, the different crosslinking routes are available producing different lignin units, and therefore the bonding motifs available to the different monomers varies, making the overall bonding patterns found in different plant lignins monomer-dependent and genetically determined.*

process [*42-44*], the physiological role of heme peroxidases remains unclear despite their ability to catalyze lignin polymerization *in vitro* [*29*]. While the full diversity of linkages between monomers have yet to be structurally characterized, some of the major motifs include include the biphenyl ether, *β*-aryl ether (*β*-O-4), biphenyl, and phenylcoumaran (*Figure 1.3*) [*29, 41*]. The methoxy groups found in coniferyl and sinapyl alcohol act as protecting groups that inhibit radical coupling at these locations, further altering the bonding patterns in a host-dependent manner based on the genetically determined monomeric concentrations. Since all three monomers share the phenolic and *β*-hydrogen sites for radical coupling, the *β*-O-4 typically represents the most prevalent bonding motif and, thus, one of the most important to deconstruct.

Lignin diversity and engineering. Studies of the lignins found in various plants have shown the lignin content of biomass to be highest in softwood, followed by hardwood, with grasses being the lowest [*32, 45, 46*]. In addition to lignin content, plants also differ in their lignin structure based on the monomer composition found in each species and have been quantified by the H:G:S ratios. The higher degree of polymerization afforded by the unmethoxylated sites in coumaryl and coniferyl alcohols compared to sinapyl alcohol has been suggested to be important for increasing the rigidity and polymerization of the lignin. As H residues are typically found to be of lower abundance, a higher G:S ratio thus yields a more crosslinked product, possibly increasing its hydrophobicity and innate rigidity innate. For example, a high concentration of G units have been found in softwood lignins whereas hardwood lignins contain approximately equal S and G units [*32, 47*]. In comparison, grasses contain a relatively high level of H units (5%) and the difficulty in grass lignin degradation has been

attributed to the highly branched nature of their lignins [*47, 48*]. Similarly, hydrolysis of softwoods has been found to be lower than those of hardwoods [*49*]. However, the overall strength of the materials found in "hardwood" and "softwood" is not solely dependent on the presence and crosslinking of lignin, as hardwoods have been found, in general, to have higher concentrations of holocellulose (cellulose + hemicellulose), and the differences in plant vascular structure leads to overall physiological differences in these materials [*45, 46*].

The engineering of monomer composition and the resulting changes in lignin structure have been an interesting area for controlling animal forage efficiency as well as deconstruction by chemical or enzymatic approaches [*50, 51*]. Several studies have focused on altering lignin structure via the availability of monomers. Experiments to alter the G:S ratio of plants have shown that lignification is relatively plastic, as growth appears normal in these engineered plants [*29*]. For example, overexpression of genes for production of sinapyl alcohol results in a lignin that is more easily pulped, due to the lower crosslinking [*50*]. However, the downregulation of biosynthetic enzymes by genetic mis-sense mutation for production of the primary three monolignols alcohol has resulted in severe growth defects, suggesting that large changes in lignin structure are detrimental to lignin structure and plant health [*52*]. Overall lignin content can also be decreased by downregulation of specific pathway enzymes, most highly from the downregulation of the hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase (HCT) which decreased lignin content by over 50% of the wild-type production and allowed for increased saccharification of the sugar biopolymers [*51*]. In addition to altering natural pathways, the incorporation of non-canonical lignin monomers has also served as a possible approach for controlling lignin structure and processing [*53*].

1.3. Fungal depolymerization of lignin

The most efficient system for lignin degradation has arisen in the family of white-rot fungi, which includes the well-studied lignin-degrading basidiomycete, *Phanerochaete chrysosporium* [*30*], as well as other white-rot fungi such as *Pleurotus ostreatus* [*54*], *Ceriporiopsis subvermispora [55, 56]*, and *Trametes versicolor* [*57, 58*]. Given the complexity of the substrate, the full complexity of the enzyme system involved in lignin depolymerization has yet to be identified but involves a number of oxidative enzymes and small molecule mediators. At this time, class II peroxidases and O_2-utilizing laccases comprise the most well-studied and abundant components of this system but other oxidative enzymes and accessory enzymes may also be involved. Depending on the organism, lignin can be degraded simultaneously (*P. chrysosporium* and *T. versicolor*), preferentially (*P. ostreatus*), or exclusively (such as *C. subvermispora*) over sugar biopolymers [*55*]. Genome analysis of lignolytic fungi further suggests that the enzyme systems differ from host to host. For example, the white-rot fungal genomes contain many copies of class II peroxidases (5-26) whereas brown-rot fungi generally lack class II peroxidases and utilize laccases and Fenton chemistry for breakdown [*55, 59-61*].

Class II peroxidases (LiPs and MnPs). The primary model developed for microbial lignin degradation has been based on the extracellular oxidative system used by *P. chrysosporium* and has thus focused on a class II peroxidases. *P. chrysosporium* primarily relies on two families of highly oxidative secreted heme peroxidases, the LiPs and MnPs. These enzymes, though relatively unusual in their oxidative capacity, act on substrates through a classic peroxidase mechanism (*Scheme 1.1*). The presence of extracellular hydrogen peroxide oxidizes the ground

$$Fe^{III} + H_2O_2 \rightarrow Fe^{IV}=O^{+\bullet} + H_2O$$
$$Fe^{IV}=O^{+\bullet} + \text{substrate} \rightarrow Fe^{IV}=O + \text{substrate}^{+\bullet}$$
$$Fe^{IV}=O + \text{substrate} \rightarrow Fe^{III} + \text{substrate}^{+\bullet} + H_2O$$

Scheme 1.1. *Mechanism of the reaction of a peroxidase with H_2O_2.*

state Fe(III) to an Fe(IV)=$O^{+\bullet}$ species, also known as compound I [*62*]. Depending on the peroxidase, this radical cation is either delocalized across the porphyrin as a π radical cation or on a nearby proximal amino acid side-chain, such as tryptophan. This highly active species is competent to oxidize the substrate by one electron, reducing the heme to Fe(IV)-oxo species (compound II). The compound II species can then oxidize a second equivalent of the substrate to return to the Fe(III) ground state [*63, 64*].

The genome of *P. chrysosporium* encodes ten copies of LiP genes, which appear to be a key component to its highly oxidative degradation system [*65-68*]. In addition to the peroxidases, secreted oxidases that produce exogenous hydrogen peroxide, such as aryl-alcohol oxidases and glyoxal oxidases, are also highly expressed, presumably to provide the oxidant for the LiPs [*69*]. Along with these enzymes, the fungus also secretes a small non-phenolic aromatic metabolite, 3,4-dimethoxybenzyl alcohol (also known as veratryl alcohol, VA), that is as a small molecule mediator for lignin degradation [*70*]. The current model for LiP function involves the reaction of LiP with hydrogen peroxide which subsequently oxidizes VA to generate the veratyl alcohol radical cation ($VA^{+\bullet}$, *Figure 1AB*) [*34*]. This extremely oxidizing species (E ~ 1.36 V) [*71*] is then proposed to act as a redox mediator, diffusing from the surface of the LiP to transfer the

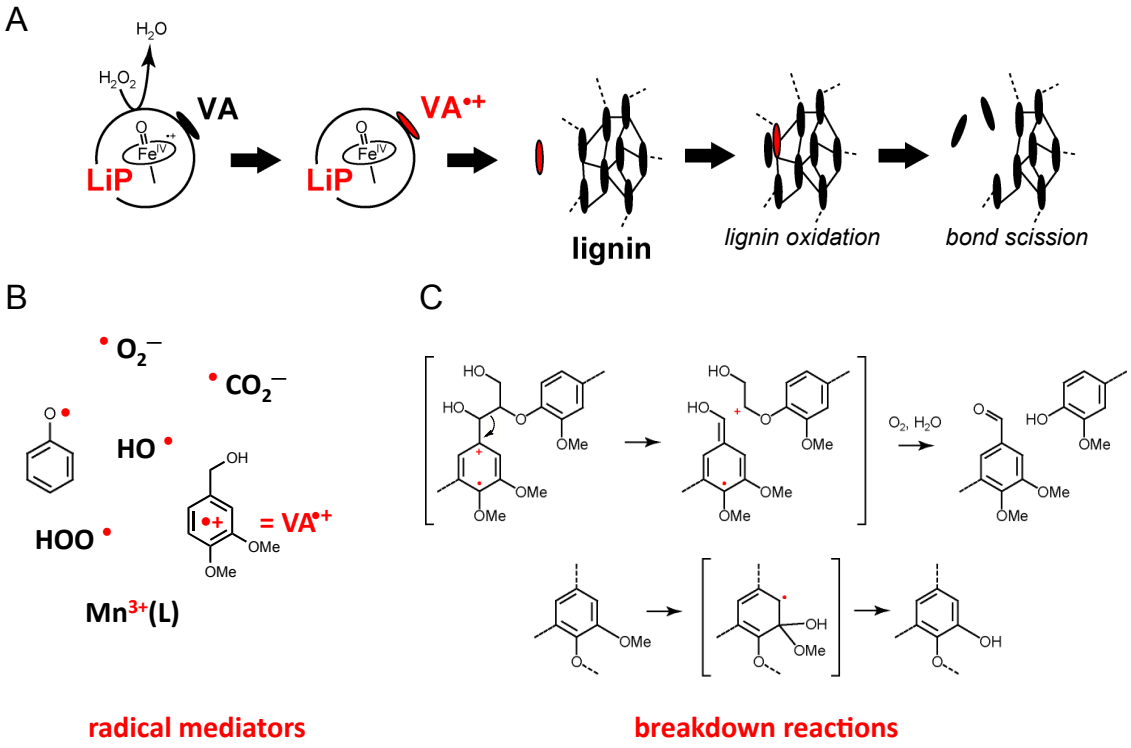


Figure 1.4. Model for LiP-mediated lignin breakdown. *(A) Overarching model for lignin breakdown relies on oxidation of the enzyme by H_2O_2, which subsequently oxidizes veratryl alcohol (VA), yielding a highly oxidative radical mediator that can diffuse to oxidize lignin. (B) Examples of lignin breakdown reactions initiated by electron abstraction by the VA radical cation include α-β cleavage and demethylation. (C) Examples of radical mediators thought to be involved in oxidative lignin breakdown.*

oxidizing equivalent to lignin. From there, it is thought that a cascade of scission reactions can occur to lead to various C–C and C–O bond-breaking events (*Figure 1.4C*) [*31*]. The oxidation of lignin and VA requires an enzyme with an unusually high oxidation potential to allow for electron abstraction from unactivated aromatic rings and α-hydrogen abstraction. Few peroxidases outside of LiP are capable of performing this reaction and it is suggested that the LiPs are more electron deficient than classic peroxidases, allowing them access to these potentials [*72*].

P. chrysosporium has also been found to secrete another family (five isozymes) of heme peroxidases, the MnPs, which are proposed to oxidize a chelated Mn^{2+} to Mn^{3+} using a similar mechanism (*Scheme 1.1*) [*73-76*]. Again, the chelated Mn ion is thought to diffuse to lignin and oxidize phenolic moieties, which remain in ~20% abundance in lignin [*77, 78*]. The MnPs have been suggested to play a very specific role in lignin degradation, as it has been noted that LiPs are both less efficient at phenolic oxidation due to enzyme inactivation and also do not oxidize Mn [*56, 79*]. Interestingly, some organisms, such as the white-rot fungus *C. subvermispora*, appear to primarily rely on MnP activity for lignin degradation; the genome of *C. subvermispora* encodes a high number of MnPs (13) as compared LiPs (1), and transcriptional profiling revealed that the presence of lignin induces a MnP response but not a LiP response [*55*]. In addition to MnPs, a new member of the class II peroxidases, the versatile peroxidases (VPs), was discovered in 1999 from the white-rot fungi, *Pleurotus eryngii* [*80*]. Interestingly, these VPs appear to be hybrid enzymes, capable of not only VA but also Mn^{2+} oxidation. In addition, VPs are also capable of oxidizing aromatic phenols without small-molecule mediators [*81*].

Laccases. Laccases are blue copper oxidases that oxidize substrates with concomitant reduction of O_2 to water. These enzymes are frequently found to be secreted from a number of lignolytic fungi and are considered to be an important player in lignin degradation for a number of species [*57*]. Laccases are generally found to only have phenolic oxidation capacity, but their substrate specificity can be changed in the presence of small-molecule redox mediators. For example, a laccase was found to be capable of oxidizing a non-phenolic *β*-O-4 lignin model dimer (*Figure 1.5*) in the presence of 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) [*82*]. However, native mediators for laccases have yet to be determined, and the physiological role of these observed activities remains uncharacterized.

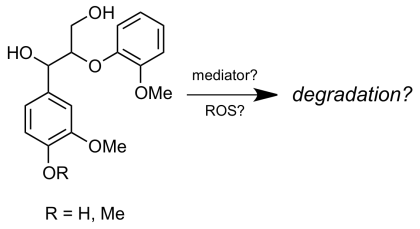


Figure 1.5. *Lignin model dimer used in assays to represent the prevalent β-O-4 lignin linkage and monitor cleavage of this bonding motif.*

Other oxidative enzymes. While peroxidases and laccases make up the most well-characterized members of the extracellular lignolytic system, other classes of oxidative enzymes could also possibly be involved. One class of enzymes that has been suggested to partake in lignin degradation are cytochrome P450 (CYPs) monooxygenases, which have the ability to use molecular oxygen to oxidize compounds for the purpose of modification, detoxification, and breakdown [*83*]. These enzymes play roles in secondary metabolite modification, hydroxylating "unactivated" alkanes and other hydrophobic compounds, xenobiotic metabolism, and potentially even in lignin breakdown [*84*]. The genome of *P. chrysosporium* encodes ~150 CYPs [*85*]; based on microarray studies, four CYPs are statistically significantly upregulated (2- to 20-fold) under lignolytic conditions and have been proposed to play a role in down-stream

processing of peroxidase-depolymerized lignin products [*86*]. Follow-up proteomics analyses revealed two membrane-bound cytochrome P450s to also be differentially upregulated under lignolytic conditions [*87*]. In addition to CYPs, some dye-decolorizing peroxidases (DyPs) have been shown to be both secreted and chemically competent to oxidize VA and break down a non-phenolic β-O-4 lignin model dimer (*Figure 1.5*) [*88-90*]. However, their physiological role in lignin degradation is unclear as they are mostly implicated in oxidative stress responses or other functions [*91, 92*].

1.4. Bacterial depolymerization of lignin

Introduction. Extensive studies on fungal lignolytic systems have revealed an impressive enzymatic system for degrading an extremely difficult and recalcitrant substrate. While fungi clearly represent the most oxidatively robust system for lignin breakdown, the ability of bacteria to interact with lignin has been reported, and it is interesting to consider their role in lignin breakdown catalyzed by microbial communities. Whereas bacteria are known to be able to efficiently degrade and utilize low molecular weight aromatic compounds (<1000 Da) [*93, 94*], including lignin degradation products from other microbes, their ability to depolymerize higher molecular weight lignin has been questioned [*95, 96*]. Nonetheless, studies on specific actinomycetes suggest that some bacteria indeed have the ability to add to the community of high molecular weight lignin degradation [*97*]. Laboratory studies have also shown that soil-dwelling bacteria can react with lignin to produce APPL, suggesting that bacterial lignin reactivity is possible and that different enzymes may be involved than those from fungi [*98*].

A role for bacteria in lignin degradation was first proposed in 1962 based on studies by Sørensen in which bacterial colonies cultured from soil produced clear spots on lignin-containing silica gel plates, and low lignin levels in those spots were verified by the addition of phloroglucinol, a lignin staining molecule [*99*]. It was then shown that *Rhodococcus opacus* (previously *Nocardia* sp.) grown under low nutrient conditions was capable of mineralizing ^{14}C-labeled maize lignin and ^{14}C-labeled synthetic lignin (dehydropolymer, DHP) to $^{14}CO_2$. Interestingly, its specificity was shown to vary towards different bonding motifs found in lignin with the high reactivity exhibited towards methoxyl group reactivity, decreasing this motif by 36% [*100*]. The Crawford laboratory was also able to characterize thirty actinomycetes and show that three strains were competent to release $^{14}CO_2$ from ^{14}C-labeled milled-wood lignocellulose labeled in either the lignin component or the cellulosic component [*101*]. While most of the lignocellulosic degradation was from carbohydrate metabolism, the bacteria were still capable of degrading up to 18% of the extracted lignin over 42 days. Additional studies also demonstrated that other soil-dwelling species, such as *Xanthamonas* sp., *Arthrobacter* sp., and *Spartina alterniflora*, could utilize lignin as the sole carbon source of growth [*102, 103*]. However, the requirement for added carbon sources, such as glucose or yeast extract [*104*], in many cases calls into question whether many bacteria would metabolize lignin in their native environment. More recent studies have focused on developing new substrates for detecting lignin reactivity in bacteria, including a fluorescently-labeled native lignin and a chemically-nitrated lignin [*38*]. Comparison of lignin degrading and non-lignin degrading bacteria and fungi in the assay showed that the relative reactivity of bacterial systems is low compared to that of *P. chrysosporium* but also identified new lignin reactive bacteria, such as *Pseudomonas putida* and *Rhodococcus jostii* RHA1.

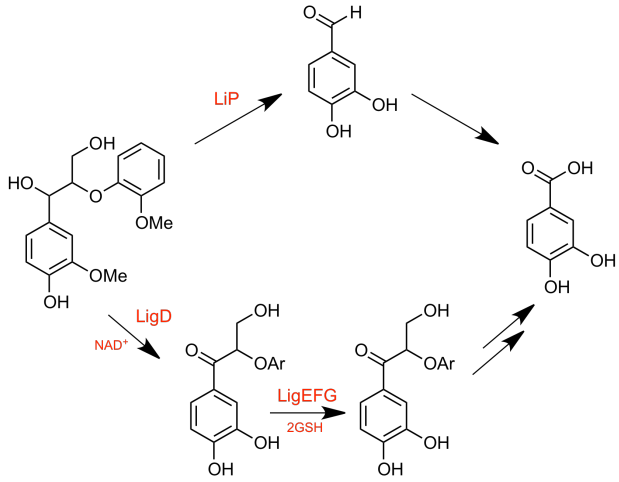
Heme peroxidases. Despite these reports, the bacterial enzymes responsible for these behaviors have yet to be identified and biochemically characterized [*37-39, 101, 105, 106*]. Studies by Ramachandra *et al.* revealed a lignocellulose-dependent peroxidase secretion profile in *Streptomyces viridosporus* T7A. Analysis of the secretome by non-denaturing PAGE followed by L-DOPA staining revealed the presence of four peroxidase isoforms [*107*]. While further characterization of one of these purified peroxidases showed evidence for non-phenolic lignin model dimer oxidation capacity [*108*], the identity of the protein remains elusive. However, the importance in identifying such proteins remains as Crawford *et al.* further correlated the extracellular peroxidase activity from *S. viridosporus* T7A and *S. chromofuscus* A2 with lignin solubilization when grown in nitrogen-free media [*98*].

Dye-decolorizing peroxidases (DyPs). DyPs were first discovered in 1999 [*88*] and shown to comprise a new family of heme peroxidases based on their lack of sequence homology to other peroxidases and classic peroxidase fold [*109*]. They also contain a distal aspartate in place of the typical histidine residue found in most heme peroxidases [*64, 109*]. The characterized DyPs have mostly been derived from fungi, some of which include characterized enzymes from *Bjerkandera adusta* and *Auricularia auricula-judae*, *Pleurotus ostreatus*, *Marasmius scorodonius*, and *Termitomyces albuminosus*. Interestingly, many have been found to be secreted [*88, 89*] and also demonstrate oxidative degradation activity against a suite of large aromatic dye molecules [*89, 110*]. After their discovery from fungi, the advent of lab-scale genome sequencing revealed the presence of bacterial DyPs. Thus, DyPs from various bacterial species were expressed and studied allowing a comparison between fungal and bacterial DyPs [*111-116*]. DyPs can be classified according to four different clades. In general, DyPs from bacteria are smaller in size and fall within the A– and B– clades, whereas fungal enzymes are typically found to cluster within the D–clade [*111, 113*].

The genome of the soil bacterium, *R. jostii* RHA1, revealed the presence of two DyP genes, *dypA* and *dypB*, residing in the A- and B-clades, respectively [*114*]. DypA was suggested to be involved in iron uptake and acquisition due to its similarity to EfeB and genomic context involving iron uptake [*114*]. DypB, however, was hypothesized to have involvement in lignin degradation. As such, deletion mutants were prepared from the native organism, producing both Δ*dypA* and Δ*dypB* strains [*111*]. Comparison studies of the mutants' lignolytic behavior resulted in lesser overall lignin modification by the organism, wherein it was slightly lower for the Δ*dypA* strain and significantly decreased in the Δ*dypB* strain. When heterologously expressed in and purified from *E. coli*, DypB, but not DypA, also showed low MnP activity, albeit with a catalytic efficiency 4-5 orders of magnitude lower than that of *P. chrysosporium* MnP. It was also determined that the presence of $MnCl_2$ increased assay reactivity 5- to 23-fold. In later studies, mutations to a proposed Mn-binding site in the enzyme produced 15-fold greater catalytic efficiency toward Mn oxidation in the enzyme [*117*]. Further mutational studies have suggested the distal pocket arginine, rather than the glutamate or aspartate, to be key in the peroxidase reactivity [*118*]. Though DypB does not appear to contain a canonical secretion signal, it was hypothesized that the presence of an encoded encapsulin molecule, with which the *dypB* gene is co-transcribed, could act as an agent for secretion [*119*]. To support this theory, the C-terminus of DyP appears to contain a ten amino acid consensus sequence that could target it to this nanocompartment. Gel filtration and dynamic light scattering (DLS) of the expressed encapsulin protein support the existence a large macromolecular assembly. After pH-induced disassembly of the encapsulin, gel filtration revealed that reassembly in the presence of DypB yielded the same

time retention, and analysis by SDS-PAGE revealed a correlation of DypB with this oligomeric peak. Interestingly, the activity of the encapsulated DypB was calculated to be seven-fold more active in the nitrated lignin assay than that of DypB alone.

Dioxygenases. Soil-dwelling bacteria have long-been hypothesized to take part in the overall environmental community lignin degradation, one of the primary supporting reasons being that many of these bacteria have expansive aromatic degradation systems [*120-122*]. Most of the work with these pathways have involved studies of the degradation of various aromatic monomers. As a result, it has been suggested that bacteria primarily uptake and digest aromatic monomers released from lignin degradation by other organisms [*120, 123*]. Recent studies of transporters have shed light on a suite of solute-binding proteins (SBPs) from ABC transporters, previously annotated as branched-chain amino-acid-binding proteins, showing their ability to bind a variety of aromatic monomeric lignin degradation products [*124, 125*]. However, whether SBPs exist for the transport of larger degradation products is unclear. Nonetheless, these aromatic degradation studies show that intracellular pathways exist that involve demethylation [*126-128*], ring-opening through dioxygenases [*120, 121, 129-131*], and conversion to primary metabolites [*120, 121, 132*]. These widely distributed reactivities give rise to the possibility for alternative breakdown pathways for lignin beyond the direct one-electron oxidation approaches utilized by fungi. For example, cleavage of the β-aryl-ether linkage can be catalyzed by known enzymes from *Sphingomonas paucimobilis*, wherein the α-hydroxyl group is oxidized to a ketone by the NAD-dependent dehydrogenase, LigD, and then glutathione-dependent enzymes, LigEFG, cleave the ketone (*Scheme 1.3*).



Scheme 1.2. *Lignin model dimer degradation by the fungal LiP is performed by a series of enzymes from the bacterium S. paucimobilis.*

Interestingly, however, Bianchetti *et al.* reported on an unusual dioxygenase from *Streptomyces* sp. SirexAA-E, a dioxygenase-carbohydrate binding module (CBM 5/12) fusion that was found to be secreted when the organism was grown on biomass or pure xylan, but not on glucose, cellulose or chitin [*133*]. A crystal structure of the full-length protein, solved to a resolution of 2.06 Å, showed a missing dimerization interface found in most non-heme dioxygenases, yielding instead a solvent-exposed depression where the catalytic center lies. Most interestingly, using a pull-down assay of insoluble substrates (synthetic lignin, chitin, and cellulose), the dioxygenase was shown to bind the synthetic lignin and chitin but not cellulose; also, with removal of the CBM (G246-WAAGTT-Tyr254-RAGD-R259), no pull-down was witnessed. Overall, the authors hypothesized that the enzyme localizes to the surface of lignin, intercepting biosynthetic intermediates, like caffeoyl-CoA, to inhibit lignin biosynthesis and, thus, plant protection.

Laccases. As fungal laccases have been implicated in playing a role in lignin degradation, this has been expanded to include some bacterial laccases. Many bacterial laccases have been heterologously expressed and biochemically characterized for general laccase activity, which has

shown ability for phenol oxidation [*134*]. More strains secreting laccases possibly involved in lignin degradation have been found, such as *Novosphingobium* sp. B-7 [*135*]. Also, the activity shown in the *P. putida* and *R. jostii* has been suggested to be laccase-dependent as activity is still seen in the developed assays absent of H_2O_2 [*38*].

1.5. Specific aims and thesis organization

In order to explore new chemical strategies for lignin degradation, we have initiated studies aimed at the discovery and characterization of oxidative and accessory enzymes in lignin-reactive soil bacteria that exhibit particularly rich activity towards the depolymerization and utilization of biomass-derived carbon sources [*136-138*]. Our interest lies in studying the biochemical logic underlying the transformation of complex substrates by living organisms with the overall goal of elucidating new molecular strategies for lignin degradation. Towards this goal, we have focused on studying the soil bacterium, *Amycolatopsis* sp. 75iv2 (formerly *Streptomcyes setonii* and *Streptomyces griseus* sp. 75vi2) as a model for bacterial lignin degradation. Questions of interest include exploring whether high oxidative bacterial enzymes that mimic or rival the activity of fungal lignolytic enzyme exist or whether due to the lesser oxidative systems, bacteria approach lignin degradation with a completely different set of enzymatic reactions.

Towards this goal, this thesis will be structured as follows: Chapter 2 discusses the identification of *A.* sp. 75iv2 as a model system as well as the assembly and analysis of its draft genome. Chapter 3 covers the identification and biochemical characterization of an extracellular catalase-peroxidase that represents the most abundant heme-containing enzyme found in the secretome under laboratory conditions. Chapter 4 describes the expression and characterization of a multi-functional dye peroxidase from *A.* sp. 75iv2 with peroxidase, Mn peroxidase, and Mn-dependent oxidase activities. Finally, Chapter 5 discusses characterization of the global responses of *A.* sp. 75iv2 to lignin using transcriptomic and proteomic analyses.

1.6. References

1. E. E. Benson, C. P. Kubiak, A. J. Sathrum and J. M. Smieja, Electrocatalytic and homogeneous approaches to conversion of CO_2 to liquid fuels, *Chem. Soc. Rev.* **2009**, *38*, 89-99.
2. T. Sakakura, J.-C. Choi and H. Yasuda, Transformation of carbon dioxide, *Chem. Rev.* **2007**, *107*, 2365-2387.
3. C. E. Laplaza and C. C. Cummins, Dinitrogen cleavage by a three-coordinate molybdenum(III) complex, *Science* **1995**, *268*, 861-863.
4. D. V. Yandulov and R. R. Schrock, Catalytic reduction of dinitrogen to ammonia at a single molybdenum center, *Science* **2003**, *301*, 76-78.
5. A. L. Feig and S. J. Lippard, Reactions of non-heme iron(II) centers with dioxygen in biology and chemistry, *Chem. Rev.* **1994**, *94*, 759-805.
6. J. T. Groves, High-valent iron in chemical and biological oxidations, *J. Inorg. Biochem.* **2006**, *100*, 434-447.
7. L. Que and W. B. Tolman, Biologically inspired oxidation catalysis, *Nature* **2008**, *455*, 333-340.

8. R. Breslow, Biomimetic chemistry: Biology as an inspiration, *J. Biol. Chem.* **2009**, *284*, 1337-1342.
9. W. B. Motherwell, M. J. Bingham and Y. Six, Recent progress in the design and synthesis of artificial enzymes, *Tetrahedron* **2001**, *57*, 4663-4686.
10. A. J. van den Wijngaard, R. D. Wind and D. B. Janssen, Kinetics of bacterial growth on chlorinated aliphatic compounds, *Appl. Environ. Microbiol.* **1993**, *59*, 2041-2048.
11. R. A. Kanaly and S. Harayama, Biodegradation of high-molecular-weight polycyclic aromatic hydrocarbons by bacteria, *J. Bacteriol.* **2000**, *182*, 2059-2067.
12. K. C. Nicolaou, N. Winssinger, J. Pastor, S. Ninkovic, F. Sarabia, Y. He, D. Vourloumis, Z. Yang, T. Li, P. Giannakakou and E. Hamel, Synthesis of epothilones A and B in solid and solution phase, *Nature* **1997**, *387*, 268-272.
13. D. E. Cane, Programming of erythromycin biosynthesis by a modular polyketide synthase, *J. Biol. Chem.* **2010**, *285*, 27517-27523.
14. J. F. Aparicio, I. Molnár, T. Schwecke, A. König, S. F. Haydock, L. Ee Khaw, J. Staunton and P. F. Leadlay, Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: Analysis of the enzymatic domains in the modular polyketide synthase, *Gene* **1996**, *169*, 9-16.
15. C. Vezina, A. Kudelski and S. N. Sehgal, Rapamycin (AY-22,989), a new antifungal antibiotic. I. Taxonomy of the producing streptomycete and isolation of the active principle, *J. Antibiot. (Tokyo)* **1975**, *28*, 721-726.
16. D. K. Ro, E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, J. M. Ndungu, K. A. Ho, R. A. Eachus, T. S. Ham, J. Kirby, M. C. Y. Chang, S. T. Withers, Y. Shiba, R. Sarpong and J. D. Keasling, Production of the antimalarial drug precursor artemisinic acid in engineered yeast, *Nature* **2006**, *440*, 940-943.
17. M. Wen, B. B. Bond-Watts and M. C. Chang, Production of advanced biofuels in engineered *E. coli*, *Curr. Opin. Chem. Biol.* **2013**, *17*, 472-479.
18. C. R. Shen and J. C. Liao, Metabolic engineering of *Escherichia coli* for 1-butanol and 1-propanol production via the keto-acid pathways, *Metab. Eng.* **2008**, *10*, 312-320.
19. J. D. Keasling, Synthetic biology for synthetic chemistry, *ACS Chem. Biol.* **2008**, *3*, 64-76.
20. V. G. Yadav and G. Stephanopoulos, Reevaluating synthesis by biology, *Curr. Opin. Microbiol.* **2010**, *13*, 371-376.
21. A. M. Weeks and M. C. Y. Chang, Constructing de novo biosynthetic pathways for chemical synthesis inside living cells, *Biochemistry* **2011**, *50*, 5404-5418.
22. DOE, The Annual Energy Outlook 2010, *U.S. Department of Energy* **2010**,
23. J. W. A. Langeveld, J. Dixon and J. F. Jaworski, Development perspectives of the biobased economy: A review, *Crop Sci.* **2010**, *50*, S142-S151.
24. P. A. M. Claassen, J. B. van Lier, A. M. Lopez Contreras, E. W. J. van Niel, L. Sijtsma, A. J. M. Stams, S. S. de Vries and R. A. Weusthuis, Utilisation of biomass for the supply of energy carriers, *Appl. Microbiol. Biotechnol.* **1999**, *52*, 741-755.
25. V. Mendu, T. Shearin, J. E. Campbell, J. Stork, J. Jae, M. Crocker, G. Huber and S. DeBolt, Global bioenergy potential from high-lignin agricultural residue, *Proc. Natl. Acad. Sci.* **2012**, *109*, 4014-4019.

26. D. L. Greene and S. Ahmad, Costs of U.S. oil dependence: 2005 update, *Oak Ridge National Laboratory* **2005**, 1-64.

27. N. Greene, Growing energy: How biofuels can help end America's oil dependence, *Natural Resources Defense Council* **2004**, 1-96.

28. M. E. Himmel, S.-Y. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, J. W. Brady and T. D. Foust, Biomass recalcitrance: Engineering plants and enzymes for biofuels production, *Science* **2007**, *315*, 804-807.

29. J. Ralph, K. Lundquist, G. Brunow, F. Lu, H. Kim, P. F. Schatz, J. M. Marita, R. D. Hatfield, S. A. Ralph, J. H. Christensen and W. Boerjan, Lignins: Natural polymers from oxidative coupling of 4-hydroxylphenylpropanoids, *Phytochem. Rev.* **2004**, *3*, 29-60.

30. T. K. Kirk and R. L. Farrell, Enzymatic "combustion": The microbial degradation of lignin, *Annu. Rev. Microbiol.* **1987**, *41*, 465-505.

31. R. ten Have and P. J. Teunissen, Oxidative mechanisms involved in lignin degradation by white-rot fungi, *Chem. Rev.* **2001**, *101*, 3397-413.

32. P. Azadi, O. R. Inderwildi, R. Farnood and D. A. King, Liquid fuels, hydrogen and chemicals from lignin: A critical review, *Ren. Sustain. En. Rev.* **2013**, *21*, 506-523.

33. M. P. Pandey and C. S. Kim, Lignin depolymerization and conversion: A review of thermochemical methods, *Chem. Eng. Technol.* **2011**, *34*, 29-41.

34. A. Khindaria, T. A. Grover and S. D. Aust, Evidence for formation of the veratryl alcohol cation radical by lignin peroxidase, *Biochemistry* **1995**, *34*, 6020-6025.

35. A. Khindaria, D. P. Barr and S. D. Aust, Lignin peroxidases can also oxidize manganese, *Biochemistry* **1995**, *34*, 7773-7779.

36. M. Ramachandra, D. L. Crawford and G. Hertel, Characterization of an extracellular lignin peroxidase of the lignocellulolytic actinomycete *Streptomyces viridosporus*, *Appl. Env. Microbiol.* **1988**, *54*, 3057-3063.

37. R. Kirby, Actinomycetes and lignin degradation, In *Advances in Applied Microbiology*, Eds. A. I. Laskin, J. W. Bennett, G. M. Gadd and S. Sariaslani, Academic Press: London, UK, **2005**, vol. 58, pp. 125-168.

38. M. Ahmad, C. R. Taylor, D. Pink, K. Burton, D. Eastwood, G. D. Bending and T. D. H. Bugg, Development of novel assays for lignin degradation: Comparative analysis of bacterial and fungal lignin degraders, *Mol. BioSyst.* **2010**, *6*, 815-821.

39. D. L. Crawford, A. L. Pometto, III and R. L. Crawford, Lignin degradation by *Streptomyces viridosporus*: Isolation and characterization of a new polymeric lignin degradation intermediate, *Appl. Environ. Microbiol.* **1983**, *45*, 898-904.

40. W.-J. Hu, S. A. Harding, J. Lung, J. L. Popko, J. Ralph, D. D. Stokke, C.-J. Tsai and V. L. Chiang, Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees, *Nat. Biotechnol.* **1999**, *17*, 808-812.

41. N. D. Bonawitz and C. Chapple, The genetics of lignin biosynthesis: Connecting genotype to phenotype, *Annu. Rev. Genet.* **2010**, *44*, 337-363.

42. M. Liang, E. Davis, D. Gardner, X. Cai and Y. Wu, Involvement of *AtLAC15* in lignin synthesis in seeds and in root elongation of *Arabidopsis*, *Planta* **2006**, *224*, 1185-1196.

43. J. Zhou, C. Lee, R. Zhong and Z.-H. Ye, MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*, *Plant Cell* **2009**, *21*, 248-266.

44. R. Vanholme, K. Morreel, J. Ralph and W. Boerjan, Lignin engineering, *Curr. Opin. Plant. Biol.* **2008**, *11*, 278-85.

45. K. K. Pandey, A study of chemical structure of soft and hardwood and wood polymers by FTIR spectroscopy, *J. Appl. Polym. Sci.* **1999**, *71*, 1969-1975.

46. J. D. McMillan, Pretreatment of Lignocellulosic Biomass, In *Enzymatic Conversion of Biomass for Fuels Production*, Eds. M. E. Himmel, J. O. Baker and R. P. Overend, American Chemical Society: Washington, DC, **1994**, vol. 566, pp. 292-324.

47. W. Boerjan, J. Ralph and M. Baucher, Lignin biosynthesis, *Ann. Rev. Plant Biol.* **2003**, *54*, 519-546.

48. H. G. Jung and D. A. Deetz, Cell wall lignification and degradability, In *Forage Cell Wall Structure and Digestibility*, Eds. H. G. Jung, D. R. Buxton, R. D. Hatfield and J. Ralph, ASA-CSSA-SSSA: Madison, WI, **1993**, vol. pp. 315-346.

49. H. Yu, G. Guo, X. Zhang, K. Yan and C. Xu, The effect of biological pretreatment with the selective white-rot fungus *Echinodontium taxodii* on enzymatic hydrolysis of softwoods and hardwoods, *Biores. Technol.* **2009**, *100*, 5170-5175.

50. S. K. Huntley, D. Ellis, M. Gilbert, C. Chapple and S. D. Mansfield, Significant increases in pulping efficiency in C4H-F5H-transformed poplars: Improved chemical savings and reduced environmental toxins, *J. Agric. Food Chem.* **2003**, *51*, 6178-6183.

51. F. Chen and R. A. Dixon, Lignin modification improves fermentable sugar yields for biofuel production, *Nat. Biotechnol.* **2007**, *25*, 759-761.

52. A. L. Schilmiller, J. Stout, J.-K. Weng, J. Humphreys, M. O. Ruegger and C. Chapple, Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in *Arabidopsis*, *Plant J.* **2009**, *60*, 771-782.

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

54. M. Mansur, M. E. Arias, J. Copa-Patiño, M. Flärdh and A. E. González, The white-rot fungus *Pleurotus ostreatus* secretes laccase isozymes with different substrate specificities, *Mycologia* **2003**, *95*, 1013-1020.

55. E. Fernandez-Fueyo*, et al.*, Comparative genomics of *Ceriporiopsis subvermispora* and *Phanerochaete chrysosporium* provide insight into selective ligninolysis, *Proc. Natl. Acad. Sci.* **2012**, *109*, 5458-5463.

56. E. Fernández-Fueyo, F. J. Ruiz-Dueñas, Y. Miki, M. J. Martínez, K. E. Hammel and A. T. Martínez, Lignin-degrading peroxidases from genome of selective ligninolytic fungus *Ceriporiopsis subvermispora*, *J. Biol. Chem.* **2012**, *287*, 16903-16916.

57. R. Bourbonnais, M. G. Paice, I. D. Reid, P. Lanthier and M. Yaguchi, Lignin oxidation by laccase isozymes from *Trametes versicolor* and role of the mediator 2,2'-azinobis(3-ethylbenzthiazoline-6-sulfonate) in kraft lignin depolymerization, *Appl. Environ. Microbiol.* **1995**, *61*, 1876-80.

58. C. S. Evans, Laccase activity in lignin degradation by *Coriolus versicolor in vivo* and *in vitro* studies, *FEMS Microbiol. Lett.* **1985**, *27*, 339-343.

59. D. Floudas*, et al.*, The paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes, *Science* **2012**, *336*, 1715-1719.

60. D. Martinez*, et al.*, Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion, *Proc. Natl. Acad. Sci.* **2009**, *106*, 1954-1959.

61. I. A. Yakovlev, A. M. Hietala, P.-E. Courty, T. Lundell, H. Solheim and C. G. Fossdal, Genes associated with lignin degradation in the polyphagous white-rot pathogen *Heterobasidion irregulare* show substrate-specific regulation, *Fungal Genet. Biol.* **2013**, *56*, 17-24.

62. B. H. Dunford, Heme Peroxidases, Wiley-VCH: Weinheim, Germany, **1999**, vol. pp. 532.

63. A. T. Smith and N. C. Veitch, Substrate binding and catalysis in heme peroxidases, *Curr. Opin. Chem. Biol.* **1998**, *2*, 269-278.

64. A. M. English and G. Tsaprailis, Catalytic structure-function relationships in heme peroxidases, *Adv. Inorg.Chem.* **1995**, *43*, 79-125.

65. R. L. Farrell, K. E. Murtagh, M. Tien, M. D. Mozuch and T. K. Kirk, Physical and enzymatic properties of lignin peroxidase isoenzymes from *Phanerochaete chrysosporium*, *Enzyme Microb. Technol.* **1989**, *11*, 322-328.

66. S. D. Aust, H. Tuisel, R. Sinclair, J. A. Bumpus, W. Ashbaugh and B. J. Brock, Lignin peroxidase H2 from *Phanerochaete chrysosporium*: Purification, characterization and stability to temperature and pH, *Arch. Biochem. Biophys.* **1990**, *279*, 158-166.

67. G. Nie, N. S. Reading and S. D. Aust, Expression of the lignin peroxidase H2 gene from *Phanerochaete chrysosporium* in *Escherichia coli*, *Biochem. Biophys. Res. Comm.* **1998**, *249*, 146-150.

68. W. A. Doyle and A. T. Smith, Expression of lignin peroxidase H8 in *Escherichia coli*: Folding and activation of the recombinant enzyme with Ca^{2+} and haem, *Biochem. J.* **1996**, *315*, 15-19.

69. S. Sato, F. A. Feltus, P. Iyer and M. Tien, The first genome-level transcriptome of the wood-degrading fungus *Phanerochaete chrysosporium* grown on red oak, *Curr. Gen.* **2009**, *55*, 273-286.

70. K. A. Jensen, K. M. C. Evans, T. K. Kirk and K. E. Hammel, Biosynthetic pathway for veratryl alcohol in the ligninolytic fungus *Phanerochaete chrysosporium*, *Appl. Environ. Microbiol.* **1994**, *60*, 709-714.

71. E. Baciocchi, M. F. Gerini, O. Lanzalunga and S. Mancinelli, Lignin peroxidase catalysed oxidation of 4-methoxymandelic acid. The role of mediator structure, *Tetrahedron* **2002**, *58*, 8087-8093.

72. C. D. Millis, D. Cai, M. T. Stankovich and M. Tien, Oxidation-reduction potentials and ionization states of extracellular peroxidases from the lignin-degrading fungus *Phanerochaete chrysosporium*, *Biochemistry* **1989**, *28*, 8484-8489.

73. M. Sundaramoorthy, H. Youngs, M. H. Gold and T. L. Poulos, High-resolution crystal structure of manganese peroxidase: Substrate and inhibitor complexes, *Biochemistry* **2005**, *44*, 6463-6470.

74. M. Tien, E. A. Pease and A. Andrawis, Manganese-dependent peroxidase from *Phanerochaete chrysosporium*, *J. Biol. Chem.* **1989**, *264*, 13531-13535.

75. M. Kusters-van Someren, K. Kishi, T. Lundell and M. H. Gold, The manganese binding site of manganese peroxidase: characterization of an Asp179Asn site-directed mutant protein, *Biochem.* **1995**, *34*, 10620-10627.

76. R. E. Whitwam, I. G. Gazarian and M. Tien, Expression of fungal Mn peroxidase in *E. coli* and refolding to yield active enzyme, *Biochem. Biophys. Res. Comm.* **1995**, *216*, 1013-1017.

77. C. Lapierre, D. Jouin and B. Monties, Lignin characterization of wheat straw samples as determined by chemical degradation procedures, In *Physico-chemical characterization of plant residues for industrial and feed use*, Eds. A. Chesson and E. R. Orskov, Elsevier Applied Science: Amsterdam, **1989**, vol. pp. 118-130.

78. H. Wariishi, K. Valli and M. H. Gold, Manganese(II) oxidation by manganese peroxidase from the basidiomycete *Phanerochaete chrysosporium*. Kinetic mechanism and role of chelators, *J. Biol. Chem.* **1992**, *267*, 23688-23695.

79. N. Chung and S. D. Aust, Inactivation of lignin peroxidase by hydrogen peroxide during the oxidation of phenols, *Arch. Biochem. Biophys.* **1995**, *316*, 851-855.

80. S. Camarero, S. Sarkar, F. J. Ruiz-Dueñas, M. J. Martínez and Á. T. Martínez, Description of a versatile peroxidase involved in the natural degradation of lignin that has both manganese peroxidase and lignin peroxidase substrate interaction sites, *J. Biol. Chem.* **1999**, *274*, 10324-10330.

81. M. Pérez-Boada, F. J. Ruiz-Dueñas, R. Pogni, R. Basosi, T. Choinowski, M. J. Martínez, K. Piontek and A. T. Martínez, Versatile peroxidase oxidation of high redox potential aromatic compounds: Site-directed mutagenesis, spectroscopic and crystallographic investigation of three long-range electron transfer pathways, *J. Mol. Biol.* **2005**, *354*, 385-402.

82. R. Bourbonnais and M. G. Paice, Oxidation of non-phenolic substrates: An expanded role for laccase in lignin biodegradation, *FEBS Lett.* **1990**, *267*, 99-102.

83. V. B. Urlacher and S. Eiben, Cytochrome P450 monooxygenases: Perspectives for synthetic application, *Trends biotechnol.* **2006**, *24*, 324-330.

84. I. G. Denisov, T. M. Makris, S. G. Sligar and I. Schlichting, Structure and chemistry of cytochrome P450, *Chem. Rev.* **2005**, *105*, 2253-2278.

85. D. Martinez, L. F. Larrondo, N. Putnam, M. D. Gelpke, K. Huang, J. Chapman, K. G. Helfenbein, P. Ramaiya, J. C. Detter, F. Larimer, P. M. Coutinho, H. B., R. Berka, D. Cullen and D. Rokhsar, Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78, *Nat. Biotechnol.* **2004**, *6*, 695-700.

86. H. Doddapaneni and J. S. Yadav, Microarray-based global differential expression profiling of P450 monooxygenases and regulatory proteins for signal transduction pathways in the white rot fungus *Phanerochaete chrysosporium*, *Mol. Gen. Genomics* **2005**, *274*, 454-466.

87. S. Shary, A. N. Kapich, E. A. Panisko, J. K. Magnuson, D. Cullen and K. E. Hammel, Differential expression in *Phanerochaete chrysosporium* of membrane-associated proteins relevant to lignin degradation, *Appl Environ Microbiol* **2008**, *74*, 7252-7257.

88. S. J. Kim and M. Shoda, Purification and characterization of a novel peroxidase from *Geotrichum candidum* Dec 1 involved in decolorization of dyes, *Appl. Env. Microbiol.* **1999**, *65*, 1029-1035.

89. C. Liers, C. Bobeth, M. Pecyna, R. Ullrich and M. Hofrichter, DyP-like peroxidases of the jelly fungus *Auricularia auricula-judae* oxidize nonphenolic lignin model compounds and high-redox potential dyes, *Appl. Microbiol. Biotechnol.* **2010**, *85*, 1869-1879.

90. Y. Sugano, DyP-type peroxidases comprise a novel heme peroxidase family, *Cell. Mol. Life Sci.* **2009**, *66*, 1387-1403.

91. A. Kaur, P. T. Van, C. R. Busch, C. K. Robinson, M. Pan, W. L. Pang, D. J. Reiss, J. DiRuggiero and N. S. Baliga, Coordination of frontline defense mechanisms under severe oxidative stress, *Mol. Syst. Biol.* **2010**, *6*, 1-16.

92. S. Létoffé, G. Heuck, P. Delepelaire, N. Lange and C. Wandersman, Bacteria capture iron from heme by keeping tetrapyrrol skeleton intact, *Proc. Natl. Acad. Sci.* **2009**, *106*, 11719-11724.

93. J. Pellinen, E. Vaisanen, M. Salkinoja-Salonen and G. Brunow, Utilization of dimeric lignin model compounds by mixed bacterial cultures, *Appl. Microbiol. Biotechnol.* **1984**, *20*, 77-82.

94. L. Eggeling and H. Sahm, Degradation of coniferyl alcohol and other lignin-related aromatic compounds by *Nocardia* sp. DSM 1069, *Arch. Microbiol.* **1980**, *126*, 141-148.

95. H. W. Kern and T. K. Kirk, Influence of molecular size and ligninase pretreatment on degradation of lignins by *Xanthomonas* sp. strain 99, *Appl. Environ. Microbiol.* **1987**, *53*, 2242-2246.

96. R. Vicuña, Bacterial degradation of lignin, *Enz. Microb. Technol.* **1988**, *10*, 646-655.

97. C. R. Taylor, E. M. Hardiman, M. Ahmad, P. D. Sainsbury, P. R. Norris and T. D. H. Bugg, Isolation of bacterial strains able to metabolize lignin from screening of environmental samples, *J. Appl. Microbiol.* **2012**, *113*, 521-530.

98. M. B. Pasti, S. R. Hagen, R. A. Korus and D. L. Crawford, The effects of various nutrients on extracellular peroxidases and acid-precipitable polymeric lignin production by *Streptomyces chromofuscus* A2 and *S. viridosporus* T7A, *Appl. Microbiol. Biotechnol.* **1991**, *34*, 661-667.

99. H. Sørensen, Decomposition of lignin by soil bacteria and complex formation between autoxidized lignin and organic nitrogen compounds, *J. Gen. Microbiol.* **1962**, *27*, 21-34.

100. J. Trojanowski, K. Haider and V. Sundman, Decomposition of ^{14}C-labelled lignin and phenols by a *Nocardia* sp., *Arch. Microbiol.* **1977**, *114*, 149-153.

101. D. L. Crawford, Lignocellulose decomposition by selected *Streptomyces* strains, *Appl. Environ. Microbiol.* **1978**, *35*, 1041-1045.

102. H. W. Kern, Bacterial degradation of dehydropolymers of coniferyl alcohol, *Arch. Microbiol.* **1984**, *138*, 18-25.

103. T. J. Kerr, R. D. Kerr and R. Benner, Isolation of a bacterium capable of degrading peanut hull lignin, *Appl. Environ. Microbiol.* **1983**, *43*, 1201-1206.

104. A. Raj, R. Chandra, M. M. K. Reddy, H. Purohit and A. Kapley, Biodegradation of kraft lignin by a newly isolated bacterial strain, *Aneurinibacillus aneurinilyticus* from the sludge of a pulp paper mill, *World J. Microbiol. Biotechnol.* **2007**, *23*, 793-799.

105. J. R. Borgmeyer and D. L. Crawford, Production and characterization of polymeric lignin degradation intermediates from two different *Streptomyces* spp., *Appl. Environ. Microbiol.* **1985**, *49*, 273-278.

106. L. Thomas and D. L. Crawford, Cloning of clustered *Streptomyces viridosporus* T7A lignocellulose catabolism genes encoding peroxidase and endoglucanase and their extracellular expression in *Pichia pastoris*, *Can. J. Microbiol.* **1998**, *44*, 364-372.

107. M. Ramachandra, D. L. Crawford and A. L. Pometto, Extracellular enzyme activities during lignocellulose degradation by *Streptomyces* spp.: A comparative study of wild-type and genetically manipulated strains, *Appl. Environ. Microbiol.* **1987**, *53*, 2754-2760.

108. H. Giroux, P. Vidal, J. Bouchard and F. Lamy, Degradation of kraft lignin by *Streptomyces viridosporus* and *Streptomyces badius*, *Appl. Environ. Microbiol.* **1988**, *54*, 3064-3070.

109. T. Sato, S. Hara, T. Matsui, G. Sazaki, S. Saijo, T. Ganbe, N. Tanaka, Y. Sugano and M. Shoda, A unique dye-decolorizing peroxidase, DyP, from *Thanatephorus cucumeris* Dec 1: Heterologous expression, crystallization and preliminary X-ray analysis, *Acta Crystallogr. D* **2004**, *60*, 149-152.

110. Y. Sugano, Y. Matsushima and M. Shoda, Complete decolorization of the anthraquinone dye Reactive blue 5 by the concerted action of two peroxidases from *Thanatephorus cucumeris* Dec 1, *Appl. Microbiol. Biotechnol.* **2006**, *73*, 862-871.

111. M. Ahmad, J. N. Roberts, E. M. Hardiman, R. Singh, L. D. Eltis and T. D. H. Bugg, Identification of DypB from *Rhodococcus jostii* RHA1 as a lignin peroxidase, *Biochemistry* **2011**, *50*, 5096-5107.

112. J. Li, C. Liu, B. Li, H. Yuan, J. Yang and B. Zheng, Identification and molecular characterization of a novel DyP-type peroxidase from *Pseudomonas aeruginosa* PKE117, *Appl. Biochem. Biotechnol.* **2012**, *166*, 774-785.

113. H. J. O. Ogola, T. Kamiike, N. Hashimoto, H. Ashida, T. Ishikawa, H. Shibata and Y. Sawa, Molecular characterization of a novel peroxidase from the cyanobacterium *Anabaena* sp. Strain PCC 7120, *Appl. Env. Microbiol.* **2009**, *75*, 7509-7518.

114. J. N. Roberts, R. Singh, J. C. Grigg, M. E. P. Murphy, T. D. H. Bugg and L. D. Eltis, Characterization of dye-decolorizing peroxidases from *Rhodococcus jostii* RHA1, *Biochemistry* **2011**, *50*, 5108-5119.

115. C. Zubieta*, et al.*, Identification and structural characterization of heme binding in a novel dye-decolorizing peroxidase, TyrA, *Proteins Struct. Funct. Bioinform.* **2007**, *69*, 234-243.

116. M. E. Brown, T. Barros and M. C. Y. Chang, Identification and characterization of a multifunctional dye peroxidase from a lignin-reactive bacterium, *ACS Chem. Biol.* **2012**, *7*, 2074-2081.

117. R. Singh, J. C. Grigg, W. Qin, J. F. Kadla, M. E. P. Murphy and L. D. Eltis, Improved manganese-oxidizing activity of DypB, a peroxidase from a lignolytic bacterium, *ACS Chem. Biol.* **2013**, *8*, 700-706.

118. R. Singh, J. C. Grigg, Z. Armstrong, M. E. P. Murphy and L. D. Eltis, Distal heme pocket residues of B-type dye-decolorizing peroxidase: Arginine but not aspartate is essential for peroxidase activity, *J. Biol. Chem.* **2012**, *287*, 10623-10630.

119. R. Rahmanpour and T. D. H. Bugg, Assembly *in vitro* of *Rhodococcus jostii* RHA1 encapsulin and peroxidase DypB to form a nanocompartment, *FEBS J.* **2013**, *280*, 2097-2104.

120. E. Masai, Y. Katayama, S. Nishikawa and M. Fukuda, Characterization of *Sphingomonas paucimobilis* SYK-6 genes involved in degradation of lignin-related compounds, *J. Ind. Microbiol. Biotechnol.* **1999**, *23*, 364-373.

121. F. Franklin, M. Bagdasarian, M. Bagdasarian and K. Timmis, Molecular and functional analysis of the TOL plasmid pWWO from *Pseudomonas putida* and cloning of genes for the entire regulated aromatic ring meta cleavage pathway, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 7458-7462.

122. A. Basu, S. K. Apte and P. S. Phale, Preferential utilization of aromatic compounds over glucose by *Pseudomonas putida* CSV86, *Appl. Environ. Microbiol.* **2006**, *72*, 2226-2230.

123. D. L. Kaplan and R. Hartenstein, Decomposition of lignins by microorganisms, *Soil Biol. Biochem.* **1980**, *12*, 65-75.

124. K. Michalska, C. Chang, J. C. Mack, S. Zerbs, A. Joachimiak and F. R. Collart, Characterization of transport proteins for aromatic compounds derived from lignin: Benzoate derivative binding proteins, *J. Mol. Biol.* **2012**, *423*, 555-575.

125. R. Pietri, S. Zerbs, D. M. Corgliano, M. Allaire, F. R. Collart and L. M. Miller, Biophysical and structural characterization of a sequence-diverse set of solute-binding proteins for aromatic compounds, *J. Biol. Chem.* **2012**, *287*, 23748-23756.

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

127. F. Kaufmann, G. Wohlfarth and G. Diekert, *O*-Demethylase from A*cinetobacterium dehalogenans*-Cloning, sequencing, and active expression of the gene encoding the corrinoid protein, *Eur. J. Biochem.* **1998**, *257*, 515-521.

128. J. B. Sutherland, Demethylation of veratrole by cytochrome P-450 in *Streptomyces setonii*, *Appl. Environ. Microbiol.* **1986**, *52*, 98-100.

129. D. Kasai, E. Masai, K. Miyauchi, Y. Katayama and M. Fukuda, Characterization of the gallate dioxygenase gene: three distinct ring cleavage dioxygenases are involved in syringate degradation by *Sphingomonas paucimobilis* SYK-6, *J. Bacteriol.* **2005**, *187*, 5067-5074.

130. S. Harayama, M. Rekik, M. Wubbolts, K. Rose, R. A. Leppik and K. N. Timmis, Characterization of five genes in the upper-pathway operon of TOL plasmid pWW0 from *Pseudomonas putida* and identification of the gene products, *J. Bacteriol.* **1989**, *171*, 5048-55.

131. T. Sonoki, Y. Otsuka, S. Ikeda, E. Masai, S. Kajita and Y. Katayama, Close association between the enzymes involved in the lignin metabolic pathway of *Sphingomonas paucimobilis* SYK-6: Interaction of *O*-demethylase (LigX) and ring fission dioxygenase (LigZ), *J. Wood Sci.* **2002**, *48*, 250-252.

132. J. Nogales, R. Macchi, F. Franchi, D. Barzaghi, C. Fernandez, J. L. Garcia, G. Bertoni and E. Diaz, Characterization of the last step of the aerobic phenylacetic acid degradation pathway, *Microbiology* **2007**, *153*, 357-365.

133. C. M. Bianchetti, C. H. Harmann, T. E. Takasuka, G. L. Hura, K. Dyer and B. G. Fox, Fusion of dioxygenase and lignin-binding domains in a novel secreted enzyme from cellulolytic *Streptomyces* sp. SirexAA-E, *J. Biol. Chem.* **2013**, *288*, 18574-18587.

134. M. C. Machczynski, E. Vijgenboom, B. Samyn and G. W. Canters, Characterization of SLAC: A small laccase from *Streptomyces coelicolor* with unprecedented activity, *Protein Sci.* **2004**, *13*, 2388-2397.

135. Y. Chen, L. Chai, C. Tang, Z. Yang, Y. Zheng, Y. Shi and H. Zhang, Kraft lignin biodegradation by *Novosphingobium* sp. B-7 and analysis of the degradation process, *Bioresource Technol.* **2012**, *123*, 682-685.

136. T. C. Balser, A. P. Kinzig and M. K. Firestone, Linking soil microbial communities and ecosystem functioning, In *The functional consequences of biodiversity: empirical progress and theoretical extensions*, Eds. A. Kinzig, S. Pacala and D. Tilman, Princeton University Press: Princeton, N. J., **2002**, vol. pp. 265-293.

137. V. Torsvik and L. Øvreås, Microbial diversity and function in soil: from genes to ecosystems, *Curr. Opin. Microbiol.* **2002**, *5*, 240-245.

138. S. Hättenschwiler, A. V. Tiunov and S. Scheu, Biodiversity and litter decomposition in terrestrial ecosystems, *Annu. Rev. Ecol. Evol. Syst.* **2005**, *36*, 191-218.

Chapter 2: Assembly of a *de novo* draft genome of *Amycolatopsis* sp. 75iv2

Portions of this work were published in the following scientific journal:

M. E. Brown, M. C. Walker, T. G. Nakashige, A. T. Iavarone and M. C. Y. Chang, Discovery and characterization of heme enzymes from unsequenced bacteria: Application to microbial lignin degradation, *J. Am. Chem. Soc.* **2011**, *133*, 18006-18009.

2.1. Introduction

Due to the complex chemistry and biology of soil, organisms that reside in this environment are often found to have evolved the ability to take advantage of the broad range of carbon sources and inorganic nutrients available in soil using an expansive array of biotransformations and metabolic processes [*1-7*]. One important carbon source is lignocellulosic biomass derived from plants, which provides a rich source of sugar- and aromatic-based carbon that is trapped in the form of complex biopolymers [*8, 9*]. We have been particularly interested in characterizing pathways for the degradation of lignin in microbial hosts beyond the well-studied lignolytic fungi [*10-13*] with the overall goal of identifying new lignin-modifying enzymes. Towards this goal, we have focused on exploring soil bacteria and their lignin reactivity, as they have been reported to react with lignin but in a potentially different manner than fungi [*14-18*]. One of the metabolic products produced from lignin by actinomycetes is acid–precipitable polymeric lignin (APPL), whose structure remains uncharacterized [*15, 19*]. However, the formation of APPL can be utilized as a biomarker for bacterial lignin degradation. We therefore monitored the growth of several soil bacteria in the presence of lignocellulose from *Miscanthus giganteus*, while chacterizing their extracellular oxidative capacity as well as their production of APPL. From these studies, we identified *Amycolatopsis* sp. 75iv2 by its maximal APPL production and extracellular peroxidase activity.

We then turned our attention to the genetic and biochemical characterization of the *A.* sp. 75iv2 secretome. As its genome was unsequenced, we utilized massively parallel next-generation sequencing technology based on the Illumina platform [*20*] to assemble a *de novo* draft genome. From analysis of the genome, we were able to both explore its genetic potential for biomass degradation and found that the genome encoded a large number of lignocellulose-degrading proteins, including secreted enzymes of the glycosyl hydrolase superfamily and a substantial number of canonical [*21-24*] and cryptic clusters for aromatic degradation. We also identified a broad range of potential candidates for its oxidative system, including a large number of peroxidases, laccases, cytochrome P450s, among others.

2.2. Materials and methods

Reagent information. Malt extract, potassium phosphate monobasic, yeast extract, cupric sulfate pentahydrate, and glycerol were purchased from EMD Biosciences (Darmstadt, Germany). Ammonium sulfate, D-glucose, magnesium chloride hexahydrate, and manganese chloride tetrahydrate were purchased from Fisher Scientific (Pittsburgh, PA). Potassium phosphate dibasic, 4-aminoantipyrene, 2,4-dichlorophenol, ammonium bicarbonate, magnesium sulfate heptahydrate, iron sulfate heptahydrate, and zinc sulfate monohydrate were purchased from Sigma-Aldrich (St. Louis, MO). Bacto™ Peptone and agar were purchased from BD (Franklin Lakes, NJ).

Bacterial strains. *Amycolatopsis* sp. 75iv2 (formerly *Streptomyces setonii* and *Streptomyces griseus* 75iv2, ATCC 39116), *Streptomyces viridosporus* (ATCC 39115), *Streptomyces badius* (ATCC 39117), and *Arthrobacter chlorophenolicus* (ATCC 700700) were purchased from the American Tissue Type Collection (Manassas, VA). *Rhodococcus opacus* (DSMZ 1069) was purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH

(Braunschweig, Germany). *Streptomyces griseus* IFO13350 was obtained from the ARS Actinobacterial Culture Collection at the United States Department of Agriculture (Peoria, IL).

Cell culture. Bacterial strains were cultured from spores (*Amycolatopsis* sp. 75iv2, *Streptomyces* spp.) or cells (*Arthrobacter cholorophenolicus* and *Rhodococcus opacus*) resuspended in 20% sterile glycerol and frozen at -80 °C. For peroxidase expression, frozen spores were spread onto YEME agar (2% w/v) plates using sterile glass beads (5 mm) and incubated at 30 °C for 3-4 d until plates were covered in spores. Spores from 3-4 d YEME agar plates were resuspended in sterile ddH_2O and used to inoculate expression media containing carbenicillin (50 μg/mL) and either *Miscanthus giganetus* lignocellulose or indulin. Otherwise, a single colony was used to inoculate the expression media. Control flasks containing either no lignin source or no spores/cells were included in the experiment to monitor for possible environmental contamination. Small-scale (50 mL) and large-scale (500 mL) cell cultures were carried out in 250-mL and 2-L baffled flasks, respectively. The cultures were incubated at 37 °C at 200 rpm, except for *Arthrobacter chlorophenolicus* (30 °C). Daily samples were removed (small-scale, 2 mL; large-scale, 40 mL).

Expression media. Growth media was prepared according to literature methods [*25, 26*]. Basal medium included (per liter): ammonium sulfate (2.64 g), potassium phosphate monobasic (2.38 g), potassium phosphate dibasic (5.65 g), magnesium sulfate heptahydrate (1.00 g), cupric sulfate pentahydrate (6.4 mg), iron sulfate heptahydrate (1.1 mg), manganese chloride tetrahydrate (7.9 mg), and zinc sulfate monohydrate (0.94 mg) with yeast extract (6.0 g). When appropriate, sterilized indulin or *Miscanthus* lignocellulose were added to a final concentration of 0.5 % (w/v).

Preparation of indulin and Miscanthus lignocellulose. Indulin was obtained as a gift from MeadWestvaco (Campbell, CA) and prepared according to literature by washing in a Soxhlet apparatus with water at 100 °C for approximately 72 h until extractions were clear [*27*]. The indulin was then acidified by washing with water (pH 2.5), filtered (Whatman No. 1 filter paper), dried at 100 °C, and ground to a powder using a mortar and pestle. Before use in cell culture, the processed indulin was heated at 135 °C for 3 h. *M. giganteus* lignocellulose, ball-milled to 2 mm, was obtained from the Energy Biosciences Institute (Berkeley, CA) and sterilized by autoclave before use.

Extracellular peroxidase activity assay. Peroxidase activity was monitored spectrophotometrically using a modified literature procedure [*28*]. For measuring peroxidase activity in cell culture, a sample (1 mL) was removed and cleared of cells and other debris by centrifugation at 9,800 × *g* for 7 min. The culture supernatant was used directly in the peroxidase assay using a SpectraMax M2 96-well plate reader (Molecular Devices; Sunnyvale, CA). Assays contained culture supernatant (100 μL), 2,4-dichlorophenol (3 mM), 4-aminoantipyrene (164 μM), and H_2O_2 (4 mM) in 50 mM potassium phosphate buffer, pH 7.0 in a total assay volume of 200 μL. Assays were initiated by the addition of H_2O_2 and monitored for an increase in absorbance at 510 nm at 25 °C. The same assay conditions were used for purified protein (10-100 μg) using a DU-800 spectrophotometer. Peroxidase activity was calculated using an extinction coefficient of 18,500 $M^{-1}cm^{-1}$ for the 4-aminoantipyrene/2,4-dichlorophenol adduct [*29*].

Precipitation and collection of APPL. After 14 d of growth in the presence of either indulin or *Miscanthus* lignocellulose, the supernatant of 10 mL of culture was separated from the

cells and the lignin source by centrifugation (9,800 × *g*) and then filtered through a 0.2 μm filter. The filtrate was then acidified with 12 M hydrochloric acid (100 μL) vortexed and allowed to precipitate for 1 hour. The precipitate was collected by centrifugation (9,800 × *g*), dried at 90 °C for 12 hours, and massed for comparison.

Isolation of *A.* sp. 75iv2 genomic DNA. *A.* sp. 75iv2 genomic DNA was isolated using a standard salting-out protocol [*30*]. *A.* sp. 75iv2 was grown in YEME (50 mL) for 40 h and collected by centifugation at 9,800 × *g* for 7 min. The cell pellet was resuspended in 5 mL of SET buffer (75 mM NaCl, 25 mM EDTA, 20 mM Tris-HCl pH 7.5) and incubated at 37 °C for 1 hour after addition of lysozyme (100 μL, 50 mg/mL in ddH_2O). Next, proteinase K (140 μL, 20 mg/mL solution in ddH_2O) and 10% SDS (600 μL) were added and incubated at 55 °C for 2 h. Sodium chloride (2 mL, 5.0 M) was then added to this solution. After cooling the sample to 37 °C, chloroform (5 mL) was added, and the solution was incubated on rotary mixer at room temperature for 30 min. Next, the sample was centrifuged at 4,500 × *g* for 15 min. The aqueous top layer was removed and the genomic DNA was precipitated by addition of isopropanol (0.6 V) and spooled onto a flame-sterilized Pasteur pipet. The spooled genomic DNA was washed in 70% ethanol, air dried, and resuspended in 10 mM Tris-HCl, 1 mM EDTA, pH 8.0 (1.0 mL). As a quality assurance, portions of the 16s rRNA gene were amplified with Pt Taq HF DNA polymerase and the EUB R933/EUB R1387 [*31*] and U1 F/U1 R [*32*] primer sets and inserted into pCR2.1-TOPO using TA cloning. 16 clones were sequenced, and all were found to match with the *Amycolatopsis* sp. 75iv2 16s rRNA gene sequence from the NCBI database.

Preparation of a paired-end genomic library for Illumina sequencing. The genomic library for Illumina sequencing was prepared using the Illumina Genomic DNA Sample Prep Kit (Illumina; Hayward, CA) with some modifications to the manufacturer specifications. Genomic DNA (25 μg) was nebulized with Ar at 35 psi for 2 min in nebulizing buffer (700 μL). The fragmented DNA was end-polished before ligation to the Paired-End Adapter Oligo Mix and separation on an agarose gel (2%). Gel bands corresponding to fragment sizes of 250-300 bp, 300-350 bp, and 350-400 bp according to the 100 bp ladder (Fermentas) were excised and used as a template (1.5 μL) for amplification with Pt Taq HF DNA polymerase with the P1 and P2 primers (14 cycles, 200 μL reaction) after gel purification. The libraries were then concentrated using the Qiagen PCR Purification Kit and eluted in Buffer EB (30 μL). The libraries were analyzed on an Agilent 2100 BioAnalyzer (Santa Clara, CA) using a DNA 1000 Series II chip (Agilent) before sample submission to (actual average sizes, 315, 390, and 460 bp; *Figure 2.2*).

Paired-end assembly of short reads. Paired-end 84-base reads were obtained from each of the three libraries on an Illumina Genome Analyzer (Hayward, CA) at the UC Davis Genome Center using the Cluster Generation Kit v2 (Illumina) and Sequencing Kit v3 (Illumina). Libraries were titrated by PCR against a known standard before sequencing. Data processing was performed on a Linux PC with dual 64 bit 2.66 GHz quad-core processors and 32 Gb of RAM. Prior to assembly, the reads were trimmed based on the average Illumina quality score for bases in a 9-base sliding window. The reads were truncated to the base where the average quality score in that window dropped below 20. Reads were trimmed to 20 bases if it would have been truncated to <20 bases based on this criterion. The short reads were assembled using VELVET, a program developed for the *de novo* assembly of short-read sequencing information [*33*]. An initial assembly was performed using the VelvetOptimizer program using default settings. Next, BWA [*34*] was used to map the short reads to this assembly to accurately determine the average length and standard deviation of the inserts for use in future assemblies. VelvetOptimizer was

then run again with a defined insert length and standard deviation so that the program would perform a complete optimization on each hash length from 39 to 79. The assembly with the largest N50 and maximum scaffold length was chosen for further assembly. Contigs that were both <100 bases and not included in a scaffold were assumed to be misassembled and therefore discarded at this time. Next, Image [*35*] was used to close gaps in the scaffolds generated by Velvet after dissolving them into their constituent contigs. Image was run iteratively on these contigs, one read library at a time. Starting with the library with the smallest insert length, Image was run using hash lengths of 71, 61, 51, 41 and 31. Image was run at each hash length until no further gaps could be closed, generating the final genome assembly.

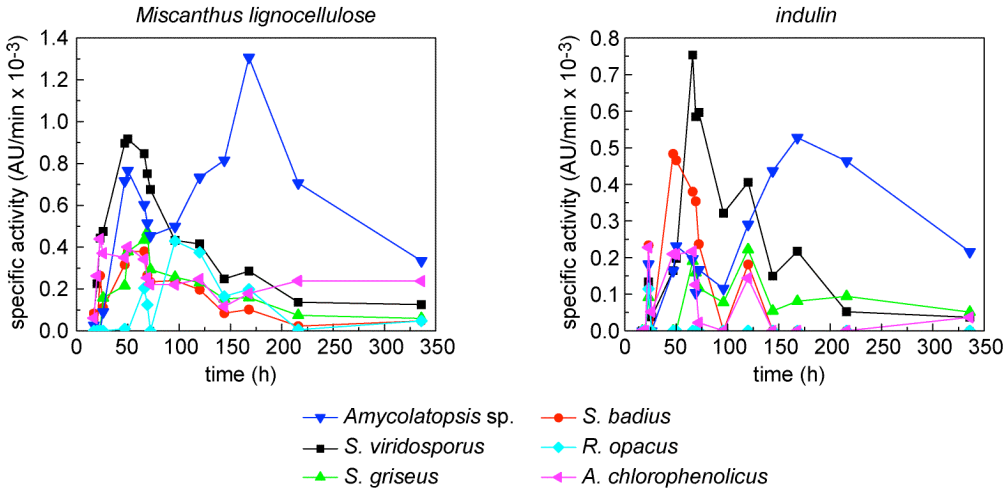
Genome annotation. Putative open reading frames (ORFs) were predicted using Glimmer3 [*36*] trained on sequences from *S. avermitilis, S. griseus, S. scabiei* and *Amycolatopsis mediterranei*. Putative functions were assigned to the ORFs using BLASTp [*37*] and Pfam [*38*]. Reported E-values are based on search of the NCBI microbial protein sequence database. Structural RNAs were identified using RNAmmer [*39*] and tRNAs were identified using tRNA-Scan [*40*]. Putative secreted proteins were identified using SignalP 3.0 using both SignalP-NN (5) and SignalP-HMM outputs [*41, 42*]. Proteins were classified as follows: (a) Yes, at least 4 SignalP–NN outputs above threshold and, (b) Probably, 2-3 SignalP-NN above threshold and identification by SignalP-HMM as secreted, (c) Possibly, 2-3 SignalP-NN outputs above the threshold with a below threshold cut-off secretion value for SignalP–HMM OR less than 2 SignalP-NN outputs above threshold and identification by SignalP-HMM as secreted, (d) Unlikely, 1 SignalP-NN output above threshold with a below threshold cut-off secretion value for SignalP-HMM, (e) No, no SignalP–NN and SignalP-HMM outputs above threshold.

2.3. Results and discussion

Identification of bacterium with most active lignin response. To identify bacteria with robust lignin reactivity, we selected several species of soil bacteria that have been reported to exhibit lignin-degrading or lignolytic-like ability – *Amycolatopsis* sp. 75iv2 (formerly known as *Streptomyces setonii* and *Streptomyces griseus* 75vi2), *Streptomyces viridosporus*, *Streptomyces griseus*, *Streptomyces badius*, *Arthrobacter cholorophenolicus*, and *Rhodococcus opacus* – for culture over 14 days in the presence of either a lignocellulosic carbon source from a perennial grass, *Miscanthus giganteus*, or indulin AT, a purified form of kraft pine lignin produced from the industry of paper pulping wherein acid hydrolysis has been performed to reduce the residual hemicelluloses, salts and nitrogen content [*43*].

As phenol oxidation is a key reaction in the model for cleavage of lignin crosslinks, we tracked the peroxidase-mediated phenol oxidation capacity in the extracellular protein fraction as well as the production of acid-precipitable polymeric lignin (APPL), the bacteria-specific lignin degradation product. From these studies, we determined that *A.* sp. 75iv2 appeared to be the most active strain by both standards (*Figure 2.1*) and consequently turned our attention to developing a rapid lab-scale pipeline for identifying candidate ligninase and accessory enzymes from the unsequenced microbial host using a combination of next-generation sequencing and functional proteomics. *A.* sp. 75iv2 is a gram-positive, soil- dwelling, filamentous actinomycete of the order Actinomycetales. The organism was originally determined to modify and partially break down lignocellulose by Antai *et al.* wherein they showed lignin and carbohydrate loss from ^{14}C-labeled lignocellulose as $^{14}CO_2$ over 12 weeks of incubation [*44*].

A



B

bacterium	**Miscanthus** *(mg per 10 mL)*	**indulin** *(mg per 10 mL)*
A. sp. 75iv2	2.6	3.6
S. viridosporus	0.6	0.7
S. griseus	1.6	1.9
S. badius	1.3	1.4
R. opacus	1.5	1.1
A. chlorophenolicus	0.5	nd

Figure 2.1. *Characterization of lignin degradation in several species of soil bacteria using either Miscanthus giganteus lignocellulose or indulin AT (kraft lignin) as a substrate. (A) Monitoring extracellular peroxidase activity using the 4-aminoantipyrine assay. (B) Mass of acid-precipitable polymeric lignin (APPL) produced by each culture after 14 d.*

Assembly of genome sequence for *A.* sp. 75iv2. Using massively-parallel sequencing based on the Illumina platform, we assembled an inexpensive *de novo* genome sequence for *A.* sp. 75iv2 from paired-end 84-base reads (*Figure 2.2*, *Table 2.1*). Assembly of the short reads using Velvet [*33*] and Image [*34*] produced a genome that spans 168 contigs with an N_{50} of 209 kB and whose size (8.4 Mb) falls within the estimated range for closely-related actinomycetes that have been sequenced, *A. mediterranei*, *S. coelicolor*, *S. griesus*, and *S. avermitilis* (8-10 MB) [*45-49*]. (*Table 2.2*) Though there was initial concern with respect to the difference in genome size between *A.* sp. 75iv2 (8.45 Mb) and its nearest sequenced phylogenetic neighbor, *A. mediterranei* (10.23 Mb), a closed genome sequence for *A.* sp. 75iv2 was subsequently published, confirming the size of the genome to be 8,442,518 bp with a G+C content of 71.9% [*50*], which agreed with the size and G+C content found here (71.8%). Putative open reading frames (ORFs), structural RNAs, and tRNAs were predicted using Glimmer3 [*36*], RNAmmer [*39*], and tRNA-Scan [*40*], respectively. This identified a total of 8,326 open-reading frames (ORFs) encoding 8,267 candidate protein-encoding sequences; this compares to the closed genome from which the number of protein-encoding sequences was determined to be 8,264 [*50*]. As to be expected, the larger genome size of *A. mediterranei* encodes significantly more ORFs (9,228) than *A.* sp. 75iv2; however, this does not act as a direct comparison as the chromosomal length of *S. avermitilis* is 9.03 Mb, but only 7,583 ORFs are encoded [*46, 49*]. Comparison of the G+C content with other closely-related actinomycetes revealed the 71.9% content to be fairly consistent and only slightly higher than less-related soil-dwelling bacteria, such as *Rhodococcus jostii* RHA1 (67% G+C) [*51*]. Unlike *R. jostii* RHA1 and many *Streptomyces* spp. which have linear chromosomes, *A. mediterranei* and *A.* sp. 75iv2 have circular chromosomal topology [*49*] which is more phylogenetically and taxonomically similar to that of *Saccharopolyspora erythraea* [*52*] and *Nocardia farcinica* [*47*].

Its recent reclassification from *S. setonii* or *S. griseus* 75vi2 into the *Amycolatopsis* genus is consistent with the observed genetic cell-wall markers. As with *A.* mediterranei and *Mycobacteria* spp., the genome of *A.* sp. 75iv2 encodes for six genes that allow for the incorporation of arabinose into the cell wall, but unlike *S. coelicolor*, there is an absence of genes for the incorporation of glycine crosslinkers in the peptidoglycan layer (*Table 2.3*, *Figure 2.2*). However, *A.* sp. 75iv2 differentiates itself from Mycobacterium by lacking the characteristic cell wall component, mycolic acid. One important genetic cluster for the production of β-keto-α-alkyl mycolic acid precursors, is the *fadD-pks-accD* cluster. As can be seen, the organization of the *fadD-pks-accD* cluster in *A.* sp. 75iv2 encodes a ketoreductase activity in the polyketide synthase domain organization not harbored in the PKS biosynthetic machinery of mycolic acid-producing bacteria. Finally, the genome of *A.* sp. 75iv2 encodes a *murE* ligase similar to that of *A. mediterranei* and *S. erythraea*, supporting the incorporation of the amino acid, meso-diaminopimilic acid (meso-DAP) into the cell wall [*49*].

A. sp. 75iv2 shows a number of secondary metabolites totaling 11 biosynthetic clusters (*Table 2.4*); however, the total number appear to be significantly less many other soil-dwelling organisms, including other other *Amycolatopsis* spp. that have been sequenced; A. mediterranei contains 26 biosynthetic gene clusters, including 11 nonribosomal peptide synthetases (NRPSs) and 4 hybrid PKS-NRPS clusters [*49*]. The genome of *R. jostii* RHA1 encodes 24 NRPSs and 7 PKSs [*51*]. The sequenced genome *S. griseus* IFO 13350 reveals the presence of 34 secondary metabolites, including 3 NRPS clusters for the biosynthesis of siderophores [*48*]. Similarly, *S. avermitilis* has been analyzed to have 30 secondary metabolite gene clusters [*46*].

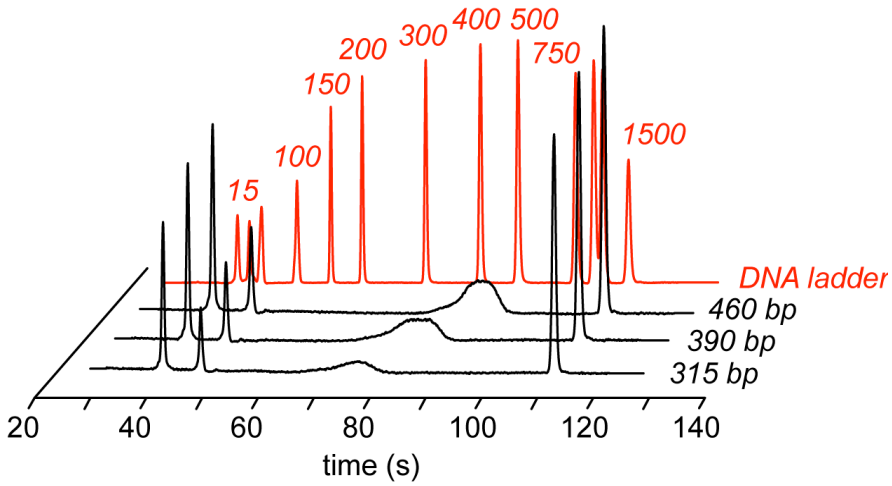


Figure 2.2. *Electropherograms of the paired-end genomic libraries prepared from A. sp. 75iv2 ATCC 39116 utilized for Illumina sequencing.*

***A.* sp. 75iv2 genome assembly statistics**	
Length (bp)	8,458,384
Number of contigs	168
N25	390,262
N50	208,574
N75	96,831
GC content (%)	71.84

Genome annotation	
Total ORFs	8,326
Conserved hypothetical	1,844 (22.1%)
Assigned function	5,741 (69.0%)
Completely hypothetical	741 (0.9%)
Average length (bp)	901
Coding density (%)	88.75

Table 2.1. *Assembly and annotation statistics for the A. sp. 75iv2 ATCC 39116 genome assembled from short reads generated by Illumina sequencing.*

	A. sp. 75iv2	*A. mediterranei*	*S. coelicolor*	*S. griseus*	*S. avermitilis*
Genome assembly size (bp)	8 458 384	10 236 715	8 667 507	8 545 929	9 025 608
Average GC content (%)	71.84	71.3	72.1	72.2	70.7
CDs number	8 326	9 228	7 769	7 138	7 583
Average CDs length (bp)	901	990	991	1 055	1 027
Coding density (%)	88.75	89.3	88.9	88.1	86.3
No. of rRNA (16S-23S-5S) operons	6	4	6	6	6
No. of tRNA genes	53	52	63	66	68

Table 2.2. *Comparison of the de novo A. 75iv2 genome assembled using 84-bp reads with the genomes of its closest sequenced relative (A. mediterranei U32) and the related model streptomycetes (S. coelicolor). Although the size of the genome assembly is smaller than A. mediterranei U32, most of the breaks in the genome fall within areas of repetitive sequence that are difficult to access with Illumina sequencing. (CDS, coding sequence)*

Gene	Predicted identity	E value	Bitscore	Location	Length (bp)	Start	Stop	Strand
Rv1049	Transketolase	0	906	A39116_contig_071:053	2106	56943	59048	+
Rv1017	Phosphribosyl pyrophosphate synthetase	2E-136	479	A39116_contig_157:080	981	90040	91020	+
Rv3806	5-Phosphoribosyl transferase	1E-78	287	A39116_contig_020:006	1014	5140	6153	+
Rv3807	Phospholipid phosphatase	3E-18	85.5	A39116_contig_020:005	522	4596	5117	+
Rv3790	Epimerase	4E-165	575	A39116_contig_020:013	1302	14810	16111	+
Rv3791	Epimerase	7E-87	314	A39116_contig_020:015	759	16108	16866	+

Table 2.3. *Predicted genes found in the A. sp. 75iv2 ATCC 39116 genome assembly involved in incorporation of arabinose into cell wall using Mycobacterium tuberculosis Rv1449 genes for identification by BLASTp [49, 53].*

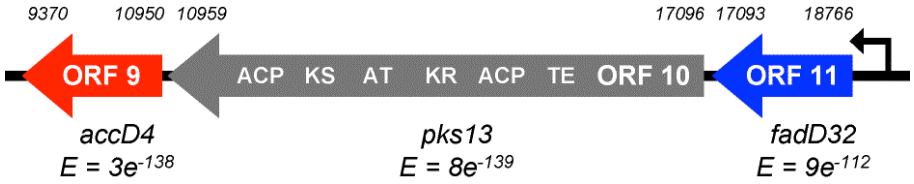


Figure 2.3. *Organization of the fadD-pks-accD cluster from contig 157 in the A. sp. 75iv2 ATCC 39116 genome assembly identified by BLASTp search with orthologs from M. tuberculosis. Analysis of this cluster from mycolic acid-producing bacteria indicates that mycolic acids are made when PKS13 has a domain organization of ACP-KS-AT-ACP-TE. Inclusion of the DH-KR (A. mediterranei U32) and KR (A. sp. 75iv2) domains in the gene encoding PKS13 and the absence of an additional FAS1 homolog (A. mediterranei U32 and A. sp. 75iv2) predicts a lack of mycolic acid production potential [49].*

Predicted identity	Location (contig_#:orf#)	Length (bp)	Start	Stop	Strand
NRPS	A39116_contig_062:292	11454	284001	295454	+
NRPS	A39116_contig_102:030	3861	31672	35532	+
PKS	A39116_contig_135:032	2895	35930	38824	+
PKS	A39116_contig_145:103	246	103378	103133	-
NRPS	A39116_contig_156:072	231	72337	72567	+
PKS	A39116_contig_157:010	6138	17096	10959	-
Chalcone synthase	None				
Terpene synthase	A39116_contig_023:040	1047	50252	49206	-
Putative terpene cyclase	A39116_contig_098:020	2187	21456	23642	+
Terpene synthase	A39116_contig_151:071	978	67912	68889	+
Lantibiotic biosynthesis protein	A39116_contig_062:261	1152	251497	252648	+
Lantibiotic biosynthesis protein	A39116_contig_141:173	1011	180499	181509	+

Table 2.4. *Predicted genes for secondary metabolite production in A. sp. 75iv2 ATCC 39116 identified by Pfam search indicates that A. sp. 75iv2 appears to contain less biosynthetic clusters for canonical secondary metabolites compared to related actinomycetes.*

S. coelicolor has over 20 biosynthetic gene clusters [*45*], and *S. bingchenggensis* has 23 secondary metabolite gene clusters [*54*]. The relative paucity of genes encoding for secondary metabolites in comparison to many other soil dwelling organisms suggests that *A.* sp. 75iv2 contains a fairly limited secondary metabolism for unknown reasons. A purely speculative notion could reside within its notion to utilize complex, relatively inaccessible substrates, leading to a lesser obligation to compete with other organisms for easy substrates to metabolize. Analysis of the genome for secondary metabolite gene clusters using antiSMASH [*55*] agreed with the relative paucity of biosynthetic gene clusters, showing the production of: bacteriocin (1 gene), ectoine (1 gene), terpene (5 gene clusters), 1 lantipeptide (1 gene cluster), products of NRPSs (2 gene clusters, of which 1 NRPS encodes siderophore production), PKS (2 gene clusters; one type I and one type II), and 2 "other." We deciphered one of the "other" (Amy39116DRAFT_2931) genes to potentially encode an NRPS (A39916_contig_102:030); we deciphered the second "other" (Amy39116DRAFT_3421) to potentially encode a PKS (A39116_contig_135:032). AntiSMASH, however, was efficient at predicting 6 new previously unpredicted biosynthetic cluster products (bacteriocin, ectoine, 1 siderophore, 2 terpenes, and polyketide) but did not predict 1 PKS (A39116_contig_145:103) one NRPS (A39116_contig_156:072) and 1 lantibiotic synthetase (A39116_contig_141:173).

A. sp. 75iv2 also maintains a rich collection of genes for biomass deconstruction with many predicted secreted enzymes within the glycosyl hydrolase superfamily as well as a considerable number of clusters for aromatic degradation (*Tables 2.5* and *2.6*, *Figure 2.4*) [*21, 23, 56*]. Regarding the small molecule aromatic degradation pathways found in bacteria, a significant number of studies surrounding *Sphingomonas paucimolis* and *Pseudomonas putida* have revealed that many soil-dwelling bacteria have complex systems for degradation of downstream lignin degradation products and similar environmental toxins [*21-24, 57*]. These pathways entail processing of small dimeric or monomeric aromatic compounds through demethylation, ring oxidation, ring opening, and further downstream enzymatic activity to produce products such as acetaldehyde and pyruvate that can be utilized by the cell. Analysis of the genome of *A.* sp. 75iv2 reveals a considerable number of clusters for aromatic degradation through various *meta*- and *ortho*-ring cleavage pathways (*Table 2.6*, *Figure 2.4*); these include many well-characterized pathways such as the protocatachuate *ortho*-cleavage pathway (*pca* genetic pathway) for the degradation of *p*-hydroxybenzoate, the TOL pathway (*xyl* genes) for the degradation of toluene and catechol, and much of the meta-ring cleavage pathway found in *Sphingomonas paucimobilis* (*lig* genes). Furthermore, the organism also encodes many other clusters for potential aromatic degradation. This abundance of aromatic degradation pathways supports that this organism is capable of metabolizing downstream products of lignin degradation into those used for carbon and energy. Nonetheless, it is suggested that many bacteria are only capable of utilizing only small molecule components of partially- or largely-degraded lignin [*58, 59*], and where these small molecule degradation pathways found in *A.* sp. 75iv2 suggest utilization of lignin components, they do not assure complete lignin reactivity, degradation, and/or utilization.

With regard to lignin reactivity, *A.* sp. 75iv2 possesses several potential oxidative systems based on the presence of predicted secreted laccases, peroxidases, and peroxide-generating enzymes (*Table 2.7*). Laccases have been suggested to take part in phenolic oxidation for partial depolymerization of lignin, and studies with unnatural redox mediators, such as ABTS, have shown certain laccases to even have non-phenolic oxidation activity, suggesting that laccases could take a larger role in lignin degradation; however, the natural redox mediators utilized for

Table 2.5. *Predicted carbohydrate-degrading enzymes from the glycosyl hydrolase and amylase superfamilies identified and organized by Pfam search. (CMB, carbohydrate binding module)*

Predicted identity	Pfam	Organism	E value	Bitscore	Location (contig_#:orf#)	Length (bp)	Start	Stop	Strand	Secreted	CBM
β-Glucosidase	Glyco_hydro_1	*Amycolatopsis mediterranei* U32	8E-178	627	A39116_contig_059:042	1413	40228	38816	-	No	No
β-Glucosidase	Glyco_hydro_1	*Saccharopolyspora erythraea* NRRL 2338	3E-170	602	A39116_contig_020:108	1251	110377	111627	+	No	No
β-Glucosidase	Glyco_hydro_1	*Streptomyces* sp. AA4	0	653	A39116_contig_055:056	1275	61084	59810	-	No	No
Glycoside hydrolase family 2	Glyco_hydro_2	*Deinococcus deserti* VCD115	0	649	A39116_contig_059:105	1818	107977	106160	-	No	No
Glycoside hydrolase family 2	Glyco_hydro_2_N	*Amycolatopsis mediterranei* U32	0	802	A39116_contig_156:036	3324	33563	36886	+	Yes	No
β-Glucosidase	Glyco_hydro_3	*Amycolatopsis mediterranei* U32	0	884	A39116_contig_035:091	1854	87509	85656	-	Yes	No
β-Glucosidase	Glyco_hydro_3	*Arthrobacter chlorophenolicus* A6	6E-177	625	A39116_contig_136:090	2358	81196	78839	-	No	No
β-Glucosidase	Glyco_hydro_3	*Rhodococcus jostii* RHA1	0	792	A39116_contig_153:012	2157	13206	11050	-	No	No
β-Glucosidase-like glycosyl hydrolase	Glyco_hydro_3	*Saccharomonospora viridis* DSM 43017	0	747	A39116_contig_082:245	1764	249844	251607	+	Yes	No
β-Glucosidase-like glycosyl hydrolase	Glyco_hydro_3	*Saccharomonospora viridis* DSM 43017	0	1124	A39116_contig_145:125	2313	128508	130820	+	No	No
β-Glucosidase-like glycosyl hydrolase	Glyco_hydro_3	*Saccharomonospora viridis* DSM 43017	8e-120	434	A39116_contig_062:071	1164	71514	70351	-	Yes	No
Glycosyl hydrolase family protein	Glyco_hydro_3	*Arthrobacter aurescens* TC1	0	720	A39116_contig_132:213	2361	218840	216480	-	No	No
Glycoside hydrolase family 12	Glyco_hydro_12	*Kribbella flavida* DSM 17836	4E-24	115	A39116_contig_007:160	633	151796	151164	-	No	No
Glucoamylase-like glycosyl hydrolase	Glyco_hydro_15	*Amycolatopsis mediterranei* U32	0	872	A39116_contig_007:060	1785	55804	57588	+	No	No
Glucan 1,4-α-glucosidase	Glyco_hydro_15	*Saccharopolyspora erythraea* NRRL 2338	0	767	A39116_contig_145:116	2106	116548	118653	+	No	No
Trehalose 6-phosphatase	Glyco_hydro_15	*Saccharomonospora viridis* DSM 43017	0	1456	A39116_contig_082:139	2532	136403	138934	+	No	No
β-1,3-Glucanase	Glyco_hydro_16	*Streptomyces* sp. SPB7	2E-112	409	A39116_contig_144:140	816	192238	191423	-	Yes	No
Glycoside hydrolase family protein	Glyco_hydro_16	*Mycobacterium vanbaalenii* PYR-1	5E-24	115	A39116_contig_023:123	807	135747	136553	+	Yes	No
Glycoside hydrolase family 16	Glyco_hydro_16	*Coraliomargarita akajimensis* DSM 45221	2E-32	144	A39116_contig_151:012	1149	13635	12487	-	Yes	No
Chitinase	Glyco_hydro_18	*Amycolatopsis mediterranei* U32	2E-174	616	A39116_contig_017:133	1194	138395	137202	-	Yes	No
Chitinase A precursor	Glyco_hydro_18	*Saccharopolyspora erythraea* NRRL 2338	2E-112	409	A39116_contig_060:080	963	81104	80142	-	Yes	No
β-N-Acetylhexosaminidase	Glyco_hydro_20	*Streptomyces* sp. AA4	0	717	A39116_contig_026:018	1569	17147	18715	+	Yes	No
β-N-Acetylhexosaminidase	Glyco_hydro_20	*Streptomyces* sp. AA4	0	638	A39116_contig_098:043	1377	49252	47876	-	Yes	No
Glycoside hydrolase family 25	Glyco_hydro_25	*Streptomyces violaceusniger* TU 4113	4E-46	189	A39116_contig_019:006	714	4843	4130	-	Yes	No
Glycoside hydrolase family 25	Glyco_hydro_25	*Streptomyces* sp. AA4	7E-67	257	A39116_contig_026:003	615	1832	2446	+	No	No

Glycoside hydrolase family 25	Glyco_hydro_25	*Streptomyces violaceusniger* TU 4113	5E-71	271	A39116_contig_057:270	822	268033	268854	+	Yes	No
Putative β-galactosidase	Glyco_hydro_42	*Streptomyces* sp. AA4	0	930	A39116_contig_146:033	1974	46358	44385	-	No	No
Glycoside hydrolase family protein	Glyco_hydro_43	*Kribbella flavida* DSM 17836	4E-81	305	A39116_contig_124:179	1005	173040	174044	+	Yes	No
Glycoside hydrolase family 57	Glyco_hydro_57	*Streptomyces* sp. AA4	0	749	A39116_contig_059:348	1482	354264	355745	+	No	No
Glycosyl hydrolase	Glyco_hydro_65	*Amycolatopsis mediterranei* U32	0	1250	A39116_contig_132:037	2373	32886	35258	+	No	No
Glycoside hydrolase family protein	Glyco_hydro_76	*Actinosynnema mirum* DSM 43827	2E-110	403	A39116_contig_166:006	990	4706	5695	+	No	No
4-α-Glucanotransferase	Glyco_hydro_77	*Streptomyces* sp. AA4	0	907	A39116_contig_023:004	1941	8540	6600	-	No	No
α-1,2-Mannosidase	Glyco_hydro_92	*Streptomyces* sp. ACTE	0	1045	A39116_contig_141:106	2385	106858	104474	-	Yes	No
Putative α-1,2-mannosidase	Glyco_hydro_92	*Amycolatopsis mediterranei* U32	0	1289	A39116_contig_035:119	3096	114976	111881	-	Unlikely	No
Glycosyl hydrolase BNR-repeat protein	BNR	*Hyphomicrobium denitrificans* ATCC 51888	7e-178	626	A39116_contig_124:111	1116	106543	105428	-	Unlikely	No
Glycosyl hydrolase (secreted)	-	*Saccharopolyspora erythraea* NRRL 2338	8e-128	460	A39116_contig_060:081	1149	82266	81118	-	Yes	No
Secreted glycosyl hydrolase	CBM_6/PKD	*Saccharomonospora viridis* DSM 43017	0	692	A39116_contig_023:015	3279	23013	19735	-	Yes	Yes
Malto-oligosyltrehalose trehalohydrolase	Alpha-amylase	*Streptomyces* sp. AA4	0	859	A39116_contig_023:001	1665	2118	454	-	No	No
Maltooligosyl trehalose synthase	Alpha-amylase	*Streptomyces* sp. AA4	0	1123	A39116_contig_023:002	2271	4427	2157	-	No	No
Glycogen debranching enzyme (glgX)	Alpha-amylase	*Streptomyces* sp. AA4	0	1222	A39116_contig_023:003	2118	6541	4424	-	No	Yes
Maltose α-D-glucosyltransferase	Alpha-amylase	*Actinosynnema mirum* DSM 43827	0	1126	A39116_contig_102:079	1812	91648	89837	-	No	No
Glycogen branching enzyme	Alpha-amylase	*Streptomyces* sp. AA4	0	1250	A39116_contig_102:077	2187	88447	86261	-	No	Yes
Glycogen debranching enzyme (glgX)	Alpha-amylase	*Streptomyces* sp. AA4	0	1168	A39116_contig_102:082	2136	96466	94331	-	No	Yes
α-Amylase catalytic region	Alpha-amylase	*Streptomyces* sp. AA4	0	1064	A39116_contig_102:080	1914	93664	91751	-	No	No
α-Amylase catalytic region	Alpha-amylase	*Kribbella flavida* DSM 1783	8e-159	563	A39116_contig_023:005	1266	8598	9863	+	No	No
α-Glucosidase	Alpha-amylase	*Streptomyces viridochromogenes* DSM 4073	0	654	A39116_contig_151:163	1527	153683	155209	+	No	No

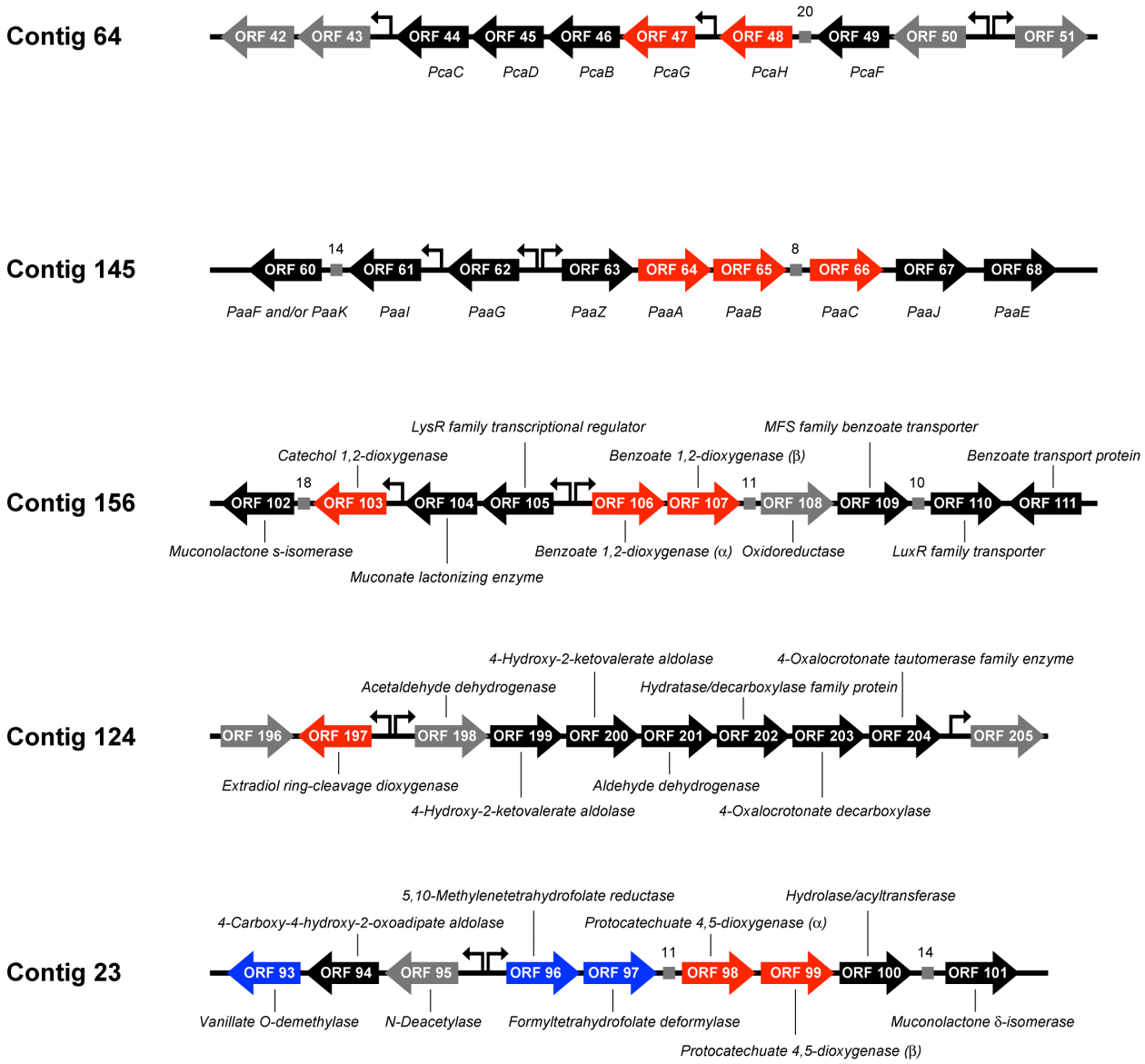


Figure 2.4. *Organization of clustered genes with predicted aromatic degradation function on A39116_contig_064 (protocatechuate pathway), S39116_contig_145 (phenylacetate pathway), A39116_contig_156, A39116_contig_124, and A39116_contig_23. Not shown from Table S6DE: A39116_contig_19, A39116_contig_121, A39116_contig_124, and A39116_contig_132. Promoters are indicated where the spacing between genes is >30 bp or at strand switchpoints. Operons are indicated when stop and start sites are <3 bp apart. Gray boxes with distances are shown when the distance between genes is between 3-30 bp. (Red, aromatic ring-cleaving enzymes; black, potential aromatic degradation function based on known pathways or superfamily; gray, unknown function with respect to aromatic degradation.)*

Table 2.6. *Potential aromatic degradation pathways in Amycolatopsis sp. 75iv2 ATCC 39116 identified by BLASTp search based on canonical ortho- and meta-cleavage pathways as well as Pfam search for ring-cleaving oxygenases and oxidative/tetrahydrofolate-dependent demethylases.*

(A). Protocatechuate (Pca) ortho cleavage pathway candidates identified from BLASTp search with sequences from Pseudomonas putida ATCC 33015 [23].

Gene	General Function	Pfam	E value	Bitscore	Location (contig_#:orf#)	Length (bp)	Start	Stop	Strand	Secreted?
pobA	*p*-Hydroxybenzoate hydroxylase		9E-116	411	A39116_contig_060:135	1182	129352	130533	+	Possibly
			1E-055	211	A39116_contig_023:062	1179	72638	73816	+	No
pcaG	Protocatechuate 3,4-dioxygenase (α)	Dioxygenase_C	9E-031	127	A39116_contig_064:047	546	52714	52169	-	No
pcaH	Protocatechuate 3,4-dioxygenase (β)	PCDO_beta_N	1E-057	217	A39116_contig_064:048	744	53531	52788	-	No
pcaB	Cycloisomerase		2E-064	240	A39116_contig_064:046	1326	52176	50851	-	Unlikely
pcaC	Decarboxylase		2E-035	141	A39116_contig_064:044	402	50096	49695	-	No
pcaD	β-Ketoadipate enolactone hydrolase		1E-051	197	A39116_contig_064:045	762	50854	50093	-	No
			2E-022	100	A39116_contig_027:040	789	38768	39556	+	No
			8E-017	81.6	A39116_contig_059:180	846	179459	178614	-	No
pcaIJ	β-Ketoadipate succinyl-coA transferase		1E-056	213	A39116_contig_146:078	642	89240	89881	+	No
pcaF	β-Ketoadipate CoA thiolase		5E-090	325	A39116_contig_064:049	1176	54726	53551	-	No
			2E-074	273	A39116_contig_007:193	1212	181905	183116	+	Unlikely
			4E-061	229	A39116_contig_076:036	1227	36111	34885	-	No
			4E-060	226	A39116_contig_008:011	1152	11213	12364	+	No
			2E-059	224	A39116_contig_128:048	1221	47357	48577	+	No
			3E-058	220	A39116_contig_165:050	1209	45465	44257	-	Unlikely
			5E-058	219	A39116_contig_144:104	1134	154477	153344	-	No
			1E-055	211	A39116_contig_151:250	1212	244946	243735	-	No
			4E-047	183	A39116_contig_132:238	1173	247962	246790	-	No
			1E-046	181	A39116_contig_017:054	1164	56747	57910	+	No
			1E-045	178	A39116_contig_151:040	1188	39425	38238	-	No
			2E-043	171	A39116_contig_023:035	1107	44815	45921	+	No
			9E-034	139	A39116_contig_057:197	1272	205648	204377	-	No

(B) TOL meta cleavage pathway candidates identified from BLASTp search with sequences Pseudomonas putida mt-2 ATCC 33015.

Gene	General Function	Pfam	E value	Bitscore	Location (contig_#:orf#)	Length (bp)	Start	Stop	Strand	Secreted
xylA	Xylene oxygenase		2E-041	164	A39116_contig_124:014	1041	13726	14766	+	No
			2E-040	160	A39116_contig_007:126	1005	115183	116187	+	No
			4E-038	152	A39116_contig_156:108	2691	109276	111966	+	No
			3E-037	150	A39116_contig_134:009	1047	8013	6967	-	Unlikely
			2E-023	104	A39116_contig_060:018	1173	18923	20095	+	No
			1E-016	82	A39116_contig_017:073	1023	76833	77855	+	Unlikely
			3E-016	80.9	A39116_contig_121:044	2196	48818	51013	+	No
xylB	Benzyl alcohol dehydrogenase		9E-032	132	A39116_contig_124:261	999	251093	250095	-	No
			3E-037	150	A39116_contig_017:168	1116	172228	171113	-	No
xylC	Benzaldehyde dehydrogenase		2E-048	188	A39116_contig_156:127	1449	130540	131988	+	No
			2E-057	218	A39116_contig_124:262	1500	252751	251252	-	No
			1E-049	191	A39116_contig_124:201	1431	193152	194582	+	No
			3E-030	127	A39116_contig_124:250	684	237922	237239	-	No
			9E-061	229	A39116_contig_017:122	1611	125321	126931	+	Possibly
			2E-049	191	A39116_contig_017:199	1515	202234	200720	-	No
xylD	Benzoate oxygenase		3E-019	87.8	A39116_contig_156:106	1350	107413	108762	+	No
xylL	Carboxylate dehydrogenase		4E-048	184	A39116_contig_156:108	2691	109276	111966	+	No
			2E-019	88.6	A39116_contig_092:038	753	35887	36639	+	Yes
			3E-016	78.2	A39116_contig_028:003	732	2486	1755	-	Unlikely
xylE	Dioxygenase	Glyoxalase	1E-026	115	A39116_contig_124:212	1023	203956	202934	-	No
		Glyoxalase	2E-025	110	A39116_contig_156:155	966	158240	157275	-	No
		Glyoxalase	1E-024	108	A39116_contig_040:048	1023	44885	45907	+	No
xylG	Dehydrogenase		1E-055	211	A39116_contig_156:127	1449	130540	131988	+	No
			6E-091	329	A39116_contig_124:201	1431	193152	194582	+	No
			4E-077	283	A39116_contig_124:262	1500	252751	251252	-	No
			6E-028	119	A39116_contig_124:250	684	237922	237239	-	No
			6E-062	233	A39116_contig_017:199	1515	202234	200720	-	No
			4E-054	206	A39116_contig_017:122	1611	125321	126931	+	Possibly
xylH	Isomerase				None					
xylF	Hydrolase		2E-034	140	A39116_contig_017:067	846	70787	69942	-	Unlikely
			3E-028	119	A39116_contig_132:265	810	276878	276069	-	No
			4E-028	119	A39116_contig_042:009	816	12210	11395	-	No
			5E-028	119	A39116_contig_121:046	798	53308	52511	-	No

(C) Candidate orthologs of Pseudomonas paucimobilis (formerly Sphingomonas paucimobilis SYK-6 ATCC 29837) aromatic degradation genes [57] identified by BLASTp.

Gene	General function	Pfam	E value	Bitscore	Location (contig_#:orf#)	Length (bp)	Start	Stop	Strand	Secreted
desA	Syringate demethylase	GCV_T	4E-112	399	A39116_contig_023:093	1404	106015	104612	-	No
ligA	Protocatechuate 4,5-dioxygenase (α)	LigA	4E-022	97.8	A39116_contig_023:098	345	109305	109649	+	No
ligB	Protocatechuate 4,5-dioxygenase (β)	LigB	9E-088	317	A39116_contig_023:099	849	109642	110490	+	No
		LigB	5E-018	85.9	A39116_contig_156:047	861	50066	50926	+	Unlikely
ligD	Cα-dehydrogenase		4E-016	80.1	A39116_contig_151:264	834	256237	255404	-	No
ligE	β-Etherase				None					
ligF	β-Etherase				None					
ligM	Vanillate demethylase	GCV_T	0	656	A39116_contig_023:093	1404	106015	104612	-	No
ligW	5-Carboxyvanillate decarboxylase		2E-043	171	A39116_contig_125:002	1047	3112	2066	-	Unlikely
			9E-038	151	A39116_contig_040:207	978	192474	191497	-	No
			1E-020	94.7	A39116_contig_019:084	948	89536	90483	+	No
			6E-019	89.4	A39116_contig_040:433	1071	414855	413785	-	No
			8E-016	79	A39116_contig_019:071	1005	76085	77089	+	No
ligX	DDVA *O* -Demethylase (biphenyl)				None					
ligY	OH-DDVA meta-cleavage hydrolase		2E-048	187	A39116_contig_019:112	1026	120015	121040	+	No
ligZ	OH-DDVA ring opening (biphenyl)				None					

(D) Additional candidate aromatic ring-opening genes identified using Pfam search of families comprising known ring-cleaving oxygenases. Candidates are prioritized based on clustering (also included) with other genes potentially involved in aromatic degradation processes

Predicted identity	Pfam	Organism	E value	Bitscore	Location (contig_#:orf#)	Length (bp)	Start	Stop	Strand
Contig 145 (Putative phenylacetate degradation pathway: Reannoted by BLASTp with MetaCyc naming convention)									
Phenylacetate degradation protein (PaaK and/or PaaF)	AMP-binding	*Saccharomonospora viridis* DSM 43017	0	770	A39116_contig_145:060	1347	57434	56088	-
Phenylacetate degradation protein (PaaI)	4HBT	*Saccharomonospora viridis* DSM 43017	6e-54	212	A39116_contig_145:061	405	57852	57448	-
Enoyl-CoA hydratase/isomerase (PaaG)	ECH	*Thermomonospora curvata* DSM 43183	2e-66	255	A39116_contig_145:062	741	58661	57921	-
Phenylacetate degradation protein (PaaZ)	Aldedh/MaoC_ dehydratas	*Saccharomonospora viridis* DSM 43017	0	1149	A39116_contig_145:063	2046	58774	60819	+
Phenylacetyl-CoA oxygenase subunit (PaaA)	PaaA_PaaC	*Streptomyces* sp. AA4	7e-160	566	A39116_contig_145:064	945	60816	61760	+

Phenylacetyl-CoA oxygenase subunit (PaaB)	PaaB	*Streptomyces* sp. AA4	4e-45	183	A39116_contig_145:065	288	61757	62044	+
Phenylacetyl-CoA oxygenase subunit (PaaC)	PaaA_PaaC	*Saccharomonospora viridis* DSM 43017	8e-131	469	A39116_contig_145:066	885	62052	62936	+
Phenylacetate degradation protein (PaaJ)	DUF59	*Streptomyces* sp. AA4	8e-67	255	A39116_contig_145:067	489	62933	63421	+
Phenylacetyl-CoA oxygenase subunit (PaaE)	FAD_binding_6/Fer2	*Saccharomonospora viridis* DSM 43017	2e-162	575	A39116_contig_145:068	1083	63423	64505	+
Contig 156 (Potential benzoate degradation pathway)									
Muconolactone δ-isomerase	MIase	*Saccharomonospora viridis* DSM 43017	4e-39	163	A39116_contig_156:102	282	104284	104003	-
Putative catechol 1, 2-dioxygenase	Dioxygenase_N	*Nocardia farcinica* IFM 10152	3e-131	471	A39116_contig_156:103	846	105151	104306	-
Muconate lactonizing enzyme	MR_MLE	*Saccharopolyspora erythraea* NRRL 2338	0	686	A39116_contig_156:104	1104	106291	105188	-
LysR family transcriptional regulator	LysR_substrate	*Saccharopolyspora erythraea* NRRL 2338	6e-141	503	A39116_contig_156:105	903	107190	106288	-
Benzoate 1, 2-dioxygenase (α)	Ring_hydroxyl_A	*Streptomyces* sp. AA4	0	834	A39116_contig_156:106	1350	107413	108762	+
Benzoate 1, 2-dioxygenase (β)	Ring_hydroxyl_B	*Saccharopolyspora erythraea* NRRL 2338	1e-86	321	A39116_contig_156:107	507	108759	109265	+
Oxidoreductase FAD-binding region	FAD_binding_6	*Saccharopolyspora erythraea* NRRL 2338	0	1340	A39116_contig_156:108	2691	109276	111966	+
MFS family benzoate membrane transporter	Sugar_tr	*Rhodococcus jostii* RHA1	1e-152	543	A39116_contig_156:109	1275	112061	113335	+
LuxR family transcriptional regulator	GerE	*Saccharopolyspora erythraea* NRRL 2338	0	1035	A39116_contig_156:110	2589	113345	115933	+
Benzoate membrane transporter (BenE)	BenE	*Rhodococcus jostii* RHA1	3e-148	528	A39116_contig_156:111	1236	117163	115928	-
Contig 124									
Extradiol ring-cleavage dioxygenase (β)	LigB	*Frankia symbiont* of Datisca glomerata	1e-96	356	A39116_contig_124:197	1071	191325	190255	-
Acetaldehyde dehydrogenase	Semialdhyde_dh	*Streptomyces hygroscopicus* ATCC 53653	4e-83	310	A39116_contig_124:198	696	191367	192062	+
4-Hydroxy-2-ketovalerate aldolase	HMGL-like	*Streptomyces hygroscopicus* ATCC 53653	2e-112	408	A39116_contig_124:199	888	192059	192946	+
4-Hydroxy-2-ketovalerate aldolase	DmpG_comm	*Streptomyces hygroscopicus* ATCC 53653	2e-26	120	A39116_contig_124:200	225	192931	193155	+
Aldehyde dehydrogenase	Aldedh	*Mycobacterium intracellulare* ATCC 13950	0	699	A39116_contig_124:201	1431	193152	194582	+
Hydratase/decarboxylase family protein	FAA_hydrolase	*Streptomyces hygroscopicus* ATCC 53653	7e-109	396	A39116_contig_124:202	780	194579	195358	+
4-Oxalocrotonate decarboxylase	FAA_hydrolase	*Streptomyces hygroscopicus* ATCC 53653	4e-103	377	A39116_contig_124:203	762	195355	196116	+
4-Oxalocrotonate tautomerase	Tautomerase	*Alicyclobacillus acidocaldarius* DSM 446	3e-11	70.9	A39116_contig_124:204	237	196107	196343	+
Contig 132									
Oxidoreductase	Flavin_Reduct	*Rhodococcus jostii* RHA1	7e-44	179	A39116_contig_132:264	498	275936	275439	-
2-Hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase	Abhydrolase_1	*Rhodococcus jostii* RHA1	5e-42	174	A39116_contig_132:265	810	276878	276069	-
Dienelactone hydrolase	DLH	*Rhodopseudomonas palustris* BisB5	4e-26	120	A39116_contig_132:266	483	277393	276911	-
3-(2,3-Dihydroxyphenyl)propionate dioxygenase	LigB	*Mycobacterium* sp. JLS	7e-86	320	A39116_contig_132:267	954	278622	277669	-

(E) Additional candidate aromatic demethylase genes identified using Pfam search of families comprising oxidative and THF-dependent enzymes. Candidates are prioritized based on clustering (also included) with other genes potentially involved in aromatic degradation processes.

Predicted identity	Pfam	Organism	E value	Bitscore	Location (contig_#:orf#)	Length (bp)	Start	Stop	Strand
Contig 124									
Carboxymuconolactone decarboxylase	CMD	*Streptosporangium roseum* DSM 43021	1e-45	185	A39116_contig_124:129	558	123260	122703	-
HxlR family transcriptional regulator	HxlR	*Streptosporangium roseum* DSM 43021	7e-27	122	A39116_contig_124:130	273	123403	123675	+
Putative vanillate *O*-demethylase oxygenase	Rieske	*Rhodococcus opacus* B4	6E-148	528	A39116_contig_124:131	1047	123763	124809	+
Hypothetical protein SeryN2_22237	-	*Saccharopolyspora erythraea* NRRL 2338	5e-15	83.2	A39116_contig_124:132	219	124806	125024	+
Phthalate 4,5-dioxygenase reductase subunit	FAD_binding_6/Fer2	*Saccharopolyspora erythraea* NRRL 2338	4e-124	447	A39116_contig_124:133	963	125021	125983	+
Contig 19									
Vanillate *O*-demethylase oxidoreductase	FAD_binding_6/Fer2	*Rhodococcus jostii* RHA1	6E-087	325	A39116_contig_019:043	1056	47649	46594	-
Putative short chain dehydrogenase	adh_short	*Streptomyces* sp. AA4	6e-102	373	A39116_contig_019:044	726	48403	47678	-
Putative ring-cleavage dioxygenase (β)	Ring_hydroxyl_B	*Nocardia farcinica* IFM 10152	3e-64	246	A39116_contig_019:045	501	48900	48400	-
Putative ring-cleavage dioxygenase (α)	Ring_hydroxyl_A	*Streptomyces* sp. AA4	0	783	A39116_contig_019:046	1278	50170	48893	-
Hypothetical protein StAA4_36196	-	*Streptomyces* sp. AA4	2e-55	217	A39116_contig_019:047	381	50563	50183	-
Aromatic ring hydroxylase	Acyl-CoA_dh_2	*Rhodococcus jostii* RHA1	2E-121	439	A39116_contig_019:048	1134	51726	50593	-
Non-clustered demethylases									
Vanillate *O*-demethylase oxidoreductase	NAD_binding_1/Fer2	*Rhodococcus jostii* RHA1	5e-95	351	A39116_contig_007:143	930	128026	128955	+

Table 2.7. Predicted oxidative systems in Amycolatopsis sp. 75iv2 ATCC 39116 potentially related to lignolytic function, including peroxide generation, peroxidases, and multicopper oxidases.

A. Peroxide generation

Predicted identity	Pfam	Organism	E value	Bitscore	Location (contig_#:orf#)	Length (bp)	Start	Stop	Strand	Secreted
Lysyl oxidase-like protein	Lysyl_oxidase	*Amycolatopsis mediterranei* U32	3E-013	80.5	A39116_contig_020:079	1197	84226	85422	+	Yes
Lysyl oxidase-like protein	Lysyl_oxidase	*Amycolatopsis mediterranei* U32	7E-012	75.1	A39116_contig_157:036	741	47206	46466	-	Yes
Amine oxidase	Amino_oxidase	*Tsukamurella paurometabola* DSM 20162	0	643	A39116_contig_040:400	1371	383880	385250	+	Possibly
Amine oxidase	Amino_oxidase	*Mycobacterium* sp. KMS	3E-175	619	A39116_contig_124:187	1503	178918	180420	+	No
Primary-amine oxidase	Cu_amine_oxid	*Geodermatophilus obscurus* DSM 43160	0	980	A39116_contig_019:063	1875	66278	68152	+	No
Amine oxidase, flavin-containing	Amino_oxidase	*Streptosporangium roseum* DSM 43021	8E-087	325	A39116_contig_040:337	1239	319854	321092	+	No
L-Amino-acid oxidase	Amino_oxidase	*Saccharopolyspora erythraea* NRRL 2338	3E-124	449	A39116_contig_019:057	1293	59026	60318	+	No
Tyramine oxidase	Cu_amine_oxid	*Streptomyces roseosporus* NRRL 15998	0	858	A39116_contig_102:049	1917	57977	59893	+	No
Galactose oxidase	Glyoxal_oxid_N	*Frankia symbiont* of *Datisca glomerata*	6E-082	310	A39116_contig_007:209	2325	197922	200246	+	No

B. Peroxidases

Predicted identity	Pfam	Organism	E value	Bitscore	Location (contig_#:orf#)	Length (bp)	Start	Stop	Strand	Secreted
Dyp-type peroxidase family protein	Dyp_perox	*Streptomyces* sp. AA4	2E-170	602	A39116_contig_157:054	1248	64076	65323	+	Yes
Dyp-type peroxidase family protein	-	*Ktedonobacter racemifer* DSM 44963	3E-69	266	A39116_contig_024:013	1344	15841	17184	+	No
Dyp-type peroxidase family protein	-	*Cyanothece* sp. PCC 7822	2E-68	265	A39116_contig_064:103	1395	107476	108870	+	No
Catalase/peroxidase HPI (Amyco1)	peroxidase	*Micromonospora* sp. ATCC 39149	0	1339	A39116_contig_019:106	2277	115024	112748	-	No
Heme peroxidase	An_peroxidase	*Streptomyces ghanaensis* ATCC 14672	0	864	A39116_contig_024:010	1821	12319	14139	+	No
Peroxidase	-	*Cytophaga hutchinsonii* ATCC 33406	6E-95	351	A39116_contig_024:013	1344	15841	17184	+	No
Heme peroxidase	An_peroxidase	*Polaromonas* sp. JS666	1E-83	313	A39116_contig_145:037	933	35395	34463	-	No
Iron-dependent peroxidase-like protein	-	*Rhizobium leguminosarum* bv. trifolii WSM1325	8E-144	514	A39116_contig_064:103	1395	107476	108870	+	No
Alkylhydroperoxidase AhpD core	CMD	*Streptomyces* sp. AA4	1E-42	175	A39116_contig_151:037	474	35881	35408	-	Unlikely
Alkylhydroperoxidase	CMD	*Streptomyces viridochromogenes* DSM 40736	4E-52	206	A39116_contig_023:197	447	205777	206223	+	No
Peroxidase (*bpoC*)	Abhydrolase_1	*Streptomyces* sp. AA4	1E-115	419	A39116_contig_062:290	804	282935	283738	+	No
Chloroperoxidase	Abhydrolase_1	*Thermobifida fusca* YX	8E-116	419	A39116_contig_136:024	834	16997	17830	+	No
Chloroperoxidase	Abhydrolase_1	*Mycobacterium* sp. MCS	3E-136	490	A39116_contig_124:194	825	185496	186320	+	No

Glutathione peroxidase	GSHPx	*Streptomyces* sp. AA4	1E-67	258	A39116_contig_084:044	492	38512	38021	-	No

C. Multi-copper oxidases

Predicted identity	Pfam	Organism	E value	Bitscore	Location (contig_#:orf#)	Length (bp)	Start	Stop	Strand	Secreted?
Putative multicopper oxidase	Cu-oxidase	*Kitasatospora setae* KM-6054	8E-129	465	A39116_contig_062:302	1461	314929	316389	+	Yes
Predicted multicopper oxidase	Cu-oxidase	*Saccharomonospora viridis* DSM 43017	5E-117	424	A39116_contig_144:169	921	222722	223642	+	Yes
Multicopper oxidase type 2	Cu-oxidase	*Salinispora arenicola* CNS-205	0	778	A39116_contig_062:081	1860	80097	78238	-	Unlikely
Multi-copper oxidoreductase	Cu-oxidase	*Amycolatopsis mediterranei* U32	4E-091	338	A39116_contig_132:061	702	63084	62383	-	No

such activity are not known [*60-62*]. As discussed earlier, the high potential secreted peroxidases such as the LiPs, MnPs, and VPs, are believed to play the role of initial oxidation and even complete degradation of the native lignins, and it is seen that genome of *A.* sp. 75iv2 encodes a number of potential, unstudied peroxidases. The genome also encodes three members of an interesting yet understudied class of peroxidases, the dye-decolorizing peroxidases (DyPs), of which fungal members have been found to oxidize high redox-potential non-phenolic aromatic compounds [*63-65*], and one bacterial member from *R. jostii* RHA1 has been shown to have the ability to modify lignin [*66*]. However, also as noted earlier, the approach that organisms take in complex target manipulation differs drastically between organisms, so as it is clear that the genome for *A.* sp. 75iv2 encodes many secreted oxidative and peroxide-generating enzymes, it must be recognized that this bacterium may either utilize unexpected and unknown pathways for lignin modification, or alternatively although the lignin degradation pathways may prove similar, the enzymes may not prove orthologous in structure or sequence, limiting utilization of genome mining for candidate searching.

Interestingly, the genome of *A.* sp. 75iv2 encodes 29 cytochrome P450 (CYP) enzymes, a class of oxygenases that have the ability to use molecular oxygen to oxidize compounds for the purpose of modification, detoxification, and breakdown of xenobiotics. These enzymes have been found to play roles in secondary metabolite modification and potentially even in lignin breakdown [*67*]. The genome of the white-rot fungus, *Phanerochaete chrysosporium*, serves as a classic model system for lignolytic microbes and also contains a large number of CYPs (~150) [*68*]. While the physiological significance of CYPs in lignin breakdown remains unclear, several CYPs have been found to be upregulated under lignolytic conditions at the RNA [*69*] and protein [*70*] level. Bacterial genomes frequently encode only a handful of CYPs (0-4 [*7, 46, 71*]) but some species such as *A.* sp. 75iv2 or *S. avermitilis*, which encodes 33 CYP genes. As a result, it has been suggested that these are involved in degradation of small molecules encountered in the environment. Of the CYPs found in *A.* sp. 75iv2, none appear to be located within secondary metabolite biosynthetic clusters. Thus, the role of the many CYPs in *A.* sp. 75iv2 remains unclear. Based on a preliminary analysis of clustered genes, at least six CYP genes reside either up- or downstream of dioxygenases and other enzymes potentially involved in small molecule aromatic degradation. In addition, we have identified two interesting CYP clusters, one of which includes a DyP, a heme peroxidase, and a catalase (*Figure 2.5*).

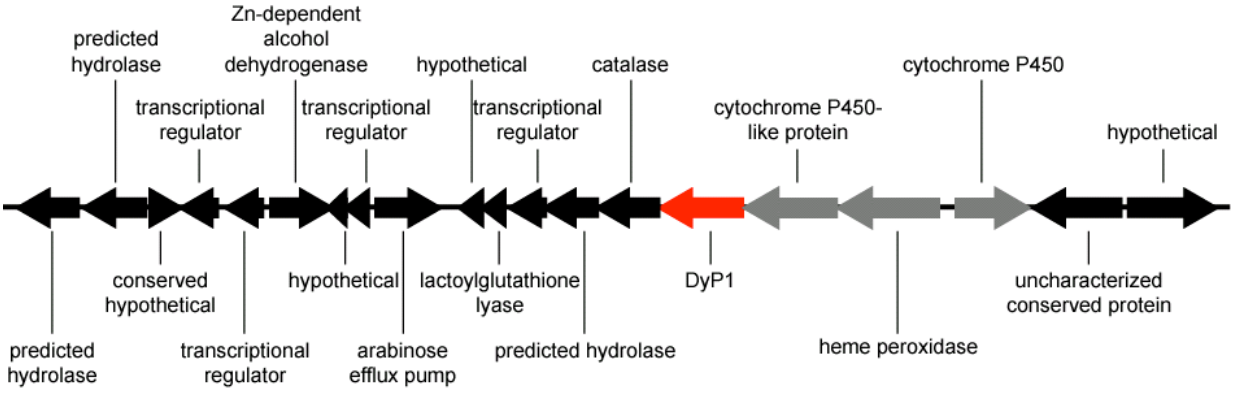


Figure 2.5. *Organization of clustered oxidative genes including a DyP (DyP1), a catalase, 2 P450s, and a heme peroxidase.*

Table 2.8. *Predicted cytochromes in Amycolatopsis sp. 75iv2 ATCC 39116 identified by Pfam and BLASTp search.*

A. Cytochrome P450s

Predicted identity	Pfam	Organism	E value	Bitscore	Location (contig_#:orf#)	Length (bp)	Start	Stop	Strand	Secreted
Cytochrome P450	P450	*Truepera radiovictrix* DSM 17093	7e-46	188	A39116_contig_024:009	1257	12065	10809	-	Yes
Cytochrome P450	P450	*Saccharomonospora viridis* DSM 43017	4e-131	471	A39116_contig_023:168	1239	181389	180151	-	probably
Cytochrome P450	P450	*Streptomyces* sp. AA4	6e-75	285	A39116_contig_151:082	1212	80998	79787	-	No
Cytochrome P450	P450	*Rhodococcus opacus* B4	0	657	A39116_contig_157:069	1212	80219	79008	-	No
Cytochrome P450	P450	*Streptomyces* sp. AA4	0	723	A39116_contig_109:001	1446	1703	258	-	No
Cytochrome P450	P450	*Streptomyces* sp. AA4	7e-155	550	A39116_contig_049:017	1155	18888	17734	-	No
Cytochrome P450	P450	*Streptomyces* sp. AA4	3e-172	608	A39116_contig_062:525	1233	538684	539916	+	No
Cytochrome P450	P450	*Streptomyces* sp. AA4	2e-146	522	A39116_contig_062:532	1185	546852	545668	-	No
Cytochrome P450	P450	*Streptomyces ghanaensis* ATCC 14672	4e-45	185	A39116_contig_040:129	1104	123581	122478	-	No
Cytochrome P450	P450	*Saccharomonospora viridis* DSM 43017	1e-176	622	A39116_contig_019:005	1320	5564	6883	+	No
Cytochrome P450	P450	*Streptomyces* sp. AA4	1e-125	453	A39116_contig_020:084	1140	89806	88667	-	No
Cytochrome P450	P450	*Mycobacterium* sp. KMS	1e-123	446	A39116_contig_023:090	1308	101999	100692	-	No
Cytochrome P450	P450	*Amyoclatopsis mediterranei* U32	1e-137	494	A39116_contig_017:368	1221	376414	375194	-	No
Cytochrome P450	P450	*Streptomyces griseoflavus* Tu4000	1e-104	384	A39116_contig_136:077	2037	66737	64701	-	No
Cytochrome P450 hydroxylase	P450	*Saccharopolyspora erythraea* NRRL 2338	3e-154	548	A39116_contig_048:036	1191	36870	35680	-	No
Cytochrome P450 hydroxylase	P450	*Streptomyces* sp. AA4	2e-137	492	A39116_contig_137:006	1095	6738	5644	-	No
Cytochrome P450 hydroxlyase	P450	*Streptomyces refuineus* subsp.	6e-50	202	A39116_contig_124:269	1206	260916	259711	-	No
Cytochrome P450 hydroxylase	P450	*Streptomyces* sp. E14	4e-146	522	A39116_contig_136:069	1203	59635	60837	+	No
Cytochrome P450 monooxygenase	P450	*Gordonia bronchialis* DSM 43247	0	695	A39116_contig_007:125	1224	113958	115181	+	No
Cytochrome P450 monooxygenase	P450	*Saccharopolyspora erythraea* NRRL 2338	2e-103	379	A39116_contig_017:155	1278	159148	160425	+	No
Cytochrome P450 monooxygenase	P450	*Streptomyces pristinaespiralis* ATCC 25486	5e-85	318	A39116_contig_040:283	1203	270362	271564	+	No
Cytochrome P450 CYP116	P450	*Rhodococcus equi* ATCC 33707	0	1236	A39116_contig_121:036	2313	39719	42031	+	Possibly
Cytochrome P450 CYP116	P450	*Rhodococcus jostii* RHA1	3e-123	445	A39116_contig_040:034	1302	36108	34807		No
Cytochrome P450 CYP125	P450	*Streptomyces* sp. AA4	0	744	A39116_contig_017:053	1227	56525	55299	-	No
Cytochrome P450 CYP125	P450	*Saccharopolyspora erythraea* NRRL 2338	3e-144	515	A39116_contig_064:134	1236	144017	142782	-	No
Cytochrome P450 CYP199	P450	*Rhodococcus jostii* RHA1	7e-159	563	A39116_contig_102:060	1191	72391	71201	-	No
Cytochrome P450 CYP136	P450	*Rhodococcus jostii* RHA1	3e-172	609	A39116_contig_120:034	1359	31034	32392	+	No

Putative cytochrome P450	-	*Saccharopolyspora erythraea* NRRL 2338	8e-146	520	A39116_contig_165:086	1113	82340	83452	+	No
Putative cytochrome P450	-	*Saccharopolyspora erythraea* NRRL 2338	2e-11	71.2	A39116_contig_135:065	249	70955	71203	+	Possibly

B. Other cytochromes

Predicted identity	Pfam	Organism	E value	Bitscore	Location (contig_#:orf#)	Length (bp)	Start	Stop	Strand	Secreted
Cytochrome c oxidase (I)	COX1	*Actinosynnema mirum* DSM 43827	2e-92	341	A39116_contig_028:001	570	1057	488	-	Possibly
Cytochrome c oxidase (I)	COX1	*Actinosynnema mirum* DSM 43827	0	925	A39116_contig_035:048	1776	42707	44482	+	No
Cytochrome c oxidase (I)	COX1	*Streptomyces* sp. AA4	4e-121	437	A39116_contig_059:401	693	402288	401596	-	No
Cytochrome c oxidase (I)	COX1	*Streptomyces* sp. AA4	9e-107	389	A39116_contig_092:001	612	28	639	+	No
Cytochrome c oxidase (I)	COX1	*Saccharomonospora viridis* DSM 43017	0	927	A39116_contig_132:008	1782	6672	4891	-	No
Putative cytochrome c oxidase (II)	COX2	*Streptomyces* sp. AA4	9e-125	449	A39116_contig_017:386	900	392387	391488	-	Unlikely
Cytochrome menaquinol oxidase (II)	Cyto_ox_2	*Rhodococcus opacus* B4	3e-118	428	A39116_contig_120:016	1014	14500	15513	+	No
Cytochrome d ubiquinol oxidase (II)	Cyto_ox_2	*Streptomyces* sp. AA4	1e-141	506	A39116_contig_049:036	1008	33864	34871	+	No
Heme/copper-type cytochrome/quinol	COX3	*Saccharomonospora viridis* DSM 43017	2e-98	361	A39116_contig_017:381	621	388870	388250	-	No
Cytochrome *c* oxidase	Cyt_c_ox_	*Streptomyces* sp. AA4	2E-41	182	A39116_contig_017:385	420	391435	391016	-	No
Cytochrome oxidase assembly protein	COX15-	*Streptomyces* sp. AA4	2e-127	458	A39116_contig_071:027	978	28646	29623	+	Possibly
Cytochrome c class I	Cytochrom	*Streptomyces* sp. AA4	3e-119	431	A39116_contig_017:380	819	388195	387377	-	Yes
Cytochrome c-type biogenesis protein	Cytochrom	*Streptomyces* sp. AA4	4e-138	494	A39116_contig_062:163	990	151806	152795	+	No
Cytochrome b membrane protein	Cytochrom	*Streptomyces* sp. AA4	0	952	A39116_contig_017:378	1671	386211	384541	-	Unlikely
Cytochrome bd quinol oxidase (I)	Bac_Ubq_	*Saccharomonospora viridis* DSM 43017	0	847	A39116_contig_049:035	1476	32379	33854	+	Possibly
Cytochrome bd ubiquinol oxidase (I)	Bac_Ubq_	*Thermobispora bispora* DSM 43833	0	676	A39116_contig_120:015	1419	13072	14490	+	No
Cytochrome c-type biogenesis protein	DsbD	*Streptomyces* sp. AA4	2e-96	355	A39116_contig_062:161	774	149431	150204	+	Yes
Cytochrome b SQR/fumarate reductase	Sdh_cyt	*Salinispora tropica* CNB-440	4e-67	257	A39116_contig_076:003	699	3451	2753	-	Unlikely

2.4. Conclusions

To close, comparison of growths of potential lignin-degrading bacteria in the presence of lignin sources revealed the highest extracellular oxidative capacity as well as production of APPL to reside in *A.* sp. 75iv2; thus, this organism was chosen as the model organism for study. Using massively-parallel sequencing based on the Illumina platform, we assembled a *de novo* genome sequence for *A.* sp. 75iv2, and analysis of the genome of *A.* sp. 75iv2 revealed the presence of not only expected oxidative enzymes but also carbohydrate breakdown enzymes; it also revealed the presence of genes encoding potential demethylases and P450 monooxygenases for lignin-related compound reactivity which could act synergistically for the degradation of biomass by uncapping and adding new phenolic sites. With this tool in hand, we can begin to explore the function and reactivity of the full oxidative system of *A.* sp. 75iv2, including laccases and DyPs, and their synergy with other enzyme families.

2.5. References

1. R. Dixon and D. Kahn, Genetic regulation of biological nitrogen fixation, *Nature Rev. Microbiol.* **2004**, *2*, 621-631.
2. A. J. van den Wijngaard, R. D. Wind and D. B. Janssen, Kinetics of bacterial growth on chlorinated aliphatic compounds, *Appl. Environ. Microbiol.* **1993**, *59*, 2041-2048.
3. R. A. Kanaly and S. Harayama, Biodegradation of high-molecular-weight polycyclic aromatic hydrocarbons by bacteria, *J. Bacteriol.* **2000**, *182*, 2059-2067.
4. G. Chen, H. Zhu and Y. Zhang, Soil microbial activities and carbon and nitrogen fixation, *Res. Microbiol.* **2003**, *154*, 393-398.
5. P. Cornelis, Iron uptake and metabolism in *pseudomonads*, *Appl. Microbiol. Biotechnol.* **2010**, *86*, 1637-1645.
6. S. Uroz, C. Calvaruso, M.-P. Turpault and P. Frey-Klett, Mineral weathering by bacteria: Ecology, actors and mechanisms, *Trends Microbiol.* **2009**, *17*, 378-387.
7. M. T. Alam, M. H. Medema, E. Takano and R. Breitling, Comparative genome-scale metabolic modeling of actinomycetes: The topology of essential core metabolism, *FEBS Lett.* **2011**, *585*, 2389-2394.
8. M. E. Himmel, S.-Y. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, J. W. Brady and T. D. Foust, Biomass recalcitrance: Engineering plants and enzymes for biofuels production, *Science* **2007**, *315*, 804-807.
9. C. Somerville*, et al.*, Toward a Systems Approach to Understanding Plant Cell Walls, *Science* **2004**, *306*, 2206-2211.
10. T. K. Kirk and R. L. Farrell, Enzymatic "combustion": The microbial degradation of lignin, *Annu. Rev. Microbiol.* **1987**, *41*, 465-505.
11. R. ten Have and P. J. Teunissen, Oxidative mechanisms involved in lignin degradation by white-rot fungi, *Chem. Rev.* **2001**, *101*, 3397-413.
12. A. Khindaria, D. P. Barr and S. D. Aust, Lignin peroxidases can also oxidize manganese, *Biochemistry* **1995**, *34*, 7773-7779.

13. A. Khindaria, T. A. Grover and S. D. Aust, Evidence for formation of the veratryl alcohol cation radical by lignin peroxidase, *Biochemistry* **1995**, *34*, 6020-6025.

14. D. L. Crawford, Lignocellulose decomposition by selected *Streptomyces* strains, *Appl. Environ. Microbiol.* **1978**, *35*, 1041-1045.

15. D. L. Crawford, A. L. Pometto, III and R. L. Crawford, Lignin degradation by *Streptomyces viridosporus*: Isolation and characterization of a new polymeric lignin degradation intermediate, *Appl. Environ. Microbiol.* **1983**, *45*, 898-904.

16. R. Kirby, Actinomycetes and lignin degradation, In *Advances in Applied Microbiology*, A. I. Laskin, J. W. Bennett, G. M. Gadd and S. Sariaslani, Academic Press: London, UK, **2005**, vol. 58, pp. 125-168.

17. L. Thomas and D. L. Crawford, Cloning of clustered *Streptomyces viridosporus* T7A lignocellulose catabolism genes encoding peroxidase and endoglucanase and their extracellular expression in *Pichia pastoris*, *Can. J. Microbiol.* **1998**, *44*, 364-372.

18. M. Ahmad, C. R. Taylor, D. Pink, K. Burton, D. Eastwood, G. D. Bending and T. D. H. Bugg, Development of novel assays for lignin degradation: Comparative analysis of bacterial and fungal lignin degraders, *Mol. BioSyst.* **2010**, *6*, 815-821.

19. J. R. Borgmeyer and D. L. Crawford, Production and characterization of polymeric lignin degradation intermediates from two different *Streptomyces* spp., *Appl. Environ. Microbiol.* **1985**, *49*, 273-278.

20. E. R. Mardis, Next-generation DNA sequencing methods, *Annu. Rev. Genom. Human Genet.* **2008**, *9*, 387-402.

21. E. Masai, Y. Katayama, S. Nishikawa and M. Fukuda, Characterization of *Sphingomonas paucimobilis* SYK-6 genes involved in degradation of lignin-related compounds, *J. Ind. Microbiol. Biotechnol.* **1999**, *23*, 364-373.

22. F. Franklin, M. Bagdasarian, M. Bagdasarian and K. Timmis, Molecular and functional analysis of the TOL plasmid pWWO from *Pseudomonas putida* and cloning of genes for the entire regulated aromatic ring meta cleavage pathway, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 7458-7462.

23. J. I. Jiménez, B. Minambres, J. L. Garcia and E. Diaz, Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440, *Environ. Microbiol.* **2002**, *4*, 824-841.

24. L. N. Ornston, The conversion of catechol and protocatechuate to β-ketoadipate by *Pseudomonas putida*. II. Enzymes of the protocatechuate pathway, *J. Biol. Chem.* **1968**, *241*, 3787-3794.

25. T. G. Pridham and D. Gottlieb, The utilization of carbon compounds by some Actinomycetales as an aid for species determination, *J. Bacteriol.* **1948**, *56*, 107-114.

26. M. Ramachandra, D. L. Crawford and G. Hertel, Characterization of an extracellular lignin peroxidase of the lignocellulolytic actinomycete *Streptomyces viridosporus*, *Appl. Env. Microbiol.* **1988**, *54*, 3057-3063.

27. H. Giroux, P. Vidal, J. Bouchard and F. Lamy, Degradation of kraft lignin by *Streptomyces viridosporus* and *Streptomyces badius*, *Appl. Environ. Microbiol.* **1988**, *54*, 3064-3070.

28. J. M. B. Macedo, L. M. F. Gottschalk and E. P. S. Bon, Lignin peroxidase and protease production by *Streptomyces viridosporus* T7A in the presence of calcium carbonate, *Appl. Biochem. Biotech.* **1999**, *79*, 735-744.

29. D. C. Yee, D. Jahng and T. K. Wood, Enhanced expression and hydrogen peroxide dependence of lignin peroxidase from *Streptomyces viridosporus* T7A, *Biotechnol. Prog.* **1996**, *12*, 40-46.

30. T. Kieser, K. F. Chater, M. J. Bibb, M. J. Buttner and D. A. Hopwood, Practical *Streptomyces* Genetics, The John Innes Foundation: Norwich, **2000**, vol. pp.

31. N. Ji, B. Peng, G. Wang, S. Wang and X. Peng, Universal primer PCR with DGGE for rapid detection of bacterial pathogens, *J. Microbiol. Methods* **2004**, *57*, 409-413.

32. J. Lu, C. Perng, S. Y. Lee and C. Wan, Use of PCR with universal primers and restriction endonuclease digestions for detection and identification of common bacterial pathogens in cerebrospinal fluid, *J. Clin. Microbiol.* **2000**, *38*, 2076-2080.

33. D. R. Zerbino and E. Birney, Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs, *Genome Res.* **2008**, *18*, 821-829.

34. H. Li and R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, *Bioinformatics* **2009**, *25*, 1754-1760.

35. I. J. Tsai, T. D. Otto and M. Berriman, Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps, *Genome Biol.* **2010**, *11*, R41.

36. A. L. Delcher, K. A. Bratke, E. C. Powers and S. L. Salzberg, Identifying bacterial genes and endosymbiont DNA with Glimmer, *Bioinformatics* **2007**, *23*, 673-679.

37. S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, *Nucleic Acids Res.* **1997**, *25*, 3389-3402.

38. A. Bateman, E. Birney, R. Durbin, S. R. Eddy, K. L. Howe and E. L. L. Sonnhammer, The Pfam protein families database, *Nucleic Acids Res.* **2000**, *28*, 263-266.

39. K. Lagesen, P. Hallin, E. A. Rødland, H.-H. Stærfeldt, T. Rognes and D. W. Ussery, RNAmmer: Consistent and rapid annotation of ribosomal RNA genes, *Nucleic Acids Res.* **2007**, *35*, 3100-3108.

40. P. Schattner, A. N. Brooks and T. M. Lowe, The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs, *Nucleic Acids Res.* **2005**, *33*, W686-W689.

41. J. Dyrløv Bendtsen, H. Nielsen, G. von Heijne and S. Brunak, Improved prediction of signal peptides: SignalP 3.0, *J. Mol. Biol.* **2004**, *340*, 783-795.

42. H. Nielsen, J. Engelbrecht, S. Brunak and G. von Heijne, Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites, *Protein Eng.* **1997**, *10*, 1-6.

43. S. H. Beis*, et al.*, Fast pyrolysis of lignins, *BioRes.* **2010**, *5*, 1408-1424.

44. S. P. Antai and D. L. Crawford, Degradation of softwood, hardwood, and grass lignocelluloses by two *Streptomyces* strains, *Appl. Environ. Microbiol.* **1981**, *42*, 378-380.

45. S. D. Bentley*, et al.*, Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2), *Nature* **2002**, *417*, 141-147.

46. H. Ikeda, J. Ishikawa, A. Hanamoto, M. Shinose, H. Kikuchi, T. Shiba, Y. Sakaki, M. Hattori and S. Omura, Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*, *Nat. Biotech.* **2003**, *21*, 526-531.

47. J. Ishikawa, A. Yamashita, Y. Mikami, Y. Hoshino, H. Kurita, K. Hotta, T. Shiba and M. Hattori, The complete genomic sequence of *Nocardia farcinica* IFM 10152, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 14925-14930.

48. Y. Ohnishi, J. Ishikawa, H. Hara, H. Suzuki, M. Ikenoya, H. Ikeda, A. Yamashita, M. Hattori and S. Horinouchi, Genome sequence of the streptomycin-producing microorganism *Streptomyces griseus* IFO 13350, *J. Bacteriol.* **2008**, *190*, 4050-4060.

49. W. Zhao*, et al.*, Complete genome sequence of the rifamycin SV-producing *Amycolatopsis mediterranei* U32 revealed its genetic characteristics in phylogeny and metabolism, *Cell Res.* **2010**, *20*, 1096-1108.

50. J. R. Davis*, et al.*, Genome sequence of *Amycolatopsis* sp. strain ATCC 39116, a plant biomass-degrading actinomycete, *J. Bacteriol.* **2012**, *194*, 2396-2397.

51. M. P. McLeod*, et al.*, The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse, *Proc. Natl. Acad. Sci.* **2006**, *103*, 15582-15587.

52. M. Oliynyk, M. Samborskyy, J. B. Lester, T. Mironenko, N. Scott, S. Dickens, S. F. Haydock and P. F. Leadlay, Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL23338, *Nat. Biotech.* **2007**, *25*, 447-453.

53. J.-C. Camus, M. J. Pryor, C. Medigue and S. T. Cole, Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv, *Microbiology* **2002**, *148*, 2967-2973.

54. X.-J. Wang*, et al.*, Genome sequence of the milbemycin-producing bacterium *Streptomyces bingchenggensis*, *J. Bacteriol.* **2010**, *192*, 4526-4527.

55. K. Blin, M. H. Medema, D. Kazempour, M. A. Fischbach, R. Breitling, E. Takano and T. Weber, antiSMASH 2.0 - a versatile platform for genome mining of secondary metabolite producers, *Nucleic Acids Res.* **2013**, *41*, W204-W212.

56. F. C. Franklin, M. Bagdasarian, M. M. Bagdasarian and K. N. Timmis, Molecular and functional analysis of the TOL plasmid pWWO from *Pseudomonas putida* and cloning of genes for the entire regulated aromatic ring *meta* cleavage pathway, *Proc. Natl. Acad. Sci.* **1981**, *78*, 7458-7462.

57. T. Sonoki, Y. Otsuka, S. Ikeda, E. Masai, S. Kajita and Y. Katayama, Close association between the enzymes involved in the lignin metabolic pathway of *Sphingomonas paucimobilis* SYK-6: Interaction of *O*-demethylase (LigX) and ring fission dioxygenase (LigZ), *J. Wood Sci.* **2002**, *48*, 250-252.

58. H. W. Kern and T. K. Kirk, Influence of molecular size and ligninase pretreatment on degradation of lignins by *Xanthomonas* sp. strain 99, *Appl. Environ. Microbiol.* **1987**, *53*, 2242-2246.

59. R. Vicuña, Bacterial degradation of lignin, *Enz. Microb. Technol.* **1988**, *10*, 646-655.

60. A. Leonowicz, N. Cho, J. Luterek, A. Wilkolazka, M. Wojtas-Wasilewska, A. Matuszewska, M. Hofrichter, D. Wesenberg and J. Rogalski, Fungal laccase: Properties and activity on lignin, *J. Basic Microbiol.* **2001**, *41*, 185-227.

61. H.-D. Youn, Y. C. Hah and S.-O. Kang, Role of laccase in lignin degradation by white-rot fungi, *FEMS Microbiol. Lett.* **1995**, *132*, 183-188.

62. R. Bourbonnais, D. Leech and M. G. Paice, Electrochemical analysis of the interactions of laccase mediators with lignin model compounds, *Biochim. Biophys. Acta* **1998**, *1379*, 381-390.

63. Y. Sugano, R. Muramatsu, A. Ichiyanagi, T. Sato and M. Shoda, DyP, a unique dye-decolorizing peroxidase, represents a novel heme peroxidase family, *J. Biol. Chem.* **2007**, *282*, 36652-36658.

64. S. J. Kim and M. Shoda, Purification and characterization of a novel peroxidase from *Geotrichum candidum* Dec 1 involved in decolorization of dyes, *Appl. Env. Microbiol.* **1999**, *65*, 1029-1035.

65. C. Liers, C. Bobeth, M. Pecyna, R. Ullrich and M. Hofrichter, DyP-like peroxidases of the jelly fungus *Auricularia auricula-judae* oxidize nonphenolic lignin model compounds and high-redox potential dyes, *Appl. Microbiol. Biotechnol.* **2010**, *85*, 1869-1879.

66. M. Ahmad, J. N. Roberts, E. M. Hardiman, R. Singh, L. D. Eltis and T. D. H. Bugg, Identification of DypB from *Rhodococcus jostii* RHA1 as a lignin peroxidase, *Biochemistry* **2011**, *50*, 5096-5107.

67. I. G. Denisov, T. M. Makris, S. G. Sligar and I. Schlichting, Structure and chemistry of cytochrome P450, *Chem. Rev.* **2005**, *105*, 2253-2278.

68. D. Martinez*, et al.*, Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78, *Nat. Biotechnol.* **2004**, *6*, 695-700.

69. H. Doddapaneni and J. S. Yadav, Microarray-based global differential expression profiling of P450 monooxygenases and regulatory proteins for signal transduction pathways in the white rot fungus *Phanerochaete chrysosporium*, *Mol. Gen. Genomics* **2005**, *274*, 454-466.

70. S. Shary, A. N. Kapich, E. A. Panisko, J. K. Magnuson, D. Cullen and K. E. Hammel, Differential expression in *Phanerochaete chrysosporium* of membrane-associated proteins relevant to lignin degradation, *Appl Environ Microbiol* **2008**, *74*, 7252-7257.

71. D. F. V. Lewis and A. Wiseman, A selective review of bacterial forms of cytochrome P450 enzymes, *Enz. Microb. Technol.* **2005**, *36*, 377-384.

Chapter 3: *Identification and biochemical characterization of two heme-containing proteins from the secretome of Amycolatopsis sp. 75iv2*

Portions of this work were published in the following scientific journal:

M. E. Brown, M. C. Walker, T. G. Nakashige, A. T. Iavarone and M. C. Y. Chang, Discovery and characterization of heme enzymes from unsequenced bacteria: Application to microbial lignin degradation, *J. Am. Chem. Soc.* **2011**, *133*, 18006-18009.

3.1. Introduction

While the discovery of microbes that display lignin reactivity continues rapidly, far less information is available on the molecular identification and biochemical characterization of bacterial enzymes that may be involved in metabolizing lignin [*1-3*]. Lignin is a complex substrate and likely requires an array of enzymes for its breakdown and deconstruction. We were therefore interested in exploring the diversity of proteins present during growth of lignin-reactive bacteria with this substrate. Given the larger size of lignin compared to bacterial dimensions, most enzymes involved in initiating biomass deconstruction are thought to reside in the extracellular environment and hence the secreted protein fraction. We thus set out to begin the characterizing the secretome of *Amycolatopsis* sp. 75iv2 as a model bacterial host, based on our previous studies that showed *A.* sp. 75iv2 to be the most active of the bacterial species screened with respect to formation of acid-precipitable polymeric lignin (APPL) and capacity for extracellular phenol oxidation (*Chapter 2*).

Analysis of the draft genome assembly of *A.* sp. 75iv2 revealed the presence of several potential oxidative systems, including secreted laccases, peroxidases, and peroxide-generating enzymes (*Chapter 2*). We focused first on examining the hemoprotein profile of the secretome as heme peroxidases have been reported to have been secreted from soil-dwelling organisms and linked to bacterial lignin degradation. The most well-studied system is represented by *Streptomyces viridosporus* T7A from which one of these extracellular peroxidases was purified to relative homogeneity and shown to be biochemically competent for the degradation of non-phenolic lignin model dimers [*4*]. However, the sequence and identity of this enzyme has yet to be reported; furthermore, assignment of its function by heterologous expression of the gene candidate yields ambiguous results [*5*]. By profiling the *A.* sp. 75iv2 secretome in the presence of either *Miscanthus* lignocellulose or a desulfonated kraft lignin (indulin AT), we found two abundant and extracellular heme proteins, named Amyco1 and Amyco2. Further biochemical characterization of Amyco1 and Amyco2 allowed the assignment of their function as a catalase-peroxidase and catalase, respectively. While competent to oxidize phenolic sites, we found that Amyco1 was unable to degrade non-phenolic model dimers, which suggests *A.* sp. 75iv2 has weak overall oxidative capacity compared to fungi or that other oxidative enzymes may be primarily responsible for lignin reactivity.

3.2. Materials and Methods

Materials and methods. Terrific broth (TB), LB Broth Miller (LB), LB Agar Miller, potassium phosphate monobasic, yeast extract, cupric sulfate pentahydrate, and glycerol were purchased from EMD Biosciences (Darmstadt, Germany). Isopropyl β-D-1-thiogalactopyranoside (IPTG), D-glucose, dithiothreitol (DTT), Tris-HCl, phenylmethanesulfonyl fluoride (PMSF), carbenicillin (Cb), sodium chloride, ethylenediaminetetraacetic acid (EDTA), isopropanol, sodium acetate, acetonitrile, ammonium sulfate, hydrogen peroxide, methanol, and manganese chloride tetrahydrate were purchased from Fisher Scientific (Pittsburgh, PA). Lysozyme, potassium phosphate dibasic, 4-aminoantipyrene, 2,4-dichlorophenol, ammonium bicarbonate, δ-aminolevulinic acid, magnesium sulfate heptahydrate, iron sulfate heptahydrate, 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid (ABTS), trichloroacetic acid (TCA), zinc sulfate monohydrate, N,N,N′,N′-tetramethylethylenediamine, sodium dithionite, hemin chloride, methyl viologen, anthraquinone-2,6-disulfonic acid, 2-hydroxy-1,4-naphthoquinone, 2,5-dihydroxy-1,4-benzoquinone,

tetramethyl-*p*-benzoquinone (duroquinone), 1,2-naphthoquinone, $K_3[Fe(CN)_6]$, and 3,3′,5,5′-tetramethylbenzidine were purchased from Sigma-Aldrich (St. Louis, MO). Formic acid was purchased from Acros Organics (Morris Plains, NJ). Polyacrylamide, electrophoresis grade sodium dodecyl sulfate (SDS), gel filtration standard and ammonium persulfate were purchased from Bio-Rad Laboratories (Hercules, CA). All PCR amplifications for cloning were carried out using Phusion polymerase (New England BioLabs; Ipswich, MA). Restriction enzymes, Antarctic phosphatase, and T4 DNA ligase were purchased from New England Biolabs (Ipswich, MA). Deoxynucleotides (dNTPs), Platinum Taq High-Fidelity polymerase (Pt Taq HF), and PCR2.1-TOPO TA cloning kit were purchased from Invitrogen (Carlsbad, CA). DNase was purchased from Fermentas (Glen Burnie, Maryland). DNA was isolated using the QIAprep Spin Miniprep Kit, QIAquick PCR Purification Kit, and QIAquick Gel Extraction Kit (Qiagen; Valencia, CA). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), resuspended at a stock concentration of 100 μM in 10 mM Tris-HCl, pH 8.5, and stored at either 4 °C for immediate usage or -20 °C for longer term usage. Following plasmid construction, all cloned inserts were sequenced at Quintara Biosciences (Berkeley, CA). Liquid chromatography-tandem mass spectrometry (LC-MS/MS) was performed at the Vincent J. Coates Proteomics/Mass Spectrometry Laboratory at UC Berkeley. For preparation of mobile phases for LC-MS/MS, acetonitrile (Fisher Optima grade, 99.9%) and formic acid (Pierce, 1 mL ampules, 99+%) were purchased from Fisher Scientific (Pittsburgh, PA) and water was purified to a resistivity of 18.2 MΩ·cm (at 25 °C) using a Milli-Q Gradient ultrapure water purification system (Millipore, Billerica, MA). All spectra except for those involved in the potentiostatic titration and the stopped-flow experiments were collected on a Beckman Coulter DU-800 UV-Visible spectrophotometer (Fullterton, CA). Spectra monitoring potentiostatic titration were collected on a Varian Cary Bio 300 UV-Visible Spectrophotometer (Agilent Technologies; Santa Clara, CA), and stopped-flow spectra were collected using a Hi-Tech Scientific double mixing stopped-flow system equipped with a diode array detector and KinetAsyst software (Hi-Tech Scientific; Bradford-on-Avon, United Kingdom).

Bacterial strains. *E. coli* DH10B-T1^R was used for plasmid construction and BL21(de3) and RP523 (Yale *E. coli* Genetic Stock Center) were used for heterologous protein production. *Amycolatopsis* sp. 75iv2 (formerly *Streptomyces setonii* and *Streptomyces griseus* 75vi2, ATCC 39116) was purchased from the American Tissue Type Collection (Manassas, VA).

Cell culture. *Amycolatopsis* sp. 75iv2 was cultured cells resuspended in 20% sterile glycerol and frozen at -80 °C. For peroxidase expression, frozen cells were spread onto YEME agar (2% w/v) plates using sterile glass beads (5 mm) and incubated at 37 °C for 3-4 d until plates were covered in spores. Mycelia mat from 3-4 d YEME agar plates was resuspended in sterile ddH_2O and used to inoculate expression media containing carbenicillin (50 μg/mL) and either *Miscanthus* lignocellulose or indulin. Otherwise, a single colony was used to inoculate the expression media. Control flasks containing either no lignin source or no spores/cells were included in the experiment to monitor for possible environmental contamination. Small-scale (50 mL) and large-scale (500 mL) cell cultures were carried out in 250-mL and 2-L baffled flasks, respectively. The cultures were incubated at 37 °C at 200 rpm. Daily samples were removed (small-scale, 2 mL; large-scale, 40 mL). For large-scale *A.* sp. 75iv2 cultures, carbenicillin (50 μg/mL) was added to all flasks daily to prevent contamination.

Expression media. Growth media was prepared according to literature methods [*4, 6*]. Basal medium included (per L): ammonium sulfate (2.64 g), potassium phosphate monobasic (2.38

g), potassium phosphate dibasic (5.65 g), magnesium sulfate heptahydrate (1.00 g), cupric sulfate pentahydrate (6.4 mg), iron sulfate heptahydrate (1.1 mg), manganese chloride tetrahydrate (7.9 mg), and zinc sulfate monohydrate (0.94 mg) with yeast extract (6.0 g). When appropriate, sterilized indulin or *Miscanthus giganteus* lignocellulose was added to a final concentration of 0.5 % (w/v).

Preparation of indulin and Miscanthus giganteus lignocellulose. Indulin was obtained as a gift from MeadWestvaco (Campbell, CA) and prepared according to literature by washing in a Soxhlet apparatus with water at 100 °C for approximately 72 h until extractions were clear [*7*]. The indulin was then acidified by washing with water (pH 2.5), filtered (Whatman No. 1 filter paper), dried at 100 °C, and ground to a powder using a mortar and pestle. Before use in cell culture, the processed indulin was heated at 135 °C for 3 h. *Miscanthus giganteus* lignocellulose, ball-milled to 2 mm, was obtained from the Energy Biosciences Institute (Berkeley, CA) and sterilized by autoclave before use.

Peroxidase activity assay. Peroxidase activity was monitored spectrophotometrically using a modified literature procedure [*8*]. For measuring peroxidase activity in cell culture, a sample (1 mL) was removed and cleared of cells and other debris by centrifugation at 9,800 × *g* for 7 min. The culture supernatant was used directly in the peroxidase assay using a SpectraMax M2 96-well plate reader (Molecular Devices; Sunnyvale, CA). Assays contained culture supernatant (100 µL), 2,4-dichlorophenol (3 mM), 4-aminoantipyrene (164 µM), and H_2O_2 (4 mM) in 50 mM potassium phosphate buffer, pH 7.0 in a total assay volume of 200 µL. Assays were initiated by the addition of H_2O_2 and monitored for an increase in absorbance at 510 nm at 25 °C. The same assay conditions were used for purified protein (10-100 µg) using a DU-800 spectrophotometer. Peroxidase activity was calculated using an extinction coefficient of 18,500 $M^{-1}cm^{-1}$ for the 4-aminoantipyrene/2,4-dichlorophenol adduct [*9*].

Catalase activity assay. Catalase activity of purified protein was monitored using a modified literature protocol [*10*]. Reactions were performed in 50 mM potassium phosphate, pH 7.0 at 25 °C and initiated by the addition of H_2O_2 (20 mM) in a total assay volume of 1 mL. Activity was monitored by a decrease in absorbance at 240 nm and calculated using an extinction coefficient of 43.6 $M^{-1}cm^{-1}$ for H_2O_2.

SDS-PAGE with overlaid heme and silver stain. The detection of extracellular heme-containing proteins from culture growths was performed using a modified literature protocol for a gel-based heme stain [*11*]. Samples (1 mL) were removed from cell culture and passed through a 0.2 µm filter. Each sample was then concentrated 20-fold from 600 µL to 30 µL before adding Laemmli sample buffer without β-mercaptoethanol (10 µL; 0.25 M Tris-HCl, 2% SDS, 40% glycerol, 0.04% bromophenol blue) and incubating at room temperature for 15 min. The sample was analyzed by SDS-PAGE at 4 °C on either an 8% or 12% gel at 200V for approximately 60-80 min. The gel was then soaked at room temperature in the dark with 3,3′,5,5′-tetramethylbenzidine (2 mM) in 0.25 M sodium acetate, pH 5.0 containing 30% (v/v) methanol for 2 h, with manual shaking every 15 to 30 min. H_2O_2 (60 mM) was added to initiate the development of blue bands within 15 min and stopped by rinsing with ddH_2O. To visualize non-heme containing proteins, the SilverQuest staining kit (Invitrogen; Carlsbad, CA) was used on heme-stained gel according to manufacturer instructions.

In-gel tryptic digest of heme-stained protein bands. Clean SDS-PAGE gels were prepared taking care to avoid keratin contamination (gloves and hair net). Bands of interest were

excised within 8 h after separation and staining and diced to approximately 1 mm pieces using a clean razor blade for extraction. The pieces were placed in a microfuge tube (0.6 mL) and washed 5× with 1:1 acetonitrile: 25 mM aqueous ammonium bicarbonate by immersing the gel piece, vortexing for 10 min, and removing the supernatant. Gel samples were vacuum-centrifuged to dryness and reduced by addition of DTT (10 mM) in 25 mM aqueous ammonium bicarbonate containing 10% acetonitrile (approximately 40 μL), mixing by briefly vortexing and centrifuging, and incubating at 55°C for 1 h. After cooling to room temperature, the supernatant was removed and discarded. Protein thiols were then alkylated by immersing the gel pieces in a solution of iodoacetamide (55 mM) in 25 mM aqueous ammonium bicarbonate. Samples were briefly vortexed and centrifuged and allowed to incubate at room temperature in the dark for 45 minutes. The supernatant was removed and discarded and samples were washed first with 25 mM aqueous ammonium bicarbonate and then with a 1:1 acetonitrile: 25 mM aqueous ammonium bicarbonate solution. This wash protocol was repeated, and the gel pieces were vacuum-centrifuged to dryness. Samples were then digested with mass-spectrometry grade Trypsin Gold (Promega; Madison, WI) by adding 1 gel volume of trypsin (volume based on size of gel fragment, typically 5-15 μL of 12.5 ng/μL trypsin in 25 mM aqueous ammonium bicarbonate) and incubating on ice for 30 min. Excess trypsin was withdrawn using a pipet and samples were then incubated overnight at 37 °C. Digested protein samples were extracted from the gel piece by addition of ddH_2O (3 gel volumes) followed by vortexing for 10 min and then sonication for 5 min. The supernatant from this initial extraction was removed and saved to pool with supernatants from subsequent extractions. The gel piece was then extracted three times by vortexing and sonication in a solution of 50 acetonitrile: 45 ddH_2O: 5 formic acid (3 gel volumes). The supernatant from these extractions were added to the initial extraction with ddH_2O. The extracted samples were concentrated to 10 μL by vacuum-centrifugation for LC-MS/MS analysis.

Preparation of secretome tryptic digests. Samples (30 mL) from 3-day old cultures were centrifuged at 9,800 × *g* for 7 min to remove cells and *Miscanthus* particles. The supernatant was then filtered with Whatman No. 1 filter paper, followed by passage through a 0.2 μm filter. Proteins were precipitated by addition of TCA to a final concentration of 10%, mixing by vortexing, and incubating at -20 °C overnight. Precipitated protein was pelleted by centrifugation at 12,000 × *g* and the supernatant removed. The pellet was resuspended and washed in acetone (300 μL) chilled to 4 °C. The samples were centrifuged again and the pellets were allowed to air dry. The dried pellets were resuspended in 5 mM potassium phosphate, pH 6.5. The mixture was vortexed to dissolve proteins after addition of 1.0 M Tris-HCl (5 μL), 100 mM DTT (5 μL), and urea (36 mg) to the suspension. The samples were then incubated at 56 °C for 1 h after which 25 mM aqueous ammonium bicarbonate (700 μL) and methanol (140 μL) were added. Mass spectrometry-grade Trypsin Gold (50 μL of a 100 μg/mL stock solution in 50 mM sodium acetate, pH 5.0) was added to the sample, which as then incubated overnight at 37 °C. Samples were then dried by vacuum centrifugation and subjected to 3 cycles of resuspension in ddH_2O (300 μL) and drying by vacuum centrifugation. In the final cycle, the sample was removed from the vacuum centrifuge when 100 μL of sample remained. Trifluoroacetic acid (0.3 μL) was then added to each sample. Samples were desalted using Omix C_{18} pipet tips (Varian Inc.; Palo Alto, CA) and eluted into 85 acetonitrile: 15 water containing 0.1% (v/v) formic acid (100 μL). These samples were concentrated by vacuum centrifugation to 10 μL for LC-MS/MS analysis.

Characterization of *Amycolatopsis* sp. 75iv2 proteins by LC-MS/MS. Trypsin-digested proteins were analyzed using a tandem mass spectrometer that was connected in-line with an

ultraperformance liquid chromatograph (UPLC). Peptides were separated using a nanoAcquity UPLC (Waters Corporation; Milford, MA) equipped with C_{18} trapping (180 µm × 20 mm) and analytical (100 µm × 100 mm) columns and a 10 µL sample loop. Solvent A was 99.9% water/0.1% formic acid and solvent B was 99.9% acetonitrile/0.1% formic acid (v/v). Sample solutions contained in 0.3 mL polypropylene snap-top vials sealed with septa caps (Wheaton Science; Millville, NJ) were loaded into the nanoAcquity autosampler compartment prior to analysis. Following sample injection (10 µL), trapping was performed for 5 min with 100% A at a flow rate of 3 µL/min. The injection needle was washed with 500 µL each of solvents A and B after injection to avoid cross-contamination between samples. The elution program consisted of a linear gradient from 20% to 33% B over 70 min, a linear gradient to 95% B over 0.33 min, isocratic conditions at 95% B for 5.67 min, a linear gradient to 1% B over 0.33 min, and isocratic conditions at 1% B for 12.67 min, at a flow rate of 500 nL/min. The analytical column and sample compartment were maintained at 35 ºC and 8 ºC, respectively.

The column was connected to a NanoEase nanoelectrospray ionization (nanoESI) emitter mounted in the nanoflow ion source of a Q-Tof Premier quadrupole time-of-flight mass spectrometer (Waters Corporation). The tip of the nanoESI emitter was positioned approximately 3 mm from the sampling cone aperture. The nanoESI source parameters were as follows: nanoESI capillary voltage 2.3 kV, nebulizing gas (nitrogen) pressure 0.15 mbar, sample cone voltage 30 V, extraction cone and ion guide voltages 4 V, and source block temperature 80 ºC. No cone gas was used. The collision cell contained argon gas at a pressure of 8×10^{-3} mbar. The Tof analyzer was operated in "V" mode. Under these conditions, a mass resolving power of 1.0×10^4 (measured at m/z = 771, full width at half-maximum height) was routinely achieved, which was sufficient to resolve the isotopic distributions of the singly and multiply charged precursor ions and fragment ions measured in this study. Thus, the ion mass and charge could be determined independently with the ion charge determined from the reciprocal of the spacing between adjacent isotope peaks in the m/z spectrum. External mass calibration was performed immediately prior to analysis using solutions of sodium formate. Survey scans were acquired in the positive ion mode over the range m/z = 400-1800 using a 0.95 s scan integration and a 0.05 s interscan delay. In the data-dependent mode, up to five precursor ions exceeding an intensity threshold of 30 cps were selected from each survey scan for tandem mass spectrometry (MS/MS) analysis. Real-time deisotoping and charge state recognition were used to select 2+, 3+, 4+, and 5+ charge state precursor ions for MS/MS. Collision energies for collisionally activated dissociation (CAD) were automatically selected based on the mass and charge state of a given precursor ion. MS/MS spectra were acquired over the range m/z = 50-2000 using a 0.45 s scan integration and a 0.05 s interscan delay. Ions were fragmented to achieve a minimum total ion current (TIC) of 40,000 cps in the cumulative MS/MS spectrum for a maximum of 3 s. To avoid the occurrence of redundant MS/MS measurements, real-time dynamic exclusion was used to preclude re-selection of previously analyzed precursor ions over an exclusion width of ± 0.15 m/z unit for 300 s.

Mass spectrometry data analysis. The data resulting from LC-MS/MS analysis of trypsin-digested proteins were processed using ProteinLynx Global Server software v 2.3 (Waters Corporation), which performed background subtraction (threshold 35% and fifth order polynomial), smoothing (Savitzky-Golay, 10 times, over three channels), and centroiding (top 80% of each peak and minimum peak width at half height four channels) of the mass spectra and MS/MS spectra. The processed data were searched against the *Amycolatopsis* sp. 75iv2 protein database generated from the *de novo* genome assembly. The following criteria were used for the

database search: precursor ion mass tolerance 50 ppm, fragment ion mass tolerance 0.1 Da, and the following variable post-translational modifications: Asn/Gln deamidation and Met oxidation. The identification of at least three consecutive fragment ions from the same series, i.e., b- or y-type fragment ions, was required for assignment of a peptide to an MS/MS spectrum. The protein identifications were validated by manual inspection of the MS/MS spectra to verify the presence of b and y-type fragment ions that uniquely identify the tryptic peptides.

Construction of plasmids. Expression plasmids for His_{10}-Amyco1, His_{10}-TEV-Amyco1, and His_{10}-Amyco2 were constructed in *E. coli* DH10B-$T1^R$ using standard techniques. The genes encoding Amyco1 and Amyco2 were amplified from genomic DNA using the following primer sets: Amyco1 F1 and Amyco1 R100 (Amyco1), Amyco1 F100 and Amyco1 R100 (TEV-Amyco1), and Amyco2 F1 and Amyco2 R1 (Amyco2). Amyco1 and Amyco2 were inserted into the NdeI-XbaI sites of pCWOri-HisN to add an N-terminal His_{10}-affinity tag. TEV-Amyco1 was inserted into the NdeI-XbaI sites of pCWOri-HisN adding an N-terminal His_{10}-affinity tag with a TEV-cleavage site to remove the affinity tag and leave a residual N-terminal GA insertion. All constructs were verified by sequencing.

Gene synthesis and construction of an expression plasmid for His_{10}-Amyco1GK. The synthetic gene encoding Amyco1 ortholog from *Geobacillus kaustophilus* was optimized for *E. coli* class II codon usage (Gene Designer 2.0, DNA 2.0) and synthesized using PCR assembly. Gene2Oligo (http://berry.engin.umich.edu/gene2oligo) was used to convert the gene sequence into primer sets using default optimization settings (Appendix). To assemble the synthetic gene, each primer was added at a final concentration of 1 μM to the first PCR reaction with Pl Taq HF DNA polymerase under standard conditions. The following thermocycler program was used for the first assembly reaction: 95 ^{o}C for 5 min; 95 ^{o}C for 30 s; 55 ^{o}C for 2 min; 72 ^{o}C for 10 s; 40 cycles of 95 ^{o}C for 15 s, 55 ^{o}C for 30 s, 72 ^{o}C for 20 s plus 3 s/cycle; these cycles were followed by a final incubation at 72 ^{o}C for 5 min. The second assembly reaction (50 μL) utilized 16 μL of the unpurified first PCR reaction as template with standard reagents for Pt Taq HF. The thermocycler program for the second PCR was: 95^{o}C for 30 s; 55 ^{o}C for 2 min; 72^{o}C for 10 s; 40 cycles of 95 ^{o}C for 15 s, 55 ^{o}C for 30 s, 72 ^{o}C for 80 s; these cycles were followed by a final incubation at 72 ^{o}C for 5 min. Again, the second PCR reaction (16 μL) was transferred into fresh reagents and run using the same program. The product DNA smear at the appropriate size was gel purified and used as a template for amplification (50 μL) with rescue primers (Amyco1GK F2 and Amyco1GK R62) using Phusion DNA polymerase. The resulting rescue product was inserted into pCR2.1-TOPO by TA cloning (Invitrogen) and verified by sequencing before subcloning into the NdeI-XbaI sites of pCWOri-HisN to add an N-terminal His_{10}-affinity tag.

Purification of His_{10}-Amyco1, His_{10}-Amyco2, and His_{10}-Amyco1GK from *E. coli* BL21(de3). TB (500 mL) containing carbenicillin (50 μg/mL) in a 2 L-baffled shake flask was inoculated to OD_{600} = 0.05 with an overnight TB culture of freshly-transformed *E. coli* BL21(de3) containing the appropriate expression plasmid. The cultures were grown at 37 °C at 200 rpm in a rotary shaker until OD_{600} = 0.7, at which time IPTG (0.5 mM) and α-aminolevulinic acid (65 μg/mL) were added. After inducing protein expression, the culture temperature was dropped to 25 °C for an additional 20 h. Cell pellets were harvested by centrifugation at 9,800 × *g* for 7 min and stored at -80 °C. Frozen cell pellets were thawed and resuspended at 5 mL/g cell paste with Buffer A (50 mM potassium phosphate, 300 mM sodium chloride, 10 mM imidazole, pH 8.0) supplemented with PMSF (0.5 mM), DTT (1 mM), and DNase (2 U/g cell paste). The cell paste was homogenized before lysis by passage through a

French Pressure cell (Thermo Scientific; Waltham, MA) at 14,000 psi. The lysate was centrifuged at 15,300 × *g* for 30 min at 4 °C to separate the soluble and insoluble fractions. The soluble lysate was passed over a Ni-NTA agarose column (Qiagen, 10 mL) at a flow rate of 2 mL/min on an ÄKTApurifier FPLC (GE Healthcare Life Sciences; Piscataway, NJ). Purification of the individual proteins was carried out as follows: (1) His_{10}-Amyco1: wash = 5 column volumes of Buffer A, elution = linear gradient of 0% to 100% Buffer B (50 mM potassium phosphate, 300 mM sodium chloride, 250 mM imidazole, 1 mM DTT, pH 8.0) over 30 column volumes followed by 5 column volumes of Buffer B, (2) His_{10}-Amyco2: wash = 5 column volumes of Buffer A followed by a linear gradient of 0% to 50% Buffer B over 20 column volumes, elution = 5 column volumes of Buffer B, (3) His_{10}-Amyco1GK: wash = 20 column volumes of Buffer A, elution = 10 column volumes Buffer B. For all three proteins, fractions were pooled based on the A_{280} elution profile and concentrated with an Amicon filtration device using a YM10 membrane (Millipore Corporation; Billerica, MA). The protein was then dialyzed 3×10^{-2} against 50 mM potassium phosphate, pH 7.0 using a 10 kD MWCO Spectra/Por cellulose ester dialysis membrane (Spectrum Laboratories; Rancho Dominguez, CA) for a 10^{-6} final dilution of the original buffer. His_{10}-Amyco1: k_{cat} = 0.86 s^{-1} (peroxidase) and 1,020 s^{-1} (catalase) for heme occupancy of 0.26. His_{10}-Amyco2: k_{cat} = not detectable (peroxidase) and 24,800 s^{-1} (catalase) for heme occupancy of 0.26. The concentrations of His_{10}-Amyco1, His_{10}-Amyco2, and His_{10}-Amyco1GK were calculated using the calculated extinction coefficients at 280 nm for their respective amino acid sequences (His_{10}-Amyco1, 157,790 $M^{-1}cm^{-1}$; His_{10}-Amyco2, 94,770 $M^{-1}cm^{-1}$; His_{10}-Amyco1GK, 172,230 $M^{-1}cm^{-1}$). The protein was stored at -80 °C after addition of glycerol (10% (v/v) final concentration).

Purification of GA-Amyco1 from *E. coli* RP523. TB (500 mL) containing carbenicillin (50 μg/mL) and hemin chloride (30 μg/mL) in a 2 L-baffled shake flask was inoculated to OD_{600} = 0.05 with an overnight TB culture of freshly-transformed *E. coli* RP523 containing pCWOri-His_{10}-TEV-Amyco1. Protein expression and Ni-NTA purification of His_{10}-TEV-Amyco1 was carried out as described above for His_{10}-Amyco1 except that hemin chloride (30 μg/mL) was added at induction in place of ALA. After concentration by Amicon filtration, the protein was incubated with His_{10}-TEV protease (1:100 wt TEV/wt substrate) during dialysis overnight against Buffer A supplemented with 1 mM DTT using a 10 kD MWCO Spectra/Por cellulose ester dialysis membrane. TEV and un-digested Amyco1 were removed by passing the mixture over a Ni-NTA agarose column (10 mL) and washing off unbound proteins with Buffer A (10 mL) containing DTT (1 mM). GA-Amyco1 was dialyzed 3×10^{-2} against 20 mM Tris, 50 mM potassium chloride, pH 7.5 using a 10 kD MWCO Spectra/Por cellulose ester dialysis membrane for a 10^{-6} final dilution of the original buffer. The protein was stored at -80 °C after addition of glycerol (10% (v/v) final concentration).

pH-rate profile for His_{10}-Amyco1. The pH-rate profile for the peroxidase activity of His_{10}-Amyco1 was determined at 25 °C using the 2,4-DCP oxidation assay with a variation of Britton and Robinson buffers in place of the potassium phosphate buffer [*12*]. For all samples, the buffer system contained boric acid (50 mM), sodium acetate (50 mM), and potassium phosphate (50 mM) and were then adjusted to the desired pH every half pH unit between 3.5 and 9.0. The final assay contained His_{10}-Amyco1 (0.5 nmol), H_2O_2 (4 mM), 2,4-DCP (3 mM), and 4-AAP (0.16 mM) in a total assay volume of 1 mL.

ICP-OES analysis. Serial dilutions of stock solutions (1,000 mg/L) in ddH_2O were used to prepare six solutions between 50 μg/L and 1,000 μg/L to generate a standard curve for ferric

nitrate and manganese sulfate, respectively. His_{10}-Amyco1 and His_{10}-Amyco2 were dialyzed (3×10^{-2} for a 10^{-6} final dilution of the original buffer) for 24 h in 10 mM sodium sulfate, pH 7.0 supplemented with EDTA (1 mM) to chelate excess metals in solution. GA-Amyco1 was dialyzed against 20 mM Tris, 50 mM potassium phosphate, 1 mM EDTA, pH 7.5. The proteins were diluted to a concentration intended to yield 400-600 μg metal/L (assuming 100% cofactor occupancy). Samples were analyzed on an Optima 7000 DV ICP-OES (Perkin Elmer; Fremont, CA) using argon as the carrier gas at a sample flow rate of 1 mL/min (RF power, 1200 watts; plasma gas flow, 15 L/min; auxiliary gas flow; 0.2 L/min; nebulizer gas flow, 0.8 L/min). His_{10}-Amyco1 and GA-Amyco1 were analyzed solely for the presence of Fe. His_{10}-Amyco2 was analyzed for the presence of both Fe and Mn. From these studies, the extinction coefficients of the Soret bands for GA-Amyco1 and His_{10}-Amyco2 were determined to be 117,300 $M^{-1}cm^{-1}$ at 406 nm and 211,700 $M^{-1}cm^{-1}$ at 413 nm, respectively.

Spectrochemical redox titration of GA-Amyco1. GA-Amyco1 purified from *E. coli* RP523 was concentrated to ~25 mg/mL in a total volume of 0.5 mL. The protein was oxidized to the aquo-ferryl species by addition of potassium ferricyanide (stock solution, 100 mM) to a final concentration of 5 mM followed by incubation on ice for 20 min. The potassium ferricyanide was removed from the oxidized protein by passage over a disposable PD-10 desalting column (GE Healthcare; Piscataway, NJ) equilibrated with 100 mM potassium phosphate, 50 mM potassium chloride, pH 7.0. In a 3.5 mL cuvette adapted to fit a pH 1100 Series potentiometer (Oakton; Vernon Hills, IL), the protein, a stir bar, and 35 μL of a mixture of redox mediators (10 μM each; methyl viologen, −440 mV; anthraquinone-2,6-disulfonic acid, −184 mV; 2-hydroxy-1,4-naphthoquinone, −137 mV; 2,5-dihydroxy-1,4-benzoquinone, −60 mV; tetramethyl-*p*-benzochinon (Duroquinone), 5 mV; 1,2-naphthoquinone, 157 mV; ferricyanide, 356 mV) and buffer were added to a total volume of 3.5 mL. The reaction stirred at room temperature and purged of air by sparging with water-saturated Ar for 1 h. An oxidized spectrum was taken and the entire mixture was fully reduced to the ferrous state with sodium dithionite (~25-30 μL of a 100 mM stock) to remove remaining O_2. The mixture was once again oxidized with the addition of potassium ferricyanide (4 μL of a 100 mM stock solution). This was then reduced by step-wise additions of 1-5 μL aliquots of sodium dithionite (10 mM), allowing the potential to equilibrate, and collecting the protein absorbance data from 500-750 nm. The fraction of reduced Amyco1 was monitored by the $\Delta A_{561\ nm}$ and plotted against the potential vs SHE to determine midpoint reduction potential by fitting the curve to the following equation [*13*]:

$$Y = \frac{e^{-96500\times(m_0-m_1)/2477.572}}{1+e^{-96500\times(m_0-m_1)/2477.572}}$$

Purification and reconstitution of GA-Amyco1 from *E. coli* BL21(de3) for stopped-flow kinetic studies. Expression and purification of His_{10}-TEV-Amyco1, TEV digestion, and isolation and dialysis of the GA-Amyco1 protein was carried out as described above for *E. coli* RP523. GA-Amyco1 purified from BL21(de3) (heme occupancy, 0.13) was then reconstituted by addition of a 10-fold molar excess of hemin chloride (6.5 mg/mL in DMSO) followed by overnight incubation at 4 °C on an orbital shaker. Unbound hemin was removed by passing the mixture over DEAE sepharose (Sigma-Aldrich, 10 mL), washing with 50 mL of 20 mM Tris, 10 mM sodium chloride, pH 7.5, and elution with 20 mM Tris, 400 mM sodium chloride, pH 7.5. To remove residual hemin and soluble GA-Amyco1 aggregates, the reconstituted GA-Amyco1 was further purified using a HiLoad 16/20 Superdex 200 prep grade size exclusion

chromatography column (GE Healthcare Life Sciences; Piscataway, NJ) equilibrated with 20 mM Tris, 10 mM potassium chloride, pH 7.5. This procedure yielded quantitative reconstitution of the purified GA-Amyco1 protein according to the Soret band absorbance (peroxidase, k_{cat} = 3.3 s^{-1}; catalase, k_{cat} = 6,080 s^{-1}).

Stopped-flow optical studies of GA-Amyco1. The reaction of heme-reconstituted GA-Amyco1 purified from *E. coli* BL21(de3) with H_2O_2 was characterized by stopped flow UV-visible spectroscopy on a Hi-Tech Scientific double mixing stopped-flow SF-61DX2 system equipped with a diode array detector and KinetAsyst software (Hi-Tech Scientific; Bradford-on-Avon, United Kingdom) using a modified literature protocol [*14*]. The heme-reconstituted GA-Amyco1 was first exchanged by dialysis into 50 mM potassium phosphate buffer, pH 7.0 and loaded into syringe A (concentration, 20 μM). After loading H_2O_2 (5 mM) into syringe B, the system was allowed to equilibrate to 5 °C. GA-Amyco1 was then rapidly mixed with H_2O_2 to yield a final protein concentration of 10 μM and H_2O_2 concentration of 2.5 mM. Spectra (300 total scans) were logarithmically acquired from 300-700 nm as a function of time over 73.8 seconds after an initial delay period of 75.8 ms. The identity of the first species was assigned as a compound-I state based on observed spectral shifts and is proposed to decay to a compound II-like state based on previous work with heme-reconstituted KatG, which does not initially contain the M-Y-W adduct [*14-16*]. At 25 ºC, the production of the initial oxidized species (I) occurred before the first time point (8 ms) and only the formation of decay of the compound II-like state (II) was detected. Upon cooling to 5 ºC, the rise and subsequent decay of both compound I-like and II-like states could be visualized.

Synthesis of veratylglycerol-β-guaiacol ether (1). Lignin model dimer **1** was synthesized using a modified literature preparation [*17*]. Briefly, potassium iodide (0.82 g, 4.93 mmol) and potassium carbonate (6.81 g, 49.3 mmol) were added to 2-methoxyphenol (5.03 g, 40.5 mmol) in acetone (75 mL). After slow addition of ethyl bromoacetate (9.81 g, 58.7 mmol) over a minute, the reaction was refluxed overnight at 75 °C in air. When the reaction was complete by TLC, work up following the literature procedure yielded ethyl(2-methoxy-phenoxy) acetate (7.78 g, 91% yield). Ethyl(2-methoxy-phenoxy) acetate (2 g, 9.51 mmol) was then condensed with 3,4-dimethoxybenzaldehyde (1.58 g, 9.51 mmol) using LDA generated *in situ* (2.61 mL of diisopropylamine and 4.5 mL of 2.5 M butyllithium solution in hexanes) under N_2 and worked up accordingly. The purified product (1.67 g, 47% yield) was obtained after flash chromatography (silica, 1:2 ethyl acetaete:hexanes). The condensation product (1.65 g, 4.38 mmol) was reduced by sodium borohydride (1.66 g, 43.8 mmol) in a 3:1 THF:H_2O mixture (48 mL) overnight to yield **1** (1.37 g, 94% yield) as a viscous, light yellow oil with an overall yield of 40%. ^{1}H-NMR (400 MHz, $CDCl_3$, 25°C) δ (ppm): (both *erythro* and *threo* diasteromers); 3.420-3.938 (2H, m, γ-CH_2-), 3.866 (3H, s, -OCH_3), 3.872 (3H, s, -OCH_3), 3.880 (3H, s, -OCH_3), 3.980-4.156 (1H, m, *β*-CH-), 4.962 (1H, d, $J_{\alpha\beta}$ = 5.6 Hz, α-CH-), 6.822-7.119 (7H, Ar-H). HRESI-MS: calculated for [M+Na^+] 357.132 Da, found 357.1319 Da

Synthesis of guaiacylglycerol-β-guaiacol ether (2). Model compound **2** was synthesized from the ethyl(2-methoxy-phenoxy) acetate intermediate (2 g, 9.51 mmol) by condensing with benzylvanillin (2.3 g, 9.51 mmol) using LDA generated *in situ* (2.61 mL of diisopropylamine and 4.5 mL of 2.5 M butyllithium solution in hexanes) according to the literature protocol [*17*]. The purified product (1.72 g, 40% yield) was obtained after flash chromatography (silica, 1:1 ethyl acetaete:hexanes) and reduced with sodium borohydride (1.42 g, 37.4 mmol) in a 3:1 THF:H_2O mixture (40 mL) overnight at quantitative yield (1.53 g, 100% yield). Overnight

deprotection of the reduced product (1.53 g, 3.73 mmol) with 10% Pd/C (0.91 g) in methanol (24 mL) under an atmosphere of H_2 yielded lignin model dimer **2** (1.05 g, 88% yield) as a slightly gray oil at an overall yield of 32%. ^{1}H-NMR (400 MHz, $CDCl_3$, 25°C) δ (ppm): (both *erythro* and *threo* diasteromers); 3.465-3.939 (2H, m, γ-CH-), 3.890 (3H, s, -OCH_3), 3.919 (3H, s, -OCH_3), 4.005-4.152 (1H, m, *β*-CH-), 4.974 (1H, d, $J_{\alpha\beta}$ = 5.2 Hz, α-CH-), 5.615 (1H, s, Ar-OH), 6.817-7.068 (7H, Ar-H). HRESI-MS: calculated for [M+Na^+] 343.116 Da, found 343.1158 Da.

Assay for monitoring degradation of lignin model compounds. Reactions of Amyco1 (~2 nmol) with lignin model compounds **1** and **2** (1 mM) were carried out in 50 mM sodium acetate, pH 5.0 with ABTS (0.5 mM) added as a redox mediator. All reactions were initiated with addition of H_2O_2 (100 mM final concentration). Reactions were allowed to stir overnight at room temperature after which any precipitate formed overnight was removed by centrifugation at 20,817 × *g* for 5 min, and the cleared solution was analyzed by RP-HPLC on a 1200 Series HPLC coupled to a G1315D diode array detector (Agilent; Santa Clara, CA).

*Analysis of reactivity with **1**.* Samples (20 µL) were chromatographed on an Eclipse XDB-C_{18} column (5 µm, 9.4 × 250 mm; Zorbax; Santa Clara, CA) using 5% acetic acid (A) and acetonitrile as the mobile phase (0.5 mL/min) using the following gradient: 0-2.5 min, hold at 25% B; 2.5-6.0 min, linear gradient from 25%-75% B; 6.0-12.0 min, hold at 75% B; 12.0-15.0 min, linear gradient from 75%-25% B.

*Analysis of reactivity with **2**.* Samples (20 µL) were chromatographed on an Eclipse XDB-C_{18} column (3.5 µm, 3.0 × 150 mm; Zorbax; Santa Clara, CA) with 5% acetic acid and acetonitrile (35%) as the mobile phase at 1.0 mL/min over 35 min.

3.3. Results and Discussion

With the interest of analyzing the extracellular oxidative proteome, we carried out a functional secretome analysis of *Amycolatopsis* sp. 75iv2. The extracellular peroxidase activity was tracked over 10 days in order to find the peak production timepoints in cultures grown in the presence of *Miscanthus giganteus* lignocellulose (*Figure 3.1A*). This peroxidase activity did not appear to be lignin-dependent and that extracellular peroxidase activity was consistently present in the secretome. Nonetheless, on days of peak extracellular activity, the secretome was concentrated and separated on and SDS-PAGE gel, and a heme-staining indicated the strong presence of a heme-containing doublet. The heme-staining was followed by silver-staining and indicated that the secreted protein profile was relatively simple and that one- dimensional separation was sufficient to isolate tryptic digests of the heme-stained protein bands. Mass spectrometry of the tryptic digests confirmed the presence of two major extracellular hemoproteins, which were named Amyco1 and Amyco2 (*Figure 3.1B-D*). A full proteomic analysis of the secretome also indicated that Amyco1 and Amyco2 were the two most likely extracellular heme protein candidates based on their relatively high abundance. It also indicated that laccases or dye decolorizing peroxidases (DyPs) were not detected in the extracellular protein fraction under these growth conditions (*Table 3.1*).

The genes encoding Amyco1 and Amyco2 were cloned from *A.* sp. 75iv2 genomic DNA, heterologously expressed in *Escherichia coli*, and purified to homogeneity (*Figures 3.2* and *3.3*).

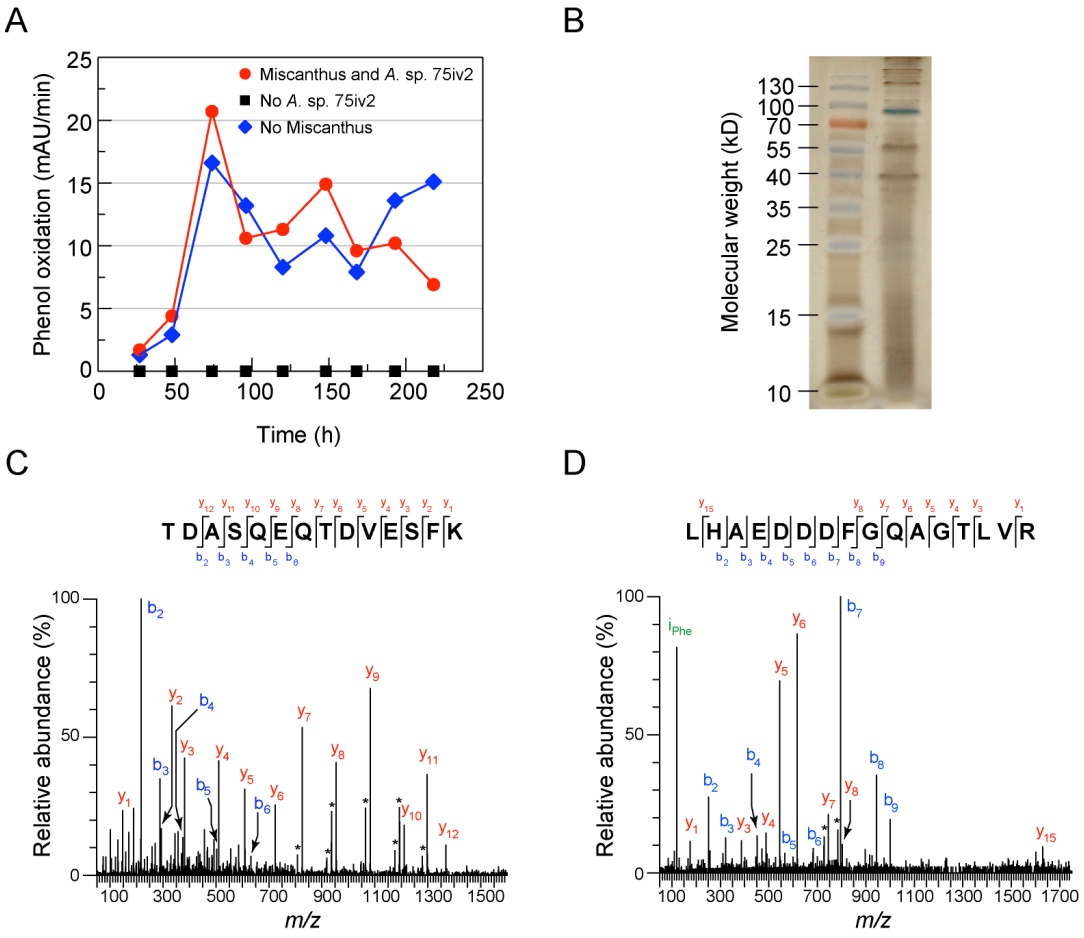


Figure 3.1. *(A) Time course for A. sp. 75iv2 growth monitoring extracellular peroxidase activity. (B) Overlaid heme- and silver-stained SDS-PAGE gel showing the extracellular protein fraction at 3 d. (C) Representative tandem mass spectrum of a tryptic peptide derived from Amyco1, a heme peroxidase identified in the secretome of A. sp. 75iv2. (D) Representative tandem mass spectrum of a tryptic peptide derived from the Amyco2 catalase.*

Table 3.1. *List of extracellular Amycolatopsis sp. 75iv2 proteins prepared from cultures grown in the presence of Miscanthus giganteus lignocellulose identified by peptide fingerprinting.*

Number of peptides	Protein ID
15	TAP domain protein [*Streptomyces* sp. AA4]
13	β-Glucosidase-like glycosyl hydrolase [*Saccharomonospora viridis* DSM 4301]
11	Dihydrolipoamide dehydrogenase [*Streptomyces albus* J1074]
11	Glutamine synthetase [*Streptomyces* sp. AA4]
11	Peptidase S8 and S53 subtilisin kexin sedolisin [*Streptomyces* sp. AA4]
11	Transaldolase [*Streptomyces* sp. AA4]
10	Lysozyme precursor [*Streptomyces* sp. AA4]
10	Dihydrolipoamide dehydrogenase [*Streptomyces* sp. AA4]
10	Ketol-acid reductoisomerase [*Actinosynnema mirum* DSM 43827]
10	ABC transporter ligand-binding protein [*Streptomyces* sp. AA4]
10	Putative γ-glutamyltranspeptidase [*Streptomyces roseosporus*]
9	Extracellular solute-binding protein [*Streptomyces* sp. AA4]
9	Chitinase [*Streptomyces* sp. AA4]
9	**Catalase [*Saccharopolyspora erythraea* NRRL 2338]**
8	Phosphoenolpyruvate carboxykinase (GTP) [*Saccharomonospora viridis*]
7	2-Isopropylmalate synthase [*Saccharomonospora viridis* DSM 43017]
7	Ribose-phosphate pyrophosphokinase [*Streptomyces* sp. AA4]
7	Lysozyme M1 precursor [*Streptomyces* sp. AA4]
6	α-1,2-Mannosidase [*Streptomyces* sp. AA4]
6	Malate synthase [*Streptomyces* sp. AA4]
6	Citrate synthase I [*Streptomyces* sp. AA4]
6	L-Rhamnose isomerase [*Actinosynnema mirum* DSM 43827]
6	Membrane alanyl aminopeptidase [*Streptomyces* sp. AA4]
5	δ-1-Pyrroline-5-carboxylate dehydrogenase [*Actinosynnema mirum*]
5	Secreted trypsin-like serine protease [*Streptomyces* sp. AA4]
5	**Catalase/peroxidase HPI [*Micromonospora* sp. ATCC 39149]**
5	β-N-Acetylhexosaminidase [*Streptomyces* sp. AA4]
5	Transduction system [*Saccharomonospora viridis* DSM 43017]
5	Ricin B lectin [*Chitinophaga pinensis* DSM 2588] (E=2e-14)
4	Oxidoreductase [*Mycobacterium kansasii* ATCC 12478]
4	Aminopeptidase N [*Saccharomonospora viridis* DSM 43017]
4	Hypothetical protein StAA4_35073 [*Streptomyces* sp. AA4]
4	Hypothetical protein StAA4_09415 [*Streptomyces* sp. AA4]
4	Desuccinylase-like deacylase [*Saccharomonospora viridis* DSM 43017]
4	β-Lactamase [*Streptomyces* sp. AA4]
3	5'-nucleotidase-like protein [*Streptomyces* sp. AA4]
3	Hypothetical protein RER_59650 [*Rhodococcus erythropolis* PR4]
3	Glucose-6-phosphate isomerase [*Streptomyces* sp. AA4]
3	Isocitrate dehydrogenase [*Streptomyces* sp. AA4]
3	Endoribonuclease L-PSP [*Streptomyces* sp. AA4]
3	Hypothetical protein StAA4_15790 [*Streptomyces* sp. AA4]
3	Single-strand binding protein [*Saccharomonospora viridis* DSM 43017]
3	Amidohydrolase [*Streptomyces* sp. AA4]
3	Phosphoglycerate kinase [*Saccharomonospora viridis* DSM 43017]
3	Dipeptidyl aminopeptidase/acylaminoacyl peptidase [*Cellulomonas flavigena*]
3	Extracellular solute-binding protein family 5 [*Streptomyces* sp. AA4]
3	Periplasmic binding protein [*Streptomyces* sp. AA4]
3	Secreted chitin binding protein [*Streptomyces coelicolor* A3(2)]
3	Aconitate hydratase [*Streptomyces* sp. AA4]
3	Glycine dehydrogenase [*Streptomyces* sp. AA4]
3	Putative lipoprotein [*Streptomyces scabiei* 87.22]
2	Two-component transcriptional regulator, winged helix family [*Kribbella flavida*]

2	Dipeptidyl aminopeptidase/acylaminoacyl peptidase [*Saccharomonospora viridis*]
2	ATP synthase F1 subcomplex (α) [*Saccharomonospora viridis*]
2	Putative carboxymuconolactone dehydrogenase family protein [*Streptomyces* sp. AA4]
2	Probable enoyl-CoA hydratase [*Streptomyces* sp. AA4]
2	Amidase [*Streptomyces* sp. AA4]
2	Chlorite dismutase [*Streptomyces* sp. AA4]
2	Hypothetical protein Svir_36020 [*Saccharomonospora viridis* DSM 43017]
2	Sugar transporter sugar-binding protein [*Streptomyces coelicolor* A3(2)]
2	Glucose/sorbosone dehydrogenase [*Saccharomonospora viridis* DSM 43017]
2	Phosphoserine aminotransferase, putative [*Saccharomonospora viridis* DSM 43017]
2	Extracellular solute-binding protein family 3 [*Streptomyces* sp. AA4]
2	Transaldolase [*Streptomyces* sp. AA4]
2	Methanol:NDMA oxidoreductase [*Amycolatopsis methanolica*]
2	Flagellar basal body-associated protein FliL [*Saccharomonospora viridis*]
2	20S proteasome A and B subunits [*Streptomyces* sp. AA4]
2	Flavoprotein disulfide reductase [*Streptomyces* sp. AA4]
2	β-Lactamase domain-containing protein [*Stackebrandtia nassauensis*]
2	Hypothetical protein Sros_1892 [*Streptosporangium roseum* DSM 43021]
2	Superoxide dismutase [Fe-Zn] 1 (FeSOD I) [*Streptomyces* sp. AA4]
1	Vitamin B12-dependent ribonucleotide reductase [*Streptomyces* sp. AA4]
1	Hypothetical protein Amir_4530 [*Actinosynnema mirum* DSM 43827]
1	Glyoxalase/bleomycin resistance protein/dioxygenase [*Arthrobacter chlorophenolicus* A6]
1	RraA family protein [*Saccharomonospora viridis* DSM 43017]
1	Hypothetical protein Svir_27240 [*Saccharomonospora viridis* DSM 43017]
1	Glyoxalase family protein [*Frankia alni* ACN14a]
1	Glu/Leu/Phe/Val dehydrogenase dimerisation region [*Streptomyces* sp. AA4]
1	UDP-galactopyranose mutase [*Saccharomonospora viridis* DSM 43017]
1	Dihydrodipicolinate reductase [*Streptomyces* sp. AA4]
1	Hypothetical protein StAA4_36786 [Streptomyces sp. AA4]
1	Homogentisate-1,2-dioxygenase [*Streptomyces* sp. AA4]
1	ATP synthase, F1 (β) [*Saccharomonospora viridis* DSM 43017]
1	Uroporphyrin-III C/tetrapyrrole (Corrin/Porphyrin)
1	Putative deacetylase [*Streptosporangium roseum* DSM 43021]
1	DNA-directed RNA polymerase (α) [*Streptomyces* sp. AA4]
1	*S*-adenosyl-L-homocysteine hydrolase [*Streptomyces* sp. AA4]
1	Aldose-1-epimerase [*Streptomyces* sp. AA4]
1	Fructose-bisphosphate aldolase [*Saccharomonospora viridis* DSM 43017]
1	Hypothetical protein ELI_09840 [*Erythrobacter litoralis* HTCC2594]
1	Oxidoreductase [*Rhodococcus erythropolis* PR4]
1	Lipase [*Streptomyces avermitilis* MA-4680]
1	N-succinyldiaminopimelate aminotransferase [*Streptomyces* sp. AA4]
1	Metal-dependent hydrolase, β-lactamase superfamily III [*Saccharomonospora viridis*]
1	Conserved hypothetical protein [*Streptomyces* sp. Mg1]
1	Succinate dehydrogenase subunit A [*Saccharomonospora viridis* DSM 43017]
1	Putative hydrolase [*Streptomyces* sp. AA4]
1	Aminopeptidase Y [*Saccharomonospora viridis* DSM 43017]
1	Putative aldehyde dehydrogenase [*Nocardia farcinica* IFM 10152]
1	Carbon monoxide dehydrogenase, medium chain [*Streptomyces* sp. AA4]
1	Permease [*Streptomyces sviceus* ATCC 29083]

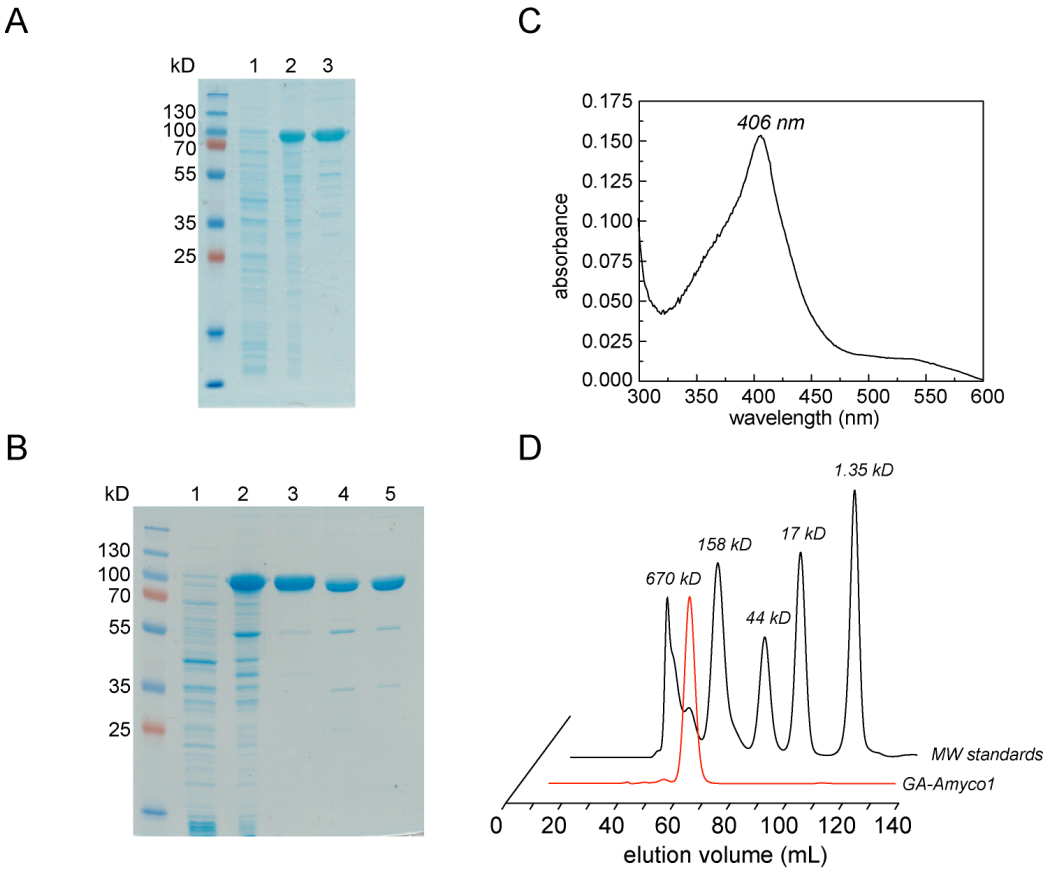


Figure 3.2. Purification and characterization of Amyco1. *(A) SDS-PAGE analysis of His_{10}-Amyco1 purification. Pre-induction sample (lane 1), post-induction sample (lane 2), Ni-NTA column eluant (lane3). (B) SDS-PAGE analysis of GA-Amyco1 purification. Pre-induction sample (lane 1), post-induction sample (lane 2), Ni-NTA column eluant (lane 3), GA-Amyco1 after treatment with TEV protease to remove His_{10}-tag (lane 4), flow-through GA-Amyco1 from the second Ni-NTA column (lane 5). (C) UV-visible spectrum of purified Amyco1.(D) Size-exclusion chromatography of GA-Amyco1. The apparent molecular weight of GA-Amyco1 was calculated to be 194 kD based on a standard curve generated from molecular weight standards (Bio-RAD) and is consistent with the formation of a dimer ($MW_{monomer}$ = 83.7 kD).*

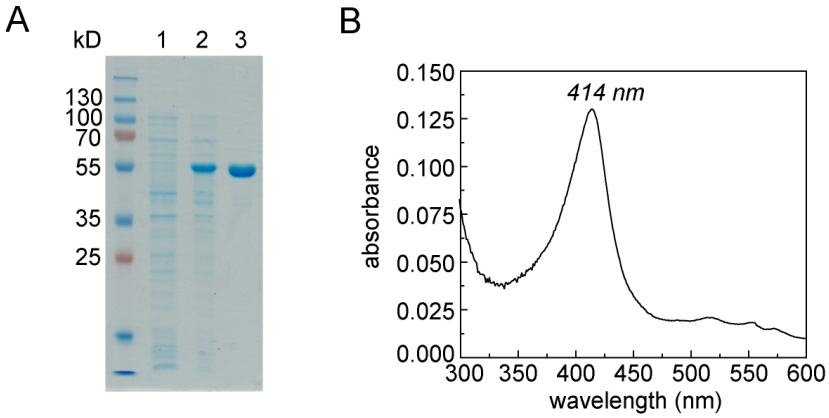
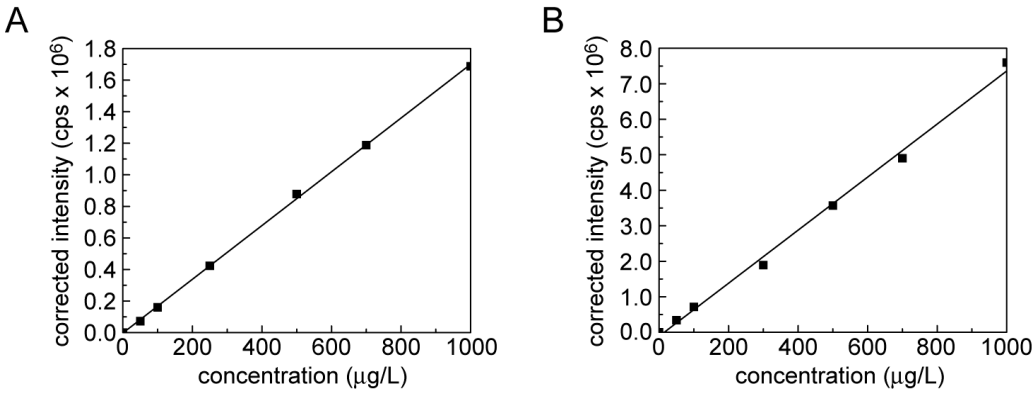


Figure 3.3. Purification and characterization of His_{10}-Amyco2. *(A) SDS-PAGE analysis of His_{10}-Amyco2 purification. Pre-induction sample (lane 1), post-induction sample (lane 2), Ni-NTA column eluant (lane 3). (B) UV-visible spectrum of purified His_{10}-Amyco2.*



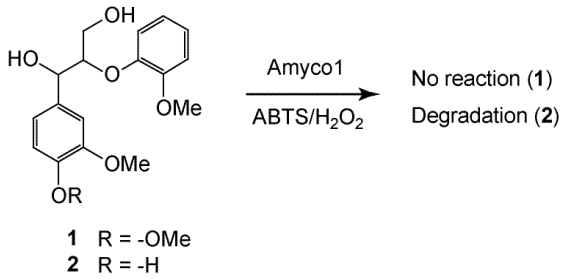
C

	strain	Fe	Mn
His_{10}-Amyco1	BL21(de3)	0.24	-
His_{10}-Amyco2	BL21(de3)	0.26	not detected
GA-Amyco1	RP523	0.46	-

Figure 3.4. *ICP-OES analysis of Amyco1 and Amyco2. Standard curves for (A) Fe and (B) Mn. (C) Metal ion equivalents measured per protein monomer.*

Further biochemical characterization indicates that Amyco1 exists as a homodimer based on size-exclusion chromatography, which measures an apparent molecular mass of 194 kD (*Figure 3.2D*). Characterization by UV-visible spectroscopy and inductively coupled plasma-optical emission spectrometry (ICP-OES) showed that both proteins contain an iron-porphyrin cofactor (*Figure 3.4*). The functions of Amyco1 and Amyco2 were assigned respectively as bifunctional catalase-peroxidase and catalase based on sequence homology and were validated using *in vitro* biochemical assays. As a catalase, Amyco2 may be involved in protection against oxidative damage to the host organism. Indeed, the white-rot fungus, *Phanerochaete chrysosporium*, demonstrates among the highest activities for peroxide-activated lignin degradation and continually secretes catalases during cell growth [*18*].

Based on the phenol oxidation reactivity of Amyco1, we carried out additional experiments for its characterization. Reactions of purified Amyco1 with synthetic lignin model dimers **1** and **2** (*Scheme 3.1*) indicate that Amyco1 is competent to carry out degradation of lignin at unprotected phenolic sites but not fully-protected methoxylated sites. Furthermore, it was found that the presence of the radical mediator, ABTS, greatly improved the reactivity of Amyco1 with



Scheme 3.1. *Reactivity of Amyco1 with lignin model dimers **1** and **2***

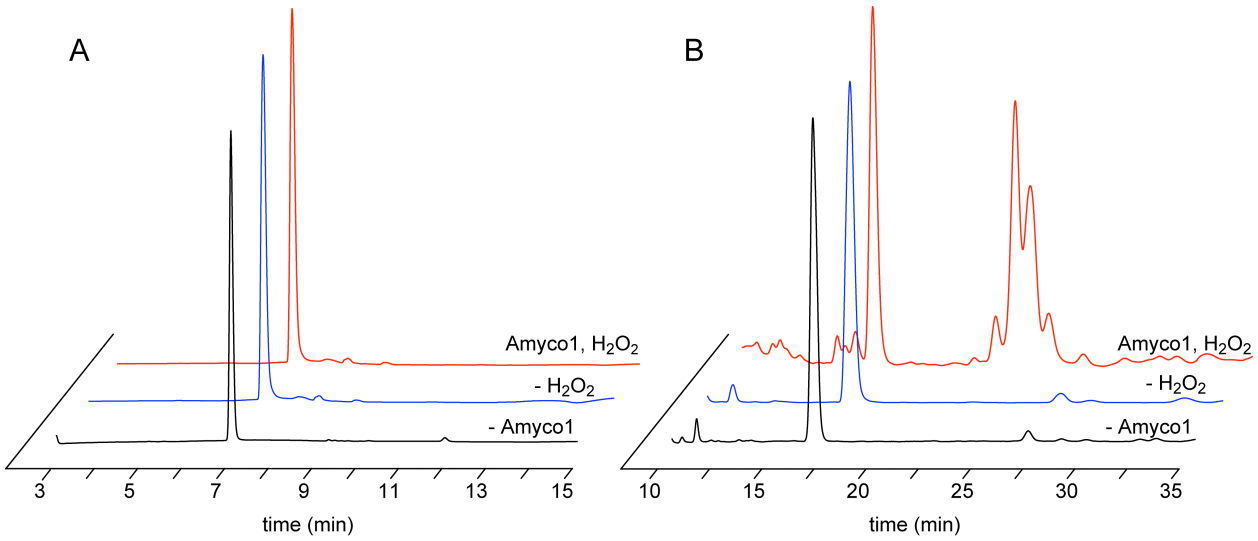


Figure 3.5. *Reaction of His_{10}-Amyco1 with synthetic lignin model dimers and **1** (A) and **2** (B) monitored by RP-HPLC at $A_{254\ nm}$..*

the phenolic model dimmers. The presence of the ABTS did not induce a reaction with the model dimer **2**. Separation of the reaction mixture by reversed-phase-HPLC revealed a mixture of products (*Figure 3.5*).

Additional studies showed that the pH-rate profile for Amyco1 was centered at pH 5, which is consistent with the lower pH of soil in the presence of organic acids produced by decaying plant matter as well as the low pH optimums for fungal lignin peroxidases (*Figure 3.6A*) [*19*]. Using a spectrochemical redox titration, the enzyme was chemically reduced stepwise under water-saturated argon to monitor its midpoint reduction potential for the Fe(II)-Fe(III) couple. It was discovered that the midpoint reduction potential of Amyco1 (-171 mV) was found to fall within the normal range of most other heme peroxidases (*Figure 3.7BC*) [*20, 21*]. Heme peroxidases with unusually oxidative Fe(III)-(FeIII) couples happen to be the LiPs and MnPs, with potentials of approximately -120 and -90 mV, respectively [*22*]. Their high oxidative capacity has been attributed to these values [*20*].

To further biochemically characterize Amyco1, rapid mixing of heme-reconstituted Amyco1 with excess H_2O_2 captured the formation of two species by stopped-flow UV-visible spectroscopy whose spectral characteristics [*14-16*] led to their assignment as the two- and one-

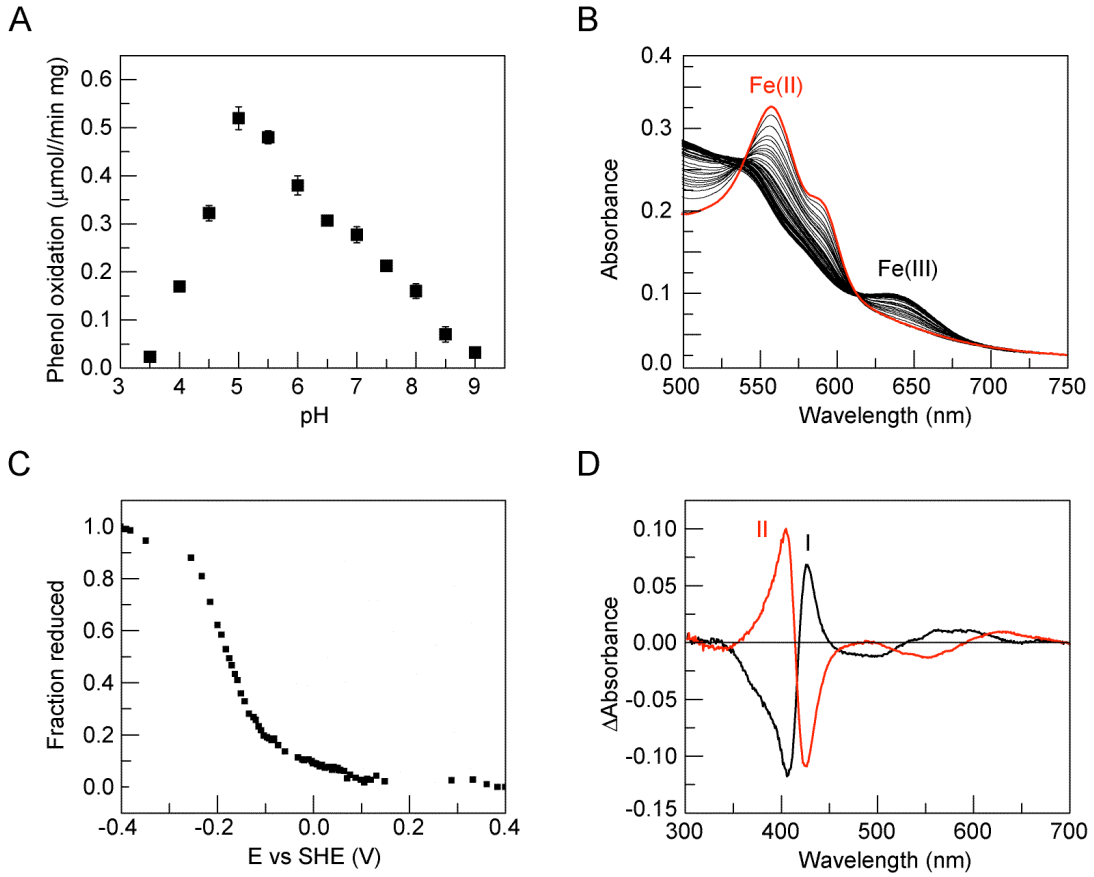


Figure 3.6. *Biochemical characterization of Amyco1. (A) pH-rate profile for Amyco1. Data are mean ± s.d. (n = 3). (B) Changes in the electronic absorption spectrum indicating stepwise reduction of the ferric state to the ferrous state during redox titration. (C) Spectrochemical redox titration curve for the one-electron reduction of Fe(III)-Amyco1. The fraction of enzyme reduced was calculated based on* $DA_{561\ nm}$*. (D) Difference spectra indicating formation of species assigned to compound I- and compound II-like states formed upon rapid mixing with H2O2 at 5ºC.*

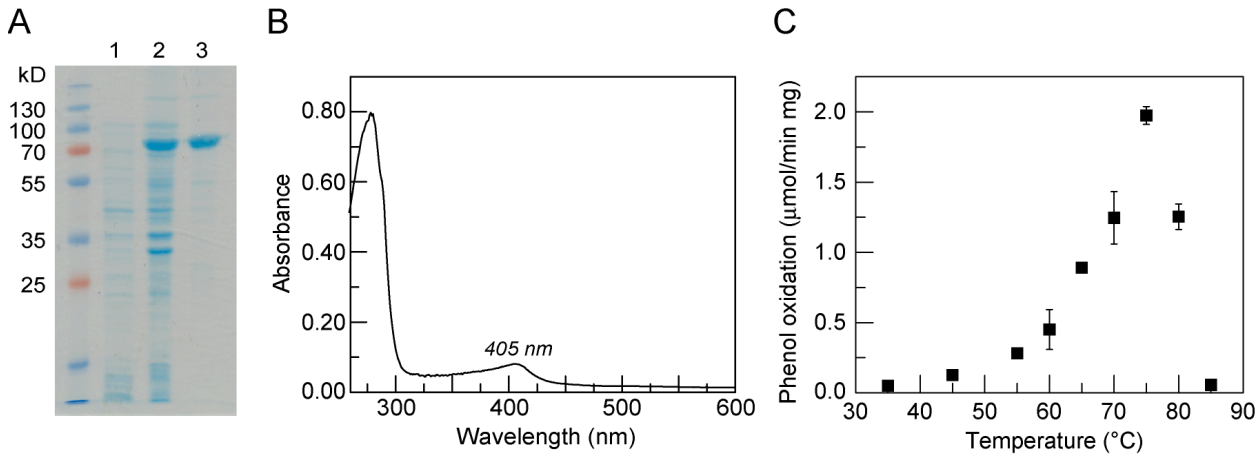


Figure 3.7. *Purification and characterization of His_{10}-Amyco1GK. (A) SDS-PAGE analysis of His_{10}-Amyco1GK purification. Pre-induction sample (lane 1), post-induction sample (lane 2), Ni-NTA column eluant (lane 3). (B) UV-visible spectrum of purified His_{10}-Amyco1GK. (C) Temperature-rate profile for an Amyco1 ortholog from Geobacillus kaustophilus. Data are mean ± s.d. (n = 3).*

electron oxidized compound I- and compound II-like states, respectively (*Figure 3.6D*), which supports a model in which each Amyco1 active site is competent to oxidize two equivalents of phenol per reaction cycle. Thus, this biochemical characterization revealed the catalase-peroxidase from *A.* sp. 75iv2 to encompass attributes very characteristic of this class of enzymes.

Finally, a homology search through sequenced metagenomes and genomes revealed that closely-related Amyco1 orthologs are widely distributed through bacteria and archaea. The gene of an Amyco1 ortholog from the thermophilic *Geobacillus kaustophilus* was synthesized, heterologously expressed in *E. coli*, and purified. The activity of this Amyco1 ortholog from *G. kaustophilus*, AmycoGK, was confirmed by biochemical characterization of the enzyme, which showed similar activity to Amyco1, but also a temperature optimum near the growth temperature of the host organism at 75 °C (*Figure 3.7*).

3.4. Conclusions

We have described a general pipeline to rapidly discover enzyme candidates involved in multi-step transformations of complex substrates in whole organisms. We have applied this strategy to studies of lignin reactivity and identified and characterized two new extracellular heme proteins from *Amycolatopsis* sp. 75iv2, a previously unsequenced soil bacterium valued for its ability to degrade lignocellulose and produce APPL. The two new extracellular proteins were identified by mass spectrometry to be a catalase and a catalase-peroxidase. Heterologous expression and purification of these enzymes allowed for biochemical characterization, showing the catalase-peroxidase, Amyco1, to be competent for phenol rather than aromatic ring oxidation. As high potential ring oxidation is expected to be necessary for efficient lignin degradation, it was concluded that Amyco1 is not the sole protein responsible for lignin modification by this organism; analysis of the genome reveals that the organism contains potential demethylases for lignin-related compounds [*23*], which could act synergistically for the degradation of biomass by uncapping new phenolic sites. Thus, with the uncapping of new phenols, Amyco1 has great potential to add to the overall lignolytic secretome. In addition to functionally identified candidates, we can now also begin to explore the function and reactivity of the full oxidative

system of *A.* sp. 75iv2, including laccases and DyP peroxidases, and their synergy with other enzyme families.

3.5. References

1. M. Ahmad, C. R. Taylor, D. Pink, K. Burton, D. Eastwood, G. D. Bending and T. D. H. Bugg, Development of novel assays for lignin degradation: Comparative analysis of bacterial and fungal lignin degraders, *Mol. BioSyst.* **2010**, *6*, 815-821.
2. M. Ahmad, J. N. Roberts, E. M. Hardiman, R. Singh, L. D. Eltis and T. D. H. Bugg, Identification of DypB from *Rhodococcus jostii* RHA1 as a lignin peroxidase, *Biochemistry* **2011**, *50*, 5096-5107.
3. D. L. Crawford, A. L. Pometto, III and R. L. Crawford, Lignin degradation by *Streptomyces viridosporus*: Isolation and characterization of a new polymeric lignin degradation intermediate, *Appl. Environ. Microbiol.* **1983**, *45*, 898-904.
4. M. Ramachandra, D. L. Crawford and G. Hertel, Characterization of an extracellular lignin peroxidase of the lignocellulolytic actinomycete *Streptomyces viridosporus*, *Appl. Env. Microbiol.* **1988**, *54*, 3057-3063.
5. L. Thomas and D. L. Crawford, Cloning of clustered *Streptomyces viridosporus* T7A lignocellulose catabolism genes encoding peroxidase and endoglucanase and their extracellular expression in *Pichia pastoris*, *Can. J. Microbiol.* **1998**, *44*, 364-372.
6. T. G. Pridham and D. Gottlieb, The utilization of carbon compounds by some Actinomycetales as an aid for species determination, *J. Bacteriol.* **1948**, *56*, 107-114.
7. H. Giroux, P. Vidal, J. Bouchard and F. Lamy, Degradation of kraft lignin by *Streptomyces viridosporus* and *Streptomyces badius*, *Appl. Environ. Microbiol.* **1988**, *54*, 3064-3070.
8. J. M. B. Macedo, L. M. F. Gottschalk and E. P. S. Bon, Lignin peroxidase and protease production by *Streptomyces viridosporus* T7A in the presence of calcium carbonate, *Appl. Biochem. Biotech.* **1999**, *79*, 735-744.
9. D. C. Yee, D. Jahng and T. K. Wood, Enhanced expression and hydrogen peroxide dependence of lignin peroxidase from *Streptomyces viridosporus* T7A, *Biotechnol. Prog.* **1996**, *12*, 40-46.
10. J. Visick and S. Clarke, RpoS- and OxyR-independent induction of HPI catalase at stationary phase in *Escherichia coli* and identification of *rpoS* mutations in common laboratory strains, *J. Bacteriol.* **1997**, *179*, 4158-4163.
11. P. E. Thomas, D. Ryan and W. Levin, An improved staining procedure for the detection of the peroxidase activity of cytochrome P-450 on sodium dodecyl sulfate polyacrylamide gels, *Anal. Biochem.* **1976**, *75*, 168-176.
12. F. Xu, W. Shin, S. H. Brown, J. A. Wahleithner, U. M. Sundaram and E. I. Solomon, A study of a series of recombinant fungal laccases and bilirubin oxidase that exhibit significant differences in redox potential, substrate specificity, and stability, *Biochim. Biophys. Acta - Prot. Struct. Mol. Enzymol.* **1996**, *1292*, 303-311.
13. P. L. Dutton, Redox potentiometry: Determination of midpoint potentials of oxidation-reduction components of biological electron-transfer systems, *Methods Enzymol.* **1978**, *54*, 411-435.

14. S. Chouchane, I. Lippai and R. S. Magliozzo, Catalase-peroxidase (*Mycobacterium tuberculosis* KatG) catalysis and isoniazid activation, *Biochemistry* **2000**, *39*, 9975-83.

15. S. Yu, S. Girotto, X. Zhao and R. S. Magliozzo, Rapid formation of compound II and a tyrosyl radical in the Y229F mutant of *Mycobacterium tuberculosis* catalase-peroxidase disrupts catalase but not peroxidase function, *J. Biol. Chem.* **2003**, *278*, 44121-44127.

16. R. A. Ghiladi, G. M. Knudsen, K. F. Medzihradszky and P. R. Ortiz de Montellano, The Met-Tyr-Trp cross-link in *Mycobacterium tuberculosis* catalase-peroxidase (KatG): Autocatalytic formation and effect on enzyme catalysis and spectroscopic properties, *J. Biol. Chem.* **2005**, *280*, 22651-63.

17. T. Katayama, F. Nakatsubo and T. Higuchi, Syntheses of arylglycerol-β-aryl ethers, *Mokuzai Gakkaishi* **1981**, *27*, 223-230.

18. S.-I. Kwon and A. J. Anderson, Catalase activities of *Phanerochaete chrysosporium* are not coordinately produced with ligninolytic metabolism: catalases from a white-rot fungus, *Curr. Microbiol.* **2001**, *42*, 8-11.

19. M. Tien, T. K. Kirk, C. Bull and J. A. Fee, Steady-state and transient-state kinetic studies on the oxidation of 3,4-dimethoxybenzyl alcohol catalyzed by the ligninase of *Phanerocheate chrysosporium* Burds, *J. Biol. Chem.* **1986**, *261*, 1687-1693.

20. C. D. Millis, D. Cai, M. T. Stankovich and M. Tien, Oxidation-reduction potentials and ionization states of extracellular peroxidases from the lignin-degrading fungus *Phanerochaete chrysosporium*, *Biochemistry* **1989**, *28*, 8484-8489.

21. B. H. Dunford, Heme Peroxidases, Wiley-VCH: Weinheim, Germany, **1999**, vol. pp. 532.

22. G. Battistuzzi, M. Bellei, C. A. Bortolotti and M. Sola, Redox properties of heme peroxidases, *Arch. Biochem. Biophys.* **2010**, *500*, 21-36.

23. J. B. Sutherland, Demethylation of veratrole by cytochrome P-450 in *Streptomyces setonii*, *Appl. Environ. Microbiol.* **1986**, *52*, 98-100.

Chapter 4: *Characterization of the dye-decolorizing peroxidases from Amycolatopsis sp. 75iv2*

Portions of this work were published in the following scientific journal:

M. E. Brown, T. Barros and M. C. Y. Chang, Identification and characterization of a multifunctional dye peroxidase from a lignin-reactive bacterium, *ACS Chem. Biol.* **2012**, *7*, 2074-2081

4.1. Introduction

Dye-decolorizing peroxidase (DyPs) are a relatively small but distinct class of heme peroxidases that are implicated in the extracellular and oxidative degradation of aromatics [*1-9*]. Crystallographic studies showed that the DyP from *Bjerkandera adusta* shared low structural and sequence homology with other known classes of peroxidases and also contained an unusual glutamate/aspirate as its distal active site residues [*10*]. After their discovery in 1999, DyPs have since been classified into four primary clades, comprising the bacterial (A-type and B-type) and fungal (the D-type) DyPs. In general, the A- and B-type enzymes are smaller and show lower activity in classic activity assays while the D-type enzymes are significantly larger and more active. In comparison, the remaining C-type branch is bacterial in origin but shares characteristics with the fungal D-type enzymes based on size and phylogenetic relationships.

Interestingly, some DyPs have been found to oxidize veratryl alcohol (VA) as well as the non-phenolic β-O-4 lignin model dimer in the absence of a mediator (*Scheme 1.2*) [*6*]. While this reactivity suggests that DyPs could be chemically competent to react with lignin, the physiological significance of these observations is unclear. The most convincing case for the involvement of DyPs in lignin breakdown comes from the soil bacterium, *Rhodococcus jostii* RHA1, whose genome contains two genes encoding an A-type (*dypA*) and B-type (*dypB*) DyP [*11*]. The generation of the respective deletion mutants, Δ*dypA* and Δ*dypB*, showed that the Δ*dypB* strain possessed significantly less lignolytic activity when compared to the wild-type strain [*12*]. *In vitro* studies further showed that heterologously-expressed DypB showed low MnP activity [*12*] and that DypB could possibly be exported from the cell using an encapsulin despite the absence of a canonical secretion signal [*13*].

Mining of the *Amycolatopsis* sp. 75iv2 genome reveals the presence of 3 *dyp* genes: *dyp*1 (C-type), *dyp*2 (C-type), and *dyp*3 (A-type). Interestingly, *dyp1* exists in a cluster of heme-containing proteins (*Figure 2.5*). DyP1-3 as well as the heme-containing cluster of which DyP1 is part, were expressed heterologously in *E. coli* and purified. DyP1 and DyP3 were partially characterized to have lower phenolic oxidation activity, though the low activity of DyP1 is attributed to the poor heme loading of the protein. Of the remaining heme proteins in the *dyp1* cluster, only expression of one out of the three candidates was successful. As the most phylogenetically distinct enzyme, we focused on further biochemical and structural characterization of DyP2. These studies showed that it had high peroxidase and manganese peroxidase activity, reminiscent of both the fungal DyPs as well a the versatile peroxidases (VPs) involved in lignin degradation. Furthermore, it was found to also possess a second mode of manganese-dependent oxidase reactivity that expands its substrate range to include non-phenolic compounds.

4.2. Materials and Methods

Commercial materials. Terrific broth (TB), LB Broth Miller (LB), LB Agar Miller, potassium phosphate monobasic, potassium chloride, and glycerol were purchased from EMD Biosciences (Darmstadt, Germany). Isopropyl β-D-1-thiogalactopyranoside (IPTG), D-glucose, dithiothreitol (DTT), Tris-HCl, phenylmethanesulfonyl fluoride (PMSF), carbenicillin (Cb), sodium chloride, cupric sulfate, MOPS, HEPES, ethylene glycol, sodium acetate, acetonitrile,

hydrogen peroxide, and manganese chloride tetrahydrate were purchased from Fisher Scientific (Pittsburgh, PA). Potassium phosphate dibasic, 4-aminoantipyrene, 2,4-dichlorophenol (DCP), α-aminolevulinic acid, 2,2′-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS), Reactive Black 5, polyethylene glycol 4000, methoxymandelic acid, anisaldehyde, $MgSO_4 \cdot 7H_2O$, $CaCl_2 \cdot 2H_2O$, thiamine·HCl, $FeSO_4$, N,N,N′,N′-tetramethylethylenediamine, sodium dithionite, amino acids, ammonium molybdate, boric acid, cobalt chloride, zinc sulfate monohydrate, methyl viologen, anthraquinone-2-sulfonic acid, 2-hydroxy-1,4-naphthoquinone, toluylene blue, *p*-benzoquinone, and $Ru(NH_3)_6Cl_3$ were purchased from Sigma-Aldrich (St. Louis, MO). Imidazole, 1H-[1,2,4]oxadiazole[4,3-a]quinoxalin-1-one (ODQ), and formic acid were purchased from Acros Organics (Morris Plains, NJ). 3′-chloroindophenol was purchased from TCI-SU (Tokyo, Japan). Reactive Blue 5 was purchased from International Laboratory USA (South San Francisco, CA). Polyacrylamide, electrophoresis grade sodium dodecyl sulfate (SDS), gel filtration standard, and ammonium persulfate were purchased from Bio-Rad Laboratories (Hercules, CA). All PCR amplifications for cloning were carried out using Platinum Taq HF DNA polymerase (Invitrogen; Carlsbad, CA). Restriction enzymes, Antarctic phosphatase, and T4 DNA ligase were purchased from New England Biolabs (Ipswich, MA). Deoxynucleotides (dNTPs) were purchased from Invitrogen (Carlsbad, CA). DNase was purchased from Fermentas (Glen Burnie, Maryland). DNA was isolated using the QIAprep Spin Miniprep Kit, QIAquick PCR Purification Kit, and QIAquick Gel Extraction Kit (Qiagen; Valencia, CA). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), resuspended at a stock concentration of 100 μM in 10 mM Tris-HCl, pH 8.5, and stored at either 4°C for immediate usage or -20°C for longer term usage. For preparation of mobile phases for LC-MS/MS, acetonitrile (Fisher Optima grade, 99.9%) was purchased from Fisher Scientific.

Bacterial strains. *E. coli* DH10B-T1^R and BL21(de3)-T1^R were used for plasmid construction and heterologous protein production, respectively. *Amycolatopsis* sp. 75iv2 ATCC 39116 was purchased from the American Tissue Type Collection (Manassas, VA).

Plasmid construction. Expression plasmid for His_{10}-TEV-DyP2 was constructed using standard techniques. The gene encoding DyP2 was amplified from *A.* sp. 75iv2 genomic DNA with Platinum Taq HF DNA polymerase using the AmycoDyP2 F2 and AmycoDyP2 R1 primers, adding an internal TEV protease site, and inserted into the NdeI-HindIII sites of pCWOri-HisN to add an N- terminal His_{10} tag. The gene encoding DyP1, was amplified from *A.* sp. 75iv2 genomic DNA with Phusion DNA polymerase using the AmycoDyP1 F1 and AmycoDyP1 R1 primers and inserted into the NdeI-HindIII sites of pCWOri-HisN to add an N-terminal His_{10} tag (*Appendix*). The gene encoding DyP3 was amplified from *A.* sp. 75iv2 genomic DNA with Phusion DNA polymerase using the AmycoDyP3 F1 and AmycoDyP3 R1 primers, adding an internal TEV protease site, and inserted into the NdeI-XbaI sites of pCWOri-HisN to add an N- terminal His_{10} tag. The genes encoding Amyco3, Amyco4, and Amyco5 were amplified from *A.* sp. 75iv2 genomic DNA with Phusion DNA polymerase using, respectively, the Amyco F3 and Amyco R3, Amyco F4 and Amyco R4, and Amyco F5 and Amyco R5 primers (*Appendix*) and inserted into the NdeI-XbaI sites of pCWOri-HisN to add an N-terminal His_{10} tag. The resulting constructs were verified by sequencing (Quintara Biosciences; Berkeley, CA).

Phylogenetic analysis and sequence alignments. Three *A.* sp. 75iv2 DyP sequences were identified in the genome sequence and used to construct a phylogenetic tree [*14, 15*]. For DyP3, the signal sequence predicted by signalP 4.0 [*16*] was removed for analysis. Representative DyP

sequences were identified from the four known DyP clades by searching the Uniprot database [*17*] using sequences of each of the structurally-characterized members (A-clade: *Escherichia coli* str. K-12, EfeB-K12; B-clade: *Bacteroides thetaiotaomicron* VPI-5482, btDyP; C-clade: *Nostoc* sp. PCC 7120, AnaPX; D-clade: *Bjerkandera adusta* Dec1, DyP) with UniRef50 [*18*] to yield homologs with <50% sequence identity. Of these DyPs, 10-12 sequences were chosen from each clade for analysis. To increase the breadth of C-clade DyPs, an additional 13 *Amycolatopsis* sp. 75iv2 DyP2 homologs were identified using BLAST [*19*] and included in the sequence alignment. A structure-based alignment was generated from these sequences using Expresso [*20*] on the T-COFFEE Multiple Sequence Alignment Server [*21*]. The alignment output was then analyzed with MEGA 5.0 [*22*] using a maximum likelihood statistical method with a nearest-neighbor-interchange strategy for 423 positions, while allowing for deletion of gaps that exist in <50% of the sequences. The confidence was evaluated with 500 bootstraps. A compressed phylogenetic tree incorporating only structurally- and/or biochemically-characterized DyPs was also constructed by the same method with 405 positions used for analysis.

Purification of His_{10}-DyP1, His_{10}-TEV-DyP2, His_{10}-TEV-DyP3, and GA-DyP2. TB (1 L) containing carbenicillin (50 μg/mL) in a 2.8 L-baffled Fernbach flask was inoculated to OD_{600} = 0.05 with an overnight culture of freshly-transformed *E. coli* BL21(de3) grown in TB. For expression of His_{10}-DyP1, cells were also expressing pRARE2. The cultures were grown at 37 °C at 200 rpm in a rotary shaker until OD_{600} = 0.7, at which time the culture was cooled on ice for 15 min before adding IPTG (0.5 mM) and α-aminolevulinic acid (65 μg/mL). After inducing protein expression, the culture was allowed to incubate for an additional 20 h at 25°C. Cell pellets were harvested by centrifugation at 9,800 × *g* for 7 min and stored at -80 °C. Frozen cell pellets were thawed and resuspended at 5 mL/g cell paste with Buffer A (50 mM potassium phosphate, 300 mM sodium chloride, 10 mM imidazole, pH 8.0) supplemented with PMSF (0.5 mM), DTT (1 mM), and DNase (2 U/g cell paste). The cell paste was homogenized before lysis by passage through a French Pressure cell (Thermo Scientific; Waltham, MA) at 14,000 psi. The lysate was centrifuged at 15,300 × *g* for 30 min at 4 °C to separate the soluble and insoluble fractions. The soluble lysate was passed over a Ni-NTA agarose column (Qiagen, 10 mL) at a flow rate of 2 mL/min on an ÄKTApurifier FPLC (GE Healthcare Life Sciences; Piscataway, NJ). The protein was isolated by elution with 5 column volumes of Buffer B (50 mM potassium phosphate, 300 mM sodium chloride, 250 mM imidazole, 1 mM DTT, pH 8.0) after washing with 25% Buffer B (20 column volumes). Fractions were pooled based on the A_{280} elution profile and concentrated with an Amicon filtration device using a YM10 membrane (Millipore Corporation; Billerica, MA).

The concentrations of His_{10}-DyP1 and His_{10}-TEV-Dyp3 were estimated using the calculated extinction coefficients at 280 nm for their respective amino acid sequences (His_{10}-DyP1, 42,400 $M^{-1}cm^{-1}$; His_{10}-TEV-DyP3, 40,575 $M^{-1}cm^{-1}$). The proteins waere stored at -80°C after addition of glycerol (10% (v/v) final concentration).

To prepare GA-DyP2, His_{10}-TEV-DyP2 protein was incubated overnight with His_6-TEV (1:75 mg TEV/mg DyP2) during dialysis in 10 kD MWCO Spectra/Por cellulose ester tubing (Spectrum Laboratories; Rancho Dominguez, CA) against Buffer A supplemented with DTT (1 mM) to remove the N-terminal His_{10}-tag. His_6-TEV and undigested His_{10}-TEV-DyP2 were removed by passing the mixture over a Ni-NTA agarose column (10 mL) and washing off unbound protein with Buffer A (10 mL) containing DTT (1 mM).

DyP2s were then further purified by size exclusion chromatography using a HiLoad 16/20 Superdex 200 prep grade size exclusion chromatography column (GE Healthcare Life Sciences; Piscataway, NJ) equilibrated with Buffer C (20 mM Tris, 50 mM potassium chloride, pH 7.5). To remove bound imidazole, the protein was then dialyzed 3×10^{-2} against Buffer C using a 10 kD MWCO Spectra/Por cellulose ester dialysis membrane for a 10^{-6} final dilution of the original buffer. The extinction coefficient of His_{10}-TEV-DyP2 (Reinheitzahl ratio, Rz = 1.6-1.8) at 280 nm was measured to be 61,800 $M^{-1}cm^{-1}$ by protein acidolysis (AAA Service Laboratory; Damascus, OR), which was also used to approximate the concentration of GA-DyP2. The protein was stored at -80 °C after addition of glycerol (10% (v/v) final concentration).

ICP-OES analysis. His_{10}-TEV-DyP2 was analyzed for Fe content using literature protocol [*14*]. From this analysis, the extinction coefficient of the Soret band for His_{10}-TEV-DyP2 was determined to be 113,300 $M^{-1}cm^{-1}$ at 404 nm. Based on the Fe content and concentration measured by analysis, the heme loading of His_{10}-TEV-DyP2 was determined to be 99.7%.

Dynamic light scattering analysis. His_{10}-TEV-DyP2 (0.5 mg/mL in Buffer C, 30 μL) was filtered to remove particulate matter by centrifugation (2 min, 20817 × *g*) through a Durapore PVDF membrane (0.1 μM, Millipore). Dynamic light scattering data (n = 20) was collected for His_{10}-TEV-DyP2 on a DynaPro Titan (20% maximum laser power; Wyatt Technology Corporation; Goleta, CA) at 589 nm and 25°C over a 1 s time interval, and ASTRA was used to analyze and average the data using a Raleigh spheres model to obtain the hydrodynamic radius and particle diameter.

Redox potential titrations. Redox potential titrations were carried out on His_{10}-TEV-DyP2 in a 1.6 mL cell consisting of a 1 cm semi-micro quartz magnetic stir cell (Starna Cells, Inc; Atascadero, CA) and stir bar. His_{10}-TEV-DyP2 was deoxygenated in an anaerobic glove box (Controlled Atmosphere Chamber; Plas Labs; Lansing, MI) under a 90 Ar:10 H atmosphere by passing an aliquot of protein (400 μL) over a NAP-5 desalting column (GE Healthcare Life Sciences) equilibrated with titration buffer (Ar-sparged Buffer C supplemented with 5% ethylene glycol). His_{10}-TEV-DyP2 (0.7 mg/mL) was transferred into the cell followed by the Ar-sparged redox mediator cocktail (toluylene blue, 5 μM; hexaammineruthenium(III) chloride, 16.5 μM; *p*-benzoquinone, 16.5 μM; 3′-chloroindophenol, 10 μM; methyl viologen, 7.5 μM; anthraquinone-2-sulfonic acid, 10 μM; 2-hydroxy-1,4-naphthoquinone, 10 μM) [*23*]. The cuvette was then outfitted with the electrodes and sealed from air before removing from the glove box. Reaction was maintained under N_2 for the duration of the experiment. For spectroelectrochemical titrations, His_{10}-TEV-DyP2 (1.5 mL) was analyzed with a platinum gauze working electrode, platinum counter electrode, and Ag/AgCl reference electrode. The BASi Epsilon potentiostat (West Lafayette, IN) was set to -550 mV vs Ag/AgCl to reduce His_{10}-TEV-DyP2 and increased stepwise to -150 mV vs Ag/AgCl, allowing each potential to stabilize for 60 min before spectra were collected. For spectrochemical titrations, His_{10}-TEV-DyP2 (1 mL) was analyzed using an oxidation-reduction potential microelectrode (MI-800, Microelectrodes, Inc.; Bedford, NH) and Ag/AgCl reference microelectrode (MI-402, Microelectrodes, Inc.). The oxidant solution was prepared in the glove box by dissolving solid ODQ in Ar-sparged DMSO (10 mM) and then diluting into titration buffer (0.5 mM). The reaction potential was monitored with a basic pH/mV/ORP meter (Orion 420A plus, Thermo Scientific; Waltham, MA). Protein was initially reduced step-wise using sodium dithionite introduced via syringe until the potential stabilized and UV-Vis spectrum showed the protein to be fully reduced. After reduction, ODQ (0.2-1 μL; 0.5 mM) was titrated in with stirring to oxidize His_{10}-TEV-DyP2. The reaction was allowed to

equilibrate for 45 min after each addition. UV-visible spectra were collected at each potential on an Agilent 8453 diode-array spectrophotometer equipped with an Agilent 8909A stirring module (Santa Clara, CA). The fraction of reduced DyP2 was monitored by $\Delta A_{560\ nm}$ and the midpoint reduction potential was determined by fitting to the following equation [*24*]:

$$Y = \frac{e^{-96500\times(m_0-m_1)/2477.572}}{1+e^{-96500\times(m_0-m_1)/2477.572}}$$

Purification of selenomethionine-labeled His_{10}-TEV-DyP2. LB (50 mL) containing carbenicillin (50 μg/mL) in a 250 mL-baffled shake flask was inoculated to $OD_{600\ nm}$ = 0.05 with an overnight culture of freshly-transformed *E. coli* BL21(de3) pCWOri-His_{10}-TEV-DyP2 grown in LB. The cultures were grown at 37 °C at 200 rpm in a rotary shaker until $OD_{600\ nm}$ = 0.8, after which the culture cell pellets were harvested by centrifugation at 9,800 × *g* for 7 min. The cell pellet was washed (10 mL) and resuspended (50 mL) with MOPS-M9 media supplemented with carbenicillin (MOPS-M9-Cb). The cell suspension (5 mL) was used to inoculate MOPS-M9-Cb (500 mL) and incubated at 37 °C at 200 rpm. When the $OD_{600\ nm}$ reached 1.0, leucine (25 mg), isoleucine (25 mg), valine (25 mg), phenylalanine (50 mg), lysine (50 mg), threonine (50 mg), and selenomethionine (75 mg) were added to the culture. After an additional incubation of 20 min at 37 °C, the culture was cooled on ice for 15 min before adding α-aminolevulinic acid (65 μg/mL) and IPTG (0.5 mM). The cultures were then incubated for an additional 23 h at 25 °C for protein production. Cell pellets were stored and SeMet-His_{10}-TEV-DyP2 (Rz = 2.02) was purified as described for His_{10}-TEV-DyP2 except that the final dialysis buffer was supplemented with TCEP (1 mM).

Crystallization and structure determination of His_{10}-TEV-DyP2. All reservoir solution components were filter sterilized by passage through 0.2 μM filtration prior to usage. His_{10}-TEV-DyP2 and SeMet-His_{10}-TEV-DyP2 were dialyzed against 3 × 1 L Buffer C over 24 h to remove glycerol for a 10^{-9} final dilution of the original buffer. The protein was then concentrated 6-7 mg/mL using a 10 kD MWCO AmiconUltra-0.5 mL centrifugal filter unit (Millipore). Protein crystals were obtained using the hanging drop vapor diffusion method by combining equal volumes of protein solution and reservoir solution containing 0.1 M sodium acetate, pH 4.9-5.2, 22 mM $MnCl_2$, and 8.0-8.5% polyethylene glycol 4000. Long rectangular crystals formed within days but maturation over >2 months yielded higher quality crystals. The crystals were cryoprotected by briefly soaking in 3:1 (v/v) well solution:ethylene glycol and looped with angled-tip microloops (200μm, M5-L18SP-200, MiTeGen; Ithaca, NY). Data were collected at Beamline 8.3.1 at the Advanced Light Source (Lawrence Berkeley National Laboratory) and processed using the XDS [*25*] and CCP4 [*4, 26*] software packages with 5% of the reflections flagged for R_{free} calculations. Phases were determined by a combination of a partial molecular replacement solution using *B. adusta* Dec1 DyP (PDB ID 2D3Q) as search model and the location of the selenium atoms determined by SAD using Phaser [*27*] and a 3.0 Å resolution dataset from a SeMet-His_{10}-TEV-DyP2 crystal. Electron density maps were improved by density modification and non-crystallographic symmetry averaging using Parrot [*28*] before building the initial model using Buccaneer [*29*]. The final model was then generated through multiple rounds of manual building in COOT [*30*] and refinement using the PHENIX package [*31*] using 2.25 Å resolution data from a His_{10}-TEV-DyP2 crystal (Accession codes**:** DyP2 PDB ID: 4G2C).

Spectrophotometric activity assays. All assays were performed at 25 °C using a Beckman DU-800 spectrophotometer. Kinetic parameters (k_{cat} and K_M) were determined in Origin (OriginLab, Northampton, MA) by nonlinear curve fitting to the data using the Michaelis-Menten equation: $v_0 = v_{max}$ [S]/ K_M + [S] where v_0 is the initial rate and [S] is the substrate concentration. Data are reported as the mean ± s.e. (n = 3-6) as determined from curve fitting. Error in the k_{cat}/K_M parameter was obtained from propagation of error from the individual kinetic terms. Peroxidase activity was also compared for 2,4-dichlorophenol, Reactive Blue 5, and ABTS in the presence of $MnCl_2$ (1 mM) and found to be unchanged.

2,4-Dichlorophenol oxidation assay. 2,4-Dichlorophenol oxidation (7.5 μM-6 mM) was assayed as previously reported [*14*] with a lower concentration of H_2O_2 (0.5 mM) and initiated with DyP2 (84 nM).

Reactive Blue 5 decolorization assay. Reactive Blue 5 decolorization activity was monitored using a modified literature protocol [*32*]. Reactions (1 mL) were performed in 50 mM sodium acetate at pH 4.5 containing H_2O_2 (0.5 mM) and Reactive Blue 5 (10-200 μM) and initiated with His_{10}-TEV-DyP2 (41 nM). Activity was monitored by a decrease in absorbance at 600 nm (ε_{600} = 8 $mM^{-1}cm^{-1}$). Kinetic parameters were determined using a range of Reactive Blue 5.

ABTS oxidation assay. ABTS oxidation was monitored using a modified literature protocol [*33*]. Reactions (1 mL) were performed in 50 mM sodium acetate at pH 4.5 containing H_2O_2 (0.5 mM) and ABTS (2.5 μM-2.5 mM) and initiated with His_{10}-TEV-DyP2 (2.3 nM). Activity was monitored by an increase in absorption at 420 nm (ε_{420} = 36 $mM^{-1}cm^{-1}$). His_{10}-TEV-DyP2 was found to be unstable at low concentration so new dilutions were made every hour to maintain activity.

Reactive Black 5 decolorization assay. Reactive Black 5 decolorization activity was monitored using a modified literature protocol [*34*]. Reactions (1 mL) were performed in 50 mM sodium acetate at pH 4.5 containing H_2O_2 (0.5 mM) and Reactive Black 5 (1-16 μM) and initiated with His_{10}-TEV-DyP2 (84 nM). Activity was monitored by a decrease in absorbance at 598 nm (ε_{598} = 30 $mM^{-1}cm^{-1}$).

Mn^{2+}-malonate oxidation assay. Mn^{2+} oxidation was monitored using a modified literature procedure [*35*]. Reactions (1 mL) were performed in 50 mM sodium malonate at pH 4.5 containing H_2O_2 (0.5 mM) and $MnCl_2$ (0.05-3 mM) and initiated with His_{10}-TEV-DyP2 (41 nM). Activity was monitored by an increase in absorption at 270 nm (ε_{270} = 11.59 $mM^{-1}cm^{-1}$).

pH-rate profile. Peroxidase activity for Reactive Blue 5 was monitored spectrophotometrically as described [*14*] with a variation of Britton and Robinson buffers in place of the respective buffers.

HPLC assay for monitoring peroxide-dependent degradation of lignin model compounds. Lignin model dimers, veratrylglycerol-*β*-guaiacol ether (**1**) and guaiacylglycerol-*β*-guaiacol ether (**2**), were prepared as previously reported [*14*] to test reactivity of DyP2. Reactions (1 mL) were performed in 50 mM sodium acetate, pH 4.5 containing H_2O_2 (1 mM) and lignin model compounds (1 mM) and initiated with addition of His_{10}-TEV-DyP2 (0.2 μM). When added, $MnCl_2$ was supplemented to a final concentration of 1 mM. All reactions were allowed to stir overnight at room temperature. The solution was centrifuged at 20,817 × *g* for 5

min to remove any particulate matter, and the cleared solution was analyzed by RP-HPLC on an Agilent 1200 Series HPLC coupled to a G1315D diode array detector.

*Analysis of reactivity with **1**.* Samples (20 μL) were chromatographed on an Eclipse XDB-C_{18} column (3.5 μm, 3.0 × 150 mm; Zorbax; Santa Clara, CA) using 0.1 % formic acid (A) and acetonitrile (B)as the mobile phase (0.5 mL/min) using the following gradient: 0-2.0 min, hold at 10% B; 2.0-4.0 min, linear gradient from 10%-40% B; 4.0-8.0 min, hold at 40% B; 8.0-10 min, linear gradient from 40%-75% B; 10.0-13.0 min, hold at 75% B; 13.0-15.0 min, linear gradient from 75%-10% B.

*Analysis of reactivity with **2**.* Samples (20 μL) were chromatographed on an Eclipse XDB-C_{18} column (3.5 μm, 3.0 × 150 mm; Zorbax; Santa Clara, CA) using 0.1 % formic acid (A) and acetonitrile (B) as the mobile phase (0.5 mL/min) using the following gradient: 0-2.5 min, hold at 25% B; 2.5-6.0 min, linear gradient from 25%-75% B; 6.0-12.0 min, hold at 75% B; 12.0-15.0 min, linear gradient from 75%-25% B.

Assays for monitoring DyP2 oxidase activity. DyP2 oxidase reactivity was assayed using a modified literature assay for oxidative decarboxylation of 4-methoxymandelic acid (MMA) to anisaldehyde [*36*]. Reactions (1 mL) were performed in 50 mM sodium acetate, pH 4.5 containing H_2O_2 (1 mM), $MnCl_2$ (5 mM), and MMA (2.5 mM) and initiated with addition of His_{10}-TEV-DyP2 (0.1 μM). Reactions were analyzed by RP-HPLC as described for **2**. For the boiled protein control, His_{10}-TEV-DyP2 was incubated at 100 °C for 15 min to denature protein before adding to the assay mixture. Mn^{2+} dependence was measured by varying $MnCl_2$ concentration from 0-5.0 mM. To verify the identity of the product as anisaldehyde, the reaction mixture was extracted with 500 μL ethyl acetate and compared to an anisaldehyde standard (12.5 mM) using GC-MS. The oven program was as follows: 50 °C for 2 min, ramp to 300 °C at 15 °C min^{-1}, 300 °C for 15 min. The analyte was monitored using a DSQII single-quadrupole mass spectrometer (Thermo Scientific) using full scan mode (*m/z* 100-1050, 7465 amu s^{-1}; 1 mL helium min^{-1}).

Anaerobic reactions. Anaerobic reactions prepared in an anaerobic glove-box (Controlled Atmosphere Chamber; Plas Labs; Lansing, MI) under a 90 Ar:10 H atmosphere. Reaction buffer, Buffer C, water, and ethanol were first sparged with Ar (5 min/mL) to remove oxygen prior to introduction to glove box. MMA and $MnCl_2$ were brought into the glove box as solids and dissolved with Ar-sparged ethanol (MMA, 0.5 M) or protein buffer ($MnCl_2$, 0.5 M). DyP2 was deoxygenated by passing an aliquot of protein (250 μL) over a NAP-5 desalting column (GE Healthcare Life Sciences) equilibrated with Ar-sparged Buffer C. Reactions in 1.5 mL vials were initiated by addition of His_{10}-TEV-DyP2 (0.1 μM), tightly sealed with septa (Sigma-Aldrich, 10 mm), removed from the glove box, and allowed to stir overnight at room temperature. Reactions were quenched by heating at 100 °C for 15 min without exposure to air. Any precipitate formed in the reaction was removed by centrifugation and analyzed by RP-HPLC as previously described for **2**.

pH-rate profile. DyP2 activity was measured by RP-HPLC as previously described using the following buffer system to avoid Mn^{2+} precipitation in phosphate buffer above pH 4.5: pH 3.5-4.5 (Britton and Robinson buffers), pH 5.0-6.0 (sodium acetate, 50 mM), pH 6.5-7.5 (HEPES, 50 mM).

4.3. Results and Discussion

Phylogenetic analysis of the DyPs from *Amycolatopsis* sp. 75iv2. DyPs can be classified according to four different clades. In general, DyPs from bacteria are smaller in size and fall within the A- and B- clades, whereas fungal enzymes are typically found to cluster within the D-clade [*5, 12*]. A DyP phylogenetic tree was constructed by searching the UniRef50 database using the structurally-characterized DyPs that represent each clade to obtain sequences with <50% identity. Phylogenetic analysis of the DyPs from *A.* sp. 75iv2 indicates that two belong to the C-clade (DyP1-2), which are clustered with DyPs away from other prokaryotes but related to fungal D-type DyPs (*Figure 4.1*). Interestingly, many of the other C-type DyPs belong to related actinomycetes that have also been reported to produce APPL and can be found clustered in the genome with biomass-processing enzymes (*Table 4.1*, *Figure 4.2*). However, DyP2 belongs to an especially divergent and sparsely populated branch of the C-clade. Consequently, additional DyP2 homologs were identified by BLAST to fill this branch and were found to consist of only a handful of sequences that mostly originate from cyanobacteria (*Figure 4.1*, *Table 4.1*). The third DyP from *A.* sp 75iv2, DyP3, can be grouped with other canonical bacterial DyPs, such as the one found to be partially responsible for the lignolytic behavior of *Rhodococcus jostii* RHA1 [*11, 12*].

Expression and initial characterization of other DyPs and related heme proteins *A.* sp 75iv2. While we focused on DyP2 for these studies, the genome of *A.* sp. 75iv2 also encodes two other DyPs, *dyp1* (C-type) and *dyp3* (A-type with a secretion signal). Interestingly, *dyp1* is clustered with four other heme-containing proteins that include a catalase (downstream), two cytochrome P450s, and another heme peroxidase (upstream) (*Figure 4.2*). Thus, these proteins were cloned and heterologously expressed in *E. coli* with an N-terminal His_{10}–tag. Both DyP1 and DyP3 were purified by nickel affinity chromatography, and UV-visible spectrometry revealed a typical Soret absorption peak, supporting their characterization as heme-containing proteins (*Figure 4.3A-D*). While DyP3 showed a reasonable Rz value (0.76; 0.31 heme/monomer based on estimated Soret absorbance of 100,000 M^{-1} cm^{-1} and the calculated ε_{280} for DyP3), the heme-loading of DyP1 was very poor (Rz = .06; estimated 0.031 heme/monomer). Monitoring the oxidation of 2,4-dichlorophenol, the specific activities of DyP1 and DyP3 were calculated to be approximately 0.000559 μmol mg^{-1} min^{-1} and 0.07 μmol mg^{-1} min^{-1}, respectively. Efforts were also made to biochemically characterize Amyco3-5, which were found to be clustered with DyP1. Amyco3 (heme peroxidase, JGI locus tag Amy_03135), Amyco4 (cytochrome P450, JGI locus tag Amy_03134) and Amyco5 (catalase, GI locus tag Amy_03132) were cloned from *A.* sp. 75iv2 genomic DNA for heterologous expression in *E. coli*. Initial expression studies indicate that of these three proteins, only Amyco3 shows significant overproduction levels (*Figure 4.3E*).

Biochemical characterization of the C-type DyP2 from *A.* sp. 75iv2. Based on its unusual phylogeny and the sequence context of other C-type DyPs from lignin-reactive actinomycetes, we cloned and heterologously expressed DyP2 from *A.* sp 75iv2 with an N-terminal His_{10}–tag containing a linker with a Tobacco Etch Virus (TEV) protease site in *Escherichia coli* (*Figure 4.4*). The nickel affinity chromatography-purified DyP2 protein was initially characterized by size-exclusion chromatography (SEC) to be both a monomeric and oligomeric species (n = 4-6) in solution (*Figure 4.4*). Further analysis showed that the monomeric form contained no bound heme cofactor and its removal by SEC significantly increased the specific activity of DyP2.

After SEC, it appeared as if a bound imidazole remained associated with the heme cofactor

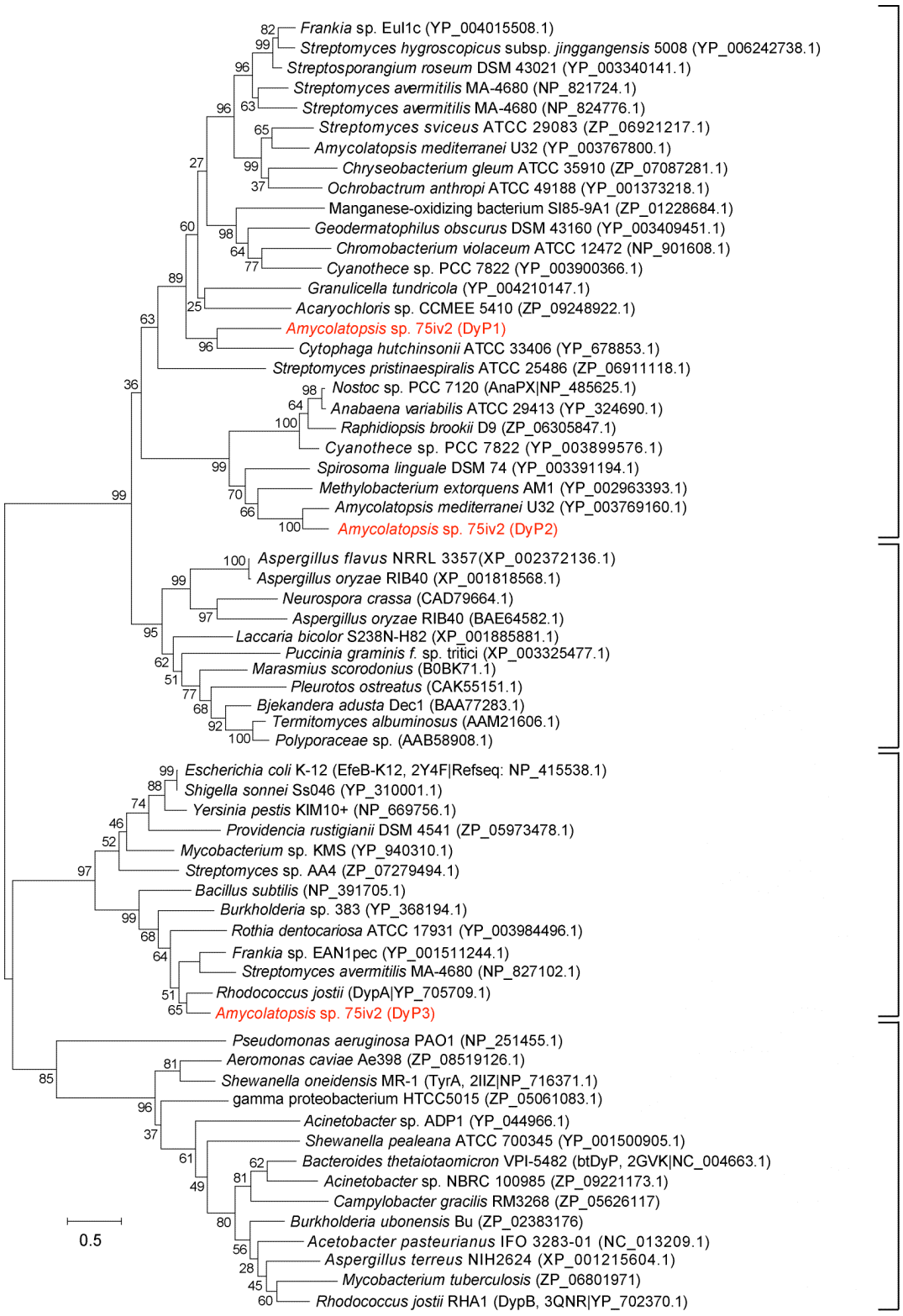
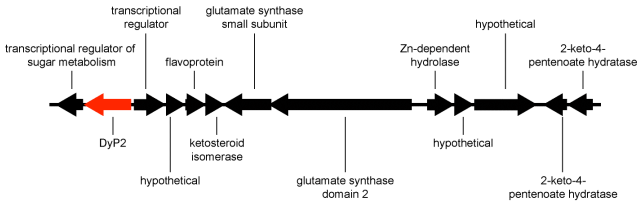


Figure 4.1. Phylogenetic tree for DyPs. *An alignment of UniProt and NCBI sequences was performed using T-COFFEE Expresso. A maximum likelihood-nearest neighbor interchange tree was then constructed in MEGA 5.0 and tested with 500 bootstrap replicates. Bootstrap values are listed near each branch. RefSeq numbers are listed in parentheses.*

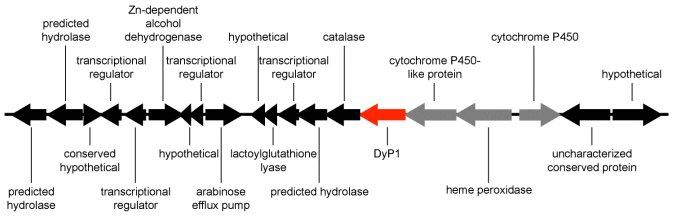
	organism	identity (%)	E-value
DyP2	*Amycolatopsis thermoflava* N1165	96	0.0
	Amycolatopsis mediterranei U32	65	e-171
	Rhizobium leguminosarum bv. *trifolii* WSM 1325	58	e-143
	Methylobacterium extorquens AM1	45	3e-92
	Methylobacterium extorquens DSM 13060	45	1e-91
	Hymenobacter roseosalivarius AA-718	41	4e-79
	Raphidiopsis brookii D9	38	6e-76
	Spirosoma linguale DSM 74	39	3e-75
	Nostoc sp. PCC 7120	37	3e-73
	Anabaena variabilis ATCC 29413	37	3e-72
	Cyanothece sp. PCC 8801	37	2e-71
	Cyanothece sp. PCC 8802	37	2e-71
	Cyanothece sp. PCC 7822	39	2e-68
	Cyanothece sp. ATCC 51142	36	3e-38
DyP1	*Amycolatopsis* sp. 75iv2 DyP1	30	1e-26
	Plesiocystis pacifica SIR-1	29	4e-36
	Moorea producta 3L	31	1e-27
	Nitrosospira multiformis ATCC 25196	29	6e-27
	Granulicella tundricola	32	6e-26
	Nitrosomonas europaea ATCC 19718	35	1e-25
	Streptomyces hygroscopicus subsp. *jinggangensis* 5008	32	2e-25
	Acaryochloris sp. CCMEE 5410	29	2e-25
	Frankia sp. EUN1f	31	5e-25
	Geodermatophilus obscurus DSM 43160	31	1e-24
	Cyanothece sp. PCC 7822	27	1e-23
	Streptosporangium roseum DSM 43021	29	1e-23
	Frankia sp. EuI1c	30	1e-23
	Sorangium cellulosum 'So ce 56'	28	6e-23
	Cytophaga hutchinsonii ATCC 33406	27	6e-23
	Vibrio nigripulchritudo ATCC 27043	26	7e-23
	Ktedonobacter racemifer DSM 44963	30	2e-22
	Frankia sp. CN3	29	4e-22
	Myxococcus xanthus DK 1622	28	9e-22
	Cyclobacterium marinum DSM 745	26	9e-22
	Candidatus Solibacter usitatus Ellin6076	28	4e-21
	Amycolatopsis mediterranei U32	29	5e-21
	Corallococcus coralloides DSM 2259	28	1e-20
	Dinoroseobacter shibae DFL 12	33	2e-20
	Methylomonas methanica MC09	27	4e-20
	Ochrobactrum anthropi ATCC 49188	27	7e-20
	Nakamurella multipartita DSM 44233	28	1e-19
	Uncultured marine microorganism HF4000_APKG1C9	28	1e-19
	Chromobacterium violaceum ATCC 12472	27	1e-19
	Streptomyces avermitilis MA-4680	30	1e-19
	Chryseobacterium gleum ATCC 35910	27	2e-19
	Cyanothece sp. PCC 7424	25	2e-19
	Mobilicoccus pelagius NBRC 104925	29	1e-18
	Enhydrobacter aerosaccus SK60	30	3e-18
	Mycobacterium sp. JLS	30	3e-18
	Aurantimonas manganoxydans SI85-9A1	27	3e-18
	Mycobacterium sp. MCS	30	3e-18
	Streptomyces sviceus ATCC 29083	30	3e-17
	Janthinobacterium sp. Marseille	28	4e-17

Table 4.1. List of C-type DyP1 and DyP2 orthologs. *C-type DyPs were identified using a BLAST search with the DyP2 sequence, which identified both DyP1 and DyP2 homologs. Sequence identity and E-values with respect to DyP2 are reported.*

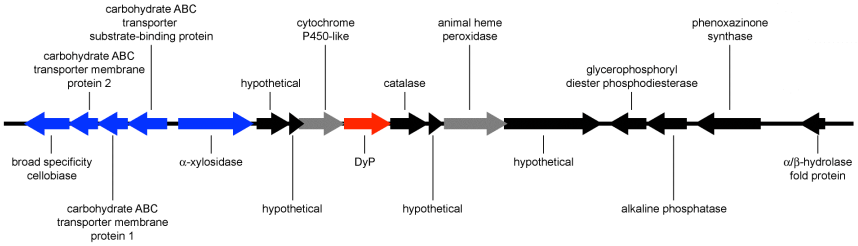
Amycolatopsis sp. 75iv2 (DyP2)



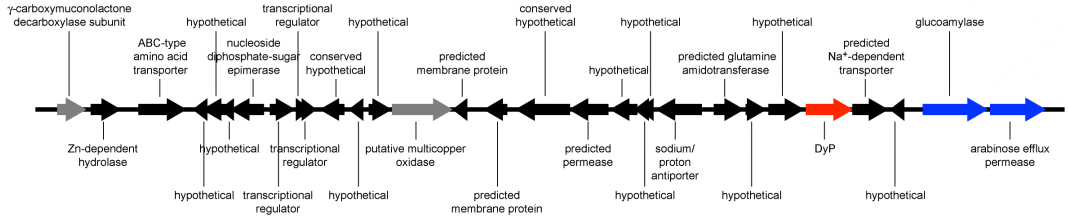
Amycolatopsis sp. 75iv2 (DyP1)



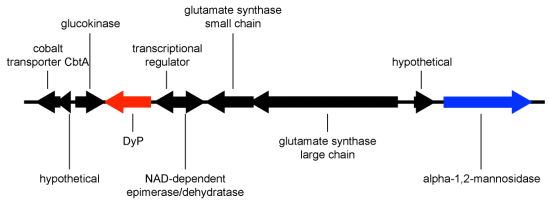
Streptomyces hygroscopicus subsp. *jinggangensis* 5008 (YP_006249142.1)



Frankia sp. CN3 (ZP_09167357.1)



Amycolatopsis mediterranei U32 (YP_003769160.1)



Streptomyces avermitilis MA-4680 (NP_821724.1)

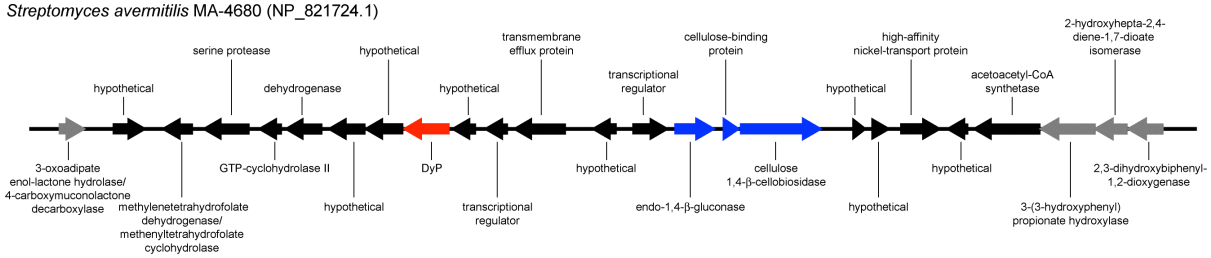


Figure 4.2. Sequence contexts of C-type DyPs from related actinomycetes. *The genes encoding C-type DyPs in related actinomycetes are found to be clustered with other genes with potential lignocellulose degradation or carbohydrate metabolic function.*

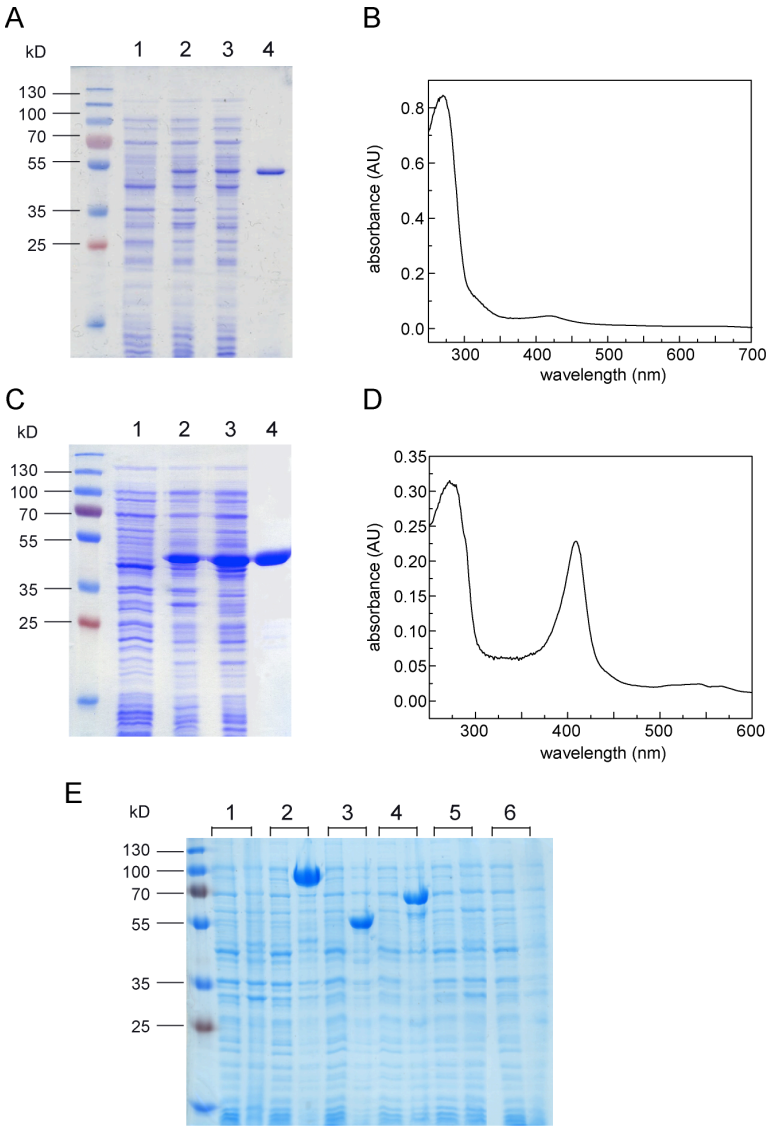


Figure 4.3. Expresion and purification of DyP1, DyP3, and Amyco3-5*. (A) SDS-PAGE analysis of His_{10}-DyP1 purification. Pre-induction sample (lane 1), post-induction sample (lane 2), soluble cell lysate (lane 3), purified product (lane 4). (B) UV-visible spectrum of His_{10}-DyP1. (C) SDS-PAGE analysis of His_{10}-DyP3 purification. Pre-induction sample (lane 1), post-induction sample (lane 2), soluble cell lysate (lane 3), purified product (lane 4). (D) UV-visible spectrum of His_{10}-DyP3. (E) SDS-PAGE analysis of pre- and post-induction samples for Amyco3-5 as compared with well-expressed proteins, Amyco1-2: (1) empty plasmid; (2) Amyco1; (3) Amyco2; (4) Amyco3; (5) Amyco4; (6) Amyco5. Conditions for protein expression were the same as those for DyP2 expression.*

by UV-visible spectroscopy (*Figure 4.4D*). Consequently, DyP2 was dialyzed until the λ_{max} of the Soret band blue-shifted into the expected range for a high-spin Fe^{3+}-hemoprotein bound proximally to a histidine (λ_{max} ~ 400-410 nm). The introduction of these two steps to the purification of DyP2 increased the peroxidase activity of DyP2 by 3,500-fold compared to the enzyme as directly isolated after affinity chromatography. Using a combination of inductively coupled plasma–optical emission spectroscopy (ICP-OES) and protein acidolysis, the heme occupancy of the purified DyP2 was determined to be >99% (*Figure 4.5*). Since removal of the His_{10}-tag by cleavage at the intervening TEV protease site did not appear to change the oligomeric state by SEC or increase specific activity (*Figure 4.6*), our biochemical and structural studies were carried out on His_{10}-DyP2.

In the DyP family, both monomeric [*2, 3, 8*] and oligomeric [*2, 4, 5, 7, 37*] members have been reported. Dynamic light scattering experiments were carried out to further characterize the oligomeric state of the active DyP2 enzyme in solution (*Figure 4.7*). The measured hydrodynamic radius is consistent with a solution state that is larger than a monomer based on the range of sizes observed for structurally-characterized DyPs [*10, 11, 37*]. Additional characterization of purified DyP2 by UV–visible spectroscopy indicates that DyP2 is isolated in a high–spin Fe^{3+} resting state (*Figure 4.4*). The Fe^{3+}/Fe^{2+} midpoint reduction potential was measured both by spectrochemical and spectroelectrochemical titration to be –85 ± 13 mV vs NHE with Nernstian behavior (*Figure 4.8AB*). Although this potential is more positive than many other heme peroxidases, which are characteristically in the range of approximately -120 to -320 mV, it is close to that observed for manganese peroxidases [*38*]. The high oxidative capacity of the LiPs and MnPs and their more positive Fe^{3+}/Fe^{2+} midpoint reduction potentials have been attributed to their electon-deficient hemes [*39*]. These classic lignolytic enzymes have been shown to reside extracellularly [*40*]. Analysis of the DyP2 sequence does not show the existence of a canonical signal sequence for secretion (*Figure 4.9*). However, the secretory machinery for actinomycetes is not as well characterized and the low pH optimum for DyP2 (*Figure 4.8C*) along with the observation that many DyPs have been isolated from the secreted protein fraction [*1, 4, 6, 8*] imply that DyP2 may play a role in extracellular oxidation chemistry.

DyP2 displays versatile peroxidase activity with a broad substrate scope. In order to explore its substrate scope, DyP2 was tested against a panel of peroxidase substrates including aromatics, azo dyes, anthroquinone dyes, and Mn^{2+}. DyP2 was found to be highly active in phenol oxidation using a standard 2,4-dichlorophenol peroxidase assay (*Table 4.*2, *Figure 4.10*). An additional signature of DyPs is their ability to degrade anthroquinone (Reactive Blue 5) as well as azo (ABTS) dyes [*1, 6, 7, 9*]. We therefore characterized DyP2 with respect to both classes of substrates and found high specific activity towards Reactive Blue 5 and ABTS, with k_{cat}/K_M's that are more similar to the C- and D-type DyPs compared to the significantly lower values observed in the characterized bacterial A- and B-type DyPs (*Tables 4.2* and *4.3*). Furthermore, DyP2 was found to be competent for the decolorization of Reactive Black 5, an azo dye considered to have a high oxidation potential due to the relative paucity of enzymes that have been shown to degrade this molecule [*41*]. The ability of DyP2 to degrade lignin model dimers containing the β-O-4 linkage [*42*] was assayed with two model substrates with (**1**) or without (**2**) a phenolic site (*Figure 4.11*). These studies showed that DyP2 could rapidly break down **1** in the absence of redox mediators, but that the lower-potential phenolic site (ΔE = 0.6-0.8 V vs NHE) [*43*] is required for this activity as no degradation of **2** was observed under the same conditions (*Figure 4.11*).

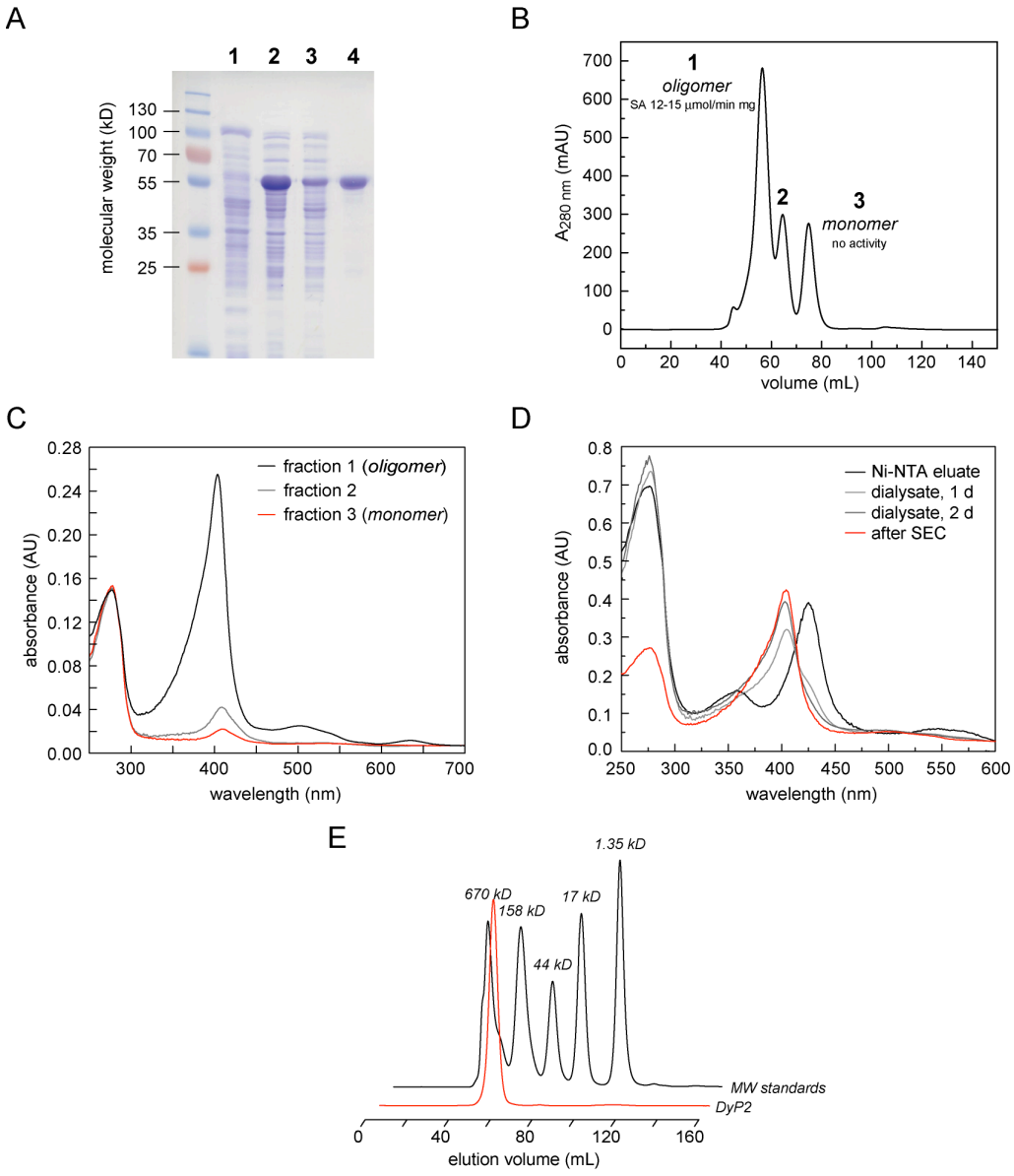


Figure 4.4. Purification of His_{10}-TEV-DyP2. *(A) SDS-PAGE analysis of His_{10}-TEV-DyP2 purification. Pre-induction sample (lane 1), post-induction sample (lane 2), soluble lysate sample (lane 3), Ni-NTA column eluate (lane 4). (B) Size-exclusion chromatogram of Ni-NTA eluate. Measurable activity in the oxidation of 2,4-dichlorophenol at pH 7 is found only in the heme-bound oligomeric state. (C) UV-visible spectra of different SEC fractions. (D) UV-visible spectra showing the shift of the Soret with dialysis indicating removal of bound imidazole as well as increase in the Rz value after isolation of the oligomer. (E) Size-exclusion chromatogram of His10-TEV-DyP2 (5 mg/mL) using a Superdex 200 column estimates a MWapp of 308 kD in 20 mM Tris, 50 mM KCl, pH 7.5 based on a linear fit to the molecular weight standards.*

A

amino acid	composition (mole %) analyzed	expected	% error
Asp	13.118	10.1	29.9
Thr	5.887	6.0	-1.9
Ser	4.167	4.5	-7.4
Glu	6.675	10.1	-33.9
Pro	7.264	7.4	-1.8
Gly	10.306	10.5	-1.9
Ala	11.052	11.3	-2.2
Val	6.126	6.0	2.1
Met	1.471	2.1	-29.9
Ile	1.363	1.4	-2.6
Leu	9.930	9.7	2.4
Tyr	2.288	2.1	9.0
Phe	6.041	4.7	28.5
His	5.070	5.3	-4.3
Lys	3.411	3.3	3.4
Arg	5.831	5.7	2.3
DyP2 concentration (μM)			15.5

B

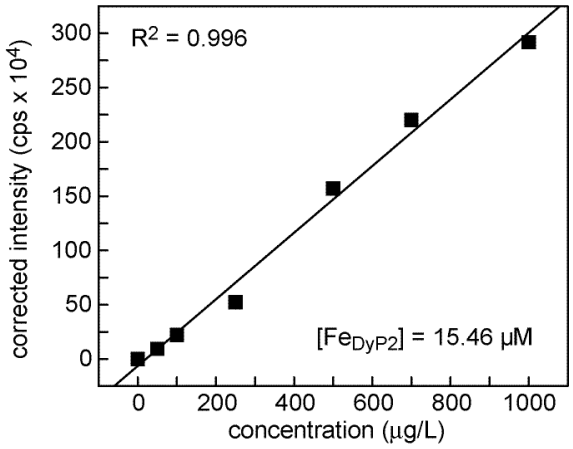
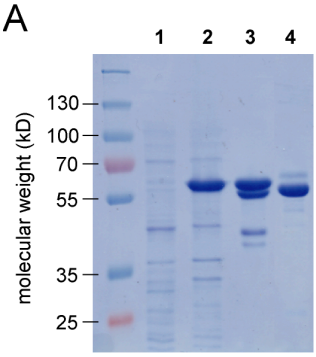
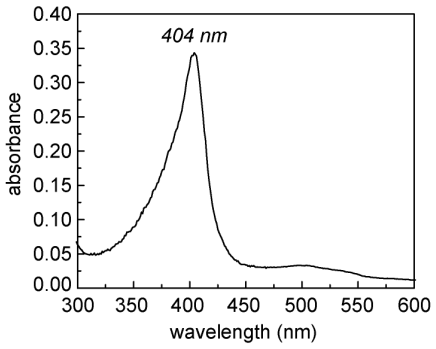


Figure 4.5. Analysis of heme occupancy of purified His_{10}-TEV-DyP2. *(A) The concentration of His_{10}-TEV-DyP2 was determined by total protein acidolysis. For predicted amino acid concentrations, N and Q were converted to D and E, respectively. (B) Fe content of His_{10}-TEV-DyP2 was measured by ICP-OES and shows full heme occupancy.*

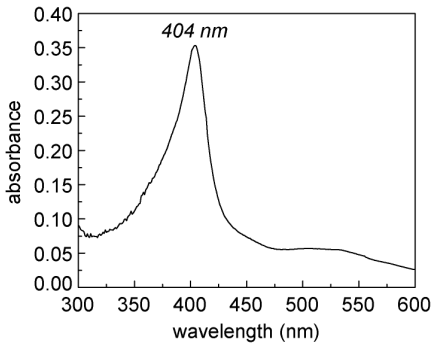
A



B



C



D

construct	SA (μmol mg^{-1} min^{-1})
His_{10}-TEV-DyP2	104
GA-DyP2	32

Figure 4.6. Comparison of His_{10}-TEV-DyP2 and GA-DyP2. *(A) SDS-PAGE analysis of His_{10}-TEV-DyP2 cleavage. Pre-induction sample (lane 1), post-induction sample (lane 2), post incubation with TEV (lane 3), SEC product (lane 4). (B) UV-visible spectrum of His_{10}-TEV-DyP2. (C) UV-visible spectrum of GA-DyP2. (D) Specific activities for 2,4-DCP oxidation at pH 7 with 4 mM H_2O_2.*

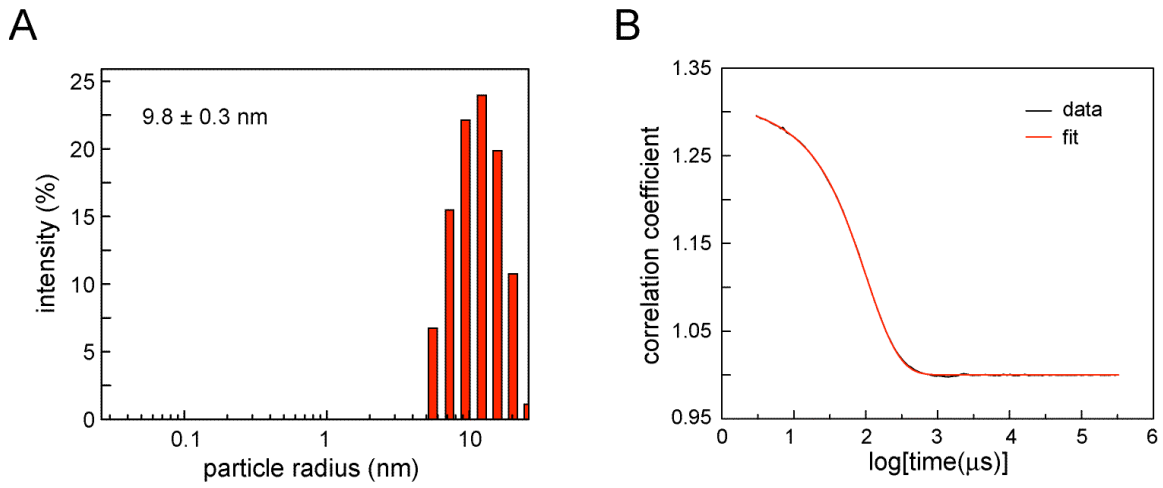


Figure 4.7. Biochemical characterization of DyP2 from A. sp 75iv2. *(A) Dynamic light scattering measurements of DyP2 in 20 mM Tris, 50 mM KCl, pH 7.5 indicate that its size in solution is consistent with an oligomer. (B) ASTRA fit of dynamic light scattering data from a monodisperse His10-TEV-DyP2 (0.5 mg/mL) sample in 20 mM Tris, 50 mM KCl, pH 7.5.*

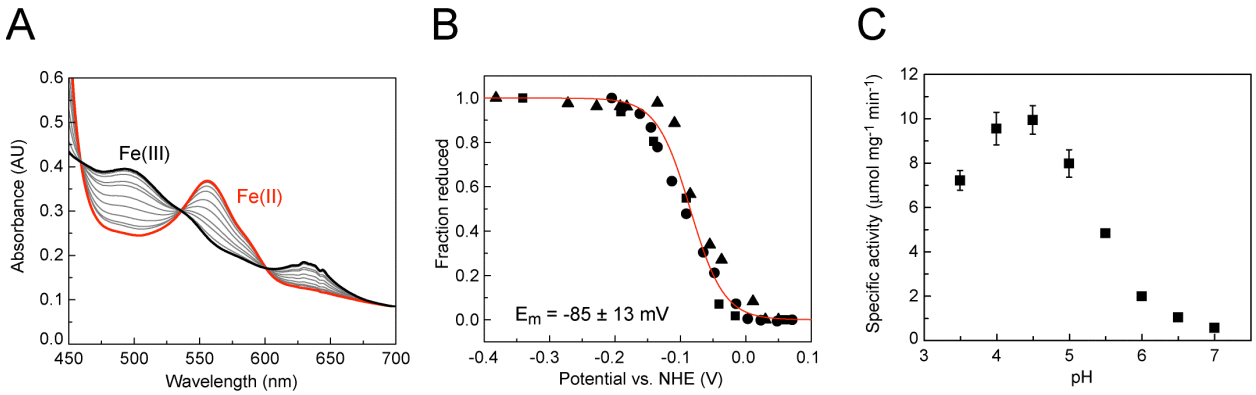
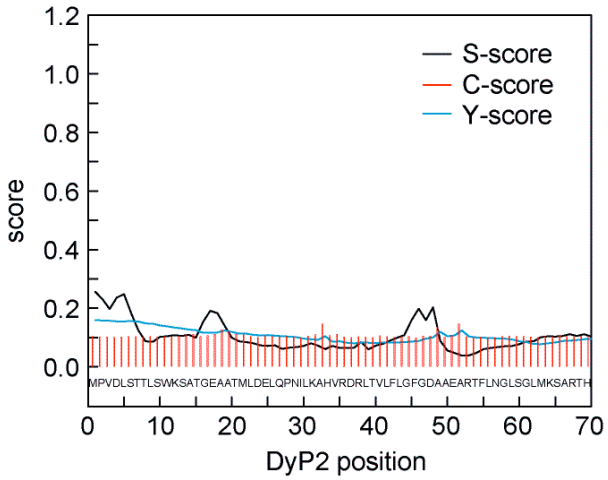


Figure 4.8. Biochemical characterization of DyP2 from A. sp 75iv2. *(A) Changes in the electron absorption spectrum indicating the stepwise oxidation of Fe2+–DyP2 (red) to Fe3+–DyP2 (black). (B) Three individual redox titrations of DyP2 monitored by ΔA560 nm. Data are reported as mean ± s.d. (C) pH–rate profile for the H2O2–dependent oxidation of Reactive Blue 5 by DyP2.*



	position	*value*	*cutoff*	*signal peptide?*
C_{max}	33	0.142		
Y_{max}	1	0.159		
S_{max}	1	0.256		
S_{max}	1-0	0.000		
D-score	1-0	0.097	0.45	NO

Figure 4.9. *Signal sequence analysis of DyP2. SignalP 4.0 was used to analyze the N-terminus of DyP2 for the presence of a signal sequence. The S-score and C-score respectively rank the likelihood of a signal peptide and cleavage site, while Y-max and the D-score represent a combined score for these properties. The low overall score for DyP2 in all of these categories indicate that DyP2 is unlikely to contain a canonical signal sequence.*

However, the most interesting behavior of DyP2 with respect to substrate scope is its high Mn^{2+} oxidation capacity. By analogy to fungal systems, manganese peroxidase activity has been thought to be important for microbial lignin degradation, taking advantage of the high Mn content of wood to produce the Mn^{3+} oxidant [*40*]. Indeed, a low level of Mn peroxidase activity was observed in DypB from *R. jostii* RHA1 [*11*]. We therefore tested DyP2 for its ability to oxidize Mn^{2+} to Mn^{3+} in the present of H_2O_2 and found a surprisingly high activity with a k_{cat}/K_M (k_{cat} = 24 ± 1 s^{-1}; K_M = 0.21 ± 0.03 mM) that is only 10-fold lower from the versatile peroxidase from *Pleurotus eryngii* and 100–fold lower than the Mn peroxidase from *Phanerochaete chrysosporium* (*Tables 4.2* and *4.3* [*41, 44, 45*]), which belongs to one of the most active lignolytic systems found to date. Although other heme-dependent peroxidases have been found to display this versatility [*45*], it has not yet been observed in any of the characterized DyPs.

DyP2 also demonstrates oxidase reactivity. Since no apparent reaction of DyP2 was observed with the non-phenolic lignin dimer **2**, we turned our attention to another lignin peroxidase substrate, MMA (**3**) [*46*], which can spontaneously decarboxylate upon ring oxidation and thereby increase the sensitivity for detecting these high potential oxidation events (ΔE > 1.4 V vs NHE) [*36*] that have been observed in fungal lignin breakdown pathways (*Figure 4.12*). In this assay, we observe a low rate of conversion of **3** into the anisaldehyde product (**4**) by DyP2 (*Figure 4.12*). The identity of **4** was confirmed by extraction of the DyP2 reaction product into ethyl acetate followed by GC–MS analysis compared to an authentic standard (*Figure 4.13*).

Further studies of this phenomenon have shown that this process is actually independent of added H_2O_2 and requires O_2 and Mn^{2+} instead (*Figures 4.12* and *4.14*). Indeed, the oxidative decarboxylation of **3** is catalyzed quantitatively by DyP2 in the presence of O_2 and Mn^{2+} alone. Controls omitting Mn^{2+} and O_2 show that both are required for the conversion of **3** to **4**. Further characterization of this reaction indicates that the pH maximum occurs around pH 5 and that the K_M for Mn^{2+} (0.76 ± 0.07 mM) is similar to that measured for the Mn peroxidase activity (*Table 4.3*, *Figure 4.14*). Thus, in addition to a high and versatile peroxidase activity, DyP2 also has an oxidase mode that is dependent on Mn^{2+}. Interestingly, the addition of Mn^{2+} is insufficient to convert the Fe^{3+}-resting state of DyP2 to an Fe^{2+}-state under aerobic conditions (*Figure 4.15*) that would be classically competent to interact with O_2 and a more complex catalytic cycle is possible that has been described for other peroxidases [*47, 48*]. More importantly, the oxidase mode of reactivity expands the substrate scope of DyP2 and allows for reaction with more difficult substrates either by direct oxidation or H atom abstraction.

Structural characterization of DyP2 from *A.* sp 75iv2. As the multiple reaction modes and high activity of DyP2 as a peroxidase, Mn peroxidase, and oxidase is exceptional for the DyP family, we next focused on structural studies to further elucidate its function. We crystallized DyP2 and solved its structure at 2.25 Å resolution by a combination of partial molecular replacement using the DyP from *Bjerkandera adusta* Dec1 [*10*] and experimental phasing obtained from selenomethionine–labeled DyP2 crystals (*Figure 4.16A*, *Table 4.4*). Two DyP2 monomers are present in the asymmetric unit as an apparent dimer (*Figure 4.17*), which is smaller than the size measured in solution (*Figures 4.4* and *4.7*). Although several DyPs are reported to be dimers [*4, 7, 37*], the DyP2 dimer observed in crystals appears to be non–native and induced by crystallization conditions.

Despite the large interface (970 $Å^2$ per monomer), the energy for dimer formation calculated by PISA [*50*] is small (ΔG ~ –0.6 kcal/mol) and is stabilized mainly by solvent molecules and

			k_{cat}/ K_M ($M^{-1}s^{-1}$)			
Clade	**Protein**	**Source**	**ABTS**	**Reactive Blue 5**	**Reactive Black 5**	**Mn^{2+}**
A	DypA	*Rhodococcus jostii* RHA1	$(2.0 \pm 0.1) \times 10^3$	–	–	ND
B	DypB	*Rhodococcus jostii* RHA1	$(2.4 \pm 0.1) \times 10^3$	–	–	$(2.5 \pm 0.01) \times 10^1$
	DyPPa	*Pseudomonas aeruginosa*	–	2.2×10^2	–	–
	TyrA	*Shewanella oneidensis*	–	7.0×10^4	–	–
C	**DyP2**	***A.* sp. 75iv2**	**$(6.6 \pm 0.9) \times 10^6$**	**$(7.1 \pm 0.9) \times 10^5$**	**$(1.6 \pm 0.1) \times 10^5$**	**$(1.2 \pm 0.2) \times 10^5$**
	AnaPX	*Nostoc* sp. PCC 7120	–	1.2×10^7	–	ND
D	DyP	*Bjerkandera adjusta* Dec1	–	4.8×10^6	–	–
	AjP I	*Auricularia auricula-judae*	1.8×10^7	5.0×10^6	5.7×10^5	ND
	AjP II	*Auricularia auricula-judae*	1.6×10^7	1.7×10^7	4.1×10^5	ND
	TAP	*Termitomyces albuminosus*	2.5×10^7	–	–	ND
	LiP	*Phanerochaete chrysosporium*	$(5.6 \pm 0.4) \times 10^5$	–	ND	ND
	MnP	*Phanerochaete chrysosporium*	ND	–	ND	$(1.2 \pm 0.1) \times 10^7$
	VP	*Pleurotus eryngii*	$(9.4 \pm 0.8) \times 10^6$	–	$(1.7 \pm 0.1) \times 10^6$	$(1.8 \pm 0.1) \times 10^6$

[a] Comparison of DyP2 kcat/KM's with other characterized DyPs [*2, 5-8, 11*] and canonical fungal peroxidases involved in lignin degradation.[*41, 44, 45, 49*] (ND, not detected; –, not determined).

Table 4.2. Peroxidase and Mn peroxidase activity of DyP2. [a]

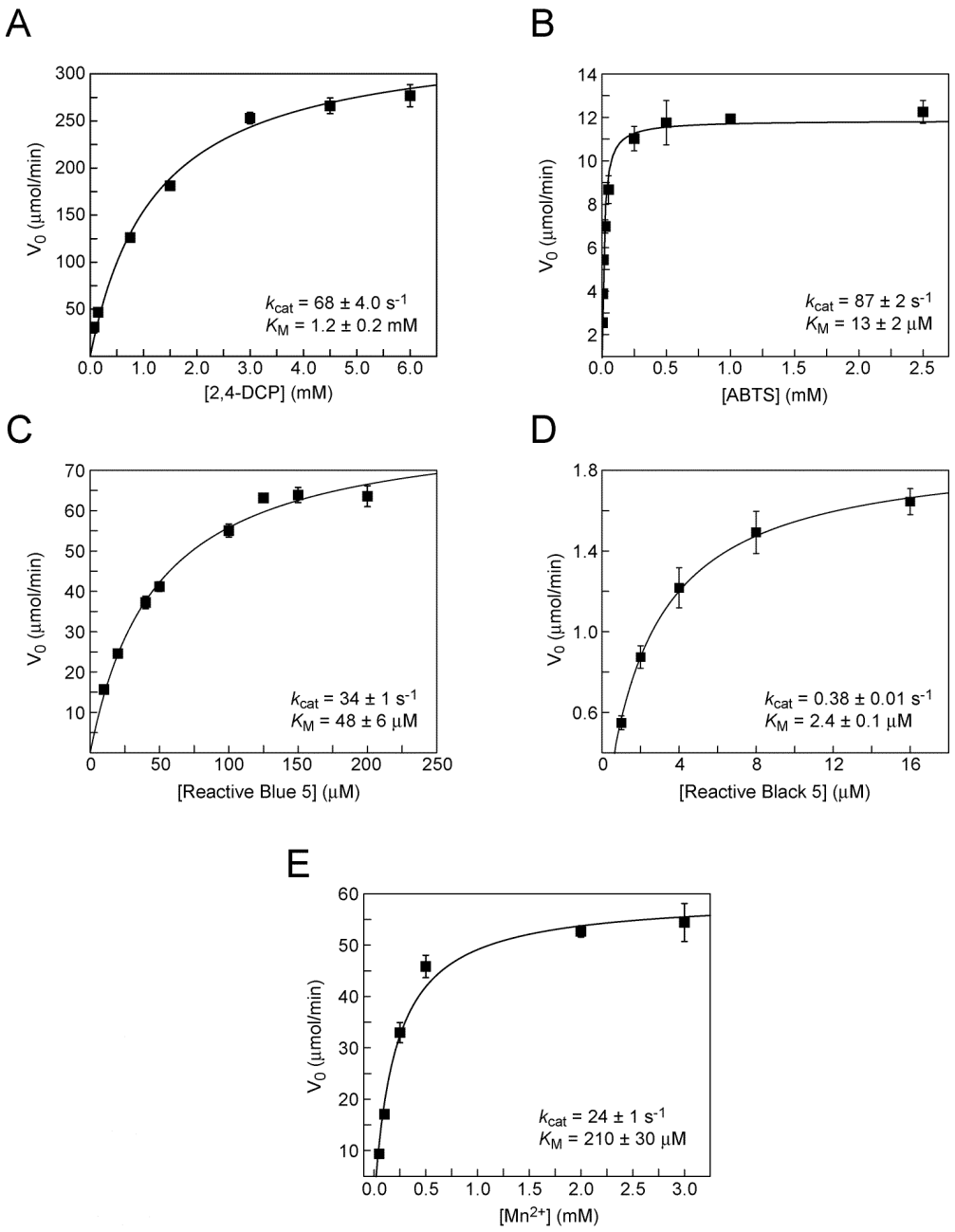


Figure 4.10. Dose-response curves for DyP2 peroxidase activity. *(A) 2,4-dichlorophenol. (B) 2,2´-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid). (C) Reactive Blue 5. (D) Reactive Black 5. (E) Mn^{2+}-malonate.*

Table 4.3. *Peroxidase kinetic parameters of characterized DyPs. Comparison of individual kinetic parameters for Amycolatopsis sp. 75iv2 DyP2 with literature values for other characterized DyPs and canonical heme peroxidases involved in lignin degradation (ND, not detected; -, not determined)--*

		ABTS		Reactive Blue 5		Reactive Black 5		Mn^{2+}		
protein	**source**	**k_{cat} (s^{-1})**	**K_M (μM)**	**k_{cat} (s^{-1})**	**K_M (μM)**	**k_{cat} (s^{-1})**	**K_M (μM)**	**k_{cat} (s^{-1})**	**K_M (μM)**	**reference**
DypA	*R. jostii* RHA1	16.8 ± 0.04	(8.2 ± 0.5) × 10^3	-	-	-	-	ND	-	[*11*]
DypB	*R. jostii* RHA1	55 ± 2	(2.3 ± 0.2) × 10^4	-	-	-	-	0.59 ± 0.04	(2.4 ± 0.2) x 10^4	[*11*]
DyPPa	*P. aeruginosa*	-	-	2.4 × 10^{-2}	107	-	-	-	-	[*7*]
TyrA	*S. oneidensis*	-	-	5.9	84	-	-	-	-	[*7*]
DyP2	***A. sp. 75iv2***	87 ± 2	13 ± 2	34 ± 1	48 ± 6	0.38 ± 0.01	2.4 ± 0.1	24 ± 1	210 ± 30	this work
AnaPX	*N.* sp. PCC 7120	-	-	384	3.6	-	-	ND	-	[*5*]
DyP	*B. adjusta* Dec1	-	-	260	54	0.97	-	-	-	[*8*]
AjP I	*A. auricula-judae*	368	20	114	23	3	5	ND	-	[*6*]
AjP II	*A. auricula-judae*	322	20	256	15	3	7	ND	-	[*6*]
TAP	*T. albuminosus*	698	28.4	-	-	-	-	ND	-	[*2*]
LiP	*P. chrysosporium*	13 ± 0.2	23 ± 1.6	-	-	ND	-	ND	-	[*49*]
MnP	*P. chrysosporium*	ND	-	-	-	ND	-	220	18 ± 2	[*41, 44*]
VP	*P. eryngii*	220 ± 30	540 ± 50	-	-	3.6	4.0	118	19	[*45*]

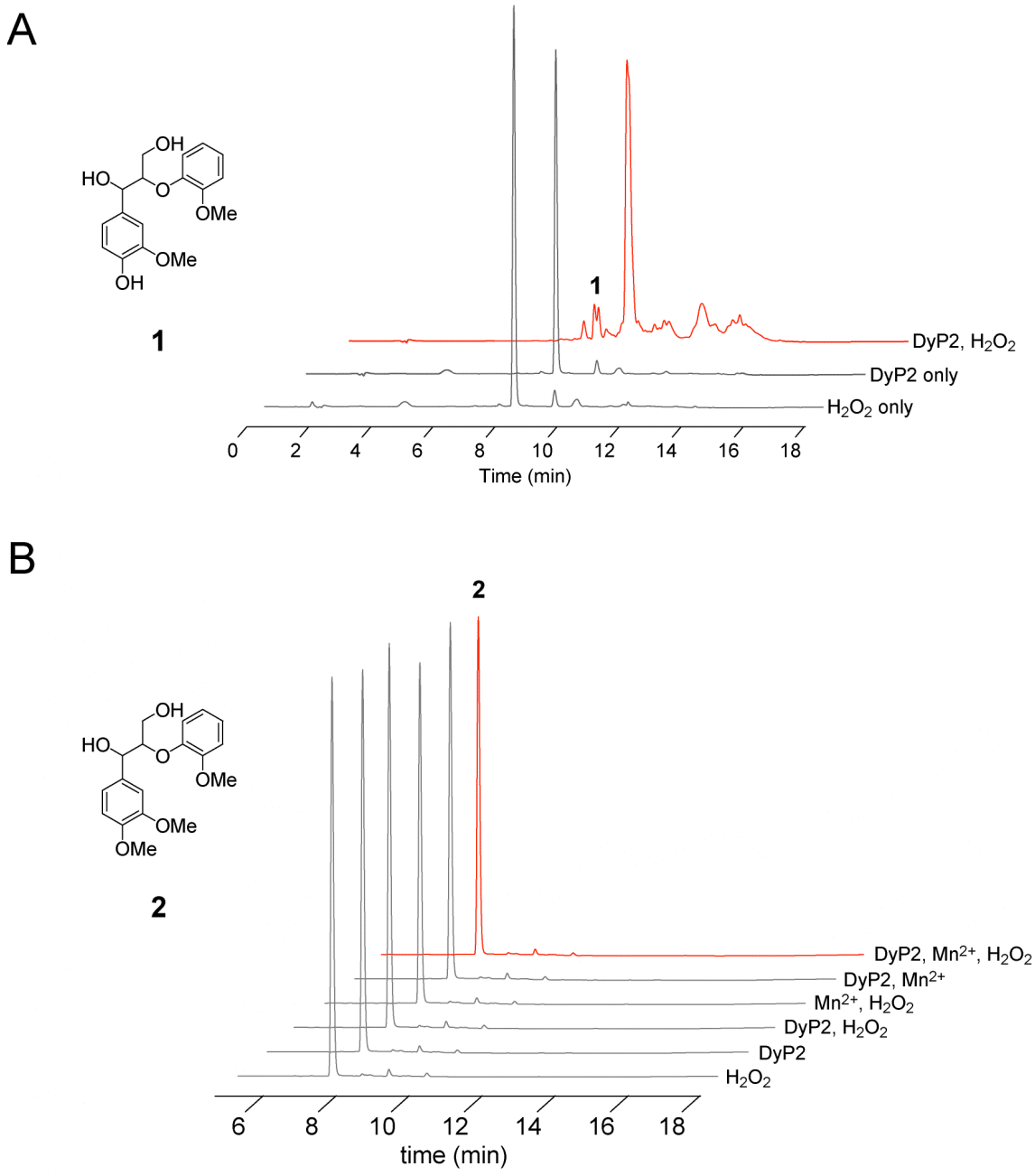


Figure 4.11. Peroxidase reactivity of DyP2 from A. sp. 75iv2 with lignin model dimers. *(A) HPLC traces showing the degradation of phenolic lignin model dimer (**1**) by DyP2 peroxidase activity monitored at A254 nm. (B) The reaction of DyP2 with non-phenolic lignin model dimer (**2**) was analyzed by $A_{254\ nm}$ using HPLC. No peroxidase- or Mn peroxidase-dependent degradation of **2** was observed under these conditions.*

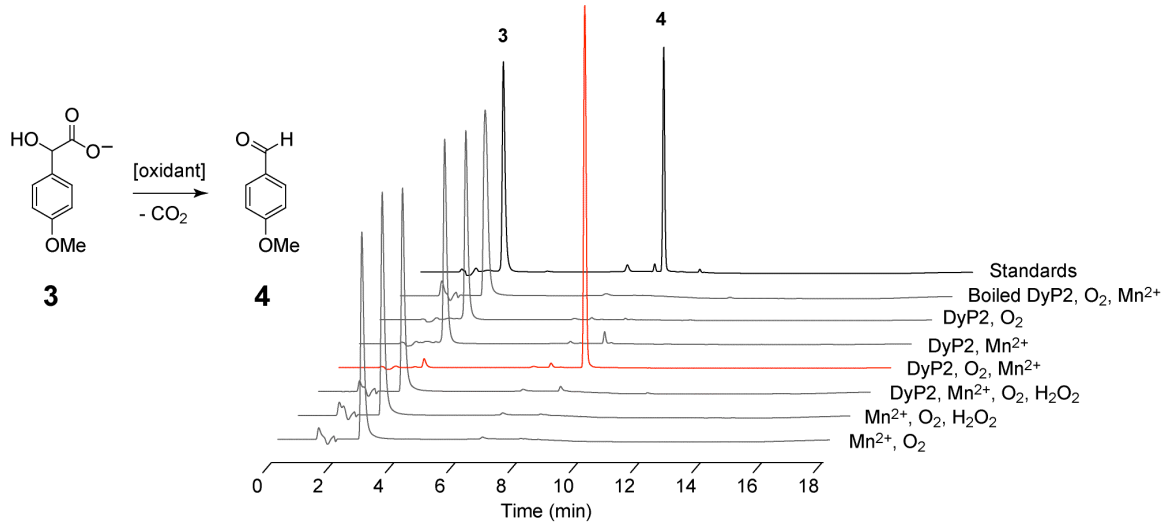


Figure 4.12. Oxidative reactivity of DyP2 reaction with non-phenolic substrate. *Oxidative decarboxylation of 4–methoxymandelic acid (**3**) to anisaldehyde (**4**). HPLC traces showing the O2- and Mn-dependent production of 4 catalyzed by DyP2 monitored by A247 nm.*

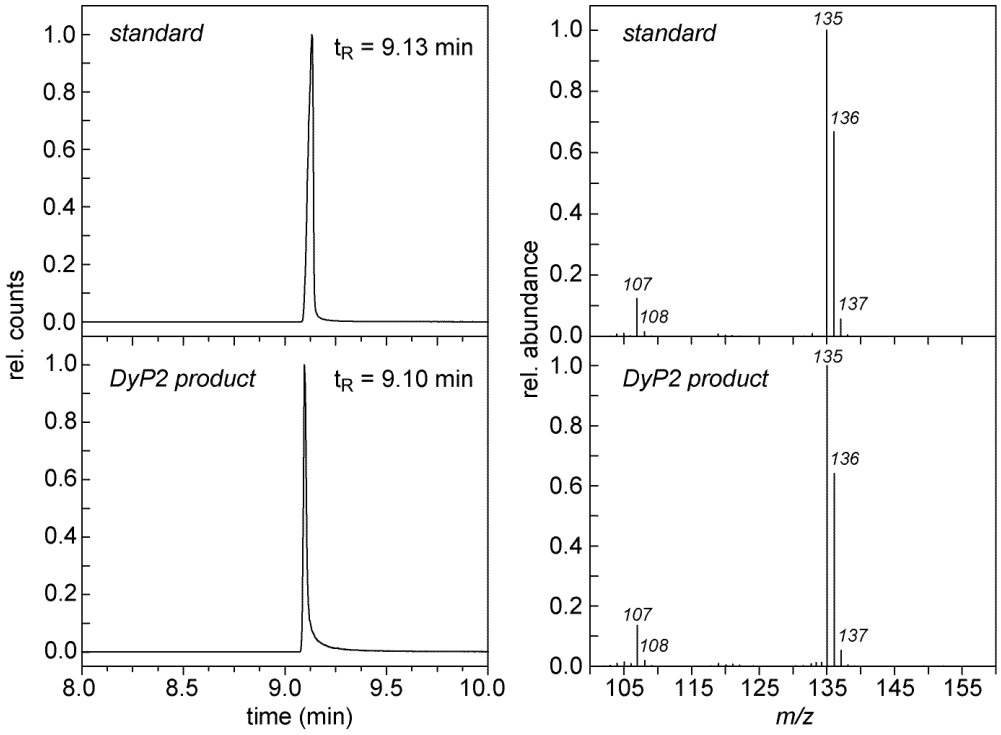


Figure 4.13. GC-MS analysis of the MMA reaction product of DyP2. *The product of the reaction of DyP2 with 4-methoxymandelic acid (MMA, **3**), O_2, and $MnCl_2$ was extracted into ethyl acetate and verified to be anisaldehyde (AAD, **4**) by comparison to an authentic standard (12.5 mM) using GC-MS.*

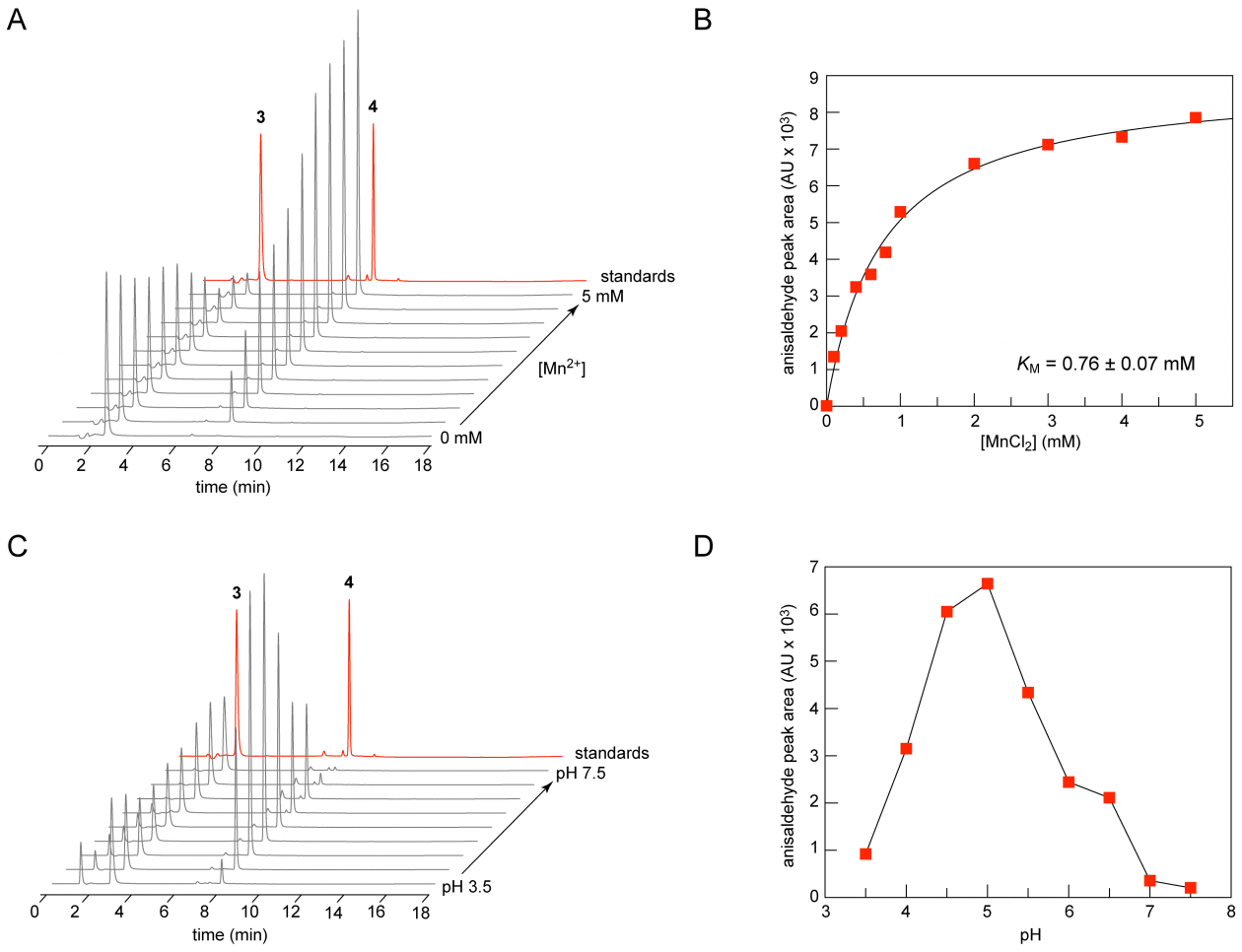


Figure 4.14. Mn- and pH-dependence of the DyP2 oxidase reaction. *The Mn- and pH-dependence for the conversion of MMA (**3**) to AAD (**4**) catalyzed by DyP2 in the presence of O_2 and $MnCl_2$ was monitored by $A_{247\ nm}$ using HPLC. (A) Chromatograms showing product formation as a function of $MnCl_2$ concentration (0-5 mM). (B) Plot of integrated area of the product peak (**4**). (C) Chromatograms showing product formation as a function of pH in the presence of 5 mM $MnCl_2$. (D) Plot of integrated area of the product peak (**4**) reveals the peak of DyP2 oxidase activity to reside between pH 4.5-5.0.*

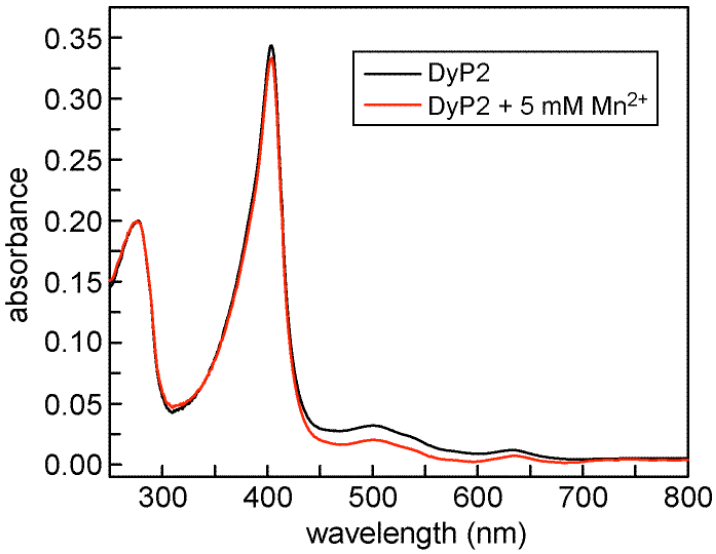


Figure 4.15. UV-visible spectra monitoring addition of $MnCl_2$ to Fe^{3+}-DyP2. *Addition of $MnCl_2$ (5 mM) to His_{10}-TEV-DyP2 was monitored by UV-visible spectroscopy and shows very little change in the electronic absorption spectrum upon the addition of Mn^{2+}.*

polar rather than hydrophobic interactions between the monomers (*Figure 4.17*). Therefore, the DyP2 oligomer observed in solution at pH 7.5 appears to be relatively weak and dissociable under crystallization conditions at pH 5. Since the maximal activity of DyPs, including DyP2, is typically found below pH 5 [*3, 5-7*], it is unclear whether the monomer, oligomer, or an equilibrium mixture represents the physiological form.

The overall α/β ferredoxin-like fold of the DyP2 protomer is shared with other structurally characterized DyPs [*3, 11, 37*] but differs in the arrangements of loops and α-helices that surround the β–barrel core (*Figure 4.16BC*). These loops appear to also create a deeper active site pocket in DyP2 (*Figures 4.16BC* and *4.18*) compared to the bacterial DypB from *R. jostii* RHA1, which has been suggested to be the origin of high peroxidase activity of the C- and D-type DyPs against a wide range of substrates [*11*]. The active site of DyP2 shows that the heme is ligated to H321 as the proximal residue with the conserved Arg (R346) and Asp (D190) in the distal pocket (*Figure 4.16D*), which mutational studies have shown to be important for DyP peroxidase activity [*3, 51*]. In addition, the distal heme pocket contains an ambiguous oxygen species (*Figure 4.19*). Based on the resolution of the data, this electron density has been assigned to a single mobile water molecule. However, the elongated density in this region of both the $2F_o–F_c$ and the $F_o–F_c$ maps leave the possibility that a diatomic oxygen adduct with an O–O distance of 1.3 Å may be present in the active site instead (*Figure 4.19*).

The most striking feature of the DyP2 crystal structure is the existence of a bound Mn ion in a binding site formed by E258, E273, E284, and a structured water held in place by a hydrogen bond to the main chain carbonyl of E273 (*Figure 4.16D*). Although the final DyP2 model does not show the E273 side–chain to be involved in Mn binding, the electron density maps indicate that it is fluxional and that a second conformation where the carboxylic acid moiety is directly coordinated to the Mn ion is also possible (*Figure 4.20*). This binding site does not appear to be common among DyPs based on sequence analysis and appears to be limited to a handful of other C-type DyPs (*Figure 4.21*). However, peroxidases from other families have been shown or predicted to contain Mn binding sites. For example, the crystal structure of the MnP from *P. chrysosporium* shows a Mn ion to be coordinated by three carboxylates, one heme propionate, as well as two waters [*52*]. The versatile peroxidases from *Pleurotus eryngii* have been shown to have a similar Mn binding site to that of the MnP [*34*]. Although the Mn site in DyP2 shows fluxionality, the $K_M(Mn^{2+})$ of DyP2 is only 10-fold higher than that of MnP. Indeed, the versatile peroxidase from *P. eryngii* also may share this fluxionality in the coordination of one of the glutamate side-chains to the Mn and demonstrates the same affinity for Mn^{2+} as MnP [*34*]. We also find a potential redox-active amino acid (Y188) in between the heme and Mn pocket that could serve to communicate between these sites, which are separated by 15 Å. Overall, the existence of a distinct binding site for Mn demonstrates that the Mn-dependent behavior of DyP2 both as a peroxidase and oxidase likely plays a key role in its catalytic function.

	SeMet-DyP2	native DyP2
data collection		
wavelength (Å)	0.9796	1.116
resolution (Å)[a]	48.92-3.01 (3.17-3.01)	49.04-2.25 (2.37-2.25)
space group	$P2_12_12_1$	$P2_12_12_1$
cell dimensions		
a, *b*, *c* (Å)	49.892, 77.672, 249.484	48.131, 77.952, 252.353
α, β, γ (deg)	90, 90, 90	90, 90, 90
no. of total reflections[a]	284795 (38677)	200541 (22364)
no. of unique reflections[a]	20191 (2850)	45753 (6301)
multiplicity[a]	7.6 (7.2)	4.4 (3.5)
completeness (%)[a]	99.7 (97.8)	98.9 (95.0)
mean $I/\sigma I$[a]	16.9 (3.2)	13.0 (2.4)
R-merge[a]	0.168 (0.917)	0.088 (0.556)
Wilson B ($Å^2$)	80.4	33.13
refinement		
R_{work} (R_{free})		0.1763 (0.2368)
no. of atoms		7832
protein		7071
heme		244
water		399
RMSD		
bond lengths (Å)		0.008
bond angles (deg)		1.11
Ramachandran favored (%)		96
Ramachandran outliers (%)		0.22
MolProbity clashscore [*53*]		18.8
average B-factor		33.8
macromolecules		33.6
solvent		35.7

[a] Values in parentheses are for the highest resolution shell.

Table 4.4. Data collection and refinement statistics for the His_{10}-TEV-DyP2 crystal structure. *The initial model for DyP2 was built using a partial molecular replacement solution using B. adusta Dec1 DyP (PDB ID 2D3Q) as a search model and experimental phases determined from SeMet-His_{10}-TEV-DyP2 (3.0 Å resolution) using Phaser. The final model was generated using a 2.25 Å resolution dataset for His_{10}-TEV-DyP2 through multiple rounds of manual building in COOT and refinement using the PHENIX package.*

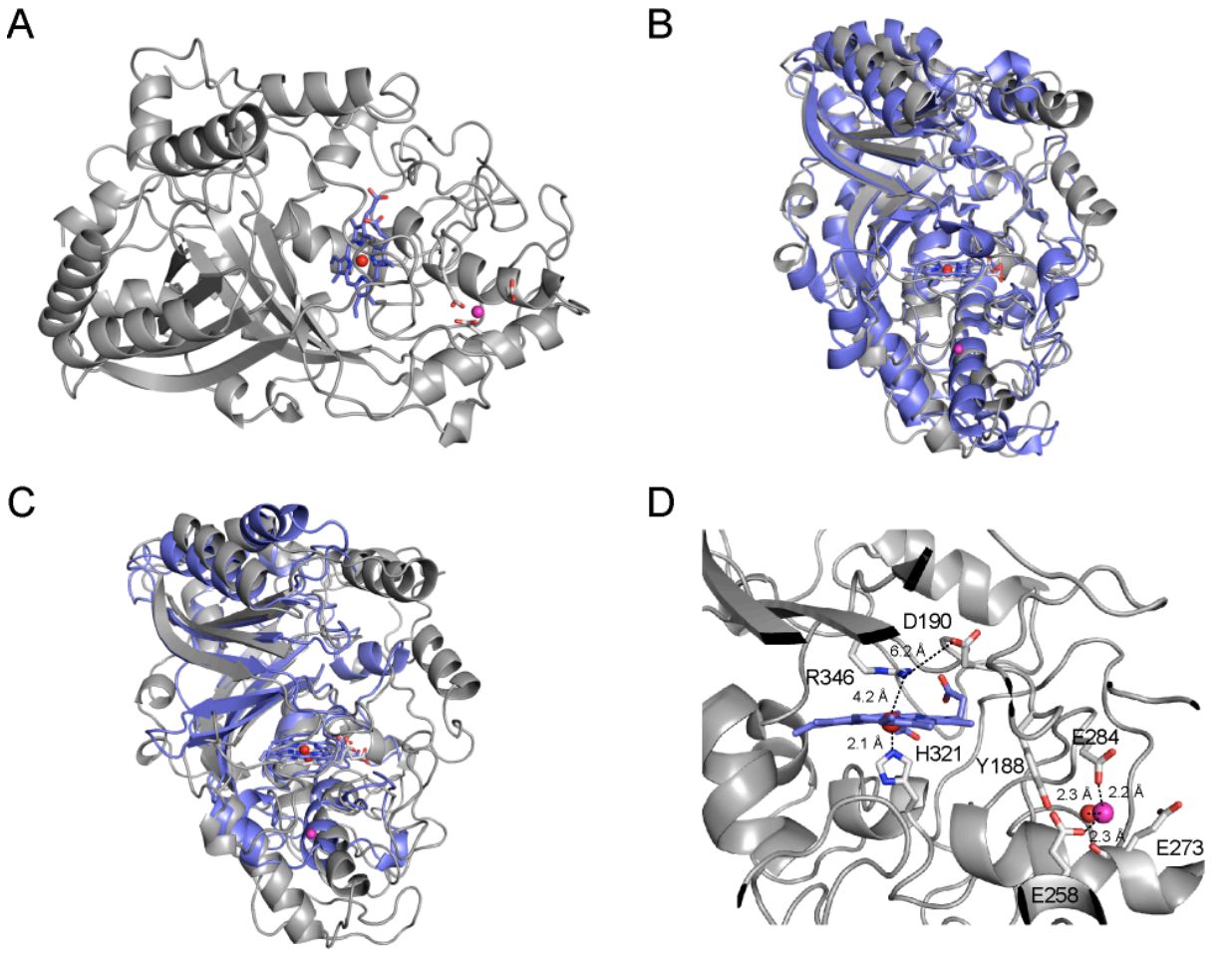


Figure 4.16. Structure of DyP2 from A. sp. 75iv2. *(A) Ribbon representation of DyP2 showing the overall fold of a DyP2 monomer and the heme site (Fe, red) and bound Mn ion (magenta), which are 15 Å apart. (B) Structural overlay of one DyP2 monomer (grey) with the D–type B. adusta Dec1 DyP (blue, PDB ID 2D3Q, 28% sequence identity) shows an RMSD of 2.7 Å for 417 aligned Cα atoms. (C) Structural overlay of one DyP2 monomer (grey) with the B–type DypB from R. jostii RHA1 (blue, PDB ID 3QNR). (D) View of the DyP2 active site of DyP2 showing the proximal H321 and distal pocket residues, D190 and R346. A Mn binding site near to the heme active site (15 Å apart) is formed by E258, E273, E284, and a structured water held in place by the main chain carbonyl group of E273 (2.7 Å). E273 is also found in a second conformation where it is directly coordinated to the Mn ion. The Y188 is also shown and is located 3.9 Å from the Mn ion and 4.8 Å from the heme site.*

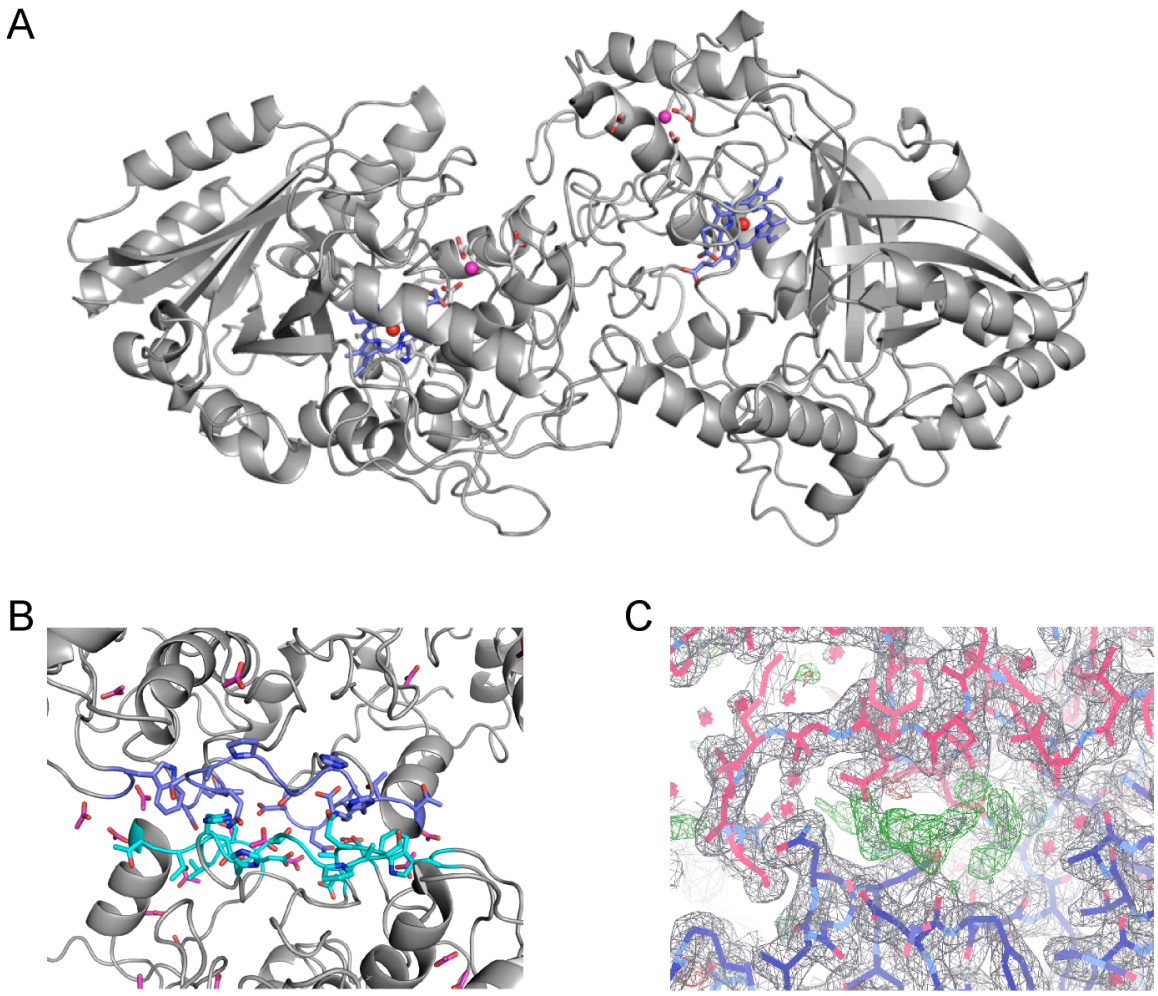


Figure 4.17. DyP2 dimer interface. *The DyP2 dimer appears to have crystallized in a non-native dimeric state. (A) Ribbon representation of DyP2 showing the DyP2 dimer with the heme site (red) and bound Mn ion (magenta). (B) Ribbon representation of the polar dimer interface formed by acetate molecules (pink) and protein loops containing residues 210-216, 293-301, 310-311, and 431-432. (C) Electron density map for the dimer interface structure at 1.3s shows unmodeled density that likely represents a PEG molecule.*

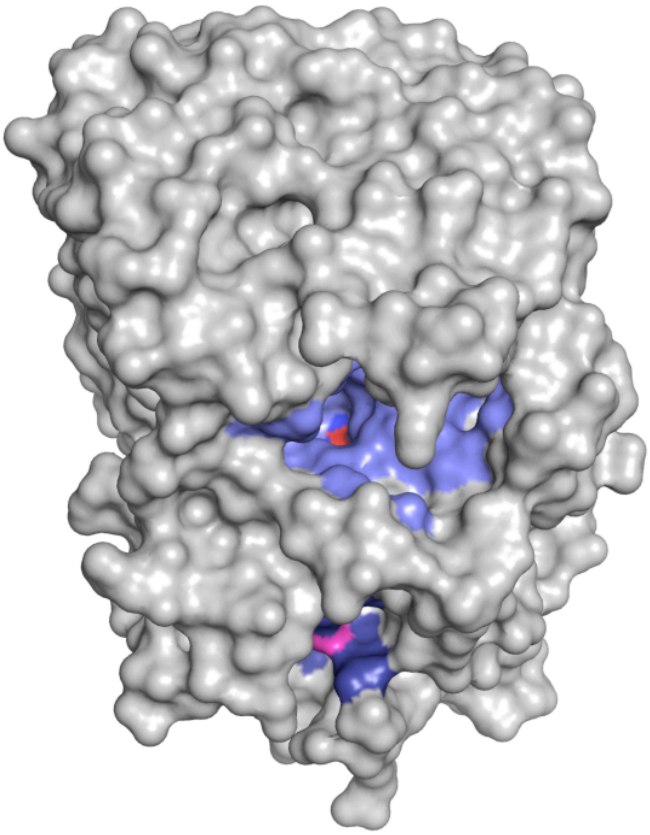


Figure 4.18. DyP2 heme and Mn binding site channels. *Surface rendering of DyP2 showing the deep chanels leading to the heme site (light blue; Fe, red) and Mn binding pocket (dark blue; Mn, magenta).representation of DyP2 showing the DyP2 dimer with the heme site (red) and bound Mn ion (magenta).*

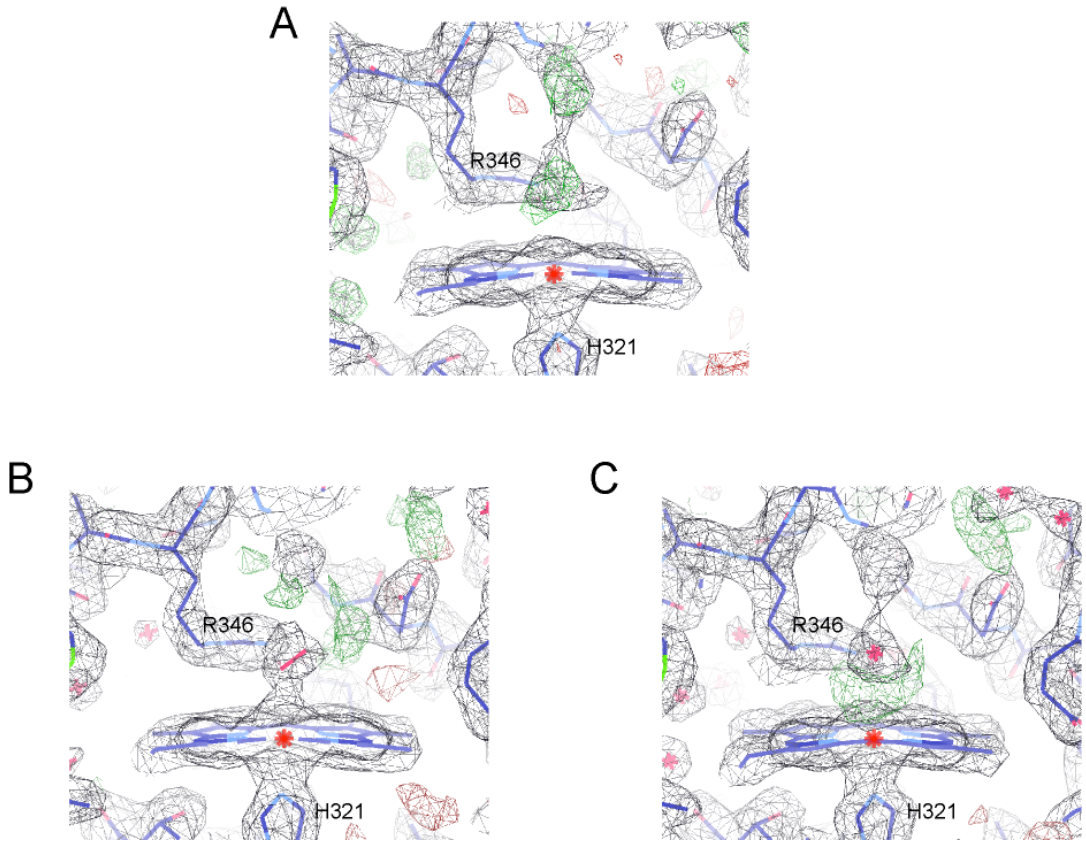


Figure 4.19. Analysis of the distal oxygenic species. *An elongated electron density is observed within the distal pocket of the heme active site, which could either represent a mobile water or a diatomic oxygen species. (A) Electron density map of active site at 1.2σ, unbiased for solvent molecules. (B) Electron density map of the active site at 1.2σ refined with a diatomic oxygen species included in the model (Fe–O distance = 2.7 Å). (C) Electron density map of the active site at 1.2σ refined with a mobile water included in the model. (Fe–O distance = 3.8 Å) (green, positive peaks in the Fo-Fc map; red, negative peaks in the Fo-Fc map)*

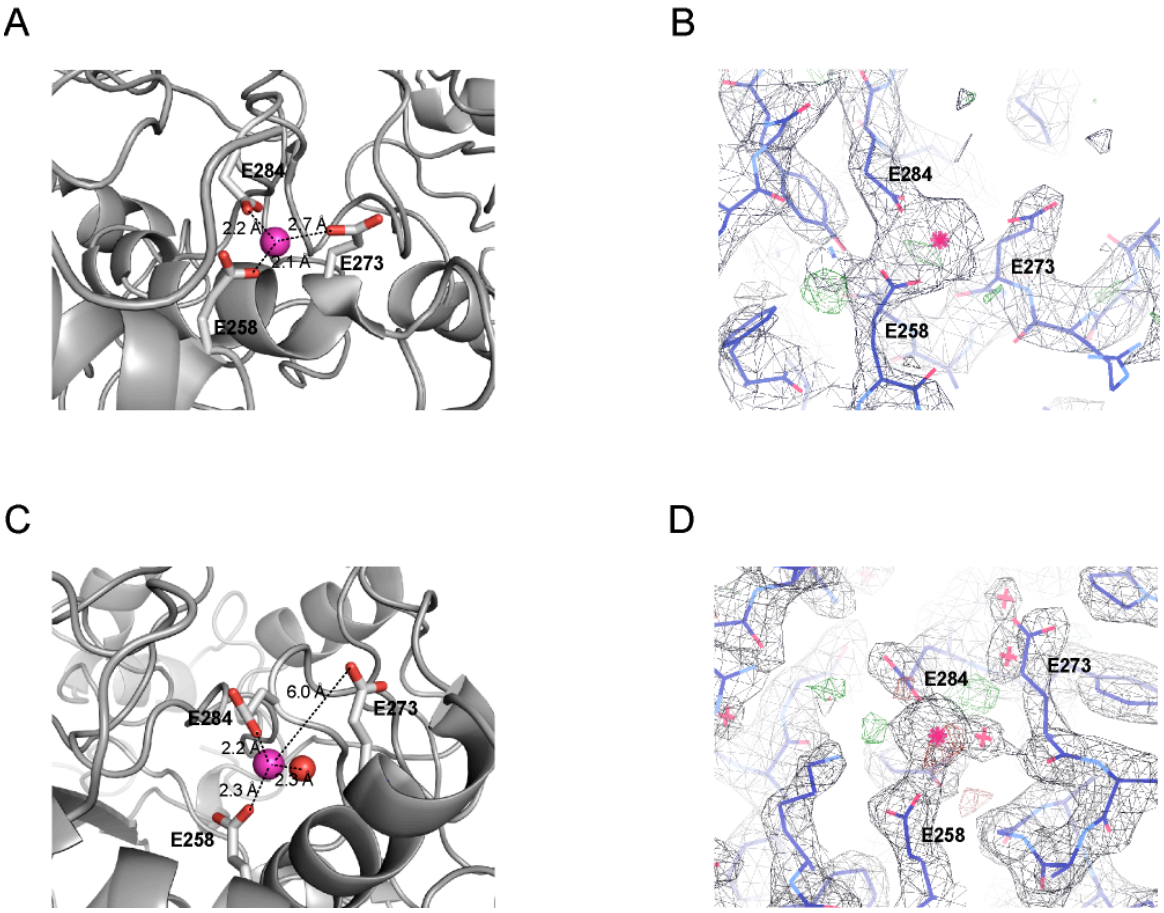


Figure 4.20. The DyP2 Mn binding site. *E273 is a fluxional residue that appears to exist in a second conformation where the carboxylate side-chain can coordinate directly to Mn. (A) Mn binding site in the SeMet-His$_{10}$-TEV-DyP2 structure. (B) Electron density map for the SeMet-His$_{10}$-TEV-DyP2 structure at 1.1s. Electron density from a water molecule (green) coordinated by the Mn and E273 main-chain carbonyl group may exist but is difficult to distinguish from the bound Mn at this resolution. (C) Mn binding site in the native His$_{10}$-TEV-DyP2 structure. (D) Electron density map for the native His$_{10}$-TEV-DyP2 structure at 1.3s.*

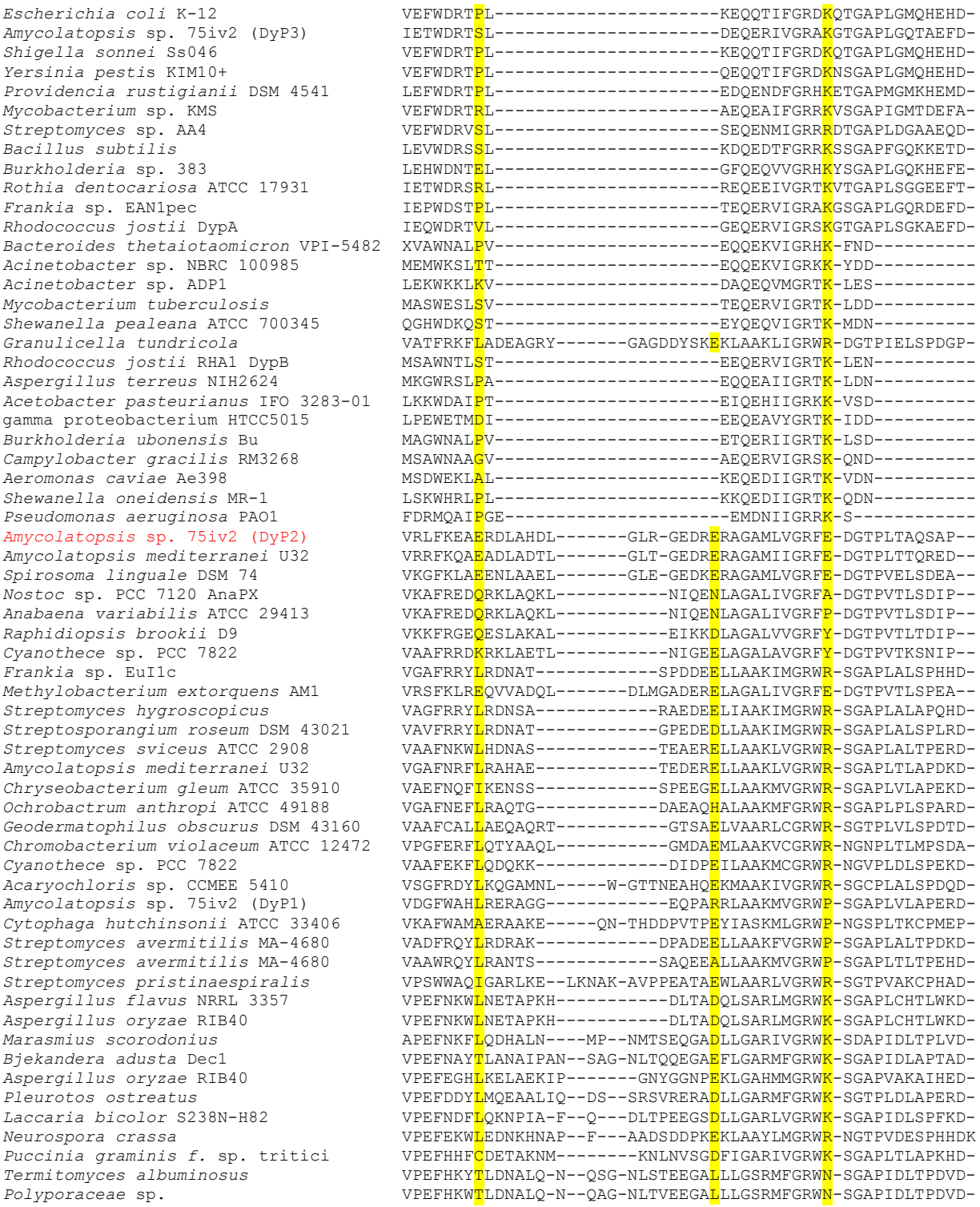


Figure 4.21. Sequence alignments for Mn-binding pocket residues. *Mn-binding site residues in DyP2 (E258, E273, E284) are found to be conserved in only a limited number of organisms.*

4.4. Conclusions

In summary, we have identified and characterized DyP2 from *Amycolatopsis* sp. 75iv2 and found that it is a versatile and multifunctional member of the DyP family. As a peroxidase, it demonstrates high activity against a broad range of peroxidase substrates, including the high–potential Reactive Black 5. In addition, DyP2 demonstrates Mn peroxidase activity near the same order of magnitude as canonical Mn peroxidases and versatile peroxidases that are involved in fungal pathways for lignin breakdown and take advantage of the high Mn content of wood and other plant material. The observed high activity differs from other characterized bacterial DyPs but is consistent with its unusual phylogeny shared with DyPs from related actinomycetes that demonstrate the capacity to depolymerize biomass. DyP2 also demonstrates a second mode of oxidase reactivity that was predicted for the MsP1 and MsP2 DyPs [*4*] but may be more mechanistically complex than canonical oxidases given its Fe^{3+} resting state. Interestingly, the oxidase mode is also dependent on Mn^{2+}, with a similar $K_M(Mn^{2+})$ as the Mn peroxidase activity, and broadens its substrate scope to include more challenging substrates. Crystallographic studies show that a distinct Mn binding site exists and indicate that Mn has an essential physiological role in DyP2 function.

4.5. References

1. Y. Sugano, DyP-type peroxidases comprise a novel heme peroxidase family, *Cell. Mol. Life Sci.* **2009**, *66*, 1387-1403.
2. T. Johjima, M. Ohkuma and T. Kudo, Isolation and cDNA cloning of novel hydrogen peroxide-dependent phenol oxidase from the basidiomycete *Termitomyces albuminosus*, *Appl. Microbiol. Biotechnol.* **2003**, *61*, 220-225.
3. Y. Sugano, R. Muramatsu, A. Ichiyanagi, T. Sato and M. Shoda, DyP, a unique dye-decolorizing peroxidase, represents a novel heme peroxidase family, *J. Biol. Chem.* **2007**, *282*, 36652-36658.
4. M. Scheibner, B. Hülsdau, K. Zelena, M. Nimtz, L. de Boer, R. Berger and H. Zorn, Novel peroxidases of *Marasmius scorodonius* degrade β-carotene, *Appl. Microbiol. Biotechnol.* **2008**, *77*, 1241-1250.
5. H. J. O. Ogola, T. Kamiike, N. Hashimoto, H. Ashida, T. Ishikawa, H. Shibata and Y. Sawa, Molecular characterization of a novel peroxidase from the cyanobacterium *Anabaena* sp. Strain PCC 7120, *Appl. Env. Microbiol.* **2009**, *75*, 7509-7518.
6. C. Liers, C. Bobeth, M. Pecyna, R. Ullrich and M. Hofrichter, DyP-like peroxidases of the jelly fungus *Auricularia auricula-judae* oxidize nonphenolic lignin model compounds and high-redox potential dyes, *Appl. Microbiol. Biotechnol.* **2010**, *85*, 1869-1879.
7. J. Li, C. Liu, B. Li, H. Yuan, J. Yang and B. Zheng, Identification and molecular characterization of a novel DyP-type peroxidase from *Pseudomonas aeruginosa* PKE117, *Appl. Biochem. Biotechnol.* **2012**, *166*, 774-785.
8. S. J. Kim and M. Shoda, Purification and characterization of a novel peroxidase from *Geotrichum candidum* Dec 1 involved in decolorization of dyes, *Appl. Env. Microbiol.* **1999**, *65*, 1029-1035.

9. Y. Sugano, Y. Matsushima and M. Shoda, Complete decolorization of the anthraquinone dye Reactive blue 5 by the concerted action of two peroxidases from *Thanatephorus cucumeris* Dec 1, *Appl. Microbiol. Biotechnol.* **2006**, *73*, 862-871.

10. T. Sato, S. Hara, T. Matsui, G. Sazaki, S. Saijo, T. Ganbe, N. Tanaka, Y. Sugano and M. Shoda, A unique dye-decolorizing peroxidase, DyP, from *Thanatephorus cucumeris* Dec 1: Heterologous expression, crystallization and preliminary X-ray analysis, *Acta Crystallogr. D* **2004**, *60*, 149-152.

11. J. N. Roberts, R. Singh, J. C. Grigg, M. E. P. Murphy, T. D. H. Bugg and L. D. Eltis, Characterization of dye-decolorizing peroxidases from *Rhodococcus jostii* RHA1, *Biochemistry* **2011**, *50*, 5108-5119.

12. M. Ahmad, J. N. Roberts, E. M. Hardiman, R. Singh, L. D. Eltis and T. D. H. Bugg, Identification of DypB from *Rhodococcus jostii* RHA1 as a lignin peroxidase, *Biochemistry* **2011**, *50*, 5096-5107.

13. R. Rahmanpour and T. D. H. Bugg, Assembly *in vitro* of *Rhodococcus jostii* RHA1 encapsulin and peroxidase DypB to form a nanocompartment, *FEBS J.* **2013**, *280*, 2097-2104.

14. M. E. Brown, M. C. Walker, T. G. Nakashige, A. T. Iavarone and M. C. Y. Chang, Discovery and characterization of heme enzymes from unsequenced bacteria: Application to microbial lignin degradation, *J. Am. Chem. Soc.* **2011**, *133*, 18006-18009.

15. J. R. Davis*, et al.*, Genome sequence of *Amycolatopsis* sp. strain ATCC 39116, a plant biomass-degrading actinomycete, *J. Bacteriol.* **2012**, *194*, 2396-2397.

16. J. Dyrløv Bendtsen, H. Nielsen, G. von Heijne and S. Brunak, Improved prediction of signal peptides: SignalP 3.0, *J. Mol. Biol.* **2004**, *340*, 783-795.

17. T. U. Consortium, Reorganizing the protein space at the Universal Protein Resource (UniProt), *Nucleic Acids Res.* **2012**, *40*, D71-D75.

18. W. Li, L. Jaroszewski and A. Godzik, Clustering of highly homologous sequences to reduce the size of large protein databases, *Bioinformatics* **2001**, *17*, 282-283.

19. S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, *Nucleic Acids Res.* **1997**, *25*, 3389-3402.

20. F. Armougom, S. Moretti, O. Poirot, S. Audic, P. Dumas, B. Schaeli, V. Keduas and C. Notredame, Expresso: Automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee, *Nucleic Acids Res.* **2006**, *34*, W604-W608.

21. C. Notredame, D. G. Higgins and J. Heringa, T-coffee: A novel method for fast and accurate multiple sequence alignment, *J. Mol. Biol.* **2000**, *302*, 205-217.

22. K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei and S. Kumar, MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Mol. Biol. Evol.* **2011**, *28*, 2731-2739.

23. R. Makino, S.-y. Park, E. Obayashi, T. Iizuka, H. Hori and Y. Shiro, Oxygen binding and redox properties of the heme in soluble guanylate cyclase, *J. Biol. Chem.* **2011**, *286*, 15678-15687.

24. P. L. Dutton, Redox potentiometry: Determination of midpoint potentials of oxidation-reduction components of biological electron-transfer systems, *Methods Enzymol.* **1978**, *54*, 411-435.
25. W. Kabsch, XDS, *Acta Crystallogr. D* **2010**, *66*, 125-132.
26. Collaborative, The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D* **1994**, *50*, 760-763.
27. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, Phaser crystallographic software, *J. Appl. Crystallogr.* **2007**, *40*, 658-674.
28. K. Y. J. Zhang, K. Cowtan and P. Main, Combining constraints for electron-density modification, *Methods Enzymol.* **1997**, *277*, 53-64.
29. M. D. Winn*, et al.*, Overview of the CCP4 suite and current developments, *Acta Crystallogr. D* **2011**, *67*, 235-242.
30. P. Emsley and K. Cowtan, Coot: Model-building tools for molecular graphics, *Acta Crystallogr. D* **2004**, *60*, 2126-2132.
31. P. D. Adams*, et al.*, PHENIX: A comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D* **2010**, *66*, 213-221.
32. Y. Sugano, Y. Ishii and M. Shoda, Role of H164 in a unique dye-decolorizing heme peroxidase DyP, *Biochem. Biophys. Res. Comm.* **2004**, *322*, 126-132.
33. R. Bourbonnais and M. G. Paice, Oxidation of non-phenolic substrates: An expanded role for laccase in lignin biodegradation, *FEBS Lett.* **1990**, *267*, 99-102.
34. M. Pérez-Boada, F. J. Ruiz-Dueñas, R. Pogni, R. Basosi, T. Choinowski, M. J. Martínez, K. Piontek and A. T. Martínez, Versatile peroxidase oxidation of high redox potential aromatic compounds: Site-directed mutagenesis, spectroscopic and crystallographic investigation of three long-range electron transfer pathways, *J. Mol. Biol.* **2005**, *354*, 385-402.
35. H. Wariishi, K. Valli and M. H. Gold, Manganese(II) oxidation by manganese peroxidase from the basidiomycete *Phanerochaete chrysosporium*. Kinetic mechanism and role of chelators, *J. Biol. Chem.* **1992**, *267*, 23688-23695.
36. E. Baciocchi, M. F. Gerini, O. Lanzalunga and S. Mancinelli, Lignin peroxidase catalysed oxidation of 4-methoxymandelic acid. The role of mediator structure, *Tetrahedron* **2002**, *58*, 8087-8093.
37. C. Zubieta*, et al.*, Crystal structures of two novel dye-decolorizing peroxidases reveal a β-barrel fold with a conserved heme-binding motif, *Proteins Struct. Funct. Bioinform.* **2007**, *69*, 223-233.
38. G. Battistuzzi, M. Bellei, C. A. Bortolotti and M. Sola, Redox properties of heme peroxidases, *Arch. Biochem. Biophys.* **2010**, *500*, 21-36.
39. C. D. Millis, D. Cai, M. T. Stankovich and M. Tien, Oxidation-reduction potentials and ionization states of extracellular peroxidases from the lignin-degrading fungus *Phanerochaete chrysosporium*, *Biochemistry* **1989**, *28*, 8484-8489.
40. R. ten Have and P. J. Teunissen, Oxidative mechanisms involved in lignin degradation by white-rot fungi, *Chem. Rev.* **2001**, *101*, 3397-413.

41. A. Heinfling, F. J. Ruiz-Dueñas, M. J. Martínez, M. Bergbauer, U. Szewzyk and A. T. Martínez, A study on reducing substrates of manganese-oxidizing peroxidases from *Pleurotus eryngii* and *Bjerkandera adusta*, *FEBS Lett.* **1998**, *428*, 141-146.

42. T. Katayama, F. Nakatsubo and T. Higuchi, Syntheses of arylglycerol-β-aryl ethers, *Mokuzai Gakkaishi* **1981**, *27*, 223-230.

43. K. Wei, S.-W. Luo, Y. Fu, L. Liu and Q.-X. Guo, A theoretical study on bond dissociation energies and oxidation potentials of monolignols, *J. Mol. Struct. THEOCHEM* **2004**, *712*, 197-205.

44. I. C. Kuan, K. A. Johnson and M. Tien, Kinetic analysis of manganese peroxidase. The reaction with manganese complexes., *J. Biol. Chem.* **1993**, *268*, 20064-20070.

45. L. Banci, S. Camarero, A. Martínez, M. Martínez, M. Pérez-Boada, R. Pierattelli and F. Ruiz-Dueñas, NMR study of manganese(II) binding by a new versatile peroxidase from the white-rot fungus *Pleurotus eryngii*, *J. Biol. Inorg. Chem.* **2003**, *8*, 751-760.

46. M. Tien and D. Ma, Oxidation of 4-methoxymandelic acid by lignin peroxidase, *J. Biol. Chem.* **1997**, *272*, 8912-8917.

47. I. Yamazaki and L. H. Piette, The mechanism of aerobic oxidase reaction catalyzed by peroxidase, *Biochim. Biophys. Acta* **1963**, *77*, 47-64.

48. A. Paszczynski, V.-B. Huynh and R. Crawford, Comparison of ligninase-I and peroxidase-M2 from the white-rot fungus *Phanerochaete chrysosporium*, *Arch. Biochem. Biophys.* **1986**, *244*, 750-765.

49. E. Fernández-Fueyo, F. J. Ruiz-Dueñas, Y. Miki, M. J. Martínez, K. E. Hammel and A. T. Martínez, Lignin-degrading peroxidases from genome of selective ligninolytic fungus *Ceriporiopsis subvermispora*, *J. Biol. Chem.* **2012**, *287*, 16903-16916.

50. E. Krissinel and K. Henrick, Inference of macromolecular assemblies from crystalline state, *J. Mol. Biol.* **2007**, *372*, 774-797.

51. R. Singh, J. C. Grigg, Z. Armstrong, M. E. P. Murphy and L. D. Eltis, Distal heme pocket residues of B-type dye-decolorizing peroxidase: Arginine but not aspartate is essential for peroxidase activity, *J. Biol. Chem.* **2012**, *287*, 10623-10630.

52. M. Sundaramoorthy, H. Youngs, M. H. Gold and T. L. Poulos, High-resolution crystal structure of manganese peroxidase: Substrate and inhibitor complexes, *Biochemistry* **2005**, *44*, 6463-6470.

53. I. W. Davis*, et al.*, MolProbity: All-atom contacts and structure validation for proteins and nucleic acids, *Nucleic Acids Res.* **2007**, *35*, W375-W383.

Chapter 5: *Using cell profiling to explore the molecular basis for lignin metabolism by Amycolatopsis sp. 75iv2*

5.1. Introduction

While the role of fungi in the environmental degradation of lignin has been well established [*1, 2*], it remains unclear to what extent that bacteria participate in this process. Previous studies have shown that some species of soil bacteria are competent to partially metabolize ^{14}C-labeled lignin [*3-5*] or synthetic lignin to $^{14}CO_2$ [*6, 7*] or to grow on lignin on a sole carbon source [*4-6*]. Furthermore, some initial information indicates that these organisms may display differing specificity towards different structural motifs found in lignin based on motif-specific ^{13}C-labelling [*6, 7*]. However, we remain limited in our molecular understanding of how bacteria may react with lignin. Work on streptomycetes has correlated the existence of an extracellular peroxidase lignin solubilization and formation of acid-precipitable polymer lignin (APPL), which is thought to be a signature of bacterial lignin metabolism [*8-10*]. Interestingly, where this peroxidase can degrade non-phenolic model lignin dimmers, the identity of the protein sequence remains unpublished. More recent studies of *Rhodococcus jostii* RHA-1 has also implicated a heme-containing dye peroxidase (DyP) to be directly involved with its observed lignolytic behavior [*11-13*]. Based on these these observations, it appears as if bacterial lignin degradation may also proceed through peroxidase-dependent oxidative processes. Given their weak oxidative power compared to fungi, however, we may expect a larger suite of accessory enzymes that may be involved in creating more sites for degradation or enzymes with alternative product distribution.

We set out to characterize *Amycolatopsis* sp. 75iv2 as a model bacterial host for lignin metabolism. Our previous studies have shown that this strain is quite active in the formation of APPL and secretion of extracellular peroxidases [*14*]. Furthermore, its genome encodes a number of heme peroxidases and other oxidative enzymes that could be involved in the modification of lignin structure. Indeed, we recently identified and characterized an unusual DyP with multiple modes of activity, including peroxidase, Mn peroxidase, and Mn-dependent oxidase [*15*]. While the genome of *A.* sp. 75iv2 encodes a large number of clusters for aromatic degradation, similar to related actinomycetes, it still remains unclear whether the modifications on lignin made by *A.* sp. 75iv2 can lead to its breakdown and utilization as a carbon source. In this study, we show that *A.* sp. 75iv2 is indeed able to utilize lignin as a carbon source for growth and that extracellular peroxidase activity is indeed induced under minimal media conditions. We also characterize the overall response of *A.* sp. 75iv2 using both transcriptional and protein profiling. Further analysis of the secreted protein fraction also reveals a change in the heme-containing protein fraction, which is derived from multiple proteins.

5.2. Materials and Methods

Materials. Yeast extract, malt extract, glycerol, potassium phosphate monobasic, cupric sulfate pentahydrate, potassium chloride, and acetic acid were purchased from EMD Biosciences (Darmstadt, Germany). Dextrose, carbenicillin, ammonium sulfate, manganese chloride tetrahydrate, and calcium chloride anhydrous, hexanes, methanol, hydrogen peroxide, Tris-HCl, bromophenol blue, sodium acetate, sodium chloride, dithiothreitol (DTT), urea, tetrahydrofuran, ethanol, acetone, Tris-HCl, and phenylmethanesulfonyl fluoride (PMSF) were purchased from Fisher Scientific (Pittsburgh, PA). Magnesium sulfate heptahydrate, zinc sulfate monohydrate, iron sulfate heptahydrate, potassium phosphate dibasic, iron chloride, manganese sulfate monohydrate, zinc sulfate heptahydrate, calcium chloride dihydrate, boric acid, vanillin, ethyl acetate, vanillin-(ring-$^{13}C_6$), malonic-2-^{13}C-acid, horseradish peroxidase, xylose, 2,4-

dichlorophenol, 4-aminoantipyrene, 3,3′,5,5′-tetramethylbenzidine (TMBD), trichloroacetic acid (TCA), ammonium bicarbonate, iodoacetamide, lysozyme, β-mercaptoethanol (BME), piperidine, aniline, pyridine, lithium aluminum hydride, methyl sulfoxide (DMSO), N,N,N′,N′-tetramethylethane-1,2-diamine (TEMED), and butylated hydroxytoluene (BHT) were purchased from Sigma-Aldrich (St. Louis, MO). Formic acid was purchased from Acros Organics (Morris Plains, NJ). Hydrochloric acid was purchased from Macron Chemicals (Charlotte, NC). Malonic acid was purchased from Spectrum Chemical Manufacturiung Corp. (Gardena, CA). Methanol-d_4 was purchase from Cambridge Isotope Labs, Inc. (Andover, MA). Polyacrylamide, electrophoresis grade sodium dodecyl sulfate (SDS) and ammonium persulfate were purchased from Bio-Rad Laboratories (Hercules, CA). T4 phosphonucleotide kinase, Klenow large fragment, T4 DNA polymerase, Klenow large fragment, Klenow large fragment (Exo-), and T4 DNA ligase were purchased from New England Biolabs (Ipswich, MA). Deoxynucleotides (dNTPs), Platinum Taq High-Fidelity polymerase (Pt Taq HF), 5× first strand buffer, RNAse H, and PCR2.1-TOPO TA cloning kit were purchased from Invitrogen (Carlsbad, CA). DNase was purchased from Fermentas (Glen Burnie, Maryland). Casamino acids, agar, and peptone were purchased from BD Bioscences (Franklin Lakes, NJ). Tris(2-carboxyethyl)phosphine (TCEP·HCl) was purchased from Thermo Scientific (Waltham, MA) Trypsin Gold and ProteaseMax were purchased from Promega (Madison, WI). Complete EDTA-free protease inhibitor cocktail was purchased from Roche Applied Science (Penzburg, Germany). RNAprotect bacteria reagent and QIAquick PCR Purification Kit were purchased from Qiagen (Valencia, CA). Superase-In and MICROB*Express* Kit were purchased from Ambion Life Technologies (Grand Island, NY). GC-rich hexamers were purchased from Gene Link, Inc. (Hawthorne, NY). Ethanol-extracted and dioxane-extracted lignins were obtained from Stefan Bauer at the Energy Biosciences Institute (Berkeley, CA) [*16*].

Bacterial strains. *Amycolatopsis* sp. 75iv2 (formerly *Streptomyces setonii* and *Streptomyces griseus* 75iv2, ATCC 39116) was purchased from the American Tissue Type Collection (Manassas, VA). *E. coli* DH10B-T1^R was used for plasmid construction and BL21(de3) was used for heterologous protein production.

Synthesis of unlabeled ferulic acid. Ferulic acid was synthesized according to literature protocol [*17*] by condensation of vanillin (2 g, 13.1 mmol) with malonic acid (1.56 g, 15 mmol) by refluxing for 2.5 h at 100 °C in pyridine (2.7 mL) containing aniline (6 drops) and piperidine (6 drops). At this time, water (17.5 mL) was added and the mixture was allowed to cool to room temperature before removing organic solvent by rotary evaporation. The remaining aqueous solution was acidified with HCl (5% (v/v) final concentration). Ethyl acetate (25 mL) was added to the acidified aqueous solution and the organic layer was washed with 5% (v/v) HCl to remove residual pyridine. The organic layer was collected, dried over $MgSO_4$, and evaporated to dryness using rotary evaporation. The residue was dissolved in ethyl acetate (6.5 mL) and the ferulic acid was precipitated as a white solid by addition of hexanes (43.5 mL) to a final concentration of 13% (v/v). After filtering, the solid was dried *in vacuo* (1.45 g, 57.2% yield).

Synthesis of unlabeled coniferyl alcohol. Ferulic acid (1.45 g, 7.5 mmol) was first esterified by refluxing overnight at 80 °C in methanol (9 mL) acidified with HCl (12 M, 140 μL, 180 mM final concentration). The solvent was then removed by rotary evaporation before reducing the methyl ester was then reduced to coniferyl alcohol. The methyl ester was dissolved in dry THF (15 mL) under N_2 atmosphere and cooled in an ice bath with stirring while a solution of lithium aluminum hydride (95%, 11.25 mmol) in dry THF (10 mL) was added dropwise over

30 min. After the addition was complete, the reaction was allowed to warm up to room temperature and stir overnight. At this time, the reaction was determined to be nearly complete using thin layer chromatography (silica, 1:1 ethyl acetate:hexanes). The purified product (760 mg, 56% yield) was obtained by flash chromatography (silica, 1:1 ethyl acetate:hexanes). ^{1}H NMR (400 MHz, methanol-d_4) δ 6.98 (s, 1H), 6.83 (d, J = 8.4 Hz, 1H), 6.74 (d, J = 10.4 Hz, 1H), 6.49 (d, J = 15.6 Hz, 1H), 6.209 – 6.140 (dt, J = 16 Hz, 6 Hz, 1H), 4.18 (d, J = 6.0 Hz, 2H), 3.84 (s, 3H).

Synthesis of (*E*)-3-(4-hydroxy-3-methoxy[$^{13}C_{ring}$]phenyl)[2-^{13}C]acrylic acid (^{13}C-labeled ferulic acid). Synthesis of ^{13}C-labeled ferulic acid was carried out as described on smaller scale. Vanillin-(ring-$^{13}C_6$) (0.5 g, 3.16 mmol) was condensed with malonic-2-^{13}C-acid (0.39 g, 3.7 mmol) in pyridine (0.6 mL) in the presence of piperidine (2 drops) and aniline (2 drops), yielding ferulic-ring-$^{13}C_6$-β-^{13}C-acid (0.44 g, 69% yield). ^{1}H NMR (400 MHz, methanol-d_4) δ 7.64 – 7.53 (m, 1H), 7.35 – 6.96 (dm, J = 157.8 Hz, 1H), 7.24 – 6.84 (dq, J = 158.5, 8.1 Hz, 1H), 6.98 – 6.59 (dm, J = 159.22 Hz, 1H), 6.29 (ddd, J = 160.7, 15.9, 5.2 Hz, 1H), 3.87 (d, J = 3.6 Hz, 3H). ^{13}C NMR (151 MHz, methanol-d_4) δ 169.51 (dd, J = 73.5, 7.1 Hz), 150.39 – 146.77 (m), 126.31 (td, J = 59.0, 6.8 Hz), 122.49 (tt, J = 58.6, 5.4 Hz), 116.22 – 113.71 (m), 111.40 – 108.98 (m), 54.99 (dt, J = 4.7, 2.2 Hz). HRESI-MS: calculated for [M-H$^+$] 200.0741 Da, found 200.0739 Da.

Synthesis of 4-((*E*)-3-hydroxy[2-^{13}C]prop-1-enyl)-2-methoxy[$^{13}C_{ring}$]phenol (^{13}C-labeled coniferyl alcohol). ^{13}C-ferulic acid (0.657 g, 4.16 mmol) was esterified and reduced as described above to produce ^{13}C-labeled coniferyl alcohol (175 mg, 43% yield). ^{1}H NMR (600 MHz, methanol-d_4) δ 7.15 – 6.76 (dm, J = 155.0 Hz, 1H), 7.01 – 6.62 (dm, J = 157.0 Hz, 1H), 6.94 – 6.55 (dm, J = 157.0, 1H), 6.52 – 6.43 (m, 1H), 6.39 – 5.93 (ddq, J = 151.0, 16.6, 5.8 Hz, 1H), 4.22 (m, 2H), 3.82 (d, J = 3.8 Hz, 3H). ^{13}C NMR (151 MHz, methanol-d_4) δ 148.19 – 147.05 (m), 145.96 (ddt, J = 74.0, 67.6, 6.8 Hz), 130.50 (m), 129.14 (td, J = 59.3, 6.8 Hz), 125.59 (t, J = 4.8 Hz), 119.51 (tt, J = 58.9, 5.4 Hz), 115.27 – 114.25 (m), 109.07 (ddq, J = 70.1, 59.8, 4.9 Hz), 62.49 (dd, J = 47.7, 6.3 Hz), 54.93 (m). HRESI-MS: calculated for [M-H$^+$] 186.0949 Da, found 186.0951 Da.

Preparation of synthetic lignin. Synthetic lignin (dehydropolymer, DHP) was synthesized from coniferyl alcohol according to literature procedure [*18, 19*]. Coniferyl alcohol (380 mg, 2.11 mmol) was first dissolved in ethanol (11 mL) and diluted to a total volume of 250 mL with MilliQ water. The coniferyl alchol solution was transferred into an additional funnel. A solution of hydrogen peroxide (125 mL, 0.038%) was then added to a second addition funnel. These two solutions were simultaneously introduced drop-wise over 4 h to a solution of horseradish peroxidase (2 mg in 30 mL MilliQ water), while slowly stirring in a 1 L 2-neck round-bottomed flask. After the addition was complete, this mixture was allowed to polymerize further for an additional 48 hours. The precipitated polymer was collected by for 10 min at 9,800 × g (4 °C) and the polymer was lyophilized (unlabeled, 266 mg; ^{13}C-labeled, 250 mg).

Re-precipitation of DHP. DHP (250 mg) was thoroughly resuspended in MilliQ water (3 mL). Ethanol (100%, 9 mL) was then added and the mixture was vortexed until completely solvated. Next, MilliQ water (24 mL) was added to the solution, leading to immediate precipitation of the polymer after vortexing. The polymer was collected by centrifugation for 10 min at 9,800 × g (4 °C) and lyophilized to yield monomer-free DHP (unlabeled, 207 mg, 83 % yield; ^{13}C-labeled, 223 mg, 89 % yield).

Analysis of DHP by gel permeation chromatography. DHP (~1.5 mg) was dissolved in THF (1.5 mL) stabilized with BHT (250 ppm) and centrifuged for 1 min at 20,817 × *g* to remove any particulate matter. The sample (100 μL) was chromatographed (1 mg/mL, 30 °C) on two sequential Varian Mesopore columns (300 × 7.5 mm ID, 3 μm porosity; Agilent Technologies; Santa Clara, CA) using a GPC 50 Plus system (Agilent Technologies; Santa Clara, CA) equipped with a UV detector. Lignin was monitored at 280 nm. Polystyrene molecular weight standards (EasiVial PS-L; Agilent Technologies) were used to calibrate the system allowing an approximate linear range of 0.16 kDa to 38.5 kDa for calculation of the DHP molecular weight peak (Mp) and size distribution.

***A.* sp. 75iv2 cell culture.** *A.* sp. 75iv2 was streaked to single colonies onto a YEME [*20*] agar (1.5% (w/v)) plate from a frozen 20% glycerol stock stored at -80°C. After incubation at 37 °C for 3 d, yellow-white mycelial colonies were produced. A single colony was then inoculated into YEME (5 mL) and incubated at 37 °C with shaking (200 rpm) for 3 d. The liquid culture (50 μL) was spread using sterile glass beads (5 mm) onto YEME plates and incubated at 37 °C for 3 d to produce white mycelial mats. For a 1 L culture, mycelia from four plates were scraped with the wide edge of a sterile spatula into the appropriate media (10 mL) and resuspended to relative homogeneity by pipetting. This solution was then used to inoculate the appropriate media (50 mL) supplemented with carbenicillin (50 μg/mL) and various carbon sources (0.05% (w/v); dioxane-extracted lignin, ethanol-extracted lignin, DHP, glucose, glycerol, or xylose) in 250-mL baffled flasks with 0.2 um sterile filter tops. Cultures were then incubated at 37 °C with shaking (200 rpm). Additional aliquots of carbenicillin (50 μg/mL) were added every 3 d to prevent contamination. Samples were removed at various time points to monitor culture growth by OD_{600} (0.5 mL), RNA-seq (2 × 1.5 mL), secretome proteomics studies (5 mL), extracellular peroxidase activity (1 mL), and secretome heme-stain SDS-PAGE (1 mL). Cell viability and culture contamination was also monitored regularly by spreading a sample (50 μL) onto YEME plates) using sterile glass beads. Serial dilution of the culture to 10^{-5} with sterile water yielded single *A.* sp. 75iv2 colonies with no evidence of contamination. Two colonies from each culture condition were resuspended in DMSO and heated at 98°C for 10 min to yield sample (5 μL) as template for amplification of the 16s rRNA with Pt-Taq HiFidelty using EubF1 and EubR1 primers as previously described [*14*]

Amino acid supplemented minimal media (SMM). Media was prepared according to a modified literature protocol by lowering the ammonium sulfate content to 10% of the original recipe [*20*]. Media contained (per 1 L): ammonium sulfate (0.2 g, 1.5 mM), casamino acids (5 g), magnesium sulfate heptahydrate (0.6 g, 2.4 mM), zinc sulfate monohydrate (0.36 mg, 2.0 μM), iron sulfate heptahydrate (1 mg, 3.6 μM), manganese chloride tetrahydrate (1 mg, 5.0 μM), and anhydrous calcium chloride (1 mg, 9.0 μM). Final volume prior to sterilization via autoclave was 985 mL. After sterilization, filter-sterilized 0.1 M phosphate buffer, pH 6.8 (15 mL) made as a 1:1 solution of 0.05 M sodium phosphate monobasic and 0.05 M potassium phosphate dibasic was added.

Minimal media (MM). Media was prepared according to a modified literature protocol with lowered nitrogen content [*7*]. Media contained (per 1 L): potassium phosphate monobasic (0.4 g, 2.9 mM), ammonium sulfate (0.2 g, 1.5 mM), magnesium sulfate heptahydrate (0.2 g, 0.8 mM), iron chloride (15 mg, 92.5 μM), manganese sulfate monohydrate (0.5 mg, 3 μM), zinc sulfate heptahydrate (1 mg, 3.5 μM), calcium chloride dihydrate (0.5 mg, 3.4 μM),

cupric sulfate pentahydrate (1 mg, 4.0 μM), potassium chloride (0.1 mg, 1.3 μM), boric acid (0.5 mg, 8.0 μM). Prior to sterilization by autoclave, the pH was adjusted to 7.0 using hydrochloric acid.

Peroxidase activity assay. Peroxidase activity was monitored spectrophotometrically using a modified literature procedure [*21*]. For measuring peroxidase activity in cell culture, a sample (1 mL) was removed and cleared of cells and lignin by centrifugation for 5 min at 9,800 × *g* (4°C). The culture supernatant was used directly in the peroxidase assay using a Beckman DU-800 spectrophotometer (Beckman Coulter Inc.; Pasadena, CA). Assays contained culture supernatant (500 μL), 2,4-dichlorophenol (3 mM), 4-aminoantipyrene (164 μM), and H_2O_2 (4 mM) in 50 mM potassium phosphate buffer, pH 7.0 in a total assay volume of 1 mL. Assays were initiated by the addition of H_2O_2 and monitored for an increase in absorbance at 510 nm at 25 °C. Peroxidase activity was calculated using an extinction coefficient of 18,500 $M^{-1}cm^{-1}$ for the 4-aminoantipyrene/2,4-dichlorophenol adduct [*22*]. For monitoring peroxidase activity in fractions from protein purification, the same reaction conditions were used in a total volume of 200 μL using SpectraMax M2 96-well plate reader (Molecular Devices; Sunnyvale, CA).

Heme analysis by Tricine-SDS-PAGE. The detection of extracellular heme-containing proteins from culture growths was performed using a modified literature protocol for a gel-based heme stain [*14, 23*]. For secretome analysis, samples (1 mL) were removed from cell culture and passed through a 0.2 μm filter. Each sample was then concentrated 40-fold 25 μL before adding Laemmli sample buffer without *β*-mercaptoethanol (8.3 μL; 0.25 M Tris-HCl, 2% SDS, 40% glycerol, 0.04% bromophenol blue) and incubating at room temperature for 15 min. The sample was analyzed by tricine-SDS-PAGE at 4 °C on a 10% gel at 200V for approximately 60 min. The gel was then soaked at room temperature in the dark on an orbital rotator with 3,3′,5,5′-tetramethylbenzidine (2 mM) in 0.25 M sodium acetate, pH 5.0 containing 30% (v/v) methanol for 1 to 2 h. H_2O_2 (60 mM) was added to initiate the development of blue bands within 15 min and stopped by rinsing with ddH_2O.

For protein fractions from secretome fractionation, samples (1.5 mL) were concentrated 50-fold to 30 μL using a 0.5 mL 3 kD MWCO Amicon centrifugal filtration device (Millipore Corporation; Billerica, MA) before adding Laemmli sample buffer without β-mercaptoethanol (10 μL) and incubating at room temperature for 15 min. The samples were separated by 10% tricine-SDS-PAGE at 4 °C and 100V for approximately 80 min. The gel was then soaked at room temperature in the dark on an orbital rotator with 3,3′,5,5′-tetramethylbenzidine (2 mM) in 0.25 M sodium acetate, pH 5.0 containing 30% (v/v) methanol for 1 h, with manual shaking every 15 to 30 min. H_2O_2 (60 mM) was added to initiate the development of blue bands within 15 min and stopped by rinsing with ddH_2O. For total protein content, the SilverQuest staining kit (Invitrogen; Carlsbad, CA) was used on heme-stained gel according to manufacturer instructions.

Isolation of mRNA. Culture samples (2 × 1.5 mL) were collected after 24 h of growth and centrifuged for 5 min at room temperature at 20,817 × *g*. The cell pellet was collected and resuspended in RNAprotect bacteria reagent (1 mL) according to manual instructions. The sample was incubated at room temperature for 5 min before pelleting cells by centrifugation for 5 min at room temperature at 20,817 × *g*, discarding the supernatant, and flash-freezing the pellet in liquid nitrogen for storage at -80 °C. mRNA was isolated according to a modified literature procedure described below [*24*].

Cell pellets were briefly thawed on ice and resuspended without vortexing with a solution of lysozyme (2 mg/mL) in in DEPC-treated TE buffer (100 μL). This mixture was allowed to incubate at room temperature for 3 min after which buffer RLT supplemented with 1% β-mercaptoethanol (350 μL, Qiagen RNeasy Mini Kit) was immediately added and the sample vortexed. The sample was passaged through a Qiashredder homogenizer spin column by centrifugation for 2 min at 20,817 × *g* (4 °C). The supernatant (~450 μL) of the filtrate was carefully removed by pipette so as to not disturbed the pellet and then mixed with ethanol (250 μL). RNA was isolated and purified using the Qiagen RNeasy Mini Kit according to manufacturer protocol and eluted into RNase-free water (40 μL). The concentration of the isolated nucleic acid was approximated based on A_{260nm} using a Nanodrop 1000 Spectrophotomer (Thermo-Fisher; Waltham, MA). Genomic DNA was then removed by enzymatic hydrolysis with RNase-free DNase I (Fermentas) according to manufacturer protocol (per 1 μg of nucleic acid; final volume, 20 μL): 10 × reaction buffer with $MgCl_2$ (2 μL), Superase-In (1 μL), water, and RNase-free DNase I (1 uL, 1 U). The sample was at 37°C for 30 min before purification with the Qiagen RNeasy Mini Kit following manufacturer protocol.

Total RNA (40 μL) was quantified by A_{260} before rRNA removal using the MICROB*Express* Kit (Ambion). RNA was precipitated with ethanol by mixing with 0.1 vol sodium acetate (3 M) and 0.02 vol glycogen (5 mg/mL), flash frozen in liquid N_2, and pelleted by centrifugation for 30 min at 20,817 × *g* (4 °C). Pellets were washed with ice-cold 70% ethanol, centrifuged, and allowed to air-dry just until liquid was no longer visible before resuspending in RNase-free water (10 μL). Two round of rRNA pull-down were then performed according to a modified manufacturer protocol. Total RNA (10 μg in ≤ 10 μL) was mixed with Binding Buffer (200 μL) and vortexed gently to mix. The Capture Oligo Mix (4 μL) was added and vortexted gently again to mix. After heating at 70 °C for 10 min, annealing was carried out by incubating for 15 min at 37 °C. The Oligo MagBeads (50 μL) were then equilibrated and incubated with the RNA sample for 15 min at 37 °C. A magnet was used to pull down the Oligo MagBeads while transferring the supernatant into a new tube using a pipette. The Oligo MagBeads were washed by resuspending in Wash Solution (100 μL) pre-incubated at 37 °C, pulling down with a magnet, and removal of the supernatant. The two supernatant samples were pooled and passaged through a second round of rRNA depletion after precipitating the RNA to concentrate the sample as previously described. After the second round including a final RNA precitation step, the sample was resuspended RNase-free water (10.5 μL). To monitor rRNA removal as well as sample quality and concentration, RNA (1 μL) was denatured at 70 °C for 2 min and analyzed on a RNA HighSens Analysis chip using an Experion 700-7000 (Bio-rad).

Preparation of cDNA libraries for RNA sequencing (RNA-seq). To hydrolyze the mRNA for library construction, the mRNA (9 μL) was mixed with Ambion Fragmentation Reagent (1 μL) and allowed to incubate for 3 min and 15 s at 70 °C. Samples were immediately placed on ice before adding Stop Solution (1 μL) and mixing by pipetting. Fragmented mRNA was precipitated and recovered by ethanol precipitation as described above. Synthesis of cDNA was carried out using SuperScript III Reverse Transcriptase (200 U, Invitrogen) for first strand synthesis. The reaction (20 μL) contained the fragmented mRNA (10.5 μL), GC-rich random hexamer primers (200 ng), RNase-free dNTPs (0.5 mM), and DMSO (5%). The mixture incubated for 5 min at 65°C and then immediately placed on ice for 2 min. Next, the first strand buffer (4 μL, 5×) and DTT (2 μL, 100 mM) were added, and the mixture was incubated for 2

min at 25°C. After addition of SuperScript III (200 U), the reaction was incubated for 10 min at 25°C, 50 min at 55°C, and finally 15 min at 70°C to heat-inactivate. Second-strand was carried out by directly adding to this mixture: *E. coli* DNA polymerase I (42.5 U, Invitrogen), RNAse H (2.5 U, Invitrogen), and dNTPs (0.3 mM) in Buffer 2 (10 μL, NEB) for a final reaction volume of 100 μL. This reaction was incubated for 2.5 h at 16°C and then purified using the QIAquick PCR Purification Kit following manufacturer protocol.

End repair was performed on the cDNA library using T4 phosphonucleotide kinase (50 U, NEB), Klenow large fragment (5 U, NEB), T4 DNA polymerase (6 U, NEB), and dNTPs (0.4 mM) in T4 DNA ligation buffer (10 μL, NEB) for a total volume of 100 μL. After incubating for 30 min at 20 °C, the cDNA was purified using the QIAquick PCR Purification Kit and eluted with water (41 μL). An overhanging A was then added by incubating the cDNA for 30 min at 37°C in the presence of Klenow large fragment (Exo-) (15 U, NEB) and dATP (0.2 mM) in Buffer 2 (5 μL, NEB) and purifying to a volume of 25 μL with the QIAquick Minelute PCR Purification Kit. Adapters were ligated immediately afterwards by incubating the cDNA library at room temperature for 30 min with adapter mix (0.3 μM), T4 DNA ligase (2000 U, NEB) and T4 DNA ligase buffer for a final volume of 30 μL. The adapter-ligated library was isolated using the QIAquick MinElute PCR Purification Kit and eluted using water (15 μL). The library was enriched by 18 rounds of PCR amplification using Platinum Taq High-Fidelity Polymerase and the PE F1 and PE R1 primers in the presence of DMSO (10%) due to the high GC-content of the *A.* sp. 75iv2 genome. Libraries were analyzed on an Agilent Bioanalyzer (Santa Clara, CA) using an Agilent DNA 1000 Kit and submitted for sequencing on a HiSeq2500 Illumina sequencing platform (San Diego, CA) to the UC Davis Genome Center (Davis, CA). Two biological replicates were sequenced from the control and dioxane-extracted lignin growths while a single library was sequenced from the DHP growth.

Analysis of RNA-seq data. Read libraries (50 bp) were pre-processed to remove adapter contamination using Scythe and low quality bases using Sickle (UC Davis Bioinformatics Core, https://github.com/ucdavis-bioinformatics). These processed reads were then aligned to the high quality draft genome of *A.* sp. 39116 [*25*] using BWA [*26*], allowing for no more than 3 mismatched bases. The number of reads per open reading frame were then tabulated using HTseq-count [*27*] based on the annotation of the genome. Differential expression analysis was performed using DESeq2 [*27*] on all predicted protein coding open reading frames. Open reading frames were considered to be differentially expressed if they exhibited a greater than 2-fold change and exhibited a p-value < 0.05. Functional enrichment analysis was performed using BiNGO [*28*] based on COG categories identified in the annotation of the genome. The enrichment of open reading frames encoding signal peptide-containing proteins and transmembrane proteins was determined using Fisher's Exact Test and considered significant with a p-value < 0.05.

Preparation of intracellular protein samples. For intracellular cytosolic proteomics, cell pellets at 24 and 48 h (~50 mg wet cell pellet) were resuspended in buffer (0.8 mL; 50 mM Tris, 50 mM NaCl, pH 7.5) supplemented with 1 mM DTT, Complete Protease Inhibitor Cocktail, and triton-100 (0.1%). Cells were then lysed using sonication on ice (5 × 10 sec, 1 min resting periods). Insoluble lysate was separated using centrifugation (20 min at 14,000 rpm). Soluble lysate fraction was precipitated using trichloroacetic acid (TCA) by adding a stock solution of TCA (100% w/v) to a final concentration 20%. This was allowed to sit on ice for 20 min before

centrifugation (4 °C, 30 min at 20,817 × *g*) to pellet the intracellular soluble proteinaceous fraction. The pellet was then washed of residual TCA as well as precipitated lignin by resuspending the pellet in ice-cold acetone mixture (90% acetone, 0.01 M HCl), allowing to sit on ice for 5 min before pelleting the protein fraction (4 °C, 30 min at 20,817 × *g*), and allowing it to dry at room temperature only until no visible acetone remains. It was then immediately resuspended as below.

Preparation of secretome samples. Samples for extracellular protein profiling (4.5 mL) were collected and concentrated 18-fold to 250 μL using a 3 kDa MWCO Amicon centrifugal filtration device. Proteins were then precipitated by addition of trichloroacetic acid (final concentration; 15 % (v/v)). This mixture was incubated on ice for 20 min and the protein fraction was collected as a pellet by centrifugation for 30 min at 20,817 × *g* (4 °C). Pellets were washed with acetone (90%) and 0.01 M HCl to remove residual salt and lignin and centrifuged again before air-drying for 5 min.

Preparation of tryptic digests. Each protein pellet (~100-200 μg) was resuspended in PBS buffer (30 μL; 137 mM sodium chloride, 2.7 mM potassium chloride, 10 mM sodium phosphate dibasic, 2 mM potassium phosphate monobasic, pH 7.35) supplemented with urea (8 M) and vortexed to dissolve proteins. ProteaseMax (30 μL, 0.2% (w/v) in 100 mM ammonium bicarbonate) was added followed by vortexing vigorously to mix. Ammonium bicarbonate (40 μL, 100 mM) was added next to yield a final volume of 100 μL. TCEP·HCl (10 μL, 110 mM) was added with vortexing to reduce disulfide bonds and the sample was incubated for 30 min at 55°C. Thiols were then protected by incubating the mixture with iodoacetamide (2.5 μL, 500 mM) at room temperature in the dark for 30 min. PBS (120 μL) was added followed by introduction of additional ProteaseMax (1.2 μL, 1% (w/v) in 100 mM ammonium bicarbonate). The sample was vortexed and briefly spun down. After addition of mass spectrometry-grade Trypsin Gold (4 μL, 2.5 mg/mL in 50 mM acetic acid), the sample was incubated overnight at 37 °C. Samples were acidified with formic acid (final concentration; 5% (v/v)) and vortexed to precipitate residual lignin. The insoluble particular matter was then removed after centrifuging for 30 min at 20,817 × *g* (room temperature). Samples were frozen and stored at -20 °C until further analysis.

Analysis by Multidimensional Protein Identification Technology (MudPIT). The trypsin-digested peptide mixtures were concentrated by loading onto a C_{18} reversed-phase capillary column (deactivated fused silica column with 0.25 mm diameter packed with 4 cm of C_{18} resin (125Å bulk packing; Phenomenex; Torrance, CA) attached to an inline-microfilter assembly) and then separated using a biphasic (strong cation exchange (3 cm of Partisphere SCX resin; 5 micron, 250 x 4.6 mm; VWR; Radnor, PA)/reverse phase resin (10 cm C_{18}) packed into a 13 cm nanocapillary Deactivated Fused Silica columns with 0.1 mm diameter SCX 5 micron, 250 x 4.6 mm capillary column. Using literature procedure, analysis was performed by two-dimensional liquid chromatography (2D-LC) separation in combination with tandem mass spectrometry as previously described [*29*]. Peptides were eluted in a 5-step MudPIT experiment (using 0%, 10%, 25%, 80%, and 100% 3 min salt bumps with 500 mM ammonium acetate) [*30*], and data were collected in an ion trap mass spectrometer, LTQ (Thermo Scientific) set in a data-dependent acquisition mode using dynamic exclusion (60 s). For this, a single MS survey (ms1) scan was followed by 7 ms2 scans. The ms2 spectra data were then extracted from the raw file using RAW Xtractor (version 1.9.9.2) and searched using the SEQUEST algorithm (Version 3.0) [*31*] against

a custom-made database containing the proteome of *Amycolatopsis* sp. 75iv2. Additionally, each of these entries was reversed and appended to the database for assessment of false-discovery rates. SEQUEST searches allowed for oxidation of methionine residues (16 Da), static modification of cysteine residues (57 Da-due to alkylation), no enzyme specificity and a mass tolerance set to ± 1.5 Da for precursor mass and ± 0.5 Da for product ion masses. The resulting ms2 spectra matches were assembled and filtered using DTASelect (version2.0.27) [*32*]. A quadratic discriminant analysis was used to achieve a maximum peptide false positive rateof 1% as previously described [*32, 33*].

Spectral counts for each protein were normalized based on average spectral count of the sample according to literature [*34*]. Upregulated and downregulated proteins were then filtered using a cut-off for the fold-change (>2-fold). The Student's t-test was used to calculated the *p*-value calculated, of which proteins with $p < 0.05$ were considered to be statistically significantly differentially regulated. Average values were used to calculate ratios for –fold regulation.

NMR studies of metabolized lignin. Lignin modification growths were performed using the same methods as previously described (see above) using 50 mL cultures and 0.05% lignin in both SMM and MM media conditions. Cultures were grown for 28 d, removing 1 mL every three d to measure culture viability using OD_{600} and culture contamination by spreading a sample (50 μL) onto YEME plates using sterile glass beads. After growths were complete, 20 mL of culture and controls were centrifuged (10 min at 9,800 × *g*, 4 °C) to separate solids (lignin and cells) and supernatant and then processed independently. The solids and supernatant were first lyophilized separately in preparation for NMR. For analysis of solids, the pellet was directly extracted with DMSO (600 μL × 3) until the organic layer was colorless. These samples were then pooled, lyophilized, and resuspended directly for NMR analysis in DMSO-d_6 (600 μL). Extraction controls with known quantities of lignin ensured that >90% of lignin was extracted, even in the presence of added cells. The same protocol was used for extraction of the supernatant. Samples were stored at -20 °C.

The NMR spectra were acquired on a Bruker Avance 800 MHz spectrometer. HSQC experiments used a standard Bruker's pulse program (hsqcetgpsisp2.2) and had the following parameters; acquired from 10-0 ppm in F2 (^{1}H) by using 2,048 data points for an acquisition time (AQ) of 128 ms, 220-0 ppm in F1 (^{13}C) by using 512 increments (F1 acquisition time 5.8 ms) of 92 scans with a 1 s interscan delay (D1). A squared cosine-bell apodization function was applied in both dimensions. HSQC cross-peaks were assigned by combining the results and comparing them with the literature. The solvent peak was used as an internal reference (δ_C 39.5, δ_H 2.49 ppm). Volume integration of peaks in HSQC plots was accomplished using Bruker's TopSpin 3.0 software to obtain an estimate in percentages of the various lignin substructures.

Analysis of intracellular protein hydrolysate. Using a modified literature procedure [*35*], to the lyophilized solid from 28 d SMM growths, including the cells (grown in the absence of lignin) as well as the cells + ^{13}C-labeled DHP, water was added (MQ,1 mL) and a pipette was used to make the culture homogeneous. To this, hydrochloric acid (1 mL, 12 M) was added, and in a bomb, each mixture was allowed to stir at 110 °C for 24 h. This was allowed to cool to room temperature before opening, and the samples were dried overnight using a stream of air. Each sample was then resuspended in THF (1 mL), after which *N*-methyl-*N*-tert-butyldimethylsilyltrifluoroacetamide (MTBSTFA, 1 mL) was added, and the samples were allowed to incubate at 60 °C for 2hr. Solid was removed from the samples by centrifugation (5

min at 20,817 × *g*) and samples were analyzed by GC-MS exactly according to literature with one modification: 5 µL were inject with a split ratio of 1:10 [*35*].

Native purification of extracellular peroxidases. Two growth conditions with either DHP (0.05% (w/v) or ethanol-extracted lignin (0.05% (w/v); EL), were used to generate secretomes for purification of extracellular peroxidases. For the DHP growth, a control (no DHP) growth was also performed along-side as a control for contaminating proteins. Both purifications followed the same procedure until the last purification step. *A.* sp 75iv2 cultures (500 mL DHP culture, 10 × 50 mL cultures; 750 mL EL culture, 15 × 50 mL cultures) were grown for 48 h in the presence of lignin in SMM as previously described. After removing cells by centrifugation for 10 min at 9,800 × *g* (4 °C), the supernatant was collected and passed through a 0.2 µm filter. PMSF (0.5 mM in ethanol) was added to the filtrate and the solution was concentrated to 100 mL using a 3 kDa MWCO Amicon filtration membrane (Millipore Corporation). The total peroxidase activity of the supernatant (50 µL) was measured assaying the oxidation of 2,4-DCP (total volume per well, 200 µL) as described above in a SpectraMax M2 96-well plate reader (Molecular Devices; Sunnyvale, CA). The concentrated supernatant was loaded onto a 5 mL HiTrapQ HP anion exchange column (GE Healthcare Life Sciences) at a flow rate of 3 mL/min on an ÄKTApurifier FPLC equilibrated with Buffer A (20 mM Tris, 10 mM NaCl, pH 7.5). The protein mixture was washed with 2 col vol of Buffer A before eluting in a linear gradient over 20 col vol to 100 % Buffer B (20 mM Tris, 1 M NaCl, pH 7.5). Fractions were then analyzed for peroxidase activity (50 µL of eluate) using a 96-well plate reader. Peroxidase-active fractions were pooled and concentrated to 2 mL with a 3 kD MWCO AmiconUltra centrifugal filtration device. After centrifugation for 1 min at 20,817 × *g* (4 °C) to remove particulate matter, the sample was then chromatgraphed (1 mL/min) on a HiLoad 16/20 Superdex 200 prep grade column (GE Healthcare Life Sciences) using Buffer C (1.5 col vol; 20 mM Tris, 150 mM potassium chloride, pH 7.5).

DHP growth. The fractions with peroxidase activity were pooled and concentrated to 1 mL with a 3 kD MWCO AmiconUltra centrifugal filtration device. Half of this sample (500 µL) was concentrated further to 60 µL for tryptic digest. The other half (500 µL) was centrifuged for 1 min at 20,817 × *g* (4 °C) to remove particulate matter and chromatgraphed (1 mL/min) using a HiLoad 16/20 Superdex 75 prep grade column (GE Healthcare Life Sciences) in Buffer C (1.5 col vol; 20 mM Tris, 150 mM potassium chloride, pH 7.5). Fractions with peak activity were pooled, concentrated to 60 µL for tryptic digest.

EL growth. Fractions with peroxidase activity exchanged into Buffer A by three rounds of concentration and dilution of the sample with Buffer A to 2 mL. This sample was then loaded onto a MonoQ 10/100 GL anion exchange column (10 mL; GE Healthcare Life Sciences) equilibrated with Buffer A. The column was washed with 2 col vol of Guffer A before separation using a linear gradient (3 mL/min) over 40 col vol to 100% Buffer B. Fractions (5 mL) were collected and those high peroxidase activity were analyzed separately. For fractions with high activity, part (3 mL) was concentrated to 60 µL using a 3 kD MWCO AmiconUltra centrifugal filtration device for tryptic digest. The other part (1.5 mL) was concentrated to 30 µL using a 0.5 mL 3 kD MWCO Amicon centrifugal filtration device and analyzed by heme stain as previously described.

Preparation of trypsin-digested samples for LC-MS/MS. To 60 µL of protein sample, ammonium bicarbonate (40 µL, 100 mM) was added followed by reduction of disulfide bonds at

60 °C for 30 min. with TCEP (10 μL, 110 mM). The thiols were then protected by iodoacetamide (2.5 μL, 500 mM) by incubation in the dark (rt, 30 minutes) after which mass spectrometry-grade Trypsin Gold (2 μL, 2.5 mg/mL in 50 mM acetic acid) was added and the samples were allowed to incubate overnight at 37 °C. Samples were then acidified with formic acid (final concentration 5%) and centrifuged to remove insoluble particulate matter (1 min, 20,817 × *g*, rt) and sequenced by LC-MS/MS after de-salting using Varian C_{18} spec tips (A57203; Agilent Technologies; Santa Clara, CA). Liquid chromatography tandem mass spectrometry (LC-MS/MS was performed by the Vincent J. Coates Proteomics/Mass Spectrometry Laboratory at UC Berkeley.

Analysis of trypsin-digested samples by LC-MS/MS. A nano LC column was packed in a 100 μm inner diameter glass capillary with an emitter tip. The column consisted of 10 cm of Polaris c18 5 μm packing material (Varian). The column was loaded by use of a pressure bomb and washed extensively with buffer A (5% acetonitrile/ 0.02% heptaflurobutyric acid (HBFA)). The column was then directly coupled to an electrospray ionization source mounted on a Thermo-Fisher LTQ XL linear ion trap mass spectrometer. An Agilent 1200 HPLC equipped with a split line so as to deliver a flow rate of 300 nl/min was used for chromatography. Peptides were eluted with a 90 minute gradient from 100% buffer A to 60% buffer B (80% acetonitrile/ 0.02% HBFA). The programs SEQUEST and DTASELECT were used to identify peptides and proteins from the *A.* sp. 75iv2 database with common contaminants appended [*31, 33*].

Construction of plasmids. Expression plasmids for genes of interest (*Table 5.1*) were constructed using Gibson cloning [*36*]. The genes were amplified from *A.* sp. 75iv2 genomic DNA with Phusion DNA polymerase using the appropriate primers and inserted into the NdeI-XbaI sites of pCWOri-HisN to add an N-terminal His_{10} tag (*Appendix*). Due to its small predicted protein size, Gibson cloning was used to insert SIL4067 with an N-terminal MBP (amplified using the same procedure from a plasmid containing MBP) into pCWOri-HisN using the NdeI-XbaI sites. The resulting constructs were verified by sequencing (Quintara Biosciences; Berkeley, CA).

Gene ID	Gene annotation	Primers
Amy39116DRAFT_1438	Conserved repeat domain, CRD	CRD1438-F1 , CRD1438-R1
Amy39116DRAFT_4067	Subtilisin inhibitor-like, SIL	SIL4067-F1 , SIL4067-R1 , SIL4067-MBP-F1 , SIL4067-MBP-R1
Amy39116DRAFT_8034	F420 oxidoreductase, F420	F420_8034-F1 , F420_8034-R1
Amy39116DRAFT_6733	Hypothetical_6733	hypo6733-F1 , hypo6733-R1
Amy39116DRAFT_5362	Fe-hydroxamate transporter, FeH1	FeH1-F1 , FeH1-R1
Amy39116DRAFT_0442	Fe-hydroxamate transporter, FeH2	FeH2-F1 , FeH2-R1
Amy39116DRAFT_5265	WXG100 family type VII secretion target, WXG1	WXG1-F1 , WXG1-R1
Amy39116DRAFT_5266	WXG100 family type VII secretion target, WXG2	WXG2-F1 , WXG2-R1
Amy39116DRAFT_3005	Zn-dependent hydrolase, Zn1	Zn1-F1 , Zn1-R1
Amy39116DRAFT_5583	Glycosidase, glyco1	Glyco1-F1 , Glyco1-R1
Amy39116DRAFT_7326	Hypothetical, Amycohypo2	AmycoHypo2-F1 , AmycoHypo2-R1
Amy39116DRAFT_6540	Hypothetical, Amycohypo3	AmycoHypo3-F1 , AmycoHypo3-R1
Amy39116DRAFT_1157	Hypothetical, Amycohypo4	AmycoHypo4-F1 , AmycoHypo4-R1
Amy39116DRAFT_0703	Metal-sulfur cluster biosynthetic enzyme, MS1	AmycoMS1-F1 , AmycoMS1-R1
Amy39116DRAFT_4561	Metal-dependent protease, Prot1	AmycoProt1-F1 , AmycoProt1-R1

Table 5.1. Plasmids constructed for screening gene candidates.

5.3. Results and Discussion

Growth of *Amycolatopsis* sp. 75iv2 using lignin as a carbon source. In our previous studies, we did not observe differences with respect to the extracellular peroxidase activity when *A.* sp. 75iv2 was grown on rich media. We thus turned our attention to utilizing nutrient-limited conditions including lowered nitrogen and carbon content in an effort to induce lignin-dependent physiological changes. Towards this goal, two different minimal medias were developed either containing no additional carbon source (minimal media, MM) [*7*] or casamino acids (supplemented minimal media, SMM) [*20*]. These media were then supplemented with three different types of lignin – ethanol-extracted lignin, dioxane-extracted lignin, and synthetic lignin (dehydropolymer, DHP). The extracted lignin sources were obtained from ball-milled *Miscanthus giganteus* lignocellulose and contain <2% contamination with residual sugar material [*16*]. Thus, DHP was synthesized by the polymerization of coniferyl alcohol using horseradish peroxidase [*18, 19*] as a third lignin type in order to control for carbohydrate contamination. Further characterization of DHP by 2D-NMR revealed a simpler spectrum compared to extracted lignin but also indicated the presence of common bonding motifs found in native lignins (*Figure 5.1*) [*16, 37*]. However, the spectrum also revealed the presence of contaminating monomer. Therefore, a precipitation method was developed in order to remove residual monomer contamination by dissolution in 75% ethanol followed by addition of water to 25% ethanol. The precipitated polymer was then collected by centrifugation and lyophilized. Analysis by gel permeation chromatography (GPC) in comparison with ethanol extracted lignin using polystyrene standards for molecular weight approximation (*Figure 5.2AB*) showed that the approximate molecular weight of the DHP (2.0 kDa) to be very similar to that of the extracted lignins, which range from 1.8 to 2.2 kDa (*Figure 5.2C*) [*16*]. GPC analysis also showed that the ethanol precipitation was very effective at removing monomer contamination and also slightly shifted the molecular weight distribution by an additional 100 Da (*Figure 5.3*).

A. sp. 75iv2 was then grown in SMM in the presence and absence of the three different lignin types (0.05% (w/v)) and monitored for a physiological response to lignin over 3 d (*Figure 5.4*). As a comparison, glucose, xylose, and glycerol (0.05% (w/v)) were also used as carbon sources for comparison of growth on six-carbon, five-carbon, and reduced sugar sources, respectively. Interestingly, a significant increase in *A.* sp. 75iv2 growth was found in the presence of all three types of lignin compared to the sugar controls within 24 h of inoculation (*Figure 5.4A*). The increased growth of *A.* sp. 75iv2 was further verified by plating the cell culture, which confirmed that the increase in OD_{600} was directly related to an increase in cell count (*Figure 5.4B*). The colonies were also identified as *A.* sp. 75iv2 rather than a contaminating microbial strain both by visual inspection and by 16S rRNA sequencing. Moreover, the increased growth was accompanied by an induction of an extracellular peroxidase response that peaked at 48 h and whose magnitude depended on the lignin source (*Figure 5.4C*). In this case, cell growth in the presence of ethanol-extracted lignin or DHP resulted in an 80 to 100-fold increase in peroxidase activity, whereas addition of dioxane-extracted lignin led to a smaller 20-fold increase. Analysis of the secretome by heme-staining of SDS-PAGE gels further revealed the existence of a change in the heme protein profile with the induction of low molecular weight proteins (*Figure 5.4D*). Based on these studies, we also analyzed the growth of *A.* sp. 75iv2 in MM with lignin as a sole carbon source over 28 d, as growth was significantly slower, which demonstrated that the growth differences were also found under these conditions (*Figure 5.5*).

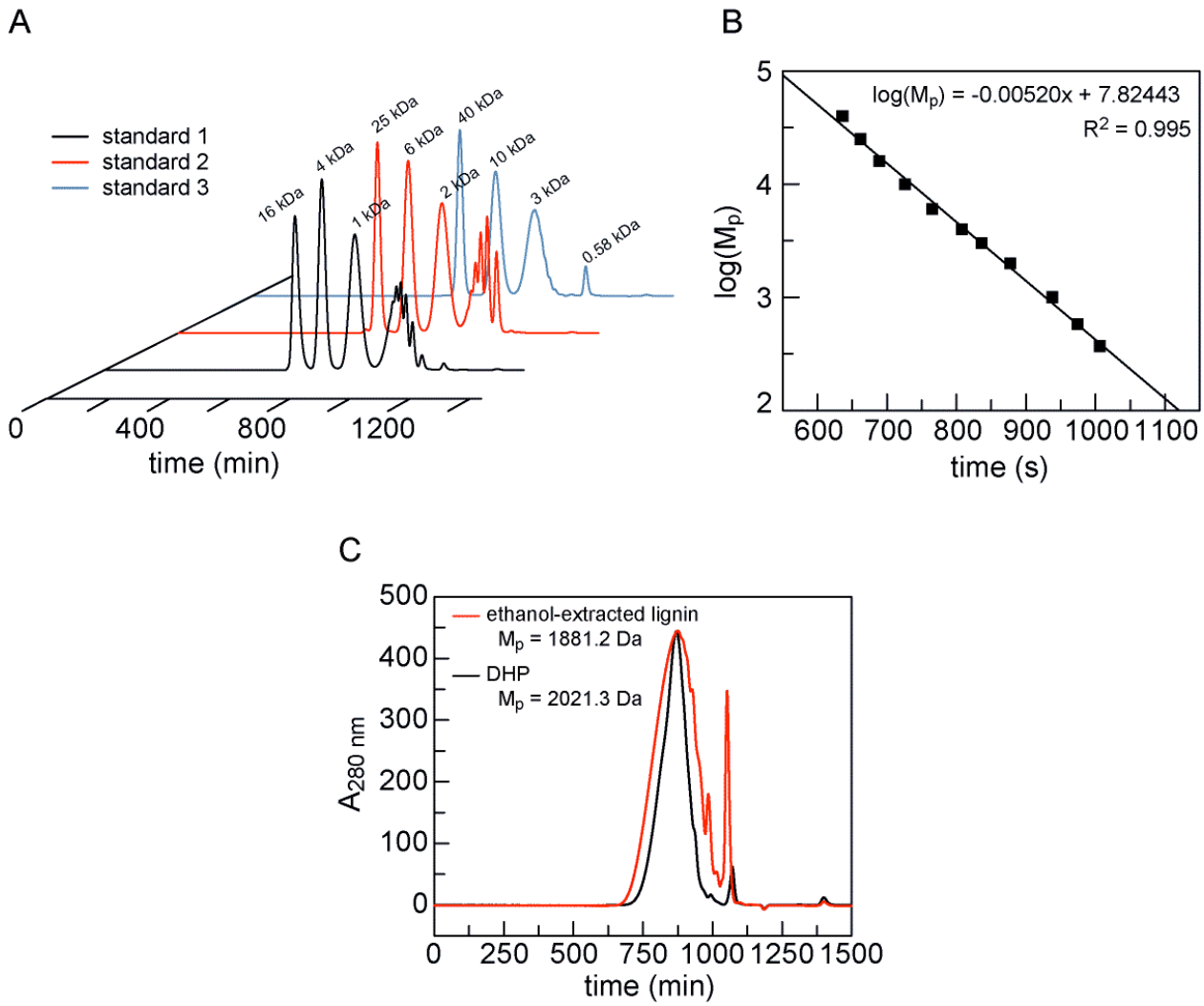


Figure 5.1. Characterization of DHP by gel permeation chromatography. *(A) Chromatograms of polystyrene molecular weight standards monitoring at 260 nm. (B) Standard curve for the polystyrene molecular weight standards. (C) Comparison of approximate molecular weight peaks of ethanol-extracted lignin and DHP shows the M_p of DHP to be similar to that of extracted lignins.*

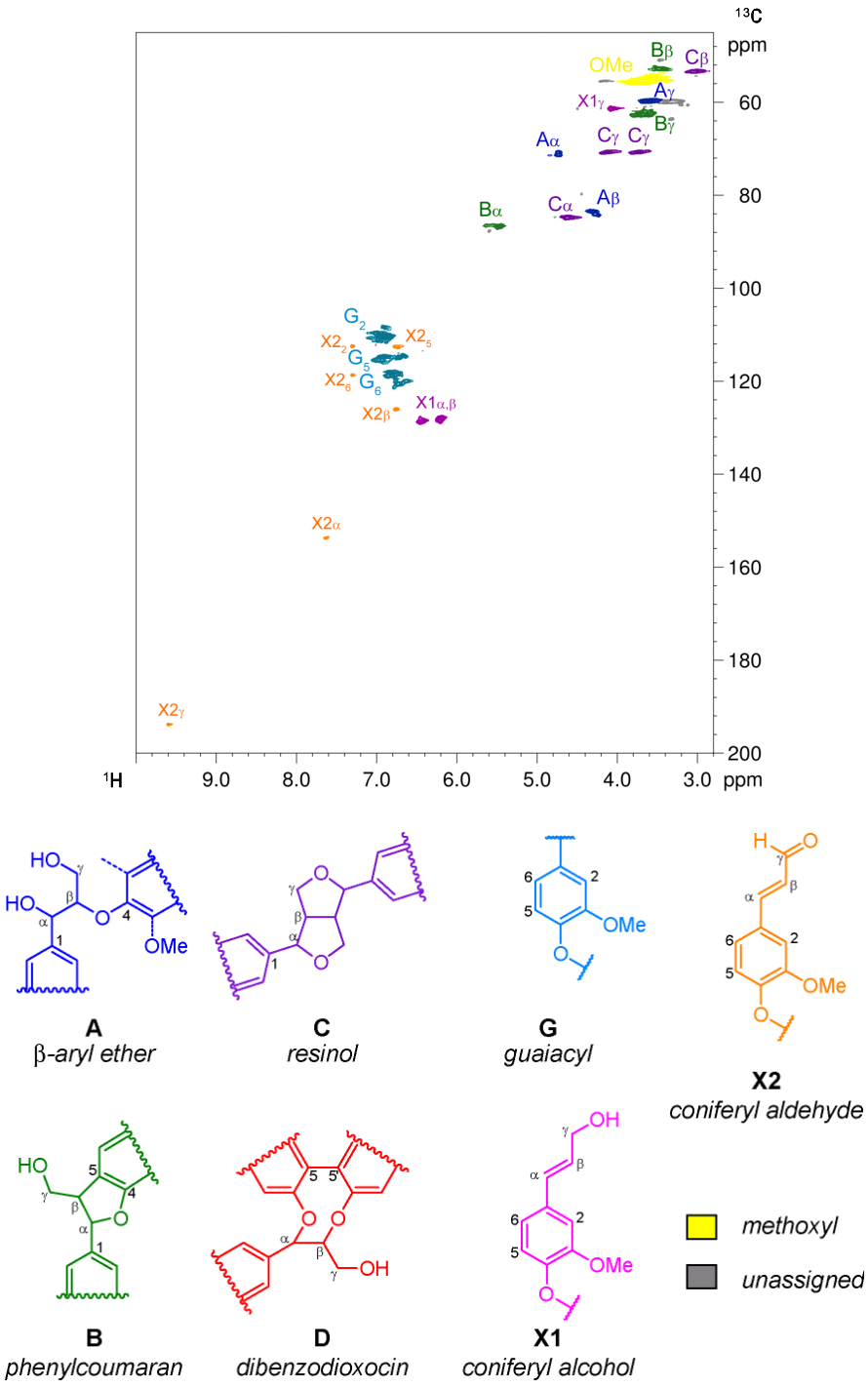


Figure 5.2. $^{1}H/^{13}C$ HSQC spectrum of DHP. *DHP was dissolved in DMSO-d_6 (20 mg/mL) and analyzed on a Bruker AV-600 spectrometer. Peaks were assigned based on literature values for the chemical shifts [16, 38] and are designated by color.*

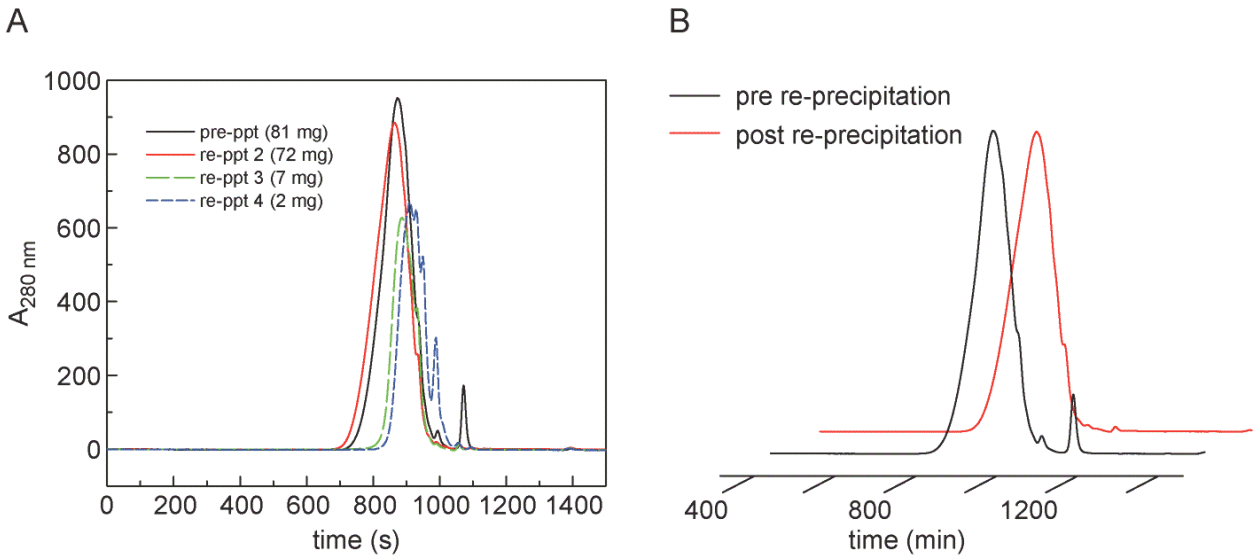


Figure 5.3. Characterization of DHP re-precipitation by gel permeation chromatography. *(A) Chromatograms of serial DHP precipitation fractions show conditions precipitate different Mp fractions of DHP (fraction 2, 25% ethanol; fraction 3, 11% ethanol; fraction 4, 8% ethanol) at 280 nm. (B) Standard curve for re-precipitation (using final 25% ethanol conditions) shows removal of contaminating monomer.*

Characterizing the transcriptional response of *A.* sp. 75iv2 to lignin by RNA sequencing. As a significant physiological response to lignin with respect to cell growth, extracellular peroxidase activity and a secreted heme protein profile was observed, we sought to characterize changes in the *A.* sp. 75iv2 transcriptional profile related to utilization of lignin as a carbon source. Given issues of mRNA quality associated with long-term growths, cells were cultured with SMM, using samples with no second carbon source as a control. Dioxane-extracted lignin and synthetic lignin (DHP) were selected as the lignin sources because of the structural modification of lignin observed in ethanol-extracted samples [*16*] and the large increase in extracellular peroxidase activity, respectively. After 24 h of growth on these substrates, total RNA was extracted and cDNA libraries were prepared after two rounds of rRNA depletion (*Figure 5.6*). The resulting libraries were sequenced using the HiSeq Illumina platform to obtain 50 base single-end reads. Reads from these libraries were mapped to the draft genome of *A.* sp. 75iv2 [*25*] using BWA [*26*] and the number of reads per putative open reading frame (ORF) was determined using HTseq-count [*27*].

The presence of dioxane-extracted lignin induced a strong transcriptional response with 1427 differentially-expressed ORFs (p-value < 0.05 and fold-change > 2), of which 762 ORFS were upregulated and 665 were downregulated (*Figure 5.7, Appendix Table A3.1*). This profile was very similar to the one obtained by growth on DHP, so we therefore focused on a more detailed analysis of the dioxane-extracted lignin (*Figures 5.8* and *5.9, Appendix Tables A3.1* and *A3.2*). Overall, the largest change in gene expression appears to be related to iron assimilation. Indeed, the most significantly upregulated ORF (140-fold increase) encodes a non-ribosomal peptide synthase (NRPS), which is located in a very highly upregulated cluster that includes siderophore-related enzymes (*Table 5.2*). Furthermore, four of the ten Fe^{3+}-hydroxamate transporters encoded in the genome are significantly upregulated (4- to 28-fold) and none are downregulated.

Two extremely large clusters (39 and 40 ORFs) along with six large clusters (with >15 ORFS) were also found to be significantly upregulated due to the presence of lignin (*Table 5.2ab*). Interestingly, both of the extremely large clusters encode a high number of proteins with unknown function (25% of each cluster) as well a number of redox-active enzymes, such as monooxygenases and cytochrome oxidases. These two clusters also include a number of universal stress proteins, including ten of the total seventeen universal stress protein genes found in the *A.* sp. 75iv2 genome. Outside of these two clusters, an additional two universal stress proteins are upregulated in response to lignin. As *uspA* mutants in *E. coli* have shown increased sensitivity to carbon starvation [*39*], it is interesting that upregulation occurs in the presence of lignin, suggesting that universal stress protein transcription is not caused by a limited carbon source in this growth. Finally, the A-type dye peroxidase (DyP3) is also highly upregulated (6.5-fold) within one of these larger clusters; however, the identity of the cellular role of this cluster is unclear. Also apparent within these two large highly upregulated clusters was the high quantity of genes encoding proteins with transmembrane helices as predicted by TMHMM [*40*]. The entire genome and transcriptome were analyzed, revealing that ORFs with predicted transmembrane helices were statistically significantly enriched among the ORFs upregulated by the presence of lignin (201 of 762 transcripts, 26.4%) and statistically significantly under represented among those downregulated (108 out of 665 transcripts, 16.2%) compared to those found in the entire genome (1745 out of 8265 genes, 21.1%). This observation is interesting

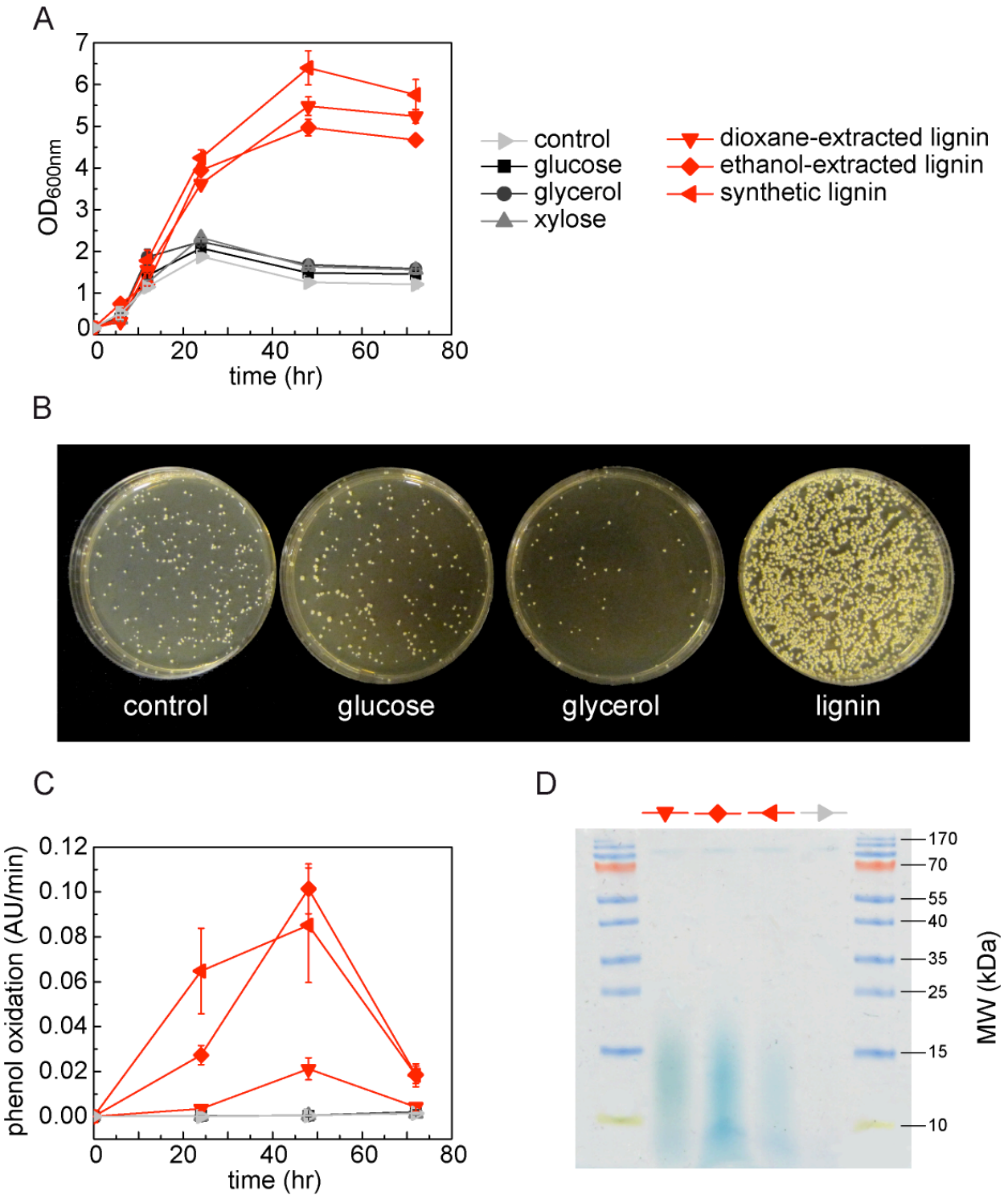


Figure 5.4. Growth of A. sp. 75iv2 on lignin in SMM. *(A) Growth in supplemented minimal media (SMM) containing casamino acids as a carbon source monitored by OD_{600}. Dioxane-extracted, ethanol-extracted, and synthetic (DHP) lignins were added as a second carbon source and compared to a control with no added carbon and carbohydrate controls consisting of glucose, glycerol, and xylose. All second carbon sources were included at 0.05% (w/v). (B) Culture samples were removed after 48 h of growth on SMM and plated on YEME agar after a 10^5-fold dilution to correlate the increase in OD_{600} with the increase in growth of A. sp. 75iv2. (C) Extracellular peroxidase activity monitored by the oxidation of 2,4-dichlorophenol in the presence of exogenous H_2O_2. (D). Heme-stain SDS-PAGE of the A. sp. 75iv2 secretome after 48 h of cell growth in the presence of lignins compared to the control with no second carbon source.*

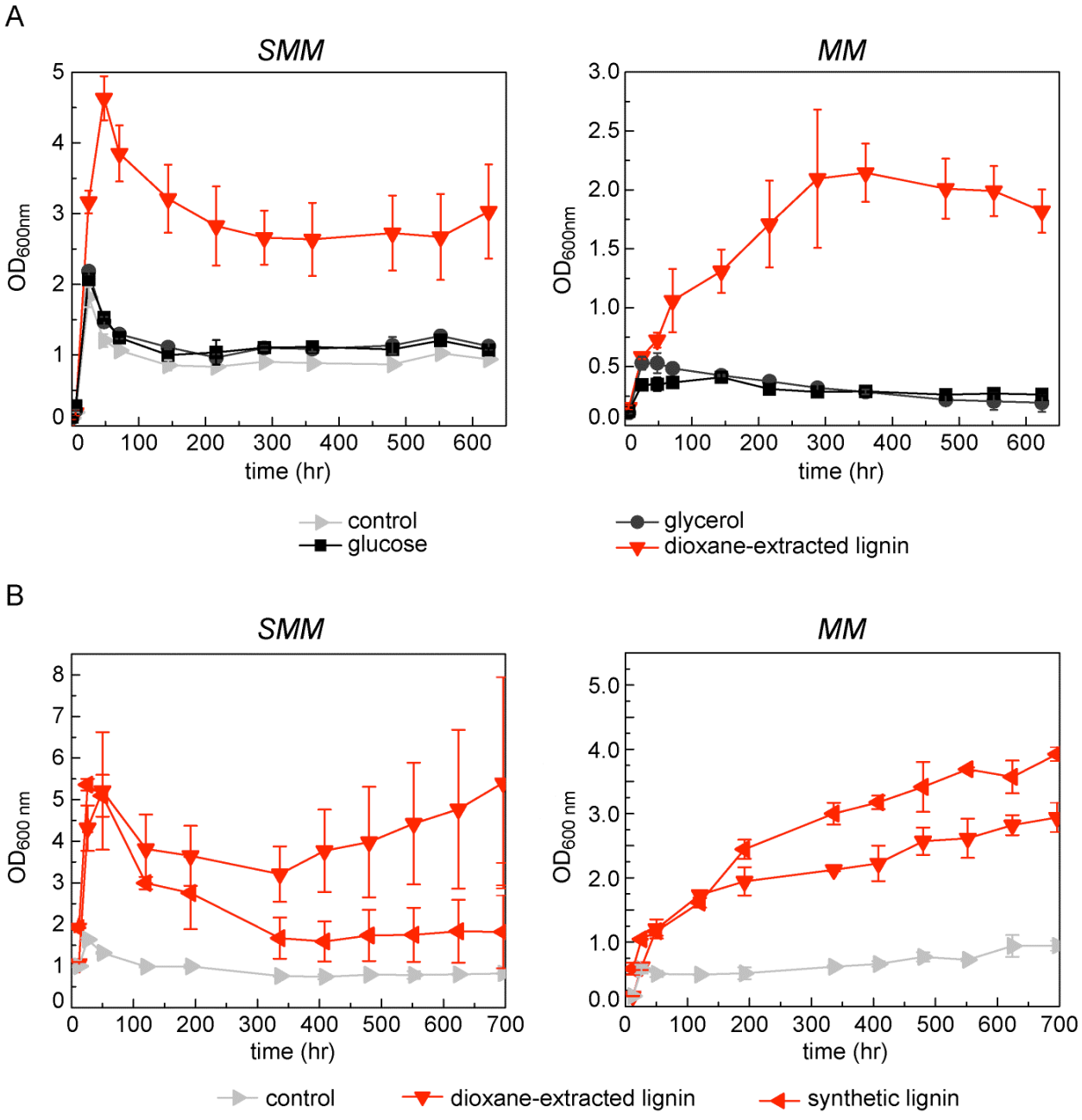


Figure 5.5. Long-term growths of A. sp. 75iv2 on lignin in SMM and MM. *(A) Growth in supplemented minimal media (SMM) and minimal media (MM). The SMM contains casmino acids as a carbon source monitored by OD_{600}. Dioxane-extracted lignin was added as a second carbon source and compared to a control with no added carbon and sugar controls consisting of glucose and glycerol. All second carbon sources were included at 0.05% (w/v). In the MM growths, lignin was added as the sole carbon source and compared to a control with no added carbon and carbohydrate controls consisting of glucose and glycerol. All carbon sources were included at 0.05% (w/v). (B) Growth in SMM and MM monitored by OD_{600} comparing long-term growth on dioxane-extracted lignin to growth on synthetic lignin as previously stated. All second carbon sources were included at 0.05% (w/v).*

given that lignin is an extracellular substrate and we would expect genes involved in its metabolism to be extracellular in location. Therefore, the genome and transcriptome were also analyzed for ORFs predicted to have secretion signals by SignalP 4.1 [*41, 42*]. However, there was not a statistically significant over representation of transcripts with signal sequences due to the presence of lignin.

Only two large clusters were downregulated in response to lignin. One cluster is composed of many unknown proteins and a cytochrome P450. The other cluster appears to involve short-chain carbohydrate modification; also within this cluster is an ORF predicted to be a phenylacetic acid-responsive transcriptional repressor as well as an ORF predicted to be an uncharacterized protein dealing with aromatic catabolism. The most downregulated gene in the transcriptional analysis is an acyl-coA synthetase (315-fold downregulated). Also, 58-fold downregulated is the catalase, suggesting the cell is decreasing H_2O_2 disproportionation and, thus, increasing H_2O_2 for reactivity.

Changes in the expression of small molecule aromatic degradation pathways that could be involved in metabolism of lignin-derived monomers were also observed. A number of aromatic-catabolism ORFs were found to be upregulated due to the presence of lignin: a *p*-hydroxybenzoate hydroxylase (*pobA*) ortholog (AmyDRAFT39116_6443, 4-fold), a xylene oxygenase (*xylA*) ortholog (AmyDRAFT39116_1410, 3-fold), a C_α-dehydrogenase (*ligD*), ortholog (along with its cluster AmyDRAFT39116_7271-_7276, 4-fold), and a phthalate 4,5-dioxygenase ortholog (AmyDRAFT39116_3552, 18-fold). Many more ORFs were downregulated due to the presence of lignin: a β-ketoadipate enolactone hydroxylase (*pcaD*) ortholog (AmyDRAFT39116_6905, 8-fold), a β-ketoadipate succinyl-CoA transferase (*pcaIJ*) ortholog (AmyDRAFT39116_7764, 26-fold), a β-ketoadipate CoA thiolase (*pcaF*) (AmyDRAFT39116_0134, 5.3-fold) ortholog, and a dienlactone hydrolase ortholog (AmyDRAFT39116_3331, 5-fold). Also, a cluster previously predicted to be a phenylacetate degradation pathway (AmyDRAFT39116_7836 through _7844) was found to be downregulated 6.5 to 9-fold. Although this response was surprising, it may indicate that only specific clusters are responsible for the assimilation of aromatic compounds derived from dioxane-extracted lignin or DHP, which is made solely from coniferyl alcohol.

BiNGO [*28*] was used to analyze the functional enrichment of differentially expressed genes based on clusters of orthologs groups (COGs), allowing for overall basic identification of global cellular response to the presence of lignin (*Figure 5.10*). From this analysis, it is immediately apparent that proteins involved in carbohydrate transport and metabolism are highly underrepresented in both the up- and downregulated transcriptional response to lignin. Proteins involved in translation are highly over represented among the upregulated ORFs while those involved in transcription are highly under represented; however, among the downregulated ORFS, protein involved in transcription are highly overrepresented. Finally, as to be expected, the transcripts involved in inorganic ion transport and metabolism are significantly over represented among the upregulated ORFs.

Characterizing changes in the intracellular protein profile of *A.* sp. 75iv2 in response to lignin. Changes in the intracellular protein profile in SMM cultures grown in the presence and absence of dioxane-extracted lignin were also characterized. Biomass samples were collected at 24 and 48 h and processed to yield tryptic peptides that were then analyzed using MudPIT [*30*], which identified a total of 1237 proteins. Proteins showing <5 spectral counts were eliminated,

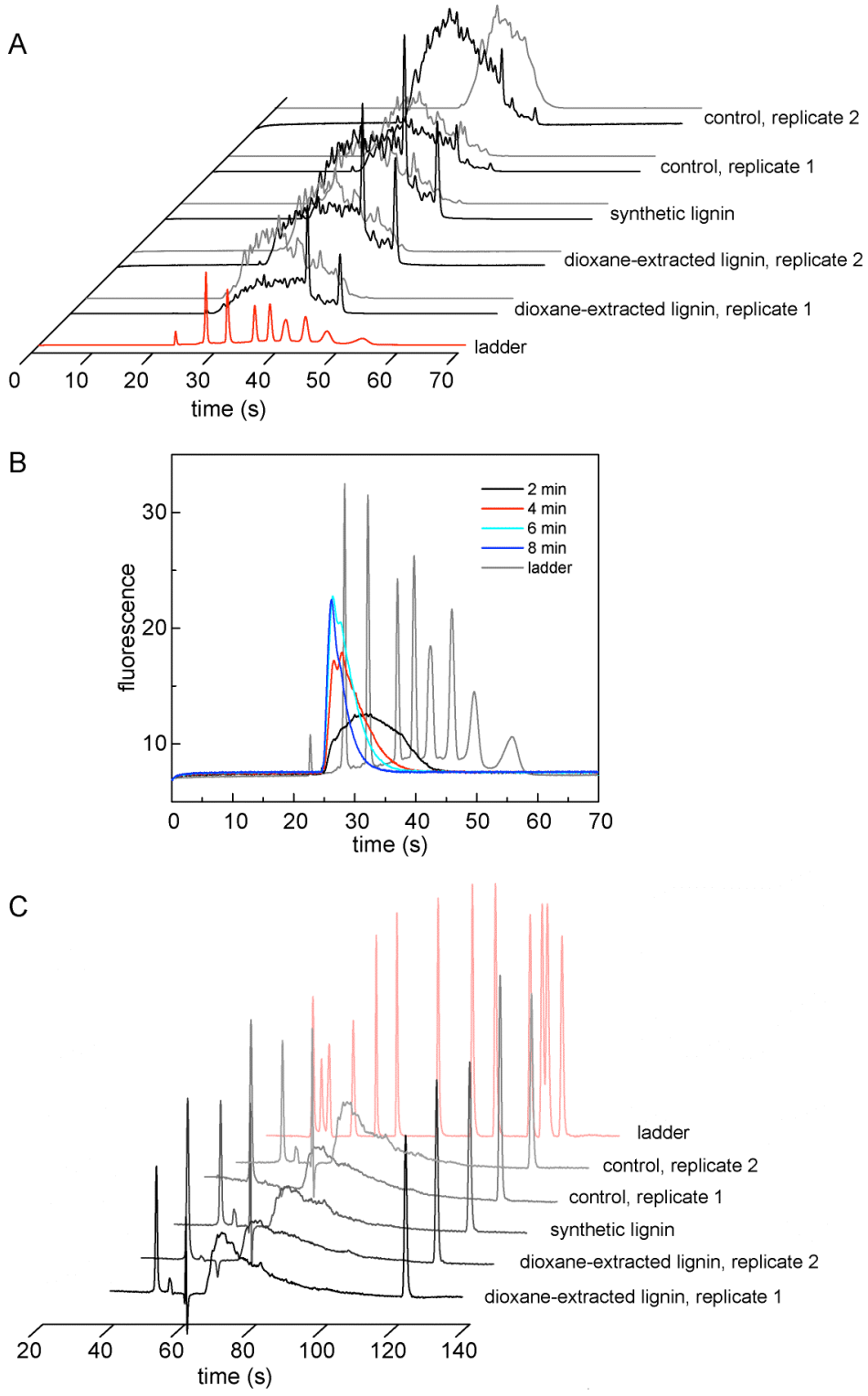


***Figure 5.6*. Characterization of the RNA and cDNA libraries.** *(A) Total g-DNA-depleted RNA preparations (black) and rRNA-depleted mRNA preparations (gray) after two rounds of rRNA depletion at 24 hrs. (B) Fragmentation control of total RNA identifying two peaks (sizing to approximately 145 and 190 nucelotides) that appear after hydrolysis for 4 minutes. After hydrolysis for 8 minutes only the peak representing 145 nucleotides remains. Bioanalyzer traces of cDNA libraries show an approximate average size of theadapter-ligated libraries to be about 225 bp.*

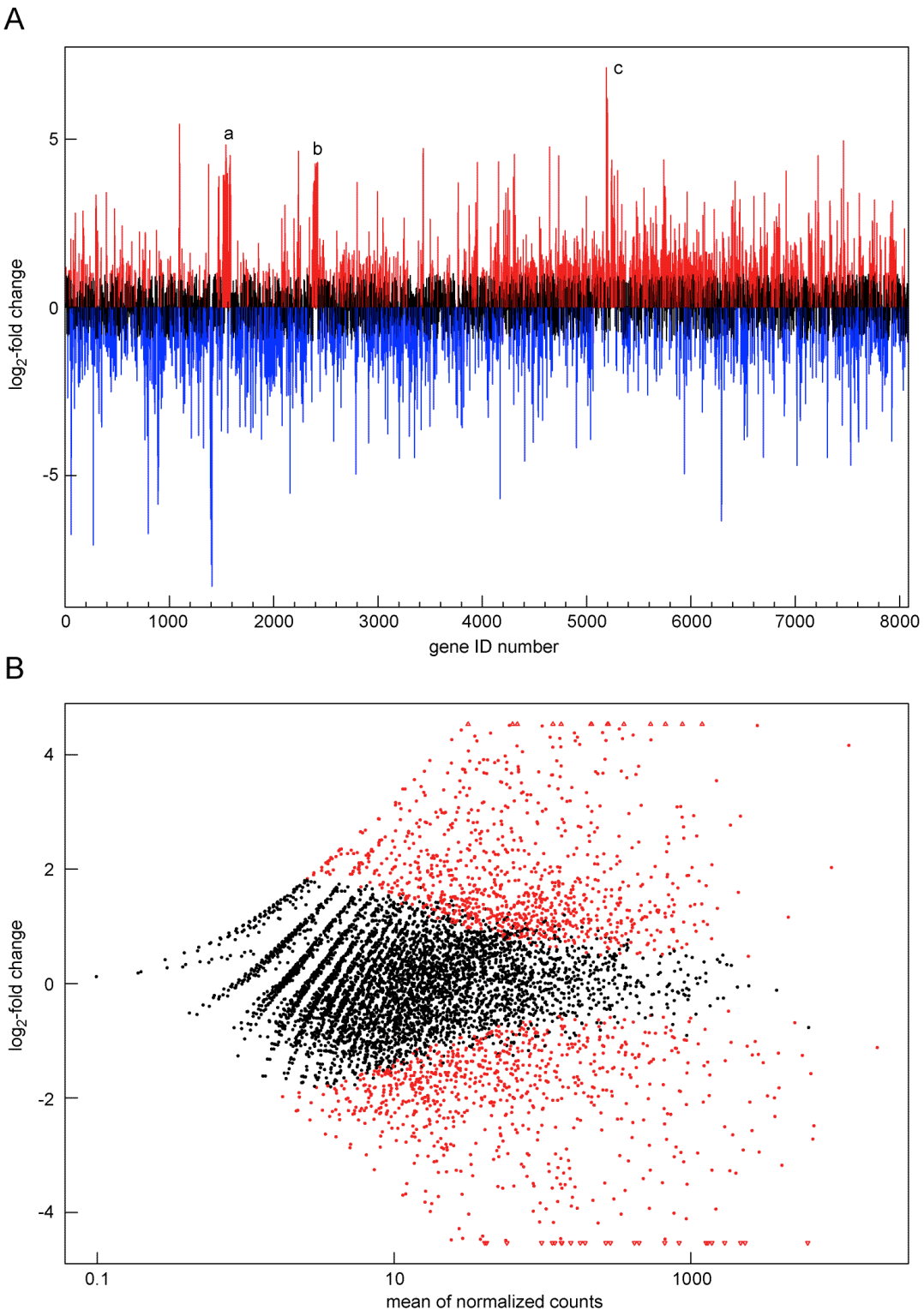


***Figure 5.7*. Characterization of the transcriptional response to dioxane-extracted lignin at 24 h.** *Differential expression landscape of A. sp. 75iv2 with respect to the presence of dioxane-extracted lignin compared to the no lignin control. Upregulated (red) and downregulated (blue) ORFs are highlighted. Large clusters of upregulated ORFs are also indicated (a-c). (B) MA-plot shows DESeq2's normalization procedures are sufficient to deal with biases in lane-to-lane variation from sequencing by the Illumina platform.*

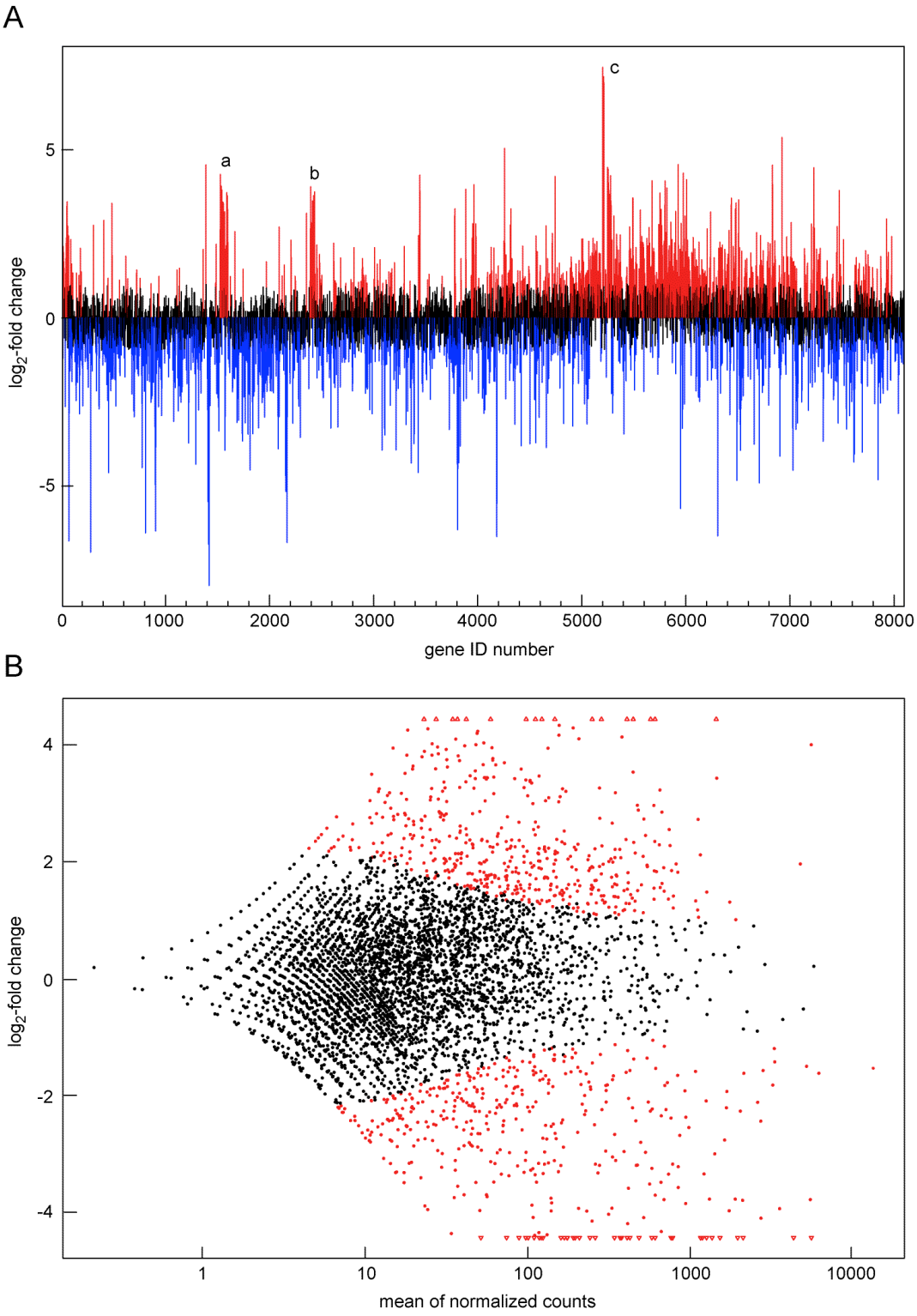


Figure 5.8. Characterization of the transcriptional response to DHP at 24 h. *(A) Differential expression landscape of A. sp. 75iv2 (775 ORFS with p-value < 0.05 and fold-change > 2) with respect to the presence of DHP compared to the no lignin control. Upregulated (420, red) and downregulated (355, blue) ORFs are highlighted. Large clusters of upregulated ORFs are also indicated (a-c). (B) MA-plot shows DESeq2's normalization procedures are sufficient to deal with biases in lane-to-lane variation from sequencing by the Illumina platform.*

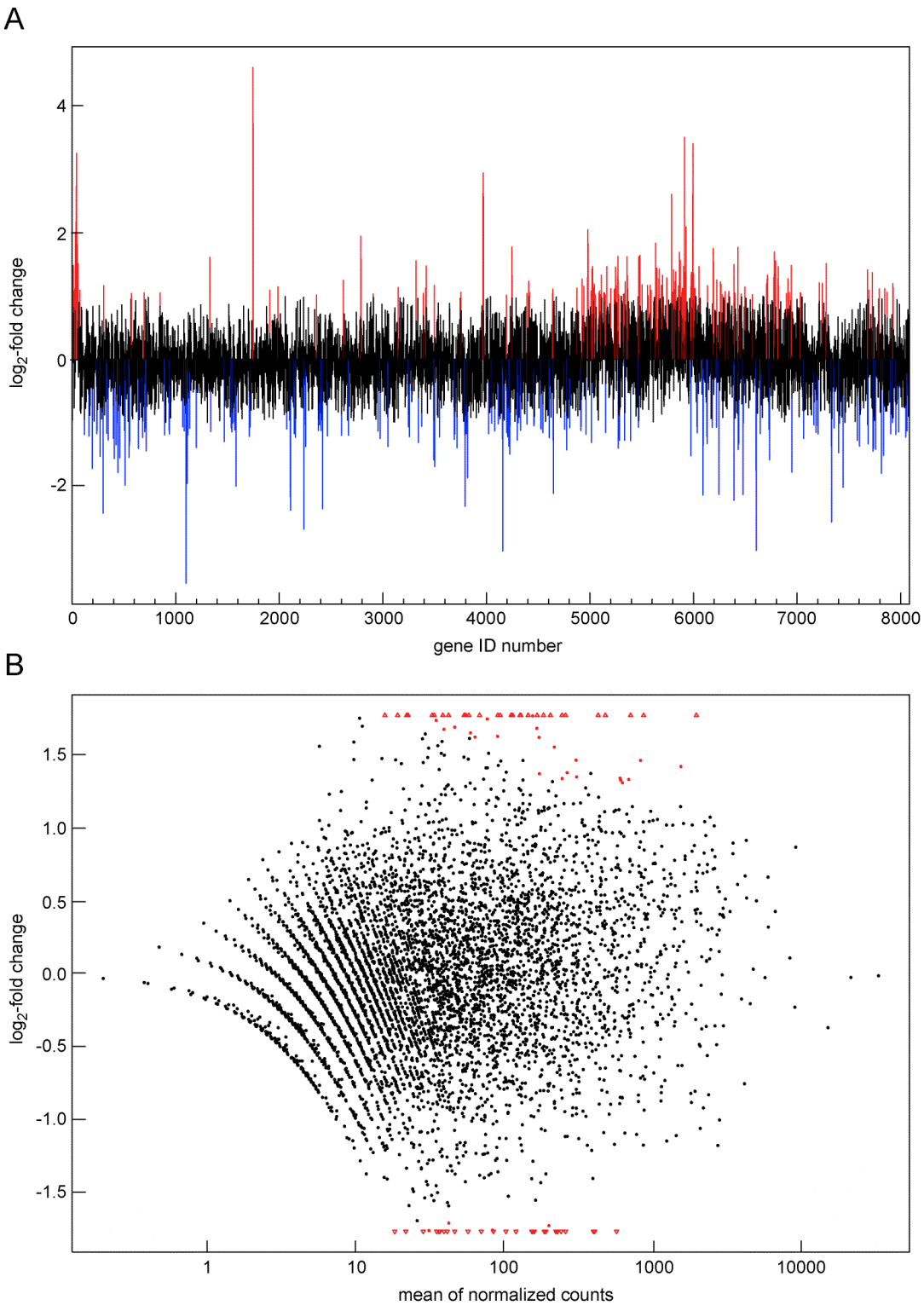


Figure 5.9. Comparison of the transcriptional response to dioxane-extracted lignin and DHP at 24 h. *A) Differential expression landscape A. sp. 75iv2 (63 ORFS with p-value < 0.05 and fold-change > 2) with respect to the presence of dioxane lignin compared to DHP. Uupregulated (38, red) and downregulated (25, blue) ORFs are highlighted. (B) MA-plot shows DESeq2's normalization procedures are sufficient to deal with biases in lane-to-lane variation from sequencing by the Illumina platform.*

Table 5.2 *Upregulated transcriptional clusters due to the presence of dioxane-extracted lignin. Statistically significant upregulated genes are based on >2-fold increase in transciptional regulation and a p-value <0.05 as calculated by the student's T-Test. SS = predicted signal sequence as predicted by SignalP 4.1. TMH = predicted transmembrane helices as predicted by TMHMM.*

cluster a

gene ID	gene annotation	Log_2-fold change	P-value adj	SS?	TMH?
Amy39116DRAFT_1576	hypothetical protein	1.368	2.59E-04	no	no
Amy39116DRAFT_1578	hypothetical protein	3.916	1.17E-09	no	yes
Amy39116DRAFT_1579	dehydrogenases with different specificities	3.670	6.36E-22	no	no
Amy39116DRAFT_1580	ribosome-associated protein Y (PSrp-1)	3.705	2.22E-10	no	no
Amy39116DRAFT_1581	transcriptional regulator, effector-binding domain/component	3.861	1.79E-07	no	no
Amy39116DRAFT_1582	alkane 1-monooxygenase	2.324	2.78E-05	no	yes
Amy39116DRAFT_1584	ferredoxin reductase	1.983	3.59E-02	no	no
Amy39116DRAFT_1586	thiol reductant ABC exporter, CydD subunit	3.748	4.27E-10	no	yes
Amy39116DRAFT_1587	cytochrome bd-I oxidase, subunit II (cydB)	3.330	1.61E-07	no	yes
Amy39116DRAFT_1588	cytochrome bd-type quinol oxidase, subunit I	3.923	6.89E-17	no	yes
Amy39116DRAFT_1589	predicted transcriptional regulator, contains C-terminal CBS domains	2.274	1.09E-06	no	no
Amy39116DRAFT_1590	hypothetical protein	2.096	2.13E-02	no	no
Amy39116DRAFT_1591	hypothetical protein	2.650	4.28E-09	no	yes
Amy39116DRAFT_1593	carbamate kinase	1.951	1.47E-02	no	no
Amy39116DRAFT_1594	ornithine carbamoyltransferase	3.748	1.23E-08	no	no
Amy39116DRAFT_1595	arginine deiminase	2.497	1.74E-05	no	no
Amy39116DRAFT_1596	amino acid transporters	3.508	3.13E-06	no	yes
Amy39116DRAFT_1597	hypothetical protein	3.696	4.31E-07	no	yes
Amy39116DRAFT_1598	zinc-binding alcohol dehydrogenase family protein	3.403	9.19E-05	no	no
Amy39116DRAFT_1599	universal stress protein UspA and related nucleotide-binding proteins	2.259	4.74E-07	no	no
Amy39116DRAFT_1600	nitroreductase family protein	2.087	7.22E-08	no	no
Amy39116DRAFT_1601	hypothetical protein	2.451	1.69E-03	no	no
Amy39116DRAFT_1602	hypothetical protein	2.255	1.87E-05	no	yes
Amy39116DRAFT_1603	hypothetical protein	4.346	1.11E-27	no	no
Amy39116DRAFT_1604	universal stress protein UspA and related nucleotide-binding proteins	4.261	1.67E-16	no	no
Amy39116DRAFT_1605	hypothetical protein	4.388	2.58E-12	no	no
Amy39116DRAFT_1606	universal stress protein UspA and related nucleotide-binding proteins	4.828	3.20E-23	no	no
Amy39116DRAFT_1607	acyl-CoA synthetase (NDP forming)	4.451	3.97E-22	no	no
Amy39116DRAFT_1608	high affinity sulphate transporter 1	2.997	3.64E-05	no	yes
Amy39116DRAFT_1609	universal stress protein UspA and related nucleotide-binding proteins	2.028	3.18E-04	no	no
Amy39116DRAFT_1610	transposase and inactivated derivatives	2.752	9.85E-17	no	no
Amy39116DRAFT_1611	hypothetical protein	2.799	2.02E-05	no	yes
Amy39116DRAFT_1612	universal stress protein UspA and related nucleotide-binding proteins	2.063	1.97E-03	no	no
Amy39116DRAFT_1613	hypothetical protein	1.719	7.57E-05	no	no
Amy39116DRAFT_1616	hypothetical protein	1.578	4.28E-02	no	no
Amy39116DRAFT_1617	predicted xylanase/chitin deacetylase/polysaccharide deacetylase	2.725	1.97E-14	yes	no
Amy39116DRAFT_1619	monooxygenase	2.109	1.45E-02	no	no
Amy39116DRAFT_1620	hypothetical protein	3.856	5.59E-06	no	no
Amy39116DRAFT_1621	hypothetical protein	3.978	4.48E-08	no	no
Amy39116DRAFT_1622	predicted phosphoribosyltransferases	2.820	4.50E-05	no	no

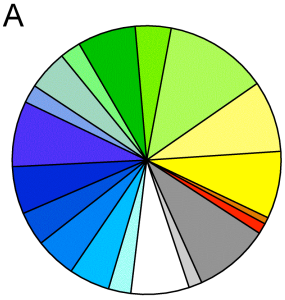
cluster b

gene ID	gene annotation	log_2-fold change	P-value adj	SS?	TMH?
Amy39116DRAFT_2467	hypothetical protein	2.338	3.91E-07	yes	yes
Amy39116DRAFT_2468	hypothetical protein	2.158	2.96E-03	no	no
Amy39116DRAFT_2469	hypothetical protein	3.049	1.83E-09	no	no
Amy39116DRAFT_2470	predicted signal-transduction protein, contains cAMP-binding and CBS domains	2.974	2.78E-13	no	no
Amy39116DRAFT_2471	universal stress protein UspA and related nucleotide-binding proteins	3.096	6.22E-12	no	no
Amy39116DRAFT_2472	hypothetical protein	3.455	4.09E-09	no	yes
Amy39116DRAFT_2473	hypothetical protein	3.193	9.71E-12	yes	no
Amy39116DRAFT_2474	glyceraldehyde-3-phosphate dehydrogenase, type I	3.673	1.45E-08	no	no
Amy39116DRAFT_2475	fructose 1,6-bisphosphatase	2.673	2.94E-03	no	no
Amy39116DRAFT_2476	Zn-dependent proteases	2.117	3.25E-02	no	yes
Amy39116DRAFT_2477	ATP-dependent metalloprotease FtsH	3.714	5.82E-15	no	yes
Amy39116DRAFT_2478	hypothetical protein	3.749	1.59E-08	yes	no
Amy39116DRAFT_2479	acyltransferase, WS/DGAT/MGAT	3.039	1.93E-05	no	no
Amy39116DRAFT_2480	glycosyltransferase	3.009	5.71E-10	no	no
Amy39116DRAFT_2481	galactose-1-phosphate uridylyltransferase	1.825	6.21E-03	no	no
Amy39116DRAFT_2482	immunoglobulin-like domain of bacterial spore germination	2.149	5.28E-03	no	yes
Amy39116DRAFT_2485	ABC-type multidrug transport system, permease component	2.564	4.71E-03	no	yes
Amy39116DRAFT_2488	predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	1.582	2.39E-02	no	no
Amy39116DRAFT_2489	Zn-dependent oxidoreductase	4.265	6.75E-08	no	no
Amy39116DRAFT_2490	ATPase, P-type (transporting), HAD superfamily, subfamily IC	3.617	3.70E-14	no	yes
Amy39116DRAFT_2491	ATPase, P-type (transporting), HAD superfamily, subfamily IC	3.679	3.45E-18	no	yes
Amy39116DRAFT_2492	CBS-domain-containing membrane protein	3.186	3.02E-20	no	no
Amy39116DRAFT_2493	DoxX	3.106	3.94E-16	no	yes
Amy39116DRAFT_2497	universal stress protein UspA and related nucleotide-binding proteins	2.788	1.10E-03	no	no
Amy39116DRAFT_2499	heavy metal translocating P-type ATPase	3.702	1.02E-11	no	yes
Amy39116DRAFT_2500	hypothetical protein	2.293	5.74E-03	no	yes
Amy39116DRAFT_2501	acyltransferase, WS/DGAT/MGAT	3.686	2.92E-09	no	no
Amy39116DRAFT_2502	hypothetical protein	4.183	6.81E-16	no	yes
Amy39116DRAFT_2503	phosphoenolpyruvate synthase/pyruvate phosphate dikinase	4.281	1.46E-20	no	no
Amy39116DRAFT_2504	universal stress protein UspA and related nucleotide-binding proteins	3.091	1.25E-20	no	no
Amy39116DRAFT_2505	transcriptional regulator	2.717	1.05E-10	no	no
Amy39116DRAFT_2506	universal stress protein UspA and related nucleotide-binding proteins	3.241	5.46E-13	no	no
Amy39116DRAFT_2507	nitroreductase family	3.098	3.91E-10	no	no
Amy39116DRAFT_2508	flavodoxin	3.799	2.32E-46	no	no
Amy39116DRAFT_2509	universal stress protein UspA and related nucleotide-binding proteins	3.423	1.78E-08	no	no
Amy39116DRAFT_2510	hypothetical protein	3.901	1.06E-09	no	yes
Amy39116DRAFT_2511	hypothetical protein	4.310	1.35E-16	no	yes
Amy39116DRAFT_2512	CBS-domain-containing membrane protein	3.348	8.35E-09	no	no
Amy39116DRAFT_2513	anti-anti-sigma factor	3.679	1.10E-10	no	no

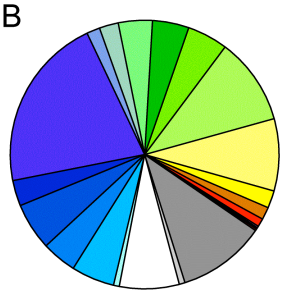
cluster c

gene ID	gene annotation	log_2-fold change	P-value adj	SS?	TMH?
Amy39116DRAFT_5362	ABC-type Fe3+-hydroxamate transport system, periplasmic component	4.778	1.96E-18	yes	no
Amy39116DRAFT_5363	amino acid adenylation domain, putative NRPS	7.125	1.10E-52	no	no
Amy39116DRAFT_5364	amino acid adenylation domain	6.249	1.29E-41	no	no
Amy39116DRAFT_5366	antibiotic synthesis protein MbtH	3.683	2.39E-06	no	no
Amy39116DRAFT_5367	predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	3.911	1.89E-07	no	no
Amy39116DRAFT_5368	siderophore export protein, MFS transporter	2.568	9.30E-07	no	yes
Amy39116DRAFT_5369	peptide arylation enzyme, enterobactin synthase subunit E	3.897	1.57E-11	no	no
Amy39116DRAFT_5370	ABC-type Fe3+-siderophore transport system, permease component	2.861	3.52E-04	no	yes
Amy39116DRAFT_5371	ABC-type enterobactin transport system, permease component	3.291	5.09E-05	no	yes
Amy39116DRAFT_5372	salicylate synthase	6.189	7.18E-22	no	no
Amy39116DRAFT_5373	siderophore biosynthesis protein, monooxygenase	5.634	4.69E-18	no	no
Amy39116DRAFT_5374	methionyl-tRNA formyltransferase	5.775	4.28E-24	no	no

A



B



C

COG category	Genome	Up	Down
RNA processing and modification	6 (0.082)	0 (0.0)	1 (0.16)
Chromatin structure and dynamics	2 (0.027)	0 (0.0)	1 (0.16)
Energy production and conversion	584 (7.97)	57 (9.06)	65 (10.62)
Cell cycle control, cell division, chromosome partitioning	47 (0.64)	10 (1.59)	4 (0.65)
Amino acid transport and metabolism	539 (7.35)	44 (7.00)	44 (7.19)
Nucleotide transport and metabolism	118 (1.61)	17 (2.70)	4 (0.65)
Carbohydrate transport and metabolism	587 (8.01)	31 (4.93)	32 (5.23)
Coenzyme transport and metabolism	303 (4.13)	31 (4.93)	26 (4.25)
Lipid transport and metabolism	448 (6.11)	26 (4.13)	35 (5.72)
Translation, ribosomal structure and biogenesis	218 (2.97)	36 (5.72)	19 (3.10)
Transcription	1018 (13.88)	50 (7.95)	129 (21.08)
Replication, recombination and repair	224 (3.06)	14 (2.23)	10 (1.63)
Cell wall/membrane/envelope biogenesis	255 (3.48)	30 (4.77)	14 (2.29)
Cell motility	5 (0.068)	0 (0.0)	0 (0.0)
Posttranslational modification, protein turnover, chaperones	149 (2.03)	16 (2.54)	25 (4.09)
Inorganic ion transport and metabolism	259 (3.53)	44 (7.00)	27 (4.41)
Secondary metabolites biosynthesis, transport and catabolism	397 (5.41)	27 (4.29)	30 (4.90)
General function prediction only	1117 (15.24)	78 (12.40)	64 (10.46)
Function unknown	526 (7.17)	54 (8.59)	53 (8.66)
Signal transduction mechanisms	379 (5.17)	51 (8.11)	13 (2.12)
Intracellular trafficking, secretion, and vesicular transport	44 (0.60)	5 (0.79)	9 (1.47)
Defense mechanisms	106 (1.45)	8 (1.27)	6 (0.98)
Extracellular structures	1 (0.0136)	0 (0.0)	1 (0.16)
Total	**7332**	**629**	**612**

Figure 5.10. COG categories for genes differentially expressed with respect to growth in the presence of dioxane-extracted lignin. *COG categories utilized were identified in the published genome obtained from JGI. COG categories represented by genes that are (A) upregulated and (B) downregulated in the presence of dioxane lignin at 24 h. (C) Comparison of COG category representation in the differentially expressed genes with that of the entire genome. The number of open reading frames represented by each COG is given, and the percentage of total genes with COG categories is in parentheses. Values that are statistically significantly over- or under-represented (p-value > 0.05) are given in black.*

yielding 548 proteins. Upregulated and downregulated proteins were then identified by using a cut-off for the fold-change (>2-fold) and *p*-value calculated using the Student's t-test ($p < 0.05$). Average values were used to calculate ratios for fold-change. Using these criteria, 189 and 128 proteins were found up-regulated and down-regulated, respectively, in response to the dioxane-extracted lignin at 24 h (*Appendix Tables A3.3* and *A3.4)*. After 48 h, 133 up-regulated and 129 down-regulated proteins were identified (*Appendix Tables A3.5* and *A3.6*). Overlap was high between these categories, indicating that the overall protein profile was similar between 24 h and 48 h of growth on lignin. However, it appears as if there is a fine regulation of the intracellular proteomics, as proteins with similar roles appear to be at different times. For example, even though there are four Fe^{3+}-hydroxamate transporters differentially expressed at 24 h and five at 48 h, the identity of only three of these are the same. Similarly, different dioxygenases are expressed at 24 h from those expressed at 48 h. Overall, for the proteins found in highest abundance, there was a lowered expression from 24 to 48 h; however, for many of the less abundant proteins, the amplitude of the response remained approximately unchanged.

Overall, we found that the intracellular protein profile closely reflecting the transcriptional response observed in the RNA-seq dataset. Again, we found that members of the putative siderophore biosynthetic and transport cluster were significantly upregulated at 24 h (NRPS domain, AmyDRAFT39116_5365, 13-fold; dioxygenase, AmyDRAFT39116_5362, found only in dioxane-extracted ligning samples). At 48 h, we again find the up-regulation of the NRPS domain and dioxygenase in response to lignin but also begin to observe the onset of production of Fe^{3+}-hydroxamate transporters (AmyDRAFT39116_5658 and _7128). As before, both up- and down-regulation of aromatic degradation proteins occurred. Aromatic degradation proteins found upregulated at 24 h included a dioxygenase (AmyDRAFT39116_4182, 3.3-fold), a catechol-degradation pathway enzyme (AmyDRAFT39116_6535, 2.7-fold), and a phenylacetate-CoA ligase (AmyDRAFT39116_7836, 2.4-fold). In the down-regulated protein category, we found a 3-oxoadipate enol-lactonase (AmyDRAFT39116_7642, 4.2-fold), a phenylacetate-CoA oxygenase (AmyDRAFT39116_7841, found only in no lignin samples), and a dioxygenase (AmyDRAFT39116_4123, 7-fold). This response was only partially maintained in the 48 h sample as some of these aromatic degradation proteins were no longer part of the proteome (such as AmyDRAFT39116_6535, AmyDRAFT39116_7836, and AmyDRAFT39116_7642). Other observed changes include the decrease in protein levels of a catalase (AmyDRAFT39116_6493, 6-fold), ferredoxin (AmyDRAFT39116_5013, 6-fold) and a ferritin-like protein (AmyDRAFT39116_5851, found only no lignin samples) that could possibly be interpreted to be involved in the increased utilization of reactive oxygen species and iron. These changes in protein profile generally agree with the transcriptional response found, as the catalase was found to be downregulated 60-fold and the ferritin-like protein was found downregulated 4-fold at the mRNA level. A number of proteins with unknown function were also found to be expressed, and in general, there was a fairly high correlation between these proteins and upregulated transcripts, though these were not transcribed in the two most highly transcribed clusters due to the presence of lignin.

Analyzing the functional enrichment of differentially expressed proteins based on clusters of orthologs groups (COGs) using a hypergeometric test to deduce statistical significance (*Table 5.3*) shows a statistically significant overrepresentation of proteins that are involved in translation as well as post-translational modification. There is also an underrepresentation of enzymes involved in transcription in both the up- and down-regulated proteins.

COG category	Genome	Up	Down
RNA processing and modification	6 (0.1)	0 (0.0)	0 (0.0)
Chromatin structure and dynamics	2 (0.0)	0 (0.0)	0 (0.0)
Energy production and conversion	584 (8.0)	25 (12.8)	10 (8.7)
Cell cycle control, cell division, chromosome partitioning	47 (0.6)	0 (0.0)	1 (0.9)
Amino acid transport and metabolism	539 (7.4)	23 (11.8)	13 (11.3)
Nucleotide transport and metabolism	118 (1.6)	8 (4.1)	2 (1.7)
Carbohydrate transport and metabolism	587 (8.0)	7 (3.6)	4 (3.5)
Coenzyme transport and metabolism	303 (4.1)	13 (6.7)	8 (7.0)
Lipid transport and metabolism	448 (6.1)	16 (8.2)	12 (10.4)
Translation, ribosomal structure and biogenesis	218 (3.0)	**15 (7.7)**	8 (7.0)
Transcription	1018 (13.9)	**8 (4.1)**	**6 (5.2)**
Replication, recombination and repair	224 (3.1)	**0 (0.0)**	2 (1.7)
Cell wall/membrane/envelope biogenesis	255 (3.5)	4 (2.1)	1 (0.9)
Cell motility	5 (0.1)	0 (0.0)	0 (0.0)
Posttranslational modification, protein turnover, chaperones	149 (2.0)	**11 (5.6)**	6 (5.2)
Inorganic ion transport and metabolism	259 (3.5)	9 (4.6)	3 (2.6)
Secondary metabolites biosynthesis, transport and catabolism	397 (5.4)	13 (6.7)	7 (6.1)
General function prediction only	1117 (15.2)	20 (10.3)	19 (16.5)
Function unknown	526 (7.2)	12 (6.2)	10 (8.7)
Signal transduction mechanisms	379 (5.2)	8 (4.1)	2 (1.7)
Intracellular trafficking, secretion, and vesicular transport	44 (0.4)	2 (1.0)	1 (0.9)
Defense mechanisms	106 (1.4)	1 (0.5)	0 (0.0)
Extracellular structures	1 (0.0)	0 (0.0)	0 (0.0)
Total	**7332**	**195**	**115**

Table 5.3. COG categories for proteins differentially expressed with respect to growth in the presence of dioxane-extracted lignin at 24 h. *Comparison of COG category representation in the differentially expressed proteins with that of the entire genome. The number of open reading frames represented by each COG is given, and the percentage of total genes with COG categories is in parentheses. Values that are statistically significantly over- or under-represented based on the hypergeometric test.*

	control	DL	DHP
Number of proteins with > 10 counts	61	83	81
> 2-fold upregulation (with p-value < 0.05)	0	53 (32)	45 (31)
> 2-fold downregulation (with p-value < 0.05)	0	35 (27)	33 (24)
> 2-fold upregulation (with p-value < 0.05) of DHP vs DL	—	0	15 (7)
> 2-fold downregulation (with p-value < 0.05) of DHP vs DL	—	0	8 (8)

Table 5.4. MUDPIT statistics overview. *Number of proteins up- and downregulated over 2-fold due to growth in the presence of lignin. Proteins with p-values < 0.05 are noted within parentheses to quantify the total of statistically significantly differentially secreted proteins.*

Table 5.5. Upregulation of proteins induced by the presence of dioxane-extracted lignin (DL). *Avg = average of the triplicate samples. SEM = standard error mean. Ratio = ratio of the +DL/control. p-value is calculated by the student's T-Test with a cut-off of < 0.05 for statistical significance. Proteins with p-values > 0.05 were included at the bottom of the table to be conside-wred as of interest. Signal Sequence (SS) was predicted by SignalP 4.1. ND = not determined.*

		no lignin control		+ DL				
gene ID	**gene annotation**	**Avg**	**SEM**	**Avg**	**SEM**	**Ratio**	**p-value**	**SS?**
Amy39116DRAFT_6163	5''-nucleotidase/2'',3''-cyclic phosphodiesterase and related esterases	0.000	0.000	4.353	0.315	ND	0.000	yes
Amy39116DRAFT_1438	conserved repeat domain	0.000	0.000	2.847	0.356	ND	0.001	yes
Amy39116DRAFT_7877	Subtilisin-like serine proteases	0.000	0.000	4.019	0.687	ND	0.004	yes
Amy39116DRAFT_8034	probable F420-dependent oxidoreductase, MSMEG_2906 family	0.000	0.000	1.393	0.239	ND	0.004	no
Amy39116DRAFT_0442	ABC-type Fe3+-hydroxamate transport system, periplasmic component	0.000	0.000	1.253	0.230	ND	0.006	yes
Amy39116DRAFT_7128	ABC-type Fe3+-hydroxamate transport system, periplasmic component	0.000	0.000	2.191	0.419	ND	0.006	yes
Amy39116DRAFT_4254	aminopeptidase N, Streptomyces lividans type	0.000	0.000	0.517	0.125	ND	0.014	no
Amy39116DRAFT_5583	Beta-glucosidase-related glycosidases	0.000	0.000	0.711	0.172	ND	0.015	yes
Amy39116DRAFT_7041	phosphopyruvate hydratase	0.000	0.000	0.867	0.230	ND	0.020	no
Amy39116DRAFT_5658	ABC-type Fe3+-hydroxamate transport system, periplasmic component	0.000	0.000	1.035	0.285	ND	0.022	yes
Amy39116DRAFT_5148	Uncharacterized ABC-type transport periplasmic/surface lipoprotein	0.000	0.000	1.316	0.366	ND	0.023	yes
Amy39116DRAFT_7167	hypothetical protein	0.000	0.000	0.786	0.243	ND	0.032	yes
Amy39116DRAFT_3844	hypothetical protein	0.000	0.000	0.658	0.212	ND	0.036	yes
Amy39116DRAFT_4067	Subtilisin inhibitor-like.	0.000	0.000	5.267	1.737	ND	0.039	yes
Amy39116DRAFT_2144	malate dehydrogenase	0.000	0.000	1.093	0.368	ND	0.041	no
Amy39116DRAFT_4012	Trehalose utilisation./PKD domain./Domain of Unknown Function	0.000	0.000	0.692	0.264	ND	0.059	yes
Amy39116DRAFT_5362	ABC-type Fe3+-hydroxamate transport system, periplasmic component	0.201	0.201	18.306	3.017	91.209	0.004	yes
Amy39116DRAFT_6208	Lyzozyme M1 (1,4-beta-N-acetylmuramidase)	0.131	0.131	4.967	0.954	37.904	0.007	yes
Amy39116DRAFT_8199	Chitinase	0.138	0.138	2.083	0.600	15.094	0.034	yes
Amy39116DRAFT_3052	Alpha-lytic protease prodomain.	0.610	0.359	8.705	1.501	14.259	0.006	yes
Amy39116DRAFT_6549	Uncharacterized protein conserved in bacteria	0.532	0.355	6.980	0.170	13.109	0.000	yes
Amy39116DRAFT_5424	parallel beta-helix repeat (two copies)	0.269	0.135	3.479	0.733	12.932	0.013	yes
Amy39116DRAFT_6578	Glycine/serine hydroxymethyltransferase	0.207	0.207	2.652	0.649	12.816	0.023	no
Amy39116DRAFT_6119	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	0.570	0.167	3.798	0.292	6.663	0.001	yes
Amy39116DRAFT_2929	ABC-type xylose transport system, periplasmic component	0.138	0.138	0.761	0.188	5.512	0.056	yes
Amy39116DRAFT_2534	hypothetical protein	1.785	0.247	9.711	2.058	5.441	0.019	no
Amy39116DRAFT_3272	RHS repeat-associated core domain	2.319	0.419	12.407	2.822	5.349	0.024	yes
Amy39116DRAFT_6209	Lyzozyme M1 (1,4-beta-N-acetylmuramidase)	12.716	3.716	51.309	3.482	4.035	0.002	no
Amy39116DRAFT_3823	rhamnose ABC transporter, rhamnose-binding protein	1.043	0.102	3.683	0.245	3.533	0.001	yes
Amy39116DRAFT_7219	hypothetical protein	1.105	0.205	3.621	0.925	3.276	0.057	no
Amy39116DRAFT_6115	Phosphoenolpyruvate carboxykinase (GTP)	0.839	0.034	2.326	0.426	2.772	0.025	no
Amy39116DRAFT_6548	Beta-lactamase class C and other penicillin binding proteins	2.889	0.262	7.955	1.070	2.753	0.010	yes
p-value > 0.05								
Amy39116DRAFT_6550	Beta-glucosidase-related glycosidases	0.000	0.000	11.601	4.519	ND	0.062	yes
Amy39116DRAFT_7439	NAD-dependent aldehyde dehydrogenases	0.000	0.000	1.514	0.661	ND	0.084	no
Amy39116DRAFT_1161	conserved repeat domain	0.000	0.000	1.215	0.573	ND	0.101	yes
Amy39116DRAFT_5534	Uncharacterized protein conserved in bacteria	0.000	0.000	0.532	0.270	ND	0.120	yes
Amy39116DRAFT_7437	acetate--CoA ligase	0.000	0.000	0.741	0.394	ND	0.133	no
Amy39116DRAFT_5965	PASTA domain./Protein kinase domain.	0.000	0.000	1.760	0.984	ND	0.148	no
Amy39116DRAFT_7441	Alcohol dehydrogenase, class IV	0.301	0.301	1.492	0.600	4.956	0.151	no
Amy39116DRAFT_5918	Uncharacterized protein conserved in bacteria	0.439	0.261	1.244	0.365	2.834	0.147	yes

Table 5.6. Upregulation of proteins induced by the presence of DHP. *Avg = average of the triplicate samples. SEM = standard error mean. Ratio = ratio of the +DHP/control. p-value is calculated by the student's T-Test with a cut-off of < 0.05 for statistical significance.. Proteins with p-values > 0.05 were included at the bottom of the table to be considewred as of interest. Signal sequence (SS) was predicted by SignalP 4.1. ND = not determined.*

		no lignin control		+ DHP				
gene ID	**gene annotation**	**Avg**	**SEM**	**Avg**	**SEM**	**Ratio**	**p-value**	**SS?**
Amy39116DRAFT_1161	conserved repeat domain	0.000	0.000	0.734	0.037	ND	0.000	yes
Amy39116DRAFT_5658	ABC-type Fe3+-hydroxamate transport system, periplasmic componen	0.000	0.000	1.029	0.075	ND	0.000	yes
Amy39116DRAFT_0442	ABC-type Fe3+-hydroxamate transport system, periplasmic componen	0.000	0.000	1.587	0.151	ND	0.000	yes
Amy39116DRAFT_7439	NAD-dependent aldehyde dehydrogenases	0.000	0.000	0.461	0.052	ND	0.001	no
Amy39116DRAFT_6733	hypothetical protein	0.000	0.000	1.892	0.313	ND	0.004	yes
Amy39116DRAFT_4012	Trehalose utilisation./PKD domain./Domain of Unknown Function	0.000	0.000	1.435	0.237	ND	0.004	yes
Amy39116DRAFT_5583	Beta-glucosidase-related glycosidases	0.000	0.000	0.914	0.152	ND	0.004	yes
Amy39116DRAFT_6163	5''-nucleotidase/2'',3''-cyclic phosphodiesterase and related esterases	0.000	0.000	1.201	0.201	ND	0.004	yes
Amy39116DRAFT_7128	ABC-type Fe3+-hydroxamate transport system, periplasmic componen	0.000	0.000	1.327	0.239	ND	0.005	yes
Amy39116DRAFT_5534	Uncharacterized protein conserved in bacteria	0.000	0.000	1.200	0.236	ND	0.007	yes
Amy39116DRAFT_5965	PASTA domain./Protein kinase domain.	0.000	0.000	1.306	0.276	ND	0.009	no
Amy39116DRAFT_3844	hypothetical protein	0.000	0.000	0.614	0.131	ND	0.009	yes
Amy39116DRAFT_5148	Uncharacterized ABC-type transport periplasmic/surface lipoprotein	0.000	0.000	0.928	0.226	ND	0.015	yes
Amy39116DRAFT_1438	conserved repeat domain	0.000	0.000	4.133	1.266	ND	0.031	yes
Amy39116DRAFT_4254	aminopeptidase N, Streptomyces lividans type	0.000	0.000	1.296	0.401	ND	0.032	no
Amy39116DRAFT_4067	Subtilisin inhibitor-like	0.000	0.000	7.631	2.706	ND	0.048	yes
Amy39116DRAFT_4477	hypothetical protein	0.000	0.000	0.618	0.234	ND	0.057	no
Amy39116DRAFT_7036	Predicted periplasmic lipoprotein involved in iron transport	0.000	0.000	0.929	0.356	ND	0.059	yes
Amy39116DRAFT_5362	ABC-type Fe3+-hydroxamate transport system, periplasmic componen	0.201	0.201	14.671	3.420	73.097	0.013	yes
Amy39116DRAFT_6208	Lyzozyme M1 (1,4-beta-N-acetylmuramidase)	0.131	0.131	7.573	0.557	57.788	0.000	yes
Amy39116DRAFT_8199	Chitinase	0.138	0.138	3.329	0.410	24.126	0.002	yes
Amy39116DRAFT_5424	parallel beta-helix repeat (two copies)	0.269	0.135	4.652	1.173	17.293	0.021	yes
Amy39116DRAFT_3052	Alpha-lytic protease prodomain.	0.610	0.359	9.488	3.249	15.542	0.053	yes
Amy39116DRAFT_6549	Uncharacterized protein conserved in bacteria	0.532	0.355	7.123	0.423	13.377	0.000	yes
Amy39116DRAFT_6578	Glycine/serine hydroxymethyltransferase	0.207	0.207	1.839	0.078	8.884	0.002	no
Amy39116DRAFT_2929	ABC-type xylose transport system, periplasmic component	0.138	0.138	1.202	0.272	8.710	0.025	yes
Amy39116DRAFT_6119	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfa	0.570	0.167	3.239	0.220	5.682	0.001	yes
Amy39116DRAFT_2534	hypothetical protein	1.785	0.247	9.612	0.953	5.385	0.001	no
Amy39116DRAFT_6209	Lyzozyme M1 (1,4-beta-N-acetylmuramidase)	12.716	3.716	48.248	4.570	3.794	0.004	no
Amy39116DRAFT_3823	rhamnose ABC transporter, rhamnose-binding protein	1.043	0.102	3.742	0.292	3.590	0.001	yes
Amy39116DRAFT_6548	Beta-lactamase class C and other penicillin binding proteins	2.889	0.262	9.215	0.883	3.189	0.002	yes
p-value > 0.05								
Amy39116DRAFT_7041	phosphopyruvate hydratase	0.000	0.000	0.705	0.277	ND	0.064	no
Amy39116DRAFT_6550	Beta-glucosidase-related glycosidases	0.000	0.000	7.079	3.403	ND	0.106	yes
Amy39116DRAFT_3539	Phosphodiesterase/alkaline phosphatase D	0.000	0.000	0.798	0.412	ND	0.125	yes
Amy39116DRAFT_7877	Subtilisin-like serine proteases	0.000	0.000	0.402	0.213	ND	0.132	yes
Amy39116DRAFT_2144	malate dehydrogenase	0.000	0.000	0.421	0.234	ND	0.147	no

Table 5.7. Downregulation of proteins induced by the presence of dioxane-extracted lignin. *Avg = average of the triplicate samples. SEM = standard error mean. Ratio = ratio of the control/+DL. p-value is calculated by the student's T-Test with a cut-off of < 0.05 for statistical significance. Proteins with p-values > 0.05 were included at the bottom of the table to be considewred as of interest. Signal sequence (SS) was determined using SignalP 4.1. ND = not determined.*

		no lignin control		+ DL				
gene ID	**gene annotation**	**Avg**	**SEM**	**Avg**	**SEM**	**Ratio**	**p-value**	**SS?**
Amy39116DRAFT_5851	Ferritin-like protein	2.514	0.127	0.000	0.000	ND	0.000	no
Amy39116DRAFT_3810	hypothetical protein	4.635	0.233	0.000	0.190	ND	0.000	yes
Amy39116DRAFT_6249	Glycerophosphoryl diester phosphodiesterase	6.430	0.431	0.000	0.000	ND	0.000	yes
Amy39116DRAFT_6068	single stranded DNA-binding protein (ssb)	1.457	0.364	0.000	0.052	ND	0.040	no
Amy39116DRAFT_5013	Ferredoxin	2.120	0.206	0.089	0.089	23.690	0.001	no
Amy39116DRAFT_5303	ribosomal protein L22, bacterial type	1.519	0.762	0.087	0.446	17.415	0.026	no
Amy39116DRAFT_0666	transaldolase, mycobacterial type	3.445	0.432	0.255	0.328	13.515	0.005	no
Amy39116DRAFT_8051	Uncharacterized protein conserved in bacteria	3.457	0.783	0.302	0.426	11.452	0.000	no
Amy39116DRAFT_2986	endoribonuclease L-PSP, putative	1.622	2.231	0.168	0.447	9.674	0.035	no
Amy39116DRAFT_7616	Predicted Zn-dependent hydrolases of the beta-lactamase	14.482	0.113	1.550	0.286	9.346	0.033	yes
Amy39116DRAFT_4800	urocanate hydratase	4.864	0.547	0.573	0.000	8.491	0.015	no
Amy39116DRAFT_5876	Acetylornithine deacetylase/Succinyl-diaminopimelate des	2.459	0.550	0.304	0.145	8.086	0.020	no
Amy39116DRAFT_5308	50S ribosomal protein L3, bacterial	4.105	1.703	0.595	1.042	6.895	0.003	no
Amy39116DRAFT_7935	dihydrolipoamide dehydrogenase	4.536	0.724	0.680	0.604	6.671	0.000	no
Amy39116DRAFT_5754	Predicted aminopeptidases	2.117	0.090	0.335	0.335	6.310	0.007	yes
Amy39116DRAFT_3821	L-rhamnose isomerase, Streptomyces subtype	7.376	0.974	1.230	0.260	5.995	0.010	no
Amy39116DRAFT_6856	citrate synthase I (hexameric type)	15.621	1.323	2.683	0.235	5.821	0.005	no
Amy39116DRAFT_4463	ketol-acid reductoisomerase	5.789	3.548	1.021	1.059	5.668	0.009	no
Amy39116DRAFT_6831	phosphoserine aminotransferase, putative	6.980	1.032	1.286	0.197	5.428	0.003	no
Amy39116DRAFT_4783	Protein-disulfide isomerase	9.871	0.571	2.220	0.044	4.446	0.031	no
Amy39116DRAFT_2161	Acetylornithine deacetylase/Succinyl-diaminopimelate des	2.014	0.712	0.463	0.123	4.349	0.039	no
Amy39116DRAFT_6820	Superoxide dismutase	19.841	0.403	4.606	0.087	4.308	0.015	no
Amy39116DRAFT_4837	glutamine synthetase, type I	6.723	0.431	1.643	0.168	4.091	0.006	no
Amy39116DRAFT_4332	ribosomal protein S16	2.061	3.929	0.506	0.906	4.075	0.044	no
Amy39116DRAFT_2450	hypothetical protein	8.971	2.067	2.660	1.116	3.373	0.002	yes
Amy39116DRAFT_2887	delta-1-pyrroline-5-carboxylate dehydrogenase, group 1	15.738	0.505	4.768	0.096	3.301	0.005	no
Amy39116DRAFT_6493	Catalase	3.912	0.488	1.277	0.000	3.064	0.022	no
Amy39116DRAFT_7704	dihydrolipoamide dehydrogenase	5.638	0.522	2.047	0.122	2.754	0.000	no
p-value > 0.05								
Amy39116DRAFT_5278	ribosomal protein S4, bacterial/organelle type	2.165	0.587	0.798	0.098	2.713	0.083	no
Amy39116DRAFT_2003	proteasome, beta subunit, bacterial type	2.981	0.772	1.367	0.237	2.182	0.116	no
Amy39116DRAFT_5235	Negative regulator of beta-lactamase expression	2.954	0.943	1.095	0.203	2.698	0.072	yes

Table 5.8. Downregulation of proteins induced by the presence of DHP. *Avg = average of the triplicate samples. SEM = standard error mean. Ratio = ratio of the control/+DHP. P-value is calculated by the student's T-Test with a cut-off of < 0.05 for statistical significance. Proteins with p-values > 0.05 were included at the bottom of the table to be considewred as of interest. Signal sequence (SS) was determined using SignalP 4.1. ND = not determined.*

		no lignin control		+ DHP				
gene ID	**gene annotation**	**Avg**	**SEM**	**Avg**	**SEM**	**Ratio**	**p-value**	**SS?**
Amy39116DRAFT_5754	Predicted aminopeptidases	2.117	0.090	0.000	0.000	ND	0.000	yes
Amy39116DRAFT_5308	50S ribosomal protein L3, bacterial	4.105	0.233	0.000	0.083	ND	0.001	no
Amy39116DRAFT_4332	ribosomal protein S16	2.061	0.431	0.000	0.090	ND	0.017	no
Amy39116DRAFT_5303	ribosomal protein L22, bacterial type	1.519	0.432	0.000	0.000	ND	0.020	no
Amy39116DRAFT_5278	ribosomal protein S4, bacterial/organelle type	2.165	0.547	0.000	0.308	ND	0.021	no
Amy39116DRAFT_6068	single stranded DNA-binding protein (ssb)	1.457	0.127	0.000	0.270	ND	0.040	no
Amy39116DRAFT_3810	hypothetical protein	4.635	0.364	0.090	0.097	51.518	0.000	yes
Amy39116DRAFT_5851	Ferritin-like protein	2.514	0.206	0.270	0.077	9.315	0.002	no
Amy39116DRAFT_6249	Glycerophosphoryl diester phosphodiesterase	6.430	1.703	0.920	0.913	6.993	0.001	yes
Amy39116DRAFT_0666	transaldolase, mycobacterial type	3.445	0.113	0.500	0.548	6.887	0.006	no
Amy39116DRAFT_2986	endoribonuclease L-PSP, putative	1.622	0.550	0.319	0.073	5.092	0.050	no
Amy39116DRAFT_5876	Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase and related deacylases	2.459	0.724	0.523	0.572	4.700	0.041	no
Amy39116DRAFT_8051	Uncharacterized protein conserved in bacteria	3.457	0.783	0.741	0.387	4.663	0.000	no
Amy39116DRAFT_4800	urocanate hydratase	4.864	2.231	1.288	0.849	3.778	0.027	no
Amy39116DRAFT_4463	ketol-acid reductoisomerase	5.789	0.522	1.676	0.000	3.454	0.018	no
Amy39116DRAFT_3821	L-rhamnose isomerase, Streptomyces subtype	7.376	1.323	2.150	0.206	3.430	0.018	no
Amy39116DRAFT_4837	glutamine synthetase, type I	6.723	0.974	2.012	0.445	3.341	0.007	no
Amy39116DRAFT_5013	Ferredoxin	2.120	0.403	0.659	0.000	3.216	0.003	no
Amy39116DRAFT_6856	citrate synthase I (hexameric type)	15.621	0.587	5.055	0.000	3.090	0.011	no
Amy39116DRAFT_2887	delta-1-pyrroline-5-carboxylate dehydrogenase, group 1	15.738	1.032	5.118	0.204	3.075	0.005	no
Amy39116DRAFT_6493	Catalase	3.912	0.712	1.360	0.449	2.878	0.039	no
Amy39116DRAFT_7935	dihydrolipoamide dehydrogenase	4.536	0.488	1.710	0.000	2.653	0.002	no
Amy39116DRAFT_6831	phosphoserine aminotransferase, putative	6.980	0.571	2.783	0.315	2.508	0.009	no
Amy39116DRAFT_7704	dihydrolipoamide dehydrogenase	5.638	0.431	2.597	0.183	2.171	0.006	no
p-value > 0.05								
Amy39116DRAFT_2003	proteasome, beta subunit, bacterial type	2.981	0.772	0.969	0.172	3.076	0.064	no

Secretome analysis of *A.* sp. 75iv2. Based on the large size of the lignin substrate, we focused on examining changes in the secretome of *A.* sp. 75iv2 to explore the molecular basis of the observed physiological to lignin consisting of increased cell growth, extracellular peroxidase activity, and changes in the secreted heme profile. Towards this goal, secretome samples were collected under the same growth conditions as the RNA-seq and intracellular proteomics studies in SMM with dioxane-extracted lignin and DHP compared to a no lignin control. The soluble fraction of the culture sample was first concentrated approximately 15-fold using a 3 kDa-MWCO filter before precipitating both proteins and water-soluble lignin products with trichloroacetic acid. After further processing and tryptic digestion, the sample was then acidified with formic acid to precipitate and remove remaining lignin metabolites. The soluble mixture appeared free of contaminating lignin by UV-visible spectroscopy and was then chromatographed and analyzed by MudPIT [*29*]. A total of 457 proteins were identified in the secretome from these samples. Proteins for which no single sample contained >10 spectral counts were ignored, resulting in 83 proteins for the dioxane-extracted lignin sample, 81 proteins for the DHP sample, and 61 proteins for the no lignin control sample (*Table 5.4*). After removal of proteins that did not meet these criteria, the spectral counts for each protein were normalized based on average spectral count of the sample according to literature [*34*]. Upregulated and downregulated proteins were then filtered using a cut-off for the fold-change (>2-fold) and *p*-value calculated using the Student's t-test ($p < 0.05$).

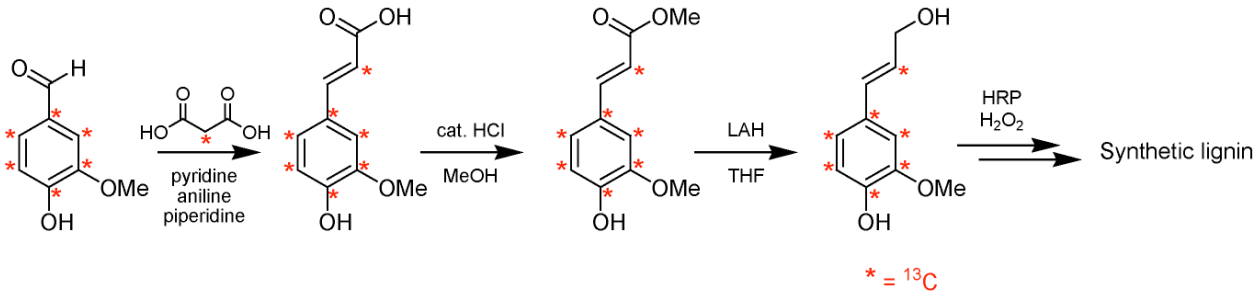
Through these studies, we found that 32 and 31 extracellular proteins were found to be statistically upregulated in the presence of dioxane-extracted lignin (*Table 5.5*), and DHP (*Table 5.6*), respectively, by these criteria. Consistent with the RNA-seq data, the periplasmic component of Fe^{3+}-hydroxamate transporters is highly upregulated in the presence of both lignin types (4 out of 10 total encoded; 70- to 90-fold for the most abundant transporter in the presence of DHP and doxane-extracted lignin, respectively; the other three transporters are only lignin-induced), one of which is the most highly upregulated proteins found in the secretome (70- to 90-fold). A predicted periplasmic lipoprotein involved in iron transport (Amy39116DRAFT_7036) is also found to be statistically significantly upregulated solely in the presence of DHP, and though upregulated in the presence of dioxane-extracted lignin, it was not considered statistically significant due to a high p-value. Of these candidates, a number of proteins with potential utility for lignocellulose degradation and utilization were identified. For carbohydrate utilization two glycosidases are found to be secreted only in the presence of lignin (Amy39116DRAFT_5583; Amy39116DRAFT_6550); similarly, the periplasmic components of sugar transporters are also found upregulated for xylose (Amy39116DRAFT_2929, 5.5- to 8.7-fold in diosane-extracted lignin and DHP, respectively) and rhamnose (Amy39116DRAFT_3823, 3.5-fold in both lignins). In addition to these proteins, a lysozyme (Amy39116DRAFT_6208, 40- to 60-fold) and a chitinase (Amy39116DRAFT_8199, ~20-fold increase) ortholog were also identified. In contrast, there were no clear oxidative or lignolytic proteins found to be upregulated under these conditions. However, genes involved in bacterial lignin modification remain cryptic and the expression seven uncharacterized or hypothetical secreted proteins were observed to be induced. Other oxidative enzymes expressed under lignin induction include NAD-dependent aldehyde dehydrogenase (Amy39116DRAFT_7439) and a malate dehydrogenase (Amy39116DRAFT_2144) both of which only expressed in the presence of either lignin.

Of the downregulated proteins (*Tables 5.7* and *5.8*), the most striking are a catalase (Amy39116DRAFT_6493, 3-fold) and two hypothetical proteins (Amy39116DRAFT_3810, 52-fold; Amy39116DRAFT_8051, 5- to 12-fold). One thing to note is that several of these

secretome proteins for cells grown without lignin include ribosomal proteins, so contamination from lysed cells may be possible.

NMR studies of dioxane-extracted and ^{13}C-labeled lignin. The characterization of structural modifications made to lignin by *A.* sp. 75iv2 could shed light on the identity of enzyme classes involved in the global metabolism of lignin. Towards this goal, we grew long-term cultures of *A.* sp. 75iv2 on dioxane-extracted lignin. After growth for 28 d, the lignin fraction was isolated by centrifugation to collect solids, including cells. After lyophilization, the solids were extracted with DMSO until the organic layer was clear. These DMSO extracts were then pooled for NMR analysis. Initial analysis by 2D-NMR seemed to implicated the loss of guaiacyl units and the associated bonding motifs (*Appendix Figure A3.4*). However, monomer analysis by GC-MS after thioacidolysis [*16*] of these samples was inconclusive (*Appendix A3.7*).

In order to facilitate these studies, we prepared coniferyl alcohol labeled on the ring and at the β-position (*Scheme 5.1*). ^{13}C-DHP was then synthesized by polymerization, precipitated to remove monomer contamination, and analyzed by GPC to give an M_p of 1.9 kDa (*Figure 5.11*). The ^{1}H/^{13}C-HSQC spectrum of the synthetic lignin gave information about the different interunit linkages present in the lignin (*Figure A3.5*). In the aliphatic region, limited resonances



Scheme 5.1 Synthesis of ^{13}C-labeled DHP. *^{13}C-labeled atoms are denoted by the red star (*)*

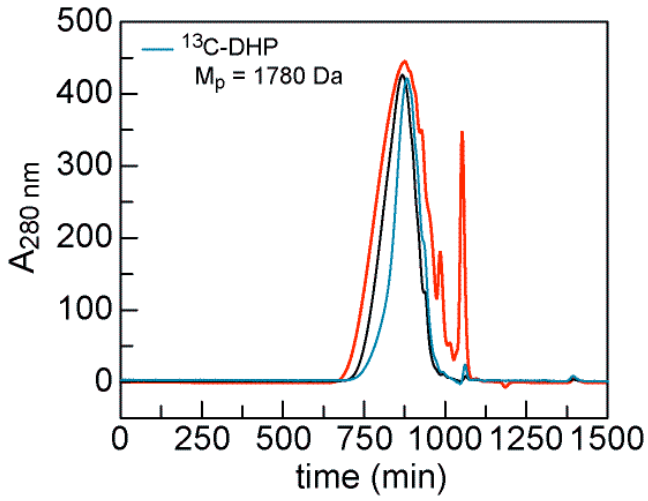


Figure 5.11. Characterization of ^{13}C-DHP by gel permeation chromatography. *Chromatograms of ^{13}C-DHP before (M_p = 1.7 kDa) and after (M_p = 1.8 kDa) ethanol precipitation compared to unlabeled DHP (Mp = 2.02 kDa), monitoring at 280 nm. (D) Comparison of approximate molecular weights of ethanol-extracted lignin (red), unlabeled DHP (black), and ^{13}C-DHP (blue) shows the M_p of ^{13}C-DHP to be similar to that of previous samples.*

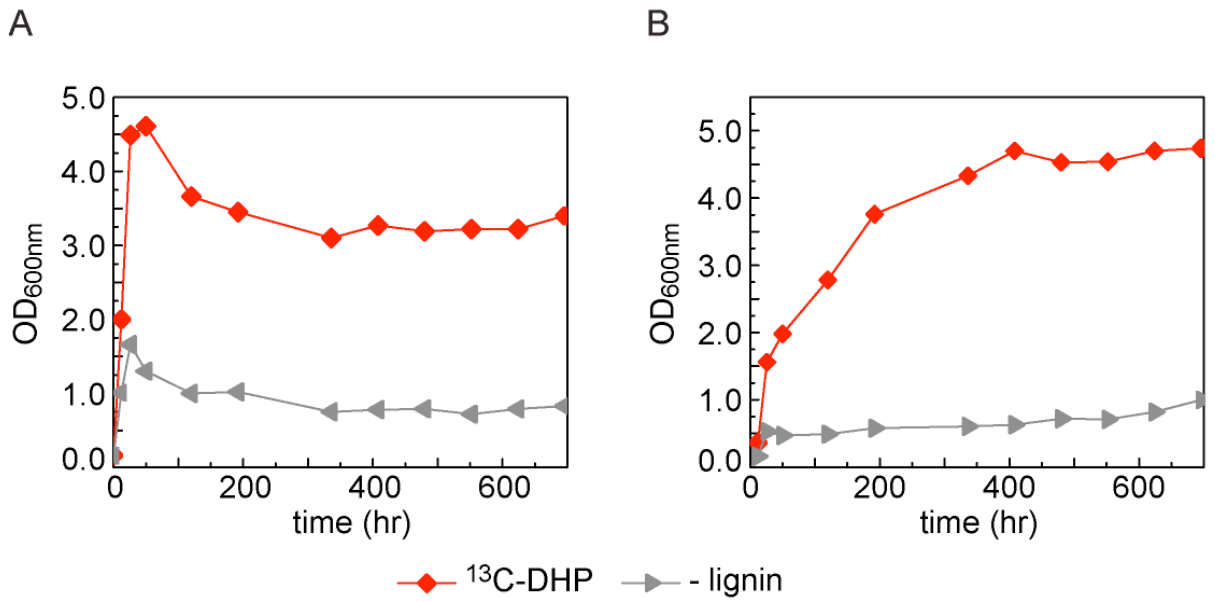


Figure 5.12. Long-term growth of A. sp. 75iv2 on ^{13}C-DHP in SMM and MM. *(A) Growth in supplemented minimal media (SMM) containing casmino acids as a carbon source monitored by OD_{600}. ^{13}C-lignin (0.05% (w/v)) was added as a second carbon source and compared to a control with no added carbon. (B) Growth in minimal media (MM) monitored by OD_{600}. ^{13}C-DHP (0.05% (w/v) was added as the sole carbon source and compared to a control with no added carbon.*

were detected, arising from the ^{13}C-labeled *β*-carbon and the unlabeled methoxyl group. The synthetic lignin was found to be rich in *β*-aryl ether (A), phenylcoumaran (B), and resinol (C) units with more modest levels of the dibenzodioxocin (D) motif. The aromatic region confirmed the ^{13}C-labeled DHP as a guaiacyl-based lignin with the presence of some cinnamyl alcohol and cinnamyl aldehyde, most likely as end groups. Growths of *A.* sp. 75iv2 on the ^{13}C-labeled DHP in SMM and MM showed the consistent lignin-dependent growth increase as previously observed with the unlabeled material (*Figure 5.12*).

Since the growth of *A.* sp. 75iv2 reaches a higher cell density in SMM, we first analyzed these samples given they likely would show greater evidence of lignin metabolism. After 28 d of growth with the ^{13}C-labeled DHP, lignin solubilization appeared to be catalyzed by *A.* sp. 75iv2 based on the darker color of the culture containing *A.* sp. 75iv2 compared to the no organism control. The culture was then centrifuged to collect the lignin and cells, lyophilized, and vortexed with DMSO-d_6 for NMR. Although the ^{1}H/^{13}C HSQC spectra appeared to support observation of increased lignin solubilization, no obvious differences in lignin structure were observed and a limited number of new resonances were found when comparing the ^{13}C-DHP incubated in the presence and absence of *A.* sp. 75iv2. More specifically, after normalization of the peak integrations to that of the DMSO solvent peak, the overall lignin content under the bacterial growth appeared to decrease by ~50-60% based on integration of several different signature peaks ranging from cinnamyl alcohol (44%) as the lowest change and G_2 and $G_{5/6}$ peaks (58%) as the highest change (*Appendix Figures A3.6* and *A3.7*). An increase in the cinnamyl aldehyde (67%) occurred due to incubation with *A.* sp. 75iv2 suggesting oxidation of cinnamyl alcohol over the 28 d period.

In the aromatic region of the supernatant samples (*Appendix Figure A3.6C-F*) we can clearly see the high intensities of the guaiacyl unit linkages, as well as three additional unidentified new peaks suggesting oxidation of the aromatic ring. These data suggest that a fraction of lignin may have been solubilized and could possibly be detected upon analysis of the supernatant. Corresponding ^{1}H/^{13}C HSQC spectral analysis of the supernatant after normalization to the solvent peak similarly reveals significant differences between the sample and the control (*Appendix Figure A3.7B*). Where there was a 1.7-fold increase in cinnamyl alcohol, there was an overall 29-fold increase in the G_2 and and a 150-fold increase in the $G_{5/6}$ peaks representing solvation of the guaiacyl substrate. Similarly, peaks representing the methoxy group only appear in the supernatant of lignin incubated with *A.* sp. 75iv2. Analysis of the control *A.* sp. 75iv2 growth in the absence of lignin reveal no peak at this location verifying the identity of this increased peak representing the methoxy group (*Appendix Figure A3.8*). The isolated solid lignin sample was also analyzed by GPC, showing that where there was an overall decrease in intensity supporting a decreased lignin content, other possible modifications include diminishment of the lower molecular weight lignins as well as a slight increase in the M_p for the sample incubated with *A.* sp. 75iv2 (*Appendix Figure A3.9*). ^{13}C-labeled DHP samples derived from *A.* sp. 75iv2 grown in MM were also analyzed, but very little change was noted in the spectrum, including the overall spectral intensity. Limited changes in the spectra were noted due to the presence of lignin, the most obvious being the presence of 6 peaks in the carbohydrate region of the spectrum, likely representing sugar motifs; however, many of the peaks were also found in the control *A.* sp. 75iv2 growth, suggesting most of these carbohydrate peaks are secreted from the organism (*Appendix Figure A3.10*).

With DHP	Activity (AU/min)	Assay volume (mL)	Volume (mL)	Total units
Unfractionated secretome	0.071	0.50	470	67
HiTrapQ	0.030	0.05	70	42
Superdex 200	0.237	0.25	14	13
Superdex 75	0.030	0.50	4	0.4
Without DHP	Activity (AU/min)	Assay volume (mL)	Volume (mL)	Total units
Unfractionated secretome	0.0004	0.50	470	0.38
HiTrapQ	0.0033	0.05	70	4.6
Superdex200	0.0044	0.25	14	0.25
Superdex 75	--	0.50	4	--

Table 5.9. Enrichment of peroxidase activity from the A. sp. 75iv2 secretome grown in the presence of DHP. *The secreted protein fraction was prepared from growth in the presence and absence of DHP after 48h. Oxidation of 2,4-dichlorophenol was used to quantify peroxidase activity. (Activity is defined as change in AU/min for 500 μL of fraction; -- represents no activity)*

	Activity (AU/min)	Assay volume (mL)	Volume (mL)	Total units
Unfractionated secretome	0.07	0.50	750	105
HiTrapQ	0.22	0.25	45	39
Superdex200	0.96	0.50	14	27
MonoQ fractions 21 and 22	0.18	0.50	10	3.5

Table 5.10. Enrichment of peroxidase activity from the A. sp. 75iv2 secretome grown in the presence of ethanol-extracted lignin. *The secreted protein fraction was prepared from growth in the presence of DHP after 48h. Oxidation of 2,4-dichlorophenol was used to quantify peroxidase activity. (Activity is defined as change in AU/min for 500 μL of fraction)*

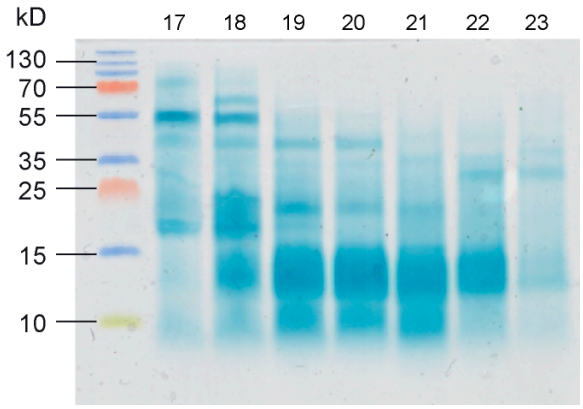


Figure 5.13. Heme stain of Enriched peroxidase activity from the A. sp. 75iv2 secretome grown in the presence of ethanol-extracted lignin. *Fractions 17-23 from final MonoQ purication represented below where peak activity resides in fractions 20-22.*

Purification of extracellular peroxidases from *A.* sp. 75iv2 grown in the presence of lignin. As the identity of the secreted peroxidases remained unclear from cell profiling studies, we turned out attention to enriching the peroxidase activity by fractioning the secretome of *A.* sp. 75iv2 grown for 48 h in the presence of either DHP or ethanol-extracted lignin as these cultures showed the highest induction level of extracellular 2,4-dichlorophenol oxidation capacity. After cell removal and concentration of the secretome, the peroxidase activity was tracked by enzymatic activity through various purification steps (*Tables 5.9* and *5.10, Appendix Figures A3.11* and *A3.12*). The first purification step consisted of anion exchange using HiTrapQ HP column (*Appendix Figures A3.11A* and *A3.12A*), followed by size-exclusion chromatograph with a Superdex 200 column (*Appendix Figures A3.11B* and *A3.12B*). With the DHP lignin samples, a final fractionation step by Superdex 75 was utilized (*Appendix Figure A3.11C*) to obtain fractions for analysis. However, these fractions remained too complex for conclusive identification of the peroxidase candidates (*Appendix Figure A3.11D*). In this case, a negative control with no lignin was also included to show that the corresponding fractions did not contain a significant level of peroxidase activity (*Table 5.9*). With the ethanol-extracted lignin, a MonoQ 10/100 column was used for the final chromatographic step (*Appendix Figure A3.12C*). Degradation of peroxidase activity was observed over 15 d fractionations; thus the time between initial concentration of the secretome and final separation and analysis by MonoQ was subsequently reduced to 3 d (*Table 5.10*). In this case, fractions were kept separate and analyzed individually rather than pooling for the next purification step. LC-MS/MS of these fractions revealed their identity to still be very complex (*Appendix Figure A3.12D)* Interestingly, the heme protein profile of these purified fractions (*Figure 5.13*) differed from that derived from the unfractionated sample (*Figure 5.4C*). Instead more than 10 heme-containing proteins were noted in the extracellular peroxidase-containing fraction (*Figure 5.13*). These bands were excised and subject to protein identification by in-gel tryptic digest followed by LC-MS/MS.

5.4. Conclusions

In minimal media conditions that more closely mimic those found in the environment, the presence of native and synthetic lignins induced a physiological response in *A.* sp. 75iv2, including increased growth, expression of extracellular peroxidases, as well as changes in the secreted heme-containing protein profile. Taken together, it appears as if A. sp. 75iv2 is capable

of assimilating and metabolizing lignin as a carbon source. Interestingly, these changes are not observed with sugar-based carbon source such as glucose, xylose, or glycerol. Cell profiling studies, including RNA-seq and proteomics studies were used to characterize the response of *A.* sp. 75iv2 and indicated a strong response with respect to Fe assimilation, potentially for insertion into heme-dependent or other oxidative enzymes. Interestingly, a large number of universal stress proteins were upregulated, which may be required to achieve a global shift in carbon metabolism. Several large gene clusters were also identified using RNA-seq that include a large number of hypothetical or uncharacterized proteins. Characterization of the secretome confirms the importance of Fe assimilation but did not yield clear candidate genes for lignin metabolism. However, further biochemical characterization and fractionation of the secretome indicates that a large number of heme-containing proteins are induced upon growth on lignin.

5.5. References

1. R. ten Have and P. J. Teunissen, Oxidative mechanisms involved in lignin degradation by white-rot fungi, *Chem. Rev.* **2001**, *101*, 3397-413.
2. T. K. Kirk and R. L. Farrell, Enzymatic "combustion": The microbial degradation of lignin, *Annu. Rev. Microbiol.* **1987**, *41*, 465-505.
3. D. L. Crawford, Lignocellulose decomposition by selected *Streptomyces* strains, *Appl. Environ. Microbiol.* **1978**, *35*, 1041-1045.
4. H. W. Kern and T. K. Kirk, Influence of molecular size and ligninase pretreatment on degradation of lignins by *Xanthomonas* sp. strain 99, *Appl. Environ. Microbiol.* **1987**, *53*, 2242-2246.
5. T. J. Kerr, R. D. Kerr and R. Benner, Isolation of a bacterium capable of degrading peanut hull lignin, *Appl. Environ. Microbiol.* **1983**, *43*, 1201-1206.
6. J. Trojanowski, K. Haider and V. Sundman, Decomposition of ^{14}C-labelled lignin and phenols by a *Nocardia* sp., *Arch. Microbiol.* **1977**, *114*, 149-153.
7. H. W. Kern, Bacterial degradation of dehydropolymers of coniferyl alcohol, *Arch. Microbiol.* **1984**, *138*, 18-25.
8. M. B. Pasti, S. R. Hagen, R. A. Korus and D. L. Crawford, The effects of various nutrients on extracellular peroxidases and acid-precipitable polymeric lignin production by *Streptomyces chromofuscus* A2 and *S. viridosporus* T7A, *Appl. Microbiol. Biotechnol.* **1991**, *34*, 661-667.
9. M. Ramachandra, D. L. Crawford and A. L. Pometto, Extracellular enzyme activities during lignocellulose degradation by *Streptomyces* spp.: A comparative study of wild-type and genetically manipulated strains, *Appl. Environ. Microbiol.* **1987**, *53*, 2754-2760.
10. M. Ramachandra, D. L. Crawford and G. Hertel, Characterization of an extracellular lignin peroxidase of the lignocellulolytic actinomycete *Streptomyces viridosporus*, *Appl. Env. Microbiol.* **1988**, *54*, 3057-3063.
11. M. Ahmad, C. R. Taylor, D. Pink, K. Burton, D. Eastwood, G. D. Bending and T. D. H. Bugg, Development of novel assays for lignin degradation: Comparative analysis of bacterial and fungal lignin degraders, *Mol. BioSyst.* **2010**, *6*, 815-821.

12. M. Ahmad, J. N. Roberts, E. M. Hardiman, R. Singh, L. D. Eltis and T. D. H. Bugg, Identification of DypB from *Rhodococcus jostii* RHA1 as a lignin peroxidase, *Biochemistry* **2011**, *50*, 5096-5107.

13. J. N. Roberts, R. Singh, J. C. Grigg, M. E. P. Murphy, T. D. H. Bugg and L. D. Eltis, Characterization of dye-decolorizing peroxidases from *Rhodococcus jostii* RHA1, *Biochemistry* **2011**, *50*, 5108-5119.

14. M. E. Brown, M. C. Walker, T. G. Nakashige, A. T. Iavarone and M. C. Y. Chang, Discovery and characterization of heme enzymes from unsequenced bacteria: Application to microbial lignin degradation, *J. Am. Chem. Soc.* **2011**, *133*, 18006-18009.

15. M. E. Brown, T. Barros and M. C. Y. Chang, Identification and characterization of a multifunctional dye peroxidase from a lignin-reactive bacterium, *ACS Chem. Biol.* **2012**, *7*, 2074-2081.

16. S. Bauer, H. Sorek, V. D. Mitchell, A. B. Ibáñez and D. E. Wemmer, Characterization of *Miscanthus giganteus* lignin isolated by ethanol organosolv process under reflux condition, *J. Agric. Food Chem.* **2012**, *60*, 8203-8212.

17. D. Gagnaire and D. Robert, A polymer model of lignin (D.H.P.) ^{13}C selectively labelled at the benzylic positions: Synthesis and NMR study, *Makromol. Chem.* **1977**, *178*, 1477-1495.

18. K. Haselwandter, O. Bobleter and D. Read, Degradation of ^{14}C-labelled lignin and dehydropolymer of coniferyl alcohol by ericoid and ectomycorrhizal fungi, *Arch. Microbiol.* **1990**, *153*, 352-354.

19. K. Haider and J. Trojanowski, Decomposition of specifically ^{14}C-labelled phenols and dehydropolymers of coniferyl alcohol as models for lignin degradation by soft and white rot fungi, *Arch. Microbiol.* **1975**, *105*, 33-41.

20. T. Kieser, K. F. Chater, M. J. Bibb, M. J. Buttner and D. A. Hopwood, Practical *Streptomyces* Genetics, The John Innes Foundation: Norwich, **2000**, vol. pp.

21. J. M. B. Macedo, L. M. F. Gottschalk and E. P. S. Bon, Lignin peroxidase and protease production by *Streptomyces viridosporus* T7A in the presence of calcium carbonate, *Appl. Biochem. Biotech.* **1999**, *79*, 735-744.

22. D. C. Yee, D. Jahng and T. K. Wood, Enhanced expression and hydrogen peroxide dependence of lignin peroxidase from *Streptomyces viridosporus* T7A, *Biotechnol. Prog.* **1996**, *12*, 40-46.

23. P. E. Thomas, D. Ryan and W. Levin, An improved staining procedure for the detection of the peroxidase activity of cytochrome P-450 on sodium dodecyl sulfate polyacrylamide gels, *Anal. Biochem.* **1976**, *75*, 168-176.

24. M. J. Barnett, C. J. Toman, R. F. Fisher and S. R. Long, A dual-genome symbiosis chip for coordinate study of signal exchange and development in a prokaryote-host interaction, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16636-16641.

25. J. R. Davis, L. A. Goodwin, T. Woyke, H. Teshima, D. Bruce, C. Detter, R. Tapia, S. Han, J. Han, S. Pitluck, M. Nolan, N. Mikhailova, M. L. Land and J. K. Sello, Genome sequence of *Amycolatopsis* sp. strain ATCC 39116, a plant biomass-degrading actinomycete, *J. Bacteriol.* **2012**, *194*, 2396-2397.

26. H. Li and R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, *Bioinformatics* **2009**, *25*, 1754-1760.

27. S. Anders and W. Huber, Differential expression analysis for sequence count data, *Genome Biol.* **2010**, *11*, 1-12.

28. S. Maere, K. Heymans and M. Kuiper, BiNGO: A cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks, *Bioinformatics* **2005**, *21*, 3448-3449.

29. H. Liu, R. G. Sadygov and J. R. Yates, A Model for random sampling and estimation of relative protein abundance in shotgun proteomics, *Anal. Chem.* **2004**, *76*, 4193-4201.

30. M. P. Washburn, D. Wolters and J. R. Yates, Large-scale analysis of the yeast proteome by multidimensional protein identification technology, *Nat. Biotechnol.* **2001**, *19*, 242-247.

31. J. K. Eng, A. L. McCormack and J. R. Yates III, An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database, *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976-989.

32. D. Cociorva, D. L. Tabb and J. R. Yates, Validation of tandem mass spectrometry database search results using DTASelect, *Curr. Protoc. Bioinformatics* **2002**, *Chapter 13*, Unit 13.4.

33. D. L. Tabb, W. H. McDonald and J. R. Yates, DTASelect and Contrast: Tools for assembling and comparing protein identifications from shotgun proteomics, *J. Proteome Res.* **2002**, *1*, 21-26.

34. B. Zhang, N. C. VerBerkmoes, M. A. Langston, E. Uberbacher, R. L. Hettich and N. F. Samatova, Detecting differential and correlated protein expression in label-free shotgun proteomics, *J. Proteome Res.* **2006**, *5*, 2909-2918.

35. M. Dauner and U. Sauer, GC-MS analysis of amino acids rapidly provides rich information for isotopomer balancing, *Biotechnol. Prog.* **2000**, *16*, 642-649.

36. D. Gibson, One-step enzymatic assembly of DNA molecules up to several hundred kilobases in size, *Nat. Meth.* **2009**, *6*, 343-345.

37. J. Ralph, K. Lundquist, G. Brunow, F. Lu, H. Kim, P. F. Schatz, J. M. Marita, R. D. Hatfield, S. A. Ralph, J. H. Christensen and W. Boerjan, Lignins: Natural polymers from oxidative coupling of 4-hydroxylphenylpropanoids, *Phytochem. Rev.* **2004**, *3*, 29-60.

38. J. Rencoret, A. Gutiérrez, L. Nieto, J. Jiménez-Barbero, C. B. Faulds, H. Kim, J. Ralph, Á. T. Martínez and J. C. del Río, Lignin composition and structure inyoung versus adult *Eucalyptus globulus* plants, *Plant Physiol.* **2011**, *155*, 667-682.

39. T. Nyström and F. C. Neidhardt, Expression and role of the universal stress protein, UspA, of *Escherichia coli* during growth arrest, *Mol. Microbiol.* **1994**, *11*, 537-544.

40. E. L. Sonnhammer, G. von Heijne and A. Krogh, A hidden Markov model for predicting transmembrane helices in protein sequences, *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **1998**, *6*, 175-182.

41. T. N. Petersen, S. Brunak, G. von Heijne and H. Nielsen, SignalP 4.0: Discriminating signal peptides from transmembrane regions, *Nat. Methods* **2011**, *8*, 785-786.

42. J. Dyrløv Bendtsen, H. Nielsen, G. von Heijne and S. Brunak, Improved prediction of signal peptides: SignalP 3.0, *J. Mol. Biol.* **2004**, *340*, 783-795.

Appendix 1: *Plasmids and oligonucleotides*

Table 1. Plamid constructs.

plasmid construct	location	primers	restriction sites	cloning method
pCWOri-HisN	294	--	--	--
pCWOri-HisN-Amyco1	476	Amyco1 F1/R100	NdeI/XbaI	ligation
pCWOri-HisN-TEV-Amyco1	539	Amyco1 F100/R100	NdeI/XbaI	ligation
pCWOri-HisN-Amyco2	479	Amyco2 F1/R1	NdeI/XbaI	ligation
pCWOri-HisN-Amyco1GK	537	Amyco1GK F2/R62	NdeI/XbaI	ligation
pCWOri-HisN-TEV-DyP2	632	AmycoDyP2 F2/R1	NdeI/HindIII	ligation
pCWOri-HisN-DyP1	445	AmycoDyP1 F1/R1	NdeI/HindIII	ligation
pCWOri-HisN-TEV-DyP3	770	AmycoDyP3 F1/R1	NdeI/XbaI	ligation
pCWOri-HisN-Amyco3	478	Amyco F3/R3	NdeI/XbaI	ligation
pCWOri-HisN-Amyco4	480	Amyco F4/R4	NdeI/XbaI	ligation
pCWOri-HisN-Amyco5	477	Amyco F5/R5	NdeI/XbaI	ligation
pCWOri-HisN-CRD1438	1268	CRD1438 F1/R1	NdeI/XbaI	Gibson
pCWOri-HisN-MBP-SIL	1269	SIL4067-F1/R1, SIL4067-MBP-F1/MBP-R1	NdeI/XbaI	Gibson
pCWOri-HisN-F420	1271	F420_8034-F1/R1	NdeI/XbaI	Gibson
pCWOri-HisN-Hypo6733	1270	hypo6733-F1/R1	NdeI/XbaI	Gibson
pCWOri-HisN-FeH1	1273	FeH1-F1/R1	NdeI/XbaI	Gibson
pCWOri-HisN-FeH2	1272	FeH2-F1/R1	NdeI/XbaI	Gibson
pCWOri-HisN-WXG1	1403	WXG1-F1/R1	NdeI/XbaI	Gibson
pCWOri-HisN-XWG2	1404	WXG2-F1/R1	NdeI/XbaI	Gibson
pCWOri-HisN-Zn1	1405	Zn1-F1/R1	NdeI/XbaI	Gibson
pCWOri-HisN-glyco1	1406	Glyco1-F1/R1	NdeI/XbaI	Gibson
pCWOri-HisN-hypo2	1406	AmycoHypo2-F1/R1	NdeI/XbaI	Gibson
pCWOri-HisN-hypo3	1408	AmycoHypo3-F1/R1	NdeI/XbaI	Gibson
pCWOri-HisN-hypo4	1409	AmycoHypo4-F1/R1	NdeI/XbaI	Gibson
pCWOri-HisN-MS1	1410	AmycoMS1-F1/R1	NdeI/XbaI	Gibson
pCWOri-HisN-Prot1	1411	AmycoProt1-F1/R1	NdeI/XbaI	Gibson

Table 2. Primers used in the construction of plasmids in Table 1.

primer	Sequence
EUB R933	gca caa gcg gtg gag cat gtg g
EUB R1387	gcc cgg gaa cgt att cac cg
U1 F	cca gca gcc gcg gta ata cg
U1 R	atc ggc tac ctt gtt acg act tc
Amyco1 F1	gagatatacatatgagcgacacccaggacaacacccccctcgagcgcg
Amyco1 R100	atatctagatcaggccaggtcgaaccggtccagttccatgaccttcgtccagg
Amyco1 F100	gcggtagtcatatggaaaacctgtattttcagggcgccatgagcgacacccaggacaacacccccctcgagcg
Amyco1GK F2	gcgagccttcatatgaaaaatcagaaccagcacaacaccag
Amyco1GK F62	ccgtgtaatctagactcgagagttcatcctacgcgcatag
Amyco2 F1	gagatatacatatgaccaagccgaccaccaacaacgtgggcatcccggtg
Amyco2 R1	atatctagatcacttgccgaacgcctcggcgaccttgtcgccgag
AmycoDyP2 F2	gagatatacatatg gaaaacctgtattttcagggt atgcctgtcgacctgtccaccacgctgtcctg
AmycoDyP2 R1	aataaagctttcacagactccgcaggaaggcaagtgacggca
AmycoDyP1 F1	gagatatacatatgaccctcgacttcagcgacacccagggcctgctcgtccgc
AmycoDyP1 R1	aataaagctttcacgccgccaggtaccgcagtgccgccac
AmycoDyP3 F1	gcggtagtcatatggaaaacctgtattttcagggcgccatggcgccggaggtcgtgccgttccacg
AmycoDyP3 R1	atatctagatcaggcgaacagcgactcgccccagtagccg
Amyco F3	gagatatacatatggccgatcaagaaaacaccacccggacatatcgcaccagcc
Amyco R3	atatctagatcatgtccgcacctcccacgcggcgaacg
Amyco F4	gagatatacatatgaccgtcgtggacgaggccgcgctcctg
Amyco R4	atatctagatcacagcagggccagccaggaccgcg
Amyco F5	gagatatacatatgaactggctgaaactgctcaacgacgcggtgctcgggctg
Amyco R5	atatctagatcaaccacggtcgtcctccgggaagaaccgctgcc
PE F1	aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatct
PE R1	caagcagaagacggcatacgagctcttccgatct
CRD1438 F1	acggtgcatccggcggcggtagtcatatggcggacgcgccggtcggctgtggttac
CRD1438 R1	gataagcttgggctgcaggtcgactctagatcacgtcggtcgtgacctcgttggaggtgtcgttggt
SIL4067-F1	ggaaaacctgtattttcagggcgccgccagcctgatgctctccctgcagga
SIL4067-R1	gataagcttgggctgcaggtcgactctagatcagaattcgaagaccccgccggactcg
SIL4067-MBP-F1	cacggtgcatccggcggcggtagtcatatgaaaa
SIL4067-MBP-R1	cctgcagggagagcatcaggctggcggcgccctgaaaatacaggttttcctcgatcccg
F420_8034-F1	acggtgcatccggcggcggtagtcatatgaccgacaccaccaggccgatccggatcg
F420_8034-R1	gataagcttgggctgcaggtcgactctagatcagccgttctgcgcgtcgcgccag
hypo6733-F1	acggtgcatccggcggcggtagtcatatgctgaacggcgacggggacaacggcc
hypo6733-R1	gataagcttgggctgcaggtcgactctagatcaggaacccttgacctgcgagcccttgt
FeH1-F1	acggtgcatccggcggcggtagtcatatgtcgtcgacgcagtcggcgtcctcagc
FeH1-R1	gataagcttgggctgcaggtcgactctagatcaacccaacttagccaacccgggcttcaacttc
FeH2-F1	acggtgcatccggcggcggtagtcatatggccgtggccgtcgacaactgcggcc
FeH2-R1	gataagcttgggctgcaggtcgactctagatcacccgccgaacccgaacccgcg
WXG1-F1	acggtgcatccggcggcggtagtcatatgtcgggcccgatgaaggtcgattacg
WXG1-R1	gataagcttgggctgcaggtcgactctagatcagccgccgaacatcttggtgatgc
WXG2-F1	acggtgcatccggcggcggtagtcatatggcaggcggctttgaaggggatccg
WXG2-R1	gataagcttgggctgcaggtcgactctagatcacaggccctcgagagcggcctgg
Zn1-F1	acggtgcatccggcggcggtagtcatatgaacgtcgtagaggactacaccggccacgtgga
Zn1-R1	gataagcttgggctgcaggtcgactctagatcaccaaccgcgctcgcgccactc
Glyco1-F1	acggtgcatccggcggcggtagtcatatggacgaggccgagcagcagccccc
Glyco1-R1	gataagcttgggctgcaggtcgactctagatcagcacgcgtgcttggcgcgcag
AmycoHypo2-F1	acggtgcatccggcggcggtagtcatatgacgaccgatcgcaatgttctgggcg
AmycoHypo2-R1	gataagcttgggctgcaggtcgactctagatcaggtcagcgaactagggtccgcagg

AmycoHypo3-F1	acggtgcatccggcggcggtagtcatatggacgaggacggcgtgaagccggc
AmycoHypo3-R1	gataagcttgggctgcaggtcgactctagatcacgtgatgttcacgatcgtgccgagcgg
AmycoHypo4-F1	acggtgcatccggcggcggtagtcatatggcgccggcgccgagcaccacgac
AmycoHypo4-R1	gataagcttgggctgcaggtcgactctagatcagcccttccagaccaccggcgaatcctcgtacg
AmycoMS1-F1	acggtgcatccggcggcggtagtcatatgaccgaggagcagacgcgcgaggg
AmycoMS1-R1	gataagcttgggctgcaggtcgactctagattagacggtgaagcccagcgcccggagc
AmycoProt1-F1	acggtgcatccggcggcggtagtcatatgcttcggatccgccgtgaactcgtcg
AmycoProt1-R1	gataagcttgggctgcaggtcgactctagattattccgtgatctccacgggctcttcggtgacg

Appendix 2: *Gene synthesis of Amyco1GK*

Table 1. List of oligonucleotides used for Amyco1GK gene assembly

primer	Sequence
Amyco1GK R2	tgatttttcatatgaaggctcgc
Amyco1GK F2	gcgagccttcatatgaaaaatcagaaccagcacaacaccag
Amyco1GK R3	aaccgtggtaaggacatttgctggtgttgtgctggttc
Amyco1GK F3	caaatgtccttaccacggttctgtcaccagctacaactcc
Amyco1GK R4	tttgttggtcgtgcgattggagttgtagctggtgacag
Amyco1GK F4	aatcgcacgaccaacaaagattggtggccgaaccaa
Amyco1GK R5	tgcaggatggacaggttcagttggttcggccaccaatc
Amyco1GK F5	ctgaacctgtccatcctgcaccagcacgaccgtaaaacta
Amyco1GK R6	attcctcatcatgcgggttagttttacggtcgtgctgg
Amyco1GK F6	acccgcatgatgaggaatttaactacgcagaagaatttcaga
Amyco1GK R7	agagcccagtaatccagtttctgaaattcttctgcgtagttaa
Amyco1GK F7	aactggattactgggctctgaaagaagacctgcgtaaact
Amyco1GK R8	tcctggctttccgtcatcagtttacgcaggtcttctttc
Amyco1GK F8	gatgacggaaagccaggactggtggccagcggatta
Amyco1GK R9	gcgggccgtagtgaccataatccgctggccaccag
Amyco1GK F9	tggtcactacggcccgctgtttatccgcatggcttgg
Amyco1GK R10	taggtgcccgcgctgtgccaagccatgcggataaaca
Amyco1GK F10	cacagcgcgggcacctaccgtatcggtgacggt
Amyco1GK R11	tagaaccaccaccgcgaccgtcaccgatacgg
Amyco1GK F11	cgcggtggtggttctactggcacgcagcgttt
Amyco1GK R12	agagttcagcggggcgaaacgctgcgtgccag
Amyco1GK F12	cgccccgctgaactcttggccggacaacgcaa
Amyco1GK R13	cggcgagctttatccaggtttgcgttgtccggcca
Amyco1GK F13	acctggataaagctcgccgcctgctgtggccaattaag
Amyco1GK R14	gctgattttgttaccgtacttttttcttaattggccacagcagg
Amyco1GK F14	aaaaagtacggtaacaaaatcagctgggccgacctgatga
Amyco1GK R15	acgttaccggccagaatcatcaggtcggccca
Amyco1GK F15	ttctggccggtaacgtggcaatcgaatctatgggc
Amyco1GK R16	ccaaaaccaatggtcttgccgcccatagattcgattgcc
Amyco1GK F16	ggcaagaccattggttttggcggcggtcgtgaagac
Amyco1GK R17	cctccgggtgccatacgtcttcacgaccgccg
Amyco1GK F17	gtatggcacccggaggaagacatctattggggcgc
Amyco1GK R18	ggccagccattctttttcagcgccccaatagatgtctt
Amyco1GK F18	tgaaaaagaatggctggcctctgaacgttattccggtgac
Amyco1GK R19	gggttttccagctcgcggtcaccggaataacgttcaga
Amyco1GK F19	cgcgagctggaaaacccgctggctgcagtacaaat
Amyco1GK R20	cgggttaacgtagatcagacccatttgtactgcagccagc
Amyco1GK F20	gggtctgatctacgttaacccggagggcccggacggta
Amyco1GK R21	gctacagggtctggcttaccgtccgggccctc
Amyco1GK F21	agccagaccctgtagcggcagcacgtgacatc
Amyco1GK R22	cgacggaaagtttcgcggatgtcacgtgctgcc
Amyco1GK F22	cgcgaaactttccgtcgcatgggtatgaacgatgaagaa
Amyco1GK R23	cgatcagcgctacggtttcttcatcgttcatacccatg
Amyco1GK F23	accgtagcgctgatcgctggcggtcacacctt
Amyco1GK R24	gccgtgtgctttgccgaaggtgtgaccgccag
Amyco1GK F24	cggcaaagcacacggcgctggcccggcatctc
Amyco1GK R25	ggctctgggcctacatgagatgccgggccagc

Amyco1GK F25	atgtaggcccagagccagaagcggcaccgatc
Amyco1GK R26	cccagaccctgtgcttcgatcggtgccgcttct
Amyco1GK F26	gaagcacagggtctgggttggatctcttcctacggtaa
Amyco1GK R27	tgtcacggcctttacctttaccgtaggaagagatccaa
Amyco1GK F27	aggtaaaggccgtgacaccattacctctggtatcgaagg
Amyco1GK R28	ggtcggcgtccatgcgccttcgataccagaggtaatgg
Amyco1GK F28	cgcatggacgccgaccccgacccagtgggataa
Amyco1GK R29	agcagacggaagtagctgttatcccactgggtcgg
Amyco1GK F29	cagctacttccgtctgctgttcgaatacgagtggaaact
Amyco1GK R30	ccgctgggcttttagtcagtttccactcgtattcgaac
Amyco1GK F30	gactaaaagcccagcgggtgcatatcagtgggaagc
Amyco1GK R31	ttcgcgcaggttcaccgcttcccactgatatgcac
Amyco1GK F31	ggtgaacctgcgcgaagaagacctggctccgg
Amyco1GK R32	acgttcggatcttcagcatccggagccaggtcttc
Amyco1GK F32	atgctgaagatccgaacgtgaaggttcctcctatgatgatg
Amyco1GK R33	gcgccaggtcggtggtcatcatcataggaggaaccttc
Amyco1GK F33	accaccgacctggcgctgcgtttcgatccagaatac
Amyco1GK R34	aacgacgagcgattttctcgtattctggatcgaaacgca
Amyco1GK F34	gagaaaatcgctcgtcgtttctacgagaatccggagga
Amyco1GK R35	gcaaatgcgtcagcgaactcctccggattctcgtaga
Amyco1GK F35	gttcgctgacgcatttgcacgcgcatggttcaaac
Amyco1GK R36	ccatatcgcggtgagtcagtttgaaccatgcgcgt
Amyco1GK F36	tgactcaccgcgatatgggcccgaagacgcgtta
Amyco1GK R37	cacttccgggcccagataacgcgtcttcgggc
Amyco1GK F37	tctgggcccggaagtgcctaaagaagactttatctggca
Amyco1GK R38	cagtcgggattgggtcttgccagataaagtcttctttagg
Amyco1GK F38	agacccaatcccgactgtagactatgagctgagcgatg
Amyco1GK R39	tgctttaatttcttcgatctcggcatcgctcagctcatagtcta
Amyco1GK F39	ccgagatcgaagaaattaaagcaaaaattctgaacagcggtct
Amyco1GK R40	accagttcggaaacggtcagaccgctgttcagaatttt
Amyco1GK F40	gaccgtttccgaactggttaagaccgcttgggcg
Amyco1GK R41	tacggaacgtagatgcagacgcccaagcggtctta
Amyco1GK F41	tctgcatctacgttccgtaattccgataaacgcggc
Amyco1GK R42	cgtgcaccatttgcaccgccgcgtttatcggaat
Amyco1GK F42	ggtgcaaatggtgcacgtatccgcctggccccg
Amyco1GK R43	cgttgacttcccaatctttctgcggggccaggcggata
Amyco1GK F43	cagaaagattgggaagtcaacgagccggagcgtctggc
Amyco1GK R44	catagacgctcagcactttcgccagacgctccggct
Amyco1GK F44	gaaagtgctgagcgtctatgaagacattcagcgtgaactg
Amyco1GK R45	gcgatggatacctttttcggcagttcacgctgaatgtctt
Amyco1GK F45	ccgaaaaaggtatccatcgcagacctgatcgtactgggt
Amyco1GK R46	tctaccgcagcagaaccacccagtacgatcaggtct
Amyco1GK F46	ggttctgctgcggtagaaaaagctgctcgcgatg
Amyco1GK R47	ggaactttaacgtcaaaaccagcatcgcgagcagctttt
Amyco1GK F47	ctggttttgacgttaaagttccgttcattccgggtcgcg
Amyco1GK R48	tctgttcttgagttgcatcaccgcgacccggaatgaac
Amyco1GK F48	gtgatgcaactcaagaacagaccgatgtggaatctttctctg
Amyco1GK R49	tcagcaaacggttccagtacagagaaagattccacatcgg
Amyco1GK F49	tactggaaccgtttgctgacggctttcgtaactatcaaaaga
Amyco1GK R50	gccgacgctgtactctttcttttgatagttacgaaagccg
Amyco1GK F50	aagagtacagcgtcggccctgaggaactgctgattg
Amyco1GK R51	ccagcagctgagctttgtcaatcagcagttcctcagg
Amyco1GK F51	acaaagctcagctgctgggcctgacggcgccaga
Amyco1GK R52	ccaccaggacggtcatttctggcgccgtcaggc

Amyco1GK F52	aatgaccgtcctggtgggtggcctgcgtgtcct
Amyco1GK R53	acggtagttagcgcccaggacacgcaggccac
Amyco1GK F53	gggcgctaactaccgtgatctgccgcacggtg
Amyco1GK R54	atgcggtcggtgaacacaccgtgcggcagatc
Amyco1GK F54	tgttcaccgaccgcattggtgtgctgactaacga
Amyco1GK R55	tgtctaccaggttaacaaagaagtcgttagtcagcacacca
Amyco1GK F55	cttctttgttaacctggtagacatgaactacgaatgggttccg
Amyco1GK R56	tcataaataccgccttcggtcggaacccattcgtagttca
Amyco1GK F56	accgaaggcggtatttatgaaattcgcgaccgtcaga
Amyco1GK R57	ccaacgtacttcgccggtctgacggtcgcgaatt
Amyco1GK F57	ccggcgaagtacgttggactgcaacccgtgttga
Amyco1GK R58	gctattagcaccgaaaatcaggtcaacacgggttgcagt
Amyco1GK F58	cctgattttcggtgctaatagcattctgcgttcctacgca
Amyco1GK R59	tcatcctgtgcgtagaactctgcgtaggaacgcagaat
Amyco1GK F59	gagttctacgcacaggatgacaaccgtgaaaaattcgttcg
Amyco1GK R60	cccacgcattgataaagtcacgaacgaatttttcacggttg
Amyco1GK F60	tgactttatcaatgcgtgggttaaagtcatgaacgctgacc
Amyco1GK R61	tttcaggtggatgtcgaaacggtcagcgttcatgactttaa
Amyco1GK F61	gtttcgacatccacctgaaacaggcaaaagaatccgtga
Amyco1GK R62	tctcgagtctagattacacggtcacggattcttttgcctg
Amyco1GK F62	ccgtgtaatctagactcgagagttcatcctacgcgcatag
Amyco1GK F63	ctatgcgcgtaggatgaac

Figure 1. Primer map for construction of synthetic gene, Amyco1GK, encoding an Amyco1 ortholog from *Geobacillus kaustophilus*. Pairs of primers are shown in alternating gray.

```
5′- gcgagccttC ATATGAAAAA TCAGAACCAG CACAACACCA GCAAATGTCC
3′- cgctcggaaG TATACTTTTT AGTCTTGGTC GTGTTGTGGT CGTTTACAGG

    TTACCACGGT TCTGTCACCA GCTACAACTC CAATCGCACG ACCAACAAAG
    AATGGTGCCA AGACAGTGGT CGATGTTGAG GTTAGCGTGC TGGTTGTTTC

    ATTGGTGGCC GAACCAACTG AACCTGTCCA TCCTGCACCA GCACGACCGT
    TAACCACCGG CTTGGTTGAC TTGGACAGGT AGGACGTGGT CGTGCTGGCA

    AAAACTAACC CGCATGATGA GGAATTTAAC TACGCAGAAG AATTTCAGAA
    TTTTGATTGG GCGTACTACT CCTTAAATTG ATGCGTCTTC TTAAAGTCTT

    ACTGGATTAC TGGGCTCTGA AAGAAGACCT GCGTAAACTG ATGACGGAAA
    TGACCTAATG ACCCGAGACT TTCTTCTGGA CGCATTTGAC TACTGCCTTT

    GCCAGGACTG GTGGCCAGCG GATTATGGTC ACTACGGCCC GCTGTTTATC
    CGGTCCTGAC CACCGGTCGC CTAATACCAG TGATGCCGGG CGACAAATAG

    CGCATGGCTT GGCACAGCGC GGGCACCTAC CGTATCGGTG ACGGTCGCGG
    GCGTACCGAA CCGTGTCGCG CCCGTGGATG GCATAGCCAC TGCCAGCGCC

    TGGTGGTTCT ACTGGCACGC AGCGTTTCGC CCCGCTGAAC TCTTGGCCGG
    ACCACCAAGA TGACCGTGCG TCGCAAAGCG GGGCGACTTG AGAACCGGCC

    ACAACGCAAA CCTGGATAAA GCTCGCCGCC TGCTGTGGCC AATTAAGAAA
    TGTTGCGTTT GGACCTATTT CGAGCGGCGG ACGACACCGG TTAATTCTTT

    AAGTACGGTA ACAAAATCAG CTGGGCCGAC CTGATGATTC TGGCCGGTAA
    TTCATGCCAT TGTTTTAGTC GACCCGGCTG GACTACTAAG ACCGGCCATT

    CGTGGCAATC GAATCTATGG GCGGCAAGAC CATTGGTTTT GGCGGCGGTC
    GCACCGTTAG CTTAGATACC CGCCGTTCTG GTAACCAAAA CCGCCGCCAG

    GTGAAGACGT ATGGCACCCG GAGGAAGACA TCTATTGGGG CGCTGAAAAA
    CACTTCTGCA TACCGTGGGC CTCCTTCTGT AGATAACCCC GCGACTTTTT

    GAATGGCTGG CCTCTGAACG TTATTCCGGT GACCGCGAGC TGGAAAACCC
    CTTACCGACC GGAGACTTGC AATAAGGCCA CTGGCGCTCG ACCTTTTGGG

    GCTGGCTGCA GTACAAATGG GTCTGATCTA CGTTAACCCG GAGGGCCCGG
    CGACCGACGT CATGTTTACC CAGACTAGAT GCAATTGGGC CTCCCGGGCC

    ACGGTAAGCC AGACCCTGTA GCGGCAGCAC GTGACATCCG CGAAACTTTC
    TGCCATTCGG TCTGGGACAT CGCCGTCGTG CACTGTAGGC GCTTTGAAAG

    CGTCGCATGG GTATGAACGA TGAAGAAACC GTAGCGCTGA TCGCTGGCGG
    GCAGCGTACC CATACTTGCT ACTTCTTTGG CATCGCGACT AGCGACCGCC

    TCACACCTTC GGCAAAGCAC ACGGCGCTGG CCCGGCATCT CATGTAGGCC
```

AGTGTGGAAG CCGTTTCGTG TGCCGCGACC GGGCCGTAGA GTACATCCGG

CAGAGCCAGA AGCGGCACCG ATCGAAGCAC AGGGTCTGGG TTGGATCTCT
GTCTCGGTCT TCGCCGTGGC TAGCTTCGTG TCCCAGACCC AACCTAGAGA

TCCTACGGTA AAGGTAAAGG CCGTGACACC ATTACCTCTG GTATCGAAGG
AGGATGCCAT TTCCATTTCC GGCACTGTGG TAATGGAGAC CATAGCTTCC

CGCATGGACG CCGACCCCGA CCCAGTGGGA TAACAGCTAC TTCCGTCTGC
GCGTACCTGC GGCTGGGGCT GGGTCACCCT ATTGTCGATG AAGGCAGACG

TGTTCGAATA CGAGTGGAAA CTGACTAAAA GCCCAGCGGG TGCATATCAG
ACAAGCTTAT GCTCACCTTT GACTGATTTT CGGGTCGCCC ACGTATAGTC

TGGGAAGCGG TGAACCTGCG CGAAGAAGAC CTGGCTCCGG ATGCTGAAGA
ACCCTTCGCC ACTTGGACGC GCTTCTTCTG GACCGAGGCC TACGACTTCT

TCCGAACGTG AAGGTTCCTC CTATGATGAT GACCACCGAC CTGGCGCTGC
AGGCTTGCAC TTCCAAGGAG GATACTACTA CTGGTGGCTG GACCGCGACG

GTTTCGATCC AGAATACGAG AAAATCGCTC GTCGTTTCTA CGAGAATCCG
CAAAGCTAGG TCTTATGCTC TTTTAGCGAG CAGCAAAGAT GCTCTTAGGC

GAGGAGTTCG CTGACGCATT TGCACGCGCA TGGTTCAAAC TGACTCACCG
CTCCTCAAGC GACTGCGTAA ACGTGCGCGT ACCAAGTTTG ACTGAGTGGC

CGATATGGGC CCGAAGACGC GTTATCTGGG CCCGGAAGTG CCTAAAGAAG
GCTATACCCG GGCTTCTGCG CAATAGACCC GGGCCTTCAC GGATTTCTTC

ACTTTATCTG GCAAGACCCA ATCCCGACTG TAGACTATGA GCTGAGCGAT
TGAAATAGAC CGTTCTGGGT TAGGGCTGAC ATCTGATACT CGACTCGCTA

GCCGAGATCG AAGAAATTAA AGCAAAAATT CTGAACAGCG GTCTGACCGT
CGGCTCTAGC TTCTTTAATT TCGTTTTTAA GACTTGTCGC CAGACTGGCA

TTCCGAACTG GTTAAGACCG CTTGGGCGTC TGCATCTACG TTCCGTAATT
AAGGCTTGAC CAATTCTGGC GAACCCGCAG ACGTAGATGC AAGGCATTAA

CCGATAAACG CGGCGGTGCA AATGGTGCAC GTATCCGCCT GGCCCCGCAG
GGCTATTTGC GCCGCCACGT TTACCACGTG CATAGGCGGA CCGGGGCGTC

AAAGATTGGG AAGTCAACGA GCCGGAGCGT CTGGCGAAAG TGCTGAGCGT
TTTCTAACCC TTCAGTTGCT CGGCCTCGCA GACCGCTTTC ACGACTCGCA

CTATGAAGAC ATTCAGCGTG AACTGCCGAA AAAGGTATCC ATCGCAGACC
GATACTTCTG TAAGTCGCAC TTGACGGCTT TTTCCATAGG TAGCGTCTGG

TGATCGTACT GGGTGGTTCT GCTGCGGTAG AAAAAGCTGC TCGCGATGCT
ACTAGCATGA CCCACCAAGA CGACGCCATC TTTTTCGACG AGCGCTACGA

GGTTTTGACG TTAAAGTTCC GTTCATTCCG GGTCGCGGTG ATGCAACTCA
CCAAAACTGC AATTTCAAGG CAAGTAAGGC CCAGCGCCAC TACGTTGAGT

AGAACAGACC GATGTGGAAT CTTTCTCTGT ACTGGAACCG TTTGCTGACG

TCTTGTCTGG CTACACCTTA GAAAGAGACA TGACCTTGGC AAACGACTGC

GCTTTCGTAA CTATCAAAAG AAAGAGTACA GCGTCGGCCC TGAGGAACTG
CGAAAGCATT GATAGTTTTC TTTCTCATGT CGCAGCCGGG ACTCCTTGAC

CTGATTGACA AAGCTCAGCT GCTGGGCCTG ACGGCGCCAG AAATGACCGT
GACTAACTGT TTCGAGTCGA CGACCCGGAC TGCCGCGGTC TTTACTGGCA

CCTGGTGGGT GGCCTGCGTG TCCTGGGCGC TAACTACCGT GATCTGCCGC
GGACCACCCA CCGGACGCAC AGGACCCGCG ATTGATGGCA CTAGACGGCG

ACGGTGTGTT CACCGACCGC ATTGGTGTGC TGACTAACGA CTTCTTTGTT
TGCCACACAA GTGGCTGGCG TAACCACACG ACTGATTGCT GAAGAAACAA

AACCTGGTAG ACATGAACTA CGAATGGGTT CCGACCGAAG GCGGTATTTA
TTGGACCATC TGTACTTGAT GCTTACCCAA GGCTGGCTTC CGCCATAAAT

TGAAATTCGC GACCGTCAGA CCGGCGAAGT ACGTTGGACT GCAACCCGTG
ACTTTAAGCG CTGGCAGTCT GGCCGCTTCA TGCAACCTGA CGTTGGGCAC

TTGACCTGAT TTTCGGTGCT AATAGCATTC TGCGTTCCTA CGCAGAGTTC
AACTGGACTA AAAGCCACGA TTATCGTAAG ACGCAAGGAT GCGTCTCAAG

TACGCACAGG ATGACAACCG TGAAAAATTC GTTCGTGACT TTATCAATGC
ATGCGTGTCC TACTGTTGGC ACTTTTTAAG CAAGCACTGA AATAGTTACG

GTGGGTTAAA GTCATGAACG CTGACCGTTT CGACATCCAC CTGAAACAGG
CACCCAATTT CAGTACTTGC GACTGGCAAA GCTGTAGGTG GACTTTGTCC

CAAAAGAATC CGTGACCGTG TAATCTAGAC TCGAGagttc atcctacgcg
GTTTTCTTAG GCACTGGCAC ATTAGATCTG AGCTCtcaag taggatgcgc

catag
gtatc

Appendix 3: *Supplementary materials for Chapter 5*

Table of Contents

Supplementary methods

Identification of peroxidase candidates for validation. Candidates for peroxidase identification were chosen based on mass spectrometry data comparing the purification of secretomes from the large-scale growths (see above) in which protein fragments chosen were identified in the spectrometry data in significantly higher quantity in the growths in the presence of DHP as compared to the growths in the absence of DHP. Proteins chosen as candidates were required to have peptides in at least the last round of purification, and only proteins with masses less than 40 kDa were considered. From this, the following genes were cloned: Amy39116DRAFT_5265 (WXG100 family type VII secretion target, "WXG1"), Amy39116DRAFT_5266 (WXG100 family type VII secretion target, "WXG2"), Amy39116DRAFT_3005 (Zn-dependent hydrolase, "Zn1"), Amy39116DRAFT_5583 (Beta-glucosidase-related glycosidase, "glyco1"), Amy39116DRAFT_7326 (hypothetical, "Amycohypo2"), Amy39116DRAFT_6540 (hypothetical, "Amycohypo3"), Amy39116DRAFT_1157 (hypothetical, "Amycohypo4"), Amy39116DRAFT_0703 (Predicted metal-sulfur cluster biosynthetic enzyme, "MS1"), Amy39116DRAFT_4561 (Predicted metal-dependent protease, "Prot1").

Screening of gene candidates from native purification. TB (50 mL) containing carbenicillin (50 μg/mL) in a 250 mL-baffled flask was inoculated to OD_{600} = 0.05 with an overnight culture of freshly-transformed *E. coli* BL21(de3) grown in TB with the plasmic of interest. The cultures were grown at 37 °C at 200 rpm in a rotary shaker until OD_{600} = 0.7, at which α-aminolevulinic acid (65 μg/mL) was added followed by adding IPTG (0.5 mM). After inducing protein expression, the culture was allowed to incubate for an additional 20 h at 20 °C. Cell pellets (of 4 mL of culture) were harvested by centrifugation at 20,817 × *g* for 5 min. Cell pellets were resuspended in 1 mL of Buffer (50 mM potassium phosphate, 50 mM sodium chloride, pH 7.5) supplemented with PMSF (0.5 mM) and lysozyme (1 mg/mL). This was stored on ice for 1 hour after which the cells were lysed by sonication while on ice (power level 5; 6 cycles; 10 sec on, 30 sec off). Insoluble lysate was removed by centrifugation (20,817 × *g*, 30 min, 4 °C) after which the soluble lysate (148 μL) was analyzed for peroxidase activity with the = 2,4-DCP assay (AU/min = change in absorbance (AU) at 510 nm/min) under standard conditions on the SpectraMax M2 96-well plate reader (Molecular Devices; Sunnyvale, CA).

Gene sequence ID	Gene annotation	Activity (AU/min)
Amy39116DRAFT_5265	WXG100 family type VII secretion target, WXG1	(<0.001)
Amy39116DRAFT_5266	WXG100 family type VII secretion target, WXG2	(<0.001)
Amy39116DRAFT_3005	Zn-dependent hydrolase, Zn1	(<0.001)
Amy39116DRAFT_5583	glycosidase, glyco1	(<0.001)
Amy39116DRAFT_7326	hypothetical, Amycohypo2	(<0.001)
Amy39116DRAFT_6540	hypothetical, Amycohypo3	(<0.001)
Amy39116DRAFT_1157	hypothetical, Amycohypo4	(<0.001)
Amy39116DRAFT_0703	metal-sulfur cluster biosynthetic enzyme, MS1	0.002
Amy39116DRAFT_4561	metal-dependent protease, Prot1	(<0.001)

Purification of His_{10}-FeH1. TB (500 mL) containing carbenicillin (50 μg/mL) in a 2 L-baffled flask was inoculated to OD_{600} = 0.05 with an overnight culture of freshly-transformed *E.*

coli BL21(de3) grown in TB. The cultures were grown at 37 °C at 200 rpm in a rotary shaker until OD_{600} = 0.7, at which time the culture was cooled on ice for 15 min before adding IPTG (0.5 mM) and α-aminolevulinic acid (65 μg/mL). After inducing protein expression, the culture was allowed to incubate for an additional 20 h at 25 °C. The cell pellet was harvested by centrifugation at 9,800 × *g* for 7 min and stored at -80 °C. The frozen cell pellet was thawed and resuspended at 5 mL/g cell paste with Buffer A (50 mM potassium phosphate, 300 mM sodium chloride, 10 mM imidazole, pH 8.0) supplemented with PMSF (0.5 mM), DTT (1 mM), and DNase (2 U/g cell paste). The cell paste was homogenized before lysis by passage through a French Pressure cell (Thermo Scientific; Waltham, MA) at 14,000 psi. The lysate was centrifuged at 15,300 × *g* for 30 min at 4 °C to separate the soluble and insoluble fractions. The soluble lysate was passed over a Ni-NTA agarose column (Qiagen, 2 mL) by gravity. The protein was isolated by elution with 10 column volumes of Buffer B (50 mM potassium phosphate, 300 mM sodium chloride, 250 mM imidazole, 1 mM DTT, pH 8.0) after washing with Buffer A (30 col vol). Fractions were pooled based on the A_{280} elution profile and concentrated with an Amicon filtration device using a YM10 membrane (Millipore Corporation; Billerica, MA).

Supplementary results: *RNA-seq and MudPIT analysis*

Table A3.1. Transcriptional response of A. sp. 75iv2 to growth in SMM in the presence of lignin. *Comparison of the log_2-fold ratio of the top 100-ranked transcripts from cells grown in the presence of dioxane-extracted lignin (DL) or DHP compared to the no lignin control after 24 h. (A) Upregulated genes. (B) Downregulated genes. (- = not statistically signicantly differentially transcribed)*

A

gene ID	gene annotation	Log_2-fold change DL vs. control	Log_2-fold change DHP vs. control
Amy39116DRAFT_5363	amino acid adenylation domain	7.12	7.47
Amy39116DRAFT_5364	amino acid adenylation domain	6.25	6.75
Amy39116DRAFT_5372	salicylate synthase	6.19	7.18
Amy39116DRAFT_5374	Methionyl-tRNA formyltransferase	5.77	5.87
Amy39116DRAFT_5373	Lysine/ornithine N-monooxygenase	5.63	7.01
Amy39116DRAFT_1131	ABC-type multidrug transport system, ATPase component	5.45	-
Amy39116DRAFT_7691	hypothetical protein	4.96	3.79
Amy39116DRAFT_1606	Universal stress protein UspA and related nucleotide-binding proteins	4.83	3.40
Amy39116DRAFT_5362	ABC-type Fe3+hydroxamate transport system, periplasmic component	4.78	5.77
Amy39116DRAFT_4804	hypothetical protein	4.77	2.74
Amy39116DRAFT_3553	Predicted metal-dependent hydrolase	4.72	3.78
Amy39116DRAFT_2325	drug resistance transporter, EmrB/QacA subfamily	4.64	-
Amy39116DRAFT_3550	Short-chain alcohol dehydrogenase of unknown specificity	4.57	3.93
Amy39116DRAFT_4455	2-isopropylmalate synthase/homocitrate synthase family protein	4.55	3.24
Amy39116DRAFT_1647	Universal stress protein UspA and related nucleotide-binding proteins	4.52	3.62
Amy39116DRAFT_7440	Uncharacterized protein conserved in bacteria	4.51	4.47
Amy39116DRAFT_4898	hypothetical protein	4.50	4.21
Amy39116DRAFT_1607	Acyl-CoA synthetase (NDP forming)	4.45	3.40
Amy39116DRAFT_3551	NAD-dependent aldehyde dehydrogenases	4.43	2.56
Amy39116DRAFT_1605	CBS-domain-containing membrane protein	4.39	3.43
Amy39116DRAFT_5924	hypothetical protein	4.38	3.72
Amy39116DRAFT_5416	hypothetical protein	4.38	4.43
Amy39116DRAFT_1603	hypothetical protein	4.35	3.45
Amy39116DRAFT_4296	Uncharacterized protein conserved in bacteria	4.33	-
Amy39116DRAFT_2511	hypothetical protein	4.31	3.75
Amy39116DRAFT_4085	Siderophore-interacting protein	4.31	3.97
Amy39116DRAFT_1132	hypothetical protein	4.29	-
Amy39116DRAFT_2503	Phosphoenolpyruvate synthase/pyruvate phosphate dikinase	4.28	3.38
Amy39116DRAFT_2489	Zn-dependent hydrolases, including glyoxylases	4.26	3.48
Amy39116DRAFT_1604	Universal stress protein UspA and related nucleotide-binding proteins	4.26	2.98
Amy39116DRAFT_1438	conserved repeat domain	4.25	4.56
Amy39116DRAFT_1645	hypothetical protein	4.24	3.64
Amy39116DRAFT_3552	Flavodoxin reductases (ferredoxin-NADPH reductases) family 1	4.24	4.25
Amy39116DRAFT_2502	hypothetical protein	4.18	3.64
Amy39116DRAFT_7439	NAD-dependent aldehyde dehydrogenases	4.17	3.98
Amy39116DRAFT_1642	Zn-dependent alcohol dehydrogenases	4.15	3.73
Amy39116DRAFT_5923	hypothetical protein	4.14	3.53
Amy39116DRAFT_5414	Predicted dehydrogenases and related proteins	4.14	3.93
Amy39116DRAFT_5470	Fatty acid desaturase	4.07	3.03
Amy39116DRAFT_7128	ABC-type Fe3+hydroxamate transport system, periplasmic component	4.05	5.38
Amy39116DRAFT_1621	hypothetical protein	3.98	3.16
Amy39116DRAFT_1588	Cytochrome bd-type quinol oxidase, subunit 1	3.92	3.92
Amy39116DRAFT_1578	Predicted membrane protein	3.92	3.94
Amy39116DRAFT_5367	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	3.91	4.00
Amy39116DRAFT_5440	Glycosyltransferases involved in cell wall biogenesis	3.90	4.23
Amy39116DRAFT_2510	hypothetical protein	3.90	-
Amy39116DRAFT_5369	Peptide arylation enzymes	3.90	4.74
Amy39116DRAFT_1538	conserved repeat domain	3.89	-
Amy39116DRAFT_1646	hypothetical protein	3.87	-
Amy39116DRAFT_4445	conserved repeat domain	3.86	2.76
Amy39116DRAFT_1581	Transcriptional regulator, effector-binding domain/component	3.86	3.16
Amy39116DRAFT_1620	hypothetical protein	3.86	-
Amy39116DRAFT_3554	Predicted flavoprotein involved in K transport	3.81	3.42
Amy39116DRAFT_2508	Flavodoxins	3.80	3.10
Amy39116DRAFT_1644	CBS-domain-containing membrane protein	3.79	3.04
Amy39116DRAFT_6627	ABC-type Mn/Zn transport systems, ATPase component	3.75	3.56
Amy39116DRAFT_2478	hypothetical protein	3.75	-
Amy39116DRAFT_1586	thiol reductant ABC exporter, CydD subunit/thiol reductant ABC exporter, CydC subunit	3.75	3.71
Amy39116DRAFT_1594	ornithine carbamoyltransferase	3.75	3.81
Amy39116DRAFT_5410	Glycosyltransferase	3.74	4.48
Amy39116DRAFT_2477	ATP-dependent metalloprotease FtsH	3.71	2.66
Amy39116DRAFT_2899	ABC-type transport system, involved in lipoprotein release, permease component	3.71	-
Amy39116DRAFT_1580	Ribosome-associated protein Y (PSrp-1)	3.71	4.14

Amy39116DRAFT_3897	Uncharacterized protein conserved in bacteria	3.70	3.25
Amy39116DRAFT_2499	ATPase, P-type (transporting), HAD superfamily, subfamily IC/heavy metal translocating P-type ATPase	3.70	3.35
Amy39116DRAFT_1597	hypothetical protein	3.70	3.12
Amy39116DRAFT_2501	acyltransferase, WS/DGAT/MGAT	3.69	3.05
Amy39116DRAFT_5366	Uncharacterized protein conserved in bacteria	3.68	-
Amy39116DRAFT_2513	anti-anti-sigma factor	3.68	2.68
Amy39116DRAFT_2491	ATPase, P-type (transporting), HAD superfamily, subfamily IC	3.68	2.57
Amy39116DRAFT_2474	glyceraldehyde-3-phosphate dehydrogenase, type I	3.67	3.90
Amy39116DRAFT_1579	Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	3.67	4.27
Amy39116DRAFT_1648	Flavodoxins	3.66	3.19
Amy39116DRAFT_1634	Predicted membrane protein (DUF2078).	3.65	2.14
Amy39116DRAFT_1651	hypothetical protein	3.65	3.24
Amy39116DRAFT_2490	ATPase, P-type (transporting), HAD superfamily, subfamily IC	3.62	2.60
Amy39116DRAFT_5934	hypothetical protein	3.60	3.23
Amy39116DRAFT_1532	hypothetical protein	3.57	-
Amy39116DRAFT_5732	Streptomyces sporulation and cell division protein, SsgA.	3.55	3.21
Amy39116DRAFT_4390	K+transporting ATPase, KdpA	3.55	4.61
Amy39116DRAFT_5413	asparagine synthase (glutamine-hydrolyzing)	3.53	4.18
Amy39116DRAFT_7671	ABC-type cobalamin/Fe3+siderophores transport systems, ATPase components	3.52	2.72
Amy39116DRAFT_1596	Amino acid transporters	3.51	3.09
Amy39116DRAFT_5412	Uncharacterized protein involved in exopolysaccharide biosynthesis	3.48	3.19
Amy39116DRAFT_2472	hypothetical protein	3.45	2.23
Amy39116DRAFT_3099	RNA polymerase sigma-70 factor, sigma-B/F/G subfamily	3.44	-
Amy39116DRAFT_6151	hypothetical protein	3.43	4.31
Amy39116DRAFT_4389	K+transporting ATPase, B subunit	3.43	5.05
Amy39116DRAFT_2509	Universal stress protein UspA and related nucleotide-binding proteins	3.42	2.76
Amy39116DRAFT_0407	Anti-sigma regulatory factor (Ser/Thr protein kinase)	3.41	2.90
Amy39116DRAFT_6916	Conserved TM helix.	3.40	2.45
Amy39116DRAFT_1598	zinc-binding alcohol dehydrogenase family protein	3.40	-
Amy39116DRAFT_4368	Aerobic-type carbon monoxide dehydrogenase, small subunit CoxS/CutS homologs	3.40	-
Amy39116DRAFT_7848	Predicted metal-dependent hydrolase of the TIM-barrel fold	3.39	-
Amy39116DRAFT_2512	CBS-domain-containing membrane protein	3.35	2.27
Amy39116DRAFT_2324	Transcriptional regulator	3.35	-
Amy39116DRAFT_1653	hypothetical protein	3.34	-
Amy39116DRAFT_0304	hypothetical protein	3.34	2.75
Amy39116DRAFT_1587	cytochrome d oxidase, subunit II (cydB)	3.33	2.82
Amy39116DRAFT_6813	Dihydrofolate reductase	3.30	-
Amy39116DRAFT_5166	Bacterial SH3 domain.	3.30	2.35

B

gene ID	gene annotation	Log_2-fold change DL vs. control	Log_2-fold change DHP vs. control
Amy39116DRAFT_1468	Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	-8.29	-8.00
Amy39116DRAFT_1463	Acyl-CoA dehydrogenases	-7.64	-6.76
Amy39116DRAFT_0277	hypothetical protein	-7.07	-7.00
Amy39116DRAFT_0060	hypothetical protein	-6.76	-6.67
Amy39116DRAFT_0809	Protease subunit of ATP-dependent Clp proteases	-6.73	-6.42
Amy39116DRAFT_6492	Fe2 uptake regulation proteins	-6.35	-6.52
Amy39116DRAFT_1466	Putative translation initiation inhibitor, yjgF family	-6.09	-5.91
Amy39116DRAFT_6493	Catalase	-5.86	-6.35
Amy39116DRAFT_0908	hypothetical protein	-5.86	-4.99
Amy39116DRAFT_4310	Predicted transcriptional regulators	-5.69	-6.53
Amy39116DRAFT_2244	Arabinose efflux permease	-5.53	-6.71
Amy39116DRAFT_0910	2-polyprenyl-6-methoxyphenol hydroxylase and related FAD-dependent oxidoreductases	-5.47	-4.98
Amy39116DRAFT_1456	Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II	-5.37	-5.27
Amy39116DRAFT_0909	Maleate cis-trans isomerase	-5.21	-5.08
Amy39116DRAFT_2243	hypothetical protein	-4.98	-5.01
Amy39116DRAFT_2889	Beta-lactamase class C and other penicillin binding proteins	-4.96	-1.89
Amy39116DRAFT_1467	Phenylacetic acid-responsive transcriptional repressor	-4.88	-4.93
Amy39116DRAFT_0911	Sugar phosphate permease	-4.83	-6.37
Amy39116DRAFT_0913	Aerobic-type carbon monoxide dehydrogenase, large subunit CoxL/CutL homologs	-4.74	-3.86
Amy39116DRAFT_7235	hypothetical protein	-4.70	-4.56
Amy39116DRAFT_7764	3-oxoacid CoA-transferase, B subunit	-4.70	-3.33
Amy39116DRAFT_4311	Methylase involved in ubiquinone/menaquinone biosynthesis	-4.62	-5.55
Amy39116DRAFT_0912	Aerobic-type carbon monoxide dehydrogenase, middle subunit CoxM/CutM homologs	-4.61	-3.54

Amy39116DRAFT_4557	Uncharacterized membrane protein (homolog of Drosophila rhomboid)	-4.57	-3.83
Amy39116DRAFT_3316	Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II	-4.49	-3.95
Amy39116DRAFT_0062	hypothetical protein	-4.48	-3.40
Amy39116DRAFT_3466	integral membrane protein, TerC family	-4.47	-4.34
Amy39116DRAFT_7532	hypothetical protein	-4.47	-3.69
Amy39116DRAFT_6906	Transcriptional regulators containing a DNA-binding HTH domain and an aminotransferase domain (MocR family) and their eukaryotic orthologs	-4.46	-4.94
Amy39116DRAFT_1458	Uncharacterized protein, possibly involved in aromatic compounds catabolism	-4.45	-2.50
Amy39116DRAFT_0905	hypothetical protein	-4.42	-3.27
Amy39116DRAFT_6494	hypothetical protein	-4.29	-3.94
Amy39116DRAFT_1388	Indolepyruvate ferredoxin oxidoreductase, alpha and beta subunits	-4.19	-2.26
Amy39116DRAFT_5069	Nucleoside-diphosphate-sugar epimerases	-4.18	-3.24
Amy39116DRAFT_2888	Proline dehydrogenase	-4.17	-1.98
Amy39116DRAFT_7533	Molecular chaperone (small heat shock protein)	-4.11	-3.56
Amy39116DRAFT_3012	2-polyprenyl-6-methoxyphenol hydroxylase and related FAD-dependent oxidoreductases	-4.03	-
Amy39116DRAFT_2887	delta-1-pyrroline-5-carboxylate dehydrogenase, group 1	-4.03	-1.66
Amy39116DRAFT_4641	cation diffusion facilitator family transporter	-4.01	-3.77
Amy39116DRAFT_2672	ABC-type sugar transport system, periplasmic component	-3.98	-3.30
Amy39116DRAFT_8167	Aerobic-type carbon monoxide dehydrogenase, large subunit CoxL/CutL homologs	-3.98	-2.17
Amy39116DRAFT_0778	hypothetical protein	-3.94	-3.76
Amy39116DRAFT_6720	thiamine biosynthesis protein ThiS	-3.94	-3.30
Amy39116DRAFT_5208	Putative stress-responsive transcriptional regulator	-3.93	-3.18
Amy39116DRAFT_1969	Cytochrome P450	-3.93	-3.29
Amy39116DRAFT_1255	Transposase and inactivated derivatives	-3.89	-3.75
Amy39116DRAFT_6747	Putative stress-responsive transcriptional regulator	-3.85	-2.34
Amy39116DRAFT_3413	Uncharacterized conserved protein	-3.84	-3.19
Amy39116DRAFT_0800	Dihydrofolate reductase	-3.82	-
Amy39116DRAFT_1336	Transcriptional regulator	-3.81	-4.38
Amy39116DRAFT_0777	hypothetical protein	-3.79	-4.48
Amy39116DRAFT_3198	Uncharacterized protein encoded in hypervariable junctions of pilus gene clusters	-3.77	-3.31
Amy39116DRAFT_1624	Rhodanese-related sulfurtransferase	-3.74	-3.96
Amy39116DRAFT_1865	Fe2□ uptake regulation proteins	-3.72	-3.48
Amy39116DRAFT_1455	Isopropylmalate/homocitrate/citramalate synthases	-3.70	-2.58
Amy39116DRAFT_1454	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	-3.68	-2.76
Amy39116DRAFT_1561	Enterochelin esterase and related enzymes	-3.68	-2.67
Amy39116DRAFT_1453	NIPSNAP.	-3.63	-2.39
Amy39116DRAFT_7920	Acyl-CoA dehydrogenases	-3.62	-4.02
Amy39116DRAFT_3935	hypothetical protein	-3.58	-4.33
Amy39116DRAFT_3013	Transcriptional regulator	-3.58	-2.27
Amy39116DRAFT_1887	Transcriptional regulator	-3.57	-4.55
Amy39116DRAFT_0360	Putative stress-responsive transcriptional regulator	-3.56	-2.73
Amy39116DRAFT_1754	hypothetical protein	-3.55	-2.83
Amy39116DRAFT_1810	Dienelactone hydrolase and related enzymes	-3.54	-3.69
Amy39116DRAFT_3618	Beta-lactamase class A	-3.54	-1.85
Amy39116DRAFT_2167	hypothetical protein	-3.52	-2.96
Amy39116DRAFT_7681	hypothetical protein	-3.51	-3.15
Amy39116DRAFT_6725	thiamine biosynthesis protein ThiC	-3.51	-2.68
Amy39116DRAFT_0061	hypothetical protein	-3.50	-2.62
Amy39116DRAFT_3925	Transcriptional regulators	-3.50	-4.51
Amy39116DRAFT_6721	Uncharacterized enzyme of thiazole biosynthesis	-3.49	-3.97
Amy39116DRAFT_3199	hypothetical protein	-3.48	-3.50
Amy39116DRAFT_0359	hypothetical protein	-3.41	-2.52
Amy39116DRAFT_2372	hypothetical protein	-3.41	-3.57
Amy39116DRAFT_5209	hypothetical protein	-3.40	-1.73
Amy39116DRAFT_7267	RNA polymerase sigma-70 factor, sigma-E family	-3.37	-2.50
Amy39116DRAFT_4734	ABC-type molybdenum transport system, ATPase component/photorepair protein PhrA	-3.36	-2.40
Amy39116DRAFT_3173	Helix-turn-helix.	-3.31	-3.96
Amy39116DRAFT_6724	phosphomethylpyrimidine kinase	-3.31	-2.72
Amy39116DRAFT_1968	hypothetical protein	-3.30	-2.51
Amy39116DRAFT_0063	Protein of unknown function (DUF3307).	-3.29	-2.85
Amy39116DRAFT_1625	Uncharacterized protein conserved in bacteria	-3.29	-2.96
Amy39116DRAFT_3888	RNA polymerase sigma factor, sigma-70 family	-3.28	-2.27
Amy39116DRAFT_4194	Protease subunit of ATP-dependent Clp proteases	-3.27	-2.75
Amy39116DRAFT_4692	Uncharacterized conserved protein	-3.27	-
Amy39116DRAFT_7264	Acetyl-CoA carboxylase, carboxyltransferase component (subunits alpha and beta)	-3.26	-3.33
Amy39116DRAFT_0779	YcfA-like protein.	-3.26	-3.19
Amy39116DRAFT_0801	Domain of unknown function (DU1801).	-3.26	-
Amy39116DRAFT_7112	transcription elongation factor GreA	-3.23	-1.84
Amy39116DRAFT_3931	Transcriptional regulator	-3.22	-4.07
Amy39116DRAFT_2028	Membrane protease subunits, stomatin/prohibitin homologs	-3.20	-2.97
Amy39116DRAFT_7839	phenylacetic acid degradation protein paaN	-3.18	-3.92

Amy39116DRAFT_0873	Uncharacterized conserved protein	-3.16	-2.34
Amy39116DRAFT_1911	Cu2+containing amine oxidase	-3.16	-3.47
Amy39116DRAFT_0333	Short-chain alcohol dehydrogenase of unknown specificity	-3.16	-2.19
Amy39116DRAFT_1172	Pyruvate/2-oxoglutarate dehydrogenase complex, dehydrogenase (E1) component, eukaryotic type, beta subunit	-3.14	-2.53
Amy39116DRAFT_7265	Acetyl/propionyl-CoA carboxylase, alpha subunit	-3.12	-3.16
Amy39116DRAFT_6297	hypothetical protein	-3.12	-2.38
Amy39116DRAFT_0907	Transcriptional regulators	-3.11	-2.39
Amy39116DRAFT_7844	phenylacetate-CoA oxygenase/reductase, PaaK subunit	-3.10	-3.04

Table A3.2. Comparison of lignin-dependent transcriptional responses of A. sp. 75iv2. *Full list of the comparison of log_2-fold ratio of transcripts from cells grown in the presence of dioxane-extracted lignin (DL) and DHP after 24 h.*

gene ID	gene annotation	Log_2-fold change DHP vs. DL
Amy39116DRAFT_1132	hypothetical protein	-3.54
Amy39116DRAFT_1131	ABC-type multidrug transport system, ATPase component	-3.31
Amy39116DRAFT_4296	Uncharacterized protein conserved in bacteria	-3.03
Amy39116DRAFT_6813	Dihydrofolate reductase	-3.02
Amy39116DRAFT_2325	drug resistance transporter, EmrB/QacA subfamily	-2.69
Amy39116DRAFT_7559	ABC-type multidrug transport system, ATPase component	-2.57
Amy39116DRAFT_0307	Enoyl-CoA hydratase/carnithine racemase	-2.43
Amy39116DRAFT_2195	Predicted transcriptional regulators	-2.38
Amy39116DRAFT_2510	hypothetical protein	-2.36
Amy39116DRAFT_6593	Predicted transcriptional regulators	-2.23
Amy39116DRAFT_2194	Uncharacterized conserved protein	-2.18
Amy39116DRAFT_6278	Transcriptional regulators	-2.15
Amy39116DRAFT_6686	benzoyl-CoA-dihydrodiol lyase	-2.14
Amy39116DRAFT_6443	4-hydroxybenzoate 3-monooxygenase	-2.14
Amy39116DRAFT_4808	Uncharacterized conserved protein with double-stranded beta-helix domain	-2.12
Amy39116DRAFT_7558	ABC-type multidrug transport system, permease component	-2.09
Amy39116DRAFT_7673	hypothetical protein	-2.02
Amy39116DRAFT_6279	drug resistance transporter, EmrB/QacA subfamily	-2.02
Amy39116DRAFT_1646	hypothetical protein	-2.00
Amy39116DRAFT_0522	Trypsin-like serine proteases, typically periplasmic, contain C-terminal PDZ domain	-1.99
Amy39116DRAFT_1146	Transcriptional regulator	-1.96
Amy39116DRAFT_2324	Transcriptional regulator	-1.83
Amy39116DRAFT_6687	benzoyl-CoA oxygenase, B subunit	-1.80
Amy39116DRAFT_6812	Uncharacterized protein conserved in bacteria	-1.76
Amy39116DRAFT_0197	Transcriptional regulator	-1.73
Amy39116DRAFT_5191	succinyl-CoA synthetase, alpha subunit	1.42
Amy39116DRAFT_5332	Ribosomal protein L10	1.46
Amy39116DRAFT_7007	ribosomal protein L25, Ctc-form	1.55
Amy39116DRAFT_5532	Protein of unknown function (DUF2596).	1.62
Amy39116DRAFT_5445	Nucleoside-diphosphate-sugar pyrophosphorylase involved in lipopolysaccharide biosynthesis	1.62
Amy39116DRAFT_5149	succinate dehydrogenase, cytochrome b556 subunit	1.68
Amy39116DRAFT_6182	Protein of unknown function (DUF1416).	1.74
Amy39116DRAFT_6632	hypothetical protein	1.76
Amy39116DRAFT_4389	K+transporting ATPase, B subunit	1.77
Amy39116DRAFT_0044	hypothetical protein	1.81
Amy39116DRAFT_0049	hypothetical protein	1.81
Amy39116DRAFT_5151	succinate dehydrogenase, flavoprotein subunit, E. coli/mitochondrial subgroup	1.83
Amy39116DRAFT_5818	hypothetical protein	1.83
Amy39116DRAFT_0038	Transcription factor WhiB.	1.86
Amy39116DRAFT_4103	ABC transporter, substrate-binding protein, aliphatic sulfonates family	1.94
Amy39116DRAFT_4098	Sulfite reductase, beta subunit (hemoprotein)	1.94
Amy39116DRAFT_2887	delta-1-pyrroline-5-carboxylate dehydrogenase, group 1	1.94
Amy39116DRAFT_0032	hypothetical protein	2.00
Amy39116DRAFT_0046	hypothetical protein	2.00
Amy39116DRAFT_5150	Succinate dehydrogenase, hydrophobic anchor subunit	2.04
Amy39116DRAFT_6115	Phosphoenolpyruvate carboxykinase (GTP)	2.09
Amy39116DRAFT_5972	NADPH-dependent glutamate synthase beta chain and related oxidoreductases	2.16
Amy39116DRAFT_0037	hypothetical protein	2.16
Amy39116DRAFT_6102	ABC-type antimicrobial peptide transport system, permease component	2.31
Amy39116DRAFT_0045	hypothetical protein	2.32
Amy39116DRAFT_4101	sulfate adenylyltransferase, small subunit	2.50
Amy39116DRAFT_4104	ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component	2.56
Amy39116DRAFT_5973	hypothetical protein	2.60
Amy39116DRAFT_6180	hypothetical protein	2.60
Amy39116DRAFT_4100	phosophoadenylyl-sulfate reductase (thioredoxin)	2.61
Amy39116DRAFT_0042	hypothetical protein	2.73
Amy39116DRAFT_4102	sulfate adenylyltransferase, large subunit	2.94
Amy39116DRAFT_0043	hypothetical protein	3.25
Amy39116DRAFT_6181	Rhodanese-related sulfurtransferase	3.40
Amy39116DRAFT_1828	ABC-type polysaccharide/polyol phosphate export systems, permease component	3.40
Amy39116DRAFT_6101	ABC-type antimicrobial peptide transport system, ATPase component	3.50
Amy39116DRAFT_1830	Predicted flavoprotein	3.71
Amy39116DRAFT_1829	daunorubicin resistance ABC transporter ATP-binding subunit	4.60

Table A3.3. List of upregulated proteins from intracellular protein samples at 24 h by MUDPIT analysis. *Comparison of the log_2-fold ratio of proteins from cells grown in the presence of dioxane-extracted lignin (DL) compared to the no lignin control after 24 h. ND = not determined.*

gene ID	gene description	control	lignin	log_2 ratio
Amy39116DRAFT_6119	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	0	70	ND
Amy39116DRAFT_0626	gamma-glutamyltranspeptidase	0	42	ND
Amy39116DRAFT_5372	salicylate synthase	0	32	ND
Amy39116DRAFT_1637	CBS-domain-containing membrane protein	0	30	ND
Amy39116DRAFT_5374	Methionyl-tRNA formyltransferase	0	20	ND
Amy39116DRAFT_5363	amino acid adenylation domain	0	19	ND
Amy39116DRAFT_5658	ABC-type Fe3+-hydroxamate transport system, periplasmic component	0	15	ND
Amy39116DRAFT_0222	glutamine synthetase, type I	0	15	ND
Amy39116DRAFT_7251	alanine dehydrogenase	0	15	ND
Amy39116DRAFT_3844	hypothetical protein	0	14	ND
Amy39116DRAFT_5396	Uncharacterized protein conserved in bacteria	0	14	ND
Amy39116DRAFT_0559	Glutathione peroxidase	0	13	ND
Amy39116DRAFT_3384	MmoB/DmpM family.	0	13	ND
Amy39116DRAFT_6249	Glycerophosphoryl diester phosphodiesterase	0	12	ND
Amy39116DRAFT_3924	Homogentisate 1,2-dioxygenase	0	12	ND
Amy39116DRAFT_0484	Cold shock proteins	0	12	ND
Amy39116DRAFT_7440	Uncharacterized protein conserved in bacteria	0	12	ND
Amy39116DRAFT_7045	Exopolyphosphatase	0	11	ND
Amy39116DRAFT_3979	isoform II.	0	11	ND
Amy39116DRAFT_2512	CBS-domain-containing membrane protein	0	11	ND
Amy39116DRAFT_7963	Cell wall-associated hydrolases (invasion-associated proteins)	0	10	ND
Amy39116DRAFT_5003	TIGR00730 family protein	0	9	ND
Amy39116DRAFT_4296	Uncharacterized protein conserved in bacteria	0	9	ND
Amy39116DRAFT_4437	3-isopropylmalate dehydratase, small subunit	0	9	ND
Amy39116DRAFT_5262	ribosomal protein L13, bacterial type	0	9	ND
Amy39116DRAFT_2725	Enoyl-CoA hydratase/carnithine racemase	0	9	ND
Amy39116DRAFT_5369	Peptide arylation enzymes	0	9	ND
Amy39116DRAFT_3383	Predicted metal-dependent hydrolase of the TIM-barrel fold	0	9	ND
Amy39116DRAFT_1579	Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	0	8	ND
Amy39116DRAFT_6593	Predicted transcriptional regulators	0	8	ND
Amy39116DRAFT_4266	ATP-binding cassette protein, ChvD family	0	8	ND
Amy39116DRAFT_0734	Isopenicillin N synthase and related dioxygenases	0	8	ND
Amy39116DRAFT_8269	Predicted nucleoside-diphosphate-sugar epimerases	0	8	ND
Amy39116DRAFT_2929	ABC-type xylose transport system, periplasmic component	0	8	ND
Amy39116DRAFT_0303	Universal stress protein UspA and related nucleotide-binding proteins	0	8	ND
Amy39116DRAFT_3380	chaperonin GroL	0	8	ND
Amy39116DRAFT_5005	succinyl-diaminopimelate desuccinylase	0	7	ND
Amy39116DRAFT_3063	SnoaL-like polyketide cyclase.	0	7	ND
Amy39116DRAFT_3138	UDP-N-acetylmuramate--alanine ligase	0	7	ND
Amy39116DRAFT_5813	Fructose-2,6-bisphosphatase	0	7	ND
Amy39116DRAFT_3000	Poly(3-hydroxybutyrate) depolymerase	0	7	ND
Amy39116DRAFT_5302	ribosomal protein S3, bacterial type	0	7	ND
Amy39116DRAFT_5235	Negative regulator of beta-lactamase expression	0	7	ND
Amy39116DRAFT_0608	translation elongation factor P	0	7	ND
Amy39116DRAFT_2471	Universal stress protein UspA and related nucleotide-binding proteins	0	7	ND
Amy39116DRAFT_4067	Subtilisin inhibitor-like.	0	7	ND
Amy39116DRAFT_4871	hypothetical protein	0	6	ND
Amy39116DRAFT_5211	GMP synthase (glutamine-hydrolyzing), C-terminal domain or B subunit/GMP synthase (glutamine-hydrolyzing), N-terminal domain or A subunit	0	6	ND
Amy39116DRAFT_5221	inosine-5''-monophosphate dehydrogenase	0	6	ND
Amy39116DRAFT_6653	Uncharacterized protein conserved in bacteria	0	6	ND
Amy39116DRAFT_4240	ATP-dependent transcriptional regulator	0	6	ND
Amy39116DRAFT_7934	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide acyltransferase (E2) component, and related enzymes	0	6	ND
Amy39116DRAFT_2798	Predicted redox protein, regulator of disulfide bond formation	0	6	ND
Amy39116DRAFT_5479	Aminopeptidase N	0	6	ND
Amy39116DRAFT_4085	Siderophore-interacting protein	0	6	ND
Amy39116DRAFT_2600	uncharacterized domain 1	0	6	ND
Amy39116DRAFT_1613	hypothetical protein	0	5	ND
Amy39116DRAFT_6718	thiamine-phosphate pyrophosphorylase	0	5	ND
Amy39116DRAFT_6602	Lysyl-tRNA synthetase (class II)	0	5	ND

Amy39116DRAFT_7928 lipoate-protein ligase B	0	5	ND
Amy39116DRAFT_3144 UDP-N-acetylmuramyl-tripeptide synthetase	0	5	ND
Amy39116DRAFT_3790 heme/flavin dehydrogenase, mycofactocin system	0	5	ND
Amy39116DRAFT_3877 Uncharacterized protein conserved in bacteria	0	5	ND
Amy39116DRAFT_0463 Protein of unknown function (DUF574).	0	5	ND
Amy39116DRAFT_2750 threonyl-tRNA synthetase	0	5	ND
Amy39116DRAFT_4002 ABC-type Fe3+-hydroxamate transport system, periplasmic component	0	5	ND
Amy39116DRAFT_4065 4-aminobutyrate aminotransferase and related aminotransferases	0	5	ND
Amy39116DRAFT_2455 Putative intracellular protease/amidase	0	5	ND
Amy39116DRAFT_4147 ribosomal protein L21	0	5	ND
Amy39116DRAFT_5365 non-ribosomal peptide synthase domain TIGR01720/amino acid adenylation domain	2	26	3.70
Amy39116DRAFT_5362 ABC-type Fe3+-hydroxamate transport system, periplasmic component	2	25	3.64
Amy39116DRAFT_7437 acetate--CoA ligase	3	32	3.42
Amy39116DRAFT_4457 Subtilase family.	6	56	3.22
Amy39116DRAFT_7439 NAD-dependent aldehyde dehydrogenases	23	206	3.16
Amy39116DRAFT_2587 Ribosomal protein S1	5	42	3.07
Amy39116DRAFT_6163 5''-nucleotidase/2'',3''-cyclic phosphodiesterase and related esterases	7	57	3.03
Amy39116DRAFT_7058 NAD-dependent aldehyde dehydrogenases	6	46	2.94
Amy39116DRAFT_6831 phosphoserine aminotransferase, putative	3	23	2.94
Amy39116DRAFT_6744 Cold shock proteins	5	37	2.89
Amy39116DRAFT_5140 adenosine deaminase	2	14	2.81
Amy39116DRAFT_0703 Predicted metal-sulfur cluster biosynthetic enzyme	2	14	2.81
Amy39116DRAFT_5673 copper ion binding protein	2	14	2.81
Amy39116DRAFT_2504 Universal stress protein UspA and related nucleotide-binding proteins	2	13	2.70
Amy39116DRAFT_7933 2-oxoglutarate dehydrogenase, E2 component, dihydrolipoamide succinyltransferase	5	31	2.63
Amy39116DRAFT_4194 Protease subunit of ATP-dependent Clp proteases	3	17	2.50
Amy39116DRAFT_4527 isocitrate lyase	5	28	2.49
Amy39116DRAFT_4317 translation elongation factor Ts	6	32	2.42
Amy39116DRAFT_5168 isocitrate dehydrogenase, NADP-dependent, eukaryotic type	7	36	2.36
Amy39116DRAFT_5187 phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase	3	15	2.32
Amy39116DRAFT_5525 Response regulators consisting of a CheY-like receiver domain and a winged-helix DNA-binding domain	2	10	2.32
Amy39116DRAFT_4463 ketol-acid reductoisomerase	5	24	2.26
Amy39116DRAFT_7140 Fumarase	3	14	2.22
Amy39116DRAFT_4561 Predicted metal-dependent protease of the PAD1/JAB1 superfamily	3	14	2.22
Amy39116DRAFT_8168 Aerobic-type carbon monoxide dehydrogenase, small subunit CoxS/CutS homologs	3	14	2.22
Amy39116DRAFT_4275 Serine/threonine protein phosphatase	2	9	2.17
Amy39116DRAFT_3499 Dihydrofolate reductase	2	9	2.17
Amy39116DRAFT_5232 Co-chaperonin GroES (HSP10)	14	62	2.15
Amy39116DRAFT_0746 Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	3	13	2.12
Amy39116DRAFT_4111 Enoyl-CoA hydratase/carnithine racemase	2	8	2.00
Amy39116DRAFT_6418 Cold shock proteins	86	330	1.94
Amy39116DRAFT_6482 aspartate kinase, monofunctional class	15	57	1.93
Amy39116DRAFT_7704 dihydrolipoamide dehydrogenase	4	15	1.91
Amy39116DRAFT_7181 ABC-type dipeptide transport system, periplasmic component	23	86	1.90
Amy39116DRAFT_5103 methylmalonyl-CoA mutase C-terminal domain/methylmalonyl-CoA mutase N-terminal domain	3	11	1.87
Amy39116DRAFT_6702 diaminobutyrate--2-oxoglutarate aminotransferase	3	11	1.87
Amy39116DRAFT_7613 Glutamate dehydrogenase/leucine dehydrogenase	10	36	1.85
Amy39116DRAFT_5578 O-Methyltransferase involved in polyketide biosynthesis	6	21	1.81
Amy39116DRAFT_0744 Predicted aminoglycoside phosphotransferase	2	7	1.81
Amy39116DRAFT_7202 adenosylhomocysteinase	7	24	1.78
Amy39116DRAFT_4182 Predicted dioxygenase of extradiol dioxygenase family	5	17	1.77
Amy39116DRAFT_0699 FeS assembly protein SufD	3	10	1.74
Amy39116DRAFT_5550 Uncharacterized conserved protein	3	10	1.74
Amy39116DRAFT_2543 probable methyltransferase	3	10	1.74
Amy39116DRAFT_0698 FeS assembly protein SufB	4	13	1.70
Amy39116DRAFT_4412 pyruvate carboxylase	5	16	1.68
Amy39116DRAFT_6003 6-phosphogluconate dehydrogenase (decarboxylating)	5	16	1.68
Amy39116DRAFT_4458 D-3-phosphoglycerate dehydrogenase	6	19	1.66
Amy39116DRAFT_8167 Aerobic-type carbon monoxide dehydrogenase, large subunit CoxL/CutL homologs	8	24	1.58
Amy39116DRAFT_5503 porphobilinogen deaminase	3	9	1.58
Amy39116DRAFT_4193 Protease subunit of ATP-dependent Clp proteases	3	9	1.58
Amy39116DRAFT_7556 Cobalamin biosynthesis protein CobN and related Mg-chelatases	3	9	1.58
Amy39116DRAFT_2007 Uncharacterized conserved protein	3	9	1.58

Amy39116DRAFT_4958 Predicted metal-dependent phosphoesterases (PHP family)	2	6	1.58
Amy39116DRAFT_7156 GTP-binding protein YchF	2	6	1.58
Amy39116DRAFT_8115 Serine proteases of the peptidase family S9A	2	6	1.58
Amy39116DRAFT_5908 Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	6	17	1.50
Amy39116DRAFT_6541 hypothetical protein	5	14	1.49
Amy39116DRAFT_3136 peroxiredoxin, OsmC subfamily	8	22	1.46
Amy39116DRAFT_5078 Acetyl-CoA carboxylase, carboxyltransferase component (subunits alpha and beta)	4	11	1.46
Amy39116DRAFT_6572 hypothetical protein	4	11	1.46
Amy39116DRAFT_6535 2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase (catechol pathway)	4	11	1.46
Amy39116DRAFT_8111 ribonucleoside-diphosphate reductase, adenosylcobalamin-dependent	4	11	1.46
Amy39116DRAFT_0254 Protein of unknown function (DUF574).	4	11	1.46
Amy39116DRAFT_4996 Protein of unknown function (DUF3117).	19	52	1.45
Amy39116DRAFT_5884 Pyruvate/2-oxoglutarate dehydrogenase complex, dehydrogenase (E1) component, eukaryotic type, beta subunit	12	32	1.42
Amy39116DRAFT_5148 Uncharacterized ABC-type transport system, periplasmic component/surface lipoprotein	3	8	1.42
Amy39116DRAFT_4438 3-isopropylmalate dehydratase, large subunit	3	8	1.42
Amy39116DRAFT_5248 D-xylulose kinase	3	8	1.42
Amy39116DRAFT_6265 fructose-bisphosphate aldolase, class II, yeast/E. coli subtype	11	29	1.40
Amy39116DRAFT_3212 polyribonucleotide nucleotidyltransferase	18	47	1.38
Amy39116DRAFT_5206 Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain	5	13	1.38
Amy39116DRAFT_4837 glutamine synthetase, type I	12	31	1.37
Amy39116DRAFT_2558 glycine cleavage system H protein	12	30	1.32
Amy39116DRAFT_2937 Arsenate reductase and related proteins, glutaredoxin family	6	15	1.32
Amy39116DRAFT_6704 ectoine hydroxylase	4	10	1.32
Amy39116DRAFT_4365 deazaflavin-dependent nitroreductase family protein	4	10	1.32
Amy39116DRAFT_4801 histidine ammonia-lyase	2	5	1.32
Amy39116DRAFT_2985 hypothetical protein	2	5	1.32
Amy39116DRAFT_8022 ABC-type Fe3+-hydroxamate transport system, periplasmic component	2	5	1.32
Amy39116DRAFT_5283 translation initiation factor IF-1	2	5	1.32
Amy39116DRAFT_5523 hypothetical protein	2	5	1.32
Amy39116DRAFT_5450 2-oxoacid:acceptor oxidoreductase, alpha subunit	2	5	1.32
Amy39116DRAFT_6079 inositol 1-phosphate synthase, Actinobacterial type	2	5	1.32
Amy39116DRAFT_3377 Uncharacterized conserved protein	2	5	1.32
Amy39116DRAFT_7269 methionine aminopeptidase, type I	2	5	1.32
Amy39116DRAFT_8229 methylmalonyl-CoA mutase N-terminal domain	19	47	1.31
Amy39116DRAFT_7836 phenylacetate-CoA ligase	5	12	1.26
Amy39116DRAFT_7242 2-oxoglutarate dehydrogenase, E1 component	13	31	1.25
Amy39116DRAFT_6205 Protein of unknown function (DUF574).	6	14	1.22
Amy39116DRAFT_4638 arginyl-tRNA synthetase	3	7	1.22
Amy39116DRAFT_6267 Predicted oxidoreductases (related to aryl-alcohol dehydrogenases)	3	7	1.22
Amy39116DRAFT_2825 Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	3	7	1.22
Amy39116DRAFT_5336 transcription termination/antitermination factor NusG	3	7	1.22
Amy39116DRAFT_2231 Pyruvate/2-oxoglutarate dehydrogenase complex, dehydrogenase (E1) component, eukaryotic type, alpha subunit	3	7	1.22
Amy39116DRAFT_7569 cobaltochelatase subunit	3	7	1.22
Amy39116DRAFT_6486 Predicted oxidoreductases (related to aryl-alcohol dehydrogenases)	10	23	1.20
Amy39116DRAFT_6248 phosphoribosylamine--glycine ligase	10	23	1.20
Amy39116DRAFT_4456 Isocitrate/isopropylmalate dehydrogenase	4	9	1.17
Amy39116DRAFT_7911 L-asparaginase II	4	9	1.17
Amy39116DRAFT_2515 ferrochelatase	4	9	1.17
Amy39116DRAFT_1981 phosphoribosyl-ATP pyrophosphohydrolase	4	9	1.17
Amy39116DRAFT_6542 acetate--CoA ligase	10	22	1.14
Amy39116DRAFT_7442 MoxR-like ATPases	11	24	1.13
Amy39116DRAFT_4512 Electron transfer flavoprotein, alpha subunit	39	85	1.12
Amy39116DRAFT_7041 phosphopyruvate hydratase	12	26	1.12
Amy39116DRAFT_2003 proteasome, beta subunit, bacterial type	6	13	1.12
Amy39116DRAFT_6326 chaperone protein DnaK	14	30	1.10
Amy39116DRAFT_5040 Acyl-CoA dehydrogenases	9	19	1.08
Amy39116DRAFT_4619 proton translocating ATP synthase, F1 alpha subunit	17	35	1.04
Amy39116DRAFT_7265 Acetyl/propionyl-CoA carboxylase, alpha subunit	20	40	1.00
Amy39116DRAFT_5883 pyruvate dehydrogenase E1 component, alpha subunit	8	16	1.00
Amy39116DRAFT_6501 DNA-binding ferritin-like protein (oxidative damage protectant)	7	14	1.00
Amy39116DRAFT_2199 Predicted lactoylglutathione lyase	5	10	1.00
Amy39116DRAFT_4620 ATP synthase, F1 delta subunit	4	8	1.00
Amy39116DRAFT_5534 Uncharacterized protein conserved in bacteria	4	8	1.00

Amy39116DRAFT_7842 phenylacetate-CoA oxygenase, PaaI subunit	3	6	1.00
Amy39116DRAFT_3668 oxidoreductase, Rxyl_3153 family	3	6	1.00
Amy39116DRAFT_5261 Ribosomal protein S9	3	6	1.00
Amy39116DRAFT_0511 Acetyl/propionyl-CoA carboxylase, alpha subunit	3	6	1.00

Table A3.4. List of downregulated proteins from intracellular protein samples at 24 h by MUDPIT analysis. *Comparison of the log_2-fold ratio of transcripts from cells grown in the presence of dioxane-extracted lignin (DL) to the no lignin control after 24 h. ND = not determined.*

gene ID	gene annotation	control	lignin	log_2ratio
Amy39116DRAFT_6843	pyridoxamine-phosphate oxidase	31	0	ND
Amy39116DRAFT_3108	hypothetical protein	25	0	ND
Amy39116DRAFT_5721	histidinol-phosphate aminotransferase	20	0	ND
Amy39116DRAFT_2049	hypothetical protein	20	0	ND
Amy39116DRAFT_0872	uncharacterized Actinobacterial protein TIGR03083	17	0	ND
Amy39116DRAFT_0104	1-deoxy-D-xylulose-5-phosphate synthase	17	0	ND
Amy39116DRAFT_4331	Predicted RNA-binding protein (contains KH domain)	16	0	ND
Amy39116DRAFT_3020	Protein of unknown function (DUF664).	16	0	ND
Amy39116DRAFT_0701	cysteine desulfurases, SufS subfamily	15	0	ND
Amy39116DRAFT_8036	Uncharacterized proteins, LmbE homologs	14	0	ND
Amy39116DRAFT_2372	hypothetical protein	12	0	ND
Amy39116DRAFT_2300	acetyl-CoA acetyltransferases	11	0	ND
Amy39116DRAFT_1388	Indolepyruvate ferredoxin oxidoreductase, alpha and beta subunits	10	0	ND
Amy39116DRAFT_0816	Inosine-uridine nucleoside N-ribohydrolase	10	0	ND
Amy39116DRAFT_8290	Uncharacterized protein conserved in bacteria	10	0	ND
Amy39116DRAFT_5979	hypothetical protein	10	0	ND
Amy39116DRAFT_7127	Uncharacterized conserved protein	9	0	ND
Amy39116DRAFT_7841	phenylacetate-CoA oxygenase, PaaH subunit	9	0	ND
Amy39116DRAFT_5276	ribosomal protein L17	9	0	ND
Amy39116DRAFT_5095	Predicted enzyme related to lactoylglutathione lyase	8	0	ND
Amy39116DRAFT_8151	Polyketide cyclase / dehydrase and lipid transport.	8	0	ND
Amy39116DRAFT_7618	Transcriptional regulator	8	0	ND
Amy39116DRAFT_7182	N-acetyl-1-D-myo-inositol-2-amino-2-deoxy-alpha-D-glucopyranoside deacetylase	7	0	ND
Amy39116DRAFT_6876	hypothetical protein	7	0	ND
Amy39116DRAFT_6720	thiamine biosynthesis protein ThiS	7	0	ND
Amy39116DRAFT_5507	Glutaredoxin and related proteins	7	0	ND
Amy39116DRAFT_0287	Lactoylglutathione lyase and related lyases	7	0	ND
Amy39116DRAFT_2367	Methylase involved in ubiquinone/menaquinone biosynthesis	7	0	ND
Amy39116DRAFT_3237	Selenocysteine lyase	7	0	ND
Amy39116DRAFT_7147	hypothetical protein	6	0	ND
Amy39116DRAFT_5901	hypothetical protein	6	0	ND
Amy39116DRAFT_3314	hypothetical protein	6	0	ND
Amy39116DRAFT_0140	Uncharacterized conserved protein	6	0	ND
Amy39116DRAFT_7655	Xanthine and CO dehydrogenases maturation factor, XdhC/CoxF family	6	0	ND
Amy39116DRAFT_2004	proteasome, alpha subunit, bacterial type	6	0	ND
Amy39116DRAFT_7225	RNA polymerase sigma-70 factor, TIGR02947 family	5	0	ND
Amy39116DRAFT_1754	hypothetical protein	5	0	ND
Amy39116DRAFT_4411	Arginase/agmatinase/formimionoglutamate hydrolase, arginase family	5	0	ND
Amy39116DRAFT_0690	NADPH:quinone reductase and related Zn-dependent oxidoreductases	5	0	ND
Amy39116DRAFT_0702	SUF system FeS assembly protein, NifU family	5	0	ND
Amy39116DRAFT_0887	Lactoylglutathione lyase and related lyases	5	0	ND
Amy39116DRAFT_7959	Transcriptional regulators	5	0	ND
Amy39116DRAFT_3487	hypothetical protein	5	0	ND
Amy39116DRAFT_3492	Glycine/D-amino acid oxidases (deaminating)	5	0	ND
Amy39116DRAFT_0472	Predicted acyl-CoA transferases/carnitine dehydratase	5	0	ND
Amy39116DRAFT_5465	NADH:ubiquinone oxidoreductase 24 kD subunit	5	0	ND
Amy39116DRAFT_6193	3-hydroxyacyl-CoA dehydrogenase	5	0	ND
Amy39116DRAFT_6184	Domain of unknown function (DUF1794).	5	0	ND
Amy39116DRAFT_6066	ribosomal protein L9	5	0	ND
Amy39116DRAFT_3653	Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	214	10	4.42
Amy39116DRAFT_7109	4-hydroxyphenylpyruvate dioxygenase	66	4	4.04
Amy39116DRAFT_8165	oxidoreductase, Rxyl_3153 family	34	3	3.50
Amy39116DRAFT_4568	Amidases related to nicotinamidase	22	2	3.46
Amy39116DRAFT_3463	Predicted metal-dependent hydrolase	28	3	3.22
Amy39116DRAFT_5537	Protein of unknown function (DUF2505).	36	4	3.17
Amy39116DRAFT_4402	Cell division initiation protein	16	2	3.00
Amy39116DRAFT_5172	Domain of unknown function (DUF1918).	23	3	2.94
Amy39116DRAFT_4123	Dioxygenases related to 2-nitropropane dioxygenase	14	2	2.81
Amy39116DRAFT_7953	response regulator receiver protein	40	6	2.74
Amy39116DRAFT_7764	3-oxoacid CoA-transferase, B subunit	20	3	2.74
Amy39116DRAFT_6871	hypothetical protein	36	6	2.58

Amy39116DRAFT_6263	Protein of unknown function (DUF3151).	123	21	2.55
Amy39116DRAFT_0147	methionine-R-sulfoxide reductase	41	8	2.36
Amy39116DRAFT_5013	Ferredoxin	71	14	2.34
Amy39116DRAFT_7112	transcription elongation factor GreA	160	33	2.28
Amy39116DRAFT_0700	FeS assembly ATPase SufC	113	25	2.18
Amy39116DRAFT_7763	3-oxoacid CoA-transferase, A subunit	9	2	2.17
Amy39116DRAFT_5829	Carboxylesterase type B	9	2	2.17
Amy39116DRAFT_3820	rhamnulose-1-phosphate aldolase/alcohol dehydrogenase	53	12	2.14
Amy39116DRAFT_7357	Long-chain acyl-CoA synthetases (AMP-forming)	13	3	2.12
Amy39116DRAFT_3459	deazaflavin-dependent nitroreductase family protein	30	7	2.10
Amy39116DRAFT_5803	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	17	4	2.09
Amy39116DRAFT_7642	3-oxoadipate enol-lactonase	17	4	2.09
Amy39116DRAFT_6870	Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II	24	6	2.00
Amy39116DRAFT_7939	Leucyl aminopeptidase	8	2	2.00
Amy39116DRAFT_2403	deazaflavin-dependent nitroreductase family protein	41	11	1.90
Amy39116DRAFT_6080	haloacid dehalogenase superfamily	22	6	1.87
Amy39116DRAFT_8090	ABC-type amino acid transpor systems, periplasmic component/domain	32	9	1.83
Amy39116DRAFT_6990	methionyl-tRNA synthetase	7	2	1.81
Amy39116DRAFT_5936	Citrate lyase beta subunit	7	2	1.81
Amy39116DRAFT_7553	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	34	10	1.77
Amy39116DRAFT_6885	Acetyltransferases	70	21	1.74
Amy39116DRAFT_6228	Predicted Zn-dependent hydrolases of the beta-lactamase fold	23	7	1.72
Amy39116DRAFT_7503	Phosphomannose isomerase	23	7	1.72
Amy39116DRAFT_4115	NADPH:quinone reductase and related Zn-dependent oxidoreductases	26	8	1.70
Amy39116DRAFT_8317	hypothetical protein	95	30	1.66
Amy39116DRAFT_0561	Zn-dependent hydrolases, including glyoxylases	19	6	1.66
Amy39116DRAFT_7287	Thiamine pyrophosphate-requiring enzymes	28	9	1.64
Amy39116DRAFT_8273	Methylase involved in ubiquinone/menaquinone biosynthesis	49	16	1.61
Amy39116DRAFT_8034	probable F420-dependent oxidoreductase, MSMEG_2906 family	48	16	1.58
Amy39116DRAFT_4436	Bacterial nucleoid DNA-binding protein	24	8	1.58
Amy39116DRAFT_4867	ABC-type Fe3+-hydroxamate transport system, periplasmic component	12	4	1.58
Amy39116DRAFT_4161	Asp-tRNAAsn/Glu-tRNAGln amidotransferase A subunit and related amidases	12	4	1.58
Amy39116DRAFT_6973	Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	9	3	1.58
Amy39116DRAFT_7659	Uncharacterized conserved protein	6	2	1.58
Amy39116DRAFT_821	L-rhamnose isomerase, Streptomyces subtype	53	18	1.56
Amy39116DRAFT_6724	phosphomethylpyrimidine kinase	35	12	1.54
Amy39116DRAFT_1835	Aldo/keto reductases, related to diketogulonate reductase	151	52	1.54
Amy39116DRAFT_5099	AIG2-like family.	61	22	1.47
Amy39116DRAFT_6632	hypothetical protein	8	3	1.42
Amy39116DRAFT_5856	Cytochrome P450	8	3	1.42
Amy39116DRAFT_7194	mannose-6-phosphate isomerase, class I	5	2	1.32
Amy39116DRAFT_4791	Dienelactone hydrolase and related enzymes	5	2	1.32
Amy39116DRAFT_6511	hypothetical protein	5	2	1.32
Amy39116DRAFT_7844	phenylacetate-CoA oxygenase/reductase, PaaK subunit	5	2	1.32
Amy39116DRAFT_5452	hypothetical protein	5	2	1.32
Amy39116DRAFT_6725	thiamine biosynthesis protein ThiC	17	7	1.28
Amy39116DRAFT_5312	ribosomal protein S7, bacterial/organelle	17	7	1.28
Amy39116DRAFT_0890	hypothetical proteinv	60	25	1.26
Amy39116DRAFT_2986	endoribonuclease L-PSP, putative	57	24	1.25
Amy39116DRAFT_7494	methylmalonyl-CoA epimerase	21	9	1.22
Amy39116DRAFT_4058	metalloprotein, YbeY/UPF0054 family	7	3	1.22
Amy39116DRAFT_8051	Uncharacterized protein conserved in bacteria	32	14	1.19
Amy39116DRAFT_6218	phosphoribosylformylglycinamidine synthase, purS protein	25	11	1.18
Amy39116DRAFT_4421	Transcriptional regulators	9	4	1.17
Amy39116DRAFT_7800	peptide deformylase	9	4	1.17
Amy39116DRAFT_6035	hypothetical protein	9	4	1.17
Amy39116DRAFT_6797	Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II	20	9	1.15
Amy39116DRAFT_4351	Predicted Zn-dependent hydrolases of the beta-lactamase fold	24	11	1.13
Amy39116DRAFT_4800	urocanate hydratase	13	6	1.12
Amy39116DRAFT_6618	Uncharacterized enzyme involved in biosynthesis of extracellular polysaccharides	200	95	1.07
Amy39116DRAFT_2098	Gluconolactonase	29	14	1.05
Amy39116DRAFT_5858	Rhodanese-related sulfurtransferase	37	18	1.04
Amy39116DRAFT_6020	thioredoxin	120	59	1.02
Amy39116DRAFT_6624	Carbonic anhydrase	42	21	1.00
Amy39116DRAFT_6770	hypothetical protein	10	5	1.00
Amy39116DRAFT_7297	Maleate cis-trans isomerase	10	5	1.00
Amy39116DRAFT_4946	Protein of unknown function (DUF3107)	6	3	1.00

Table A3.5. List of upregulated proteins from intracellular protein samples at 48 h by MUDPIT analysis. *Comparison of the log_2-fold ratio of transcripts from cells grown in the presence of dioxane-extracted lignin (DL) compared to the no lignin control after 48 h. ND = not determined.*

gene ID	gene annotation	control	lignin	log_2 ratio
Amy39116DRAFT_6163	5''-nucleotidase/2'',3''-cyclic phosphodiesterase and related esterases	0	140	ND
Amy39116DRAFT_6119	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	0	75	ND
Amy39116DRAFT_0626	gamma-glutamyltranspeptidase	0	65	ND
Amy39116DRAFT_4445	conserved repeat domain	0	50	ND
Amy39116DRAFT_3527	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	0	37	ND
Amy39116DRAFT_5051	hypothetical protein	0	28	ND
Amy39116DRAFT_5172	Domain of unknown function (DUF1918).	0	23	ND
Amy39116DRAFT_7437	acetate--CoA ligase	0	23	ND
Amy39116DRAFT_7442	MoxR-like ATPases	0	19	ND
Amy39116DRAFT_6182	Protein of unknown function (DUF1416).	0	17	ND
Amy39116DRAFT_3979	isoform II.	0	16	ND
Amy39116DRAFT_0285	hypothetical protein	0	16	ND
Amy39116DRAFT_5374	Methionyl-tRNA formyltransferase	0	16	ND
Amy39116DRAFT_5317	hypothetical protein	0	14	ND
Amy39116DRAFT_5235	Negative regulator of beta-lactamase expression	0	14	ND
Amy39116DRAFT_0484	Cold shock proteins	0	13	ND
Amy39116DRAFT_4296	Uncharacterized protein conserved in bacteria	0	12	ND
Amy39116DRAFT_6249	Glycerophosphoryl diester phosphodiesterase	0	11	ND
Amy39116DRAFT_3000	Poly(3-hydroxybutyrate) depolymerase	0	11	ND
Amy39116DRAFT_7128	ABC-type Fe3+-hydroxamate transport system, periplasmic component	0	10	ND
Amy39116DRAFT_6497	hypothetical protein	0	10	ND
Amy39116DRAFT_3524	DNA-binding ferritin-like protein (oxidative damage protectant)	0	10	ND
Amy39116DRAFT_5658	ABC-type Fe3+-hydroxamate transport system, periplasmic component	0	10	ND
Amy39116DRAFT_2504	Universal stress protein UspA and related nucleotide-binding proteins	0	10	ND
Amy39116DRAFT_1637	CBS-domain-containing membrane protein	0	9	ND
Amy39116DRAFT_3962	Aerobic-type carbon monoxide dehydrogenase, large subunit CoxL/CutL homologs	0	9	ND
Amy39116DRAFT_1438	conserved repeat domain	0	8	ND
Amy39116DRAFT_0510	Dihydrofolate reductase	0	8	ND
Amy39116DRAFT_5534	Uncharacterized protein conserved in bacteria	0	8	ND
Amy39116DRAFT_5365	non-ribosomal peptide synthase domain TIGR01720/amino acid adenylation domain	0	8	ND
Amy39116DRAFT_6050	hypothetical protein	0	8	ND
Amy39116DRAFT_7220	hypothetical protein	0	7	ND
Amy39116DRAFT_4002	ABC-type Fe3+-hydroxamate transport system, periplasmic component	0	7	ND
Amy39116DRAFT_0622	S-adenosylmethionine synthetase	0	7	ND
Amy39116DRAFT_2512	CBS-domain-containing membrane protein	0	7	ND
Amy39116DRAFT_6214	phosphoribosylformylglycinamidine synthase II	0	7	ND
Amy39116DRAFT_7233	Fructose-2,6-bisphosphatase	0	7	ND
Amy39116DRAFT_7556	Cobalamin biosynthesis protein CobN and related Mg-chelatases	0	7	ND
Amy39116DRAFT_7005	DivIVA domain	0	6	ND
Amy39116DRAFT_1538	conserved repeat domain	0	6	ND
Amy39116DRAFT_6852	Imidazolonepropionase and related amidohydrolases	0	6	ND
Amy39116DRAFT_1663	Protein-tyrosine-phosphatase	0	6	ND
Amy39116DRAFT_6583	hypoxanthine phosphoribosyltransferase	0	6	ND
Amy39116DRAFT_6483	aspartate-semialdehyde dehydrogenase (peptidoglycan organisms)	0	6	ND
Amy39116DRAFT_4266	ATP-binding cassette protein, ChvD family	0	6	ND
Amy39116DRAFT_6042	leucyl-tRNA synthetase, eubacterial and mitochondrial family	0	6	ND
Amy39116DRAFT_7348	hypothetical protein	0	6	ND
Amy39116DRAFT_2170	Beta-glucosidase-related glycosidases	0	6	ND
Amy39116DRAFT_1982	ATP phosphoribosyltransferase	0	6	ND
Amy39116DRAFT_2007	Uncharacterized conserved protein	0	6	ND
Amy39116DRAFT_7733	Ricin-type beta-trefoil lectin domain.	0	6	ND
Amy39116DRAFT_5107	Phosphomannomutase	0	5	ND
Amy39116DRAFT_5040	Acyl-CoA dehydrogenases	0	5	ND
Amy39116DRAFT_4799	amidase, hydantoinase/carbamoylase family	0	5	ND
Amy39116DRAFT_6972	N-acetyl-beta-hexosaminidase	0	5	ND
Amy39116DRAFT_6593	Predicted transcriptional regulators	0	5	ND
Amy39116DRAFT_5140	adenosine deaminase	0	5	ND
Amy39116DRAFT_4288	Secreted trypsin-like serine protease	0	5	ND
Amy39116DRAFT_0734	Isopenicillin N synthase and related dioxygenases	0	5	ND
Amy39116DRAFT_7877	Subtilisin-like serine proteases	0	5	ND

Amy39116DRAFT_2977 Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	0	5	ND
Amy39116DRAFT_3844 hypothetical protein	0	5	ND
Amy39116DRAFT_2798 Predicted redox protein, regulator of disulfide bond formation	0	5	ND
Amy39116DRAFT_0573 aspartyl-tRNA synthetase, bacterial type	0	5	ND
Amy39116DRAFT_2498 hypothetical protein	0	5	ND
Amy39116DRAFT_2517 Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	0	5	ND
Amy39116DRAFT_5452 hypothetical protein	0	5	ND
Amy39116DRAFT_5453 Geranylgeranyl pyrophosphate synthase	0	5	ND
Amy39116DRAFT_2530 MoxR-like ATPases	0	5	ND
Amy39116DRAFT_4182 Predicted dioxygenase of extradiol dioxygenase family	0	5	ND
Amy39116DRAFT_5366 Uncharacterized protein conserved in bacteria	0	5	ND
Amy39116DRAFT_3377 Uncharacterized conserved protein	0	5	ND
Amy39116DRAFT_7243 Uncharacterized protein conserved in bacteria	0	5	ND
Amy39116DRAFT_7357 Long-chain acyl-CoA synthetases (AMP-forming)	0	5	ND
Amy39116DRAFT_2040 Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide acyltransferase (E2) component, and related enzymes	0	5	ND
Amy39116DRAFT_2051 Transglycosylase-like domain./Putative peptidoglycan binding domain.	0	5	ND
Amy39116DRAFT_2087 Dihydrodipicolinate synthase/N-acetylneuraminate lyase	0	5	ND
Amy39116DRAFT_7655 Xanthine and CO dehydrogenases maturation factor, XdhC/CoxF family	0	5	ND
Amy39116DRAFT_4964 Acyl dehydratase	0	4	ND
Amy39116DRAFT_6744 Cold shock proteins	4	88	4.46
Amy39116DRAFT_5362 ABC-type Fe3+-hydroxamate transport system, periplasmic component	5	90	4.17
Amy39116DRAFT_2316 Antibiotic biosynthesis monooxygenase.	2	23	3.52
Amy39116DRAFT_2073 Predicted enzyme related to lactoylglutathione lyase	2	23	3.52
Amy39116DRAFT_7439 NAD-dependent aldehyde dehydrogenases	11	93	3.08
Amy39116DRAFT_6115 Phosphoenolpyruvate carboxykinase (GTP)	8	55	2.78
Amy39116DRAFT_6572 hypothetical protein	5	32	2.68
Amy39116DRAFT_6205 Protein of unknown function (DUF574).	2	11	2.46
Amy39116DRAFT_4457 Subtilase family.	16	83	2.38
Amy39116DRAFT_5265 WXG100 family type VII secretion target	12	61	2.35
Amy39116DRAFT_7058 NAD-dependent aldehyde dehydrogenases	2	10	2.32
Amy39116DRAFT_3924 Homogentisate 1,2-dioxygenase	4	18	2.17
Amy39116DRAFT_4463 ketol-acid reductoisomerase	2	9	2.17
Amy39116DRAFT_4172 valyl-tRNA synthetase	2	9	2.17
Amy39116DRAFT_6482 aspartate kinase, monofunctional class	6	23	1.94
Amy39116DRAFT_6181 Rhodanese-related sulfurtransferase	6	23	1.94
Amy39116DRAFT_5148 Uncharacterized ABC-type transport system, periplasmic component/surface lipoprotein	7	26	1.89
Amy39116DRAFT_6418 Cold shock proteins	33	121	1.87
Amy39116DRAFT_5168 isocitrate dehydrogenase, NADP-dependent, eukaryotic type	5	18	1.85
Amy39116DRAFT_4527 isocitrate lyase	4	14	1.81
Amy39116DRAFT_5038 TIGR03089 family protein	2	7	1.81
Amy39116DRAFT_4565 Domain of unknown function (DUF2017)	2	7	1.81
Amy39116DRAFT_2534 hypothetical protein	2	7	1.81
Amy39116DRAFT_8167 Aerobic-type carbon monoxide dehydrogenase, large subunit CoxL/CutL homologs	7	24	1.78
Amy39116DRAFT_2161 Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase and related deacylases	3	10	1.74
Amy39116DRAFT_6409 Cold shock proteins	27	87	1.69
Amy39116DRAFT_7011 UDP-N-acetylglucosamine diphosphorylase/glucosamine-1-phosphate N-acetyltransferase	2	6	1.58
Amy39116DRAFT_2516 Enoyl-[acyl-carrier-protein] reductase (NADH)	2	6	1.58
Amy39116DRAFT_7933 2-oxoglutarate dehydrogenase, E2 component, dihydrolipoamide succinyltransferase	10	29	1.54
Amy39116DRAFT_4194 Protease subunit of ATP-dependent Clp proteases	6	17	1.50
Amy39116DRAFT_3239 Protein tyrosine/serine phosphatase	4	11	1.46
Amy39116DRAFT_4561 Predicted metal-dependent protease of the PAD1/JAB1 superfamily	10	26	1.38
Amy39116DRAFT_4458 D-3-phosphoglycerate dehydrogenase	7	18	1.36
Amy39116DRAFT_6541 hypothetical protein	10	25	1.32
Amy39116DRAFT_0442 ABC-type Fe3+-hydroxamate transport system, periplasmic component	6	15	1.32
Amy39116DRAFT_3819 L-arabinose isomerase	4	10	1.32
Amy39116DRAFT_5098 Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide dehydrogenase (E3) component, and related enzymes	2	5	1.32
Amy39116DRAFT_7143 fructose-1,6-bisphosphatase, class II	2	5	1.32
Amy39116DRAFT_0698 FeS assembly protein SufB	2	5	1.32
Amy39116DRAFT_3494 Uncharacterized conserved protein	2	5	1.32
Amy39116DRAFT_6004 DNA polymerase III, beta subunit	2	5	1.32
Amy39116DRAFT_7985 Predicted amidohydrolase	6	14	1.22
Amy39116DRAFT_5901 hypothetical protein	3	7	1.22
Amy39116DRAFT_6820 Superoxide dismutase	15	34	1.18

Amy39116DRAFT_7045	Exopolyphosphatase	4	9	1.17
Amy39116DRAFT_3032	Dipeptidyl aminopeptidases/acylaminoacyl-peptidases	5	11	1.14
Amy39116DRAFT_5507	Glutaredoxin and related proteins	21	45	1.10
Amy39116DRAFT_2741	hypothetical protein	8	17	1.09
Amy39116DRAFT_0222	glutamine synthetase, type I	8	17	1.09
Amy39116DRAFT_7109	4-hydroxyphenylpyruvate dioxygenase	9	18	1.00
Amy39116DRAFT_6579	Inorganic pyrophosphatase	6	12	1.00
Amy39116DRAFT_6501	DNA-binding ferritin-like protein (oxidative damage protectant)	6	12	1.00
Amy39116DRAFT_7147	hypothetical protein	3	6	1.00
Amy39116DRAFT_8038	Lactoylglutathione lyase and related lyases	3	6	1.00
Amy39116DRAFT_3020	Protein of unknown function (DUF664)	3	6	1.00

Table A3.6. List of downregulated proteins from intracellular protein samples at 48 h by MUDPIT analysis. *Comparison of the log_2-fold ratio of transcripts from cells grown in the presence of dioxane-extracted lignin (DL) compared to the no lignin control after 48 h. ND = not determined.*

gene ID	gene annotation	control	lignin	log_2 ratio
Amy39116DRAFT_5396	Uncharacterized protein conserved in bacteria	20	0	ND
Amy39116DRAFT_0147	methionine-R-sulfoxide reductase	19	0	ND
Amy39116DRAFT_4623	F0F1-type ATP synthase, subunit a	12	0	ND
Amy39116DRAFT_3653	Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	12	0	ND
Amy39116DRAFT_0608	translation elongation factor P	12	0	ND
Amy39116DRAFT_4442	hypothetical protein	11	0	ND
Amy39116DRAFT_0701	cysteine desulfurases, SufS subfamily	10	0	ND
Amy39116DRAFT_5851	Ferritin-like protein	10	0	ND
Amy39116DRAFT_5883	pyruvate dehydrogenase E1 component, alpha subunit	9	0	ND
Amy39116DRAFT_2359	Uncharacterized conserved protein	9	0	ND
Amy39116DRAFT_6724	phosphomethylpyrimidine kinase	8	0	ND
Amy39116DRAFT_4255	Domain of unknown function (DUF477).	8	0	ND
Amy39116DRAFT_3413	Uncharacterized conserved protein	8	0	ND
Amy39116DRAFT_5980	hypothetical protein	8	0	ND
Amy39116DRAFT_2982	ribosomal protein L20	8	0	ND
Amy39116DRAFT_2323	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	8	0	ND
Amy39116DRAFT_4621	ATP synthase, F0 subunit b	7	0	ND
Amy39116DRAFT_5700	Transcriptional regulators	7	0	ND
Amy39116DRAFT_5884	Pyruvate/2-oxoglutarate dehydrogenase complex, dehydrogenase (E1) component, eukaryotic type, beta subunit	7	0	ND
Amy39116DRAFT_6066	ribosomal protein L9	7	0	ND
Amy39116DRAFT_6879	Enoyl-CoA hydratase/carnithine racemase	6	0	ND
Amy39116DRAFT_4568	Amidases related to nicotinamidase	6	0	ND
Amy39116DRAFT_4545	Peroxiredoxin	6	0	ND
Amy39116DRAFT_4690	N-methylhydantoinase A/acetone carboxylase, beta subunit	6	0	ND
Amy39116DRAFT_0745	Acyl dehydratase	6	0	ND
Amy39116DRAFT_7947	Iron-sulfur cluster assembly accessory protein	6	0	ND
Amy39116DRAFT_3492	Glycine/D-amino acid oxidases (deaminating)	6	0	ND
Amy39116DRAFT_3136	peroxiredoxin, OsmC subfamily	6	0	ND
Amy39116DRAFT_7764	3-oxoacid CoA-transferase, B subunit	6	0	ND
Amy39116DRAFT_7763	3-oxoacid CoA-transferase, A subunit	6	0	ND
Amy39116DRAFT_5673	copper ion binding protein	6	0	ND
Amy39116DRAFT_5952	glycosyltransferase, MGT family	6	0	ND
Amy39116DRAFT_8111	ribonucleoside-diphosphate reductase, adenosylcobalamin-dependent	6	0	ND
Amy39116DRAFT_5885	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide acyltransferase (E2) component, and related enzymes	6	0	ND
Amy39116DRAFT_5305	ribosomal protein L2, bacterial/organellar	6	0	ND
Amy39116DRAFT_5353	Sporulation control protein	6	0	ND
Amy39116DRAFT_5003	TIGR00730 family protein	5	0	ND
Amy39116DRAFT_5056	hypothetical protein	5	0	ND
Amy39116DRAFT_6943	large conductance mechanosensitive channel protein	5	0	ND
Amy39116DRAFT_6940	Trypsin-like serine proteases, typically periplasmic, contain C-terminal PDZ domain	5	0	ND
Amy39116DRAFT_6388	hypothetical protein	5	0	ND
Amy39116DRAFT_7838	Enoyl-CoA hydratase/carnithine racemase	5	0	ND
Amy39116DRAFT_7841	phenylacetate-CoA oxygenase, PaaH subunit	5	0	ND
Amy39116DRAFT_2950	3-phosphoshikimate 1-carboxyvinyltransferase	5	0	ND
Amy39116DRAFT_4147	ribosomal protein L21	5	0	ND
Amy39116DRAFT_0416	Aerobic-type carbon monoxide dehydrogenase, large subunit CoxL/CutL homologs	5	0	ND
Amy39116DRAFT_7670	Predicted O-methyltransferase	5	0	ND
Amy39116DRAFT_1981	phosphoribosyl-ATP pyrophosphohydrolase	5	0	ND
Amy39116DRAFT_1989	hypothetical protein	5	0	ND
Amy39116DRAFT_2229	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide acyltransferase (E2) component, and related enzymes	52	6	3.12
Amy39116DRAFT_4315	ribosome recycling factor	16	2	3.00
Amy39116DRAFT_6797	Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II	30	4	2.91
Amy39116DRAFT_5013	Ferredoxin	25	4	2.64
Amy39116DRAFT_6493	Catalase	12	2	2.58
Amy39116DRAFT_6460	DNA-binding protein, YbaB/EbfC family	40	7	2.51
Amy39116DRAFT_7265	Acetyl/propionyl-CoA carboxylase, alpha subunit	40	7	2.51
Amy39116DRAFT_6870	Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II	16	3	2.42
Amy39116DRAFT_5810	Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	20	4	2.32

Amy39116DRAFT_6175 phosphate ABC transporter, phosphate-binding protein	10	2	2.32
Amy39116DRAFT_0020 hypothetical protein	19	4	2.25
Amy39116DRAFT_0618 hypothetical protein	14	3	2.22
Amy39116DRAFT_4365 deazaflavin-dependent nitroreductase family protein	9	2	2.17
Amy39116DRAFT_6511 hypothetical protein	13	3	2.12
Amy39116DRAFT_7494 methylmalonyl-CoA epimerase	20	5	2.00
Amy39116DRAFT_3459 deazaflavin-dependent nitroreductase family protein	12	3	2.00
Amy39116DRAFT_6516 Zn-dependent hydrolases, including glyoxylases	8	2	2.00
Amy39116DRAFT_5331 ribosomal protein L7/L12	39	10	1.96
Amy39116DRAFT_4588 dihydroxyacetone kinase, DhaK subunit	15	4	1.91
Amy39116DRAFT_7266 Acyl-CoA dehydrogenases	21	6	1.81
Amy39116DRAFT_4491 Zn-dependent hydrolases, including glyoxylases	7	2	1.81
Amy39116DRAFT_0439 Methylase involved in ubiquinone/menaquinone biosynthesis	7	2	1.81
Amy39116DRAFT_7613 Glutamate dehydrogenase/leucine dehydrogenase	7	2	1.81
Amy39116DRAFT_5152 succinate dehydrogenase and fumarate reductase iron-sulfur protein	10	3	1.74
Amy39116DRAFT_4402 Cell division initiation protein	10	3	1.74
Amy39116DRAFT_7953 response regulator receiver protein	10	3	1.74
Amy39116DRAFT_5503 porphobilinogen deaminase	10	3	1.74
Amy39116DRAFT_6463 ABC-type dipeptide transport system, periplasmic component	16	5	1.68
Amy39116DRAFT_6218 phosphoribosylformylglycinamidine synthase, purS protein	36	12	1.58
Amy39116DRAFT_7927 probable F420-dependent oxidoreductase, Rv1855c family	15	5	1.58
Amy39116DRAFT_0488 Protein of unknown function (DUF2630).	12	4	1.58
Amy39116DRAFT_5296 Ribosomal protein L5	6	2	1.58
Amy39116DRAFT_7553 Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	26	9	1.53
Amy39116DRAFT_7888 pyruvate dehydrogenase E1 component, homodimeric type	20	7	1.51
Amy39116DRAFT_0748 Predicted oxidoreductases (related to aryl-alcohol dehydrogenases)	22	8	1.46
Amy39116DRAFT_4254 aminopeptidase N, Streptomyces lividans type	11	4	1.46
Amy39116DRAFT_6326 chaperone protein DnaK	60	22	1.45
Amy39116DRAFT_7889 Protein of unknown function (DUF3052).	27	10	1.43
Amy39116DRAFT_5339 Aspartate/tyrosine/aromatic aminotransferase	35	13	1.43
Amy39116DRAFT_5232 Co-chaperonin GroES (HSP10)	40	15	1.42
Amy39116DRAFT_5151 succinate dehydrogenase, flavoprotein subunit, E. coli/mitochondrial subgroup	8	3	1.42
Amy39116DRAFT_6325 Molecular chaperone GrpE (heat shock protein)	8	3	1.42
Amy39116DRAFT_0887 Lactoylglutathione lyase and related lyases	8	3	1.42
Amy39116DRAFT_0873 Uncharacterized conserved protein	8	3	1.42
Amy39116DRAFT_7883 (acyl-carrier-protein) S-malonyltransferase	8	3	1.42
Amy39116DRAFT_7659 Uncharacterized conserved protein	8	3	1.42
Amy39116DRAFT_6263 Protein of unknown function (DUF3151).	21	8	1.39
Amy39116DRAFT_6270 Uncharacterized protein conserved in bacteria	21	8	1.39
Amy39116DRAFT_0700 FeS assembly ATPase SufC	21	8	1.39
Amy39116DRAFT_2887 delta-1-pyrroline-5-carboxylate dehydrogenase, group 1	62	24	1.37
Amy39116DRAFT_5659 Uncharacterized conserved protein	43	17	1.34
Amy39116DRAFT_7273 Lactoylglutathione lyase and related lyases	10	4	1.32
Amy39116DRAFT_6316 ATP-dependent chaperone ClpB	5	2	1.32
Amy39116DRAFT_3790 heme/flavin dehydrogenase, mycofactocin system	5	2	1.32
Amy39116DRAFT_2736 DNA-binding regulatory protein, YebC/PmpR family	5	2	1.32
Amy39116DRAFT_4058 metalloprotein, YbeY/UPF0054 family	5	2	1.32
Amy39116DRAFT_5550 Uncharacterized conserved protein	5	2	1.32
Amy39116DRAFT_3873 Autotransporter adhesin	5	2	1.32
7264 Acetyl-CoA carboxylase, carboxyltransferase component (subunits alpha and Amy39116DRAFT_beta)	17	7	1.28
Amy39116DRAFT_6329 Raf kinase inhibitor-like protein, YbhB/YbcL family	48	20	1.26
Amy39116DRAFT_0938 Uncharacterized conserved protein	12	5	1.26
Amy39116DRAFT_7004 Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II	7	3	1.22
Amy39116DRAFT_6248 phosphoribosylamine--glycine ligase	7	3	1.22
Amy39116DRAFT_7800 peptide deformylase	7	3	1.22
Amy39116DRAFT_0504 Putative translation initiation inhibitor, yjgF family	7	3	1.22
Amy39116DRAFT_5918 Uncharacterized protein conserved in bacteria	67	29	1.21
Amy39116DRAFT_4195 trigger factor	16	7	1.19
Amy39116DRAFT_0656 glyceraldehyde-3-phosphate dehydrogenase, type I	59	26	1.18
Amy39116DRAFT_7007 ribosomal protein L25, Ctc-form	11	5	1.14
6618 Uncharacterized enzyme involved in biosynthesis of extracellular Amy39116DRAFT_polysaccharides	89	41	1.12
5078 Acetyl-CoA carboxylase, carboxyltransferase component (subunits alpha and Amy39116DRAFT_beta)	26	12	1.12
Amy39116DRAFT_3493 Transcription elongation factor	70	34	1.04
Amy39116DRAFT_7890 Peroxiredoxin	49	24	1.03
Amy39116DRAFT_6474 2-isopropylmalate synthase, yeast type	14	7	1.00
Amy39116DRAFT_6885 Acetyltransferases	12	6	1.00
Amy39116DRAFT_6720 thiamine biosynthesis protein ThiS	10	5	1.00

Amy39116DRAFT_5307 50S ribosomal protein L4, bacterial/organelle	10	5	1.00
Amy39116DRAFT_3923 hypothetical protein	8	4	1.00
Amy39116DRAFT_7842 phenylacetate-CoA oxygenase, PaaI subunit	6	3	1.00
Amy39116DRAFT_5844 hypothetical protein	6	3	1.00

Supplementary results: *NMR data analysis*

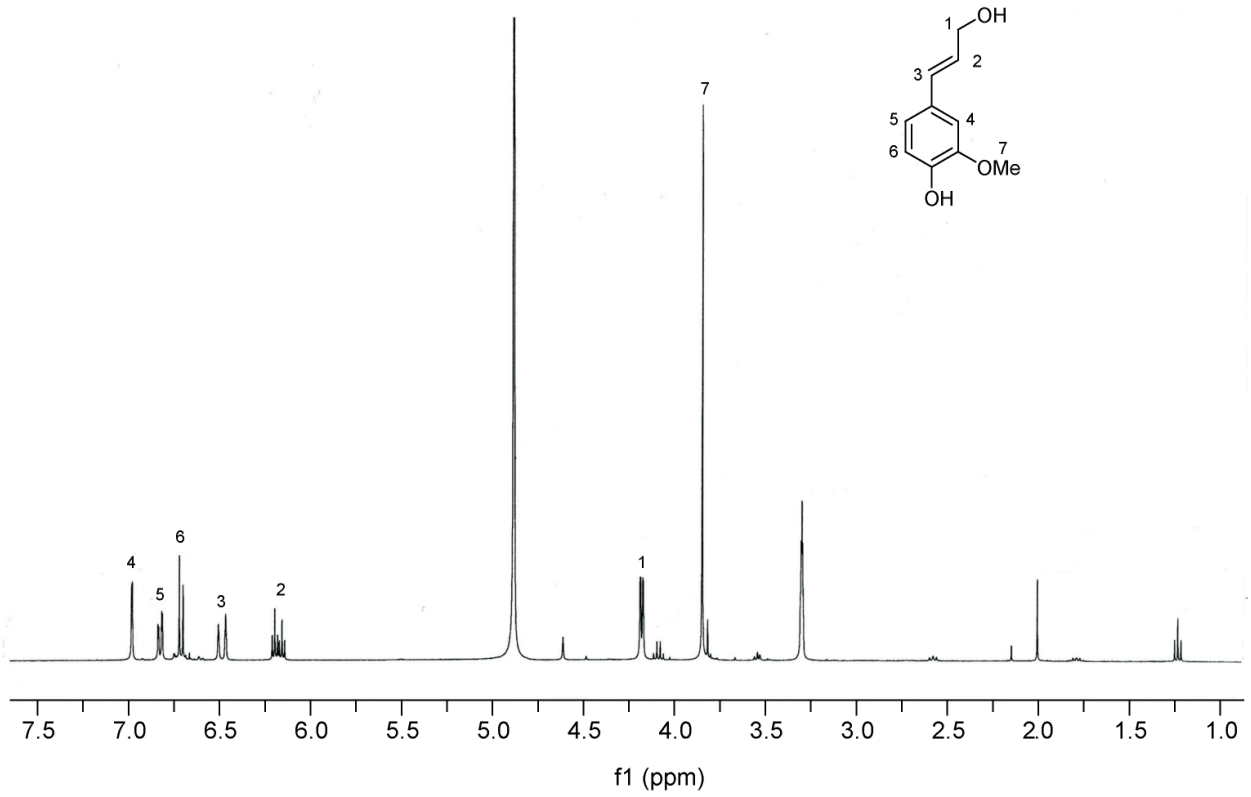


Figure A3.1 NMR spectrum of unlabeled coniferyl alcohol in methanol-d_4. *^{1}H NMR spectrum acquired on an AVQ-400.*

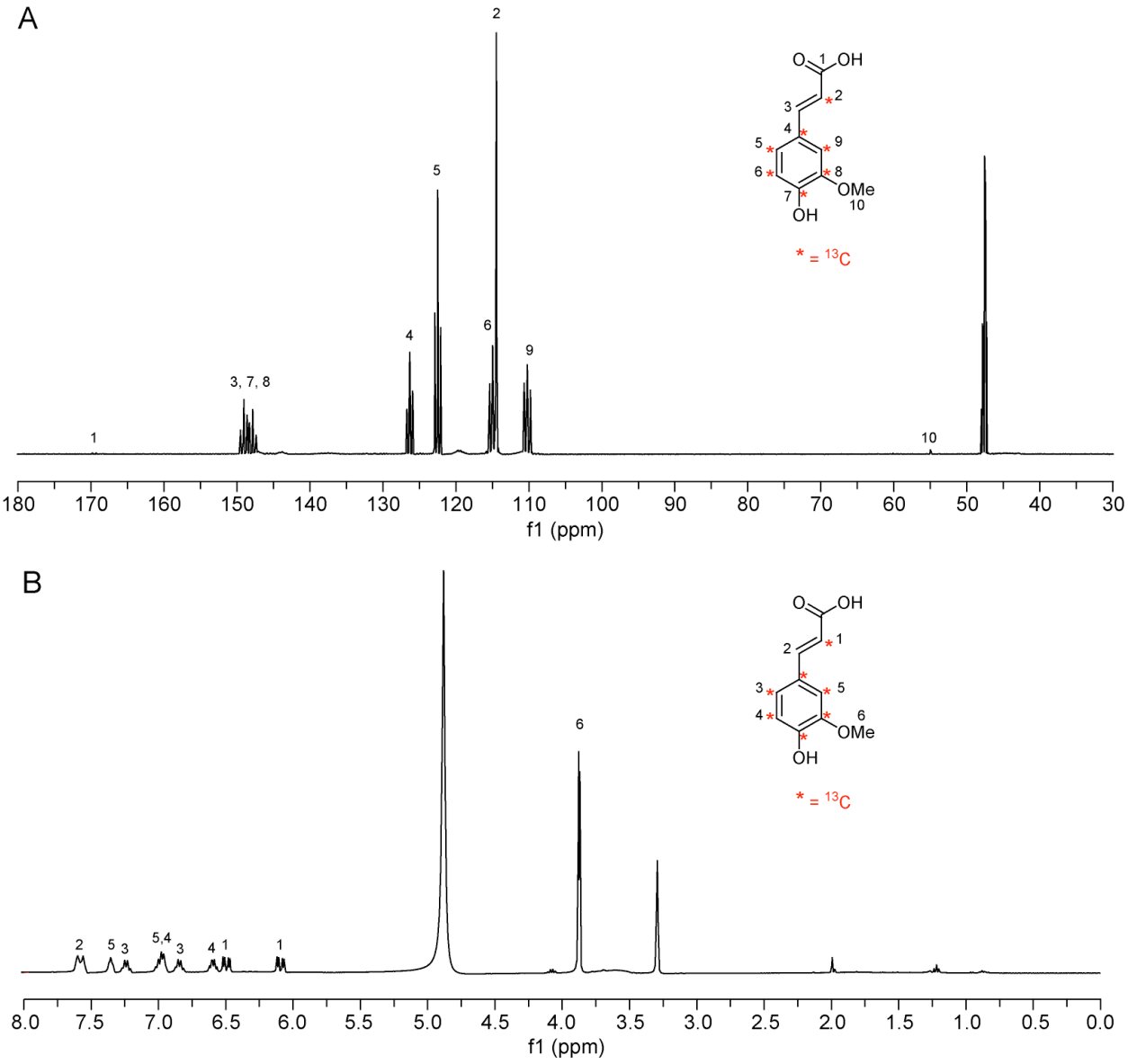


Figure A3.2. NMR spectra of ^{13}C-labeled ferulic acid in methanol-d$_4$. *A) ^{1}H NMR spectrum acquired on an AVQ-400 B) ^{13}C NMR spectrum acquired on an AV-600.*

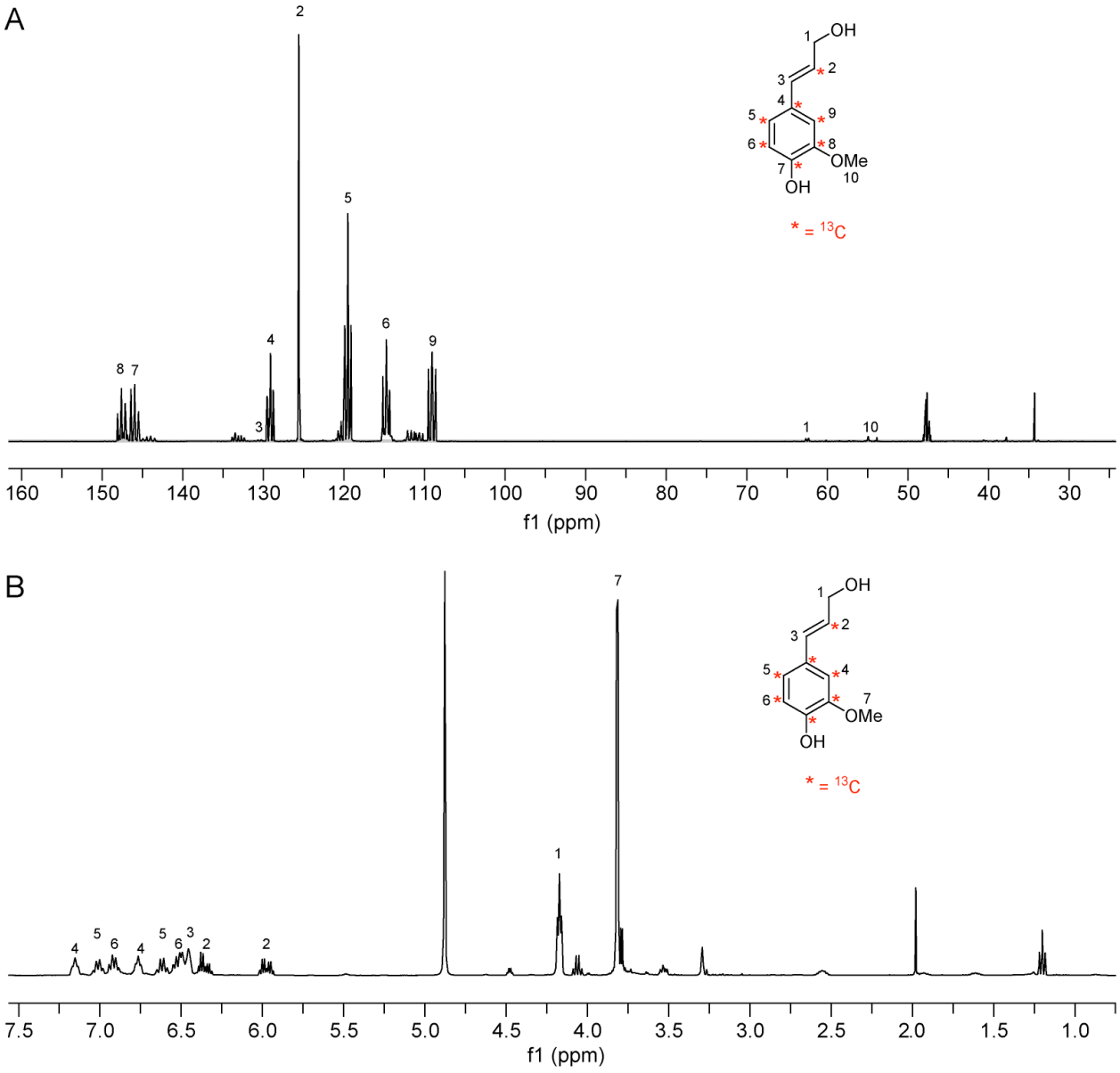


Figure A3.3. NMR spectra of ^{13}C-labeled coniferyl alcohol in methanol-d$_4$. *A) ^{1}H NMR spectrum acquired on an AVQ-400 B) ^{13}C NMR spectrum acquired on an AV-600.*

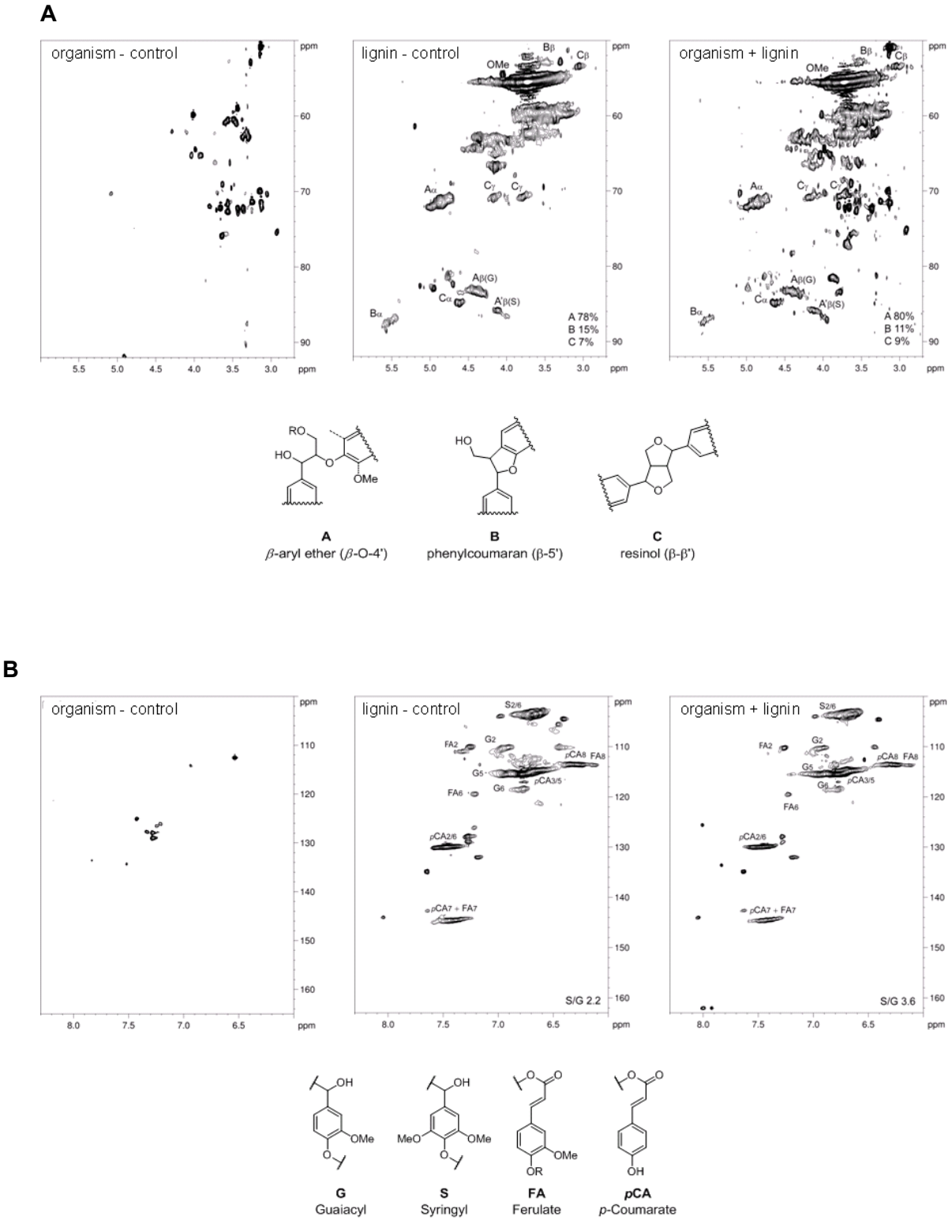


Figure A3.4. $^{1}H/^{13}C$ HSQC of the solid dioxane-extracted lignin. *Comparison of lignin incubated with and without A. sp. 75iv2 shows (A) minor changes in bond linkage percentages such as the phenylcoumaran linkage (B) as well as (B) a change in the S:G ratio due to the presence of A. sp. 75iv2.*

		ISTD	H		G		S		S:G Ratio
	Replicate	Area	Area	%	Area	%	Area	%	
with lignin	A	70839	1305	0.42	96717	31.1	212998	68.5	2.20
	B	70317	1039	0.38	84122	30.6	189467	69.0	2.25
no lignin	A	70006	3503	0.38	283560	30.9	630853	68.7	2.22
	B	77090	5450	0.42	406442	31.5	877615	68.1	2.16

Table A3.7. GC-MS analysis of lignin after thioacidolysis of dioxane-extracted lignin. *Peak areas representing the three primary lignin units (H, G, and S) were used to calculate the S:G ratios which approximate to the same values, revealing little difference in the lignin units due to the presence of A. sp. 75iv2.*

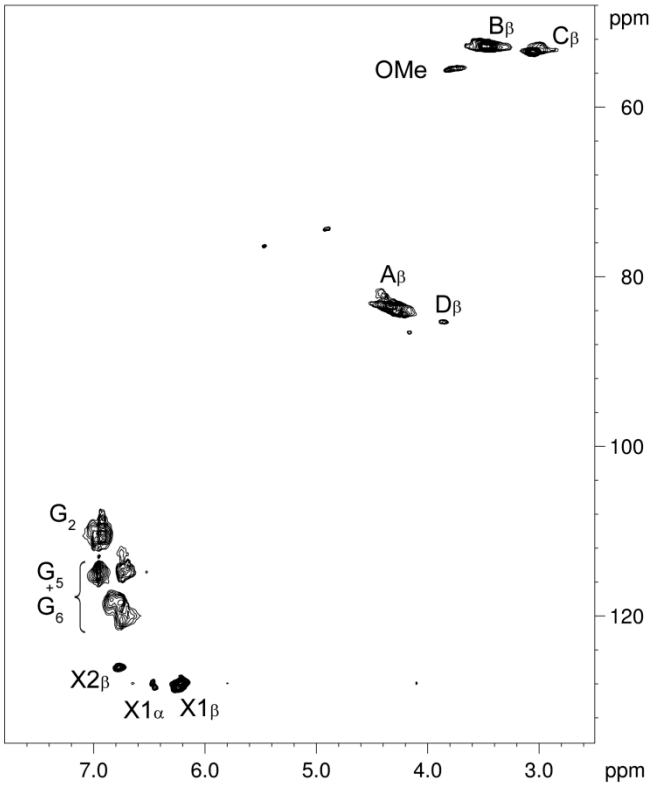


Figure A3.5. $^{1}H/^{13}C$ HSQC of the ^{13}C-labeled DHP in DMSO-d_6 after monomer removal. *The spectrum is labeled as follows: (A) β-aryl ether, (B) phenylcoumaran, (C) resinol, (D) dibenzodioxocin, (G) guaiacyl, (X1) cinnamyl alcohol, (X2) cinnamyl aldehyde, (OMe) unlabeled methoxy groups.*

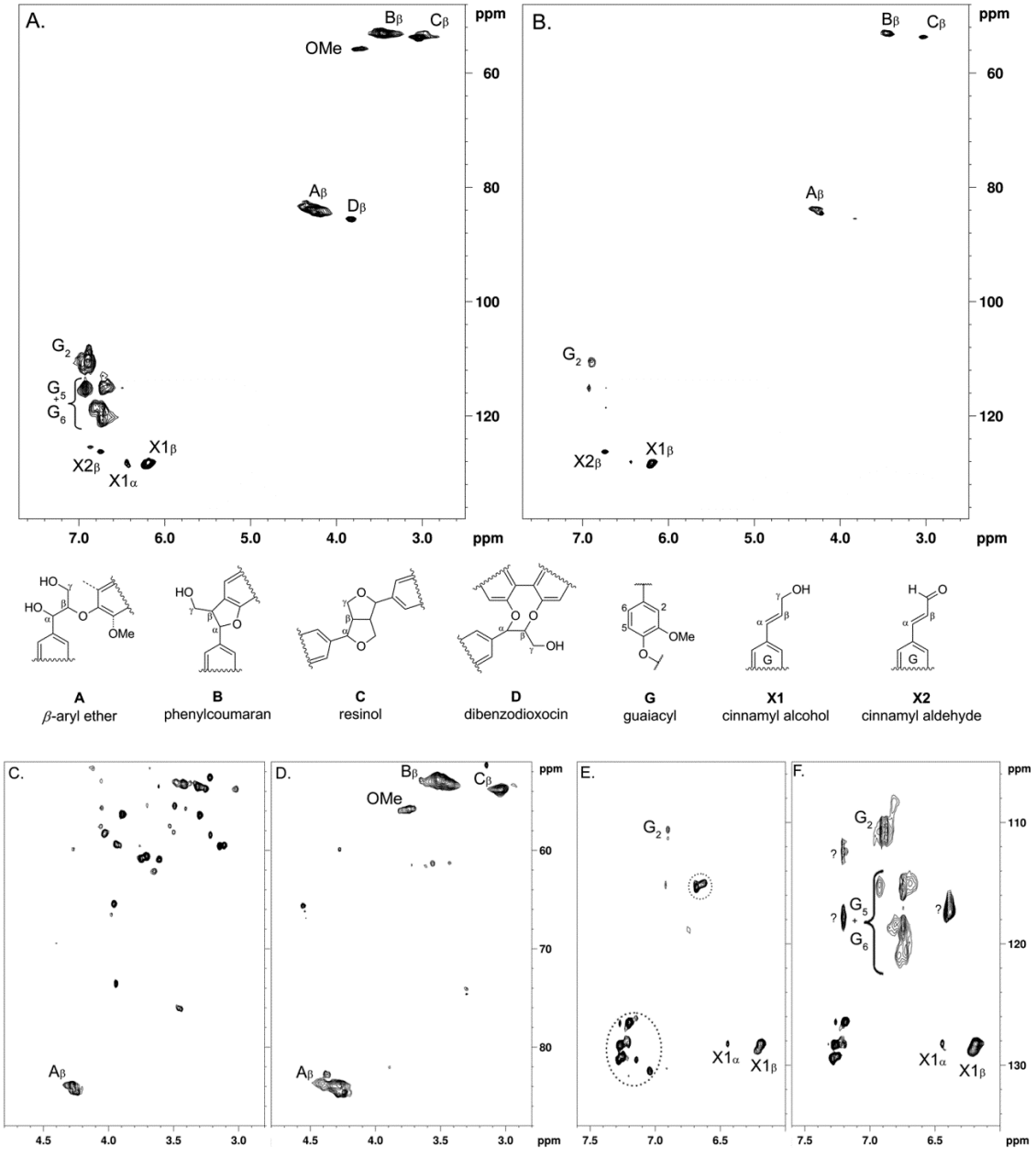
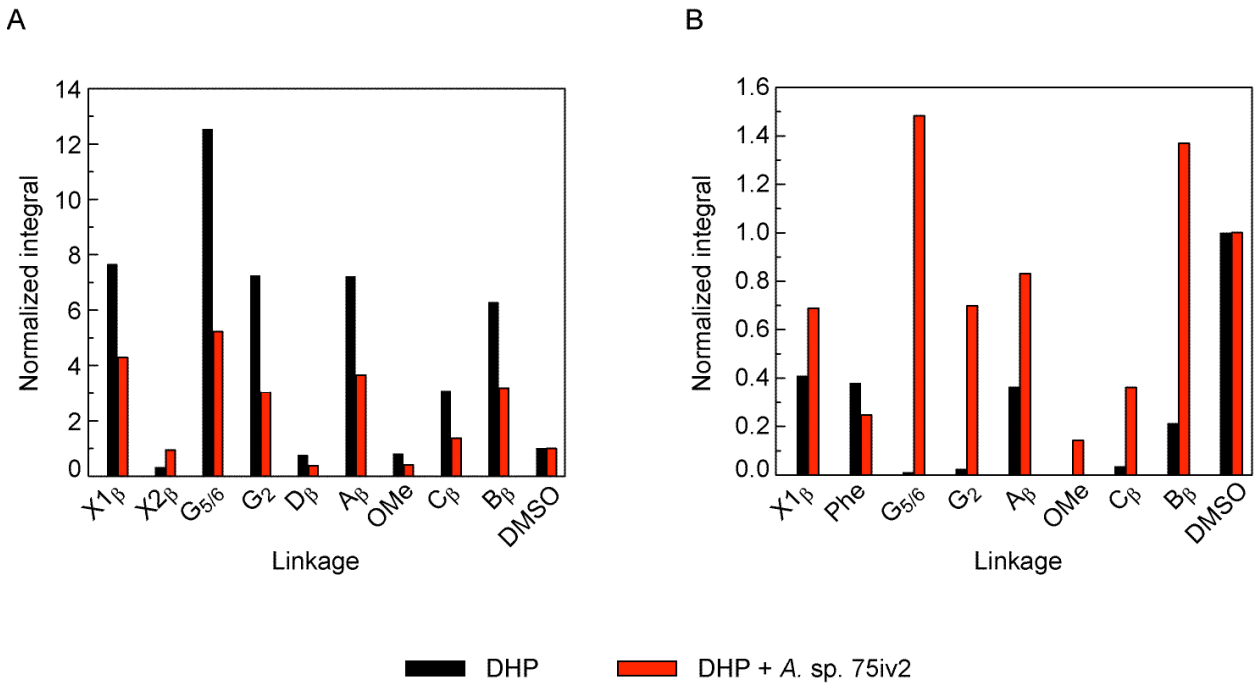


Figure A3.6. $^{1}H/^{13}C$ HSQC spectra of ^{13}C-labeled DHP from SMM samples. *The solid fraction of ^{13}C-labeled DHP in SMM (A) without and (B) with bacterial growth. Partial HSQC spectra of the supernatant fractions;^{13}C-labeled DHP in SMM (C + E) without and (D+F) with bacteria growth. Aromatic amino acids remaining in the media after the month-long growth are circled.*



C

solid fraction		supernatant fraction	
bonding motif	-fold change	bonding motif	-fold change
$X1_{\beta}$	0.56	$X1_{\beta}$	1.68
$X2_{\beta}$	3.03	Phe	0.65
$G_{5/6}$	0.42	$G_{5/6}$	149.71
G_2	0.42	G_2	29.29
D_{β}	0.50	A_{β}	2.28
A_{β}	0.51	OMe	ND
OMe	0.50	C_{β}	10.20
C_{β}	0.45	B_{β}	6.45
B_{β}	0.51		
DMSO	1.00	DMSO	1.00

Figure A3.7. Comparison of $^{1}H/^{13}C$ HSQC normalized spectral integrals for different bonding motifs of ^{13}C-labeled DHP from SMM samples. *(A) The solid fraction of ^{13}C-labeled DHP growths in SMM without (black) and (red) with bacterial growth. (B) The supernatant fraction of ^{13}C-labeled DHP growths in SMM without (black) and (red) with bacterial growth. The aromatic amino acid phenylalanine, Phe, which was found remaining in the media after the month-long growth is also quantified. (C) Quantification of – fold change of spectral integrals of bonding motifs.*

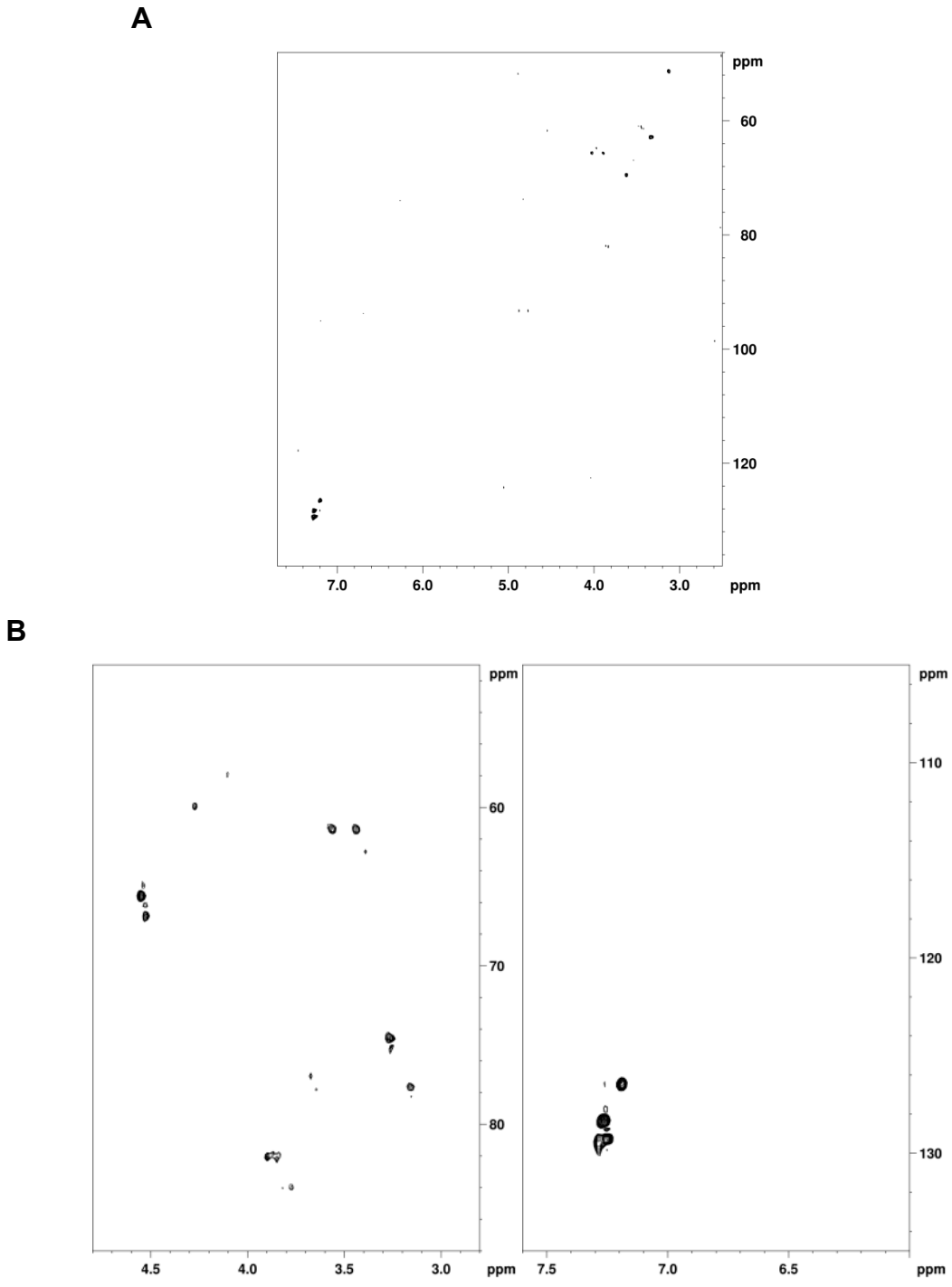


Figure A3.8. $^{1}H/^{13}C$ HSQC monitoring A. sp. 75i v2 metabolite profile in SMM in the absence of DHP. *(A) Cell pellet. (B) Supernatant. As controls, cells were grown in SMM for 28 d along-side the lignin growths. Solid cells and supernatant were separated and lyophilized exactly as the samples of lignin growths. Solvation in DMSO-d_6 was accomplished in the same fashion but with only one round of solvation, vortexing, and centrifugation to remove insoluble matter.*

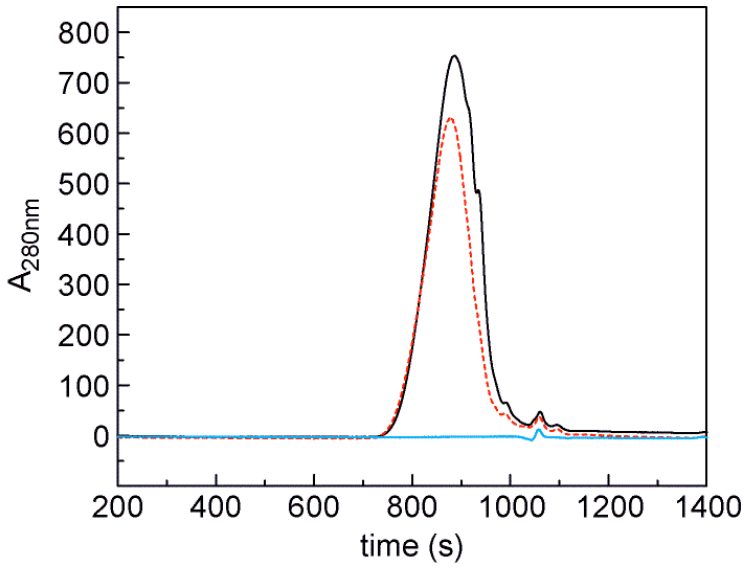


Figure A3.9. Gel permeation chromatography of solid ^{13}C-labeled DHP from growths. *GPC of solid fractions from 28 d growths (DHP control, black; DHP + culture, dashed red; culture, blue). Analysis shows limited size difference that include: less overall DHP as well as some modifications of lower molecular weight fractions, with a slight overall increase in M_p due to the presence of the culture.*

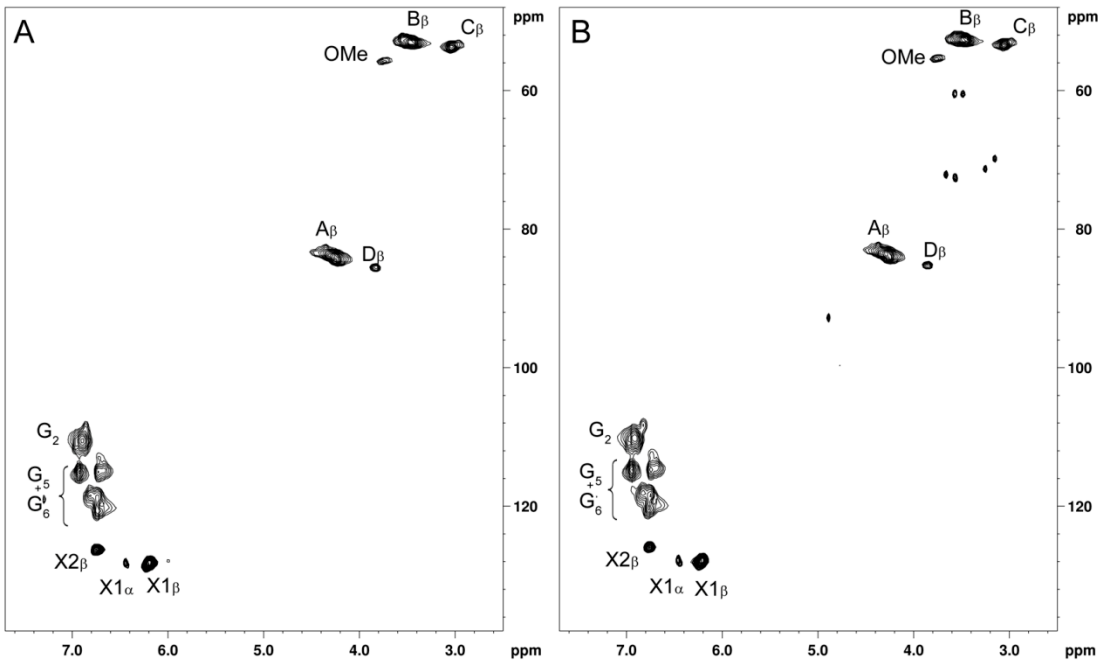
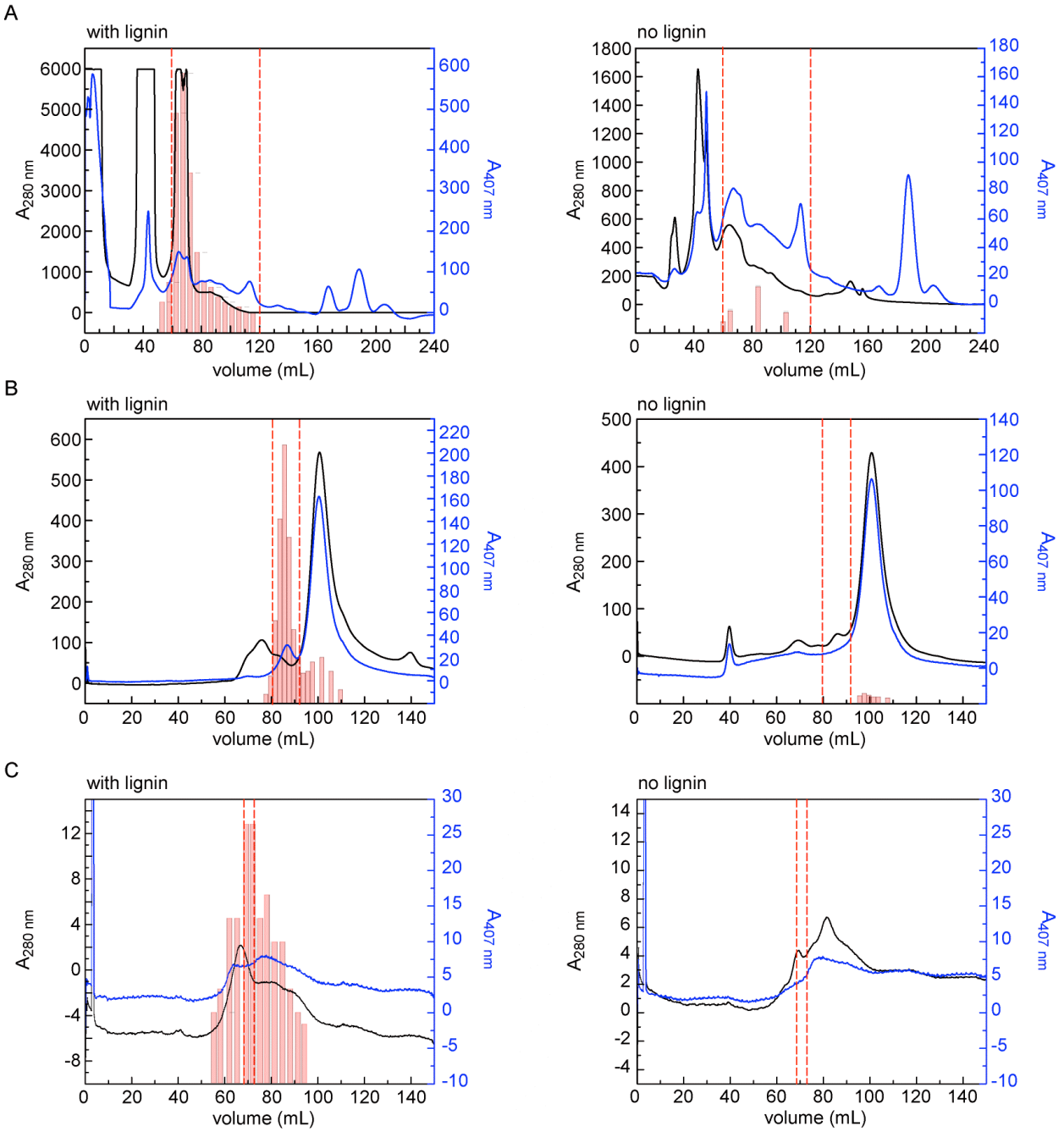


Figure A3.10. ^{1}H/^{13}C HSQC spectra for cell growth controls with ^{13}C-labeled DHP in MM. *HSQC NMR spectra of the solid fractions of ^{13}C-labeled DHP in MM (A) without and (B) with bacteria growth reveals very little modification to substrate due to the presence and growth of A. sp. 75iv2.*

Supplementary results: *Peroxidase purification data*

Figure A3.11. Native purification of peroxidase active enzyme(s) using DHP as the lignin substrate with control (no lignin) growth for comparison of candidates identified by MS/MS. *(A) Purification of the concentrated culture supernatant using the HiTrapQ column. Relative peroxidase activity (to normalized to peak activity in the +lignin purification step) is denoted by light red bar graph; no peroxidase activity was detected after the last step of purification in the no lignin control growth. High peroxidase active fractions in the with lignin growth (denoted by red-dashed lines) were pooled and concentrated. The same fractions (denoted by the dashed red lines) in the control growth were pooled and concentrated. (B) Separation of the peroxidase-active fraction using size-exclusion chromatography (SEC) with a Superdex 200 prepgrade SEC column. Half of the sample was concentrated and analyzed by LC-MS/MS after digestion with trypsin. (C) The other half was separated further by SEC using the Superdex 75 prepgrade SEC column. Peroxidase activity was spread through many fractions, but fractions with the peak activity (denoted by the dashed red lines) were pooled, digested with trypsin and analyzed by LC-MS/MS. (D) LC-MS/MS results identifying the proteins and number of ions. Proteins for expression were considered from those that quantified ions more heavily in the + lignin as opposed to the – lignin.*



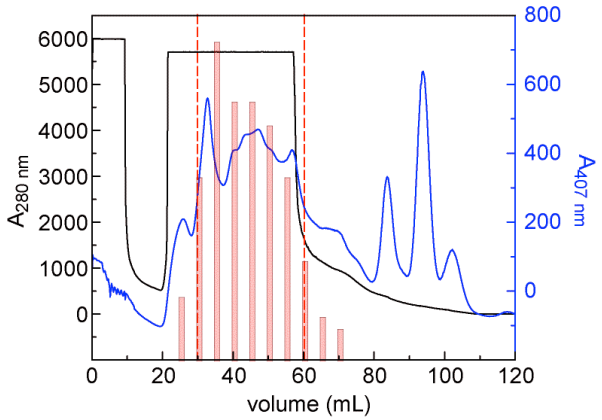
D

gene ID	gene annotation	Superdex 200 + lignin	Superdex 200 - lignin	Superdex 75 + lignin	Superdex 75 - lignin
Amy39116DRAFT_4250	ribose 5-phosphate isomerase	8	0	0	0
Amy39116DRAFT_0890	hypothetical protein	5	0	0	0
Amy39116DRAFT_7829	hypothetical protein	6	0	0	0
Amy39116DRAFT_6618	Uncharacterized enzyme involved in biosynthesis of extracellular polysaccharides	11	0	0	0
Amy39116DRAFT_5265	**WXG100 family type VII secretion target**	**5**	**0**	**2**	0
Amy39116DRAFT_2557	FOG: FHA domain	7	2	0	0
Amy39116DRAFT_5266	**WXG100 family type VII secretion target**	***6***	**0**	**4**	0
Amy39116DRAFT_3005	**Zn-dependent hydrolases, including glyoxylases**	**28**	**7**	**11**	0
Amy39116DRAFT_7494	methylmalonyl-CoA epimerase	5	0	0	0
Amy39116DRAFT_5583	**Beta-glucosidase-related glycosidases**	**36**	**0**	**5**	0
Amy39116DRAFT_7088	Predicted xylanase/chitin deacetylase	15	0	0	0
Amy39116DRAFT_5099	AIG2-like family	5	4	0	0
Amy39116DRAFT_0012	hypothetical protein	2	0	0	0
Amy39116DRAFT_5875	ABC-type amino acid transport/signal transduction systems	10	3	0	0
Amy39116DRAFT_7273	Lactoylglutathione lyase and related lyases	5	0	0	0
Amy39116DRAFT_7795	Uncharacterized conserved protein	4	0	0	0
Amy39116DRAFT_6773	Lactoylglutathione lyase and related lyases	3	0	0	0
Amy39116DRAFT_3377	Uncharacterized conserved protein	4	0	0	0
Amy39116DRAFT_7326	**Uncharacterized protein conserved in bacteria**	**2**	**0**	**9**	0
Amy39116DRAFT_6850	Galactose mutarotase and related enzymes	6	0	0	0
Amy39116DRAFT_6540	**hypothetical (L,D-transpeptidase catalytic domain)**	**15**	**0**	**6**	0
Amy39116DRAFT_7128	ABC-type Fe3+-hydroxamate transport system, periplasmic component	7	0	0	0
Amy39116DRAFT_7219	hypothetical protein	8	0	0	0
Amy39116DRAFT_6337	hypothetical protein	12	0	0	0
Amy39116DRAFT_4694	Trypsin	10	0	0	0
Amy39116DRAFT_5201	hypothetical protein	12	0	0	0
Amy39116DRAFT_4403	Guanyl-specific ribonuclease Sa	3	0	0	2
Amy39116DRAFT_6209	Lyzozyme M1 (1,4-beta-N-acetylmuramidase)	3	4	12	0
Amy39116DRAFT_2724	Peptidyl-prolyl cis-trans isomerase (rotamase)	4	0	0	0
Amy39116DRAFT_5949	Uncharacterized conserved protein	2	0	0	0
Amy39116DRAFT_2098	Gluconolactonase	2	0	0	0
Amy39116DRAFT_6326	chaperone protein DnaK	19	4	4	0
Amy39116DRAFT_0666	transaldolase, mycobacterial type	6	0	0	0
Amy39116DRAFT_5547	Periplasmic glycine betaine/choline-binding	3	0	0	0
Amy39116DRAFT_5753	Predicted metalloprotease	3	0	0	0
Amy39116DRAFT_7829	hypothetical protein	3	0	0	0
Amy39116DRAFT_6484	Cna protein B-type domain	9	0	11	0
Amy39116DRAFT_1157	**hypothetical protein**	**3**	**0**	**13**	0
Amy39116DRAFT_7181	ABC-type dipeptide transport system, periplasmic component	6	0	0	0
Amy39116DRAFT_6463	ABC-type dipeptide transport system, periplasmic component	7	0	0	0
Amy39116DRAFT_6992	Uncharacterized protein conserved in bacteria	6	6	7	9
Amy39116DRAFT_7956	Cytochrome c, mono- and diheme variants	3	0	0	0
Amy39116DRAFT_6114	hypothetical protein	2	0	0	0
Amy39116DRAFT_1812	Esterase/lipase	3	0	0	0
Amy39116DRAFT_4318	ribosomal protein S2, bacterial type	2	4	0	0
Amy39116DRAFT_5331	ribosomal protein L7/L12	9	3	0	0
Amy39116DRAFT_4477	hypothetical protein	3	0	0	0
Amy39116DRAFT_1835	Aldo/keto reductases, related to diketogulonate reductase	4	0	0	0
Amy39116DRAFT_0149	hypothetical protein	10	0	0	0
Amy39116DRAFT_6210	hypothetical protein	7	5	2	0
Amy39116DRAFT_3646	Metal-dependent hydrolase	3	0	0	0
Amy39116DRAFT_7889	Protein of unknown function (DUF3052)	3	0	0	0
Amy39116DRAFT_4294	methionine aminopeptidase, type I	3	0	0	0
Amy39116DRAFT_4195	trigger factor	4	0	0	0
Amy39116DRAFT_4457	Subtilase family.	24	24	57	75
Amy39116DRAFT_4922	isoform II	5	9	0	0
Amy39116DRAFT_7957	Rieske Fe-S protein	4	0	0	0
Amy39116DRAFT_2750	threonyl-tRNA synthetase	3	0	0	0
Amy39116DRAFT_7697	Subtilisin-like serine proteases	4	0	0	0
Amy39116DRAFT_7044	Membrane-bound lytic murein transglycosylase B	4	0	0	0
Amy39116DRAFT_6003	6-phosphogluconate dehydrogenase (decarboxylating)	6	0	0	0
Amy39116DRAFT_1938	Fibronectin type III domain	5	0	0	0
Amy39116DRAFT_0157	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	4	0	4	0
Amy39116DRAFT_2532	Cell wall-associated hydrolases (invasion-associated proteins)	3	0	0	0
Amy39116DRAFT_5965	PASTA domain./Protein kinase domain	20	5	3	5

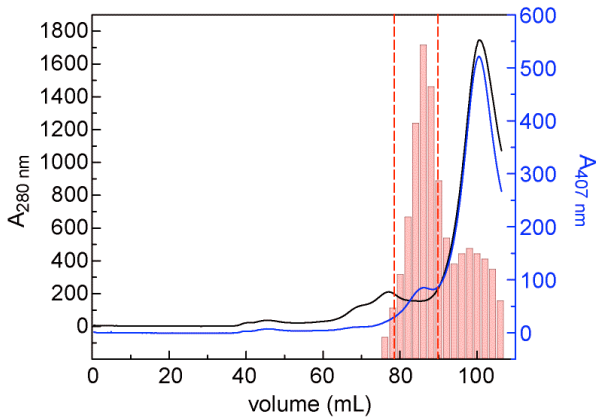
Amy39116DRAFT_4849 Predicted hydrolases or acyltransferases	3	0	3	0
Amy39116DRAFT_7036 Predicted periplasmic lipoprotein involved in iron transport	7	0	0	0
Amy39116DRAFT_6746 chaperonin GroL	3	3	0	0
Amy39116DRAFT_6119 Predicted hydrolases or acyltransferases	5	6	0	0
Amy39116DRAFT_1438 conserved repeat domain	4	0	0	0
Amy39116DRAFT_4513 Electron transfer flavoprotein	4	3	0	0
Amy39116DRAFT_2697 Predicted acyltransferases	3	0	0	0
Amy39116DRAFT_6072 hypothetical protein	4	0	0	0
Amy39116DRAFT_6550 Beta-glucosidase-related glycosidases	3	0	0	0
Amy39116DRAFT_4823 putative hydrolase	5	0	3	0
Amy39116DRAFT_7046 Subtilisin-like serine proteases	5	15	0	0
Amy39116DRAFT_5311 translation elongation factor EF-G	*8*	0	0	0
Amy39116DRAFT_8229 methylmalonyl-CoA mutase N-terminal domain	*2*	0	0	0
Amy39116DRAFT_3272 RHS repeat-associated core domain	*4*	0	0	0
Amy39116DRAFT_7503 Phosphomannose isomerase	0	2	0	0
Amy39116DRAFT_4288 Secreted trypsin-like serine protease	0	8	19	21
Amy39116DRAFT_5978 Peptidyl-prolyl cis-trans isomerase (rotamase) - cyclophilin family	0	5	2	7
Amy39116DRAFT_5276 ribosomal protein L17	0	6	4	4
Amy39116DRAFT_2602 ABC-type branched-chain amino acid transport systems	0	4	0	0
Amy39116DRAFT_4783 Protein-disulfide isomerase	0	3	3	0
Amy39116DRAFT_5235 Negative regulator of beta-lactamase expression	0	3	0	0
Amy39116DRAFT_6940 Trypsin-like serine proteases	0	6	0	0
Amy39116DRAFT_0565 Na+/H+ antiporter, bacterial form	0	3	0	0
Amy39116DRAFT_4570 hypothetical protein	0	0	6	0
Amy39116DRAFT_0703 Predicted metal-sulfur cluster biosynthetic enzyme, Paad-like	0	0	**6**	0
Amy39116DRAFT_4561 Predicted metal-dependent protease of the PAD1/JAB1 superfamily	0	0	**12**	0
Amy39116DRAFT_4351 Predicted Zn-dependent hydrolases of the beta-lactamase fold	0	0	3	0
Amy39116DRAFT_5644 hypothetical protein	0	0	3	0
Amy39116DRAFT_2315 Plant protein of unknown function (DUF946)	0	0	5	0
Amy39116DRAFT_8064 Molecular chaperone	0	0	4	4
Amy39116DRAFT_4148 ribonuclease, Rne/Rng family	0	0	8	6
Amy39116DRAFT_2482 Sporulation and spore germination	0	0	3	0
Amy39116DRAFT_2370 LGFP repeat./N-acetylmuramoyl-L-alanine amidase	0	0	4	0
Amy39116DRAFT_2450 hypothetical protein	0	0	4	0

Figure A3.12. Native purification of peroxidase active enzyme(s) using ethanol-extracted lignin as the lignin substrate with control (no lignin) growth for MS comparison. *(A) Purification of the concentrated culture supernatant using the HiTrapQ column. Relative peroxidase activity (to normalized to peak activity in the +lignin purification step) is denoted by light red bar graph.High peroxidase active fractions in the with lignin growth (denoted by red-dashed lines) were pooled and concentrated. The same fractions (denoted by the dashed red lines) in the control growth were pooled and concentrated. (B) Separation of the peroxidase-active fraction using size-exclusion chromatography (SEC) with a Superdex 200 prepgrade SEC column. High peroxidase active fractions in the with lignin growth (denoted by red-dashed lines) were pooled and concentrated. (C) Further separation of the high peroxidase fraction using MonoQ anion exchange. Peroxidase activity was spread through many fractions, but fractions with the peak activity (denoted by the red asterisk) were kept separate for analysis. (D) LC-MS/MS results identifying the proteins and number of ions of MonoQ purification fractions 20-23 (volume, 97 - 107 mL; F, fraction).*

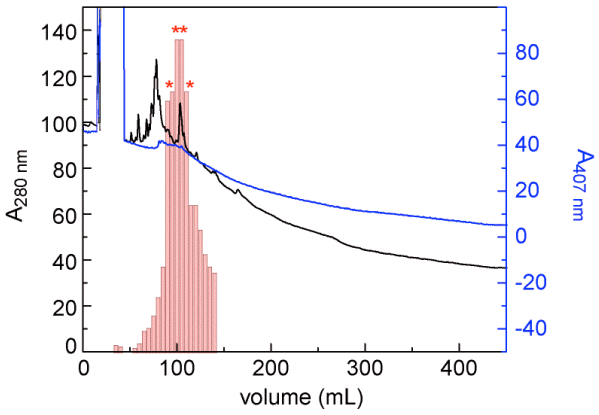
A



B



C



D

gene ID	gene annotation	F 20	F 21	F 22	F 23
Amy39116DRAFT_6548	Beta-lactamase class C and other penicillin binding proteins	58	9	0	0
Amy39116DRAFT_5875	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	84	405	556	55
Amy39116DRAFT_5236	hypothetical protein	112	0	0	0
Amy39116DRAFT_2557	FOG: FHA domain	62	13	63	6
Amy39116DRAFT_6883	hypothetical protein	22	0	0	0
Amy39116DRAFT_0216	Cutinase	23	0	0	0
Amy39116DRAFT_6540	L,D-transpeptidase catalytic domain	35	12	10	9
Amy39116DRAFT_7219	hypothetical protein	86	34	40	35
Amy39116DRAFT_4413	Subtilisin-like serine proteases	69	0	0	0
Amy39116DRAFT_5266	WXG100 family type VII secretion target	8	29	14	0
Amy39116DRAFT_7273	Lactoylglutathione lyase and related lyases	12	0	0	0
Amy39116DRAFT_5807	hypothetical protein	31	5	0	0
Amy39116DRAFT_5547	Periplasmic glycine betaine/choline-binding (lipo)protein of an ABC-type	40	0	0	0

transport system				
Amy39116DRAFT_0626 gamma-glutamyltranspeptidase	103	0	0	0
Amy39116DRAFT_5097 hypothetical protein	27	0	0	0
Amy39116DRAFT_5797 Protein of unknown function	15	0	0	0
Amy39116DRAFT_6867 hypothetical protein	38	37	6	6
Amy39116DRAFT_5265 WXG100 family type VII secretion target	4	37	15	0
Amy39116DRAFT_3131 pyridoxal phosphate enzyme, YggS family	20	0	0	0
Amy39116DRAFT_4639 Protein of unknown function (DUF3105)	14	0	0	0
Amy39116DRAFT_0149 hypothetical protein	28	0	0	0
Amy39116DRAFT_7181 ABC-type dipeptide transport system, periplasmic component	31	117	42	6
Amy39116DRAFT_6119 Predicted hydrolases or acyltransferases	19	29	42	20
Amy39116DRAFT_7957 Rieske Fe-S protein	53	37	20	15
Amy39116DRAFT_2885 hypothetical protein	2	0	0	0
Amy39116DRAFT_7088 Predicted xylanase/chitin deacetylase	19	4	3	3
Amy39116DRAFT_6618 Uncharacterized enzyme involved in biosynthesis of extracellular polysaccharides	5	16	20	31
Amy39116DRAFT_4783 Protein-disulfide isomerase	4	0	0	0
Amy39116DRAFT_7044 Membrane-bound lytic murein transglycosylase B	14	2	0	0
Amy39116DRAFT_3543 Uncharacterized conserved protein	11	0	0	0
Amy39116DRAFT_4570 hypothetical protein	9	11	5	2
Amy39116DRAFT_7024 Glucose/sorbosone dehydrogenases	11	0	0	0
Amy39116DRAFT_6072 hypothetical protein	45	18	0	0
Amy39116DRAFT_7494 methylmalonyl-CoA epimerase	2	21	11	7
Amy39116DRAFT_7142 Secreted trypsin-like serine protease	10	0	0	0
Amy39116DRAFT_8316 hypothetical protein	6	7	0	0
Amy39116DRAFT_4288 Secreted trypsin-like serine protease	4	0	0	0
Amy39116DRAFT_0657 3-phosphoglycerate kinase	10	0	0	0
Amy39116DRAFT_0465 Dienelactone hydrolase family	6	0	0	0
Amy39116DRAFT_4589 hypothetical protein	7	0	0	5
Amy39116DRAFT_7036 Predicted periplasmic lipoprotein involved in iron transport	30	24	31	41
Amy39116DRAFT_0887 Lactoylglutathione lyase and related lyases	5	0	0	2
Amy39116DRAFT_5784 hypothetical protein	3	0	0	0
Amy39116DRAFT_2315 Plant protein of unknown function (DUF946)	18	0	0	0
Amy39116DRAFT_0522 Trypsin-like serine proteases, typically periplasmic	13	0	0	0
Amy39116DRAFT_3884 hypothetical protein	7	0	0	0
Amy39116DRAFT_6701 diaminobutyrate acetyltransferase	7	0	0	0
Amy39116DRAFT_7795 Uncharacterized conserved protein	2	0	0	0
Amy39116DRAFT_7956 Cytochrome c, mono- and diheme variants	26	26	17	22
Amy39116DRAFT_4182 Predicted dioxygenase of extradiol dioxygenase family	2	0	0	0
Amy39116DRAFT_6484 Cna protein B-type domain	48	15	10	5
Amy39116DRAFT_5331 ribosomal protein L7/L12	4	0	0	0
Amy39116DRAFT_5313 ribosomal protein S12, bacterial/organelle	2	0	0	0
Amy39116DRAFT_0157 Predicted hydrolases or acyltransferases	23	28	11	8
Amy39116DRAFT_3844 hypothetical protein	34	3	4	3
Amy39116DRAFT_2901 Universal stress protein UspA and related nucleotide-binding proteins	2	0	0	0
Amy39116DRAFT_7436 Excalibur calcium-binding domain	6	0	4	3
Amy39116DRAFT_5742 hypothetical protein	26	4	0	0
Amy39116DRAFT_6326 chaperone protein DnaK	10	3	0	26
Amy39116DRAFT_6992 Uncharacterized protein conserved in bacteria	53	66	34	95
Amy39116DRAFT_4457 Subtilase family	35	22	28	18
Amy39116DRAFT_2135 ABC-type dipeptide transport system, periplasmic component	6	0	0	0
Amy39116DRAFT_2401 Thiamine pyrophosphate-requiring enzymes	3	0	0	0
Amy39116DRAFT_4318 ribosomal protein S2, bacterial type	25	32	10	2
Amy39116DRAFT_4294 methionine aminopeptidase, type I	3	0	0	0
Amy39116DRAFT_4195 trigger factor	2	2	5	5
Amy39116DRAFT_1938 Fibronectin type III domain	19	12	38	17
Amy39116DRAFT_4149 Subtilisin-like serine proteases	13	8	0	0
Amy39116DRAFT_7128 ABC-type Fe3+-hydroxamate transport system, periplasmic component	2	0	0	0
Amy39116DRAFT_0442 ABC-type Fe3+-hydroxamate transport system, periplasmic component	2	0	0	0
Amy39116DRAFT_5235 Negative regulator of beta-lactamase expression	2	0	0	0
Amy39116DRAFT_7046 Subtilisin-like serine proteases	20	13	18	13
Amy39116DRAFT_5965 PASTA domain./Protein kinase domain	5	0	0	0
Amy39116DRAFT_5870 Phospholipase C	3	6	0	0
Amy39116DRAFT_6209 Lyzozyme M1 (1,4-beta-N-acetylmuramidase)	3	0	0	0
Amy39116DRAFT_6265 fructose-bisphosphate aldolase, class II	4	0	0	0
Amy39116DRAFT_6210 hypothetical protein	9	0	0	39
Amy39116DRAFT_6167 cell envelope-related function transcriptional attenuator common domain	11	0	0	0
Amy39116DRAFT_7894 Glucoamylase and related glycosyl hydrolases	6	0	0	0
Amy39116DRAFT_4513 Electron transfer flavoprotein, beta subunit	5	0	0	4
Amy39116DRAFT_1438 conserved repeat domain	3	0	0	0
Amy39116DRAFT_5416 hypothetical protein	4	0	4	2
Amy39116DRAFT_5754 Predicted aminopeptidases	2	2	0	2

Amy39116DRAFT_2690	Methyltransferase small domain	2	0	0	0
Amy39116DRAFT_0077	hypothetical protein	9	14	4	6
Amy39116DRAFT_2750	threonyl-tRNA synthetase	5	0	0	0
Amy39116DRAFT_2482	Sporulation and spore germination	3	0	0	0
Amy39116DRAFT_2697	Predicted acyltransferases	8	0	5	0
Amy39116DRAFT_1872	ABC-type sugar transport system, ATPase component	2	0	0	0
Amy39116DRAFT_3707	Glycine/D-amino acid oxidases (deaminating)	15	10	0	21
Amy39116DRAFT_5583	Beta-glucosidase-related glycosidases	2	92	288	136
Amy39116DRAFT_2534	hypothetical protein	3	0	0	3
Amy39116DRAFT_2370	LGFP repeat./N-acetylmuramoyl-L-alanine amidase	4	3	0	0
Amy39116DRAFT_7563	4-amino-4-deoxy-L-arabinose transferase and related glycosyltransferases	6	0	0	0
Amy39116DRAFT_6550	Beta-glucosidase-related glycosidases	2	83	108	3
Amy39116DRAFT_7100	hypothetical protein	6	0	4	0
Amy39116DRAFT_5153	Uncharacterized protein containing vWA domain	2	0	0	0
Amy39116DRAFT_2950	3-phosphoshikimate 1-carboxyvinyltransferase	18	5	30	23
Amy39116DRAFT_4406	Superfamily I DNA and RNA helicases	2	0	0	0
Amy39116DRAFT_7581	cobyric acid synthase CobQ	3	0	0	0
Amy39116DRAFT_2587	Ribosomal protein S1	7	0	0	0
Amy39116DRAFT_6541	hypothetical protein	2	0	0	0
Amy39116DRAFT_7512	Predicted metalloprotease	4	0	0	0
Amy39116DRAFT_6890	LGFP repeat./PKD domain./Glucose / Sorbosone dehydrogenase	2	5	3	0
Amy39116DRAFT_2757	hypothetical protein	5	0	0	0
Amy39116DRAFT_5424	parallel beta-helix repeat (two copies)	2	178	146	0
Amy39116DRAFT_6053	hypothetical protein	5	0	0	0
Amy39116DRAFT_5311	translation elongation factor EF-G	6	0	4	3
Amy39116DRAFT_4849	Predicted hydrolases or acyltransferases	6	2	15	19
Amy39116DRAFT_4264	NAD-specific glutamate dehydrogenase	2	0	0	0
Amy39116DRAFT_7877	Subtilisin-like serine proteases	0	29	20	0
Amy39116DRAFT_5097	hypothetical protein	0	5	0	0
Amy39116DRAFT_4252	Zn-dependent hydrolases, including glyoxylases	0	9	0	0
Amy39116DRAFT_3136	peroxiredoxin, OsmC subfamily	0	2	0	0
Amy39116DRAFT_0703	Predicted metal-sulfur cluster biosynthetic enzyme	0	5	6	4
Amy39116DRAFT_7697	Subtilisin-like serine proteases	0	20	14	14
Amy39116DRAFT_0146	hypothetical protein	0	2	0	0
Amy39116DRAFT_6012	Predicted RNA-binding protein	0	6	0	7
Amy39116DRAFT_0626	gamma-glutamyltranspeptidase	0	7	0	0
Amy39116DRAFT_1630	hypothetical protein	0	6	0	2
Amy39116DRAFT_6624	Carbonic anhydrase	0	2	2	0
Amy39116DRAFT_7528	Predicted secreted acid phosphatase	0	4	0	0
Amy39116DRAFT_3167	Acetyltransferases, including N-acetylases of ribosomal proteins	0	29	0	0
Amy39116DRAFT_7176	ABC-type dipeptide transport system, periplasmic component	0	26	0	0
Amy39116DRAFT_4525	Secreted trypsin-like serine protease	0	3	29	13
Amy39116DRAFT_4910	Uncharacterized conserved protein (some members contain a vWA domain)	0	3	0	0
Amy39116DRAFT_6037	Uncharacterized protein conserved in bacteria	0	3	3	0
Amy39116DRAFT_4085	Siderophore-interacting protein	0	4	3	4
Amy39116DRAFT_3284	ABC-type dipeptide transport system, periplasmic component	0	2	0	0
Amy39116DRAFT_1108	Cytochrome P450	0	2	0	0
Amy39116DRAFT_7220	hypothetical protein	0	8	0	0
Amy39116DRAFT_7343	Branched-chain amino acid aminotransferase/4-amino-4-deoxychorismate lyase	0	2	0	0
Amy39116DRAFT_2802	Sugar kinases, ribokinase family	0	2	0	0
Amy39116DRAFT_5894	Glycosyltransferase	0	2	0	0
Amy39116DRAFT_1326	Threonine dehydrogenase and related Zn-dependent dehydrogenases	0	11	0	0
Amy39116DRAFT_0367	Signal transduction histidine kinase	0	2	0	0
Amy39116DRAFT_0727	Acetyl/propionyl-CoA carboxylase, alpha subunit	0	2	0	0
Amy39116DRAFT_4823	putative hydrolase	0	2	5	0
Amy39116DRAFT_6369	hypothetical protein	0	2	0	0
Amy39116DRAFT_1806	hypothetical protein	0	3	0	0
Amy39116DRAFT_1911	Cu2+-containing amine oxidase	0	2	0	0
Amy39116DRAFT_0008	DNA-directed DNA polymerase III (polc)	0	10	14	12
Amy39116DRAFT_6035	hypothetical protein	0	4	5	0
Amy39116DRAFT_3108	hypothetical protein	0	0	6	6
Amy39116DRAFT_7262	hypothetical protein	0	0	57	18
Amy39116DRAFT_4250	ribose 5-phosphate isomerase	0	0	23	121
Amy39116DRAFT_5407	hypothetical protein	0	0	47	0
Amy39116DRAFT_6337	hypothetical protein	0	0	95	46
Amy39116DRAFT_4572	Predicted redox protein, regulator of disulfide bond formation	0	0	18	0
Amy39116DRAFT_6703	Uncharacterized protein, possibly involved in glyoxylate utilization	0	0	6	8
Amy39116DRAFT_2798	Predicted redox protein, regulator of disulfide bond formation	0	0	4	0
Amy39116DRAFT_0478	Cytosine/adenosine deaminases	0	0	10	0
Amy39116DRAFT_7966	Uncharacterized conserved protein	0	0	7	0
Amy39116DRAFT_6361	hypothetical protein	0	0	6	13
Amy39116DRAFT_2007	Uncharacterized conserved protein	0	0	4	0

Amy39116DRAFT_1823 hypothetical protein	0	0	39	42
Amy39116DRAFT_1822 hypothetical protein	0	0	4	12
Amy39116DRAFT_4915 Predicted metalloprotease	0	0	6	20
Amy39116DRAFT_5276 ribosomal protein L17	0	0	5	11
Amy39116DRAFT_0236 Dioxygenases related to 2-nitropropane dioxygenas	0	0	3	7
Amy39116DRAFT_0234 Acyl CoA:acetate/3-ketoacid CoA transferase, alpha subunit	0	0	5	0
Amy39116DRAFT_5856 Cytochrome P450	0	0	7	18
Amy39116DRAFT_0565 Na+/H+ antiporter, bacterial form	0	0	7	0
Amy39116DRAFT_5757 Protein of unknown function (DUF2587)	0	0	4	0
Amy39116DRAFT_6114 hypothetical protein	0	0	5	0
Amy39116DRAFT_6983 aminodeoxychorismate synthase, component I, bacterial clade	0	0	2	0
Amy39116DRAFT_4373 malate synthase A	0	0	4	0
Amy39116DRAFT_1206 Predicted nucleotidyltransferases	0	0	2	0
Amy39116DRAFT_0032 hypothetical protein	0	0	2	0
Amy39116DRAFT_4889 Zn-ribbon protein, possibly nucleic acid-binding	0	0	24	52
Amy39116DRAFT_2965 DNA-3-methyladenine glycosylase (3mg)	0	0	12	35
Amy39116DRAFT_6894 Predicted acyl-CoA transferases/carnitine dehydratase	0	0	2	0
Amy39116DRAFT_3431 Glutamine synthetase	0	0	3	0
Amy39116DRAFT_8198 phosphofructokinase	0	0	5	0
Amy39116DRAFT_3000 Poly(3-hydroxybutyrate) depolymerase	0	0	3	0
Amy39116DRAFT_5799 seryl-tRNA synthetase	0	0	3	6
Amy39116DRAFT_7185 Zn-dependent proteases	0	0	2	0
Amy39116DRAFT_1166 Predicted transcriptional regulators	0	0	21	0
Amy39116DRAFT_2016 Superfamily II RNA helicase	0	0	2	0
Amy39116DRAFT_0511 Acetyl/propionyl-CoA carboxylase, alpha subunit	0	0	2	0
Amy39116DRAFT_3873 Autotransporter adhesin	0	0	5	2
Amy39116DRAFT_7326 Uncharacterized protein conserved in bacteria	0	0	0	15
Amy39116DRAFT_3005 Zn-dependent hydrolases, including glyoxylases	0	0	0	86
Amy39116DRAFT_2098 Gluconolactonase	0	0	0	48
Amy39116DRAFT_2323 Predicted hydrolases or acyltransferase	0	0	0	28
Amy39116DRAFT_0666 transaldolase, mycobacterial type	0	0	0	57
Amy39116DRAFT_5099 AIG2-like family	0	0	0	16
Amy39116DRAFT_5507 Glutaredoxin and related proteins	0	0	0	5
Amy39116DRAFT_1812 Esterase/lipase	0	0	0	7
Amy39116DRAFT_2540 Uncharacterized protein conserved in bacteria	0	0	0	7
Amy39116DRAFT_7800 peptide deformylase	0	0	0	10
Amy39116DRAFT_2002 ubiquitin-like protein Pup	0	0	0	4
Amy39116DRAFT_4922 isoform II	0	0	0	8
Amy39116DRAFT_2125 conserved hypothetical protein	0	0	0	5
Amy39116DRAFT_2597 Lactoylglutathione lyase and related lyases	0	0	0	4
Amy39116DRAFT_2555 Uncharacterized conserved protein	0	0	0	5
Amy39116DRAFT_2061 hypothetical protein	0	0	0	5
Amy39116DRAFT_1023 ABC-type nitrate/sulfonate/bicarbonate transport system	0	0	0	2
Amy39116DRAFT_4451 2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase (catechol pathway)	0	0	0	4
Amy39116DRAFT_7881 Acyl carrier protein	0	0	0	5
Amy39116DRAFT_7183 hypothetical protein	0	0	0	2
Amy39116DRAFT_4476 Phosphoglycerate dehydrogenase and related dehydrogenases	0	0	0	6
Amy39116DRAFT_5263 hypothetical protein	0	0	0	2
Amy39116DRAFT_5935 Protein kinase domain	0	0	0	17
Amy39116DRAFT_4799 amidase, hydantoinase/carbamoylase family	0	0	0	5
Amy39116DRAFT_6178 hypothetical protein	0	0	0	2
Amy39116DRAFT_6325 Molecular chaperone GrpE (heat shock protein)	0	0	0	5
Amy39116DRAFT_1778 4-hydroxy-2-oxovalerate aldolase	0	0	0	2
Amy39116DRAFT_7889 Protein of unknown function (DUF3052)	0	0	0	5
Amy39116DRAFT_6889 Transcriptional regulators	0	0	0	4
Amy39116DRAFT_5516 Nucleoside-diphosphate-sugar epimerases	0	0	0	5
Amy39116DRAFT_5310 translation elongation factor TU	0	0	0	13
Amy39116DRAFT_1201 NADPH-dependent glutamate synthase beta chain and related oxidoreductases	0	0	0	2
Amy39116DRAFT_4335 hypothetical protein	0	0	0	3
Amy39116DRAFT_5715 Site-specific recombinase XerD	0	0	0	2
Amy39116DRAFT_0816 Inosine-uridine nucleoside N-ribohydrolase	0	0	0	7
Amy39116DRAFT_1719 Thiamine pyrophosphate-requiring enzymes	0	0	0	2
Amy39116DRAFT_2762 Nucleoside-diphosphate-sugar epimerases	0	0	0	2
Amy39116DRAFT_2960 hypothetical protein	0	0	0	2
Amy39116DRAFT_6616 2-polyprenyl-6-methoxyphenol hydroxylase and related FAD-dependent oxidoreductases	0	0	0	3
Amy39116DRAFT_4445 conserved repeat domain	0	0	0	3
Amy39116DRAFT_6115 Phosphoenolpyruvate carboxykinase (GTP)	0	0	0	2
Amy39116DRAFT_0650 hypothetical protein	0	0	0	2
Amy39116DRAFT_6567 DNA topoisomerase I, bacterial	0	0	0	2