ELUCIDATING THE MECHANISM OF ENZYMES UTILIZED DURING THE BACTERIAL DEGRADATION OF LIGNIN IN THE PROTOCATECHUATE 4,5-CLEAVAGE PATHWAY

A Dissertation

by

TESSILY NICHOLLE HOGANCAMP

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Frank M. Raushel
Committee Members,	David P. Barondeau
	Thomas D. Meek
	Coran M. H. Watanabe
Head of Department,	Simon W. North

December 2019

Major Subject: Chemistry

ABSTRACT

Lignin is the most abundant renewable aromatic in the world that has great potential as a sustainable source of monomeric units for the production of bioplastics. However, the chemical and biological processing of lignin in industry is difficult due to the heterogeneity of the polymer structure. The natural recycling of lignin occurs through depolymerization by oxidoreductases secreted by fungi. Many bacterial pathways metabolize these lignin oligomers into small carbon fragments which are used in the TCA cycle. Many of the known enzymatic pathways funnel to protocatechuate (PCA) that can be further metabolized by three cleavage pathways. Our goal is to decipher the mechanistic details of the PCA 4,5-cleavage pathway from the *Sphingomonas* sp. SYK-6.

The enzymes LigU, LigJ, and LigK play an important role in the degradation, but have not been structurally or mechanistically characterized. Using NMR spectroscopy, LigU was shown to function as a 1,3-allylic isomerase where the double bond between C5-C4 of (4E)-oxalomesaconate (OMA) migrates to C4-C3 forming (3*Z*)-2-keto-4-carboxy-3-hexenedioate (KCH). LigU was further characterized to show that the stereospecific proton transfer at C5 of KCH is pro*S* and that Lys-66 and Cys-100 are critical catalytic residues. We recently determined the structure of wild-type LigU and the ligand-bound complex of the mutant LigU K66M by X-ray crystallography. The ligand-bound structure also shows that the two catalytic residues are on the same face of the ligand to support a *syn*-fashion reaction. We also determined the three-dimensional

structures of the LigJ hydratase. The structure of the product bound complex shows that CHA, 4-carboxy-4-hydroxy-2-oxoadipate is formed in the *S*-configuration. The mutant E284Q has no observable catalytic activity and this mutant was used to co-crystallize the substrate, KCH with LigJ. Based on the kinetic parameters and the structure of the bound substrate in the active site, we propose that Glu-284 acts as both the catalytic base and acid in an overall *syn*-addition. My future work will focus on characterizing the C-C cleavage of LigK aldolase by site-directed mutagenesis and monitoring changes in activity. Also, attempts at co-crystallization of mutants with substrate and substrate analogs are ongoing to determine the active site chemistry.

DEDICATION

I dedicate this dissertation to the great teachers who inspired me to dream big and then to work hard to make the dream a reality.

To the science teacher who first taught me about the Punnett square and inspired me to pursue the life sciences.

To the history teacher who tried to burn a cow patty during class and failed. You inspired me to test the status quo and to learn from failure.

To my neighbor and friend, a school teacher, who inspired me to trust my own vision and to explore the world.

ACKNOWLEDGEMENTS

First, I would like to thank my advisor Professor Frank M. Raushel. You have inspired me to set high expectations to produce quality and insightful discoveries, and I'm forever grateful to the critical research skills that you have taught me. I would like to thank my committee members, Dr. Barondeau, Dr. Meek, and Dr. Watanabe, for their support and guidance during my PhD career. I would like to thank my co-authors Dr. Mark Mabanglo and Seth Cory for your expertise in X-ray crystallography and efforts into successful papers with quality protein structures. I would like to acknowledge the staff scientists at TAMU, Dr. Howard Williams, Dr. Greg Wylie, Dr. Douglas Elliott, Dr. Bryan Tomlin, and Dr. Xu Xianzhong who helped me optimize many experiments for my projects. I would like to thank fellow graduate students Dr. Jihyun Kim for help with NMR experiments, Dongju Lee for initial crystal screens of LigU and LigJ enzymes, and Dr. Su Tang for help with X-ray diffraction experiments. I would like to recognize Andrea Scott our administrative associate that has always made work run so smoothly. Thank you Sandy Horton and Valerie McLaughlin for your support over the years. Thank you Ed Janousek for always being a friendly face at work and for the encouragement over the years. I appreciate the camaraderie over the past five years in the Raushel Lab from the staff scientists, post-docs, my peers, and undergraduate students. A special acknowledgement to fellow graduate student and my best friend at TAMU Zane Taylor for all your help in lab and the jokes and friendship that really made midnight purifications and assays not so bad. I appreciate the mentorship and friendship

of Dr. Josh Wofford, my first friend at TAMU. I would also like to acknowledge the Graduate Student Association of Chemistry (GSAC) for a wonderful club that brings the graduate students together beyond science that fosters many friendships, specifically with Dr. Greg and Shawna Waetzig and Dr. Stephen Fordham. A special note to my friend, Isita Jhulki, I'm so happy to share this grad school experience with you. Thank you to my small group family that has helped me to stay grounded to the importance of family and love and laughter. I appreciate my housemates Dr. Jeff Rose and Daniel, Kathy, and Brielle Tomkins for putting up with my weird work hours and for great life advice and fun times. I would like to give special attention to my parents, grandparents, siblings, and extended family for your love, care, and encouragement. I have had so many wonderful people in my life, and I'm so thankful for all of them. To my childhood friends, Tina Claridy and Angela Mitchell, thank you for always pushing me to great adventures. Finally, I would like to acknowledge the person who stole my heart at the age of 16, to my great husband, Josh. Thank you Josh for giving me unconditional love and support. We have already had so many great dreams come true together, and I'm so excited for the next adventure.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Frank M. Raushel, Professor David P. Barondeau, and Professor Coran M. H. Watanabe of the Department of Chemistry and Professor Thomas D. Meek of the Department of Biochemistry and Biophysics.

The nuclear magnetic resonances (NMR) data for Chapter 2 had instrumental assistance from Dr. Gregory P. Wylie, Dr. Douglas W. Elliott, Dr. Howard J. Williams, Dr. Jihyun Kim, and Dr. Xianzhong Xu. For Chapters 2, 3, and 4 chemicals not purchased commercially were synthesized by Dr. Tamari Narindoshvili. For Chapters 3 and 4 the X-ray crystallography facilities of Dr. James Sacchettini were used. The X-ray crystallography data for Chapters 3 and 4 had instrumental assistance from Dr. Viktor Grichko, Dr. Su Tang, and Dr. Inna Krieger. Initial crystal screen for chapters 3 and 4 were by Dongju Lee and Dr. Mark Mabanglo. The X-ray crystallography structures in Chapter 3 were grown and determined by Dr. Mark Mabanglo of the Department of Chemistry and were published in 2018. The X-ray crystallography structures in Chapter 4 were determined by Seth A. Cory and Dr. David P. Barondeau of the Department of Chemistry and were published in 2019. Crystal structure data for Chapters 3 and 4 were collected at Stanford Synchrotron Radiation Lightsource (SSRL). The inductively

coupled plasma mass spectrometry (ICP-MS) data for Chapter 3 had instrumental help from Dr. Bryan Tomlin.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate student was supported by a teaching assistantship from Texas A&M University for 2014-2016 and a research assistantship from Texas A&M University and the Robert A. Welch Foundation under Grant Number A-840 for 2015-2019. Its content are solely the responsibility of the authors and do not necessarily represent the official views of the Robert A. Welch Foundation and Texas A&M University.

NOMENCLATURE

ca.	approximately
mya	million years ago
UV	ultra-violet
G	guaiacyl
S	syringyl
H	*p*-hydroxyphenyl
β-O-4	β-aryl ether
α-O-4	α-aryl ether
4-O-5	diaryl ether
5-5-O-4	dibenzodioxocin
β-1	spirodienone
β-β	resonol
β-5	phenylcoumaran
TCA	tricarboxylic acid
sp.	species
DyPs	dye-decolorizing peroxidases
CcP	cytochrome c peroxidase
LiP	lignin peroxidase
MnP	manganese peroxidase
VP	versatile peroxidase

LDA	lignin-degrading auxiliary
O_2	molecular oxygen
H_2O_2	hydrogen peroxide
Fe(III)	ferric iron
Fe (IV) $=O^{\bullet+}$	Compound I
Fe (IV) =O	Compound II
ox	oxidized
H_2O	water
pI	isoelectric point
E°	electronic reduction potential
mV	milli-Volts
C	carbon
α	alpha
β	beta
γ	gamma
Å	angstrom
PDB	protein data bank
RMSD	root mean square difference
H/His	histidine
R/Arg	arginine
W/Trp	tryptophan
LRET	long-range electron transfer

Por	porphyrin ring
C/Cys	cysteine
Mn	manganese
E/Glu	glutamate
D/Asp	aspartate
M	molar
s	second
Lac	laccase
IPR	InterPro entry
kDa	kiloDaltons
T1	type one copper
M/Met	methionine
I/Ile	isoleucine
nm	nanometers
S	sulfur
Cu	copper
$A_{\parallel}$	hyperfine coupling constant
mT	milli-tesla
EPR	electron paramagnetic resonance
T2	type two copper
T3	type three copper
LMS	laccase-mediator system

ABTS	2,2’-azinbobis(3-ethylbenzthiazoline-6-sulfonate)
HBT	1-hydroxybenzotriazole
TEMPO	2,2,6,6-tetramethyl-1-piperidinyloxy
F/Phe	phenylalanine
L/Leu	leucine
e-	electron
H	hydrogen
R	substituent group
DDVA	5,5’-dehydrodivanillate
OH-DDVA	2,2’,3-trihydroxy-3’-methoxy-5,5’-dicarboxybiphenyl
NADH	nicotinamide adenine dinucleotide reduced form
FAD	flavin adenine dinucleotide
μM	micro-molar
K_m	Michaelis constant
k_{cat}	turnover number
k_{cat} / K_m	substrate specificity
DCHM-HOPDA	4,11-dicarboxy-8-hydroxy-9-methoxy-2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate
5CV	5-carboxyvanillate
CHPD	4-carboxy-2-hydroxypenta-2,4-dienoate
AHS	amidohydrolase superfamily
mer	subunits of protein quaternary structure

H_4folate	tetrahydrofolate
5-methyl-H_4folate	5-methyl-tetrahydrofolate
Y/Tyr	tyrosine
Q/Gln	glutamine
mM	milli-Molar
PCA	protocatechuate
CHMS	4-carboxy-2-hydroxymuconate-6-semialdehyde
PDC	2-pyrone-4,6-dicarboxylate
NADP+	nicotinamide adenine dinucleotide phosphate oxidized form
NADPH	nicotinamide adenine dinucleotide phosphate reduced form
OMA	(4*E*)-oxalomesaconate
KCH	(3*Z*)-2-keto-4-carboxy-3-hexendioate
CHA	(*S*)-4-carboxy-4-hydroxy-2-oxoadipate
X	residue in motif
S/Ser	serine
N/Asn	asparagine
O	oxygen
Fe	iron
COG	clusters of orthologous groups
NMR	nuclear magnetic resonance
K/Lys	lysine
DAP	diaminopimelate

RraA	ribonuclease activity A
RNA	ribonucleic acid
T/Thr	threonine
IPTG	isopropyl β- D-thiogalactopyranoside
rpm	revolutions per minute
min	minute
°C	degree Celsius
mg	milligram
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
nM	nanoMolar
TEA	triethanolamine
mL	milli-liter
μL	micro-liter
K	Kelvin
ppm	parts per million
equiv	equivalent
D_2O	deuterium oxide
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
MOPS	3-(*N*-morpholino) propanesulfonic acid
KOH	potassium hydroxide
KCl	potassium chloride
$NaBH_4$	sodium borohydride

$NaBD_4$	sodium borodeuteride
DCl	deuterium chloride
pM	pico-Molar
$\Delta\varepsilon$	differential extinction coefficient
cm	centimeter
eq	equation
v	initial velocity
E_t	total enzyme concentration
[A]	substrate concentration
K_a	Michaelis constant
mol	moles
HMG	4-hydroxy-4-methyl-2-oxoglutarate
$MgCl_2$	magnesium chloride
h	hour
Hz	hertz
^4J	four bond coupling constant in NMR
^3J	three bond coupling constant in NMR
^2J	two bond coupling constant in NMR
^1J	one bond coupling constant in NMR
OMA_{eq}	equilibrium concentration of OMA
KCH_{eq}	equilibrium concentration of KCH
PDC_o	initial concentration of PDC

PDC_{eq}	equilibrium concentration of PDC
WT	wild-type
OAA	oxaloacetate
PYR	pyruvate
A/Ala	alanine
SEC	size-exclusion chromatography
ε	extinction coefficient
ICP-MS	inductively coupled plasma mass spectrometry
ms	milli-seconds
^{45}Sc	Scandium isotope 45
Cd	cadmium
Co	cobalt
Cu	copper
Ni	nickel
Zn	zinc
HNO_3	nitric acid
ppb	parts per billion
V	volts
MΩ	mega ohm
EFI	Enzyme Function Initiative
SSN	sequence similarity network
w/v	weight per volume ratio

SSRL	Stanford Synchrotron Radiation Lightsource
TLS	translational/libration/screw model
E. coli	*Escherichia coli*
$(2F_o - F_c)$	composite omit map
CHM	4-carboxy-2-hydroxymuconate
γ-RSD	γ-resorcylate decarboxylase
MSD	6-methylsalicylate decarboxylase
IDC	5-carboxy-uracil decarboxylase
ACMSD	α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase
BSA	bovine serum albumin
nL	nano-liter
MDH	malate dehydrogenase
G/Gly	glycine
V/Val	valine
$mF_o - F_c$	omit map

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

LIST OF SCHEMES

Page

1. INTRODUCTION: ENZYMATIC DEGRADATION OF LIGNIN

1.1. Lignin Introduction

Lignin is a crucial global carbon sink constituting the second largest source of biopolymer after cellulose (1). The etymology of the word lignin originates from the Latin word *lignum*, which means "wood". The existence of lignin first appeared when aqueous plants evolved into land plants ca. 450 million years ago (mya) based on fossil records (2). The main adaptation that made this possible was the biosynthesis of lignin material. The plant phylum called tracheophyta that arranged lignin monomers into polymers within the cell walls for rigidity to help increase length and strength first appeared in the late Silurian period (443 to 416 mya) and diversified and covered the land into the early Devonian period (416 to 358 mya) (3). Ancestral analysis of similar genetic pathways in aqueous and terrestrial plants show the recruitment of enzymes able to synthesize precursors to lignin monomers, the phenylpropanoid molecules. There is a core of eight enzymes that were duplicated to make these aromatic structures that could absorb harmful UV radiation to prevent cell damage (2). Along with UV-radiation stress, land plants had to overcome the medium change from water to air. This required rigidity and a pumping system that could work against gravity to transport water and nutrients. In addition, plants needed to survive new threats of pathogens and viruses, so they needed a hard surface to withstand such threats. Ultimately, lignin was the answer to all of these new environmental threats. Lignin is a structural property of vascular

plants and some algae that distinguishes it from water based plants. Lignin is found in the walls of secondary thickened plant cells (2).

1.1.1. Chemical Structure

Lignin is composed of three phenylpropanoid monomers (guaiacyl (G), syringyl (S), and *p*-hydroxyphenyl (H)) which turn into monolignol precursors (**Figure 1**) (4). These basic building blocks are methoxylated derivatives of benzene ($R\text{-}O\text{-}CH_3$) (4). They are linked either by ether (C-O) or carbon-carbon bonds. Some common bonds include β-O-4 (β-aryl ether), α-O-4 (α-aryl ether), 4-O-5 (diaryl ether), 5-5-O-4 (dibenzodioxocin), benzodioxane, β−1 (spirodienone), β−β (resonol), β−5 (phenylcoumaran), and 5−5' bonds to form many different complex heteropolymers (**Figure 2**) (4). The β-O-4 linkage is the most abundant at 40-60% levels in softwood and hardwood (5). The distribution and pattern of monomers in lignin is randomly determined as it is radical coupling of peroxidases and laccases that joins the units together (1). The heterogeneity of lignin is the foremost reason limiting its usefulness in a biosustainable economy.

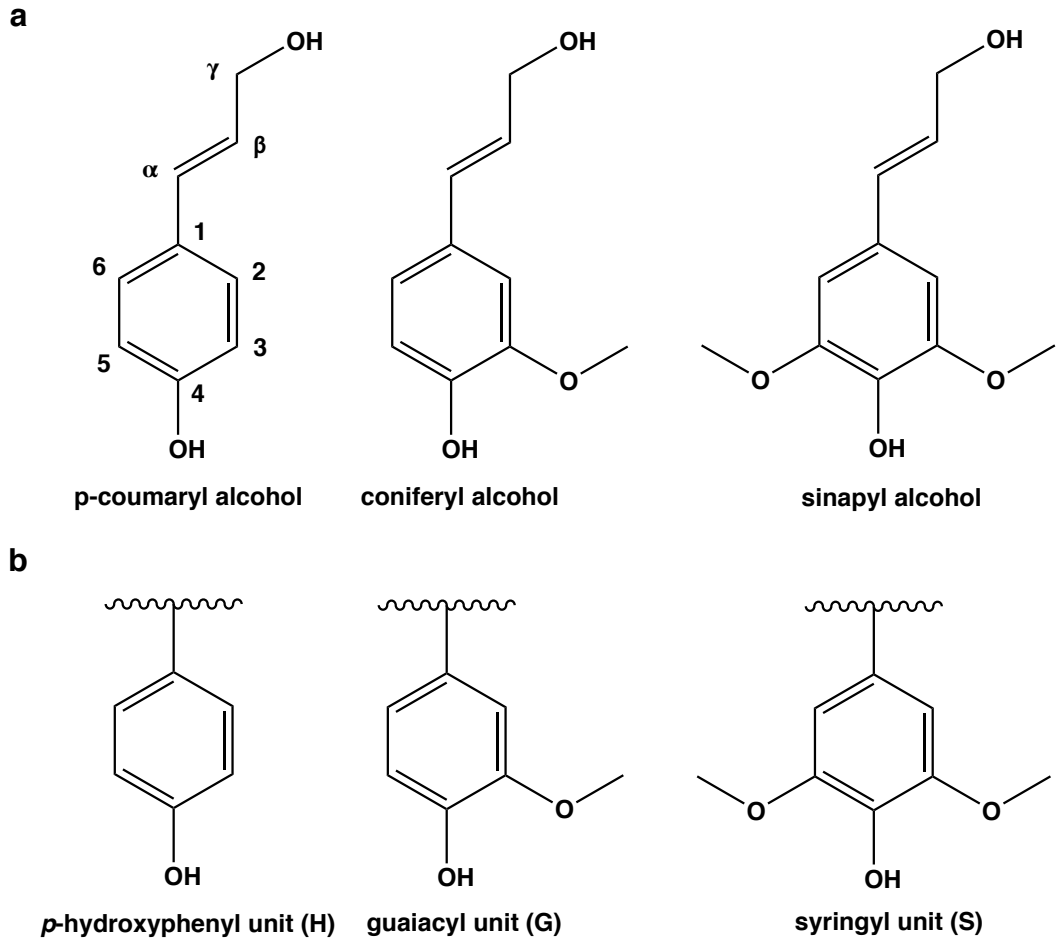


Figure 1. (a) The three monolignol alcohols of lignin. (b) The corresponding lignin units.

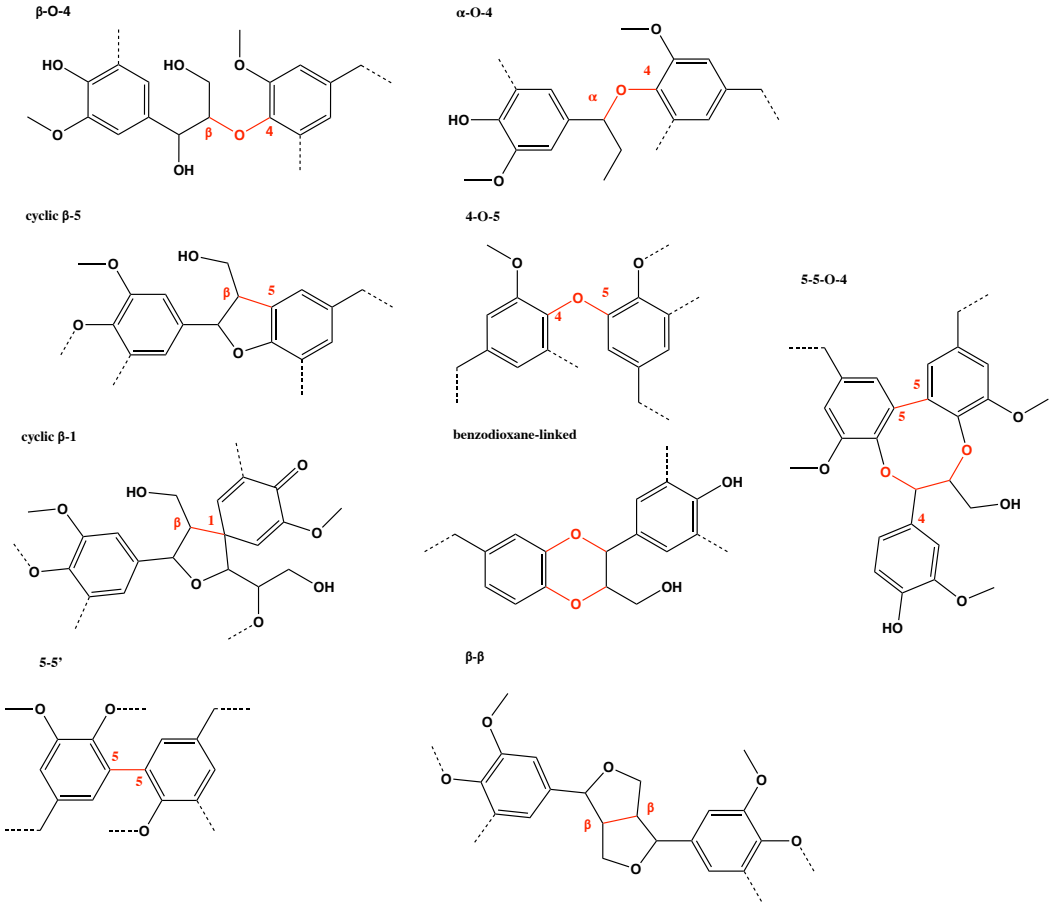


Figure 2. Some common bonds in the structure of lignin.

1.1.2. Significance and Broader Impact

The removal of lignin from wood used in industrial settings has been advancing since the 1950s in scientific articles (6). Lignin removal originated in the paper industry and was regarded as a waste. Traditionally, the focus was to extract cellulose from the lignocellulosic biomass, and lignin was just burned as a heating source for processing cellulose. However, the pressure of limited fossil fuels and the concern of greenhouse emission from using coal and petroleum, has incentivized industry and the government to support fundamental research in alternative and renewable energy (7). Therefore, to increase the competitive cost of wood as a source, lignin has been the topic of extensive research for decades (7). There are many challenges of using lignin versus petroleum as a raw source. The main challenge is the heterogeneity of lignin monomers and bond types. Cellulose has only one major bond type to break, whereas lignin is composed of more than ten types of bonds with different ratios of the three monomers and their derivatives in different plant species (8). Secondly, compared to fossil fuels, lignin has a high oxygen and moisture content that lowers the energy burning value thus making it less competitive (9). Third, since it has many oxygen bonds that are broken, these fragments are highly unstable and tend to re-polymerize and form different polymers (10). To overcome the high oxygen content and stability, hydrodeoxygenation and fast pyrolysis research has been improving the purity and stability of the fragments (11). This research has been investigating different types of catalysts to remove oxygen, which carries its own challenges, such as feasible environmental pressures that can be scaled-up and developing catalysts that do not use expensive metals but more

economically ones without comprising the cost to be competitive with nonrenewable fuels (11).

Another avenue of interest has been to use microbial engineering to develop novel bacterial strains that can depolymerized lignin fragments to a valuable intermediate via genetically modified enzymatic pathways (12). This intermediate can then be converted into a precursor for the basic units for industrial polymers like nylon, resins, lubricants, and plastics (13-15). This is known as lignin valorization for the production of value-added chemicals.

1.2. Lignin Degradation Overview

The enzymes responsible for lignin depolymerization appeared in fungi ca. 295 mya based on comparative genomics and molecular clock analyses (16). In nature, the degradation of a fallen tree to the surrounding soil typically takes between 46 to 124 years with the time dependent on many factors, like tree type and climate (17). The initial decomposition starts with breaking wood into small particles and the depolymerization of lignin, which is initiated by termites, other wood-eating arthropods, and fungi (18). The termites mechanically grind through the lignin to consume the cellulose with most of the lignin still intact (18). Fungi will generate radical species using several types of enzymes to depolymerize lignin. This allows the fungi to get access to cellulose to use as their energy source. Once lignin is broken down into individual monomers or small lignin substructures, in general, soil bacteria and animal

gut microbiota will further metabolize the aromatic compounds into short carbon chains that are then shuttled into the tricarboxylic acid (TCA) cycle, as shown in **Figure 3** (16).

Termites initiates the degradation of lignocellulose by mechanical breakdown of wood by the mouth. In the foregut, the wood particles mix with endogenous cellulases produced in the salivary glands and midgut of the termite (18). This releases monosaccharides that are absorbed in the midgut (18). The digested wood particles are then hydrolyzed by flagellates in the hindgut by cellulases and hemicellulases (18). The short chain fatty acids are resorbed, and the lignin-rich particles are excreted (18). A study from 2017 has shown that the termite does more pretreatment to lignocellulose than previously shown, especially the young termite workers (19). The subfamily of termites *Macrotermitinae* have a specific fungi called *Termitomyces* spp. in their nests, which together they are able to almost completely decompose lignocellulose (19). The *Termitomyces* sp. is a basidiomycete fungi. Termites over long periods of time have developed mutualistic relationships with multiple microbial symbionts, including the gut microbiota and the bacterial community that is part of the fungus comb (19). Current studies are identifying genes within termites and their symbiotic fungal and bacterial species that have oxidative properties. There are early reports that termites have laccase genes that code for oxidative enzymes (19).

Other animals are beneficial to the degradation of lignin including different species of arthropods, like beetles, millipedes, ants, and terrestrial isopods (20). The insect species, fungi, and soil microbiome together form a holobiont to efficiently sequester all of the energy from wood (21). The current paradigm is that the digestive

tract microbiome of arthropods allows the insects to degrade lignin in a mutualistic symbioses. However, the capabilities of omics technology and the growing interest in sustainable energy is motivating the discovery of new lignolytic genes in many diverse species. The many microbiomes around the degradation of wood, like soil and insect microbiomes, have not been isolated and many functions still remain unknown (22).

The largest wood-degrading fungal group belongs to the phylum Basidiomycota, generally known as white-rot fungi, with characterized enzymes from genes in the representative genera of *Pleurotus*, *Phanerochaete*, and *Ceriporiopsis* (23). In general, fungi secrete intra- and extracellular enzymes with cofactors and mediators that generate products able to perform radical decoupling. Since the wood cell walls are too small for the enzymes to penetrate, these high electron-potential mediators are responsible for the initial depolymerization of the large lignin structure. Recent studies have also discovered bacterial genes encoding ligninolytic enzymes, mostly dye-decolorizing peroxidases (DyPs) and laccases that belong to both actinobacteria and proteobacteria phylums (24). In the next sections, a general description will be given on different classes of peroxidases and laccases. This will be followed by a biphenyl and aromatic catabolic pathways review.

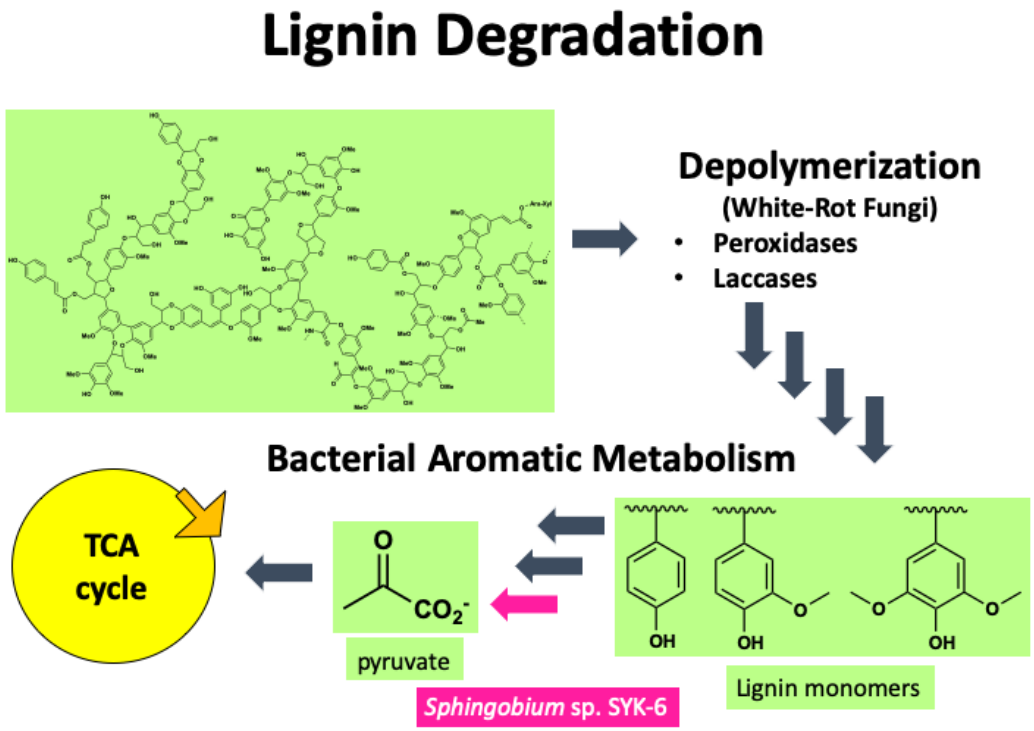


Figure 3. General overview of lignin degradation in nature. The pink arrow highlights the contribution of my research using *Sphingobium* sp. SYK-6 and its placement in the broader topic (12).

1.2.1. Peroxidases

The role of peroxidases is to depolymerize the lignin heteropolymer by cleaving the bonds shown in **Figure 2** to lignin oligomers, dimers, and monomers that can then be mineralized by bacteria into small carbon chains. The types of substrates oxidized can be generalized to different classes of lignolytic enzymes. The main contrast is whether the substrate is a phenolic or nonphenolic compound where nonphenolic substrates require a higher redox potential for bond cleavage. Heme peroxidases are one significant class of enzymes that has the function of lignin degradation. Heme peroxidases are classified into two groups based on protein sequence and structure: those that are only found in animals and those that are found in plants, fungi, and prokaryotes (25). The plant superfamily has been additionally divided into three classes based upon sequencing. Class I includes intracellular enzymes such as cytochrome c peroxidase (CcP). Class II includes the secretory fungal peroxidases whose main function is lignin degradation. Class III includes the secretory plant peroxidases like horseradish peroxidase, tobacco peroxidase, and peanut peroxidase. Class II enzymes are different from class I by post-translational modifications of glycosylation, the presence of two calcium ions in the structure, and four conserved disulfide bridges (26). The three main types of peroxidases found in fungal degradation are lignin peroxidase (LiP), manganese peroxidase (MnP), and versatile peroxidase (VP). The most recently identified class is the dye-decolorizing peroxidase (DyP) (27). There are many characteristics common to all of these peroxidases. First, they all contain a heme cofactor located in an internal cavity that is activated by hydrogen peroxide (28). Hydrogen peroxide is generated via

lignin-degrading auxiliary (LDA) enzymes like glyoxal oxidase, a member of the copper radical oxidase class, that reduces O_2 to H_2O_2 by coupling to the oxidation of aldehydes to the corresponding carboxylic acids (29). Other LDA enzymes identified are aryl alcohol oxidase, pyranose 2-oxidase, cellobiose dehydrogenase, and glucose oxidase (30).

The proposed catalytic mechanism of action begins with hydrogen peroxide traveling through a large channel to the heme center to oxidize via two electrons the resting ferric iron [Fe (III)] to form oxo-ferryl radical cation intermediate Compound I [Fe (IV) $=O^{\bullet+}$] and water (**Figure 4**) (31). Compound I is in an activated two electron enzyme form that can now oxidize the substrate through a one electron transfer step placing the radical cation onto the substrate (31). This reduces Compound I to Compound II [Fe (IV) =O]. The substrate binding site is at the opening of the large channel (31). This one electron reduction occurs again when the same or a second substrate enters causing the iron to return to its ferric resting state (31). The small channel leads to the heme propionate substrate where the manganese ion is oxidized and can diffuse out of the cell to oxidize phenolic lignin (31).

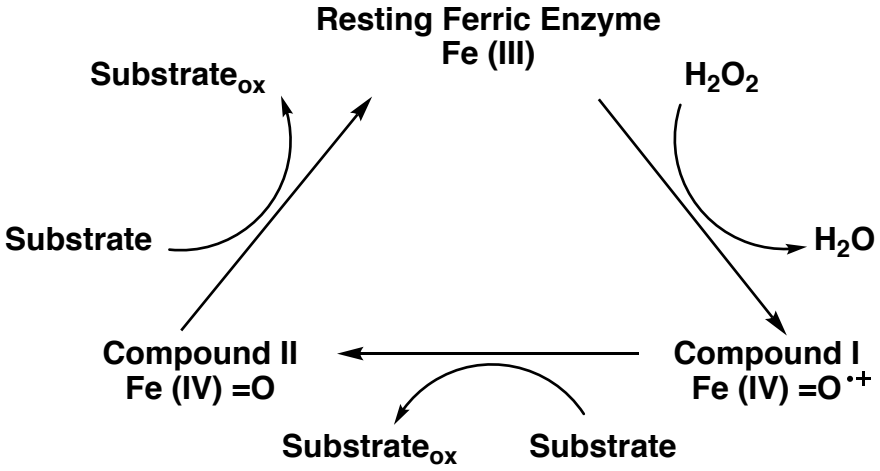


Figure 4. The general catalytic cycle of heme peroxidases (31).

1.2.1.1. Lignin Peroxidases

Lignin Peroxidase (LiP), first discovered in 1983, is a globular, monomeric glycoprotein that contains a heme prosthetic group buried in the enzyme that oxidizes phenolic and non-phenolic aromatics via heme H_2O_2 activation (**Scheme 1**). The heme is composed of one mole of iron protoporphyrin IX per mole of protein. The varying amount of glycosylation, which gives the ability to be extracellular, gives rise to many isozymes with pI range of 3.2 to 4.0 (32). The glycosylation includes the oligomannose sugar and both N- and O-glycosylation (33). LiPs use a high reduction potential ($E° = 1400$ mV) to oxidize nonphenolic structures to cation radical intermediates that nonenzymatically cleave C_α-C_β bonds (34). To date, most characterized LiPs are from the fungal species *Pleurotus eryngiii*, *Pleurotus ostreatus*, *Phanerochaete chrysosporium*, and *Arthromyces ramosus* (25).

In 1993, the first structure of LiP was determined in the white-rot fungi *Phanerochaete chrysosporium* to a resolution of 2.0 Å (PDB entry 1lga) (35). The monomer is mostly helical composed of eight major and eight minor α-helices and three short antiparallel β-sheets made of two residues per strand. It is divided into two domains by the heme plane that creates a proximal and distal side. The structural fold is similar to the intracellular class I cytochrome c peroxidase (CcP) except for an additional 67 residues on the C-terminus (residue 276 to 343) that forms an extended chain with little contact to the rest of the protein that ends at an opening to the heme active site. The RMSD between C_α-atoms of LiP and CcP (residues 15 to 275) is 1.8 Å with 18%

sequence identity. The eight cysteine residues all form disulfide bridges in LiP, where in CcP there are no bridges (**Figure 5**).

In addition to the tetradentate ligand of the porphyrin, there are two axial ligands, His-176 on the proximal side and a water molecule on the distal side. The peroxide replaces the water molecule that is further stabilized by Arg-43. The distal His-47 acts as a proton acceptor for the bound peroxide (28). Compared to other peroxidases, the iron in the heme of LiP is more electron-deficient and consequently a higher oxidation-reduction potential (36). The structure determined in 1999 to a resolution of 1.7 Å from the same species, but isozyme H2 with pI 4.15, elucidated significant structural properties about radical formation on the surface with residue Trp-171 (PDB entry 1llp) (26). The strong oxidation capabilities of LiP toward hard to cleave nonphenolic bonds of lignin come from its ability to perform chemistry on the surface of the enzyme via long-range electron transfer (LRET). The surface residue was identified as Trp-171 in the recombinant LiP isoenzyme H8 by using site-directed mutagenesis where Trp-171 was mutated to a phenylalanine and serine (**Figure 5**) (37). These two factors allows LiP to oxidize a variety of mildly activated aromatic substrates in lignin.

Scheme 1. Lignin Peroxidase Three Step Reaction. Step 1. The activation of the ferric porphyrin ring by hydrogen peroxide. Step 2. The one electron transfer activating lignin substrate. Step 3. The one electron transfer activating a second lignin substrate and enzyme returning to resting ferric state.

$$[LiP]Por\text{-}Fe^{3+} + H_2O_2 \rightarrow [LiP]Por^{\bullet+}\ [Fe^{4+} = O] + H_2O$$

$$[LiP]Por^{\bullet+}\ [Fe^{4+} = O] + Sub. \rightarrow [LiP]Por\ [Fe^{4+} = O] + Sub._{ox}^{\bullet+}$$

$$[LiP]Por\ [Fe^{4+} = O] + Sub. + 2H^{+} \rightarrow [LiP]Por\text{-}Fe^{3+} + Sub._{ox}^{\bullet+} + H_2O$$

(Adapted from Hofrichter, M. et al. (2010) (25))

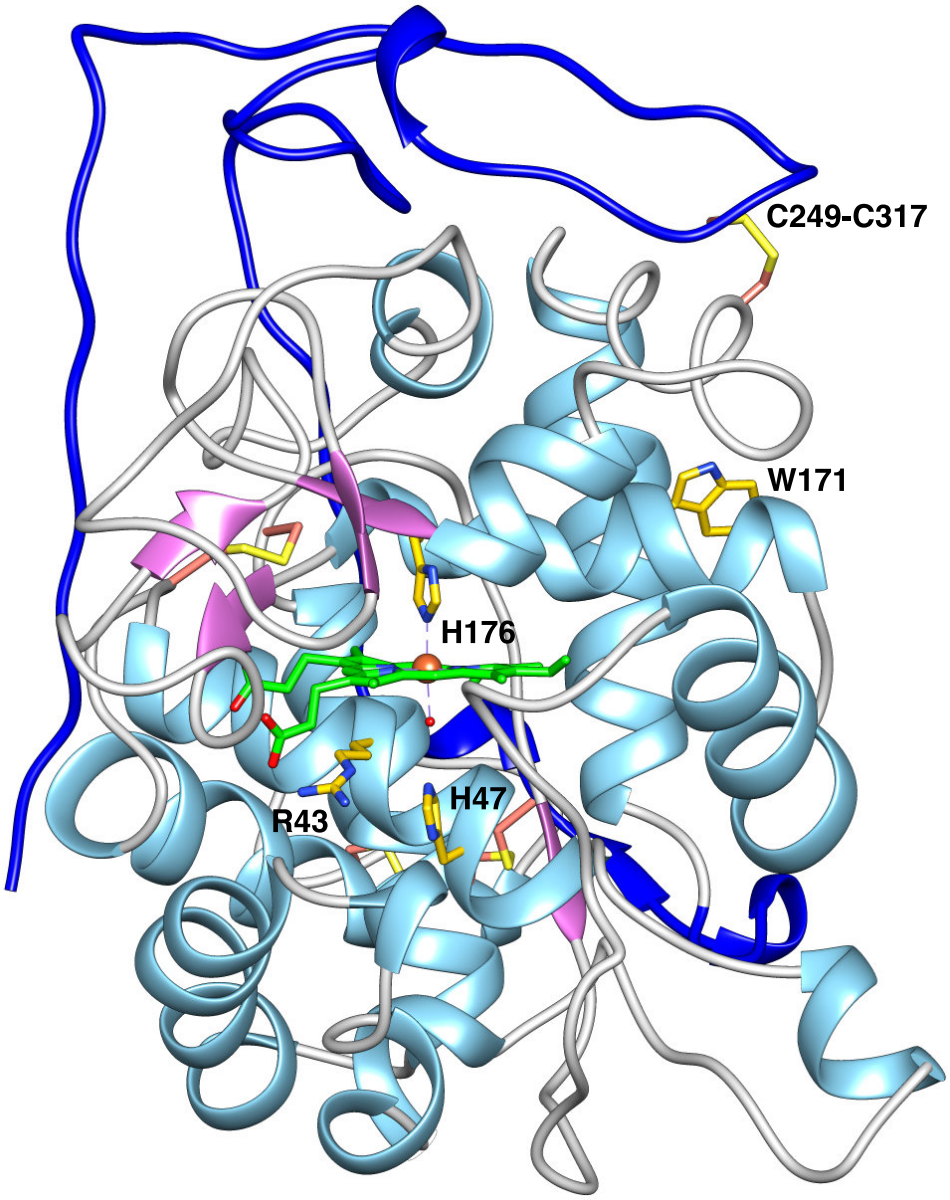


Figure 5. Ribbon model of lignin peroxidase (isozyme H2) active site from PDB entry 1llp in the species *P. chrysosporium* at 1.7 Å resolution. The α-helices are colored sky blue, the β-strands are purple, the loops are gray, and the additional C-terminus chain (residues 276 to 343) is dark blue. The porphyrin ligand is green, the iron is a bronze sphere, the residues are yellow, and the cysteine disulfide bonds are salmon color (26).

1.2.1.2. Manganese Peroxidases (MnP)

Manganese Peroxidase (MnP) is a glyosylated peroxidase that oxidizes primarily phenolic aromatic units with a redox potential of 800-900 mV (4). MnP is similar in structure to LiP in that both have one mole of iron protoporhyrin IX ring per mole of enzyme with similar properties (4). MnP has five disulfide bonds, so one more than LiP (4). The additional bridge is between Cys-341 and Cys-348, and it helps to form the Mn^{2+} binding site (**Figure 6**) (4). The Mn^{2+} is located in a cation-binding site at the surface of the protein and coordinates to Glu-35, Glu-39, Asp-179, the heme propionate oxygen, and two water molecules (4). The catalytic cycle oxidizes Mn(II) to Mn(III) by Compound I and Compound II. Then Mn(III) is the high potential intermediate that oxidizes a variety of phenolic substrates in lignin while subsequently being reduced to its resting state (4). Like LiP, the heme center of MnP is activated by H_2O_2 where it is reduced to two water molecules. Depolymerization occurs by the ability of the Mn(III) one-electron carrier to diffuse into lignin to produce phenoxy radical products. These products then randomly rearrange, cleave bonds, and nonenzymatically degrade into smaller substructures.

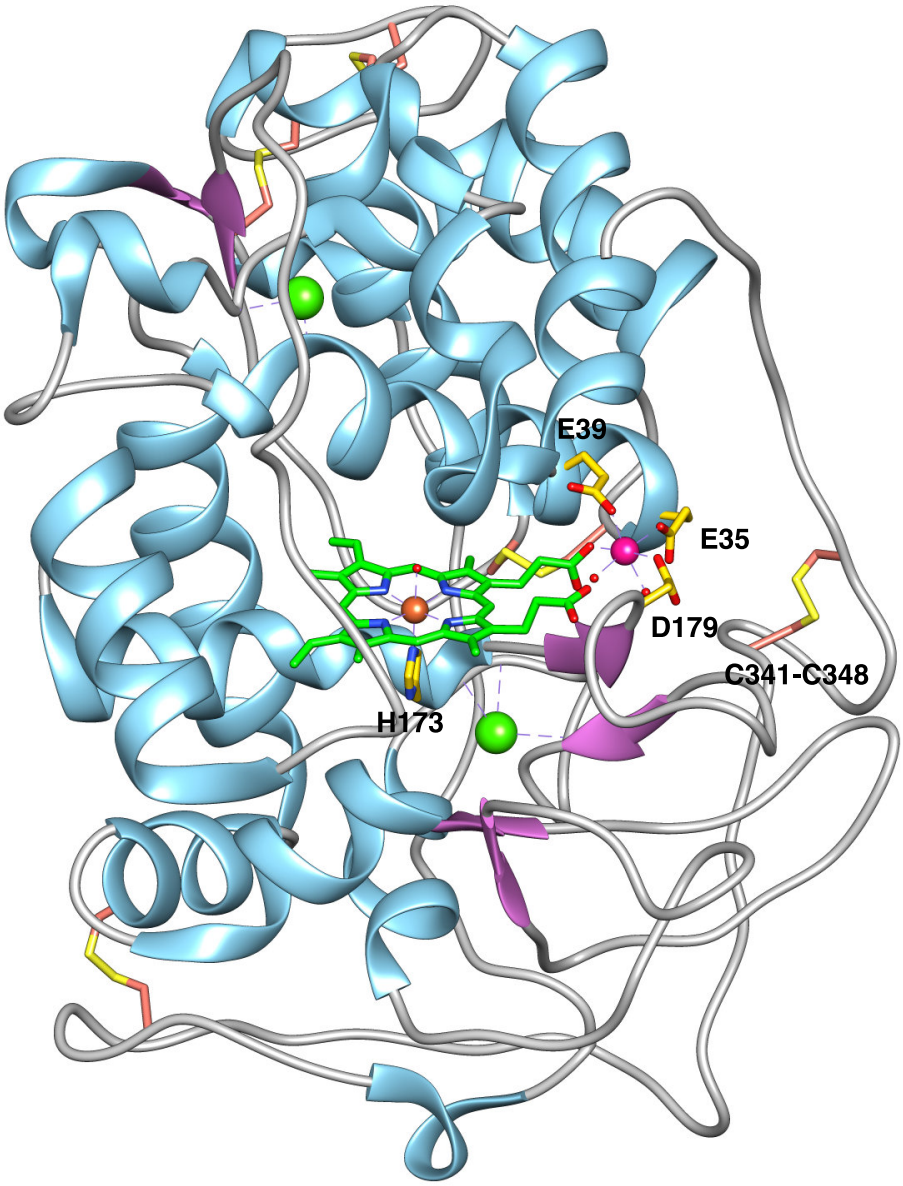


Figure 6. Ribbon model of manganese peroxidase (PDB entry 1yyd) from *P. chrysosporium* at 1.45 Å resolution (38). The α-helices are colored sky blue, the β-strands are purple, and the loops are gray. The porphyrin ligand is green, the iron ion is a bronze sphere, the calcium ions are green spheres, and the manganese ion is a pink sphere. The residues are yellow, and the cysteine disulfide bonds are salmon color.

1.2.1.3. Versatile Peroxidases (VP)

Versatile peroxidases are a hybrid between the LiP and MnP enzymes that allows a broad spectrum of substrates to be oxidized. This includes both phenolic and nonphenolic compounds and low and high redox potential substrates. It also has the ability to oxidize high redox potential dyes, specifically known for azo dye Reactive Black 5 that requires an electron potential of ~1500 mV (39). They were identified in the late 1990s, and they were first mistakenly identified as MnPs (40). It has been isolated in the following white-rot fungi species: *Pleurotus erynigii*, *Pleurotos ostreatus*, *Pleurotus pulomaius*, *Bjerkandera fumosa*, and *Bjerkandera adusta* (40). It has the ability to use two modes of oxidation. First, like LiP it has an exposed tryptophan (Trp-164) on the surface of the enzyme that creates a long-range electron transfer (LRET) pathway to the heme center to oxidize large substrates on the surface that have steric hindrance to reach the heme active site independently of a manganese cofactor (40). However, the substrate specificity for the LiP substrate veratryl alcohol, is 30-fold lower than the ability by LiP, which its value is 3×10^4 M^{-1} s^{-1} (40). The surface environment around the tryptophan is positively-charged by basic residues that favors anionic substrates, whereas the surface of LRET for LiP is negatively-charged by acidic residues (39). This requires LiP to use mediators; thus making VP a more attractive peroxidase in an industrial setting (39). The second mode of action is like MnP where it oxidizes manganese from +2 to +3 oxidation state in a coordination sphere composed of three tricarboxylate groups from glutamate and aspartate residues near the propionate units of the heme (39). Site-directed mutagenesis studies have shown that the coordination in

VP is more stable than in MnP (41). These reaction details were elucidated based on the structure of VP from *P. eryngii* (PDB entry 2boq), which was determined to a resolution of 1.3 Å (41).

1.2.1.4. Dye-Decolorizing Peroxidases

Dye-decolorizing peroxidase (DyP) were first identified in *Bjerkandera adusta* in 1999 and has since been widely observed in bacterial species (42). It is a highly stable glycosylated heme protein and has been divided into four classes. Like the other peroxidases, it is activated by H_2O_2 to catalyze the oxidation of substrate even though the structure and sequence of DyPs are different from other peroxidases (27). Structurally, it is different from other peroxidases because the distal histidine residue at the heme is an aspartate residue (27). This location is responsible for H_2O_2 binding; therefore, it is suggested that it now proceeds through an unique mechanism for activation (27). Recent studies have shown that it can oxidize substrates with molecular oxygen in the absence of hydrogen peroxide (43). Also, some types of DyPs don't need Mn (II) as a mediator (44).

1.2.2. Laccases

Laccases (Lac) contain four copper ions that perform oxidation reactions in the presence of molecular oxygen by four one-electron abstraction from phenolic substrates to produce phenolic radicals and two molecules of water as a byproduct (**Figures 7** and **8**) (28). Laccases belong to the Cupredoxin Superfamily (IPR008972), which is a large

and broad family with more than 240,000 proteins in the InterPro database (45). The signature character of the Cupredoxin Superfamily is containing a mononuclear type one copper site causing the protein to have an intense blue color (45). Structurally, the superfamily has a β-sandwich composed of seven β-strands that form two β-sheets arranged in a Greek-key β-barrel (**Figure 9**) (45). The β-strands are adjacent in space but not in sequence. Within this superfamily, laccases are part of the multi-domain family. This family has members identified with two to six domains (45). They are glycoproteins that vary in molecular weight (50-97 kDa), oligomeric state (monomeric, dimeric, and tetrameric), and isoforms (23). The carbohydrates, generally mannose, cause the enzyme to be stable (46). They are covalently linked and compose 10-45% of the molecular weight (4). They have been identified in plants, fungi, bacteria, and insects, but most studies have been on fungal laccases with over 100 fungal enzymes characterized. Bacterial laccases were first described in 1995 in *Azospirillum lipoferum*, a bacterium associated with plant root (47). It now has been identified in a number of other bacteria, which includes *Escherichia coli, Bacillus subtilis, and Pseudomonas syringae*. One of the best studied bacterial laccases is the protein CotA from *Bacillus subtilis* (48). In fungal species the enzymes are both intra- and extracellular, but in bacterial laccases, they are mostly intracellular (4).

Laccases are also identified into the multicopper oxidases (MCO) type two domain (IPR011706) (49). In this domain all enzymes have a conserved architecture of the two active sites composed of four copper ions for both bacterial and fungal laccases. There are three types of spectroscopically different copper ions. The type one (T1) blue

copper forms the mononuclear site that is near the surface, but not solvent exposed (50). The bacterial laccases at T1 has three conserved equatorial ligands (His-419, Cys-492, and His-497 (numbering based on PDB entry 1w6w from *Bacillus subtilis*)) with coordination distances of ~2.0 Å (51). One of the axial positions is vacant and the other one is Met-502 that does not participate in the coordination at a longer distance of ~3.7 Å. T1 copper therefore has a distorted bipyramidal trigonal geometry (48). The fungal laccases have the same conserved ligands that are coordinated (His-394, Cys-450, and His-455, numbering based on PDB entry 2hrg from *Trametes trogii*) to the copper ion with coordination distances of ~2.0 Å (50). In this species, the axial ligand is Ile-452 at a distance of 3.6 Å. The fungal laccases have variable residues for this position in which leucine and phenylalanine have also been identified (52). The blue color is due to the intense electronic absorption at ~600 nm due to a ligand-to-metal (Cys S → Cu^{2+}) charge-transfer transition (53). The T1 copper is also paramagnetic based on a small copper hyperfine coupling constant ($A_{\parallel} \approx$ 4-10 mT) in electron paramagnetic resonance (EPR) (53). The type two (T2) copper has no strong absorption, but it is paramagnetic based on a copper hyperfine coupling constant ($A_{\parallel} \approx$ 20 mT) in EPR (53). It is coordinated to two histidine residues (His-105 and His-422 in *B. subtilis*) and one water molecule. The other two copper ions are classified as T3 coppers where they form a dinuclear center bridged by a hydroxy group in a linear arrangement. They are characterized by an electronic absorption at 330 nm, and they are EPR silent due to antiferromagnetic coupling between the ion pair (53). The T2 and T3 copper ions form the trinuclear cluster for the second active site. Most laccases have three domains;

however, in 2009 the first two-domain laccase structure was determined (54). Currently, there are 31 structures deposited onto PDB database, with 30 being from bacterial species (55). Within the three domain monomer, domain three contains the mononuclear copper site. The trinuclear copper site is between domains one and three. For *B. subtilis* (PDB entry 1w6w and 1w8e and 1gsk) domain one is residues 2-176, connection one is residues 177-182, domain two is residues 183-340, connection two is residues 341-368, and domain three is residues 369-501 (**Figure 9**) (48, 51).

The proposed reaction mechanism for laccases is initiated by oxidation of the phenolic substrate at the T1 site where an electron is transferred to the copper center to generate the phenolic radical product (45). The oxidation state of the copper changes from a +2 to +1 state. The electron is transferred from the T1 site to the T2/T3 site by an internal His/Cys pathway where the three coppers at this site are also reduced to a +1 oxidation state. In *T. trogii*, the electron transfer pathway is composed of the T1 Cys-450 and two T3 ligands, His-449 and His-451 (50). This monoelectronic oxidation occurs a total of four times to produce at the T2/T3 site the activation and reduction of molecular oxygen to two molecules of water. The coppers then return to the resting state with a +2 oxidation state to repeat the cycle. The electrons travel ~13 Å between the two active sites (50). The T2/T3 site has two solvent channels where one channel is for molecular oxygen to enter (closer to the T3 coppers) and the other channel is for water to exit (near the T2 coppers) (51).

Laccases can easily oxidize aromatic substrates that have electron-withdrawing groups like phenols and anilines that require a mild redox potential (56). The substrate

specificity of laccases includes more substrates by employing mediators in a laccase-mediator system (LMS) (52). Mediators are low molecular weight compounds that are oxidized at the T1 site to stable high potential electron carriers (52). Mediators are then able to diffuse into lignin to oxidize compounds with an electron potential greater than ~1100 mV while subsequently being reduced by the substrate via radical coupling (57). Normally, at the T1 site, the electron potential maximum is ~760 mV and ~455 mV for fungal and bacterial laccases, respectively (52, 58, 59). Now over 100 phenolic and non-phenolic compounds of various weights are considered substrates for laccases (60). Both natural and synthetic mediators have been identified, with the synthetic mediator typically having a higher redox potential (60). However, desired properties of mediators include being non-toxic, economic, efficient, and able to maintain the redox cycle without inactivating the enzyme. Therefore, natural mediators with lower redox potential are typically more attractive since they have no unfavorable properties. The most popular synthetic mediators include 2,2'-azinbobis(3-ethylbenzthiazoline-6-sulfonate) (ABTS), 1-hydroxybenzotriazole (HBT), and the stable 2,2,6,6-tetramethyl-1-piperidinyloxy free radical (TEMPO) (**Figure 10**) (60). These mediators use different mechanism to oxidize the substrates: ABTS (E°: 1090 mV) uses electron transfer, HBT (E°: 1100 mV) uses hydrogen atom transfer, and TEMPO (E°: >1000 mV) uses ionic oxidation (60). Natural mediators produced by the fungi and bacteria include acetosyringone, acetovanillone, syringaldehyde, vanillin, *p*-coumaric acid, sinapic acid, and ferulic acid (**Figure 10**) (60).

Laccases are very attractive green biocatalyst when compared to other oxidases. First, they don't require added cofactors and the byproduct is water compared to lignin delignification peroxidases which require hydrogen peroxide. Also, there are bacterial laccases that can be exploited instead of fungal species, which on an industrial scale, are easier to grow due to greater pH range tolerance, more thermostable, and faster growth rates (60).

The utilization of these oxidases (peroxidases and laccases) in industrial applications are highly desirable with each class of enzyme having both advantages and disadvantages when compared to each other (**Table 1**). LiP has the broadest range of substrates and the largest redox potential, but maintaining the environment necessary for optimal fungal activity is challenging and costly. MnP has the advantage as being the most prominent enzyme found in fungal species genome and so using more efficient species is something that can be explored. VP is attractive in that it has natural abilities to perform both LiP and MnP reactions. DyP and Lac are naturally found in bacterial species, and so it eliminates the troubles of growing fungi on a large-scale, but their abilities to completely depolymerize lignin alone doesn't compete with fungal peroxidases (61). Therefore, the current field is actively identifying new species to hopefully find more efficient chemistry pathways and engineering current enzymes to improve abilities to find a suitable system that can be industrially scaled.

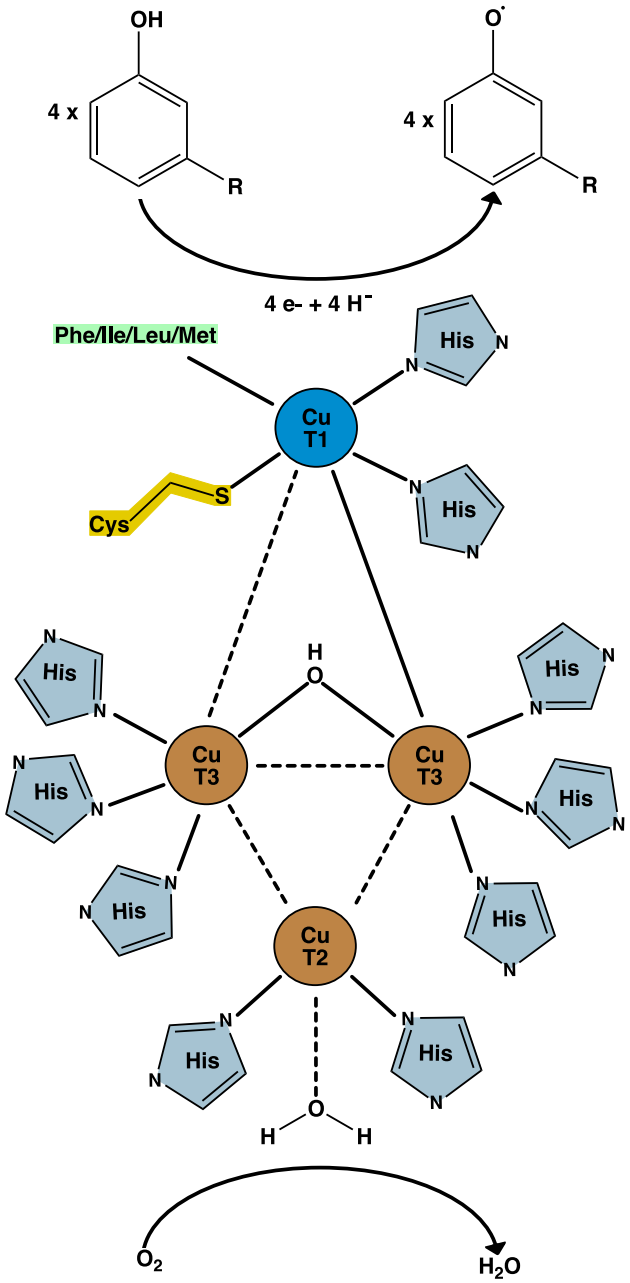


Figure 7. The active site of laccase with catalytic cycle shown (50).

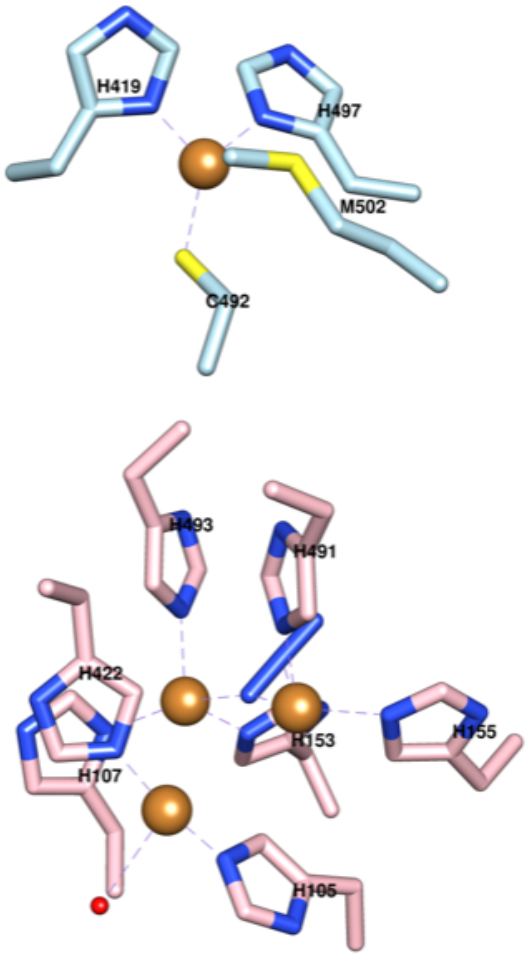


Figure 8. The model of the bacterial active site of laccase (PDB entry 1w6w) (51).

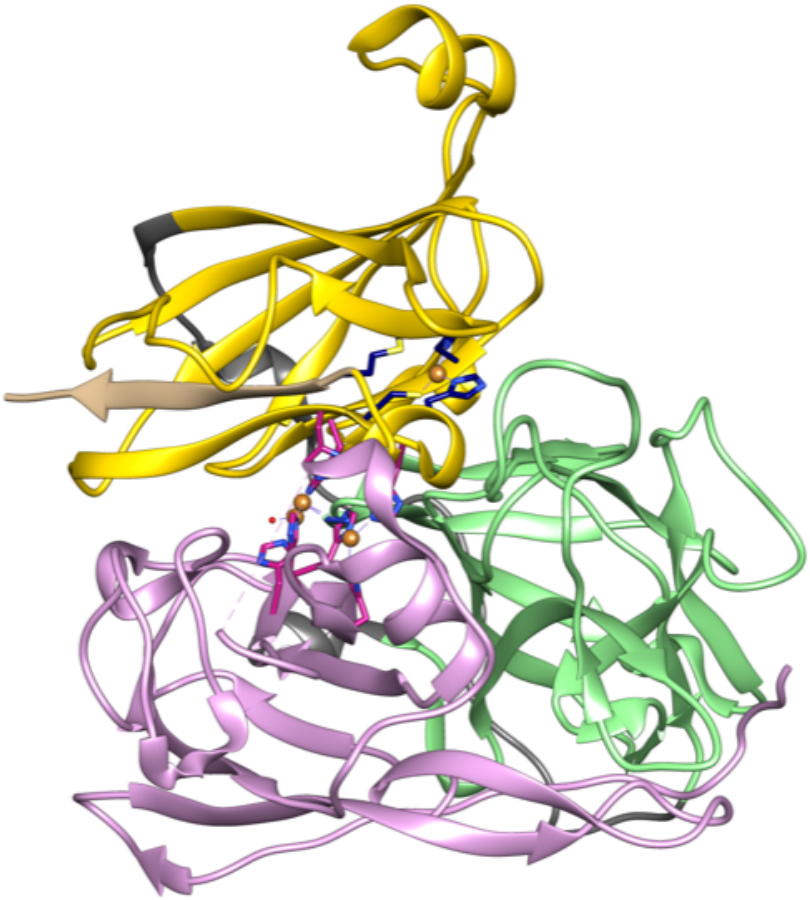
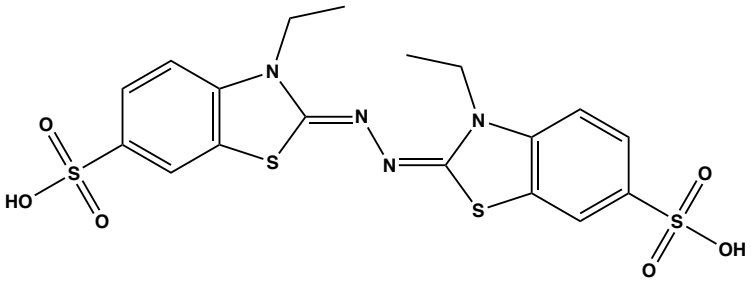
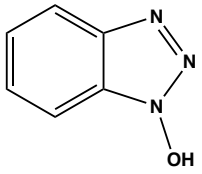


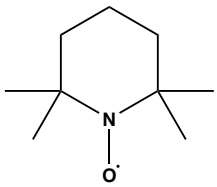
Figure 9. The model of the bacterial laccase with the three domains colored differently (PDB entry 1w6w) (51).



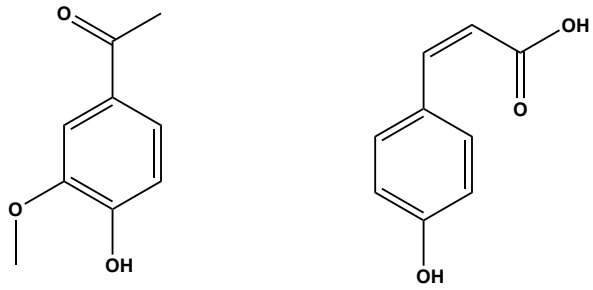
2,2'-azinbobis (3-ethylbenzthiazoline-6-sulfonate) (ABTS)



1-hydroxybenzotriazole (HBT)



2,2,6,6-tetramethyl-1-piperidinyloxy free radical (TEMPO)



vanillin

p-coumaric acid

Figure 10. Synthetic and natural mediators for laccases. (60)

Table 1. Comparison of the properties of lignin peroxidase (LiP), manganese peroxidase (MnP), versatile peroxidase (VP), dye-decolorizing peroxidase (DyP), and laccase (Lac).

	LiP	**MnP**	**VP**	**DyP**	**Laccase**
Glyoprotein	yes	yes	yes	yes	yes
Active Site	heme	heme Mn binding site	heme Mn binding site	heme	4 copper atoms
Reaction Activator	H_2O_2	H_2O_2	H_2O_2	H_2O_2 O_2	O_2
Mediators	none	Mn	Mn	Mn	phenolic compounds
Substrate	phenols, nonphenolic aromatics	phenols	Phenols, nonphenolic aromatics	Phenols, nonphenolic, veratryl alcohol	phenols
Active Site Location	LRET	Mn binding site	Mn binding site and LRET	heme	T1 copper site
pH range	2.0-5.0	2-6	3.5-6.5 (dependent on substrate)	3.2	2.0-8.5
Redox Potential Maximum	1500 mV	900 mV	1500 mV	1200 mV	400 mV (bacterial) 1000 mV (fungal) 1100 V (mediators)
Species	fungi	fungi	fungi	bacteria	bacteria fungi

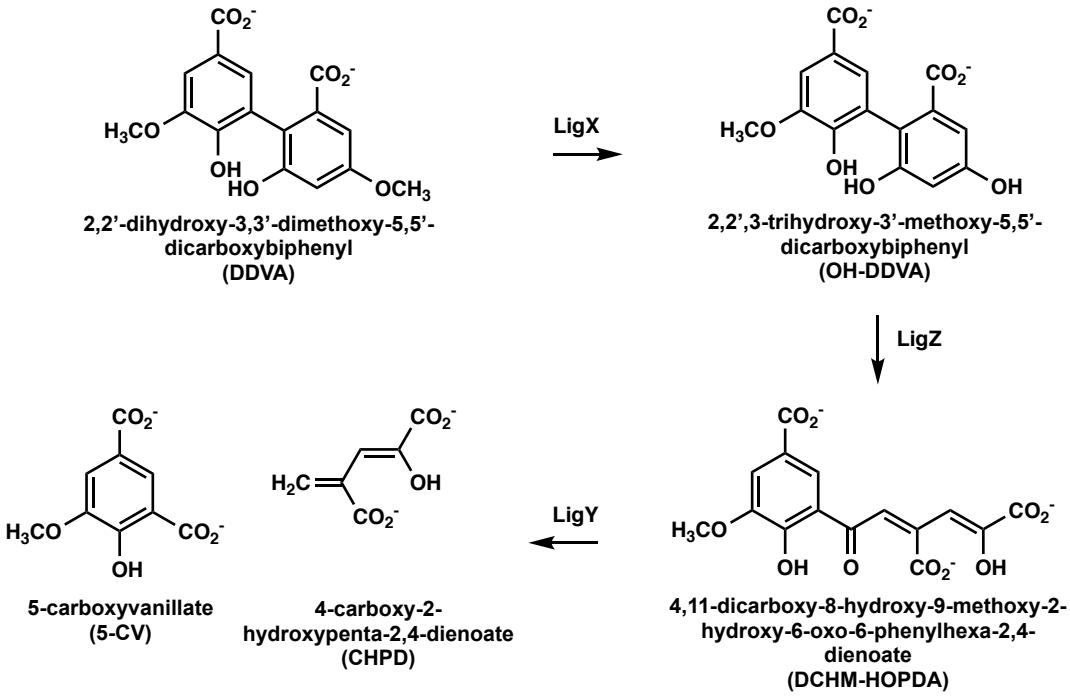
1.2.3. The Catabolic Pathway of Biphenyl (LigX, LigZ, and LigY)

Once lignin has been depolymerized by mostly white-rot fungi peroxidases and some bacterial laccases, an assortment of bacterial pathways have evolved to metabolize these lignin oligomers. One example is the biphenyl catabolic pathway of soil bacterium *Sphingobium* sp. SYK-6 (62).

Referring to **Figures 1** and **3**, the oxidation of lignin by fungal enzymes decomposes the polymer into three general types of aromatic monomers: syringyl, guaiacyl, and *p*-hydroxyphenyl units. Now there is some ambiguity in the process of degradation in substructures of lignin mainly composed of dimer units of these monomers. There are both fungal and bacterial species that have been shown to cleave these dimer bonds; however, a full characterization of these enzymatic processes have not been completely performed. These aromatic units are metabolized through different pathways with different cleavage sites. Some of these enzymatic reactions overlap in the different pathways, especially for the last few steps before the final products are made. To illustrate one of many metabolic degradation pathways of dimer cleavage, the biphenyl catabolic pathway from the bacterium *Sphingobium* sp. will be described (**Scheme 2**) (62).

The biphenyl catabolism in an α-proteobacterium, *Sphingobium* sp. SYK-6 occurs through the enzymatic reaction of LigX, LigZ, and LigY. These reactions require a C-C bond to be cleaved in a *meta*-fashion (62).

Scheme 2. Biphenyl Catabolic pathway in *Sphingobium* sp. SYK-6.



LigX is a three-component monooxygenase that requires Fe (II) to catalyze the *O*-demethylation of 5,5'-dehydrodivanillate (DDVA) to 2,2',3-trihydroxy-3'-methoxy-5,5'-dicarboxybiphenyl (OH-DDVA) (63). The genes encoding it are *ligXa, ligXc,* and *ligXd* (64). The conversion occurs in the presence of NADH. LigXa is the oxygenase component, LigXc is the ferredoxin, and LigXd is a ferredoxin reductase, which requires NADH (64). LigXa has a molecular weight of 49.7 kDa and is a homotrimer; LigXc has a molecular weight of 18 kDa and is a monomer; and LigXd has a molecular weight of 54 kDa and is a monomer (64). LigXd contains a FAD binding domain and indeed FAD is the prosthetic group that has one to one stoichiometry with the monomer of LigXd. Several substrates were assayed, but the enzymes showed specificity for DDVA O-demethylation with kinetic values of ~60 μM and 6 s^{-1}, for K_m and k_{cat}, respectively with optimal mixing ratios of LigXa, LigXc, and LigXd at 1:30:2 (64).

LigZ is a *meta*-ring cleavage extradiol dioxygenase from the *Sphingobium* sp. SYK-6 that transforms 2,2',3-trihydroxy-3'-methoxy-5,5'-dicarboxybiphenyl (OH-DDVA) into 4,11-dicarboxy-8-hydroxy-9-methoxy-2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate (DCHM-HOPDA) in the presence of molecular oxygen (65, 66). It is a 39 kDa type II extradiol dioxygenase that is Fe(II) dependent, which includes the downstream enzyme, LigB protocatechuate 4,5-dioxygenase homolog (~20% sequence identity). The substrate specificity for OH-DDVA is 2.2 x 10^7 s^{-1} M^{-1} at pH 7.5 (66).

LigY catalyzes the *meta*-cleavage hydrolysis of OH-DDVA to 5-carboxyvanillate (5CV) and 4-carboxy-2-hydroxypenta-2,4-dienoate (CHPD) (67, 68). It is a members of the amidohydrolase superfamily (AHS) with a mononuclear zinc ion

in the active site coordinated to the conserved residues His-6, His-8, His-179, and Glu-282 (68). Several other enzymes in this pathway (LigW decarboxylase, LigI hydrolase, and LigJ hydratase) are also members of the prominent superfamily that has over 100,000 unique protein sequences. It is a 37 kDa protein that contains one mole of zinc per protomer and is a hexamer in solution (68). The apo-structure of LigY was determined to a resolution of 1.9 Å and is characterized as a trimer of dimers (68). It has the significant structural fold of the AHS with a distorted $(\beta/\alpha)_8$-barrel fold. It was structural similar to LigJ hydratase with a RMSD of 1.6 Å over 309 C-α atoms; the major difference is that LigY has five additional α-helices (68). The kinetic parameters of LigY hydrolase at pH 7.5 was 9 s^{-1}, 0.38 μM, and 25 s^{-1} μM^{-1} for k_{cat}, K_m, and k_{cat}/K_m, respectively (68). The proposed reaction mechanism is an enol-keto tautomerization followed by *gem*-diol intermediate and C-C bond fission (68). Docking studies suggest that the C1 carboxylate of the substrate coordinates to the zinc ion and Arg-234 (68). It is proposed that the Arg-234 induces strain on the dienolate during the tautomerization (68). The residue Arg-72 is predicted to interact with the 4-carboxylate of the substrate and stabilizes the oxyanion hole during the tetrahedral intermediate (68). The proposed catalytic base is Glu-282, which directly aligns with other mechanisms in the AHS. It deprotonates the C2-hydroxyl of DCHM-HOPDA in the first half of the reaction and donates a proton to the dienoate leaving group in the second half of the reaction (68). The role of the conserved His-223 is to stabilize the intermediate via hydrogen bonding (68).

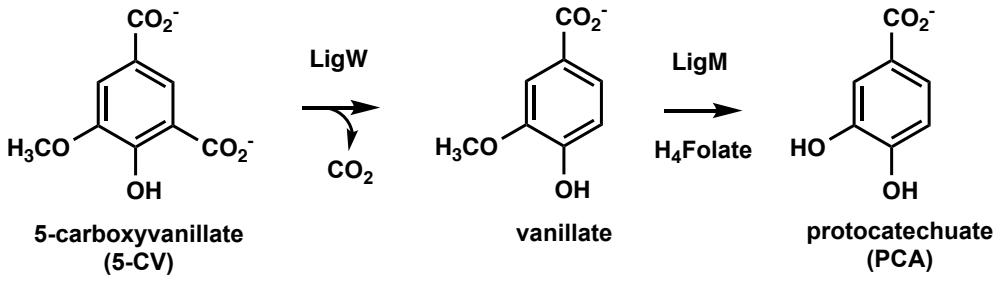
1.2.4. Prerequisite Reactions to Ring-Cleavage (LigW and LigM)

LigW catalyzes the nonoxidative decarboxylation of 5-carboxyvanillate (5CV) to vanillate in the presence of Mn (II) (69). Vanillate is formed by the cleavage of a carbon-carboxylate bond and the formation of a new carbon-hydrogen bond (69). LigW is a 38 kDa protein that is a 32-mer in solution with 1 mole of Mn(II) per subunit for *Sphingobium* sp. SYK-6 and exists as a tetramer for *Novosphingobium aromaticivorans* (69). The kinetic constants for the decarboxylation of 5CV by LigW were 2 s^{-1}, 55 μM, and 4 x 10^4 M^{-1} s^{-1} for k_{cat}, K_m, and k_{cat}/K_m, respectively (69). The proposed mechanism is initiated by the binding of substrate in the active site, which displaces two water molecules from the Mn (II) center (69). The substrate coordinates in a bidentate fashion with a hydroxyl group and carboxylate group (69). This facilitates an out-of-plane bending of the substrate by the enzyme with structural evidence to a resolution to 0.96 Å in *N. aromaticivorans* (69). The proton from the hydroxyl group is subsequently transferred to Asp-314 or directly to C5 (69). The last step is the decarboxylation that follows the protonation of C5 of the substrate (69).

The substrate protocatechuate is primed for 4,5-cleavage pathway of *Sphingobium* sp. SYK-6 by the *O*-demethylation of vanillate by LigM in the presence of the cofactor tetrahydrofolate (H_4folate) (70). The reaction is dependent on a tyrosine residue that is responsible for the binding of folate in the active site (71). LigM is also able to transfer the methyl group from 3-*O*-methylgallate to H_4folate to produce 5-methyl-tetrahydrofolate (5-methyl-H_4folate) (71). The substrate binding site is deep within the enzyme where the substrate and H_4folate are close to each other (71). Both

the substrate and cofactor enter through the same solvent channel (71). It is proposed that the substrate enters first and then H_4folate binds, which blocks the substrate until the product is made (71). This gives an ordered, sequential kinetic model (71). This reaction is dependent upon the hydroxyl group of Tyr-247 to stabilize the substrate during methyl removal by interaction with the oxygen of vanillate while H_4folate-N_5 nucleophilically attacks the methyl group (71). Following the transfer, Tyr-247 donates a proton to the demethylated substrate (71). Gln-57 aids in proton donation to Tyr-247 (71). This catalytic tyrosine-dependent demethylase is unique to demethylases (71). The *O*-demethylation of vanillate by LigM has a k_{cat} and k_{cat}/K_m of ~5 s^{-1} and 9 s^{-1} mM^{-1}, respectively at pH 8.0 (71).

Scheme 3. The reaction of LigW decarboxylase and LigM *O*-demethylase.



1.2.5. PCA 4,5-Cleavage Pathway

The PCA degradation pathway includes three different cleavage sites, 2,3-cleavage; 3,4-cleavage; and 4,5-cleavage (62). In the 4,5-cleavage pathway in the bacterium *Sphingobium* sp. SYK-6, protocatechuate is converted to pyruvate and oxaloacetate by the combined actions of LigAB, LigC, LigI, LigU, LigJ, and LigK (62). LigAB dioxygenase converts protocatechuate (PCA) into 4-carboxy-2-hydroxymuconate-6-semialdehyde (CHMS) in the presence of molecular oxygen (72). LigC dehydrogenase converts CHMS into 2-pyrone-4,6-dicarboxylate (PDC) while subsequently reducing NADP+ to NADPH (73). LigI hydrolase catalyzes the ring-opening of the lactone PDC to (4*E*)-oxalomesaconate (OMA) (74). Our recent findings have determined the role of LigU as an isomerization in the conversion of OMA into (3*Z*)-2-keto-4-carboxy-3-hexendioate (KCH) (75, 76). We also propose the catalytic mechanism of LigJ hydratase in the conversion of KCH into (*S*)-4-carboxy-4-hydroxy-2-oxoadipate (CHA) (77). The last enzymatic step is the retro-aldol cleavage of CHA into pyruvate and oxaloacetate by LigK aldolase (78-80).

LigAB protocatechuate (PCA) 4,5-dioxygenase catalyzes the extradiol aromatic ring cleavage of PCA by insertion of molecular oxygen across the C4-C5 bond of the aromatic ring to form 4-carboxy-2-hydroymuconate-6-semialdehyde (CHMS) (72). It is a member of the dioxygen-activating mononuclear non-heme iron (II) superfamily that contains a 2-His-1-carboxylate facial triad (81). It is further classified into the largely uncharacterized family of extradiol-cleaving catechol dioxygenases (81). A major distinction between the reaction being extradiol or intradiol is the oxidation state of the

iron center with Fe (II) and Fe (III), respectively. Its activity is optimum at pH 7.5 (82). LigAB has an $\alpha_2\beta_2$ subunit composition with the motif **H**X_{48}**H**X_{180}**E** and molecular weight of 97 kDa (81, 83). The α subunit is small (139 residues) and plate-like while the β subunit is large (302 residues) with an α_9/β_{11} structural fold (83). The active site is in a deep cleft in the β subunit (83). The iron (II) ion is coordinated to the conserved residues His-12, His-61, Glu-242 in a distorted trigonal pyramid that stabilizes the ion to the active site (83). It has weak interactions with Asn-59, which is not a conserved residue (83). It is suggested that the different subunit compositions of extradiol dioxygenases contributes to the different substrate specificities for enzymes in this class (83). The mechanism of action initiates when substrate binds to the Fe (II) center in a bidentate fashion with the adjacent hydroxyl groups displacing water molecules bound to the Fe (II) center in the apo form (81, 84). Once ligand is bound in the active site, a lid composed of helices from the α subunit form over the pocket making it inaccessible to solvent (83). The hydrophobic pocket makes contact with the aromatic ring of the substrate, thus aiding in the stabilization (83). Other substrates lacking the carboxylate substituent of PCA (like catechol) have little activity with LigAB (82). Therefore, the ionic interactions that the carboxylate group has with nearby residues (Ser-269 and Asn-270) seems to be essential for substrate binding (82, 83). The two Fe-O coordination with PCA have asymmetric binding, which gives different oxygen states (83). It is predicted that the longer distance is a hydroxyl oxygen while the shorter distance is an oxyanion (83). Once substrate is bound, then the molecular oxygen can coordinate to the metal center opposite to any of the amino acid ligands (83). An electron transfer

from Fe (II) to O_2 forms the Fe (III) oxygen radical species (83). An electron is then transferred from the substrate to the metal ion reforming Fe (II) and now a radical substrate (83). The radical substrate and radical oxygen form a C-O bond, followed by Creigee rearrangement, and O-O bond cleavage to give a lactone intermediate and a hydroxide ion on the metal ion (85). This hydroxide ion hydrolyzes the lactone to yield CHMS. The chemistry is stabilized by hydrogen bond interactions with two histidine residues and a tyrosine residue (83). The CHMS is in equilibrium between the chain and cyclic hemiacetal form. It is the hemiacetal form of CHMS that is predicted to be used in the next enzymatic reaction of LigC dehydrogenase (73).

LigC dehydrogenase converts CHMS into PDC while reducing NADP+ to NADPH (73). It is a homodimer with a molecular weight of 34.6 kDa with optimum activity at pH 8.0 (73). It belongs to clusters of orthologous group 673 (COG673), which is a cog of predicted dehydrogenases. Its closest homolog with structural data is WlbA from *Bordetella pertussis* with 22% sequence identity (86) and with genetic identification is iolU from *Bacillus subtilis* with 26% sequence id (87). It has an approximate K_m of 25 μM for $NADP^+$, which is 10 times lower than for NAD^+ and for CHMS in the presence of $NADP^+$ it is 26 μM (73). When the *ligC* gene was inactivated in the *Sphingobium* sp. SYK-6 cells growth on vanillate was prevented (73). PCA did accumulate in the cells with this deletion which indicated inactivity of LigAB dioxygenase (73). Therefore, it seems that PDC is an inducer for *ligAB* gene expression (73). To date, only the gene has been characterized; there is not elucidation of the structural and mechanistic details of LigC dehydrogenase.

LigI hydrolase catalyzes the ring opening of PDC by converting it to (4*E*)-oxalomesaconate (OMA) (74, 75). LigI is the first identified member of the amidohydrolase superfamily that requires no metal cofactor for catalysis (74). LigI has a molecular weight of 34 kDa and is a member of cog3618 (74). The kinetic constants for the hydrolysis of PDC at pH 8.25 are 340 s^{-1} and 9.8 x 10^6 M^{-1} s^{-1} for k_{cat} and k_{cat}/K_m, respectively (74). At pH 8.5 there is an equal mixture of substrate and product (74). The activity is monitored by the decrease in the absorption maximum of PDC at 312 nm (74). The mechanism of action of LigI hydrolase is initiated by the activation of a water molecule by conserved residue Asp-248 (74). The carbonyl carbon of PDC is the location of nucleophilic attack by the activated water molecule (74). The carbonyl group is stabilized during the reaction by interactions with His-180, His-31, and His33 (74). The cleaving of the carbon-oxygen bond opens the lactone ring to (4*E*)-OMA (74, 75).

LigU isomerase was recently identified and characterized to convert (4*E*)-OMA into (3*Z*)-2-keto-4-carboxy-3-hexendioate (KCH) (75). It was previously annotated as a tautomerase (88, 89). Results from ^1H and ^{13}C NMR spectroscopy clearly show the migration of the double bond between C4-C5 migrating to C3-C4 via a 1,3-allylic isomerization while identifying the configuration around the double bond (75). Further studies that include site-specific mutagenesis and X-ray crystallography identified Lys-66 and Cys-100 as critical residues for catalysis in a *syn*-fashion (76). The kinetic constants in the isomerization of OMA to KCH by LigU at pH 8.0 were 1300 s^{-1} and 7.7 x 10^6 M^{-1} s^{-1} for k_{cat} and k_{cat}/K_m, respectively (75). LigU has a molecular weight of 37 kDa and is a homodimer (75). Structurally, each monomer has two domains that

contains the diaminopimelate (DAP) epimerase fold (75). This fold has a central α-helix that is wrapped by antiparallel β-strands that form a β-barrel (75).

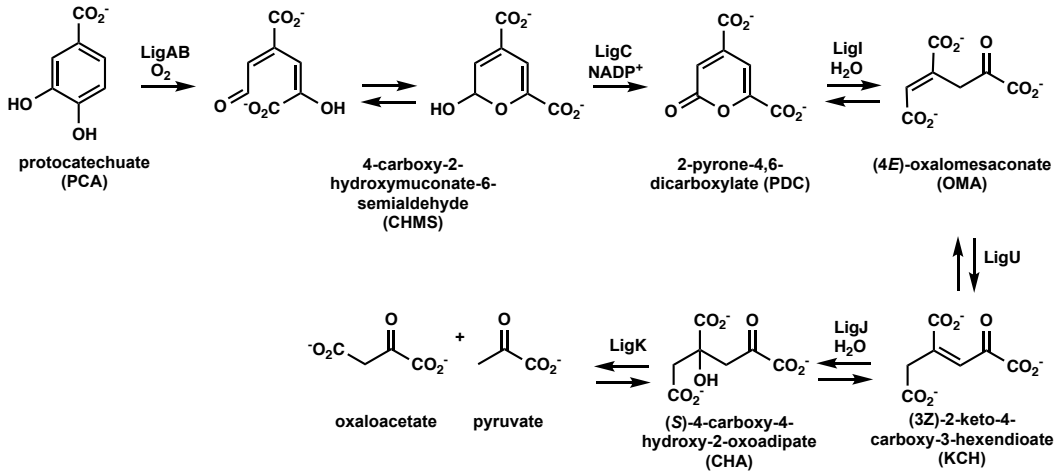
LigJ hydratase catalyzes the hydration of KCH to (*S*)-4-carboxy-4-hydroxy-2-oxoadipate (CHA) (77). It is the first member of the amidohydrolase superfamily characterized with a hydration reaction (77). It has the signature AHS structural fold, which is a distorted $(\beta/\alpha)_8$-barrel, and the catalytic base (Glu-284) is from the end of β-strand 8 (77). The mechanism of action begins with Glu-284 acting as the base to activate the water molecule for nucleophilic attack at C4 (77). The structure of wild-type LigJ with product CHA shows electron density for this addition in the *S*-configuration by X-ray crystallography (77). Kinetic analysis of the mutant E284Q showed no activity was monitored (77). Using this mutant, co-crystallization with substrate was successful, which suggested that Glu-284 was also the catalytic acid that protonated position C3 in the second half of the reaction to give a *syn*-addition reaction (77). The kinetic constants for the hydration of KCH to CHA by LigJ at pH 8.0 are 25 s^{-1} and 2.6 x 10^6 M^{-1} s^{-1} for k_{cat} and k_{cat}/K_m, respectively (77). LigJ has a molecular weight of 38 kDa and is a homodimer (77).

LigK aldolase catalyzes the final reaction in the PCA 4,5-cleavage pathway of *Sphingobium* sp. SYK-6 by the retro-aldol cleavage of CHA into pyruvate and oxaloacetate (78-80). It is a member of the regulator of ribonuclease activity A (RraA)-like superfamily (80). It is named after the first member of this family was identified in *Escherichia coli* where the protein RraA binds and modulates the RNA degradosome and inhibits RNase E activity (80). This family has an αββα sandwich fold (80). LigK

has a molecular weight of 24 kDa and is a hexamer or dimer of trimers (80). LigK is a member of the Class II pyruvate aldolases where a divalent metal ion is used to stabilize an enolate intermediate as opposed to the Class I family that utilizes a Schiff base mechanism to stabilize a carbanion intermediate (80). Class II belongs to cog0684.

Based upon previous studies on a homolog with 60% sequence identity from *Pseudomonas putida* F1, the proposed mechanism shows the second half of the reaction using a water to donate a proton to the pyruvate enolate (80). This water is charge-stabilized by Thr-145. The proposed reaction mechanism leaves a hydroxide stabilized by this threonine. Previous studies also suggest that the cleavage is activated 10-fold in the presence of inorganic phosphate (90). Current and future work is ongoing with deciphering the mechanism of action for LigK aldolase.

Scheme 4. The PCA 4,5-Cleavage Pathway of *Sphingobium* sp. SYK-6.



1.2.6. LigU and LigJ Mechanistic Elucidation

In this dissertation, the function of LigU was discovered to be a isomerization reaction that catalyzed the preferred substrate for LigJ hydratase. The next chapter will show in detail the supporting evidence. This evidence then allowed the elucidation of the mechanism of action of LigJ hydratase by site-directed mutagenesis and X-ray crystallography. The next chapter will describe structural and mechanistic details of LigU isomerase. Future work will be on the elucidation of the LigK aldolase mechanism.

1.3. References

1. Baucher, M., Monties, B., Montagu, M. V., and Boerjan, W. (1998) Biosynthesis and Genetic Engineering of Lignin. *Critical Reviews in Plant Sciences 17,* 125-197.
2. Weng, J. K. and Chapple, C. (2010) The Origin and Evolution of Lignin Biosynthesis. *New Phytologist 187,* 273-285.
3. Gibling, M. R. and Davies, N. S. (2012) Palaeozoic landscapes shaped by plant evolution. *Nature Geoscience 5,* 99-105.
4. Wong, D. W. S. (2009) Structure and Action Mechanism of Ligninolytic Enzymes. *Appl. Biochem. Biotechnol. 157,* 174-209.
5. Kamimura, N., Takahashi, K., Mori, K., Araki, T., Fujita, M., Higuchi, Y., and Masai, E. (2017) Bacterial catabolism of lignin-derived aromatics: New findings in a recent decade. *Environmental Microbiology Reports 9,* 679-705.
6. Santos, R. B., Hart, P. W., Jameel, H., and Chang, H. (2013) Wood Based Lignin Reactions Important to the Biorefinery and Pulp and Paper Industries. *BioResources 8,* 1456-1477.
7. Chio, C., Sain, M., and Qin, W. (2019) Lignin utilization: A review of lignin depolymerization from various aspects. *Renewable and Sustainable Energy Reviews 107,* 232-249.
8. Tarasov, D., Leitch, M., and Fatehi, P. (2018) Lignin-carbohydrate complexes: properties, applications, analyses, and methods of extraction: a review. *Biotechnol Biofuels 11*, 1-28.

9. Wang, H., Pu, Y., Ragauskas, A., and Yang, B. (2019) From lignin to valuable products-strategies, challenges, and prospects. *Bioresource Tech 271,* 449-461.
10. Schutyser, W., Renders, T., Bosch, S. V., Koelewijn, S. F., Beckham, G. T., and Sels, B. F. (2018) Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chem Soc Rev. 47,* 852-908.
11. Si, Z., Zhang, X., Wang, C., Ma, L., and Dong, R. (2017) An Overview on Catalytic Hydrodeoxygenation of Pyrolysis Oil and Its Model Compounds. *Catalysts 7,* 1-22.
12. Becker, J. and Wittmann, C. (2019) A field of dreams: Lignin valorization into chemicals, materials, fuels, and health-care products. *Biotechnology Advances 37,* 107360.
13. Abdelaziz, O. Y., Brink, D. P., Prothmann, J., Ravi, K., Sun, M., Garcia-Hidalgo, J., Sandahl, M., Hulteberg, C. P., Turner, C., Liden, G., and Gorwa-Grauslund, G. (2016) Biological valorization of low molecular weight lignin. *Biotech. Ad. 34,* 1318-1346.
14. Wu, W., Dutta, T., Varman, A. M., Eudes, A., Manalansan, B., Loque, D., and Singh, S. (2017) Lignin Valorization: Two Hybrid Biochemical Routes for the Conversion of Polymeric Lignin into Value-added Chemicals. *Nature Scientific Report 7,* 1-13.
15. Rorrer, N. A., Dorgan, J. R., Vardon, D. R., Martinez, C. R., Yang, Y., and Beckham, G. T. (2016) Renewable Unsaturated Polyesters from Muconic Acid. *ACS Sustainable Chem. Eng. 4,* 6867-6876.

16. Kamimura, N., Sakamoto, S., Mitsuda, N., Masai, E., and Kajita, S. (2019) Advances in microbial lignin degradation and its applications. *Current Opinion in Biotech 56,* 179-186.

17. Russell, M. B., Woodall, C. W., Fraver, S., D'Amato, A. W., Domke, G. M., and Skog, K. E. (2014) Residence Times and Decay Rates of Downed Woody Debris Biomass/Carbon in Eastern US Forests. *Ecosystems 17,* 765-777.

18. Cragg, S. M., Beckham, G. T., Bruce, N. C., Bugg, T. D. H., Distel, D. L., Dupree, P., Etxabe, A. G., Goodell, B. S., Jellison, J., McGeehan, J. E., McQueen-Mason, S. J., Schnorr, K., Walton, P. H., Watts, J. E. M., and Zimmer, M. (2015) Lignocellulose degradation mechanisms across the Tree of Life. *Current Opinion of Chemical Biology 29,* 108-119.

19. Li, H., Yelle, D. J., Li, C., Yang, M., Ke, J., Zhang, R., Liu, Y., Zhu, N., Liang, S., Mo, X., Ralph, J., Currie, C. R., and Mo, J. (2017) Lignocellulose pretreatment in a fungus-cultivating termite. *PNAS 114,* 4709-4714.

20. Bouchon, D., Zimmer, M., and Dittmer, J. (2016) The Terrestrial Isopod Microbiome: An All-in-One Toolbox for Animal-Microbe Interactions of Ecological Relevance. *Frontiers in Microbiology 7,* 1472.

21. Bredon, M., Dittmer J., Noel, C., Moumen, B., and Bouchon, D. (2018) Lignicellulose degradation at the holobiont level: teamwork in a keystone soil invertebrate. *Microbiome 6,* 1-19.

22. Jansson, J. K. and Hofmockel, K. S. (2018) The soil microbiome-from metagenomics to metaphenomics. *Current Opinion in Microbiology 43,* 162-168.

23. Andlar, M., Rezic, T., Mardetko, N., Kracher, D., Ludwig, R., and Santek, B. (2018) Lignocellulose degradation: An overview of fungi and fungal enzymes involved in lignocellulose degradation. *Eng. Life Sci. 18,* 768-778.
24. Brown, M. E. and Chang, M. C. Y. (2014) Exploring bacterial lignin degradation. *Current Opinion in Chemical Biology 19,* 1-7.
25. Hofrichter, M., Ullrich, R., Pecyna, M. J., Liers, C., and Lundell, T. (2010) New and classic families of secreted fungal heme peroxidases. *Appl Microbiol Biotechnol 87,* 871-897.
26. Choinowski, T., Blodig, W., Winterhalter, K. H., and Piontek, K. (1999) The Crystal Structure of Lignin Peroxidase at 1.70 Å Resolution Reveals a Hydroxy Group on the C^{β} of Tryptophan 171: A Novel Radical Site Formed During the Redox Cycle. *J. Mol. Biol. 286,* 809-827.
27. Sugano, Y. (2009) DyP-type peroxidases comprise a novel heme peroxidase family. *Cell. Mol. Life Sci. 66,* 1387-1403.
28. Pollegioni, L., Tonin, F., and Rosini, E. (2015) Lignin-degrading enzymes. *FEBS 282,* 1190-1213.
29. Daou, M. and Faulds, C. B. (2017) Glyoxal oxidases: their nature and properties. *World J. Microbiol. Biotechnol. 33,* 1-11.
30. Janusz, G., Pawlik, A., Sulej, J., Swiderska-Burek, U., Jarosz-Wilkolazka, A., and Paszcynski, A. (2017) Lignin degradation: microorganisms, enzymes involved, genome analysis and evolution. *FEMS Microbiology Reviews 41,* 941-962.

31. Abdel-Hamid, A. M., Solbiati, J. O., Cann, I. K. O. (2013) Insights into Lignin Degradation and its Potential Industrial Applications. *Adv. Applied Microbio. 82,* 1-28.

32. Edwards, S. L., Raag, R., Wariishi, H., Gold, M. H., and Poulos, T. L. (1993) Crystal structure of lignin peroxidase. *PNAS 90,* 750-754.

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

34. Hammel, K. E., and Cullen, D. (2008) Role of fungal peroxidases in biological ligninolysis. *Current Opinion in Plant Bio. 11,* 349-355.

35. Poulos, T. L., Edwards, S. L., Wariishi, H., and Gold, M. H. (1993) Crystallographic Refinement of Lignin Peroxidase at 2 Å*. *Journal of Bio. Chem. 268,* 4429-4440.

36. Millis, C. D., Cai, D., Stankovich, M. T., and Tien, M. (1989) Oxidation-Reduction Potentials and Ionization States of Extracellular Peroxidases from the Lignin-Degrading Fungus *Phanerochaete chrysosporium*. *Biochemistry 28,* 8484-8489.

37. Doyle, W. A., Blodig, W., Veitch, N. C., Piontek, K., and Smith, A. T. (1998) Two Substrate Interaction Sites in Lignin Peroxidase Revealed by Site-Directed Mutagenesis. *Biochemistry 37,* 15097-15105.

38. Sundaramoorthy, M., Youngs, H. L., Gold, M. H., and Poulos, T. L. (2005) High-Resolution Crystal Structure of Manganese Peroxidase: Substrate and Inhibitor Complexes. *Biochemistry 44,* 6463-6470.

39. Ravichandran, A., and Sridhar, M. (2017) Insights into the mechanism of lignocellulose degradation by versatile peroxidases. *Current Science 113,* 35-42.

40. Perez-Boada, M., Ruiz-Duenas, F. J., Pogni, R., Basosi, R., Choinowski, T., Martinez, M. J., Piontek, K., and Martinez, A. T. (2005) 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. 354,* 385-402.

41. Ruiz-Duenas, F. J., Morales, M., Perez-Boada, M., Choinowski, T., Martinez, M. J., Piontek, K., and Martinez, A. T. (2007) Manganese Oxidation Site in *Pleurotus eryngii* Versatile Peroxidase: A Site-Directed Mutagenesis, Kinetic, and Crystallographic Study. *Biochemistry 46,* 66-77.

42. Sugano, Y., Sasaki, K., and Shoda, M. (1999) cDNA Cloning and Genetic Analysis of a Novel Decolorizing Enzyme, Peroxidase Gene *dyp* from *Geotrichum candidum* Dec 1. *J. Biosci. & Bioeng. 87,* 411-417.

43. Lauber, C., Schwarz, T., Nguyen, Q. K., Lorenz, P., Lochnit, G., and Zorn, H. (2017) Identification, heterologous expression and characterization of a dye-decolorizing peroxidase of *Pleurotus sapidus*. *AMB Expr. 7,* 1-15.

44. Gonzalo, G., Colpa, D. I., Habib, M. H. M., and Fraaije, M. W. (2016) Bacterial enzymes involved in lignin degradation. *J. Biotech. 236,* 110-119.

45. Hakulinen, N., and Rouvinen, J. (2015) Three-dimensional structures of laccases. *Cell. Mol. Life Sci. 72,* 857-868.

46. Datta, R., Kelkar, A., Baraniya, D., Molaei, A., Moulick, A., Meena, R. S., and Formanek, P. (2017) Enzymatic Degradation of Lignin in Soil: A Review. *Sustainability 9,* 1-18.

47. Faure, D., Bouillant, M. L., and Bally, R. (1995) Comparative Study of Substrates and Inhibitors of *Azospirillum lipoferum* and *Pyricularia oryzae* Laccases. *Applied and Environmental Microbio. 61,* 1144-1146.

48. Enguita, F. J., Martins, L. O., Henriques, A. O., and Carrondo, M. A. (2003) Crystal Structure of a Bacterial Endospore Coat Component. *J. Bio. Chem. 278,* 19416-19425.

49. Giardina, P., Faraco, V., Pezzella, C., Piscitelli, A., Vanhulle, S., and Sannia, G. (2010) Laccases: a never-ending story. *Cell. Mol. Life Sci. 67,* 369-385.

50. Matera, I., Gullotto, A., Tilli, S., Ferraroni, M., Scozzafava, A., and Briganti, F. (2008) Crystal structure of the blue mutlicopper oxidase from the white-rot fungus *Trametes trogii* complexed with *p*-toluate. *Inorganica Chimica Acta 361,* 4129-4137.

51. Bento, I., Martins, L. O., Lopes, G. G., Carrondo, M. A., and Lindley, P. F. (2005) Dioxygen reduction by multi-copper oxidases; a structural perspective. *Dalton Trans.* 3507-3513.

52. Agrawal, K., Chaturvedi, V., and Verma, P. (2018) Fungal laccase discovered but yet undiscovered. *Bioresour. Bioprocess. 5,* 1-12.

53. Tao, L., Stich, T. A., Liou, S., Soldatova, A. V., Delgadillo, D. A., Romano, C. A., Spiro, T. G., Goodin, D. B., Tebo, B. M., Casey, W. H., and Britt, R. D. (2017) *J. Am. Chem. Soc. 139,* 8868-8877.

54. Komori, H., Miyazaki, K., and Higuchi, Y. (2009) X-ray structure of a two-domain type laccase: A missing link in the evolution of multi-copper proteins. *FEBS Letters 583,* 1189-1195.

55. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Research 28,* 235-242.

56. Chandra, R., and Chowdhary, P. (2015) Properties of bacterial laccases and their application in bioremediation of industrial wastes. *Environ. Sci. 17,* 326-342.

57. Kunamneni, A., Camarero, S., Garcia-Burgos, C., Plou, F. J., Ballesteros, A., and Alcalde, M. (2008) Engineering and Applications of fungal laccases for organic synthesis. *Microbial Cell Factories 7,* 1-17.

58. Durao, P., Bento, I., Fernandes, A. T., Melo, E. P., Lindley, P. F., and Martins, L. O. (2006) Perturbations of the T1 copper site in the CotA laccase from *Bacillus subtilis*: structural, biochemical, enzymatic and stability studies. *J. Biol. Inorg. Chem. 11,* 514-526.

59. Jones, S. M., and Solomon, E. I. (2015) Electron transfer and reaction mechanism of laccases. *Cell. Mol. Life Sci. 72,* 869-883.

60. Christopher, L. P., Yao, B., and Ji, Y. (2014) Lignin biodegradation with laccase-mediator systems. *Frontiers in Energy Research 2,* 1-13.

61. Salvachua, D., Katahira, R., Cleveland, N. S., Khanna, P., Resch, M. G., Black, B. A., Purvine, S. O., Zink, E. M., Prieto, A., Martinez, M. J., Martinez, A. T., Simmons, B. A., Gladden, J. M., and Beckham, G. T. (2016) Lignin depolymerization by fungal secretomes and a microbial sink. *Green Chem. 18,* 6046-6062.

62. Masai, E., Katayama, Y., and Fukuda, M. (2007) Genetic and Biochemical Investigations on Bacterial Catabolic Pathways for Lignin-Derived Aromatic Compounds. *Biosci. Biotechnol. Biochem. 71,* 1-15.

63. Sonoki, T., Obi, T., Kubota, S., Higashi, M., Masai, E., and Katayama, Y. (2000) Coexistence of Two Different O Demethylation Systems in Lignin Metabolism by *Sphingomonas paucimobilis* SYK-6: Cloning and Sequencing of the Lignin Biphenyl-Specific *O*-Demthylas (LigX) Gene. *Appl. Environ. Microbiol. 66,* 2125-2132.

64. Yoshikata, T., Suzuki, K., Kamimura, N., Namiki, M., Hishiyama, S., Araki, T., Kasai, D., Otsuka, Y., Nakamura, M., Fukuda, M., Katayama, Y., and Masai, E. (2014) Three-Component *O*-Demethylase System Essential for Catabolism of a Lignin-Dervied Biphenyl Compound in *Sphingobium* sp. Strain SYK-6. *Appl. Environ. Microbiol. 80,* 7142-7153.

65. Peng, X., Egashira, T., Hanashiro, K., Masai, E., Nishikawa, S., Katayama, Y., Kimbara, K., and Fukuda, M. (1998) Cloning of *Sphingomonas paucimobilis* SYK-6 Gene Encoding a Novel Oxygenase That Cleaves Lignin-Related Biphenyl and Characterization of the Enzyme. *Appl. Environ. Microbiol. 64,* 2520-2527.

66. Kuatsjah, E., Chen, H., Wither, S. G., and Eltis, L. D. (2017) Characterization of an extradiol dioxygenase involved in the catabolism of lignin-derived biphenyl. *FEBS Letters 591,* 1001-1009.

67. Peng, X., Masai, E., Katayama, Y., and Fukuda, M. (1999) Characterization of the *meta*-Cleavage Compound Hydrolase Gene Involved in Degradation of the Lignin-Related Biphenyl Structure by *Sphingomonas paucimobilis* SYK-6. *Appl. Environ. Microbiol. 65,* 2789-2793.

68. Kuatsjah, E., Chan, A. C. K., Kobylarz, M. J., Murphy, M. E. P., and Eltis, L. D. (2017) The bacterial *meta*-cleavage hydrolase LigY belongs to the amidohydrolase superfamily, not to the α/β-hydrolase superfamily. *J. Biol. Chem. 44,* 18290-18302.

69. Vladimirova, A., Patskovsky, Y., Fedorov, A. A., Bonanno, J. B., Fedorov, E. V., Toro, R., Hillerich, B., Seidel, R. D., Richards, N. G. J., Almo, S. C., and Raushel, F. M. (2016) Substrate Distortion and the Catalytic Reaction Mechanism of 5-Carboxyvanillate Decarboxylase. *J. Am. Chem. Soc. 138,* 826-836.

70. Abe, T., Masai, E., Miyauchi, K., Katayama, Y., and Fukuda, M. (2005) A Tetrahydrofolate-Dependent *O*-Demethylase, LigM, Is Crucial for Catabolism of Vanillate and Syringate in *Sphingomonas paucimobilisi* SYK-6. *J. Bacteriology 187,* 2030-2037.

71. Kohler, A. C., Mills, M. J. L., Adams, P. D., Simmons, B. A., and Sale, K. L. (2017) Structure of aryl *O*-demethylase offers molecular insight into a catalytic tyrosine-dependent mechanism. *PNAS 16,* E3205-E3214.

72. Noda, Y., Nishikawa, S., Shiozuka, K. I., Kadokura, H., Nakajima, H., Yoda, K., Katayama, Y., Morohoshi, N., Haraguchi, T., and Yamasaki, M. (1990) Molecular Cloning of the Protocatechuate 4,5- Dioxygenase Genes of *Pseudomonas paucimobilis*. *J. Bacteriol. 172,* 2704−2709.

73. Masai, E., Momose, K., Hara, H., Nishikawa, S., Katayama, Y., and Fukuda, M. (2000) Genetic and Biochemical Characterization of 4-Carboxy-2-Hydroxymuconate-6-Semialdehyde Dehydrogenase and Its Role in the Protocatechuate 4,5-Cleavage Pathway in *Sphingomonas paucimobilis* SYK-6. *J. Bacteriol. 182,* 6651−6658.

74. Hobbs, M. E., Malashkevich, V., Williams, H. J., Xu, C., Sauder, J. M., Burley, S. K., Almo, S. C., and Raushel, F. M. (2012) Structure and Catalytic Mechanism of LigI: Insight into the Amidohydrolase Enzymes of cog3618 and Lignin Degradation. *Biochemistry 51,* 3497− 3507.

75. Hogancamp, T. N., and Raushel, F. M. (2018) Functional Annotation of LigU as a 1,3-Allylic Isomerase during the Degradation of Lignin in the Protocatechuate 4,5-Cleavage Pathway from the Soil Bacterium *Sphingobium* sp. SYK-6. *Biochemistry 57,* 2837−2845.

76. Hogancamp, T. N., Cory, S. A., Barondeau, D. P., Raushel, F. M. (2019) Structure and Chemical Reaction Mechanism of LigU, an Enzyme That Catalyzes an Allylic Isomerization in the Bacterial Degradation of Lignin. *Biochemistry 58,* 3494-3503.

77. Hogancamp, T. N., Mabanglo, M. F., and Raushel, F. M. (2018) Structure and Reaction Mechanism of the LigJ Hydratase: An Enzyme Critical for the Bacterial

Degradation of Lignin in the Protocatechuate 4,5-Cleavage Pathway. *Biochemistry 57,* 5841−5850.

78. Maruyama, K. (1990) Purification and Properties of 4- Hydroxy-4-Methyl-2-Oxoglutarate Aldolase from *Pseudomonas ochraceae* Grown on Phthalate. *J. Biochem. 108,* 327−333.
79. Hara, H., Masai, E., Miyauchi, K., Katayama, Y., and Fukuda, M. (2003) Characterization of the 4-Carboxy-4-Hydroxy-2-Oxoadipate Aldolase Gene and Operon Structure of the Protocatechuate 4,5- Cleavage Pathway Genes in *Sphingomonas paucimobilis* SYK-6. *J. Bacteriol.* 185, 41−50.
80. Wang, W., Mazurkewich, S., Kimber, M. S., and Seah, S. Y. K. (2010) Structural and Kinetic Characterization of 4-Hydroxy-4- methyl-2-oxoglutarate/4-Carboxy-4-hydroxy-2-oxoadipate Aldolase, a Protocatechuate Degradation Enzyme Evolutionarily Convergent with the HpaI and DmpG Pyruvate Aldolases. *J. Biol. Chem. 285,* 36608− 36615.
81. Koehntop, K. D., and Emerson, J. P. (2005) The 2-His-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron(II) enzymes. *J. Biol. Inorg. Chem. 10,* 87-93.
82. Barry, K. P., and Taylor, E. A. (2013) Characterizing the Promiscuity of LigAB, a Lignin Catabolite Degrading Extradiol Dioxygenase from *Sphingomonad paucimobilis* SYK-6. *Biochemistry 52,* 6724-6736.

83. Sugimoto, K., Senda, T., Aoshima, H., Masai, E., Fukuda, M., and Mitsui, Y. (1999) Crystal structure of an aromatic ring opening dioxygenase LigAB, a protocatechuate 4,5-dioxygenase, under aerobic conditions. *Structure 7,* 953-965.

84. Vaillancourt, F. H., Bolin, J. T., and Eltis, L. D. (2006) The Ins and Outs of Ring-Cleaving Dioxygenases. *Critical Reviews in Biochem. and Mol. Biol. 41,* 241-267.

85. Mendel, S., Arndt, A., and Bugg, T. D. H. (2004) Acid-Base Catalysis in the Extradiol Catechol Dioxygenase Reaction Mechanism: Site-Directed Mutagenesis of His-115 and His-179 in *Escherichia coli* 2,3-Dihydroxyphenylpropionate 1,2-Dioxygenase (MhpB). *Biochemistry 43,* 13390-13396.

86. Thoden, J. B., and Holdn, H. M. (2011) Biochemical and Structural Characterization of WlbA from *Bordetella pertussis* and *Chromobacterium violaceum*: Enzymes Required for the Biosynthesis of 2,3-Diacetamido-2,3-dideoxy-D-mannuronic Acid. *Biochemistry 50,* 1483-1491.

87. Kang, D., Tanaka, K., Takenaka, S., Ishikawa, S., Yoshida, K. (2017) *Bacillus subtilis iolU* encodes an additional $NADP^+$-dependent *scyllo*-inositol dehydrogenase. *Biosci., Biotech., and Biochem. 81,* 1026-1032.

88. Nogales, J., Canales, Á., Jimeńez-Barbero, J., Serra, B., Pingarroń, J. M., García, J. L., and Díaz, E. (2011) Unravelling the Gallic acid Degradation Pathway in Bacteria: the gal Cluster from *Pseudomonas putida*. *Mol. Microbiol. 79,* 359−374.

89. Mazurkewich, S., Brott, A. S., Kimber, M. S., and Seah, S. Y. K. (2016) Structural and Kinetic Characterization of the 4-Carboxy-2- hydroxymuconate Hydratase from

the Gallate and Protocatechuate 4,5-Cleavage Pathways of *Pseudomonas putida* KT2440. *J. Biol. Chem. 291,* 7669−7686.

90. Mazurkewich, S., and Seah, S. Y. K. (2016) Investigation into the Mode of Phosphate Activation in the 4-Hydroxy-4-Methyl-2-Oxoglutarate/4-Carboxy-4-Hydroxy-2-Oxoadipate Aldolase from *Pseudomonas putida* F1. *PLOS One 11,* 1-15.

2. FUNCTIONAL ANNOTATION OF LIGU AS A 1,3-ALLYLIC ISOMERASE DURING THE DEGRADATION OF LIGNIN IN THE PROTOCATECHUATE 4,5-CLEAVAGE PATHWAY FROM THE SOIL BACTERIUM *SPHINGOBIUM* SP. SYK-6[*]

2.1. Introduction

Lignin, the woody part of plants, provides structural integrity and protection and is the second most abundant renewable carbon source that is readily available (1). Chemically, it is a complex aromatic heteropolymer connected by ether and biphenyl linkages. The biodegradation of lignin is initiated by basidiomycete fungi that secrete oxidoreductases that are used for the enzymatic cleavage of the connecting linkages (2). Soil bacteria subsequently metabolize these larger polymeric units into smaller fragments. One of these smaller fragments, vanillate, is demethylated to protocatechuate (PCA) and can be further metabolized via three catabolic pathways. In the 4,5-cleavage pathway, pyruvate and oxaloacetate are the ultimate products formed by the combined actions of the enzymes LigAB, LigC, LigI, LigU, LigJ, and LigK (3).

In the protocatechuate 4,5-cleavage pathway, we have previously determined the three-dimensional crystal structure of LigI [Protein Data Bank (PDB) entry 4D8L] from Sphingobium sp. SYK-6 and have demonstrated that this enzyme catalyzes the

---

[*] Reprinted with permission from "Functional Annotation of LigU as a 1,3-Allylic Isomerase during the Degradation of Lignin in the Protocatechuate 4,5-Cleavage Pathway from the Soil Bacterium *Sphingobium* sp. SYK-6" by Tessily N. Hogancamp and Frank M. Raushel, *Biochemistry,* 2018, *57* (19), pp 2837-2845, Copyright 2018 American Chemical Society

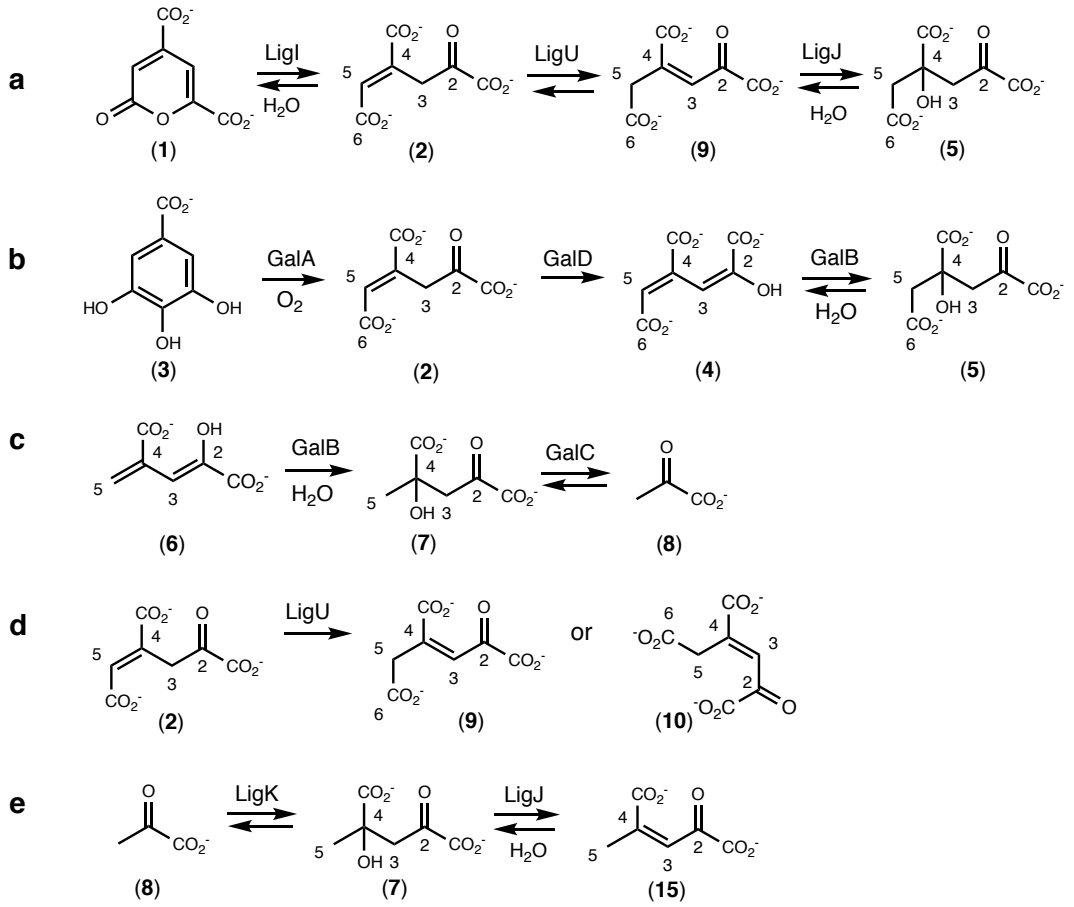
hydrolysis of 2-pyrone-4,6-dicarboxylate (PDC, 1) to oxalomesaconate (OMA, 2) as shown in **Scheme 5a** (4). This metabolite is also formed from the oxidation of gallic acid (3) by GalA in the gallate pathway from *Pseudomonas putida* KT2440 (5) In the gallate pathway, it has been proposed that OMA is subsequently tautomerized to (2*E*,4*E*)-4-carboxy- 2-hydroxymuconate (CHM, 4) by the action of GalD, before being hydrated to 4-carboxy-4-hydroxy-2-oxoadipate (CHA, 5) by GalB (**Scheme 5b**) (6,7) The structure of GalB from *P. putida* has been determined (PDB entry 5CGZ), and a chemical mechanism has been proposed for the hydration of (2*E*,4*E*)-CHM (4) to CHA (5) (7). In the PCA 4,5-cleavage pathway from *Sphingobium* sp. SYK-6, the formation of CHA (5) is catalyzed by LigJ, and the enzyme homologous to GalD is LigU with 61% sequence identity (6).

A similar set of transformations has been demonstrated in the biphenyl degradation pathway with the initial formation of 4-carboxy-2-hydroxypenta-2,4-dienoate (CHPD, 6) (7). In the proposed pathway, CHPD (6) is reversibly hydrated by GalB to 4-hydroxy-4-methyl-2-oxoglutarate (HMG, 7) before a retro-aldol cleavage catalyzed by GalC that forms two molecules of pyruvate (8) as illustrated in **Scheme 5c** (7). However, unlike the enol tautomer of CHM (4) or the corresponding keto tautomer of OMA (2), the double bond between C4 and C5 of CHPD (6) is not sufficiently activated for the addition of hydroxide to C4 and a proton to C5 (or the reverse dehydration reaction). Chemically, it is far more likely that the dehydration of HMG (7) by GalB or LigJ will occur via the initial abstraction of a proton at C3 with formation of the double bond between C3 and C4 rather than between C4 and C5. If this were the

case, then it is also be highly likely that the two putative isomerase enzymes (GalD and LigU) would actually catalyze a 1,3-allylic isomerization of the double bond between C4 and C5 in OMA (2) to a product with a double bond between C3 and C4 and thus form either stereoisomer of 2-keto-4-carboxy-3-hexene-dioate (KCH, 9 or 10) as illustrated in **Scheme 5d**.

In this paper, we have demonstrated via nuclear magnetic resonance (NMR) spectroscopy that LigU from *Sphingobium* sp. SYK-6 catalyzes the isomerization of (4*E*)-OMA (2) to (3*Z*)-KCH (9). We have also shown that the preferred substrate for the LigJ hydratase is (3*Z*)-KCH (9) rather than (4*E*)-OMA (2) or (3*Z*,4*E*)-CHM (4), as previously proposed (6,7).

Scheme 5. (a) Proposed Degradation Pathway for 2-Pyrone-4,6-dicarboxylate (1), (b) Proposed Degradation Pathway for Gallate (3), (c) Proposed Reactions Catalyzed by GalB and GalC, (d) Potential Reaction Products for the Isomerization of OMA (2) Catalyzed by LigU, and (e) Reactions Catalyzed by LigK and LigJ



2.2. Material and Methods

2.2.1. Materials

All chemicals and buffers were purchased from Sigma-Aldrich unless otherwise specified. The synthesis of 2- pyrone-4,6-dicarboxylate (PDC, 1) was conducted according to published procedures (8).

2.2.2. Cloning, Expression, and Purification of LigI, LigU, LigJ, and LigK from *Sphingobium* sp. SYK-6

The gene for LigI from *Sphingobium* sp. SYK-6 was obtained as previously described (4). The genes for *Sphingobium* sp. SYK-6 LigU (UniProt entry Q0KJL4), LigJ (UniProt entry Q9KWl6), and LigK (UniProt entry Q8RQW0) were synthesized via codon optimization for expression in *Escherichia coli* by GenScript. The genes were subsequently subcloned from a pUC57 vector into a pET-30a (+) (EMD Millipore) expression vector using NdeI as the forward restriction enzyme and XhoI as the reverse restriction enzyme. The forward primer for the subcloning of LigU was ATATATTCATATGCCGCGTCGTGATCGTAATATGGACAGCGCG, and the reverse primer was ACATCTCGAGCACGAACACAACGCCATCCATCAGTTTACGCG (Integrated DNA Technologies, IDT). The forward primer for LigJ was ATATATTCATATGATGATGATTATTGATTGTCACGGTCACTACAC, and the reverse primer was ACATCTCGAGCAGACCACGGGCTTTCAGTTTCGC (IDT).

The forward primer for the subcloning of LigK was ATATATTCATATGATGCGTGGTGCCGCTATGGGTGTGG, and the reverse primer was ACATCTCGAGCACATATTTCAGACCTTCTTTTTCCAGACGTTCACG (IDT). All of the modified genes included a C-terminal linker and a His tag as Leu-Glu-His_6.

The plasmids were transformed into BL21(DE3) competent cells (EMD Millipore). Single colonies were used to inoculate 6 mL cultures of LB, which were grown for 12 h at 37 °C. The 6 mL culture was used to inoculate 1 L of LB containing 1.0 mM kanamycin at 37 °C. At an OD_{600} of ~0.6, 1.0 mM isopropyl β- D-thiogalactopyranoside (IPTG) was added and the cells were grown overnight at 22 °C. The cells were harvested by centrifugation for 10 min at 10000 rpm. Cell pellets were resuspended in ~100 mL of binding buffer [20 mM HEPES/ KOH or 20 mM phosphate/KOH (pH 7.9), 0.50 M NaCl, and 10 mM imidazole] and lysed by sonication using a Branson Sonifier 450 with six complete cycles of on and off for 5 min of each cycle. The cells were cooled externally with ice. The cellular debris was removed by centrifugation at 13000 rpm for 30 min. The supernatant solution was collected and passed through a 0.45 μm syringe filter (VWR). The filtrate was loaded onto a 5 mL HisTrap HP column (GE Healthcare) attached to an NGC liquid chromatography system (Bio-Rad Laboratories, Inc.) equilibrated with binding buffer. The His-tagged proteins were eluted with a 0 to 40% gradient of elution buffer [20 mM HEPES/KOH or phosphate/KOH (pH 7.9), 0.25 M NaCl, and 0.50 M imidazole]. The fractions containing the purified protein were dialyzed overnight in 50 mM phosphate/KOH buffer (pH 7.9). The protein was dispensed into Eppendorf tubes and flash-frozen using

liquid nitrogen before being stored at −80 °C. The final yields were ∼50 mg of LigU, ∼30 mg of LigJ, and ∼15 mg of LigK per liter of growth medium. Protein purity was determined using sodium dodecyl sulfate−polyacrylamide gel electrophoresis.

2.2.3. Measurement of the Reaction Catalyzed by LigI

The reaction catalyzed by LigI was initially measured using anion-exchange chromatography by monitoring the absorbance at 255 nm. Each reaction mixture contained 1.0 mM PDC (1), 100 nM LigI, and 125 mM triethanolamine (TEA)/KOH (pH 8.25) in a volume of 1−3 mL. After an initial incubation period of ∼5 min, 100 μL samples were loaded onto the column and eluted with a gradient from 0 to 2 M KCl. For certain experiments, LigI was removed from the reaction mixture via filtration through a PALL Nanosep 10 kDa filter. All chromatography experiments used a 5 mL HiTrap Q HP anion-exchange chromatography column. The gradient flow rate was 4 mL/min with an equilibration, a column wash, an elution, and a second column wash phase programmed into the method.

2.2.4. Structural Analysis of the LigI-Catalyzed Hydrolysis Product

The product of the LigI-catalyzed reaction was characterized using NMR spectroscopy with a Bruker Avance III 500 MHz spectrometer equipped with an H−C−N cryoprobe at 305 K. The LigI-catalyzed reaction product was characterized by ^{1}H NMR spectroscopy in a solution that initially contained 100 mM $NaHCO_3$/HCl (pH 9.5), 0.5 μM LigI, 10% D_2O, and 10 mM PDC. The reaction mixture was incubated for 7 min

before acquisition of the NMR signal. After the data were collected, the pH of the reaction mixture was 9.22. The acquisition parameters used a Watergate W5 zggpw5 pulse sequence with gradients using double echo to suppress the water resonance calibrated at 4.653 ppm (9). The deuterium-exchange reaction was characterized by ^{1}H NMR spectroscopy by incubating 5.0 mM PDC with 250 nM LigI and 50 mM sodium bicarbonate/DCl (pD 9.8) in ~100% D_2O. Watergate suppression using a zggpw5 pulse sequence was used with a relaxation delay of 5 s. A control reaction was arranged similarly in 10% D_2O with the same collection parameters.

To stabilize the LigI reaction product, 3−5 equiv of either $NaBH_4$ or $NaBD_4$ was added to the reaction mixture to reduce the carbonyl group at C2 of OMA (2). After a 5 min incubation of LigI with PDC, 1 equiv of either $NaBH_4$ or $NaBD_4$ was added to the reaction mixture every 10 min until a total of 3−5 equiv of the reducing agent had been added. The samples were centrifuged for 10 min at 10000 rpm in an Eppendorf microcentrifuge and then lyophilized to dryness and resuspended in 600 μL of D_2O. After the NMR data were collected, the pH values of the reaction mixtures were pH 9.82 and 9.86 for the samples reduced with $NaBH_4$ and $NaBD_4$, respectively. The NMR acquisition used a 1D pulse with an f1 presaturation zgpr pulse sequence with the water resonance calibrated at 4.701 ppm for the sample reduced with $NaBH_4$ and 4.703 ppm for the sample reduced with $NaBD_4$.

The ^{13}C NMR experiments required a higher concentration of product to be formed compared to that for the ^{1}H NMR experiments; therefore, 40 mM PDC was used in a solution containing 100 mM phosphate/KOH (pH 8.0), 10 μM LigI, 10% D_2O, and a

total of 3 equiv of $NaBD_4$. Afterward, the pH measured 8.56. The ^{13}C NMR-decoupled experiment used the 1D zgpg pulse sequence with power-gated decoupling. The ^{13}C NMR coupled experiments used the 1D zggd30 pulse sequence with gated proton decoupling using a 30° flip angle.

2.2.5. Measurement of Nonenzymatic Isomerization of OMA

The nonenzymatic conversion of OMA (2) to KCH (9) was monitored at 255 nm by anion-exchange chromatography. The reaction mixture contained 1.0 mM PDC, 125 mM TEA/KOH (pH 8.2), and 100 nM LigI in a volume of 2.0 mL. The reaction mixture was incubated for 10 min before LigI was removed with two PALL Nanosep 10 kDa filters. A series of 100 µL aliquots were removed every 27 min, loaded onto the column, and eluted with a gradient from 0 to 2 M KCl. The nonenzymatic rate of isomerization at pH 8.0 was determined by fitting the area of the product to a single exponential.

2.2.6. Measurement of the LigU-Catalyzed Reaction

The catalytic activity of LigU was measured at pH 7.0 and 7.5 in 75 mM MOPS/KOH and at pH 8.0 in 75 mM TEA/KOH. The reaction mixture contained 1.0 mM PDC, 100 nM LigI, 75 mM buffer, and various concentrations of LigU (1.0 nM to 1.0 µM). The binding buffer was 20 mM MOPS/KOH (pH 7.0 and 7.5) or 20 mM TEA/KOH (pH 8.0). The elution buffer contained 2.0 M KCl. A 100 µL sample was injected onto the column at ~25 min intervals to observe the change in product formation at 255 nm using anion-exchange chromatography.

2.2.7. Structural Analysis of the LigU Isomerase Product

The product of the LigU-catalyzed reaction was characterized using ^{1}H and ^{13}C NMR spectroscopy with a Bruker Avance III 500 MHz spectrometer equipped with an H−C−N cryoprobe at 305 K. The reaction mixture contained 10 mM PDC in 100 mM sodium bicarbonate/HCl (pH 9.5) with 0.5 µM LigI and 0.5 µM LigU with 10% D_2O with an 8 min incubation period prior to the start of data collection. The acquisition parameter used a 1D zggpw5 Watergate W5 pulse sequence with gradients using double echo to suppress the water resonance calibrated to 4.653 ppm (9). The LigU product was reduced with $NaBH_4$ and $NaBD_4$ in the same manner that was used for the LigI product. The ^{1}H NMR spectrum was acquired using the 1D zg30 pulse sequence using a 30° flip angle for both reactions. The ^{13}C NMR experiments required a higher concentration of product to be formed compared to the ^{1}H NMR experiments; therefore, 40 mM PDC was used in a solution containing 100 mM phosphate/KOH (pH 8.0), 10 µM LigI, 10 µM LigU, 10% D_2O, and 3 equivalents of $NaBD_4$ added in 0.5 equiv increments over ~45 min. The ^{13}C NMR-decoupled experiment used the 1D zgpg sequence with power-gated decoupling. The ^{13}C NMR coupled experiments used the 1D zggd30 sequence with gated decoupling using a 30° flip angle.

2.2.8. Identification of the Preferred Substrate for LigJ

The reaction mixture contained 2.0 mM PDC, 100 nM LigI, and 100 nM LigU in 125 mM TEA/KOH (pH 8.0) in a volume of 500 µL and was incubated for 15 min. LigI and LigU were removed by filtration using two PALL Nanosep 10 kDa filter and one 3

kDa filter. LigJ (0.8 μM) was added to the reaction mixture, and a 100 μL aliquot was loaded onto an anion-exchange column and monitored at 255 nm for the decrease in absorbance for LigI or LigU products. The elution buffer was a gradient from 0 to 22% 10 mM TEA/KOH (pH 8.0) with 2 M KCl.

2.2.9. Measurement of the Catalytic Activity of LigU

The initial velocity of the reaction catalyzed by LigU was determined spectrophotometrically by monitoring the change in absorbance at 265 nm. The substrate was made *in situ* by the addition of 0.5 μM LigI to variable concentrations of PDC (from 0 to 0.6 mM) in 100 mM potassium phosphate/KOH (pH 8.0) in a volume of 1.0 mL. The reaction catalyzed by LigI was monitored at 312 nm until the reaction was complete. LigU (400 pM) was added and the reaction monitored at 265 nm. The change in the concentration of the substrate was determined using a differential extinction coefficient ($\Delta\varepsilon_{265}$) of 5980 $M^{-1}\,cm^{-1}$. The initial concentration of the OMA (2) was determined from the concentration of the remaining PDC (1) at 265 nm. The kinetic parameters were determined from a fit of the data to **eq 1**

$$v/E_t = k_{cat}[A]/(K_a + [A])$$

Equation 1

where v is the initial velocity, k_{cat} is the turnover number, [A] is the concentration, and K_a is the Michaelis constant.

2.2.10. Dehydration of HMG (7) by LigJ

The retro-aldol condensation by LigK converts 2 mol of pyruvate to HMG (7), and then LigJ dehydrates HMG. The reaction was monitored by ^{1}H NMR spectroscopy with a Bruker Avance III 500 MHz spectrometer equipped with an H−C−N cryoprobe at 305 K. The initial reaction mixture contained 20 mM pyruvate, 200 mM potassium phosphate/KOH (pH 8.0), 200 nM LigK, and 5 mM $MgCl_2$ and was incubated for 3 h. LigK was removed using two PALL Nanosep 10 kDa filters and one 3 kDa filter, and then the filtrate was lyophilized to dryness. One sample was resuspended in 200 mM potassium phosphate/KOD (pD 7.2) to a volume of 600 µL of D_2O. The second sample included 0.5 µM LigJ in the resuspension reaction mixture. Both samples were analyzed by NMR spectroscopy using the zggpw5 pulse program for water suppression with a relaxation delay of 5 s (9). The data were processed using a line broadening of 0.2 Hz.

2.2.11. Structure of the LigJ Substrate

Experimental support for the location of the double bond in the LigJ substrate was provided by monitoring the reaction catalyzed by LigJ in D_2O with CHA (5) using ^{1}H NMR spectroscopy. CHA was prepared by incubating 10 mM PDC with 100 nM LigI, 100 nM LigU, and 100 nM LigJ in 100 mM potassium phosphate/KOH (pH8.0) in 10% D_2O to a final volume of 600 µL. CHA was also prepared in D_2O by incubating 20 mM PDC with 100 nM LigI, 100 nM LigU, and 100 nM LigJ in 200 mM potassium phosphate/KOH (pH 8.0). The enzymes were removed using two PALL Nanosep 10 kDa filters and one 3 kDa filter. The filtrate was lyophilized to dryness and then

resuspended in 600 μL of D_2O containing 100 mM potassium phosphate/KOD (pD 7.2) with 0.5 μM LigJ. The reaction was measured by ^1H NMR spectroscopy.

2.3. Results

2.3.1. Enzyme-Catalyzed Hydrolysis of PDC by LigI

The hydrolysis of PDC (1) catalyzed by LigI from *Sphingobium* sp. SYK-6 forms one product enzymatically at pH 7.7, which was monitored by the decrease in the absorbance for PDC and the appearance of a new peak for OMA by anion-exchange chromatography at 255 nm (**Figure 11a**). PDC and the OMA (2) product elute as separate peaks with PDC having a retention time of 11.1 min and OMA eluting later at 14.4 min. The longer retention time for OMA is consistent with the increase in the number of carboxylate groups relative to that for PDC.

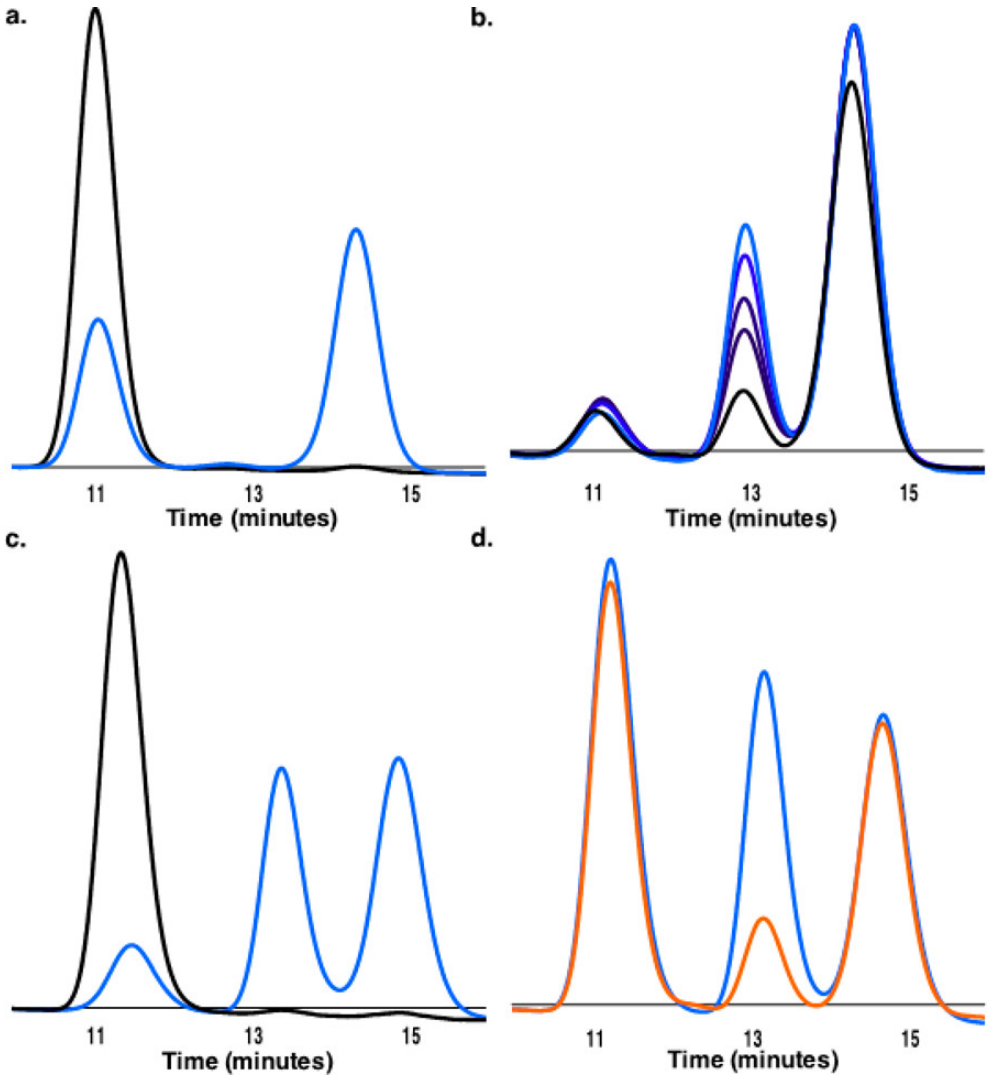


Figure 11. Anion-exchange separation of reaction products after hydrolysis of PDC (1). (a) PDC has a retention time of 11.1 min (black line). Hydrolysis of PDC with LigI illustrates the formation of the hydrolysis product (OMA, 2) at 14.4 min (blue line). (b) At pH 7.7, OMA nonenzymatically forms a new product as a function of time with a retention time of 13.0 min (black to light blue lines). (c) The reaction products formed immediately after PDC is incubated with 100 nM LigI and 100 nM LigU. PDC, KCH (9), and OMA appear at 11.1, 13.0, and 14.4 min, respectively (blue line), at pH 7.7. The reaction mixture of PDC with 100 nM LigU shows no product formation (black line). (d) Reaction mixture after PDC is incubated with LigI and LigU (blue line) at pH 7.7. Reaction mixture after PDC was incubated with LigI, LigU, and 0.8 μM LigJ (orange line) at pH 7.7.

2.3.2. NMR Analysis of the LigI Hydrolysis Product

NMR spectroscopy was used to confirm the structure of the OMA (2) product formed from the enzymatic hydrolysis of PDC (1) by LigI. The ^{1}H NMR spectrum of PDC exhibits two resonances at 7.24 and 6.75 ppm with a small ^{4}J coupling constant of 1.4 Hz (**Figure 12**). The resonance at 7.24 ppm originates from the hydrogen at C3, whereas the resonance at 6.75 ppm is from the hydrogen at C5. After LigI is added to a solution of PDC in H_2O at pH 10, the two resonances for the substrate disappear and are replaced by two new resonances at 6.69 and 4.02 ppm, with an intensity ratio of ~1:2 (**Figure 13a and 14a**). These two resonances originate from the hydrogens at C5 and C3, respectively. When the LigI-catalyzed hydrolysis of PDC is conducted in D_2O, the resonances for the two hydrogens at C3 disappear for OMA and the one hydrogen at C3 for PDC also disappears. (**Figure 15**).

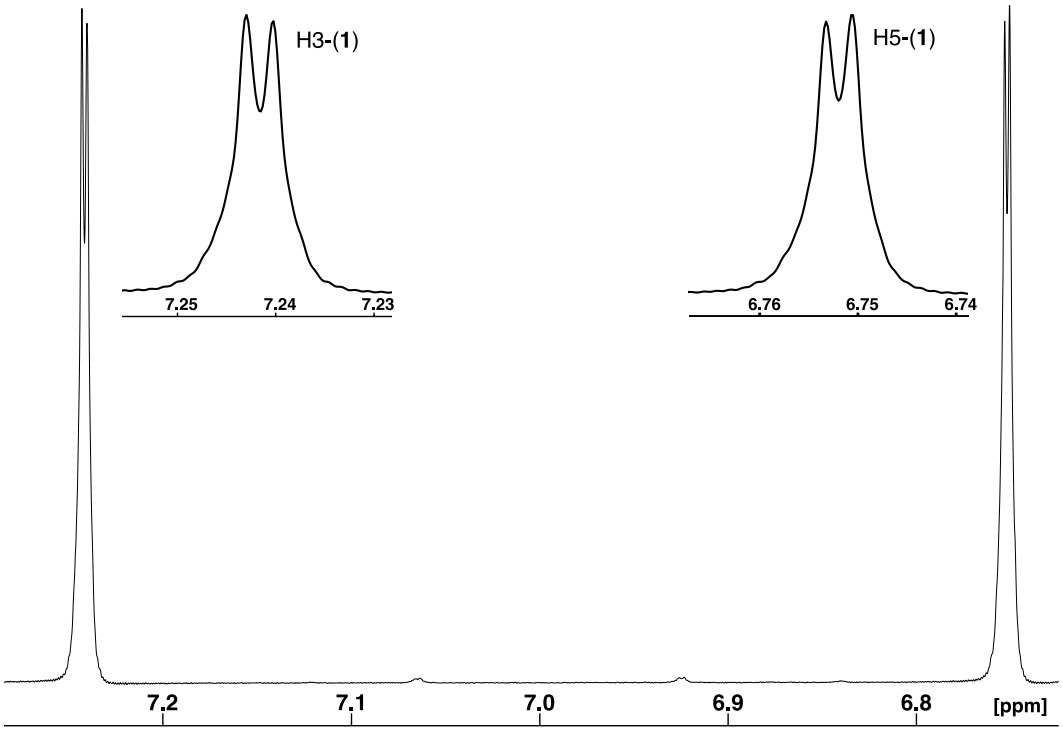


Figure 12. ^{1}H NMR spectrum of PDC (1) in 10% D2O has two resonances at 7.24 and 6.75 ppm with a small ^{4}J coupling constant of 1.4 Hz measured at pH 7.2. The resonance at 7.24 ppm originates from the hydrogen at C3, whereas the resonance at 6.75 ppm is from the hydrogen at C5. The data were collected using a WATERGATE zggpw5 pulse program with 128 scans with a relaxation delay of 5 seconds.

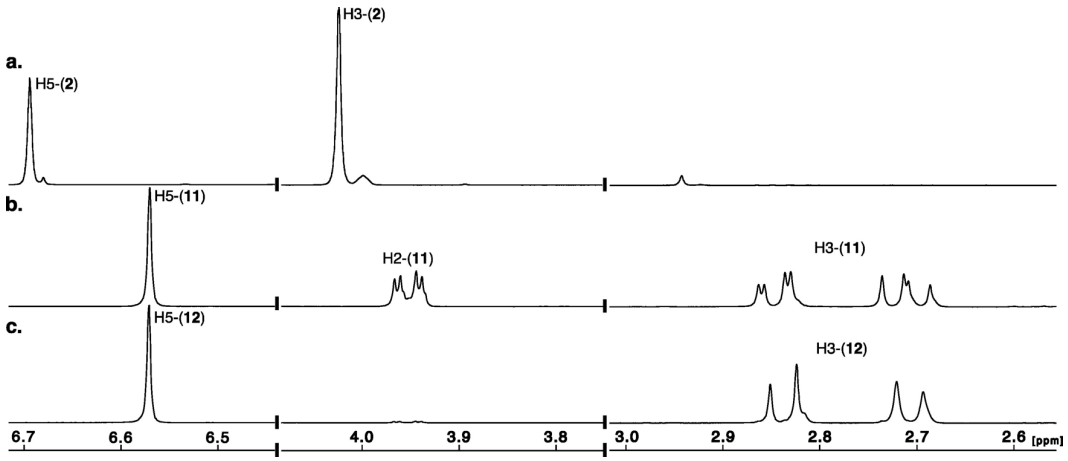


Figure 13. (a) ^{1}H NMR spectrum of OMA (2) formed enzymatically by the hydrolysis of PDC (1) with LigI. (b) NMR spectrum of the hydrolysis product after reduction with $NaBH_4$ (11). (c) NMR spectrum of the hydrolysis product after reduction with $NaBD_4$ (12).

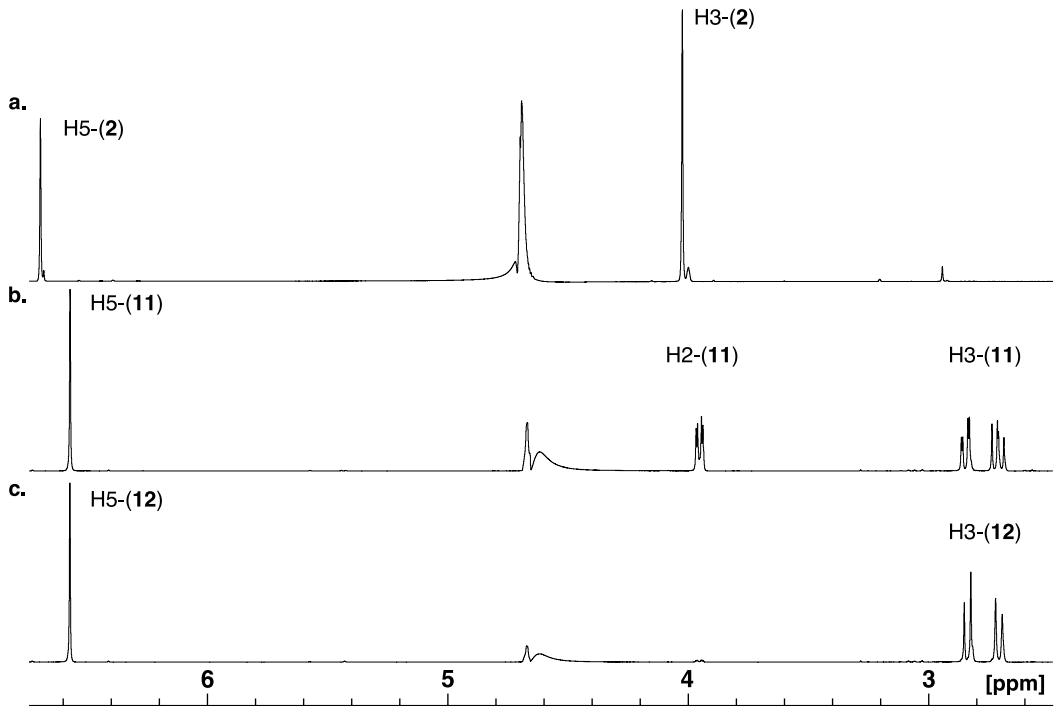


Figure 14. The full width spectra of Figure 13.

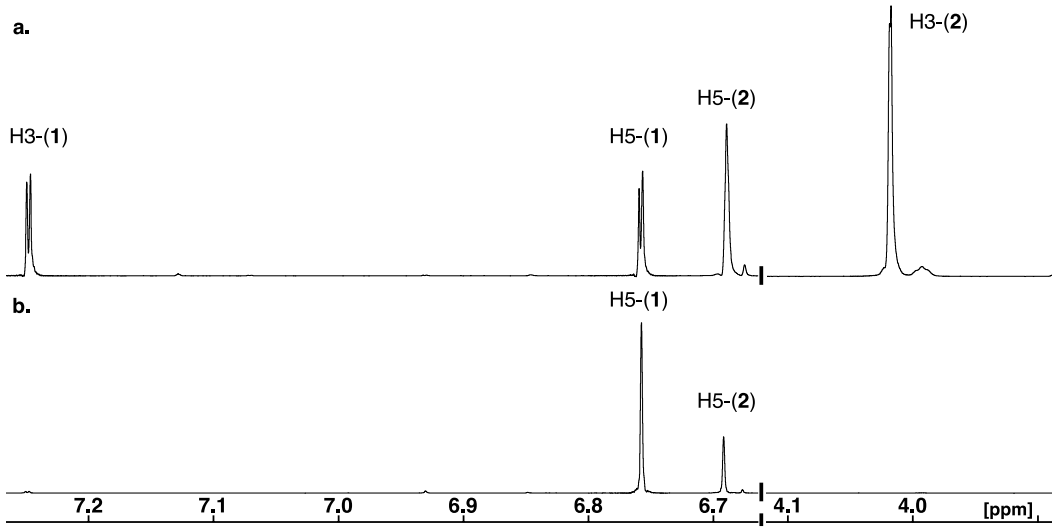
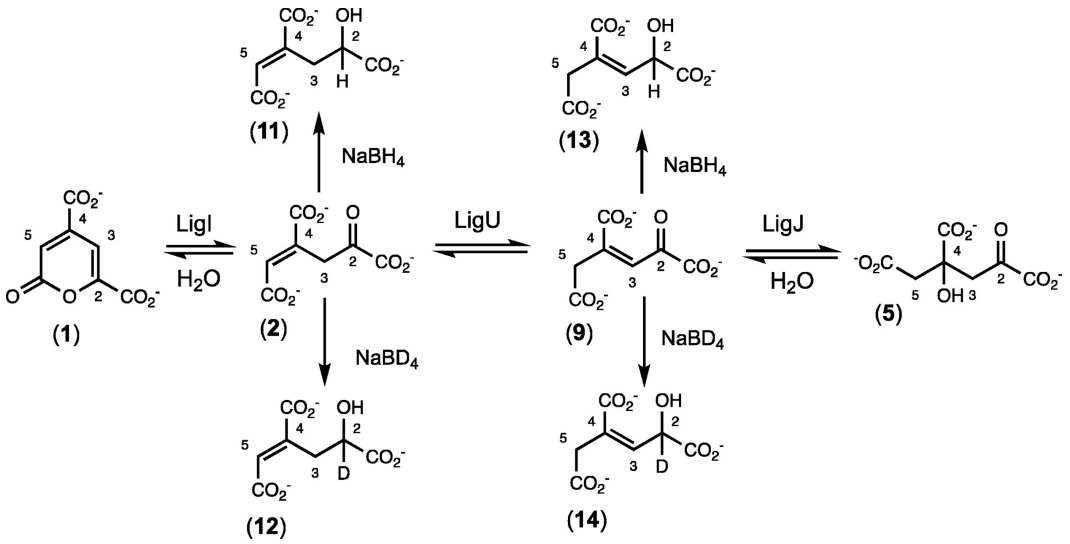


Figure 15. (a) ^{1}H NMR spectrum of PDC (1) and OMA (2) in 10% D_2O at pH measured 8.1. PDC (1) is represented by the resonance at 7.24 ppm that originates from the hydrogen at C3 and the resonance at 6.75 ppm from the hydrogen at C5 with a small ^{4}J coupling constant of 1.4 Hz. OMA (2) is represented by the resonances at 6.69 and 4.02 ppm that originates from the hydrogen at C5 and C3, respectively. (b) ^{1}H NMR spectrum of PDC (1) and OMA (2) in ~100% D_2O at pH 8.0 with a total incubation time of 23 minutes. The resonances for the hydrogens at C3 have exchanged with the solvent while the hydrogens at C5 are fixed. The data were collected using a WATERGATE zggpw5 pulse program with 128 scans with a relaxation delay of 5 seconds.

The OMA reaction product was further characterized by reduction of the carbonyl group at C2 with $NaBH_4$ (compound 11 in **Scheme 6**). After reduction, the downfield resonance of OMA shifts to 6.57 ppm for the hydrogen at C5 (**Figures 13b** and **14b**). The resonances for the two hydrogens at C3 shift to 2.85 ppm (doublet of doublets, ^2J = 13.6 Hz and ^3J = 3.0 Hz) and 2.71 ppm (doublet of doublets, ^2J = 13.6 Hz and ^3J = 11.1 Hz). The resonance for the hydrogen at C2 is at 3.95 ppm (doublet of doublets, ^3J = 11.1 and 3.0 Hz). When the reaction product is reduced by $NaBD_4$ (compound 12), the resonance at 3.95 ppm disappears and the two resonances at 2.85 and 2.71 ppm are each now doublets with a ^2J geminal coupling constant of 13.6 Hz (**Figures 13c** and **14c**). These results are fully consistent with the location of the double bond between C4 and C5 in the original OMA product (2). The configuration of the double bond between C4 and C5 was confirmed by measuring the coupling constant between the hydrogen at C5 and C3 in OMA that had been reduced with $NaBD_4$. The proton-decoupled ^{13}C resonance for C3 of 12 is a singlet at 34.1 ppm (**Figure 16a**). In the proton coupled spectrum (**Figure 16b**), the resonance for C3 is a doublet of triplets (^3J = 8.2 Hz and ^1J = 138.7 Hz). The 8.2 Hz coupling constant between the hydrogen at C5 and C3 is fully consistent with an *E* configuration for the double bond between C4 and C5, (10−12) as expected from the structure of the substrate PDC (1).

Scheme 6. The reduced products of LigI and LigU products with $NaBH_4$ and $NaBD_4$.



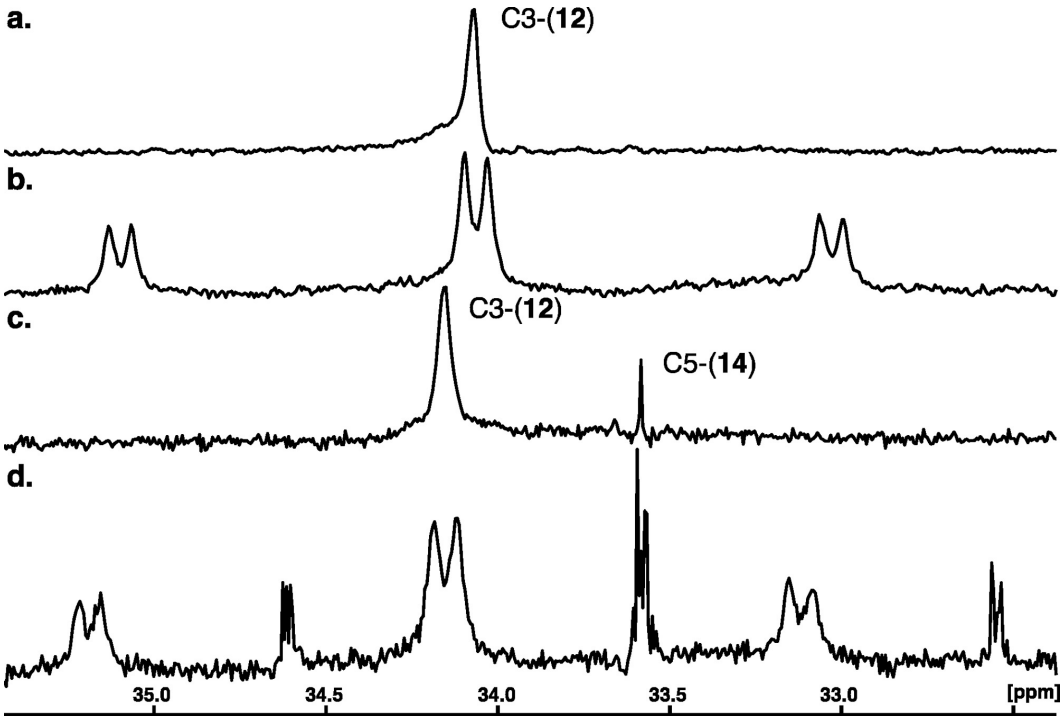


Figure 16. (a) ^{13}C proton-decoupled NMR spectrum for C3 of the LigI-catalyzed hydrolysis product after reduction with $NaBD_4$ (12). (b) ^{13}C proton-coupled NMR spectrum of the LigI-catalyzed hydrolysis product after reduction with the $NaBD_4$ product (12). (c) ^{13}C proton-decoupled NMR spectrum for C3 of the LigI-catalyzed hydrolysis product and for C5 of the LigU-catalyzed isomerization product after reduction with $NaBD_4$ (14). (d) ^{13}C proton-coupled NMR spectrum of the LigI-catalyzed hydrolysis product and for C5 of the LigU-catalyzed isomerization product after reduction with $NaBD_4$ (14).

2.3.3. Nonenzymatic Isomerization of the LigI Reaction Product

PDC (1) can be separated from the OMA (2) hydrolysis product by anion-exchange chromatography (**Figure 11a**). If the hydrolyzed reaction mixture is allowed to further incubate after removal of LigI, a second product begins to appear with a retention time of 13.0 min (**Figure 11b**). The apparent first-order rate constant for the appearance of the new product at pH 8.0 is 0.71 ± 0.02 h^{-1}. The ^{1}H NMR spectrum of the reaction mixture shows the appearance of two new resonances for the isomerized reaction product at 6.39 ppm (1H, triplet, ^{4}J = 1.5 Hz) and a doublet at 3.20 ppm (2H, ^{4}J = 1.5 Hz) with an intensity ratio of ~1:2 (**Figure 17**).

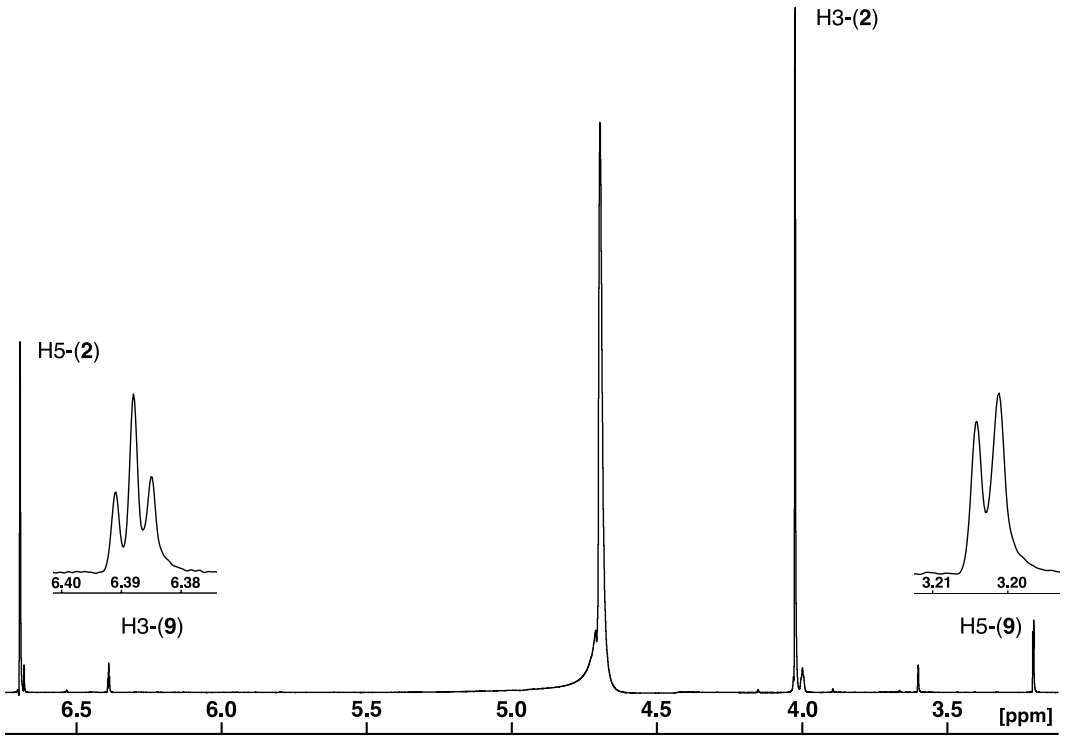


Figure 17. ^{1}H NMR spectrum of OMA (2) and the nonenzymatic conversion of OMA (2) to KCH (9) in 10% D_2O at pH 9.7 in the presence of 0.5 μM LigI. OMA (2) is represented by the resonance at 6.69 and 4.02 ppm that originates from the hydrogens at C5 and C3, respectively. The spectrum was taken after ~6 hour incubation and shows the nonenzymatic conversion of OMA (2) to KCH (9) by the appearance of two new resonances at 6.39 (triplet) and 3.20 (doublet) ppm that originates from C3 and C5, respectively, with ~1:2 ratio. KCH resonances have a small ^{4}J coupling constant of 1.5 Hz. The water resonance appears at 4.69 ppm and a resonance from an unknown origin appears at 3.60 ppm.

2.3.4. LigU Catalyzes the Isomerization of the LigI Reaction Product

The addition of LigU to the hydrolyzed reaction mixture formed by the incubation of LigI and PDC (1) accelerates the rate of formation of the isomerized reaction product. LigI and PDC were added together at pH 7.7 and the reaction products separated by high-performance liquid chromatography (HPLC) (**Figure 11a**). Under these conditions, the formation of the isomerized product is barely detectable. However, if LigI and LigU are incubated together at pH 7.7 with PDC, the isomerized reaction product rapidly forms (**Figure 11c**). No reaction was detectable in the absence of LigI at pH 7.7 (**Figure 11c**).

2.3.5. Characterization of the LigU-Isomerized Reaction Product by NMR Spectroscopy

At pH 9.5, the NMR spectrum of the reaction mixture that was formed after the incubation of LigI, LigU, and PDC (1) in H_2O exhibited two additional resonances for the LigU-isomerized product (9) at 6.39 ppm (triplet, ^4J = 1.5 Hz) for the hydrogen at C3 and a doublet at 3.20 ppm (^4J = 1.3 Hz for the two hydrogens at C5) (**Figures 18a** and **19a**). These are the same two resonances that are observed after the nonenzymatic isomerization of the LigI reaction product (**Figure 17**). The solution containing the LigI and LigU reaction products was reduced with $NaBH_4$, lyophilized, and then resuspended in D_2O. New resonances (**Figures 18b** and **19b**) were observed for the chemically reduced LigU product (13) at 5.47 ppm (doublet of triplets, ^3J = 9.1 Hz and ^4J = 1.1 Hz) for the hydrogen at C3, 4.84 ppm (doublet, ^3J = 9.0 Hz) for the hydrogen at C2, and 3.11

ppm (two doublets of doublets, ^{2}J = 15.6 Hz and ^{4}J = 1.1 Hz) for the two hydrogens at C5 at pH 8.21. When the reaction mixture was reduced with $NaBD_4$ (14), the resonance at 4.83 ppm disappeared because the hydrogen at C2 in 14 was replaced with deuterium (**Figures 18c** and **19c**). The resonance at 5.47 ppm for the hydrogen at C3 was then a broad singlet due to coupling with the deuterium at C2. The resonance at 3.11 ppm for the two hydrogens at C5 appeared as a pair of doublets due to a ^{2}J geminal coupling constant of 15.6 Hz. The coupling constants and chemical shift values for the chemically reduced product of the LigU-catalyzed reaction product are consistent with only the isomerization of the double bond in OMA (2) from C5/C4 to C4/C3 as in 2-keto-4-carboxy-3-hexenedioic acid (KCH, 9 or 10).

In an attempt to determine the stereochemistry of the double bond between C4 and C3 in the LigU-catalyzed reaction product (9 or 10), the coupling constant between the hydrogen at C3 and C5 was determined for the $NaBD_4$-reduced product (14). The proton-decoupled carbon resonances for the methylene carbon at C5 of 14 and the methylene carbon at C3 of 12 are illustrated in **Figure 16c**. The proton-coupled carbon spectrum in **Figure 16d** shows C5 of 14 at 33.6 ppm with a $^{3}J_{C5H3}$ of ~2.7 Hz and a $^{1}J_{C5H5}$ of 129 Hz. The size of the ^{3}J coupling constant is consistent with a *Z* configuration about the double bond between C3 and C4 of 14, (10−12) and thus, LigU catalyzes the isomerization of OMA (2) to the *Z* isomer of KCH (9).

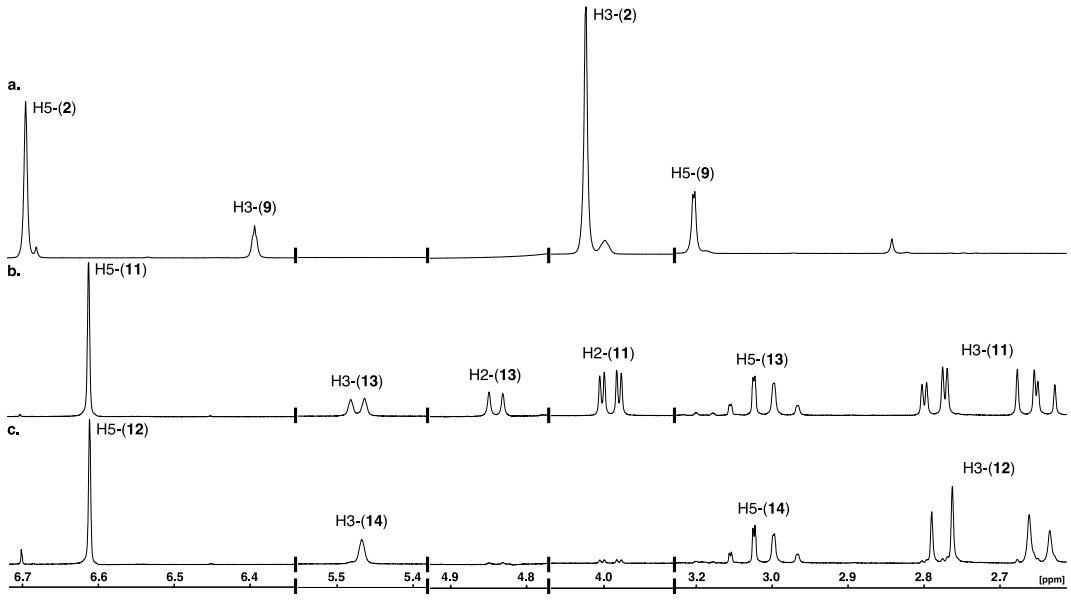


Figure 18. (a) ^{1}H NMR spectrum of reaction products 2 and 9 formed enzymatically by the hydrolysis of PDC (1) with LigI and LigU at pH 9.2. (b) NMR spectrum of the hydrolysis and isomerization products after reduction with $NaBH_4$ at pH 8.2 (13). (c) NMR spectrum of the hydrolysis and isomerization products after reduction with $NaBD_4$ at pH 8.4 (14).

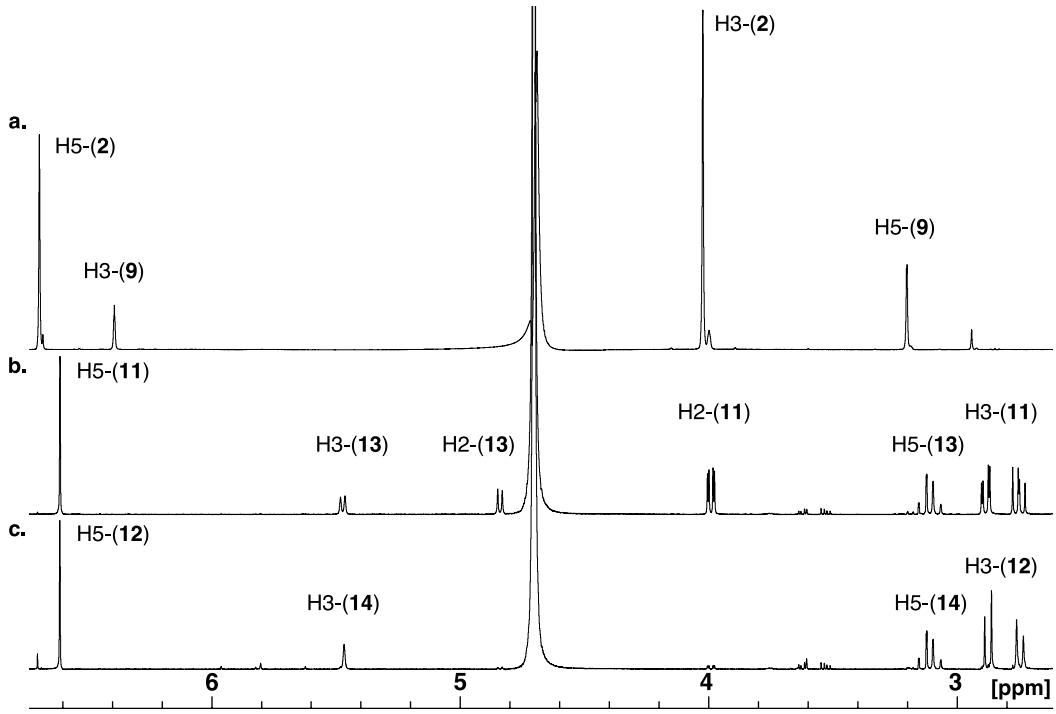


Figure 19. The full width ^{1}H NMR spectra of Figure 18.

2.3.6. Determination of the Equilibrium Constant for the LigU-Catalyzed Reaction

The equilibrium constant for the LigU-catalyzed isomerase reaction was measured using NMR spectroscopy. The initial reaction mixture contained 5.0 mM PDC, 0.4 μM LigI, 0.4 μM LigU, 10% D_2O, and 200 mM buffer. Potassium phosphate/KOH was used for pH 7.7, potassium carbonate/KOH for pH 8.9, and sodium bicarbonate/HCl for pH 10.1 in a volume of 600 μL. The reaction mixture was incubated for ~30 min before the relative concentrations of the OMA (2) and KCH (9) for the LigU- catalyzed reaction were determined by integration of the NMR signals for the hydrogens at C3 and C5. The equilibrium constants at pH 7.7, 8.9, and 10.1 were determined to be 0.27, 0.29, and 0.35, respectively, with an average value of 0.30 ± 0.03.

2.3.7. Determination of Extinction Coefficients for OMA and KCH

To measure the kinetic constants for the reaction catalyzed by LigU spectrophotometrically, the extinction coefficient for OMA (2) was determined at the isosbestic point (265 nm) relative to that of PDC (1) by allowing various amounts of PDC (from 0 to 0.5 mM) to be enzymatically hydrolyzed to equilibrium using LigI in 100 mM $NaHCO_3$/KCl buffer at pH 9.2 (**Figure 20**). An extinction coefficient of 1290 $M^{-1}\,cm^{-1}$ for OMA was determined from a plot of the absorbance at 265 nm and the initial concentration of PDC. The extinction coefficient at 265 nm for KCH (9) was obtained using reaction mixtures that initially contained PDC (from 0 to 0.5 mM), 0.5 μM LigI, and 0.5 μM LigU, in 100 mM $NaHCO_3$/HCl (pH 9.2). The equilibrium concentration of PDC (PDC_{eq}) was determined from the absorbance at 312 nm using an

extinction coefficient of 6250 M^{-1} cm^{-1} (4). The sum of the equilibrium concentrations of OMA and KCH (OMA_{eq} + KCH_{eq}) was obtained by subtraction of the equilibrium concentration of PDC from the initial concentration of PDC (PDC_o − PDC_{eq}). The equilibrium concentrations of OMA (OMA_{eq}) and KCH (KCH_{eq}) were subsequently obtained using the equilibrium ratio of OMA to KCH (KCH_{eq} / OMA_{eq} = 0.30). The extinction coefficient for KCH of 7190 M^{-1} cm^{-1} was determined from a plot of the net absorbance at 265 nm (after correcting for the amount of OMA_{eq} and PDC_{eq}) versus the equilibrium concentration of KCH.

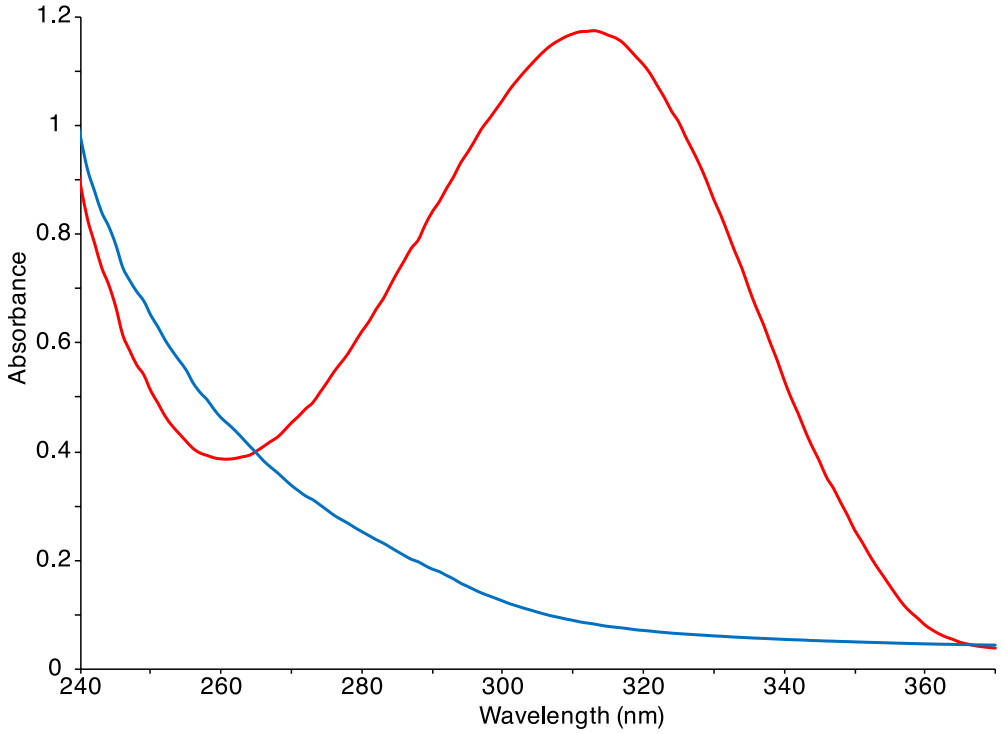


Figure 20. UV-visible absorbance spectrum of 0.3 mM PDC (red) and OMA (blue) at pH 10. The isosbestic point is at 265 nm. The assay was prepared by incubation of 0.3 mM PDC with 1.0 µM LigI in 150 mM sodium bicarbonate/HCl pH 10 to a volume total of 250 µL and measuring the absorbance at 30 °C.

2.3.8. LigU Kinetic Parameters

The kinetic constants for the LigU-catalyzed reaction were determined by following the formation of KCH (9) at 265 nm using a differential extinction coefficient of 5900 $M^{-1}\,cm^{-1}$ (7190−1290 $M^{-1}\,cm^{-1}$). The kinetic parameters for the LigU-catalyzed reaction at pH 8.0 are 170 ± 30 μM, 1300 ± 120 s^{-1}, and (7.7 ± 1.5) × $10^6\,M^{-1}\,s^{-1}$ for k_{cat}, K_m, and k_{cat}/K_m, respectively.

2.3.9. Preferred Substrate for the Reaction Catalyzed by LigJ

The preferred substrate for the LigJ-catalyzed reaction was initially determined using anion-exchange chromatography. The reaction mixture for the control experiment contained 100 nM LigI and 1 nM LigU with 1.0 mM PDC (1) in 125 mM TEA/KOH (pH 7.7) in a total volume of 500 μL. The enzymes were removed by filtration by being passed through a PALL Nanosep 10 kDa filter after an incubation period of 15 min. The filtrate was collected, and a 100 μL aliquot was loaded onto the HPLC column; the elution profiles of PDC, OMA, and KCH are shown in **Figure 11d**. LigJ (0.8 μM) was mixed with the filtrate solution; a 100 μL aliquot was loaded onto the column, and the reaction products were assessed by HPLC immediately (**Figure 11d**). The only notable change in the chromatogram is the significant decrease in the concentration of the LigU-catalyzed reaction product, KCH.

2.3.10. Hydrogen-Deuterium Exchange Catalyzed by LigJ

LigJ enzymatically hydrates KCH (9) into CHA (5) via the addition of water across the double bond between C3 and C4. The ^{1}H NMR spectrum of CHA exhibits two sets of doublets of doublets for the protons at C3 and C5, due to the new chiral center at C4. The resonances for C5 are centered at 2.51 ppm (^{2}J = 14.9 Hz), and those for C3 are centered at 3.23 ppm (^{2}J = 17.6 Hz) as shown in **Figure 21a**. If the CHA product is lyophilized and subsequently resuspended in ~100% D_2O in the presence of 0.5 μM LigJ, then the doublet of doublets centered at 3.23 ppm rapidly becomes a broadened singlet at 3.18 ppm due to the exchange of one of the two hydrogens at C3 with the solvent (**Figure 21b**). There is no noticeable exchange of the protons at C5. These results further support the conclusion that the double bond of the substrate hydrated by LigJ is located between C3 and C4 rather than between C4 and C5.

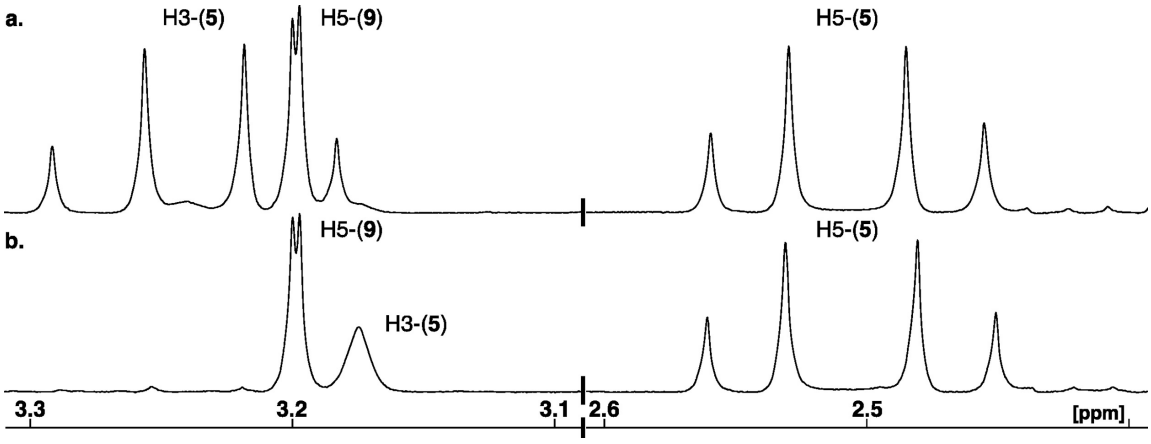


Figure 21. (a) ^{1}H NMR spectrum of a mixture of CHA (5) and KCH (9) in 10% D_2O at pH 7.2 after the addition of LigJ. (b) ^{1}H NMR spectrum of CHA and KCH at pH 7.2 after LigJ was added to the solution in ~100% D_2O.

2.3.11. LigK-Catalyzed Formation of HMG and Dehydration by LigJ

It has been shown previously that LigK catalyzes the formation of HMG (7) from two molecules of pyruvate (13,14), and that LigJ catalyzes the hydration of CHPD (6) (or an isomerized product) to HMG (7). To further support the proposal that the preferred substrate for LigJ is KCH (9), where the double bond is between C3 and C4, rather than between C4 and C5, we incubated HMG (made enzymatically by LigK from pyruvate) with LigJ. The ^{1}H NMR spectrum of the reaction mixture formed after the addition of LigK to 10 mM pyruvate at pH 8.0 is presented in **Figures 22a** and **23a**. The three hydrogens from the C5 methyl group of HMG (7) resonate at 1.31 ppm, and the two hydrogens at C3 form a doublet of doublets centered at 3.19 ppm. After the addition of LigJ to this reaction mixture, a new quartet appears at 6.39 ppm and a doublet appears at 2.02 ppm with a relative integration of ~1:3 and a coupling constant of 1.6 Hz (**Figures 22b** and **23b**) that was verified by a ^{1}H–^{1}H COSY experiment. These resonances support a structure for the dehydrated product with a conjugated double bond between C3 and C4 (compound 15) rather than between C4 and C5 as shown in **Scheme 5e**.

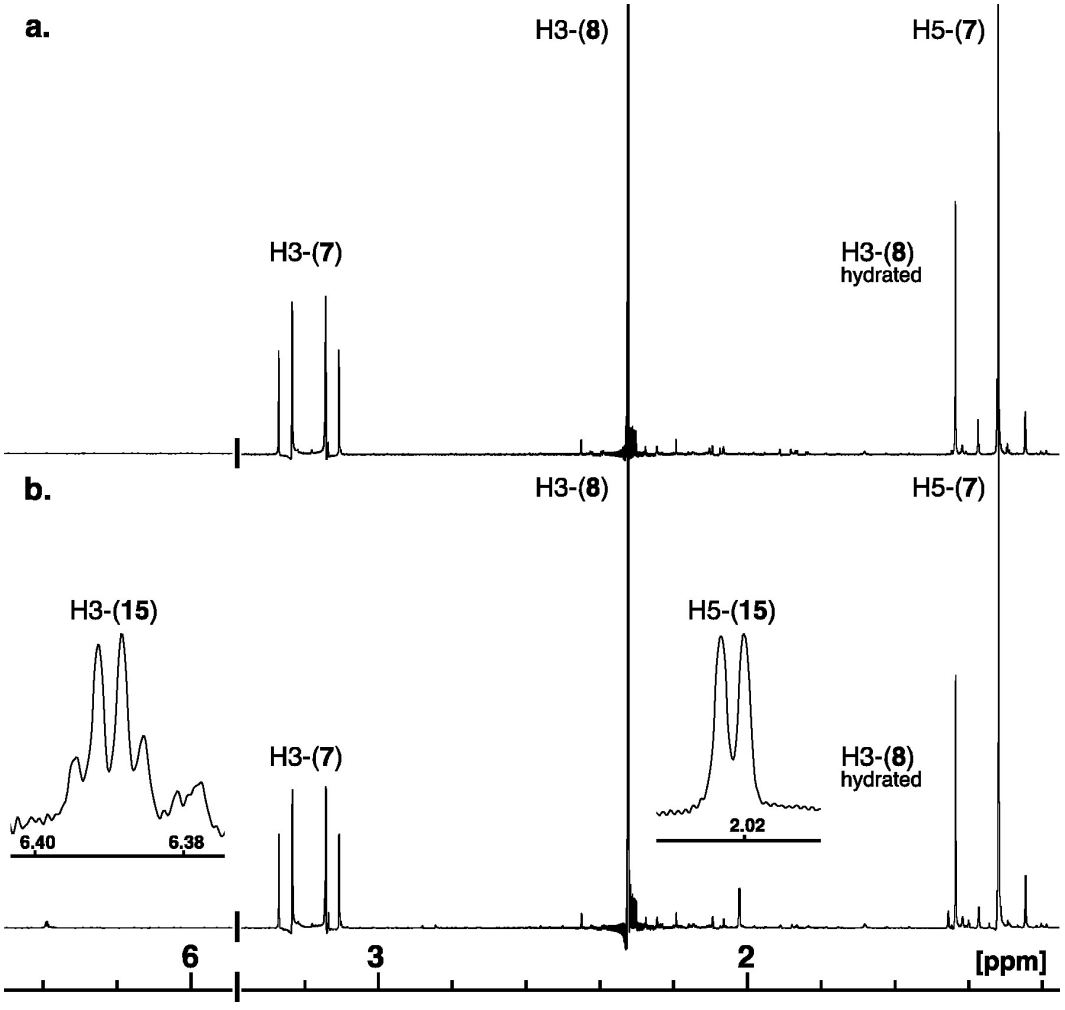


Figure 22. (a) ^{1}H NMR spectrum for the formation of HMG (7) after the addition of LigK to pyruvate at pH 7.3. (b) ^{1}H NMR spectrum of the reaction mixture after the addition of LigJ to HMG formed by the condensation of pyruvate by LigK. Insets show the formation of compound 15 via the dehydration at C3 and C4 by LigJ at pH 7.3.

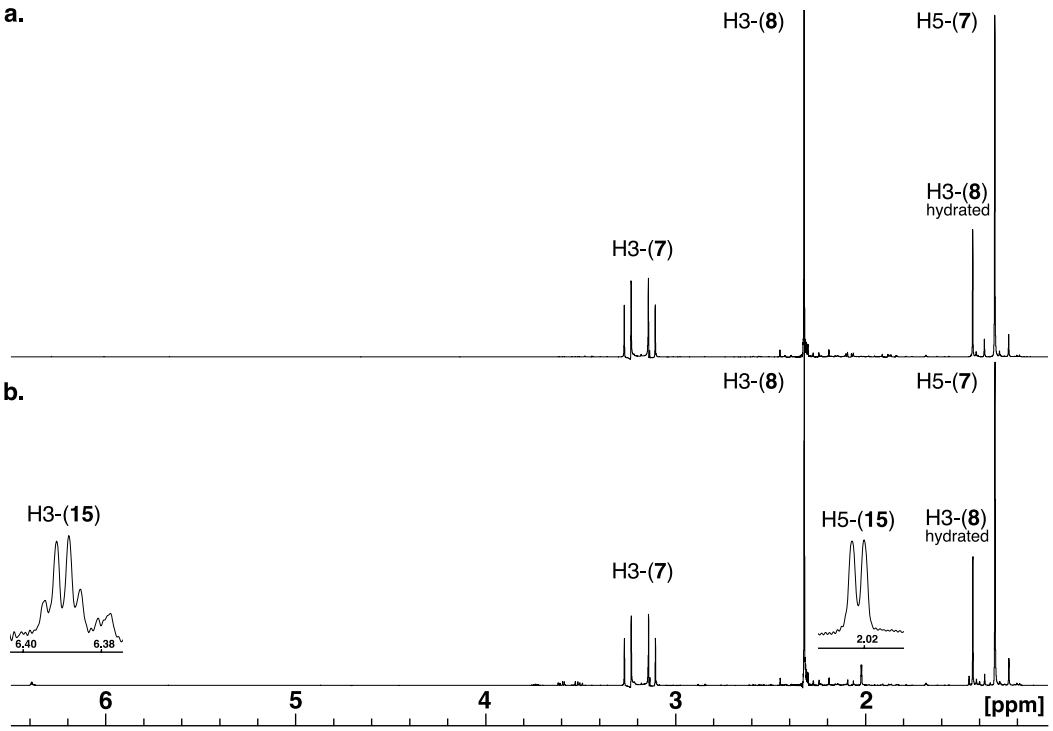


Figure 23. ^{1}H NMR full width spectra of Figure 22. The water resonance is suppressed at 4.70 ppm.

2.4. Discussion

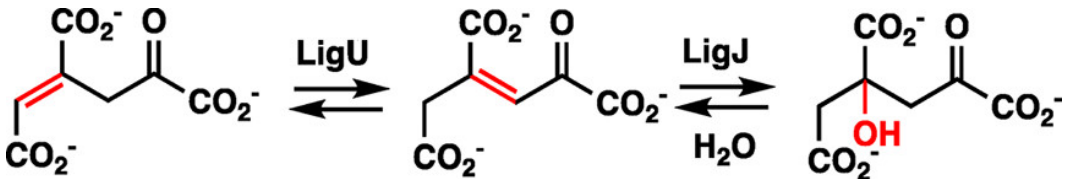
2.4.1. Functional Annotation of LigU

It has been shown previously that LigI from *Sphingobium* sp. SYK-6 catalyzes the hydrolysis of PDC (1) to OMA (2) (4). Using ^{1}H and ^{13}C NMR spectroscopy, we have confirmed the structure of the OMA product and have demonstrated that the double bond between C4 and C5 is of the *E* configuration. We have also shown by ultraviolet (UV) and NMR spectroscopy that the OMA product is somewhat unstable and isomerizes to KCH (9) where the double bond has migrated from between C4 and C5 to between C3 and C4. The nonenzymatic isomerization reaction is relatively slow at pH 8.0 ($t_{1/2}$ = 59 min) but is rapidly catalyzed by the action of LigU with a value of k_{cat} that is greater than 1000 s^{-1}. LigU thus catalyzes the isomerization of OMA (2) to the *Z* isomer of KCH (9), presumably via a 1,3- proton transfer reaction. KCH has previously been suggested as an intermediate in the bacterial degradation of 3,4,5- trimethoxycinnamic acid(15) and gallic acid (8).

LigJ has previously been shown to catalyze the formation of CHA (5) via the addition of hydroxide to C4 of the substrate (16). We have now demonstrated that the actual substrate for the catalytic activity of LigJ is the *Z* isomer of KCH (9) rather than OMA (2) or its enol tautomer. Thus, water is added to the double bond between C3 and C4 of KCH (9) rather than the double bond between C4 and C5 of OMA (2). This conclusion is supported by the preferential loss of KCH (9), relative to that of OMA, when LigJ is added to a mixture of KCH and OMA (**Figure 11d**). This conclusion is

also supported by the preferential exchange with solvent of the hydrogens at C3, relative to the C5 hydrogens, when CHA (5) is incubated with LigJ in D_2O (**Figure 21b**). The reactions catalyzed by LigU and LigJ are summarized in **Scheme 7**.

Scheme 7. The reactions catalyzed by LigU and LigJ.



2.4.2. Functional Homologues of the LigU Isomerase

The amino acid sequence of LigU is similar to the sequences of those enzymes that have been shown previously to have the diaminopimelate epimerase protein fold (17). The enzymes with a verified substrate profile and a protein sequence identity of >35% include methylaconitate isomerase (PrpF) from *Shewanella oneidensis* (35%), methylitaconate-Δ-isomerase (Mii) from *Eubacterium barkeri* (41%), and GalD from *P. putida* KT2440 (61%) (6, 18-20). The three-dimensional structure of the ligand-free

form of PrpF has been determined (PDB entry 2PVZ) in addition to that with malonate (PDB entry 5K87) or trans-aconitate bound in the active site (PDB entry 2PW0) (18, 19). The crystal structure of Mii has been determined (PDB entry 3G7K) (20), but the structure of GalD from *P. putida* KT2440 has not (6). Of interest are the fully conserved residues within the active site cleft between the two structural domains of these proteins that are postulated to be the general acid/base pair for the proton transfer step(s), which include Lys66 and Cys100 [numbering based on the LigU sequence (**Figure 24**)]. PrpF catalyzes the isomerization of (4*S*,*Z*)-4-methylaconitate to (*Z*)-2-methylaconitate through a base-catalyzed allylic rearrangement (18, 19). Mii catalyzes the isomerization of (*R*)-3-methylitaconate to dimethylmaleate (20). The structural similarities of the substrates for Mii, PrpF, and LigU are illustrated in **Scheme 8**.

LigU has previously been proposed to function as an OMA tautomerase that enzymatically converts (4*E*)-OMA (2) to (2*E*,4*E*)-CHM (4) based on ^{1}H NMR studies using the GalD homologue from *P. putida* KT2440 (6, 7). In that investigation, the addition of GalD to OMA results in the appearance of a new resonance at 6.73 ppm, which was subsequently thought to be the hydrogen attached to C3 of CHM (6, 7). The resonance for the hydrogen at C5 of CHM is suggested to appear at 6.63 ppm but is not visible because of the presence of the hydrogen from C5 of unreacted OMA that also resonates at 6.63 ppm. Unfortunately, the quality of the published NMR spectra for this experiment is insufficient to support the formation of CHM (4) or KCH (9) by the catalytic activity of GalD. In contrast, here we have clearly demonstrated that LigU catalyzes the interconversion of OMA to KCH (9) with the isomerization of the double

bond between C5 and C4 to that between C4 and C3. We have further demonstrated that LigJ catalyzes the hydration of KCH to CHA (5).

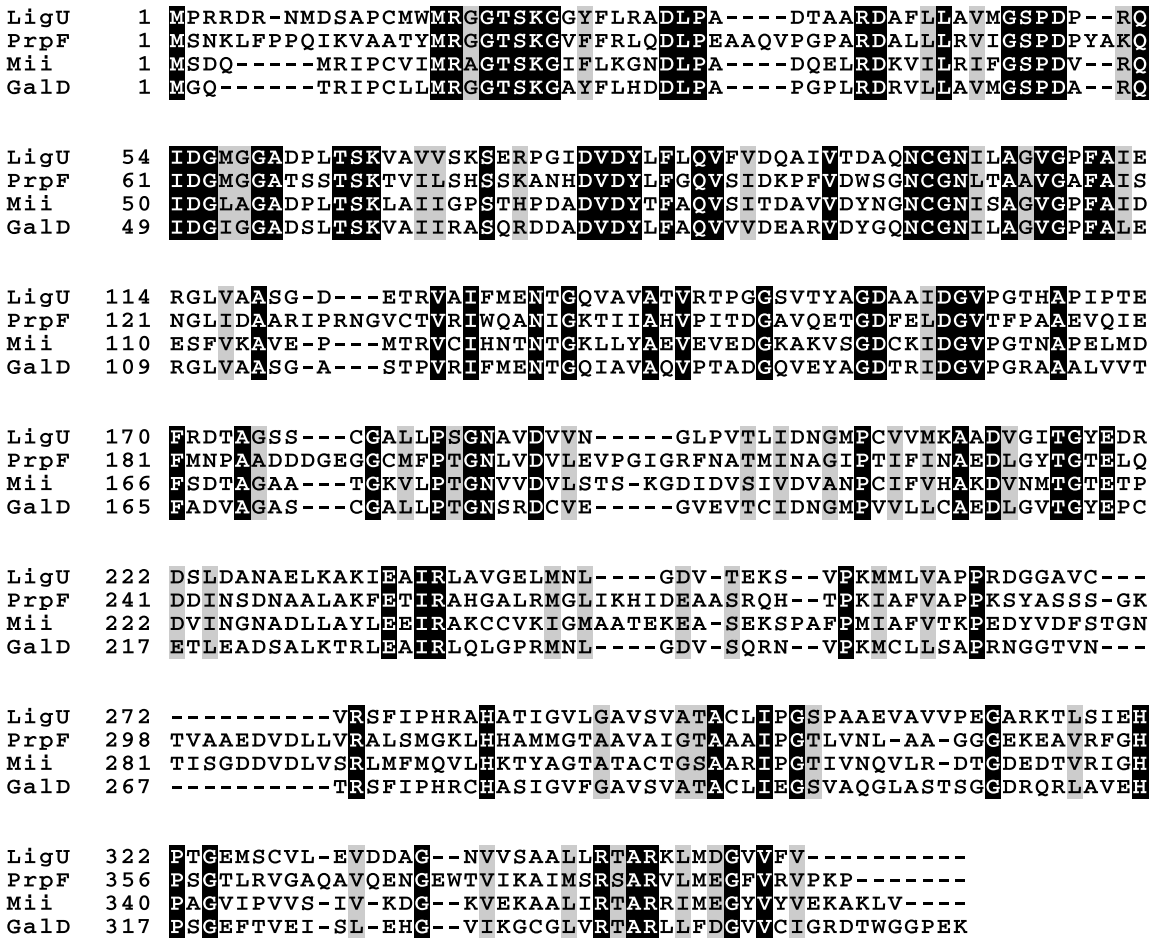
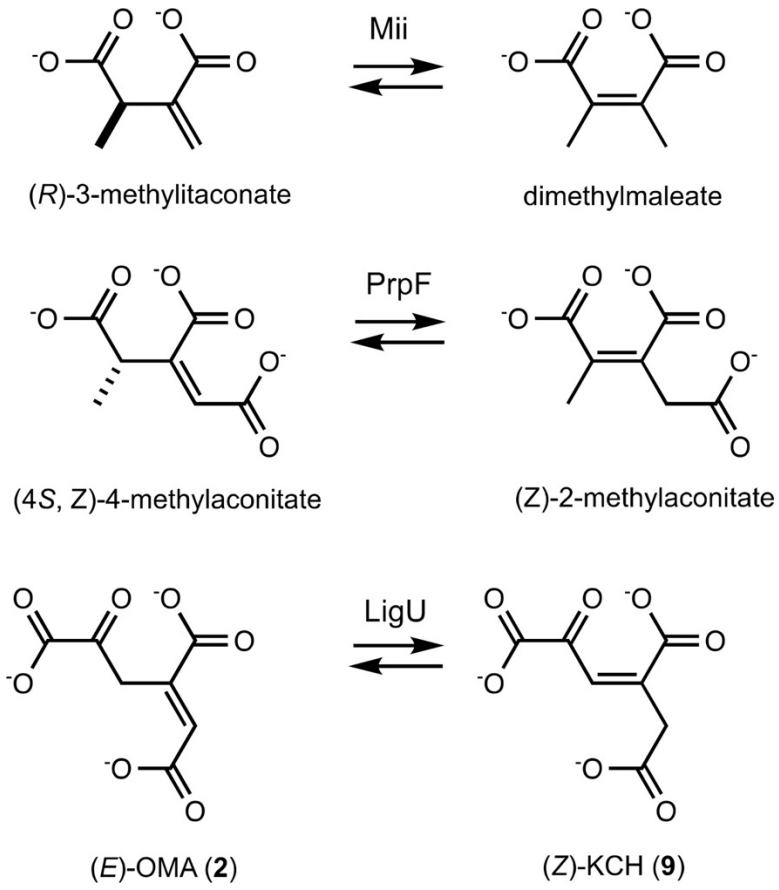


Figure 24. Sequence alignment of LigU and related proteins. LigU (UniProt entry: Q0KJL4) aligned with PrpF (UniProt entry: Q8EJW4.1), Mii (UniProt entry: Q0QLE6.1), and GalD (UniProt entry: Q88JY0.1). The black highlighted residues are identical and the gray highlighted residues are 50 to 75% identical. The sequence alignment was created using ExPASy BOXSHADE version 3.21, https://embnet.vital-it.ch/software/BOX_form.html.

Scheme 8. The structural similarities of the substrates for Mii, PrpF, and LigU.



2.5. References

1. Pollegioni, L., Tonin, F., and Rosini, E. (2015) Lignin-degrading enzymes. *FEBS J. 282,* 1190−1213.
2. Brown, M. E., and Chang, M. C. Y. (2014) Exploring bacterial lignin degradation. *Curr. Opin. Chem. Biol. 19,* 1−7.
3. Kamimura, N., Takahashi, K., Mori, K., Araki, T., Fujita, M., Higuchi, Y., and Masai, E. (2017) Bacterial catabolism of lignin-derived aromatics: New findings in a recent decade. *Environ. Microbiol. Rep. 9,* 679−705.
4. Hobbs, M. E., Malashkevich, V., Williams, H. J., Xu, C., Sauder, J. M., Burley, S. K., Almo, S. C., and Raushel, F. M. (2012) Structure and Catalytic Mechanism of LigI: Insight into the Amidohydrolase Enzymes of cog3618 and Lignin Degradation. *Biochemistry 51,* 3497−3507.
5. Nogales,J.,Canales,Á.,Jimeńez-Barbero,J.,García,J.L.,and Díaz, E. (2005) Molecular Characterization of the Gallate Dioxygenase from *Pseudomonas putida* KT2440. *J. Biol. Chem. 280,* 35382−35390.
6. Nogales,J.,Canales,Á.,Jimeńez-Barbero,J.,Serra,B.,Pingarroń, J. M., García, J. L., and Díaz, E. (2011) Unravelling the Gallic acid Degradation Pathway in Bacteria: the gal Cluster from *Pseudomonas putida*. *Mol. Microbiol. 79,* 359−374.
7. Mazurkewich, S., Brott, A. S., Kimber, M. S., and Seah, S. Y. K. (2016) Structural and Kinetic Characterization of the 4-Carboxy-2-hydroxymuconate Hydratase from

the Gallate and Protocatechuate 4,5-Cleavage Pathways of *Pseudomonas putida* KT2440. *J. Biol. Chem. 291,* 7669−7686.

8. Kersten, P. J., Dagley, S., Whittaker, J. W., Arciero, D. M., and Lipscomb, J. D. (1982) 2-Pyrone-4,6-Dicarboxylic Acid, a Catabolite of Gallic Acids in *Pseudomonas* Species. *J. Bacteriol. 152,* 1154−1162.
9. Liu, M., Mao, X., Ye, C., Huang, H., Nicholson, J. K., and Lindon, J. C. (1998) Improved WATERGATE Pulse Sequence for Solvent Suppression in NMR Spectroscopy. *J. Magn. Reson. 132,* 125−129.
10. Stubbe, J. A., and Kenyon, G. L. (1971) Analogs of Phosphoenolpyruvate. On the Specificity of Pyruvate Kinase from Rabbit Muscle. *Biochemistry 10,* 2669−2677.
11. Braun, S. (1978) High Resolution ^{13}C NMR Spectroscopy IV- Stereochemical Assignments in Butenedioic Acids and 3-Pentene-2-ones. *Org. Magn. Reson. 11,* 197−203.
12. Letcher, R. M., and Acheson, R. M. (1981) Vicinal C, H Spin Coupling Constants in Determining Alkene Stereochemistry. *Org. Magn. Reson. 16,* 316−318.
13. Maruyama, K. (1990) Purification and Properties of 4-Hydroxy-4-Methyl-2-Oxoglutarate Aldolase from *Pseudomonas ochraceae* Grown on Phthalate. *J. Biochem. 108,* 327−333.
14. Wang, W., Mazurkewich, S., Kimber, M. S., and Seah, S. Y. K. (2010) Structural and Kinetic Characterization of 4-Hydroxy-4- methyl-2-oxoglutarate/4-Carboxy-4-hydroxy-2-oxoadipate Aldolase, a Protocatechuate Degradation Enzyme

Evolutionarily Convergent with HpaI and DmpG Pyruvate Aldolases. *J. Biol. Chem. 285,* 36608−36615.

15. Donnelly, M. I., and Dagley, S. (1981) Bacterial Degradation of 3,4,5-Trimethoxycinnamic Acid with Production of Methanol. *J. Biol. Chem. 147,* 471−476.
16. Hara, H., Masai, E., Katayama, Y., and Fukuda, M. (2000) The 4-Oxalomesaconate Hydratase Gene, Involved in the Protocatechuate 4,5-Cleavage Pathway, Is Essential to Vanillate and Syringate Degradation in *Sphingomonas paucimobilis* SYK-6. *J. Bacteriol. 182,* 6950−6957.
17. Pillai, B., Cherney, M. M., Diaper, C. M., Sutherland, A., Blanchard, J. S., Vederas, J. C., and James, M. N. G. (2006) Structural insights into stereochemical inversion by diaminopimelate epimerase: An antibacterial drug target. *Proc. Natl. Acad. Sci. U. S. A. 103,* 8668− 8673.
18. Garvey, G. S., Rocco, C. J., Escalante-Semerena, J. C., and Rayment, I. (2007) The three-dimensional crystal structure of the PrpF protein of *Shewanella oneidensis* complexed with trans-aconitate: Insights into its biological function. *Protein Sci. 16,* 1274−1284.
19. Rocco, C. J., Wetterhorn, K. M., Garvey, G. S., Rayment, I., and Escalante-Semerena, J. C. (2017) The PrpF protein of *Shewanella oneidensis* MR-1 catalyzes the isomerization of 2-methyl-cis-aconitate during the catabolism of propionate via the AcnD-dependent 2- methylcitric acid cycle. *PLoS One 12,* e0188130.

20. Velarde, M., Macieira, S., Hilberg, M., Bröker, G., Tu, S., Golding, B. T., Pierik, A. J., Buckel, W., and Messerschmidt, A. (2009) Crystal Structure and Putative Mechanism of 3-Methylitaconate-Δ- isomerase from *Eubacterium barkeri*. *J. Mol. Biol. 391,* 609−620.

3. STRUCTURE AND REACTION MECHANISM OF THE LIGJ HYDRATASE: AN ENZYME CRITICAL FOR THE BACTERIAL DEGRADATION OF LIGNIN IN THE PROTOCATECHUATE 4,5-CLEAVAGE PATHWAY[*]

3.1. Introduction

The degradation of lignin, the complex polymer that provides the structural rigidity to vascular plants, is not fully understood (1, 2). In the soil bacterium *Sphingobium* sp. SYK- 6, guaiacyl, one of the monomeric units of lignin, is converted to vanillate and then demethylated to form protocatechuate by LigM, where it is further metabolized through the protocatechuate (PCA) 4,5-cleavage pathway (3, 4). LigAB (protocatechuate 4,5-dioxygenase) converts PCA, in the presence of O_2, into 4-carboxy-2-hydroxymuconate-6-semialdehyde (CHMS) (5). LigC (CHMS dehydrogenase) oxidizes CHMS with $NADP^+$ to form 2-pyrone-4,6-dicarboxylate (PDC, 1) (6). LigI (PDC hydrolase) converts PDC into (4*E*)-oxalomesaconate (OMA, 2) (7). Recently, we have shown that LigU (OMA isomerase) transforms OMA into (3*Z*)-2-keto-4-carboxy-3-hexenedioate (KCH, 3) (8). LigJ (KCH hydratase) adds water across the C4=C3 bond of KCH to form 4-carboxy-4-hydroxy-2-oxoadipate (CHA, 4) (9-11). However, the stereo-chemistry for the addition of water to CHA has not previously been determined for the reaction catalyzed by LigJ. In the final step, LigK catalyzes the retro-aldol cleavage of

---

[*] Reprinted with permission from "Structure and Reaction Mechanism of the LigJ Hydratase: An Enzyme Critical for the Bacterial Degradation of Lignin in the Protocatechuate 4,5-Cleavage Pathway" by Tessily N. Hogancamp, Mark F. Mabanglo, and Frank M. Raushel, *Biochemistry*, 2018, *57* (40), pp 5841-5850, Copyright 2018 American Chemical Society

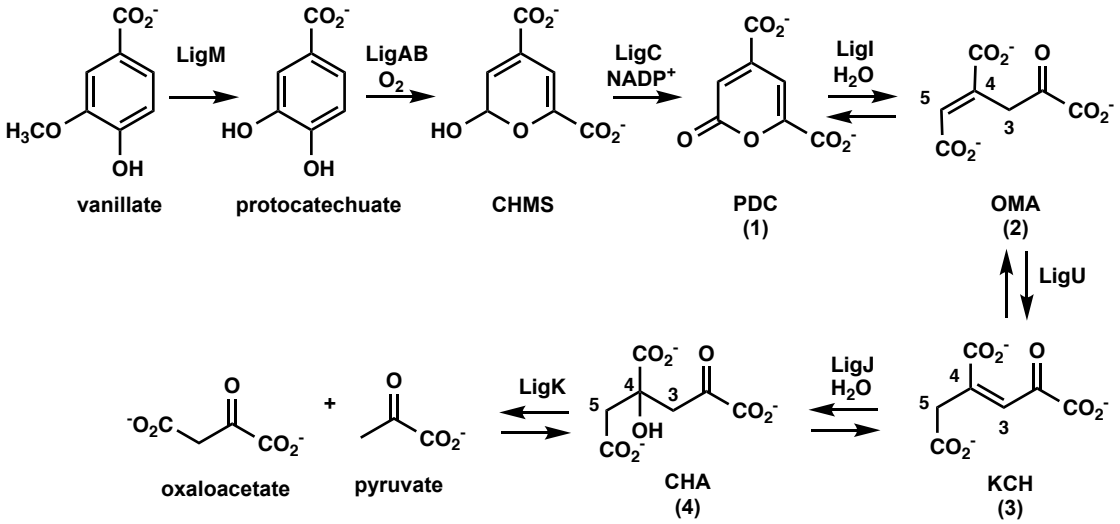
CHA to form pyruvate and oxaloacetate, and these products are ultimately consumed via the TCA cycle (**Scheme 9**) (10, 12-14).

The LigJ hydratase from *Sphingobium* sp. SYK-6 is a member of the amidohydrolase superfamily (AHS) (15). This large and prominent enzyme superfamily is currently composed of more than 100,000 unique protein sequences from both eukaryotic and prokaryotic organisms (16). The hallmark of the AHS is a distorted $(\beta/\alpha)_8$-barrel structural fold that houses a metal center at the C-terminal end of the β-barrel (17). In most cases, one or two divalent cations bind within the active site; however, a small number of examples bind up to three divalent cations, and certain enzymes, such as LigI, do not bind any divalent cations in the active site (7, 18). Nearly all enzymes in the AHS use the metal center in the active site to activate a water molecule for nucleophilic attack on an electrophilic substrate and catalyze the hydrolysis of ester and amide functional groups, including the deamination of aromatic bases. However, a small number of examples catalyze the decarboxylation of *ortho*-substituted benzoic acids (LigW) and the isomerization of uronate sugars (UxaC) (19, 20).

LigJ catalyzes the hydration of a double bond, and currently, no member of the amidohydrolase superfamily whose primary catalytic function is the hydration of a double bond has been structurally or mechanistically interrogated. The AHS is subdivided into 24 clusters of orthologous groups (COG), and LigJ is a member of cog2159 (21, 22). This COG also includes enzymes that catalyze decarboxylation reactions (LigW) and the hydrolysis of β-unsaturated ketones (LigY) (19, 23). Here we have determined the three-dimensional crystal structure of LigJ in the presence and

absence of bound ligands. With wild-type (WT) LigJ, we have succeeded in crystallizing the bound product, CHA. This has enabled the determination of the absolute stereochemistry at C4 of the product and the identification of critical residues contained within the active site. Crystallization and structure elucidation of an active site mutant have enabled the determination of the three-dimensional structure of the substrate-bound complex (KCH, 4) and a proposal for the mechanism of action of this most interesting enzyme in the AHS.

Scheme 9. The PCA 4,5-cleavage pathway in *Sphingobium* sp. SYK-6.



3.2. Material and Methods

3.2.1. Materials

All commercial chemicals used in this investigation were obtained from Sigma-Aldrich, unless otherwise stated. 2-Pyrone-4,6-dicarboxylate (PDC, 1) was synthesized according to a previously published protocol (24). Racemic 4-carboxy-4-hydroxy-2-oxoadipate (CHA, 4) was synthesized as previously described (25). Briefly, OAA was dissolved in a 3 M NaOH solution at 0 °C, and the pH adjusted to 7.5. The solution was incubated at 4 °C for 4 days and then lyophilized, yielding a mixture of the trisodium salt of 4 and $NaHCO_3$. The enzymes LigI (UniProt entry O87170), LigU (UniProt entry Q0KJL4), and LigK (UniProt entry G2IQQ8) were isolated and purified as previously described (7,8).

3.2.2. Expression and Isolation of LigJ

The open reading frame encoding LigJ from *Sphingobium* sp. SYK-6 (UniProt entry G2IQQ5) was synthesized by GenScript and ligated into a pET-30a (+) (EMD Millipore) vector and expressed in *Escherichia coli* BL21(DE3) competent cells (EMD Millipore) as described previously (8). Wild-type and mutant enzymes were purified using a C-terminal His tag and passed through a 5 mL HisTrap HP column (GE Healthcare) using an NGC liquid chromatography system (Bio-Rad Laboratories, Inc.) equilibrated with binding buffer. The binding buffer included 5 mM imidazole, 0.5 M KCl, and 20 mM HEPES/KOH (pH 7.9). The protein was eluted using a 0 to 40%

gradient of elution buffer that contained 500 mM imidazole, 0.25 M KCl, and 20 mM HEPES/KOH (pH 7.9) at a flow rate of 3 mL/min. The purity of the fractions collected was determined by sodium dodecyl sulfate−polyacrylamide gel electrophoresis. The appropriate fractions were combined and dialyzed against 3 L of 20 mM HEPES/KOH (pH 7.9) overnight at 4 °C. The dialysis membrane had a molecular weight cutoff of 3500 Da (Spectra/Por Laboratories). For wild-type LigJ, approximately 30 mg/L purified protein was obtained.

3.2.3. Mutation of LigJ

Four site-directed mutants of LigJ (E284Q, Y194F, T190A, and H223N) were constructed using the QuikChange protocol according to the manufacturer's instructions (Agilent). The forward primer for mutant E284Q (with the mutation underlined) was TCCTGTTCGGCAGCCAAATGGTTGGTGCAGTCC, and the reverse primer was GGACTGCACCAACCATTTGGCTGCCGAACAGGA (Integrated DNA technologies, IDT). The forward primer for mutant Y194F was GCGACCGGTGCCTATTTCCTGG CGGCCGACACG, and the reverse primer was CGTGTCG GCCGCCAGGAAATAGGCACCGGTCGC (IDT). The forward primer for mutant T190A was GGCTATGCACGCGGCCGGTGCCTATTACC, and the reverse primer was GGTAATAGGCACCGGCCGCGTGCATAGCC (IDT). The forward primer for mutant H223N was GCGTTTTATTATCCCGAATGGTGGCGGTGCAG, and the reverse primer was CTGCACCGCCACCATTCGGGATAATAAAACGC (IDT). The corresponding genes for each of the four mutant enzymes were sequenced to verify that

the appropriate changes to the DNA sequence were made. The four mutant enzymes were purified as described for the wild-type enzyme (8). The approximate yields of purified enzyme per liter of growth medium for each of the mutant enzymes were as follows: 4 mg for E284Q, 17 mg for Y194F, 15 mg for T190A, and 18 mg for H223N.

3.2.4. Molecular Weight Determination

The oligomeric state of LigJ was determined by size exclusion chromatography (SEC) using a 10/300 GL Superdex 200 column from GE Healthcare. The calibration of the column was conducted using the following molecular weight standards: carbonic anhydrase (29 kDa), albumin (66 kDa), alcohol dehydrogenase (150 kDa), β-amylase (200 kDa), and apoferritin (443 kDa). The elution buffer in all cases was 50 mM HEPES/KOH (pH 8.0) at a flow rate of 0.75 mL/min.

3.2.5. Determination of the Extinction Coefficient for CHA

To measure the kinetic constants for the LigJ-catalyzed reaction spectrophotometrically, the extinction coefficient for the product CHA (4) was determined at 265 nm from a plot of the absorbance at 265 nm relative to the concentration of CHA that had been synthesized chemically. At 265 nm, the extinction coefficient for CHA was determined to be 60 M^{-1} cm^{-1}.

3.2.6. Determination of the Equilibrium Constant for the LigJ Reaction

The equilibrium constant for the LigJ-catalyzed hydration reaction was measured using ^{1}H nuclear magnetic resonance (NMR) spectroscopy. A solution of 5 mM PDC was incubated with LigI, LigU, and LigJ (0.5 μM each) in 100 mM phosphate/KOH buffer (pH 8.0) until equilibrium was reached. To measure the equilibrium constant, the resonances at C3 of CHA and C5 of KCH were integrated and used to obtain the [CHA]/[KCH] ratio.

3.2.7. Measurement of Hydratase Activity

The initial velocity for the hydration of KCH to CHA by LigJ and the LigJ mutants was determined spectrophotometrically by monitoring the change in absorbance at 265 nm and 30 °C in a 96-well plate. The concentrations of LigJ used in the assays were as follows: wild type (4.0 nM), Y194F (5.0 nM), H223N (250 nM), T190A (4.0 nM), and E284Q (500 nM). The substrate KCH (3) was made *in situ* by the combined enzymatic activities of LigI (0.5 μM) and LigU (0.2 μM) to variable concentrations of PDC (from 0 to 1.4 mM) in 100 mM HEPES/KOH (pH 8.0) in a volume of 250 μL. The initial concentration of KCH was determined using the equilibrium constant for the LigU-catalyzed reaction ([KCH]/[OMA] = 0.30) after subtracting the equilibrium concentration of PDC (determined at 312 nm using an extinction coefficient ε of 6250 M^{-1} cm^{-1}) from the initial concentration of PDC (7, 8). The change in the concentration of the substrate after the addition of LigJ was determined using a differential extinction coefficient ($\Delta\varepsilon_{265}$) of 2173 M^{-1} cm^{-1}. This value was determined from the weighted

average of the extinction coefficients for KCH (16%; $\varepsilon = 7190\ M^{-1}\ cm^{-1}$), PDC (31%; $\varepsilon = 1290\ M^{-1}\ cm^{-1}$), and OMA (53%; $\varepsilon = 1290\ M^{-1}\ cm^{-1}$) of $2233\ M^{-1}\ cm^{-1}$ and the extinction coefficient of $60\ M^{-1}\ cm^{-1}$ for CHA (4) at 265 nm. The kinetic parameters were determined from a fit of the data to **eq 2**

$$v/E_t = k_{cat}(A)/(K_a + A)$$

Equation 2

where v is the initial velocity, E_t is the total enzyme concentration, k_{cat} is the turnover number, A is the substrate concentration, and K_a is the Michaelis constant.

3.2.8. Metal Analysis

The metal content of LigJ was determined with a PerkinElmer NexION 300D ICP-MS instrument. The detector was set to acquire data in analog mode with 20 sweeps, a 50 ms dwell time for each mass, and three replicates per reading. The internal standard used was ^{45}Sc from ICPMS- 71D (Inorganic Ventures). A five-point external calibration curve was used from a 10 µg/mL mixture of Cd, Co, Cu, Fe, Mn, Ni, and Zn in 2% HNO_3 to prepare serial dilutions of 10, 25, 50, 100, and 200 ppb. Metal-free solutions were prepared by filtering deionized water (Thermo Scientific Barnstead E-Pure Ultrapure Water Purification Systems, 120 V, 18.2 MΩ) with GD/X sterile 0.22 µm

PES Filter Media (Whatman GE Healthcare) and then through a column of Chelex 100 resin (Bio-Rad). The LigJ samples were prepared by changing the buffer from 20 mM HEPES/KOH (pH 7.9) to 50 mM metal-free MES/KOH (pH 6.1) by centrifugation through a Vivaspin 10 kDa molecular weight cutoff filter (GE Healthcare) at 4000 rpm until the concentration was approximately 1.0 mg/mL. In a final volume of 6 mL, 1, 2, and 3 μM solutions of LigJ were prepared in 1% HNO3. The samples were vortexed and heated to 100 °C to completely digest and solubilize the protein. The effect of added zinc on the catalytic activity of the E284Q mutant was determined by adding 0−1 mM Zn^{2+} directly to the mixture used for measuring the catalytic activity of LigJ.

3.2.9. Sequence Similarity Networks

Using the UniProt database, the sequences belonging to cog2159 were retrieved by searching "COG2159". This procedure identified 1299 unique protein sequences that were subsequently submitted to the Enzyme Function Initiative (EFI) Enzyme Similarity Tool (EST) (https://efi.igb.illinois.edu/efi-est) to generate a sequence similarity network (SSN) (26, 27). The alignment score parameter was set at 35, 40, 45, 50, 55, 60, 65, 70, 75, and 80 to generate networks at different stringencies. The SSN was viewed using Cytoscape (http://cytoscape.org) (28).

3.2.10. Crystallization and Structure Determination

The crystallization conditions for LigJ were identified from a sparse matrix screen using commercial reagents from Hampton Research. Purified LigJ (10 mg/mL)

was incubated with an equilibrium mixture of 1.3 mM OMA/KCH and then mixed in a 1:1 ratio with various screening solutions on 96-3 LVR Intelliplates using a Mosquito crystallization robot. Crystal plates were incubated at 18 °C, and crystal formation occurred within a week. The best diffracting crystals grew in a well containing 0.1 M sodium malonate (pH 5.0) and 12% (w/v) poly(ethylene glycol) 3350. Subsequent crystals of LigJ, including the E284Q mutant with the bound substrate (KCH), were obtained using the same conditions.

Well-diffracting crystals were prepared for data collection by being briefly soaked in a cryoprotectant solution containing the reservoir solution supplemented with 20% ethylene glycol and then flash-frozen in liquid nitrogen. Diffraction data were collected either in house using X-rays from a rotating anode generator (Rigaku HF-700) or from Stanford Synchrotron Radiation Lightsource (SSRL) beamline 11-1 (Table 1). Diffraction data were indexed, integrated, and scaled using HKL2000 (29). The structure of LigJ, in complex with the product CHA, was determined by molecular replacement using as a search model the 1.8 Å structure of *Rhodopseudomonas palustris* (strain ATCC BAA-98) LigJ [Protein Data Bank (PDB) entry 2GWG] that shares 62% sequence identity (30). The *Sphingobium* sp. SYK-6 LigJ structure was built using PHENIX Autobuild (31) with its amino acid sequence as a guide. Initial refinement was performed in PHENIX (31) with coordinate shaking and simulated annealing to remove model bias. Subsequent model building was performed in COOT (32), while the structures of the bound small molecules were generated using PHENIX eLBOW (31). Further refinement was performed in PHENIX with coordinate, B factor, occupancy, and

TLS parameter optimization. The refined structure of the LigJ−KCH substrate complex was then used as a search model in the structure determination of subsequent LigJ complexes using molecular replacement. The crystallographic data collection and refinement statistics for the LigJ structures are listed in **Table 2**.

	LigJ WT	E284Q	LigJ WT
Data collection			
Wavelength (Å)	1.54	0.98	1.54
Resolution (Å)	70.30-2.20 (2.32-2.20)	50.00-1.65 (1.68-1.65)	50.00-2.02 (2.05-2.02)
Space group	P 2_1	P 2_1	P 2_1
a, b, c (Å)	82.2, 78.2, 135.7	82.7, 78.3, 136.2	82.5, 78.3, 135.6
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Z (molecules/ASU)	4	4	4
Ligand identity	Zn^{2+}	Zn^{2+}, KCH	Zn^{2+}, CHA
Total/unique reflections	320,543/87,452	815,126/205,636	379,312/112,743
Multiplicity	3.7 (3.5)	4.0 (3.4)	3.4 (2.5)
Completeness (%)	99.8 (99.1)	98.5 (93.8)	99.4 (96.0)
Mean I/σI	4.9 (1.6)	23.3 (1.5)	10.7 (2.2)
R_{merge}	0.178 (0.862)	0.074 (0.867)	0.184 (0.463)
$CC_{(1/2)}$ (%)	98.3 (55.6)	99.9 (64.9)	99.4 (77.1)
Refinement			
R_{work}/R_{free}	0.2057/0.2702	0.2020/0.2661	0.1895/0.2279
No. of TLS groups	1	1	1
RMSD bonds (Å)	0.009	0.008	0.007
RMSD angles (°)	1.126	0.968	0.897
Ramachandran favored (%)	93.77	92.50	95.25
Ramachandran allowed (%)	5.12	6.70	4.22
Ramachandran outliers (%)	1.11	0.80	0.53
Average B, all atoms ($Å^2$)	31.2	24.5	47.0
Average B, ligands ($Å^2$)	Zn (32.2)	Zn (27.2) KCH (29.6)	Zn (18.8) CHA (28.5)
PDB id	6DWV	6DXS	6DXQ

Table 2. Crystallographic data and refinement statistics.

3.3. Results

3.3.1. Isolation and Purification of LigJ

The open reading frame encoding LigJ was chemically synthesized and subsequently subcloned into a pET-30a (+) vector. The plasmid was expressed in *E. coli* and purified using the C- terminal His tag. The molecular mass of LigJ, based on the elution volume from a calibrated size exclusion column, was 84.8 kDa. The expressed protein consists of 349 amino acids, which includes the His-tag linker, and a calculated monomeric molecular mass of 39.1 kDa (excluding any bound divalent cations). The protein therefore associates as a dimer in solution. On the basis of ICP-MS, the zinc content of the purified enzyme, relative to the subunit concentration, was 0.72 (**Table 3**).

enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})	Zn^{2+} Content
wild-type	25.6 ± 1.6	9.8 ± 3.1	(2.6 ± 0.8) x 10^6	0.72
E284Q	< 0.2	--	--	0.48
Y194F	3.3 ± 0.2	3.2 ± 0.8	(1.0 ± 0.3) x 10^6	0.74
T190A	10.5 ± 0.9	11.9 ± 4.7	(8.8 ± 3.6) x 10^5	0.70
H223N	0.28 ± 0.06	310 ± 110	(8.9 ± 3.6) x 10^2	0.71

Table 3. Kinetic Constants for LigJ and Selected Mutants

3.3.2. Three-Dimensional Structure of LigJ

The three-dimensional crystal structure of WT LigJ was determined with Zn^{2+} in the active site at a resolution of 2.2 Å (PDB entry 6DWV). The dimeric protein has a distorted $(\beta/\alpha)_8$-barrel structural fold in which the single zinc ion binds in the active site at the C-terminal end of the central β-barrel (**Figure 25**). Each subunit consists of 15 helices that surround the central barrel structure, which is composed of eight parallel β-strands (**Figure 26**). The dimer interface of 2252 $Å^2$ is formed from multiple hydrogen bonds and electrostatic interactions from α-helices 5 (A189− Q207), 3′ (G226−V228), 6 (P229−M241), and 7 (Q264− V273) (**Figures 25** and **26**) (33).

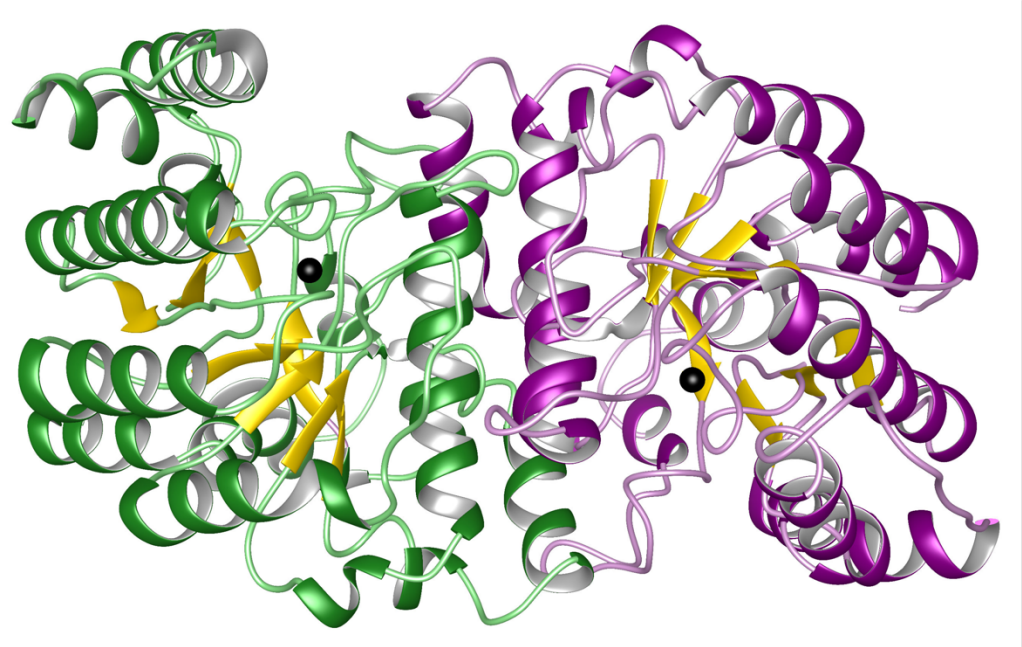


Figure 25. Ribbon structure of the dimeric form of wild-type LigJ (PDB entry 6DWV) in the absence of added ligands. One monomer is depicted with the β-strands colored gold, the α-helices green, and the loops light green. The second monomer is depicted with the β-strands colored gold, the α-helices purple, and the loops pink. The zinc ion in each active site is displayed as a black sphere. This figure was constructed using Chimera (34).

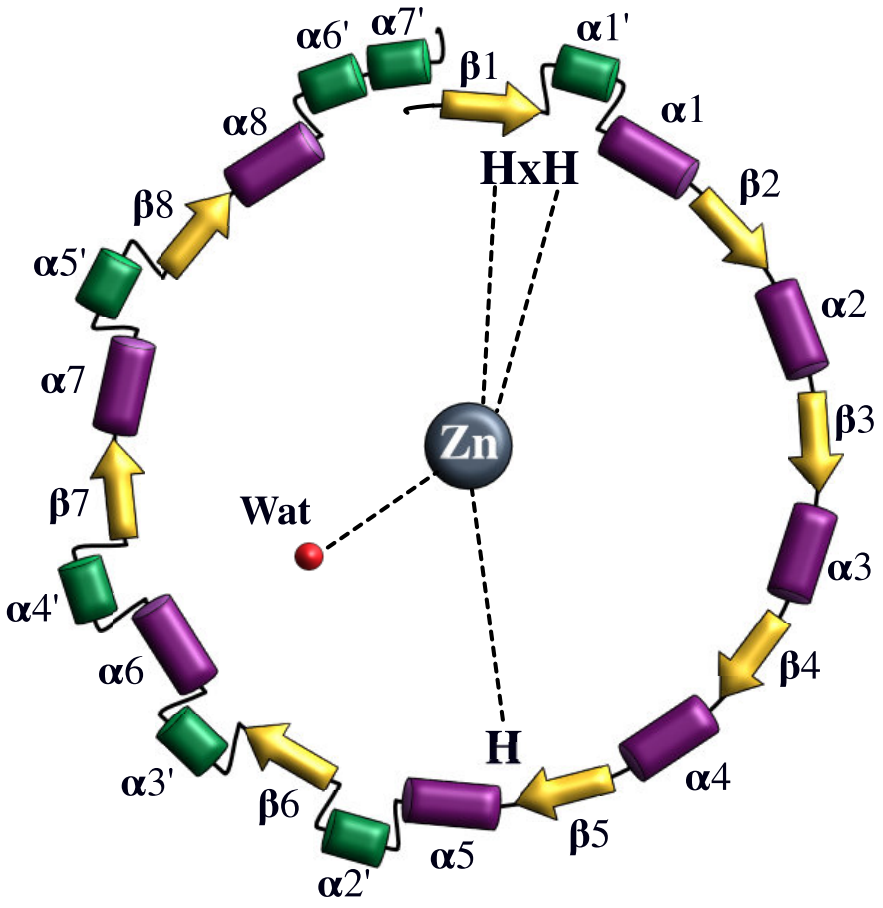


Figure 26. Schematic representation of the α-helices and β-strands that comprise the overall structure of LigJ. The metal that binds at the C- terminal end of the β-barrel is depicted as a gray sphere, and the residues that coordinate the metal ion to the protein are depicted with dashed lines. This figure was made using Affinity Designer 1.6.1 from Serif (Europe) Ltd.

3.3.3. Structure of the Metal Center

Wild-type LigJ coordinates a single zinc ion in the active site by His-8, His- 10, His-178, and one water molecule (**Figure 27**). His-8 and His-10 originate from β-strand 1, and His-178 originates from β-strand 5. The coordination geometry is that of a distorted tetrahedron. Residues Glu-284 from β-strand 8 and His-223 from β-strand 6 also interact with the bound water molecule. Each of the four subunits in the asymmetric unit has modest variations in the distances between the ligand and the metal ion. The distance from the bound water molecule to the zinc varies from 2.0 to 3.3 Å (**Figure 28**). This flexibility in the coordination geometry and structure of the mononuclear metal center in the active site of LigJ is presumably caused by the lone water molecule being displaced from the metal center by the binding of the substrate.

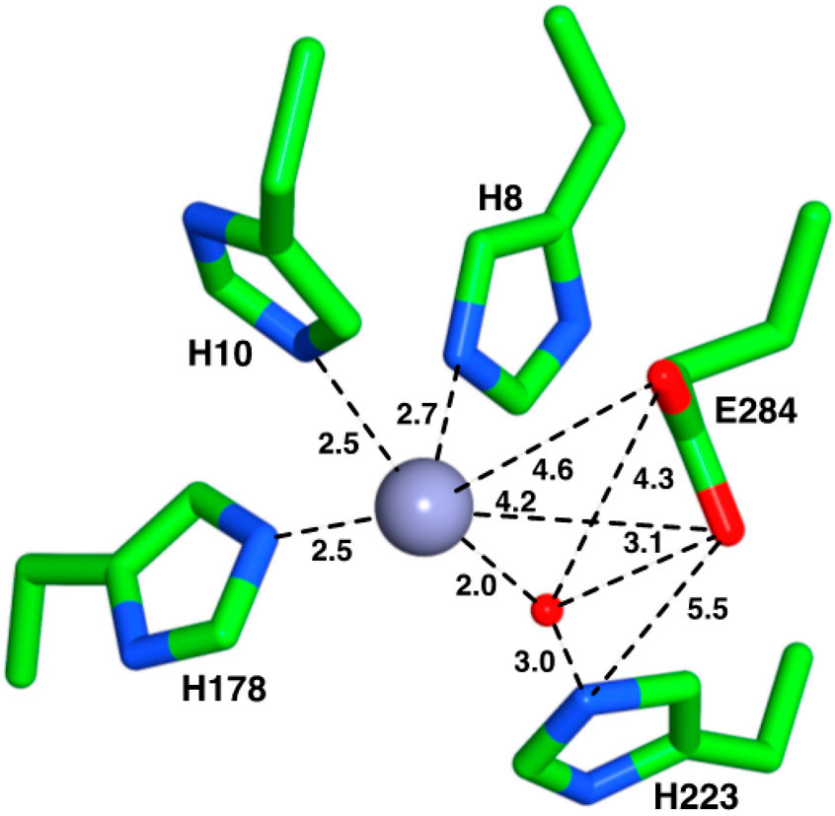


Figure 27. Mononuclear metal center in the active site of LigJ (subunit B in PDB entry 6DWV). The zinc ion is shown as a gray sphere where it is coordinated to His-8, His-10, His-178, and one water molecule in a distorted tetrahedral geometry. His-223 is a conserved residue within cog2159 and is positioned 3.0 Å from the bound water molecule and 5.5 Å from Glu-284. Glu-284 is 4.6 Å from the zinc ion and 3.1 Å from the bound water molecule. There are slight variations in the distances from the coordinating ligands to the zinc ion for the other subunits, which are shown in Figure 29. This figure was made using Chimera (34).

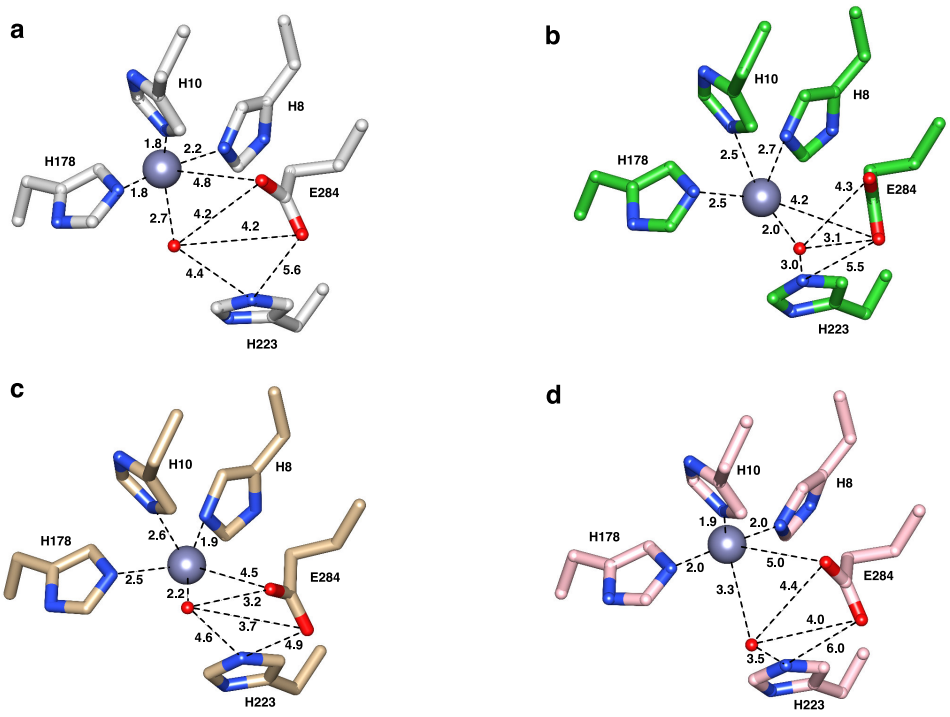


Figure 28. Mononuclear metal center in the active site of LigJ (subunits A through D in PDB id: 6DWV). The zinc ion is shown as a gray sphere where it is coordinated to His-8, His-10, His-178, and one water molecule. (a) LigJ subunit A; (b) LigJ subunit B; (c) LigJ subunit C; (d) LigJ subunit d. Each subunit has slight variations in the distances between the residues near the active site. Subunit B was used in Figure 28 and in Figure 36.

3.3.4. Structure of the Product-Bound Complex

The structure of the product-bound complex was determined by co-crystallization of wild-type LigJ in the presence of added OMA (2), the product from the enzyme-catalyzed hydrolysis of PDC (1). We have shown previously that OMA (2) non-enzymatically isomerizes to KCH (3), the true substrate for LigJ (8). In the product-bound complex of LigJ, CHA (4) is coordinated to the zinc in the active site in a bidentate fashion by the C1 carboxylate and the carbonyl oxygen from C2 (**Figure 29 a−c**). The product in the active site of LigJ has thus displaced the water molecule from the metal center (**Figure 27**). In this structure, CHA is further coordinated to the active site by interactions between the C6 carboxylate and the side chain guanidinium groups of Arg-71 and Arg-290. His-223 also forms an apparent hydrogen bond to the C1 carboxylate group of CHA. The carboxylate group attached to C4 of CHA is hydrogen bonded to Ser-73 and to the phenolic oxygen of Tyr- 194. Arg-234′ (from the adjacent subunit) interacts with the C1 carboxylate of CHA (**Figure 29c**). In the product-bound complex with CHA, the side chain carboxylate of Glu-284 is 2.2 Å from the hydroxyl group attached to C4. The stereochemistry of the enzymatically produced CHA is of the *S* configuration. In this structure, the pro-*S* hydrogen at C3 is 2.8 Å from the carboxylate group of Glu-284, whereas the pro-*R* hydrogen attached to C3 is 3.8 Å from the phenolic oxygen of Tyr-194 (**Figure 29b**). To test the importance of those residues found in the active site of LigJ that interact directly with the substrate, Glu-284, His-223, Tyr- 194, and Thr-190 were substituted to give mutants E284Q, H223N, Y194F, and T190A, respectively.

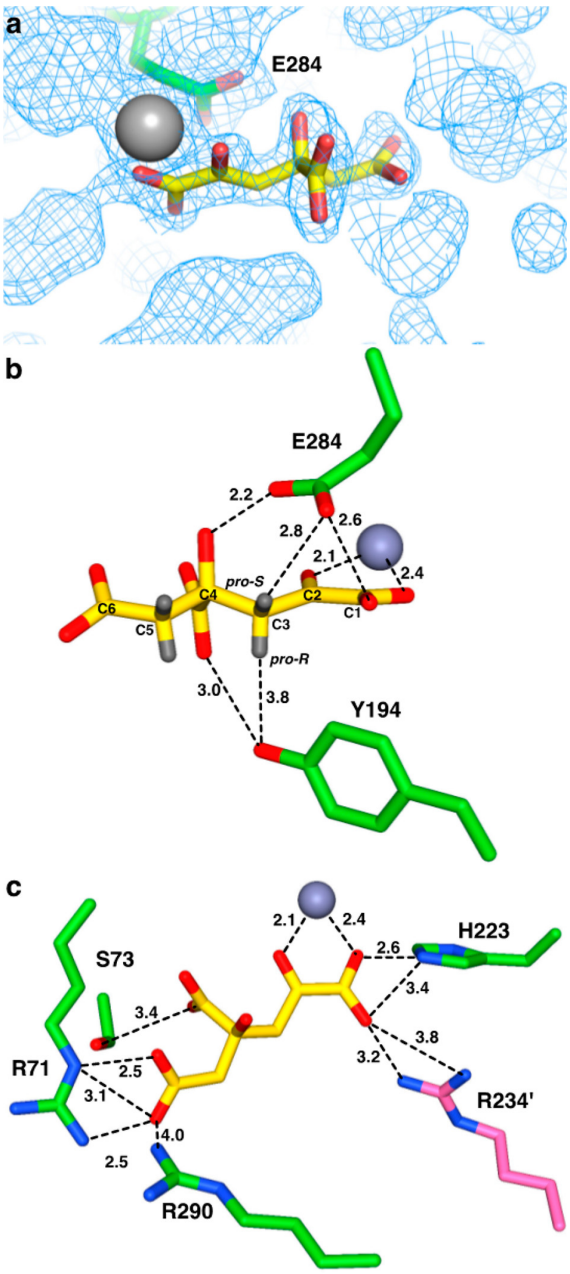


Figure 29. Model of CHA (yellow) bound in the active site of LigJ (PDB entry 6DXQ). The zinc is colored gray, and the associated distances are given in angstroms. For the sake of clarity, the three histidine residues that coordinate the zinc are not shown. (a) Simulated annealing composite omit map ($2F_o - F_c$) contoured at 1.5 σ (subunit A). (b) Orientation of Glu-284 and Tyr-194 relative to the bound substrate (subunit A). (c) Orientation of residues that contribute to the binding of the product in the active site of LigJ. Arg- 234′ is from the adjacent subunit in the dimeric complex (subunit A). This figure was made using Chimera (34).

3.3.5. Kinetic Constants for Hydration of KCH by LigJ

The hydration of KCH (3) catalyzed by LigJ to CHA (4) was monitored spectrophotometrically at pH 8 and 30 °C by following the decrease in absorbance at 265 nm in a reaction mixture that contained KCH formed from the hydrolysis of PDC (1) by LigI and the isomerization of OMA (2) by LigU. The kinetic constants for the wild-type enzyme and the mutants for those residues that were shown to interact with the bound product (E284Q, H223N, Y194F, and T190A) are listed in **Table 3**. The value of k_{cat} for the wild-type enzyme is 26 s^{-1}, and for the T190A and Y194F mutants, it is reduced it by ~2.4- and 7.8-fold, respectively. An 85-fold decrease is observed with H223N, and no activity could be obtained with E284Q. The values for k_{cat}/K_m follow the same trend, where those of mutants Y194F and T190A decrease slightly from that of the wild-type enzyme while that of mutant H223N decreases by >1000-fold. The equilibrium constant for the hydration of KCH to CHA (the [CHA]/[KCH] ratio for LigJ at pH 8.0 was determined to be 3.4). This value was determined by ^{1}H NMR spectroscopy by measuring the relative concentration of CHA (from the integrated areas for the protons at C5 and C3) to KCH (from the integrated areas for the protons at C5 and C3) (**Figure 30** and **Table 4**). The zinc content changed only for mutant E284Q relative to that of the wild type (**Table 3**). The activity of E284Q was measured in the presence of added $ZnCl_2$ (0−1 mM), but no enzymatic activity could be observed.

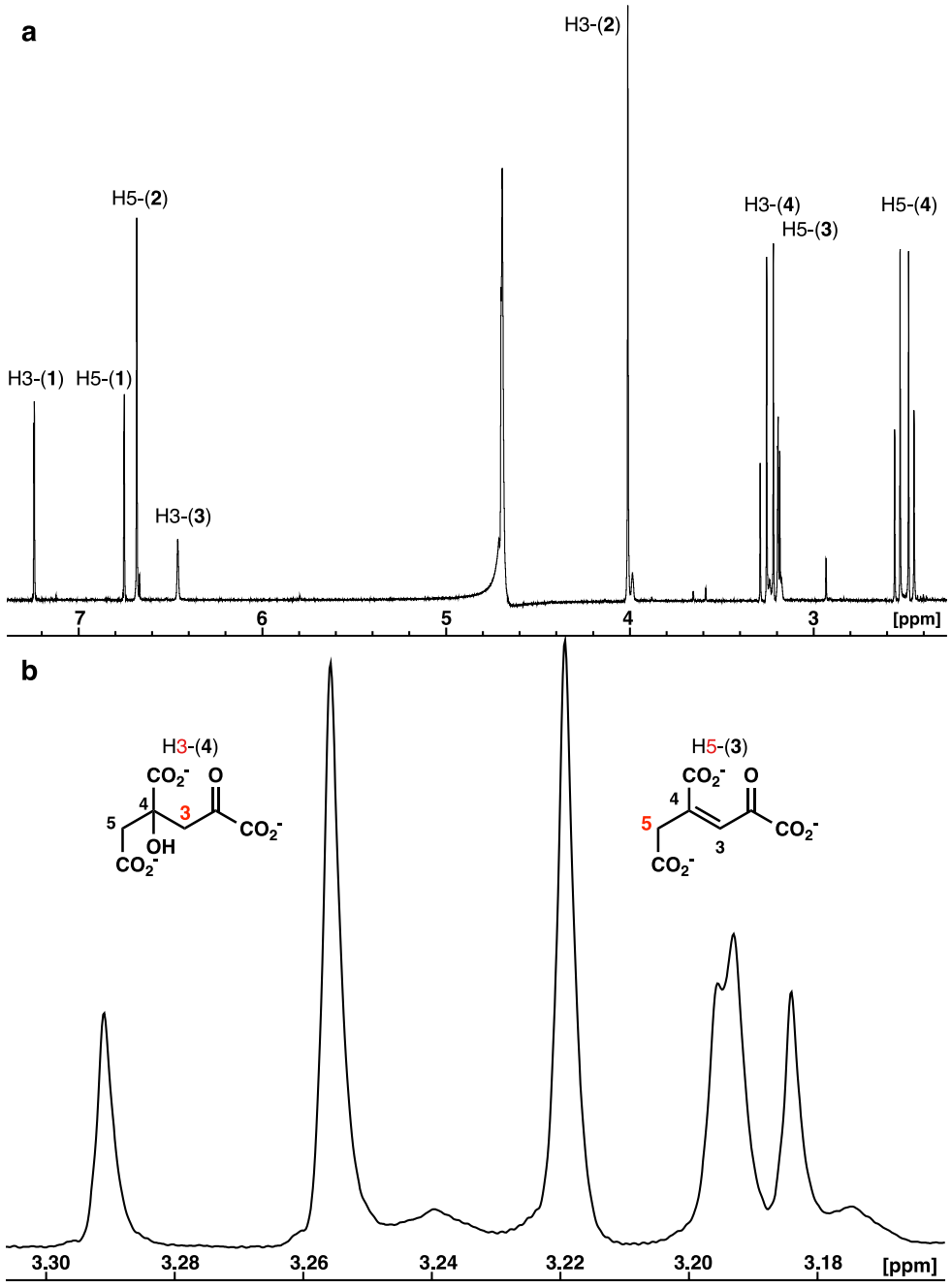


Figure 30. (a) ^{1}H NMR spectrum of the equilibrium mixture of PDC (1), OMA (2), KCH (3), and CHA (4) in 10% D_2O at pH 8.0 after the addition of 0.5 µM LigI hydrolase, LigU isomerase, and LigJ hydratase to 5 mM PDC. (b) ^{1}H NMR spectrum of (a) focused on the doublet of doublets centered at 3.24 ppm that are resonances for the two protons at C3 for CHA, and the doublet centered at 3.19 ppm is for the two protons at C5 for KCH. The equilibrium constant ([CHA]/[KCH]) is 3.4.

Chemical shift (ppm)	**Integration Area**	**Chemical**	**Proton**
7.24	0.8190	PDC	H3
6.75	0.8391	PDC	H5
6.68	1.6111	OMA	H5
6.46	0.5028	KCH	H3
4.69	-----	water suppression	--
4.01	2.7474	OMA	H3
3.24	3.2772	CHA	H3
3.19	1.0000	KCH	H5
2.93	0.1352	unknown	--
2.50	3.5712	CHA	H5

Table 4. Integration of the resonances from the ^{1}H NMR spectrum in Figure 31 for PDC, OMA, CHA, and KCH catalyzed by LigI, LigU, and LigJ at equilibrium at pH 8.0.

3.3.6. Structure of the Substrate-Bound Complex

The E284Q mutant was shown to be catalytically inactive for the hydration of KCH to OMA. This mutant was therefore chosen in an attempt to trap the substrate in the active site of LigJ. The mutant was successfully crystallized, and the substrate KCH was found in the active site (**Figure 31**). Similar to the molecular interactions exhibited by the product-bound complex (**Figure 29**), the carbonyl oxygen at C2 and the C1 carboxylate are ligated to the zinc bound in the active site at distances of 2.8 and 2.2 Å, respectively. In addition, the ionic interactions between the C6 carboxylate and the side chain guanidino groups of Arg-71 and Arg-290 are also preserved. The imidazole side chain of His-223 interacts with the C1 carboxylate of the bound KCH, which is similar to its interaction with the product OMA (**Figures 29c** and **31c**). In this structure, there is a single water molecule that is positioned for nucleophilic attack at C4 of the substrate KCH. This water molecule is 2.7 Å from the *re* face of C4, consistent with the stereochemistry for the product-bound complex (**Figure 31b**). This water molecule is hydrogen bonded to the side chain amide from the glutamine residue that has replaced Glu-284 in the mutant enzyme. The positioning of this interaction in the substrate-bound complex (**Figure 31**) is consistent with a proton abstraction of the water molecule by Glu-284 for nucleophilic attack at C4 of the substrate. In addition, the positioning of Glu-284 in the product-bound complex is consistent with the delivery of this proton to C3 during the hydration reaction with an overall *syn* addition of water to the double bond between C4 and C3 of the substrate KCH.

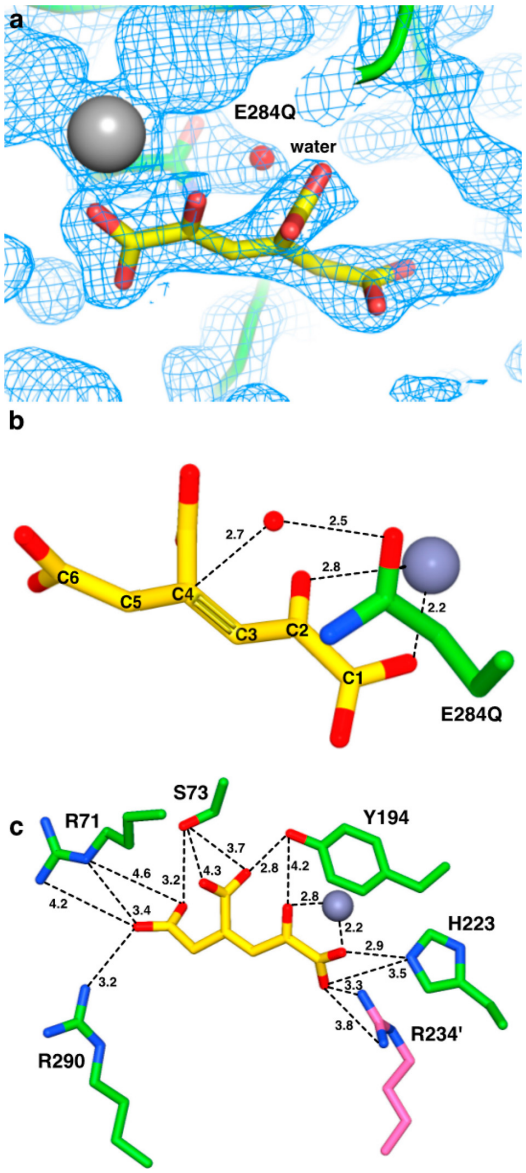


Figure 31. Model of KCH in the active site of the E284Q mutant of LigJ (subunit B of PDB entry 6DXS). The zinc is colored gray, and the associated distances are given in angstroms. For the sake of clarity, the three histidine residues that coordinate the zinc are not shown. (a) Simulated annealing composite omit map ($2F_o - F_c$) contoured at 1.0σ. (b) Orientation of Gln-284 and the bound water molecule relative to the orientation of C4 from KCH. Orientation of residues that contribute to the binding of the substrate in the active site of LigJ. Arg-234′ is from the adjacent subunit in the dimeric complex. This figure was made using Chimera and Affinity Designer 1.6.1 from Serif (Europe) Ltd (34).

3.4. Discussion

3.4.1. Role of LigJ in the PCA 4,5-Cleavage Pathway

The role of LigJ as a hydratase in the bacterial degradation of lignin has been known for more than 35 years (9). LigJ has since been identified in a number of other bacterial species with a ≥60% sequence identity with the enzyme from *Sphingobium* sp. SYK-6 (35). This enzyme catalyzes the penultimate step in the PCA 4,5-cleavage pathway before the product of this reaction is ultimately cleaved by LigK to pyruvate and oxaloacetate as precursors to the TCA cycle. Unlike the isomerization of OMA to KCH by LigU, which can occur non-enzymatically (8), the conversion of KCH to CHA does not occur in the absence of LigJ. This was shown by disruption of the *ligJ* gene in *Sphingomonas paucimobilis* sp. SYK-6 and incubation of whole cells of the mutant strain with either vanillate or syringate (11). The metabolites that were produced showed an accumulation of OMA, PDC, and an unknown product that is likely to be KCH (8, 11).

3.4.2. Structure and Reaction Mechanism of LigJ

The three-dimensional crystal structure of LigJ was determined in the presence and absence of the substrate (KCH) and product (CHA). The protein folds as a distorted $(\beta/\alpha)_8$-barrel with a single divalent cation perched at the C-terminal end of the β-barrel. The zinc in the active site is coordinated by two histidines from the end of β-strand 1 and

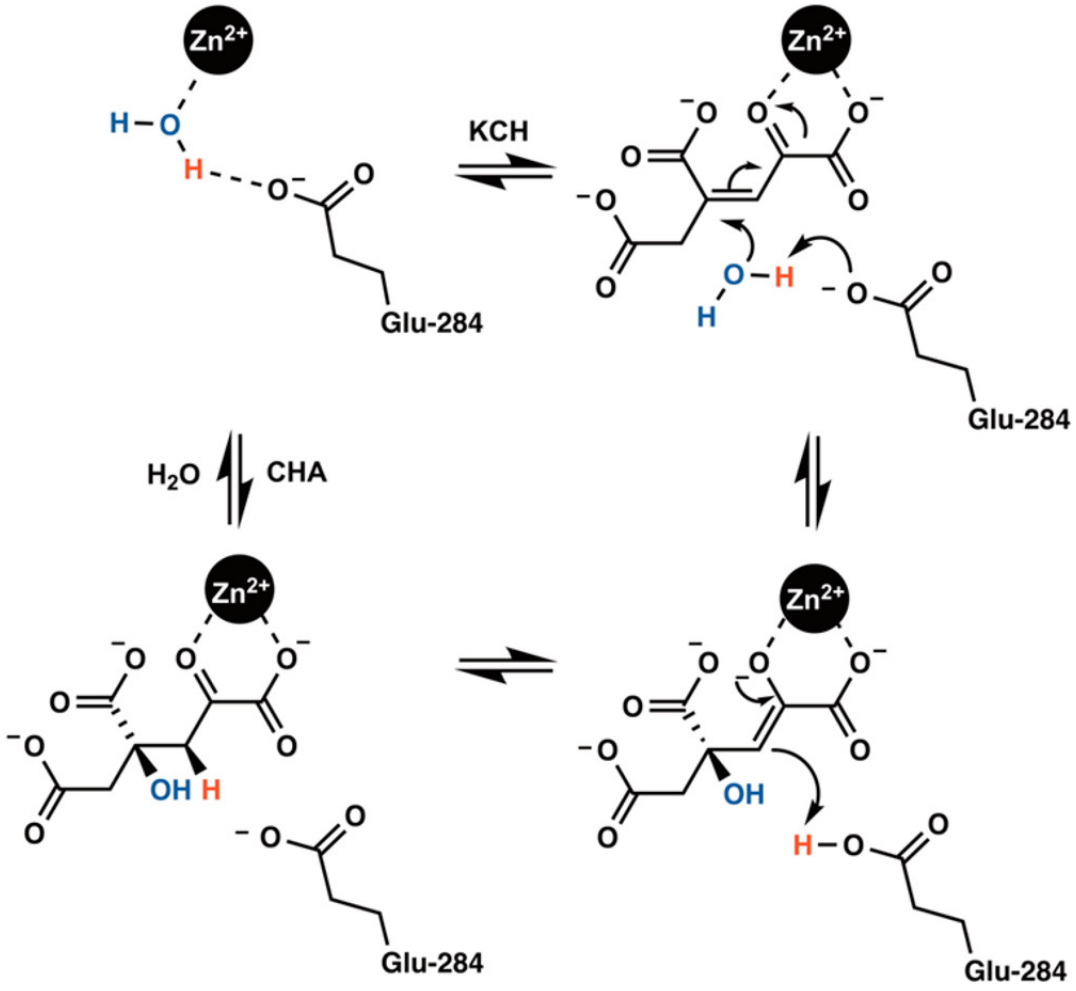
another histidine from the end of β-strand 5. A single water molecule completes the coordination scaffold in an overall tetrahedral geometry.

In the product-bound complex, the C1 carboxylate and the carbonyl oxygen attached to C2 of CHA have displaced the lone water molecule from the coordination shell of the zinc in the active site of the apo-enzyme. The coordination changes from four to five ligands now giving a distorted trigonal bipyramidal geometry. The product further interacts with the protein by ionic interactions of the guanidino groups of Arg-71 and Arg-290 with the C6 carboxylate of CHA. The C7 carboxylate hydrogen bonds with Ser-73, and the C1 carboxylate interacts with Arg-234′ from the adjacent subunit and His-223 from the end of β-strand 6. The structure of the product-bound complex has revealed that the hydroxyl group is added to C4 in the *S* configuration. The absolute configuration of the enzymatically formed CHA has not previously been determined, but the relative stereochemistry is consistent with the previously reported determination of the hydration of 4-carboxy-2-hydroxymuconate (CHM) by the enzyme GalB from *Pseudomonas putida* KT2440, the sequence of which is 12% identical to the LigJ hydratase from *Sphingobium* sp. SYK-6 (36).

The product-bound complex of LigJ with CHA illustrates the hydrogen bond interaction of Glu-284 with the hydroxyl group attached to C4. This structure suggests that Glu-284 functions as a general base in the activation of the water molecule for nucleophilic attack. In the reverse reaction, where CHA is dehydrated to KCH, a proton must be abstracted from C3. In the product-bound complex of CHA, the side chain carboxylate of Glu-284 is 2.8 Å from the pro-*S* hydrogen and the phenolic oxygen of

Tyr-194 is 3.8 Å from the pro-*R* hydrogen attached to C3. If the pro-*S* hydrogen is abstracted, then the addition of water to KCH is consistent with a *syn* addition, whereas if the pro-*R* hydrogen is abstracted, then the addition is *anti*. Mutation of Tyr-194 and Glu-284 clearly demonstrates the absolute importance of Glu-284 and the relative unimportance of Tyr-194. In the product-bound complex, H223 is not positioned to interact with either the hydroxyl group at C4 or either of the two hydrogens attached to C3. However, when H223 is mutated to an asparagine, there is a 100-fold decrease in k_{cat} and a 30-fold increase in K_m for the substrate KCH. These changes are most likely due to the loss of the ionic interaction of H233 with the C1 carboxylate group of the substrate and product (**Figures 29c** and **31c**). These results are fully consistent with the conclusion that in the forward reaction Glu-284 functions as a general base in the abstraction of a proton from water, concurrent with the nucleophilic attack at C4 of the substrate. The nucleophilic attack is facilitated by delocalization of the negative charge to the carbonyl oxygen that interacts with the zinc in the active site. In the final step of the reaction, C3 is protonated by the protonated form of Glu- 284, now acting as a general acid. The proposed mechanism is presented in **Scheme 10**.

Scheme 10. The proposed reaction mechanism for LigJ hydratase.



The proposed reaction mechanism for LigJ is fully supported by the crystal structure of LigJ determined in the presence of the substrate KCH. In this structure, the closest water molecule to C4 is hydrogen bonded to the side chain carboxylate of Glu-284. The positioning of this water molecule and the orientation of Glu-284 are fully consistent with the *syn* addition of water to C4/C3 and the role of Glu-284 as both the general acid and the general base in this transformation.

The proposed reaction mechanism for LigJ, with a *syn* addition of water and the utilization of a single catalytic residue that functions as both the general acid and the general base, has been previously proposed for the reaction mechanism of (3*S*)-methylglutaconyl-CoA hydratase (MGCH) from *Pseudomonas putida*, where (*E*)-3-methylglutaconyl-CoA is converted to (3*S*)-hydroxymethyl-glutaryl-CoA (37). This enzyme is a member of the crotonase superfamily, where related hydratases often require two glutamate residues to facilitate the overall hydration reaction. However, substrates that have a carbonyl group conjugated to the double bond to be hydrated tend to function with *syn* addition, whereas those substrates where the double bond is conjugated to a carboxylate group, such as fumarase and aconitase, tend to occur with an *anti* addition of water (38).

3.4.3. The Amidohydrolase Superfamily and cog2159

LigJ belongs to cog2159 and is a member of the AHS of enzymes. Most of the functionally characterized enzymes contained within the amidohydrolase superfamily catalyze the hydrolysis of phosphate and carboxylate esters and amides (17, 39).

Notable enzymes in this superfamily include phosphotriesterase (40), dihydroorotase (41), urease (42), and cytosine deaminase (43). In general, most of the structurally characterized enzymes in the AHS contain one or two divalent cations in the active site, and the mononuclear and binuclear metal centers are utilized to facilitate the activation of the bound water/hydroxide for nucleophilic attack on the substrate (17). However, some enzyme examples in this superfamily have been shown to contain three divalent cations in the active site such as that found in histidinol phosphate phosphatase from cog1387 (44). Another outlier in this regard is LigI from cog3618, an enzyme that catalyzes the hydrolysis of PDC (1) during the bacterial degradation of lignin (7). This enzyme does not bind divalent cations, but it uses the conserved aspartate from the end of β-strand 8 to activate the water molecule for nucleophilic attack. The conserved histidine residues from the ends of β-strands 1 and 5 function as Lewis acids for the activation of the lactone substrate (7).

Although most of the enzymes contained within the AHS are hydrolases, those enzymes of cog2159 catalyze a much more diverse set of reactions. The sequence similarity network (SSN) for cog2159 at an E value cutoff of 10^{-65} illustrates the sequence relationships among the 1299 nonredundant protein sequences (**Figure 32**). A total of 14 of these proteins have been experimentally characterized or have a three-dimensional crystal structure (**Table 5**). In the case presented here, LigJ catalyzes the hydration of a conjugated double bond. Moreover, as we have now shown, the mononuclear metal center contained within the active site of LigJ is not used to activate the water molecule. It is, however, utilized to activate the substrate via polarization of

the carbonyl oxygen. In another example from cog2159, LigW (and other closely related decarboxylation enzymes) catalyzes the decarboxylation of *o*-hydroxy benzoate substrates (**Scheme 11**) (19). LigW contains a single divalent cation in the active site, and the substrate is activated by coordination of the carboxylate group and the adjacent phenolic oxygen by the single divalent cation (**Figures 33** and **35**). Distortion of the substrate facilitates the protonation of the carbon bearing the carboxylate group by the active site aspartate residue that originates from the end of β-strand 8. Water is not involved in any phase of the reaction. Other characterized enzymes within cog2159 that also catalyze decarboxylation reactions to similar aromatic structures decorated with carboxylate and hydroxyl substituents include γ-resorcylate decarboxylase (γ-RSD) and 6-methylsalicylate decarboxylase (MSD) (45,46). The identification of 5-carboxy-uracil decarboxylase (IDC) from the fungal species *Cordyceps militaris* and *Metarhizium anisopliae* greatly diversifies the reactions of cog2159 to include a reaction step that is important for DNA methylation (47). The sequence of the LigJ hydratase is 24% identical to that of α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD) (48−50).

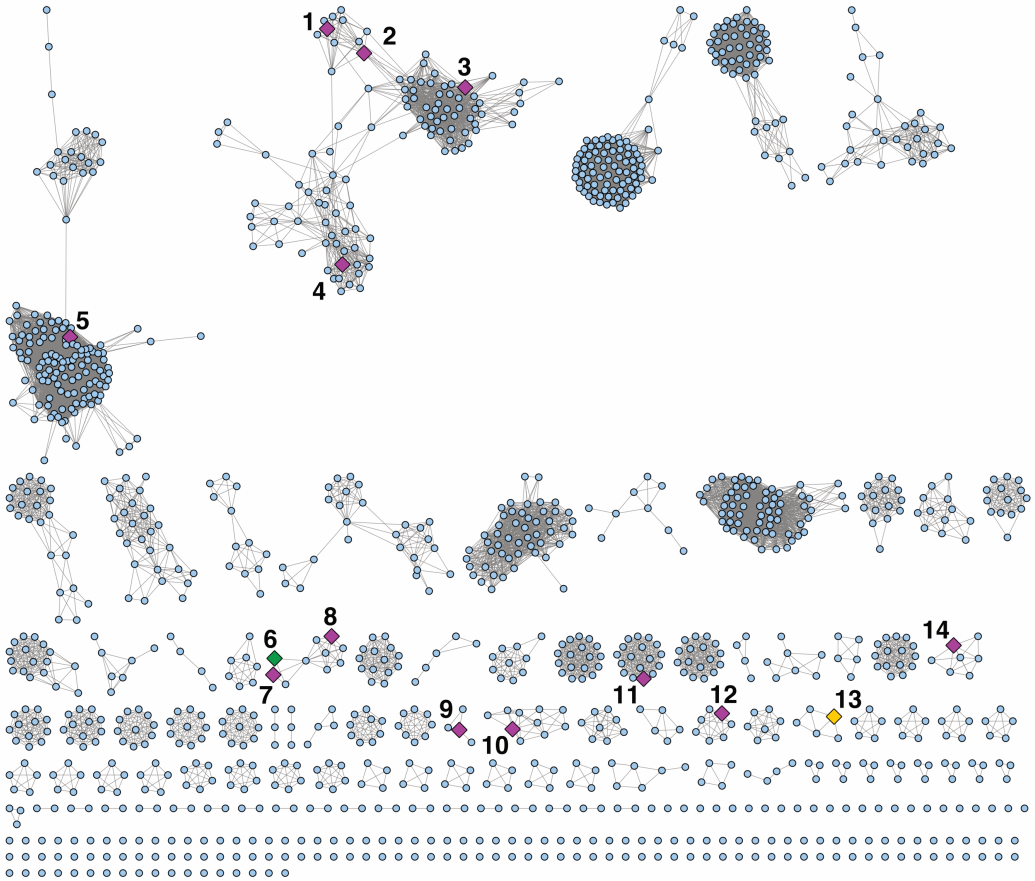
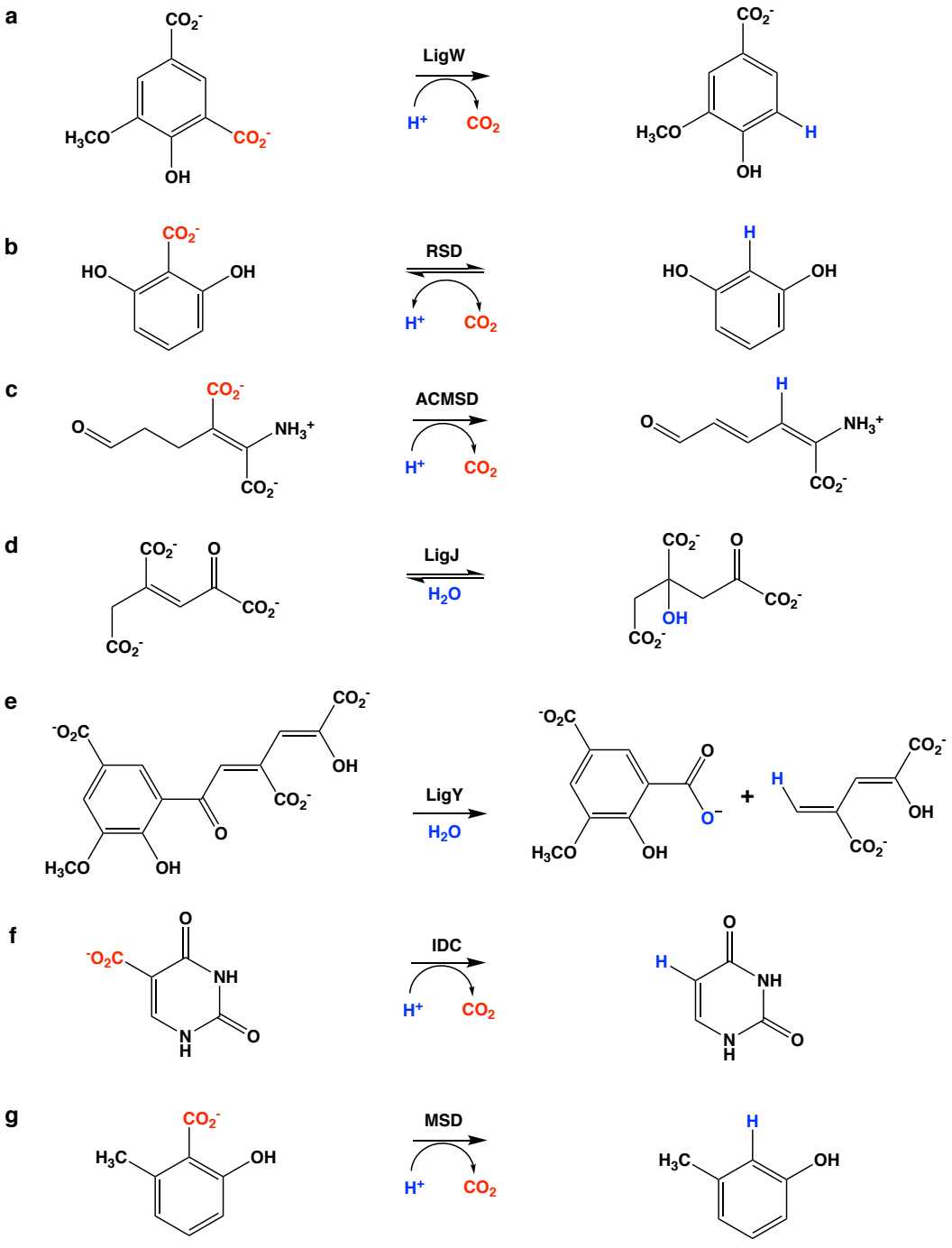


Figure 32. Cytoscape-generated sequence similarity network (SSN) of cog2159 at a BLAST E value cutoff of 10^{-65} contains 1,299 nonredundant protein sequences. The magenta diamonds represent enzymes with crystal structures deposited in the PDB. The gold diamond represents 6-methylsalicylic decarboxylase (MSD). The green diamond represents LigJ from this study. The description of each numbered node is presented in Table 5.

SSN ID	Name	Function	Organism	PDB id	Uniprot ID
1	LigW2	5-carboxyvanillate decarboxylase	*Novosphingobium aromaticivorans* (strain ATCC 700278 / DSM 12444 / CIP105152 / NBRC 16084 / F199)	4INF, 4QRN, 4QS5, 4QS6, 4QTG	Q2GA79
2	LigW	5-carboxyvanillate decarboxylase	*Sphingobium* sp. SYK-6.	4ICM	G2IN12
3	γRSD	gamma-resorcylate decarboxylase	*Polaromonas* sp. (strain JS666 / ATCC BAA-500)	3S4T, 4QRO	Q12BV1
4		uncharacterized protein	*Staphylococcus aureus* (strain Mu50 / ATCC 700699)	3NUR	A0A0H3K164
5	ACMSD	2-amino-3-carboxymuconate-6-semialdehyde decarboxylase	*Homo sapiens* *Pseudomonas fluorescens*	2WM1, 4IGM, 4IGN, 4IH3, 4OFC, 2HBV	Q8TDX5 Q83V25
6	LigJ	(3*Z*)-2-keto-4-carboxy-3-hexenedioate hydratase	*Sphingobium* sp. SYK-6.	This study	Q9KWL6
7	LigJ	annotated 4-oxalomesaconate hydratase	*Rhodopseudomonas palustris* (strain ATCC BAA-98 / CGA009)	2GWG	Q6N0R4
8	LigY	4,11-dicarboxy-8-hydroxy-9-meth- oxy-2-hydroxy-6-oxo-6-phenyl-hexa-2,4-dienoate hydrolase	*Sphingobium* sp. SYK-6.	5VN5	G2IN02
9		metal-dependent hydrolase	*Lactobacillus plantarum* (strain ATCC BAA-793 / NCIMB 8826 / WCFS1)	2F6K	F9US96
10		predicted amidohodrolase (Dihydroorotase family)	*Clostridium acetobutylicum* (strain ATCC 824 / DSM 792 / JCM 1419 / LMG5710 / VKM B-1787)	3CJP	Q97DY3
11	IDC	5-carboxyl-uracil decarboxylase	*Cordyceps militaris* (strain CM01) *Metarhizium robertsii* (strain ARSEF 23 / ATCC MYA-3075)	4HK5, 4HK6, 4HK7, 4LAN, 4LAO, 4LAK, 4LAM, 4LAL 4HJW	G3J531 E9F0X0
12		putative 4-oxalomesaconate hydratase	*Lactobacillus acidophilus* (strain ATCC 700396 / NCK56 / N2 / NCFM)	3IJ6	Q5FK79
13	MSD	6-methylsalicylic decarboxylase	*Aspergillus clavatus*		T1PRE6
14		uncharacterized protein	*Bordetella bronchiseptica* (strain ATCC BAA-588 / NCTC 13252 / RB50)	3IRS, 3K4W	A0A0H3M034

Table 5. Characterized Enzymes for Sequence Similarity Network of cog2159.

Scheme 11. Reactions Catalyzed by Enzymes in cog2159.



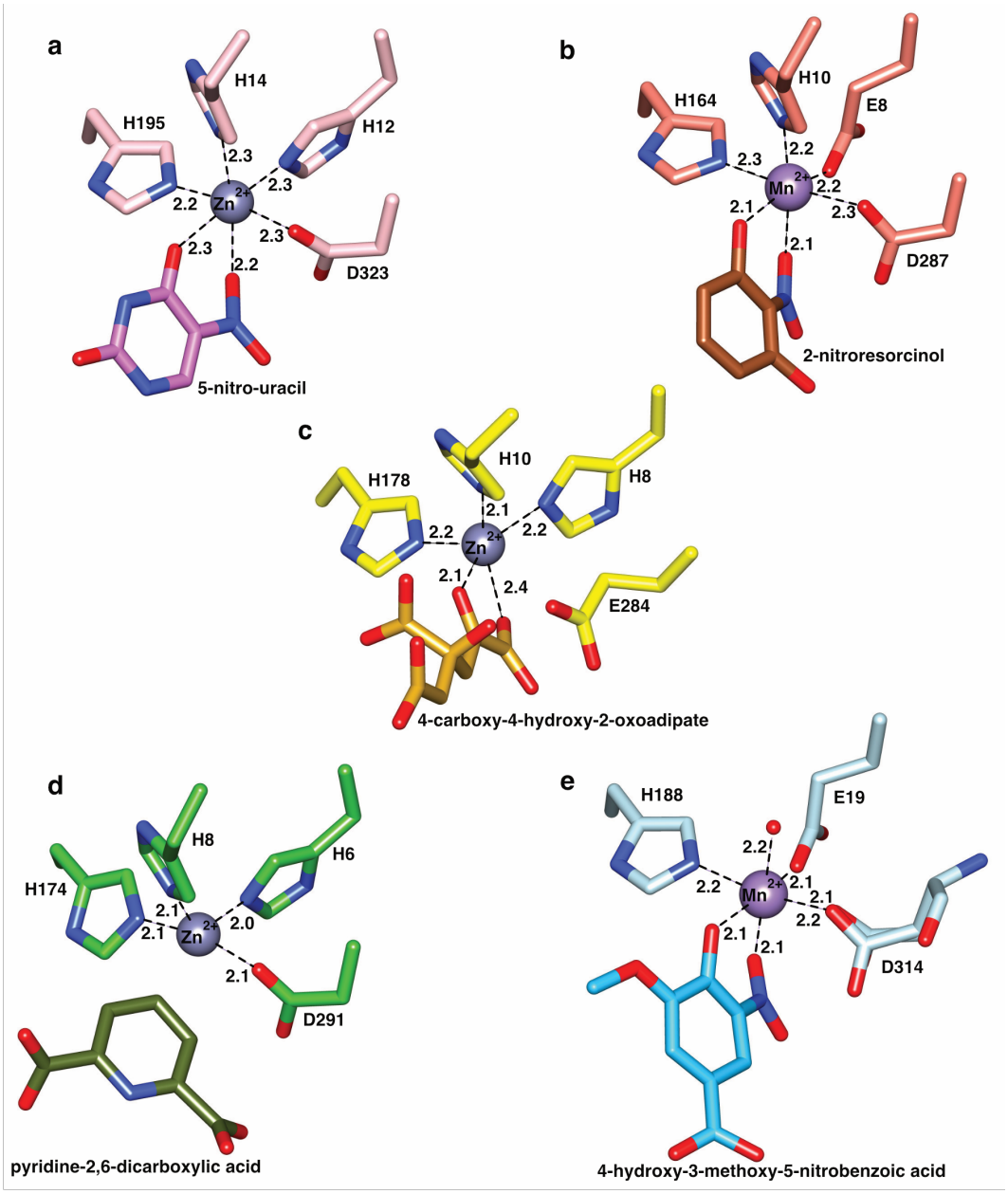


Figure 33. Active site coordination with ligands bound in the active site from enzymes of cog2159. (a) IDC decarboxylase, subunit A (PDB entry 4HK6); (b) γ-RSD decarboxylase, subunit D (PDB entry 4QRO) (c) LigJ hydratase, subunit D (PDB entry 6DXQ); (d) ACMSD decarboxylase, subunit A (PDB entry 4IH3), and (e) LigW decarboxylase, subunit B (PDB entry 4QRN).

Of particular interest to LigJ is the reaction catalyzed by LigY. LigY is also a member of cog2159 and catalyzes the hydrolysis of 4,11-dicarboxy-8-hydroxy-9-methoxy-2-hydroxy- 6-oxo-6-phenylhexa-2,4-dienoate (DCHM-HOPDA) to 5-carboxyvanillate (5CVA) and 4-carboxy-2-hydroxypenta-2,4-dienoate (CHPD) as illustrated in **Scheme 11** (23). The sequence of the enzyme from *Sphingobium* sp. SYK-6 is 37% identical to that of LigJ (**Figure 34**). The structure of LigY from *Sphingobium* sp. SYK-6 has been determined in the absence of a bound substrate or product (PDB entry 5VN5). The enzyme binds a single divalent cation in a fashion quite similar to that of LigJ (**Figure 35**). It would be of interest to determine the structure of LigY in the presence of substrates or products and to evaluate the evolutionary relationship of the catalytic reaction mechanisms for the two enzymes that are essential for the bacterial degradation of lignin to common metabolites.

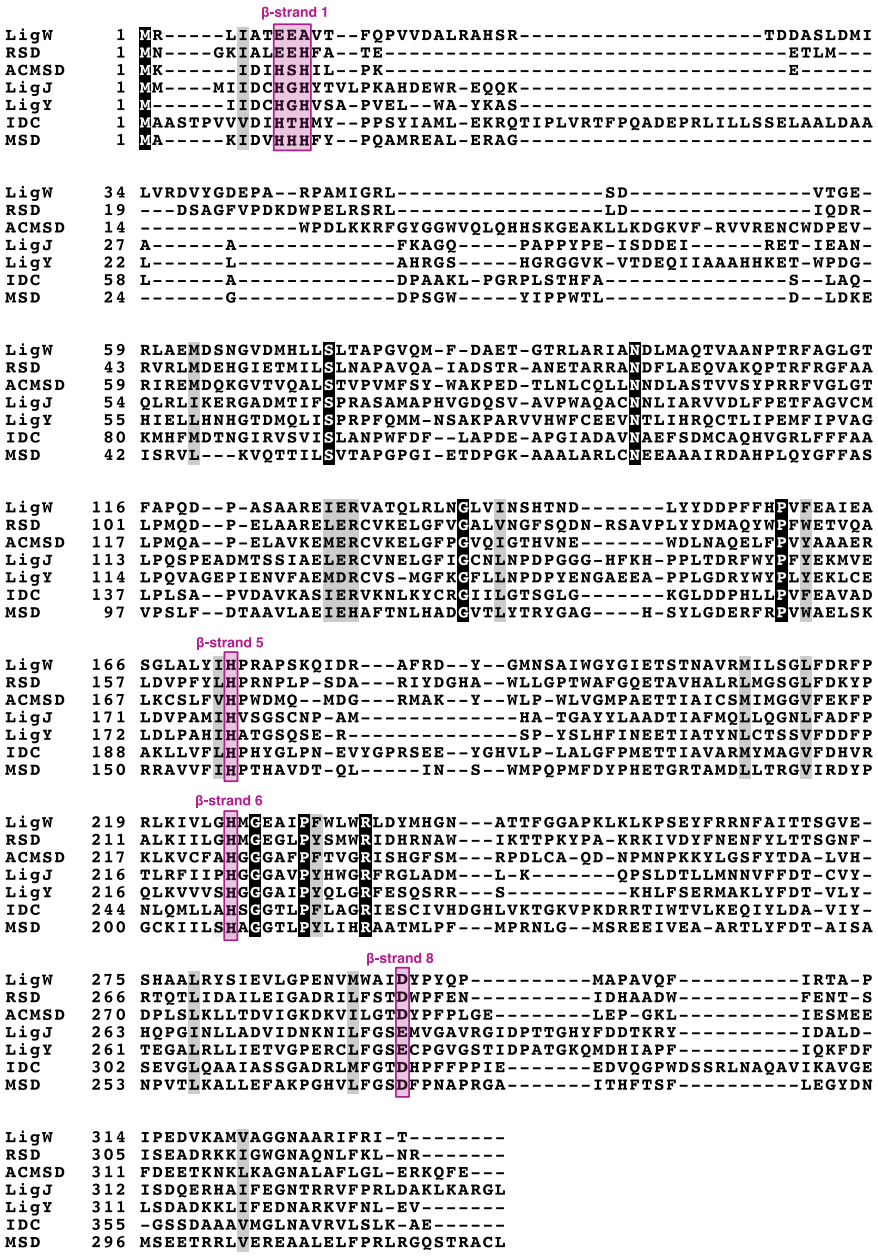


Figure 34. Sequence Alignment of LigJ (UniProt: Q9KWl6) to other characterized members of cog2159 have the following sequence identity to LigJ: LigW (UniProt: Q2GA79) is 18%; RSD (UniProt: Q12BV1) is 23%; ACMSD (UniProt: Q8TDX5) is 22%; LigY (UniProt: G2IN02) is 37%; IDC (Uniprot: G3J531) is 15%; and MSD (UniProt: T1PRE6) is 20%. These sequence identity percentages were calculated using Clustal 1.2.4 (https://www.ebi.ac.uk/Tools/msa/clustalo/). The sequence alignment was created using ExPASy BOXSHADE version 3.21, https://embnet.vital-it.ch/software/BOX_form.html.

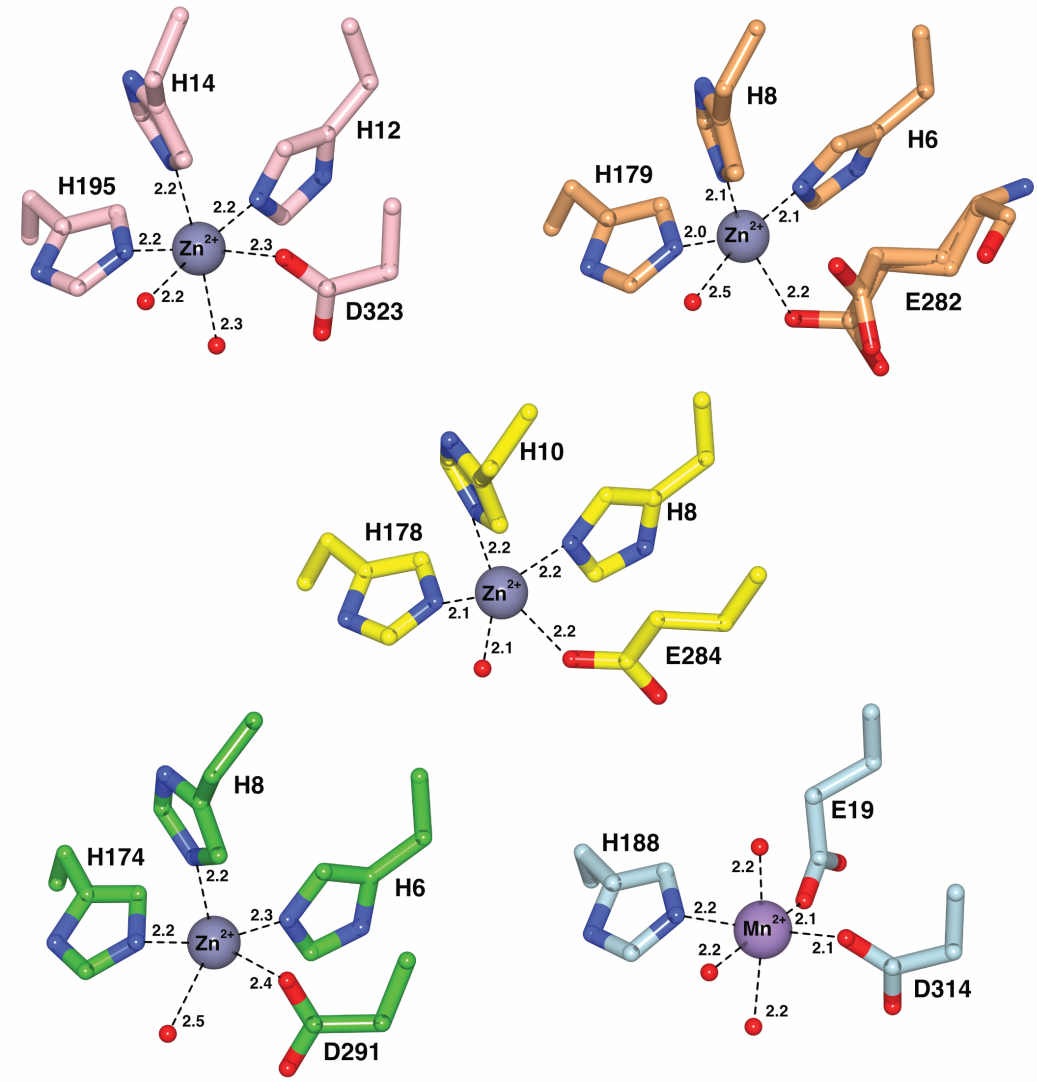


Figure 35. Active site coordination of metal ions with no other ligands bound in the active site from members of cog2159. (a) IDC decarboxylase, subunit C (PDB entry: 4HK5); (b) LigY hydrolase, subunit A (PDB entry: 5VN5); (c) LigJ hydratase, subunit B (PDB entry 6DXQ); (d) ACMSD decarboxylase, subunit A (PDB entry: 4OFC), and (e) LigW decarboxylase, subunit A (PDB entry: 4QTG).

3.5. References

1. Brown, M. E., and Chang, M. C. Y. (2014) Exploring bacterial lignin degradation. *Curr. Opin. Chem. Biol. 19,* 1−7.
2. de Gonzalo, G., Colpa, D. I., Habib, M. H. M., and Fraaije, M. W. (2016) Bacterial enzymes involved in lignin degradation. *J. Biotechnol. 236,* 110−119.
3. Nishikawa, S., Sonoki, T., Kasahara, T., Obi, T., Kubota, S., Kawai, S., Morohoshi, N., and Katayama, Y. (1998) Cloning and Sequencing of the *Sphingomonas (Pseudomonas) paucimobilis* Gene Essential for the O Demethylation of Vanillate and Syringate. *Appl. Environ. Microbiol. 64,* 836−842.
4. Kamimura, N., Takahashi, K., Mori, K., Araki, T., Fujita, M., Higuchi, Y., and Masai, E. (2017) Bacterial catabolism of lignin-derived aromatics: New findings in a recent decade. *Environ. Microbiol. Rep. 9,* 679−705.
5. Noda, Y., Nishikawa, S., Shiozuka, K. I., Kadokura, H., Nakajima, H., Yoda, K., Katayama, Y., Morohoshi, N., Haraguchi, T., and Yamasaki, M. (1990) Molecular Cloning of the Protocatechuate 4,5- Dioxygenase Genes of *Pseudomonas paucimobilis*. *J. Bacteriol. 172,* 2704−2709.
6. Masai, E., Momose, K., Hara, H., Nishikawa, S., Katayama, Y., and Fukuda, M. (2000) Genetic and Biochemical Characterization of 4-Carboxy-2-Hydroxymuconate-6-Semialdehyde Dehydrogenase and Its Role in the Protocatechuate 4,5-Cleavage Pathway in *Sphingomonas paucimobilis* SYK-6. *J. Bacteriol. 182,* 6651−6658.

7. Hobbs, M. E., Malashkevich, V., Williams, H. J., Xu, C., Sauder, J. M., Burley, S. K., Almo, S. C., and Raushel, F. M. (2012) Structure and Catalytic Mechanism of LigI: Insight into the Amidohydrolase Enzymes of cog3618 and Lignin Degradation. *Biochemistry 51,* 3497− 3507.
8. Hogancamp, T. N., and Raushel, F. M. (2018) Functional Annotation of LigU as a 1,3-Allylic Isomerase during the Degradation of Lignin in the Protocatechuate 4,5-Cleavage Pathway from the Soil Bacterium *Sphingobium* sp. SYK-6. *Biochemistry 57,* 2837−2845.
9. Maruyama, K. (1983) Enzymes Responsible for Degradation of 4-Oxalmesaconic Acid in *Pseudomonas ochraceae*. *J. Biochem. 93,* 567− 574.
10. Maruyama, K. (1985) Purification and Properties of γ- Oxalomesaconate Hydratase from *Pseudomonas ochraceae* Grown with Phthalate. *Biochem. Biophys. Res. Commun. 128,* 271−277.
11. Hara, H., Masai, E., Katayama, Y., and Fukuda, M. (2000) The 4-Oxalomesaconate Hydratase Gene, Involved in the Protocatechuate 4,5-Cleavage Pathway, Is Essential to Vanillate and Syringate Degradation in *Sphingomonas paucimobilis* SYK-6. *J. Bacteriol. 182,* 6950−6957.
12. Maruyama, K. (1990) Purification and Properties of 4- Hydroxy-4-Methyl-2-Oxoglutarate Aldolase from *Pseudomonas ochraceae* Grown on Phthalate. *J. Biochem. 108,* 327−333.
13. Hara, H., Masai, E., Miyauchi, K., Katayama, Y., and Fukuda, M. (2003) Charaterization of the 4-Carboxy-4-Hydroxy-2-Oxoadipate Aldolase Gene and

Operon Structure of the Protocatechuate 4,5- Cleavage Pathway Genes in *Sphingomonas paucimobilis* SYK-6. *J. Bacteriol. 185,* 41−50.

14. Wang, W., Mazurkewich, S., Kimber, M. S., and Seah, S. Y. K. (2010) Structural and Kinetic Characterization of 4-Hydroxy-4-methyl-2-oxoglutarate/4-Carboxy-4-hydroxy-2-oxoadipate Aldolase, a Protocatechuate Degradation Enzyme Evolutionarily Convergent with the HpaI and DmpG Pyruvate Aldolases. *J. Biol. Chem. 285,* 36608− 36615.
15. Liu, A., and Huo, L. (2014) Amidohydrolase Superfamily. In *Encyclopedia of Life Sciences*, John Wiley & Sons, Ltd., Chichester, U.K. DOI: 10.1002/9780470015902.a0020546.pub2
16. Raushel, F. M. (2016) Finding homes for orphan enzymes. *Perspectives in Science. 9,* 3−7.
17. Seibert, C., and Raushel, F. M. (2005) Structural and Catalytic Diversity within the Amidohydrolase Superfamily. *Biochemistry 44,* 6383−6391.
18. Cummings, J. A., Vetting, M., Ghodge, S. V., Xu, C., Hillerich, B., Seidel, R. D., Almo, S. C., and Raushel, F. M. (2014) Prospecting for Unannotated Enzymes: Discovery of a 3′,5′-Nucleotide Bisphosphate Phosphatase within the Amidohydrolase Superfamily. *Biochemistry 53,* 591−600.
19. Vladimirova, A., Patskovsky, Y., Fedorov, A. A., Bonanno, J. B., Fedorov, E. V., Toro, R., Hillerich, B., Seidel, R. D., Richards, N. G. J., Almo, S. C., and Raushel, F. M. (2016) Substrate Distortion and the Catalytic Reaction Mechanism of 5-Carboxyvanillate Decarboxylase. *J. Am. Chem. Soc. 138,* 826−836.

20. Yoon, S. H., Moon, T. S., Iranpour, P., Lanza, A. M., and Prather, K. J. (2009) Cloning and Characterization of Uronate Dehydrogenases from Two Pseudomonads and *Agrobacterium tumefaciens* Strain C58. *J. Bacteriol. 191,* 1565−1573.
21. Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res. 28,* 33−36.
22. Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., Geer, L. Y., and Bryant, S. H. (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res. 45,* D200−D203.
23. Kuatsjah, E., Chan, A. C. K., Kobylarz, M. J., Murphy, M. E. P., and Eltis, L. D. (2017) The bacterial meta-cleavage hydrolase LigY belongs to the amidohydrolase superfamily, not to the α/β-hydrolase superfamily. *J. Biol. Chem. 292,* 18290−18302.
24. Kersten, P. J., Dagley, S., Whittaker, J. W., Arciero, D. M., and Lipscomb, J. D. (1982) 2-Pyrone-4,6-Dicarboxylic Acid, a Catabolite of Gallic Acids in *Pseudomonas* Species. *J. Bacteriol. 152,* 1154−1162.
25. Yerabolu, J. R., Liotta, C. L., and Krishnamurthy, R. (2017) Anchimeric-Assisted Spontaneous Hydrolysis of Cyanohydrins Under Ambient Conditions: Implications for Cyanide-Initiated Selective Transformations. *Chem. - Eur. J. 23,* 8756−8765.

26. Gerlt, J. A., Allen, K. N., Almo, S. C., Armstrong, R. N., Babbitt, P. C., Cronan, J. E., Dunaway-Mariano, D., Imker, H. J., Jacobson, M. P., Minor, W., Poulter, C. D., Raushel, F. M., Sali, A., Shoichet, B. K., and Sweedler, J. V. (2011) The Enzyme Function Initiative. *Biochemistry 50,* 9950−9962.

27. Gerlt, J. A., Bouvier, J. T., Davidson, D. B., Imker, H. J., Sadkhin, B., Slater, D. R., and Whalen, K. L. (2015) Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta, Proteins Proteomics 1854,* 1019−1037.

28. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res. 13,* 2498−2504.

29. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. In *Methods in Enzymology* (Carter, C. W., Jr., Ed.) Vol. 276, pp 307−326, Academic Press, New York.

30. Forouhar, F., Abashidze, M., Jayaraman, S., Cunningham, K., Ciao, M., Ma, L., Xiao, R., Acton, T. B., Montelione, G. T., Hunt, J. F., and Tong, L. (2006) Crystal structure of 4-oxalomesaconate hydratase LigJ from *Rhodopseudomonas palustris*, Northeast Structural Genomics target RpR66. RCSB Protein Data Bank entry 2GWG.

31. Adams,P.D.,Afonine,P.V.,Bunkoćzi,G.,Chen,V.B.,Davis, I. W., Echols, N., Headd, J. J., Hung, E.-W., Kapral, G. J., Grosse- Kunstleve, R. W., McCoy, A. J., Moriarty,

N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66,* 213−221.

32. Emsley, P., and Cowtan, K. (2004) Coot model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60,* 2126−2132.
33. Krissinel, E. (2015) Stock-based detection of protein oligomeric states in jsPISA. *Nucleic Acids Res. 43,* W314−W319.
34. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera -A visualization system for exploratory research and analysis. *J. Comput. Chem. 25,* 1605−1612.
35. Masai, E., Katayama, Y., and Fukuda, M. (2007) Genetic and Biochemical Investigations on Bacterial Catabolic Pathways for Lignin-Derived Aromatic Compounds. *Biosci., Biotechnol., Biochem. 71,* 1−15.
36. Mazurkewich, S., Brott, A. S., Kimber, M. S., and Seah, S. Y. K. (2016) Structural and Kinetic Characterization of the 4-Carboxy-2-hydroxymuconate Hydratase from the Gallate and Protocatechuate 4,5-Cleavage Pathways of *Pseudomonas putida* KT2440. *J. Biol. Chem. 291,* 7669−7686.
37. Wong, B. J., and Gerlt, J. A. (2004) Evolution of Function in the Crotonase Superfamily: (3S)-Methylglutaconyl-CoA Hydratase from *Pseudomonas putida*. *Biochemistry 43,* 4646−4654.

38. Mohrig, J. R., Moerke, K. A., Cloutier, D. L., Lane, B. D., Person, E. C., and Onasch, T. B. (1995) Importance of Historical Contingency in the Stereochemistry of Hydratase-Dehydratase Enzymes. *Science 269,* 527−529.

39. Holm, L., and Sander, C. (1997) An Evolutionary Treasure: Unification of a Broad Set of Amidohydrolases Related to Urease. *Proteins: Struct., Funct., Genet. 28,* 72−82.

40. Benning, M. M., Kuo, J. M., Raushel, F. M., and Holden, H. M. (1994) Three-Dimensional Structure of Phosphotriesterase: An Enzyme Capable of Detoxifying Organophosphate Nerve Agents. *Biochemistry 33,* 15001−15007.

41. Thoden, J. B., Phillips, G. N., Jr., Neal, T. M., Raushel, F. M., and Holden, H. M. (2001) Molecular Structure of Dihydroorotase: A Paradigm for Catalysis through the Use of a Binuclear Metal Center. *Biochemistry 40,* 6989−6997.

42. Jabri, E., Carr, M. B., Hausinger, R. P., and Karplus, P. A. (1995) The Crystal Structure of Urease from *Klebsiella aerogenes*. *Science 268,* 998−1004.

43. Ireton, G. C., McDermott, G., Black, M. E., and Stoddard, B. L. (2002) The Structure of *Escherichia coli* Cytosine Deaminase. *J. Mol. Biol. 315,* 687−697.

44. Ghodge, S. V., Fedorov, A. A., Fedorov, E. V., Hillerich, B., Seidel, R., Almo, S. C., and Raushel, F. M. (2013) Structural and Mechanistic Characterization of L-Histidinol Phosphate Phosphatase from the Polymerase and Histidinol Phosphatase Family of Proteins. *Biochemistry 52,* 1101−1112.

45. Sheng, X., Patskovsky, Y., Vladimirova, A., Bonanno, J. B., Almo, S. C., Himo, F., and Raushel, F. M. (2018) Mechanism and Structure of γ-Resorcylate Decarboxylase. *Biochemistry 57,* 3167− 3175.

46. Snini, S. P., Tadrist, S., Laffitte, J., Jamin, E. L., Oswald, I. P., and Puel, O. (2014) The gene PatG involved in the biosynthesis pathway of patulin, a food-borne mycotoxin, encodes a 6-methylsalicylic acid decarboxylase. *Int. J. Food Microbiol. 171,* 77−83.

47. Xu, S., Li, W., Zhu, J., Wang, R., Li, Z., Xu, G. L., and Ding, J. (2013) Crystal structure of isoorotate decarboxylases reveal a novel catalytic mechanism of 5-carboxyl-uracil decarboxylation and shed light on the search for DNA decarboxylase. *Cell Res. 23,* 1296−1309.

48. Li, T., Iwaki, H., Fu, R., Hasegawa, Y., Zhang, H., and Liu, A. (2006) α-Amino-β-carboxymuconate-ε-semialdehyde Decarboxylase (ACMSD) Is a New Member of the Amidohydrolase Superfamily. *Biochemistry 45,* 6628−6634.

49. Martynowski, D., Eyobo, Y., Li, T., Yang, K., Liu, A., and Zhang, H. (2006) Crystal Structure of α-Amino-β-carboxymuconate- ε-semialdehyde Decarboxylase: Insight into the Active Site and Catalytic Mechanism of a Novel Decarboxylation Reaction. *Biochemistry 45,* 10412−10421.

50. Huo, L., Liu, F., Iwaki, H., Li, T., Hasegawa, Y., and Liu, A. (2015) Human α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD): A structural and mechanistic unveiling. *Proteins: Struct., Funct., Genet. 83,* 178−187.

4. STRUCTURE AND CHEMICAL REACTION MECHANISM OF LIGU, AN ENZYME THAT CATALYZES AN ALLYLIC ISOMERIZATION IN THE BACTERIAL DEGRADATION OF LIGNIN*

4.1. Introduction

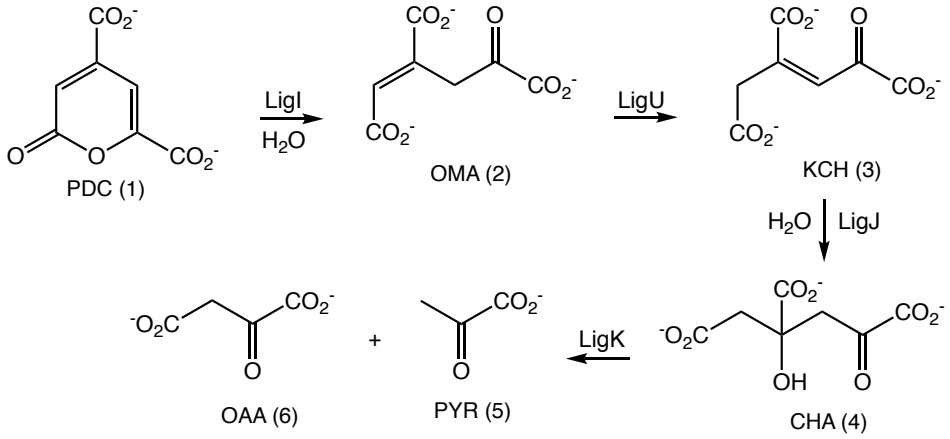
Lignin, the most abundant renewable aromatic compound in the world, is currently being pursued as a viable starting material for biopolymers. However, it is an underutilized renewable resource due to its recalcitrant nature as a complex heteropolymer. Lignin is comprised of phenylpropanoid units (guaiacyl, syringyl, and p-hydroxyphenyl precursors) that couple to one another through C-C and C-O bonds (1, 2). The natural recycling of lignin occurs through the combined actions of eukarya, archaea, and bacteria (3). In general, fungi are responsible for the depolymerization of lignin, and bacteria further metabolize the monomeric units into viable carbon sources. Many of the known enzymatic pathways funnel to the intermediate protocatechuate (PCA) where it is further metabolized by three cleavage pathways (4). In the PCA 4,5-cleavage pathway, the LigAB complex converts PCA and O_2 into 4-carboxy-2-hydroxymuconate-6-semialdehyde (CHMS) (5). LigC then oxidizes CHMS with $NADP^+$ to form 2-pyrone-4,6-dicarboxylate (PDC, 1) (6). LigI hydrolyzes PDC into (4*E*)-oxalomesaconate (OMA, 2) (7). LigU isomerizes OMA into (3*Z*)-2-keto-4-carboxy- 3-hexendioate (KCH,

---

* Reprinted with permission from "Structure and Chemical Reaction Mechanism of LigU, an Enzyme That Catalyzes an Allylic Isomerization in the Bacterial Degradation of Lignin" by Tessily N. Hogancamp, Seth A. Cory, David P. Barondeau, and Frank M. Raushel, *Biochemistry*, 2019, *58* (33), pp 3494-3503, Copyright 2019 American Chemical Society

3) (8). LigJ hydrates KCH to (*S*)-4-carboxy-4-hydroxy-2-oxoadipate (CHA, 4) (9). Finally, LigK cleaves CHA into pyruvate (PYR, 5) and oxaloacetate (OAA, 6), which are then shuttled into the TCA cycle (**Scheme 12**) (10, 11).

Scheme 12. Reactions catalyzed by LigI, LigU, LigJ and LigK.



Previously, we demonstrated that LigU isomerase catalyzes the conversion of OMA to KCH via a 1,3-allylic isomerization of the double bond from C4/C5 to C3/C4. The structural characterization of LigJ hydratase complexed with the substrate KCH (3) provided evidence supporting the migration of the double bond from C4/C5 to C3/C4 prior to the addition of water to C4 of KCH (8).

LigU belongs to cog2828, which is also known as the PrpF superfamily. PrpF is reported to catalyze the isomerization of 4-methyl-cis-aconitate to 2-methyl-cis-aconitate, an intermediate in the 2-methylcitric acid cycle (12, 13). Structurally, this family shares a common protein fold, which was first discovered for diaminopimelate epimerase (14), where there is an apparent domain duplication event and the active site is at the junction of the two domains (15, 16). The sequence similarity network (SSN) for cog2828 at an E value cutoff of 10^{-80} illustrates the sequence relationship among the 158 nonredundant protein sequences using a 40% identity threshold (**Figure 36**).

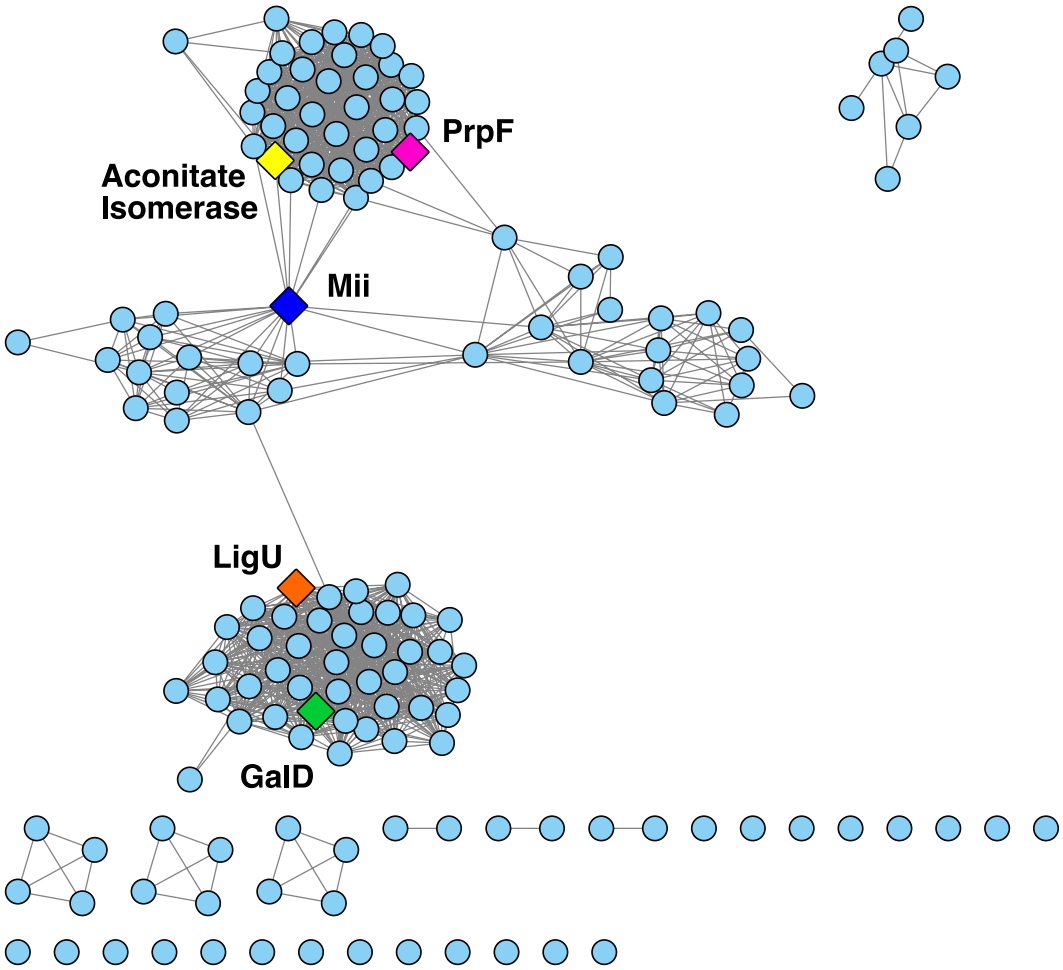


Figure 36. Sequence Similarity Network (SSN) of cog2828 at a BLAST E value cutoff of 10^{-80} contains 1,457 edges and 158 nodes generated using Cytoscape (http://cytoscape.org) (17). The list of proteins was generated using EFI-EST using a 40% identity threshold (18, 19). Cog2828 proteins were obtained from UniProt (20).

Several aspects of the reaction mechanism for LigU and the PrpF enzyme family remain unknown. To provide mechanistic details of the LigU catalyzed isomerization reaction, we have determined the three-dimensional crystal structure of LigU in the presence and absence of bound ligands. These structures, and the catalytic properties of mutant enzymes, have led to the identification of active site residues that are required for the abstraction/donation of protons at C3 and C5 of the substrate and product. The stereochemistry of the proton abstracted from C5 was determined by ^{1}H NMR spectroscopy of the reaction products when the reaction was conducted in D_2O. Based on these findings a detailed chemical reaction mechanism has been proposed for the reaction catalyzed by LigU. In this mechanism, Cys-100 abstracts the pro-*S* hydrogen from C3 of OMA, and Lys-66 donates a proton to the *si*-face of the double bond at C5 for a net suprafacial hydrogen transfer.

4.2. Materials and Methods

4.2.1. Materials

All commercial chemicals used in this investigation were obtained from Millipore Sigma, unless otherwise stated. 2-Pyrone-4,6-dicarboxylate (PDC, 1) was synthesized as previously described (21). The LigI hydrolase (UniProt entry: O87170), LigJ hydratase (UniProt entry: G2IQQ5), and LigK aldolase (UniProt entry: G2IQQ8) were isolated and purified as previously described (7-9).

4.2.2. Cloning, Expression, and Isolation of LigU

The open reading frame encoding LigU from *Novosphingobium* sp. strain KA1 (*Sphingomonas* sp. strain KA1) (UniProt entry: Q0KJL4) was chemically synthesized by GenScript and subsequently ligated into a pET- 30a(+) vector (EMD Millipore). The protein is 357 residues long plus a C-terminal linker and His tag added as Leu-Glu-His6. The gene was expressed in *Escherichia coli* BL21(DE3) (EMD Millipore) and the protein purified as previously described (8). The mutants S21A, K66Q, K66M, C100S, S254A, K257M, and H281N were constructed using the Agilent QuikChange protocol and the primers are listed in **Table 6**. The mutant enzymes were purified as described previously for the wild-type enzyme (8). To improve the purity of the enzymes used in the crystallization trials, size-exclusion chromatography was added to the purification protocol of LigU (8). Protein was injected onto a HiLoad 16/600 Superdex 200 size exclusion chromatography column (GE Healthcare) equilibrated in 20 mM HEPES/KOH, pH 7.9 and eluted using a flow rate of 0.8 mL/min. The fractions were checked for purity by Mini-PROTEAN TGX stain-free SDS-PAGE (BIO-RAD). Selected fractions were combined and concentrated to ~25 mg/mL. The combined fractions were estimated to be ≥95% pure.

Mutant	Forward Primer Reverse Primer
S21A	CGTGGCGGTACCGCCAAAGGTGGCTAC GTAGCCACCTTTGGCGGTACCGCCACG
K66Q	GACCCGCTGACCAGCCAGGTGGCGGTGGTTAG CTAACCACCGCCACCTGGCTGGTCAGCGGGTC
K66M	GACCCGCTGACCAGCATGGTGGCGGTGGTTAG CTAACCACCGCCACCATGCTGGTCAGCGGGTC
C100S	CCGATGCGCAGAACAGCGGTAACATCCTGGC GCCAGGATGTTACCGCTGTTCTGCGCATCGG
S254A	GGCGACGTGACCGAAAAGGCCGTTCCGAAAATGATG CATCATTTTCGGAACGGCCTTTTCGGTCACGTCGCC
K257M	GAGCGTTCCGATGATGATGCTGGTTGCGCCGCCG CGGCGGCGCAACCAGCATCATCATCGGAACGCTC
H281N	CGCACCGTGCGAATGCGACCATTGGTGTTC GAACACCAATGGTCGCATTCGCACGGTGCG

Table 6. The forward and reverse primers constructed for LigU mutants. The change to the primary sequence in underlined.

4.2.3. Measurement of Catalytic Activity

The initial velocity for the isomerization of OMA (2) to KCH (3) by LigU was determined spectrophotometrically by monitoring the change in absorbance at 265 nm and 30 °C in a 96-well UV-STAR (Greiner Bio-One) plate in a volume of 250 µL. OMA was made *in situ* by the addition of 1.0 µM LigI hydrolase to variable concentrations of PDC (0 to 1.2 mM) in 200 mM HEPES/KOH (pH 8.0). Control reactions were conducted in parallel using the same variable concentrations of PDC to obtain the nonenzymatic conversion of OMA to KCH. The reaction catalyzed by LigI was monitored at 312 nm until the reaction was complete. Aliquots of 10 µL of an appropriately diluted LigU solution were subsequently added and the reaction was monitored by changes in absorbance at 265 nm. The concentrations of the LigU mutants used for measurement of the catalytic properties were as follows: K66Q (1.0 µM), K66M (2.0 µM), C100S (50 nM), H281N (25 nM), S21A (0.5 nM), K257M (1.0 µM), and S254A (250 pM). The change in the concentration of OMA was determined using a differential extinction coefficient ($\Delta\varepsilon_{265}$) of 5900 M^{-1} cm^{-1} (8). The kinetic constants were determined using an iterative non-linear least square regression fit to **eq 3** where ν is the initial velocity, E_t is the total enzyme concentration, k_{cat} is the turnover number, A is the substrate concentration, and K_m is the Michaelis constant. SigmaPlot 11.0 was used to calculate the kinetic constants and standard errors.

$$\nu/E_t = k_{cat}(A)/(K_m + A)$$

Equation 3

4.2.4. Stereospecific Hydrogen Exchange at C5 of KCH

The stereospecific exchange of protons at C5 of KCH was determined using ^{1}H NMR spectroscopy with a Bruker Avance III 500 MHz spectrometer equipped with an H-C-N cryoprobe at 305 K. The reaction was conducted in 200 mM KH_2PO_4/KOD (pD 6.9), containing 5.0 mM $MgCl_2$, 6.0 mM NADH, 5.0 mM PDC, and a lyophilized enzyme cocktail of 1.0 μM each LigI, LigU, LigJ, LigK and malate dehydrogenase in a volume of 600 μL of D_2O. Collectively, these enzymes convert the initial substrate PDC (1) to malate. The reaction was monitored by ^{1}H NMR spectroscopy using the water suppression pulse program Watergate zggpw5 pulse sequence (18).

The deuterated L-malate, made from the combined activities of LigI, LigU, LigJ, LigK, and malate dehydrogenase was compared to the deuterated L-malate formed from the action of fumarase with fumarate in D_2O and also with [2,3-^{2}H]-fumarate in H_2O. The control reaction was conducted in 200 mM KH_2PO_4/KOD (pD 6.6), 5.0 mM fumaric acid, and 1.0 μM fumarase from porcine heart (Sigma-Aldrich) in a final volume of 600 μL in D_2O. The zggpw5 pulse program with water suppression was calibrated at 4.675 ppm. A fully protonated sample of 5.0 mM malate was prepared in 200 mM KH_2PO_4/KOD (pD 7.02) in a final volume of 600 μL in ~100% D_2O.

4.2.5. Hydrogen-Deuterium Exchange Experiments

The proton exchange at C5 of KCH (**3**) with solvent D_2O in the presence of LigU was performed in 200 mM potassium phosphate/KOD (pD 7.0) in a volume of 600 μL. The reaction was monitored by ^{1}H NMR spectroscopy with a Bruker Avance III 500

MHz spectrometer equipped with an H-C-N cryoprobe at 305 K. KCH was prepared by mixing 20 mM PDC (**1**) with 200 mM $NaHCO_3$/KCl, pH 9.6, in the presence of 1.0 μM LigI hydrolase and subsequent titration to pH 9.5 with KOH. The reaction was incubated for 2 h at room temperature, and then LigI was removed using PALL Nanosep 10 kDa and 3 kDa filters with an Eppendorf centrifuge. The flow-through was collected and aliquoted (150 μL) into 1.7 mL sterile microcentrifuge tubes. The samples were lyophilized to dryness and stored at -20°C until needed.

The hydrogen-deuterium exchange reactions were initiated by adding 580 μL of D_2O containing 200 mM KH_2PO_4/KOD (pD 6.6 to 7.0) to a lyophilized sample of KCH. A total of 1.0 nM LigU and 66 nM bovine serum albumin (BSA) was added to the reaction mixture and the reaction followed by ^1H NMR spectroscopy. The hydrogen-deuterium exchange reaction was monitored for ~6.4 h using the Watergate zggpw5 pulse program with water suppression (*18*). The mutants K66M, C100S, and H281N were also tested in a similar fashion. A 16 scan acquisition was collected before any enzyme was added to calibrate the instrument for each of the reactions with the three mutant proteins. The total concentration of LigU K66M, C100S, and H281N added to each reaction was 111 nM, 11 nM, and 102 nM, respectively.

4.2.6. Crystallization of LigU Wild-Type and Mutants

Crystals of LigU wild-type were prepared by sitting drop vapor diffusion (Intelli-Plate 24-4 by Hampton Research) with a drop and reservoir volume of 4 μL and 500 μL, respectively. The reservoir contained 200 mM $CaCl_2$ dihydrate, 100 mM HEPES/KOH,

pH 6.5, and 35% pentaerythritol ethoxylate (Molecular Dimensions). The drop was 2.0 µL enzyme (22.3 mg/mL in 20 mM HEPES/KOH, pH 7.5) and 2.0 µL reservoir solution. Precipitation was observed after ~24 h. Crystals were observed after incubation at 17 ºC for approximately 6 months. These crystals were supplemented with 5.0 µL of reservoir solution and harvested by a nylon loop (Hampton Research) prior to flash freezing in liquid nitrogen. Crystals of LigU C100S were prepared by sitting drop vapor diffusion with a drop and reservoir volume of 4.0 µL and 500 µL, respectively. The reservoir contained 200 mM $MgCl_2$ hexahydrate, 100 mM HEPES/KOH, pH 7.5, and 25 % (w/v) PEG 3350 (Hampton Research). The drop was 2.0 µL enzyme (10 mg/mL in 20 mM HEPES/KOH, pH 7.9) and 2.0 µL reservoir solution. Crystals with the morphology of glass plates were observed after a ~2 week incubation at 17 °C. Flash freezing of crystals was conducted in a similar fashion to wild-type crystals. Sparse matrix screening of LigU K66M in the presence of OMA was performed at 17 ºC using a Mosquito robot (TTP Labtech). Initial crystallization conditions were identified within the Wizard Classic 1/2 screen (Rigaku) containing 1000 mM sodium citrate tribasic, 100 mM sodium cacodylate/HCl, pH 6.5 with a total volume of 100 µL. The drop was 200 nL enzyme (20.4 mg/mL in 20 mM HEPES/KOH, pH 7.9), 100 nL ligand (40 mM OMA, pH 8.2), and 500 nL reservoir solution. OMA was made by incubating 40 mM PDC with ~4 µM of LigI hydrolase in 200 mM HEPES/KOH, pH 8.2 using KOH to adjust the pH of the reaction solution to pH 8.2. LigI was removed by centrifugation using PALL Nanosep 10 KDa and 3 KDa filters and the flow-through was collected. These crystals were incubated at 17°C for approximately 1 week, harvested directly

from the drop using a nylon loop, and frozen directly in liquid nitrogen. Initial LigU C100S (2.5 Å) crystals were prepared by the hanging drop vapor diffusion method. The reservoir solution contained 0.20 M $MgCl_2$, 0.10 M HEPES/KOH, pH 7.5, and 25% (w/v) PEG 3350. LigU C100S was concentrated to 10 mg/mL in 10 mM HEPES/KOH, pH 7.5, and then 2.0 µL of the protein solution was mixed with 2.0 µL of the well solution and 1.0 µL of a 10 mM solution of OMA. The crystals were stored at 17 °C and incubated for ~3 months before being submerged in reservoir solution, harvested, and plunged into liquid nitrogen.

4.2.7. Data Collection and Processing

Diffraction data was collected at Stanford Synchrotron Radiation Lightsource (SSRL) BL 14-1 on an Eiger 16M detector at 100 K. Data was integrated in either HKL2000 or iMosflm (23, 24). Scaling and merging was conducted in either HKL2000 or Aimless (23, 25). Merged observed intensities were then imported into Phenix reflection file editor where an R-free set was assigned using the default settings (2000 reflections) (26).

4.2.8. Structure Determination

Using a lower resolution dataset, initial phases for LigU (C100S) were determined by molecular replacement using chain A of PDB entry: 3G7K (methylitaconate-Δ-isomerase, *Eubacterium barkeri* with 41% sequence identity) as a search model in the Phaser-MR module of Phenix (27, 28). Several rounds of model

building and refinement were conducted using AutoBuild. Additional rounds of manual rebuilding were then conducted in Coot and refinement of the coordinates performed in phenix.refine (29, 30). A monomer from this model was used for molecular replacement against the wild type and LigU(K66M) + OMA datasets. For the high resolution LigU(C100S) dataset, the original R-free flags were transferred and extended to the higher resolution for further refinement. In the LigU (K66M) + OMA dataset, translational NCS was identified by Phenix Xtriage (26). However, the translational NCS did not appear to have an impact on molecular replacement or refinement. OMA was built in JLigand and optimized in REEL (31, 32). Non-crystallographic symmetry restraints were utilized in the early stages of refinement but were released toward the ending stages. Final refinement of coordinates was conducted in phenix.refine with the following settings enabled: xyz (reciprocal and real), individual b-factors, TLS, occupancies, X-ray/stereochemistry weights, and X-ray/ADP weights. Real space xyz refinement was disabled during the last cycle of refinement for LigU due to significant disorder in chain B of the structure. Model validation was conducted in Phenix and Molprobity prior to deposition (33). Data collection and refinement statistics can be found in **Table 7**. The RMSD calculations were determined using Chimera to align the residue pairs with refitting iterations until full population convergence with a 5.0 Å cutoff for the maximum Cα-Cα pairwise distance (34).

Structure	LigU	LigU (C100S)	LigU (K66M) + OMA
PDB entry	6P3J	6P3K	6P3H
Data Collection Parameters			
Beamline	SSRL BL14-1	SSRL BL14-1	SSRL BL14-1
Wavelength (Å)	1.1271	1.1271	1.1271
Resolution (Å)	50.00 - 2.02	50.00 - 1.88	39.53 - 1.62
Space Group	$P2_12_12_1$	C2	$P2_12_12$
a, b, c (Å)	62.78, 67.36, 161.07	97.42, 142.33, 60.35	84.88, 134.00, 168.37
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 111.73, 90.00	90.00, 90.00, 90.00
Completeness (%)[a]	97.8 (86.5)	87.3 (47.1)	97.5 (94.2)
Unique reflections	45,097	53,939	236,378
I/σ[a]	30.2 (1.1)	22.1 (2.1)	8.8 (2.6)
$CC_{1/2}$ last shell (%)	59.3	91.5	44.6
Redundancy[a]	5.2 (2.8)	2.9 (1.8)	5.1 (5.2)
R_{merge}[a]	5.7 (59.7)	4.6 (24.2)	9.1 (59.7)
R_{pim}[a]	2.6 (38.8)	3.0 (20.0)	4.4 (27.6)
Refinement			
resolution (Å)[a]	42.00 - 2.02 (2.07 - 2.02)	42.03 - 1.88 (1.92 - 1.88)	39.47 - 1.62 (1.66 - 1.62)
no. of reflections (free)[a]	44,713 (2,003)	53,866 (2,597)	236,265 (2,000)
R_{work}[a]	18.6 (30.1)	17.3 (23.2)	15.7 (28.7)
R_{free}[a]	21.2 (31.1)	21.5 (29.8)	18.7 (28.3)
Molecules in asymmetric unit	2	2	4
TLS groups	5	4	8
No. of atoms			
protein	5,144	5,140	10,505
Ions	12	6	6
Ligand	NA	NA	56
Water	224	388	1,516
B-factors			
All	61.3	43.0	29.4
Protein	61.8	43.1	28.2
Ions	59.9	37.7	35.8
Ligand	NA	NA	23.8
Water	50.0	41.9	38.4
Validation			
Ramachandran preferred (%)	96.8	96.7	98.0
Ramachandran allowed (%)	3.2	3.3	1.9
Ramachandran outliers (%)	0.0	0.0	0.1
RMSD for bond lengths (Å)	0.003	0.014	0.017
RMSD for bond angles (°)	0.596	1.249	1.394
Rotamer outliers (%)	0.9	0.6	0.8
Clash score	2.9	5.6	4.0

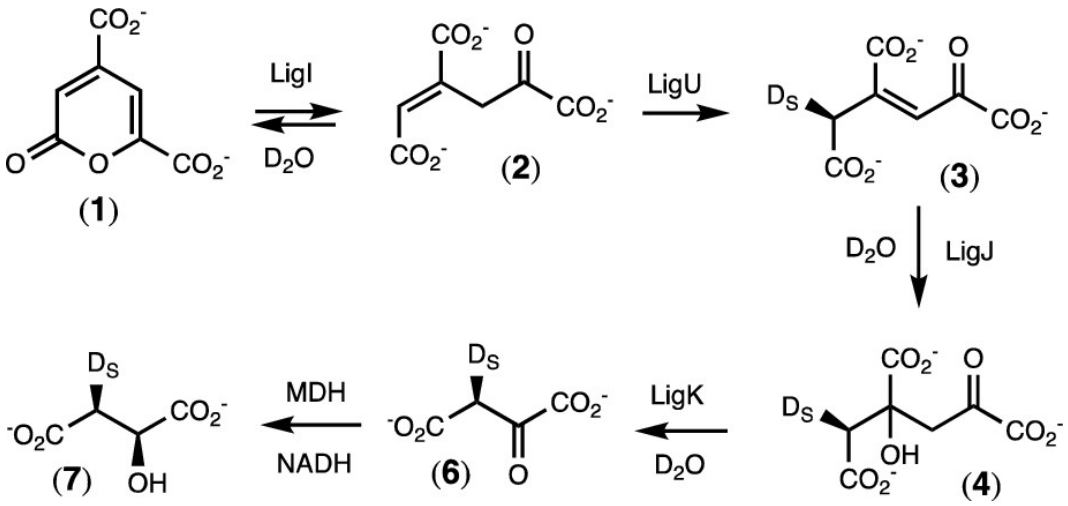
Table 7. Data collection and refinement statistics for LigU isomerase crystals.

4.3. Results

4.3.1. Stereochemistry of Proton Transfer at C5 by LigU

During the allylic isomerization catalyzed by LigU, a proton is abstracted from C3 of OMA (2) and another proton (or perhaps the same proton) is added to C5 to form KCH (3) (**Scheme 12**). To determine whether the proton transfer at C5 of KCH occurs with either the pro-*R* or pro-*S* hydrogen, the isomerization reaction was conducted in D_2O in the presence of LigU, LigJ, and LigK to drive the reaction forward toward the formation of oxaloacetate (6). Simultaneously, the oxaloacetate was reduced to L-malate (7) using NADH with malate dehydrogenase (MDH) (**Scheme 13**). C3 of the enzymatically formed L-malate corresponds with C5 of KCH, and it will be bonded to one hydrogen and one deuterium (from solvent).

Scheme 13. Stereochemistry of the LigU-Catalyzed Isomerization Reaction.



Proton NMR spectroscopy was used to determine the specific stereochemical arrangement of the hydrogen and deuterium at C3 of L-malate (7) formed from PDC in the presence of LigI, LigU, LigJ, LigK, and MDH when the reactions were conducted in D_2O. In the ^1H NMR spectrum of unlabeled L-malate, the pro-*R* and pro-*S* hydrogens at C3 are separated by 0.31 ppm (**Figure 38a**). The chemical shift assignments for these two hydrogens were determined by catalyzing the formation of L-[3-^2H]malate by the addition of D_2O to fumarate with fumarase (**Figure 38b**). It is known that fumarase catalyzes the abstraction of the pro-*R* hydrogen of L-malate at C3, and thus, the pair of partially resolved triplets at 2.309 ppm in **Figure 38b** is due to the pro-*S* hydrogen (35, 36). The L-malate formed from the LigU-catalyzed reaction in D_2O exhibits an unresolved triplet at 2.605 ppm, and thus, LigU catalyzes the exchange of the pro-*S* hydrogen from C5 of KCH (**Figure 38c**). The stereochemical transformations are summarized in **Scheme 13**.

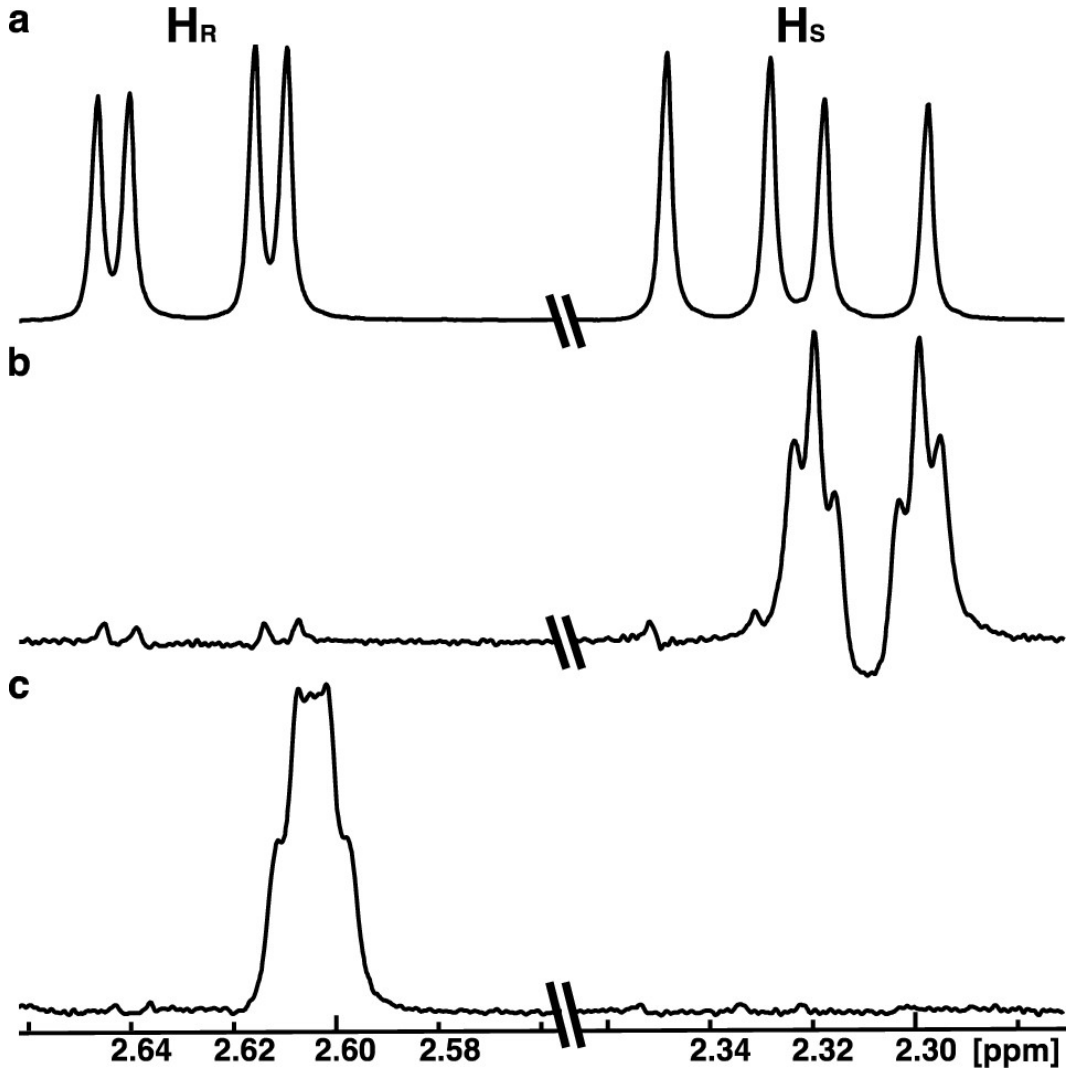


Figure 37. ^{1}H NMR spectra of L-malate for the hydrogen(s) attached to C3. (a) L-Malate purchased commercially with H_R from C3 a doublet of doublets centered at 2.629 ppm with a ^{2}J of 15.3 Hz and a ^{3}J of 3.1 Hz and H_S from C3 a doublet of doublets centered at 2.323 ppm with a ^{2}J of 15.3 Hz and a ^{3}J of 10.1 Hz. (b) L-Malate formed from the enzymatic hydration of fumarate with fumarase in D_2O with H_S from C3 a doublet of triplets centered at 2.309 ppm with a ^{3}J of 10.2 Hz. (c) L-Malate formed enzymatically from PDC (1) in the presence of LigI, LigU, LigJ, LigK, and L-malate dehydrogenase in D_2O with H_R from C3 an unresolved triplet centered at 2.605 ppm.

4.3.2. Mutation of Active Site Residues

The kinetic constants for the isomerization of OMA to KCH were determined for wild-type LigU and seven single-site mutant enzymes (**Table 8**). The residues Lys-66 and Cys-100 were mutated on the basis of previous studies, suggesting they were catalytic residues in homologous enzymes. The other four residues (Ser-21, Ser-254, Lys-257, and His-281) were chosen on the basis of homology models to be in the active site pocket. The wild-type enzyme is quite active with k_{cat} and k_{cat}/K_m values of 1300 s^{-1} and 7.7 × 10^6 M^{-1} s^{-1}, respectively. No activity ($k_{cat}/ K_m < 14$ M^{-1} s^{-1}) was observed with the K66Q, K66M, and K257M mutants, and significant losses in activity were observed for the C100S and H281N mutants.

enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
wild-type	1300 ± 120	170 ± 30	(8 ± 2) x 10^6
S21A	60 ± 5	120 ± 30	(5 ± 1) x 10^5
K66Q	--	--	≤ 4
K66M	--	--	≤ 4
C100S	1.6 ± 0.2	140 ± 50	(12 ± 3) x 10^3
S254A	150 ± 10	100 ± 20	(14 ± 2) x 10^5
K257M	--	--	≤ (14 ± 3)
H281N	4.3 ± 0.6	380 ± 110	(11 ± 2) x 10^3

Table 8. Kinetic Constants for Isomerization of OMA to KCH by LigU[a].

[a] At 30 °C, pH 8.0, and OMA concentrations from 0 to 0.8 mM. The kinetic parameters were calculated using an iterative nonlinear least-squares regression fitting analysis, and standard errors were obtained by the matrix inversion method using Sigmaplot 11.0.

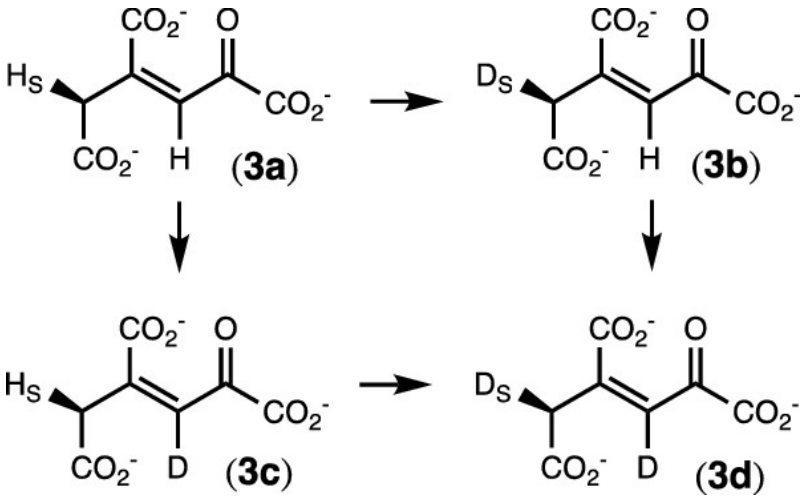
4.3.3. Hydrogen-Deuterium Exchange at C5 of KCH

In the LigU-catalyzed isomerization of OMA, a proton from C3 must be abstracted and a proton must subsequently be added to C5 to form KCH. ^{1}H NMR spectroscopy was used to monitor the exchange of hydrogens at C5 of KCH in D_2O during the LigU-catalyzed reaction in an attempt to determine which residues are responsible for these protons transfers (**Scheme 14**). In the ^{1}H NMR spectrum of KCH (3a in **Scheme 14**), the two hydrogens at C5 of KCH resonate as a doublet at 3.209 ppm. These two hydrogens are spin coupled to the hydrogen at C3 (6.862 ppm) with a ^{4}J of 1.4 Hz (**Figure 39a**). After the addition of wild-type LigU, the intensity of the doublet at 3.209 ppm decreases and is gradually replaced by an unresolved triplet at 3.191 ppm with an apparent coupling constant of ~1.9 Hz. The new triplet originates from the pro-*R* hydrogen at C5 that is spin coupled to the geminal pro-*S* deuterium of KCH (3b and 3d). Also observed is the formation of a new singlet at 3.207 ppm that partially overlaps with the original doublet centered at 3.209 ppm. This new resonance must be due to the exchange of the hydrogen at C3 (3c) of KCH with solvent deuterium and the accompanying absence of the ^{1}H−^{1}H spin coupling. The loss of the original hydrogen at C3 is possible only via the transient enzymatic formation of OMA and the subsequent non-enzymatic exchange of the labile hydrogens adjacent with the carbonyl group at C2.

The hydrogen−deuterium exchange reactions were similarly conducted with the K66M, C100S, and H281N mutants of LigU. When the K66M mutant of LigU was added to KCH, there was no observable change in the intensity of the ^{1}H NMR spectrum for the C5 protons of KCH (**Figure 39b**). However, significant exchange of a proton

attached to C5 of KCH was observed when the C100S mutant was utilized (**Figure 39c**). The intensity of the doublet for the protons attached to C5 at 3.211 ppm decreases, while the intensity of the unresolved triplet at 3.194 ppm increases until the doublet completely vanishes. There was no indication of the appearance of a new singlet slightly upfield from the doublet at 3.207 ppm as was observed with wild-type LigU. The solvent exchange reaction for the H281N mutant of LigU was similar to that of the wild-type LigU except that higher enzyme concentrations were required to observe the changes in the NMR spectrum of KCH (**Figure 39d**).

Scheme 14. Hydrogen-deuterium exchange at C5 of KCH catalyzed by LigU isomerase.



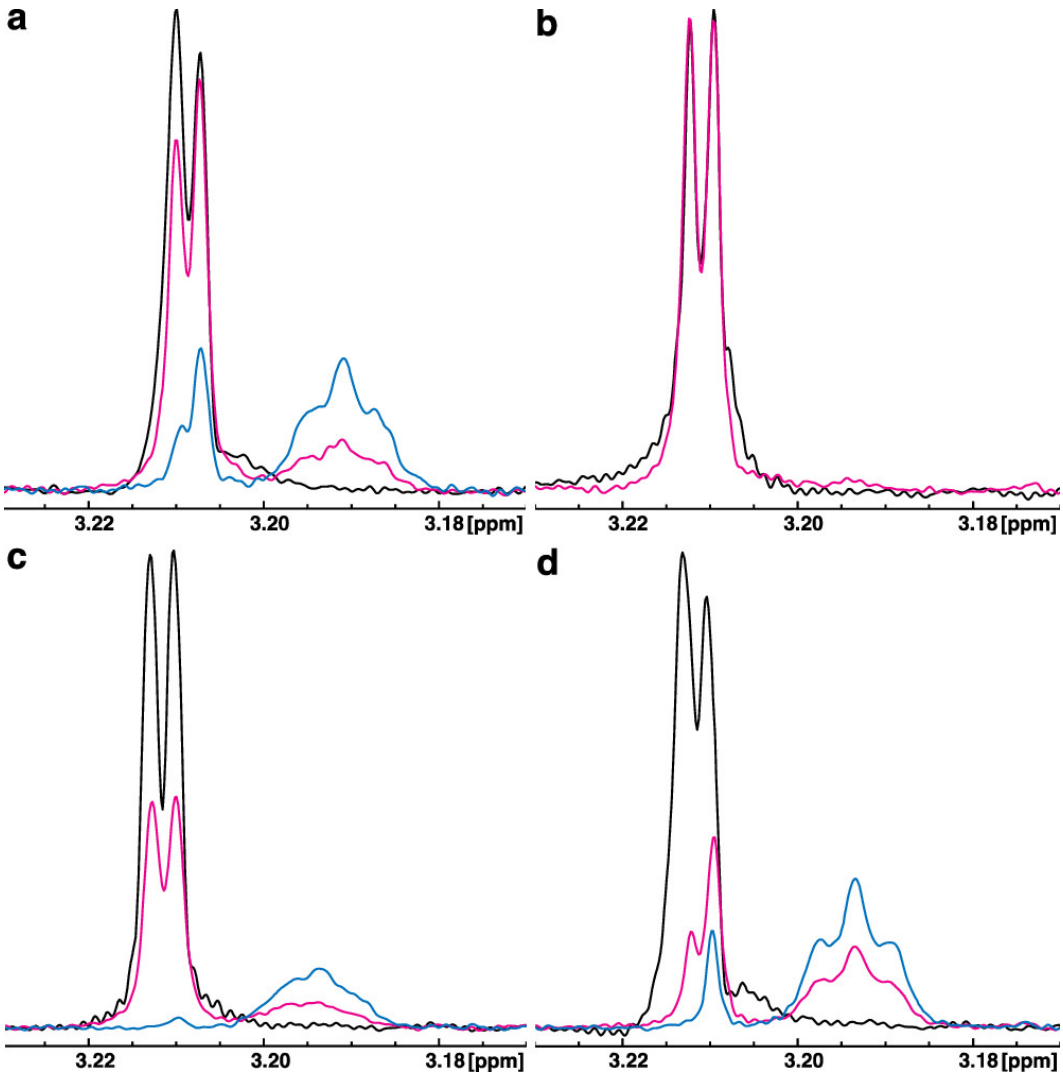


Figure 38. ^{1}H NMR spectra of 1.0 mM KCH showing the exchange of hydrogen at C5 with solvent deuterium after the addition of wild-type LigU or one of the mutant enzymes: (a) wild-type LigU (1.0 nM) before the addition of enzyme (black), after 10 min (red), and after 160 min (blue), (b) K66M (111 nM) before the addition of enzyme (black) and after 60 min (red), (c) C100S (11 nM) before the addition of enzyme (black), after 25 min (red), and after 350 min (blue), and (d) H281N (113 nM) before the addition of enzyme (black), after 20 min (red), and after 320 min (blue). Additional details are provided in the text.

4.3.4. Three-Dimensional Structure of LigU

The structure of wild-type LigU was determined to a resolution of 2.0 Å. LigU crystallized in space group $P2_12_12_1$ and contains two monomers per asymmetric unit with an RMSD of 0.7 Å over the 344 fully populated residue pairs (**Figure 40**). Each of the two subunits is composed of an N-terminal domain (residues 1−171) and a C-terminal domain (residues 172−357). The two monomers interact with one another near the N-terminal domain via residues that originate from the last β-strand (Ala-347−Phe-356) from the C-terminal domain. The protein fold of each monomer consists of 16 β-strands and 10 α-helices (**Figure 41**). Each of the two domains has a central α-helix that is wrapped by antiparallel β-strands that form a β-barrel.

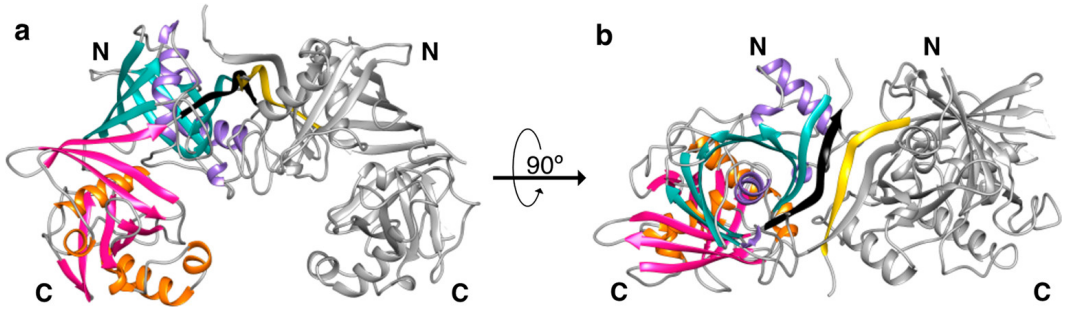


Figure 39. (a) Ribbon model of the homodimer of wild-type LigU (PDB entry 6P3J) at 2.0 Å resolution. One of the two monomers is colored silver, and the second is colored purple for the N-terminal α-helices, sea green for the N-terminal β-strands, orange for the C-terminal α-helices, and pink for the C-terminal β-strands. β-Strand 16 from chain A is colored black, and β-strand 16 from chain B is colored yellow. (b) Ribbon model rotated 90° to show the interactions from the last β-strands of each monomer. The figures were made using Chimera (37).

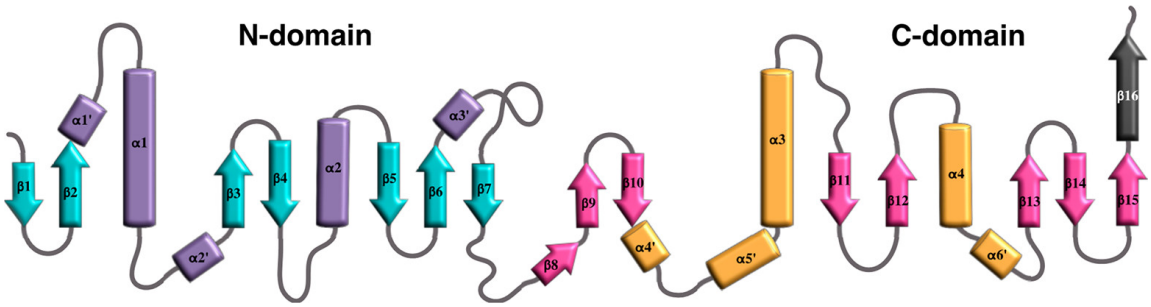


Figure 40. Topology model for the two domains of LigU chain A where the α-helices and β-strands of the N-domain are depicted as purple cylinders and sea green arrows, respectively. The α-helices and β-strands of the C-terminal domain are depicted as orange cylinders and pink arrows, respectively. β-Strand 16 is shown as a black arrow. The figure was made using Affinity Designer 1.6.1.

The N- and C-terminal domains that compose each monomeric unit are structurally similar to one another with an average RMSD of 2.1 Å over the 118 fully populated residue pairs, but the level of amino acid sequence identity between the two domains is only 10% (**Figures 42** and **43**). This value was calculated using chain A where the N-terminal domain (residues 8−171) was compared with residues from the C-terminal domain (residues 172−345). In addition to the central α-helices, α2 and α4, there are also external α-helices in each domain (**Figure 41**). These external α-helices and nearby loops create a solvent channel that leads to the active site. The active site is located at the junction of these two domains where the internal α-helices contact one another at their N- terminal ends. The putative catalytic residues, Lys-66 and Cys- 100, are located on loops within the N-terminal domain, whereas other highly conserved residues (Lys-257 and His- 281) that interact with the substrate (see below) are from the C-terminal domain.

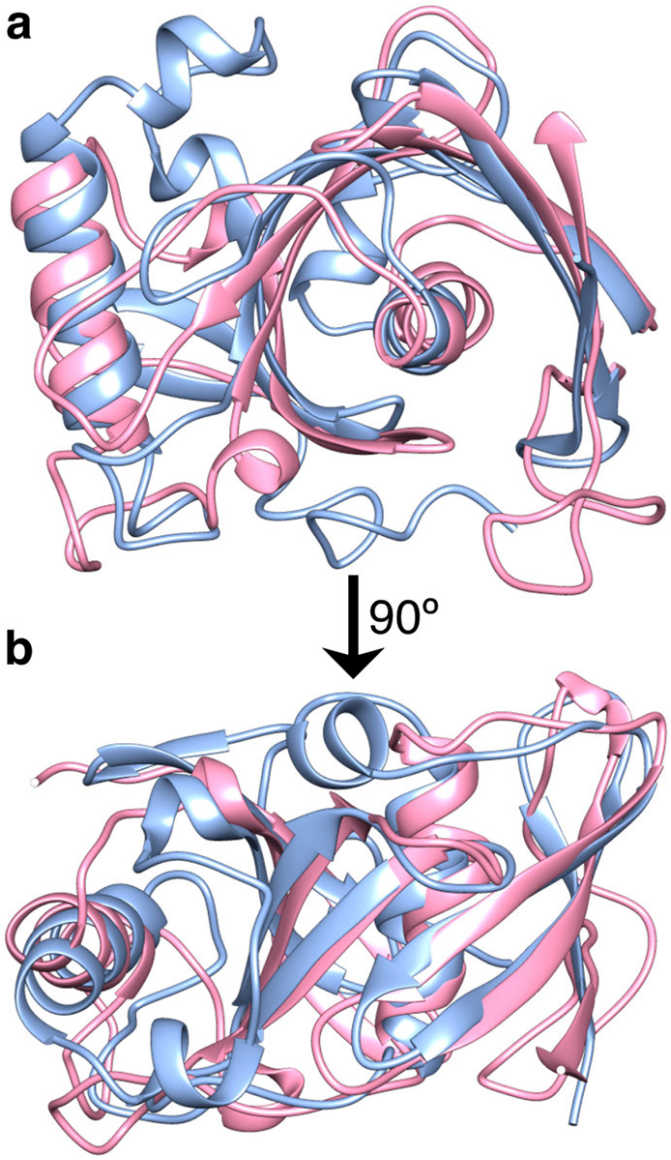


Figure 41. (a) Structural alignment of the N- and C-terminal domains of LigU. The N-terminal domain (residues 8−171) is colored pink, and the C-terminal domain (residues 172−345) is colored light blue from chain A. (b) Ribbon model rotated 90°. The figures were made using Chimera (37) and Affinity Designer 1.6.1.

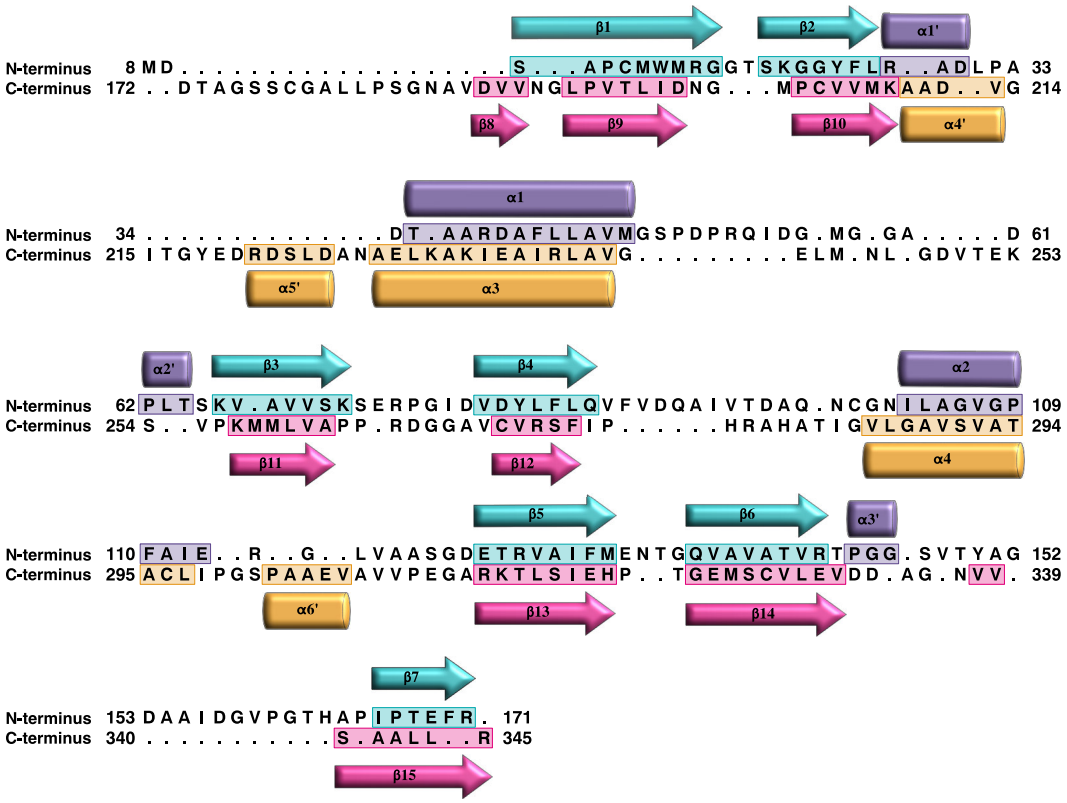


Figure 42. Sequence alignment of the N-terminal (residues 8-171) and C-terminal (residues 172-345) domains of wild-type LigU, chain A (PDB entry: 6P3J and UniProt entry: Q0KJL4). β-Strand 16 (Ala-347 through Phe-356) was omitted from the alignment calculation. The sequence alignment was created using Chimera 1.12 (37) and Affinity Designer 1.6.1.

4.3.5. Structure of the C100S Mutant of LigU

The structure of the C100S mutant of LigU was determined to a resolution of 1.9 Å by molecular replacement (PDB entry 6P3K). This mutant crystallized in space group C2 with two monomers per asymmetric unit and the following unit cell dimensions: a = 97.42 Å, b = 142.33 Å, and c = 60.35 Å (**Table 7**). The RMSD between wild-type LigU and C100S is 0.5 Å for chain A over the 351 fully populated residue pairs. There are no significant changes between the wild-type and mutant enzymes except for the orientation of the hydroxyl group in C100S, relative to the thiol group in the wild-type enzyme. The thiol of Cys-100 is hydrogen bonded to Lys-66 at a distance of 3.2 Å in the wild-type enzyme, whereas the distance between Ser-100 and Lys- 66 is 5.0 Å in the mutant enzyme (**Figure 44**).

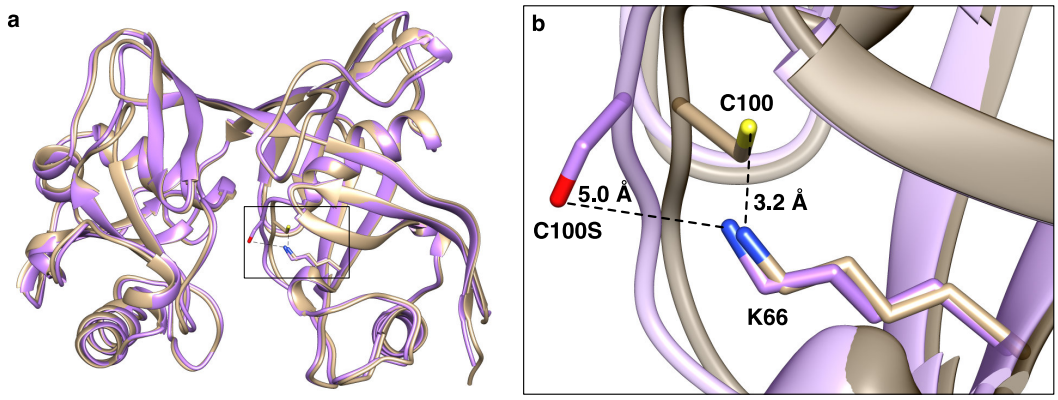


Figure 43. (a) Superposition of wild-type LigU and the C100S mutant. The wild-type enzyme is colored tan, the C100S mutant is colored in purple. (b) Inset of the active site showing the orientation change of the hydroxyl group in C100S, relative to the thiol group in the wild-type enzyme.

4.3.6. Structure of the Ligand-Bound Complex with the K66M Mutant of LigU

The K66M mutant of LigU was crystallized in the presence of 10 mM OMA (2). The protein crystallized in space group $P2_12_12$ with four monomers per asymmetric unit and the following unit cell dimensions: a = 84.88 Å, b = 134.00 Å, and c = 168.37 Å (**Table 7**). The simulated annealing omit map contoured to 3σ has defined electron density for the orientation of all functional groups of the substrate in each monomer (**Figure 45a**). Even though the electron density shows the high occupancy of the ligand, a resolution of 1.6 Å does not provide discrete peaks for individual atoms. Thus, we were unable to distinguish the double bond location and have modeled the substrate, OMA, in the active site. The modeling is supported by both functional data showing K66M has no observable activity and the equilibrium constant of 0.3 favoring OMA over the product KCH in solution (8). In the ligand-bound complex, there is a large network of hydrogen bond interactions with the backbone and side chains of nearby residues to the carboxylate and carbonyl groups of the substrate (**Figure 45b**). The C1 carboxylate group is within hydrogen bonding distance of a water molecule and the amide backbone of Gly-288, Leu-287, Val-286, and Gly-101. Additional hydrogen bonds to the C2 carbonyl oxygen are formed via interactions with the backbone nitrogens of Asn-102, Val-286, and Gly-101. The C6 carboxylate forms hydrogen bonds with two water molecules, Gln-98, and Lys-257. Lastly, the C7 carboxylate forms hydrogen bonds with a water molecule, His-281, Asn-102, and Ser-21. Overall, the substrate forms a substantial number of hydrogen bonds within the active site.

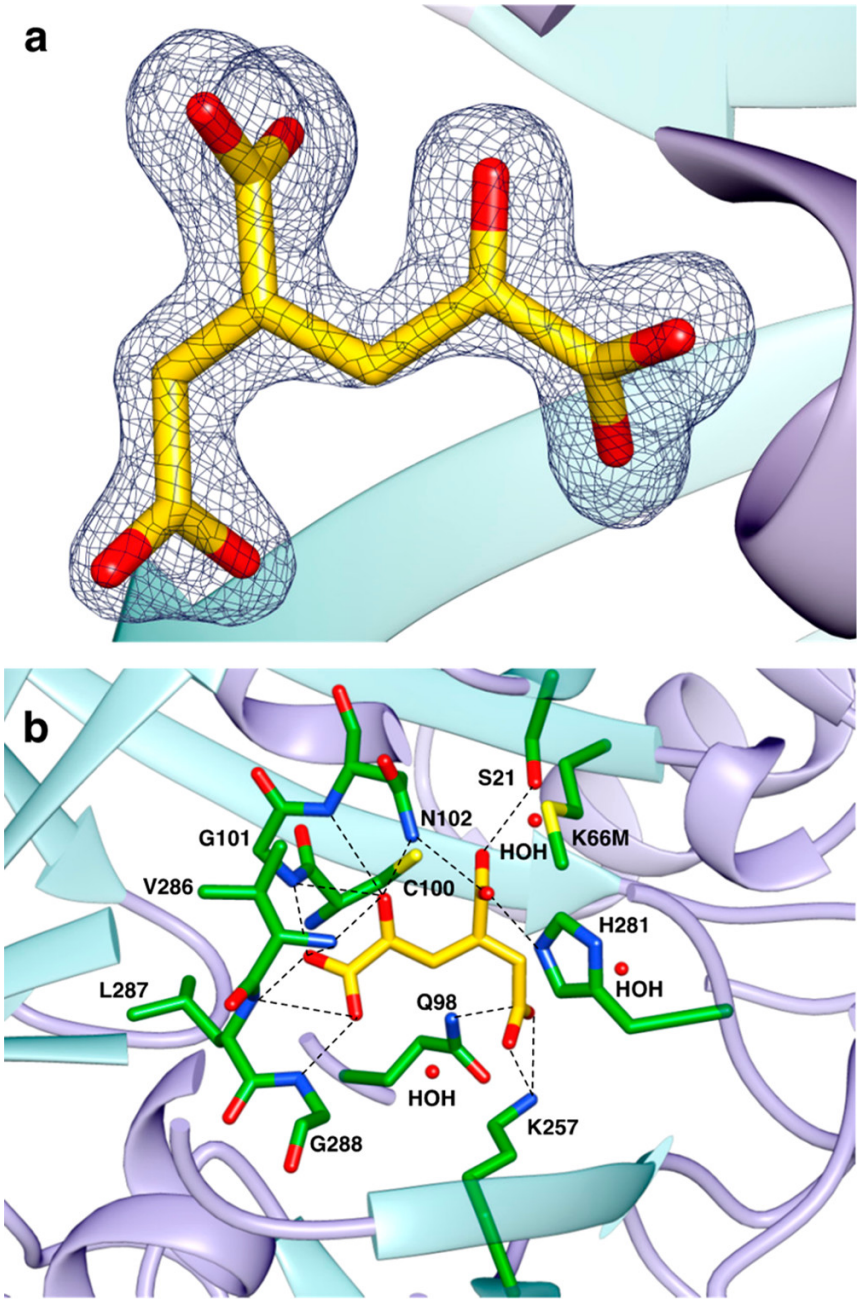


Figure 44. Model representations of LigU K66M co-crystallized with the substrate (PDB entry 6P3H, chain A) where the ligand is colored yellow. (a) Simulated annealing omit $mF_o - DF_c$ map (black mesh) contoured at 3.0 σ showing unbiased electron density for substrate OMA (2) in the active site. (b) Model showing hydrogen bonding and ionic interactions between the substrate and protein.

4.3.7. Conformational Changes in LigU Complexed with the Substrate

A conformational change between apo and ligand-bound complexes is observed. Chains A and B from wild-type LigU were aligned with chains A and C, respectively, of the K66M mutant. The RMSD across 695 α-carbon atoms of the dimer was 3.8 Å. The majority of this difference comes from the C-terminal domain, specifically α-helices α5′ (Arg- 221−Asp-225) and α3 (Ala-228−Val-241). Arg-221 from α5′ changes orientation to be more fully extended in the active site channel to occlude the opening with movements of 6.3−7.0 Å in each subunit (**Figure 46a**). Arg-279 on the loop after β-strand 12 also moves inward, helping to block the channel by moving distances of 5.2−7.9 Å in each subunit. The apo form of LigU has a large solvent channel for binding of the substrate. In the bound structure, the solvent channel disappears as the α-helices and loops from the flexible C-terminus move inward to give a closed conformation (**Figure 46b,c**). An overlay of the monomeric units is shown in **Figure 47**. The conformational differences between wild-type LigU and the K66M mutant are most likely due to the presence of the substrate bound in the active site of the K66M mutant rather than any differences caused by the mutation of Lys-66 to methionine.

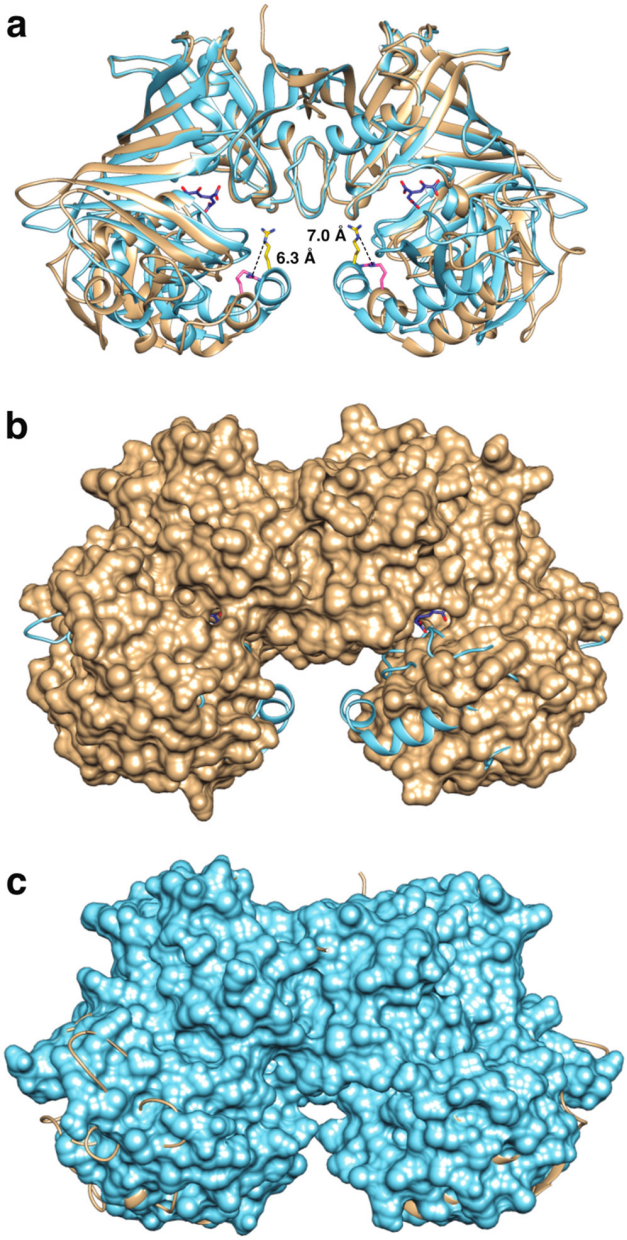


Figure 45. Model representations of LigU K66M co-crystallized with the substrate (PDB entry 6P3H, chain A) where the ligand is colored yellow. (a) Simulated annealing omit $mF_o - DF_c$ map (black mesh) contoured at 3.0 σ showing unbiased electron density for substrate OMA (2) in the active site. (b) Model showing hydrogen bonding and ionic interactions between the substrate and protein.

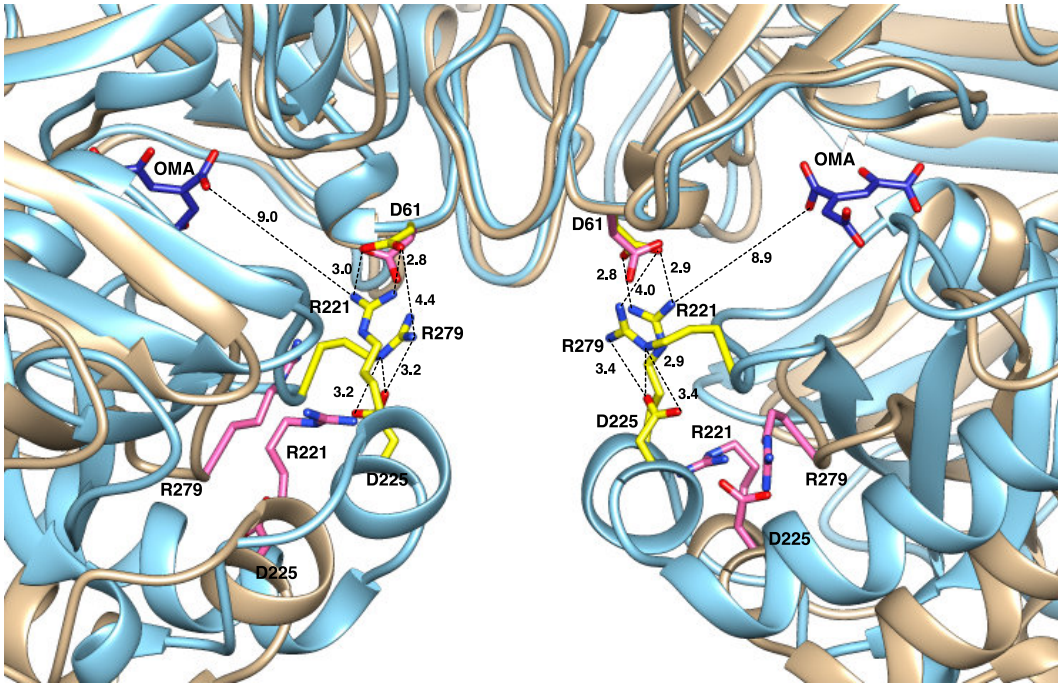


Figure 46. Superposition of wild-type LigU (chains A and B) and the K66M mutant (chains A and C) homodimers. The wild-type enzyme is colored tan with the residues colored pink, and the K66M mutant is colored sky blue with OMA colored medium blue and the residues colored yellow. The residues most responsible for closing the solvent channel are shown with ionic interactions measured in angstrom. They are Asp-61, Arg-221, Arg-279, and Asp-225.

4.4. Discussion

4.4.1. Chemical Reaction Mechanism

The mechanism of action for the 1,3-allylic isomerization of OMA (2) to KCH (3) catalyzed by LigU was addressed using NMR spectroscopy, site-directed mutagenesis, and X-ray crystallography. Six residues contained within the active site of LigU were mutated to residues that cannot function as general acid/base catalysts for the transfer of protons between C3 and C5 of OMA and KCH. Of the mutants made for this investigation, the most significant reductions in catalytic activity were obtained for Lys-66, Cys-100, Lys-257, and His-281 (**Table 8**). In the substrate-bound complex of the K66M mutant, the ε-amino group of Lys-257 makes a hydrogen bond to the C6 carboxylate and is positioned too far from either C3 or C5 to be involved in a proton transfer (**Figure 45b**). Mutants of the three remaining residues (K66, C100, and H281) were subsequently interrogated for their ability to catalyze hydrogen−deuterium exchange reactions between KCH and the solvent.

The wild-type enzyme catalyzes the exchange of one of the two prochiral hydrogens at C5 of KCH with the solvent and also facilitates the exchange of the proton at C3 via the transient formation of OMA (**Figure 39a**). The exchange pattern for the H281N mutant was identical to that observed for the wild-type enzyme. The K66M mutant, however, was unable to catalyze the exchange of the hydrogen at C5 with deuterium from the solvent (**Figure 39b**). The C100S mutant was functionally able to catalyze hydrogen−deuterium exchange at C5 but was unable to facilitate the exchange

at C3 (**Figure 39c**). These results support the conclusion that Cys-100 functions to catalyze the removal of a proton from C3 and that Lys-66 functions to donate a proton to C5 during the conversion of OMA to KCH.

The stereochemical outcomes for the proton transfers at C3 and C5 of OMA (2) were determined by the positioning of the active site residues in the three-dimensional structure of LigU complexed with the substrate and isotopic labeling of the product KCH (3). When the LigU-catalyzed reaction is conducted in D_2O, deuterium is found exclusively in the pro-*S* position at C5 (**Figure 38**). This result is fully consistent with the presumptive positioning of Lys-66 in the active site of the K66M mutant of LigU where the terminal methyl group of the methionine is 3.8 Å from the *si*-face of the double bond at C5 (**Figure 48**). The stereochemical outcome for the proton transfer at C3 could not be determined directly, but the only residue in the active site that is positioned to abstract a proton from C3 of OMA is Cys-100. The thiol group of Cys-100 is 3.4 Å from the pro-*S* hydrogen at C3. Because Lys-66 and Cys-100 are on the same face of the bound substrate, the 1,3-allylic isomerization catalyzed by LigU is suprafacial (**Scheme 15**). In the proposed reaction mechanism, the thiolate of Cys-100 abstracts the pro-*S* hydrogen from C3 of OMA in a reaction that is stabilized by charge delocalization to the C2 carbonyl. KCH is formed by the protonation of C5 via proton transfer from the side chain of Lys-66.

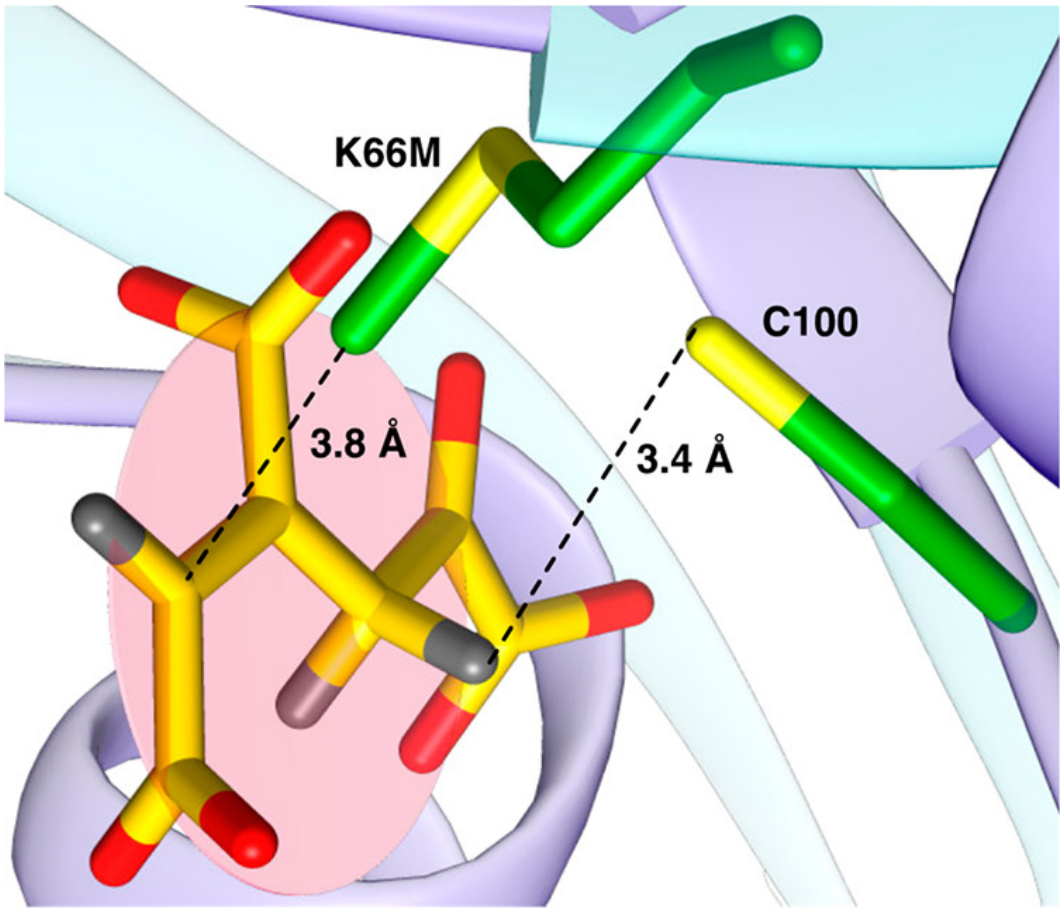
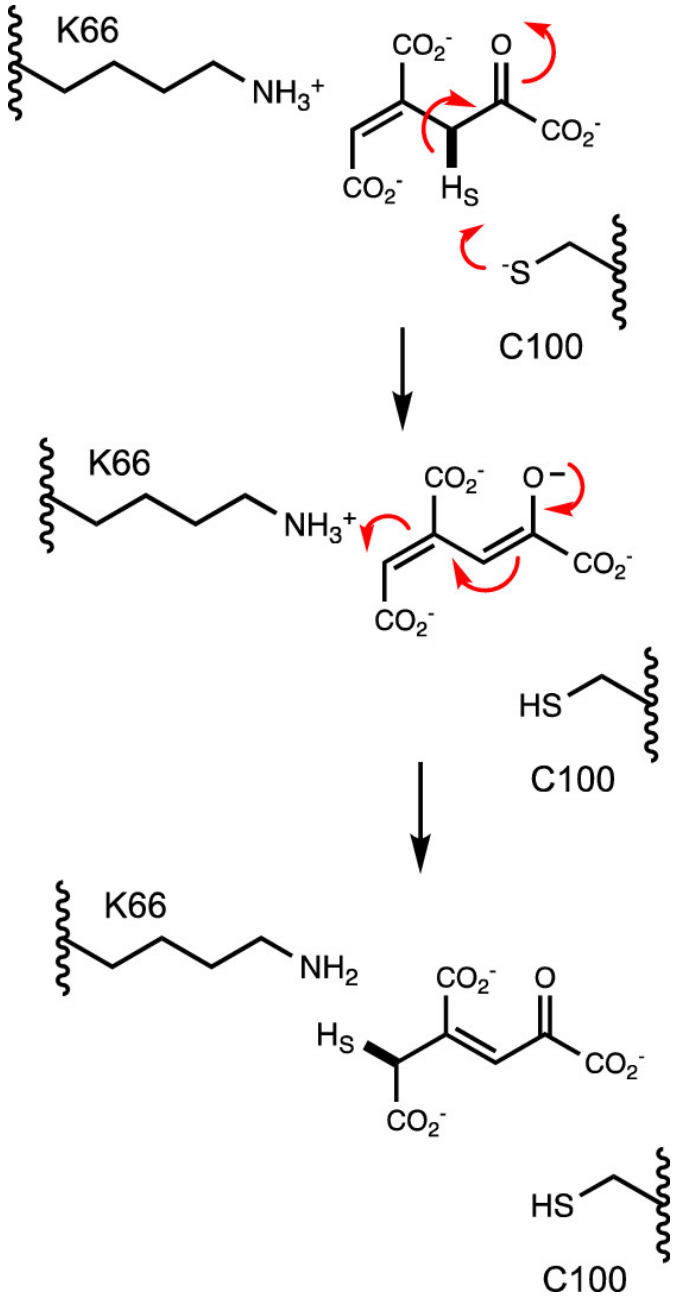


Figure 47. Model representation of the K66M mutant of LigU co-crystallized with OMA (PDB entry 6P3H, chain A) where OMA is colored yellow.

Scheme 15. Proposed Reaction Mechanism of LigU.



4.4.2. Comparisons between LigU and PrpF

PrpF and aconitate isomerase catalyze the allylic rearrangement of *cis*-aconitate to *trans*-aconitate (12, 38, 39). PrpF from *Shewanella oneidensis* MR-1 has also been reported to catalyze the isomerization of (*S*)-4-methyl-*cis*-aconitate to 2-methyl-*cis*-aconitate (13). Klinman and Rose demonstrated that aconitate isomerase from *Pseudomonas putida* catalyzes the 1,3-allylic isomerization of *cis*-aconitate via the abstraction of the pro-*S* hydrogen from C4 and suprafacial donation of a proton to C2 (39). These results are fully in accord with the results obtained with LigU (**Scheme 15**). The three-dimensional structure of PrpF has been determined with *trans*-aconitate bound in the active site (PDB entry 2PW0) (12). A structural comparison of the bound ligands in the active sites of these two enzymes is illustrated in **Figures 49** and **50**. The C1 carboxylate group of *trans*-aconitate is positioned like the C7 carboxylate group of OMA and interacts with the side chain of His-317. The C5 carboxylate of *trans*-aconitate overlays quite closely with the C6 carboxylate of OMA. The side chains of the proposed catalytic residues (Lys-73 and Cys-107) are aligned quite closely with those for LigU and are positioned on the same face of the ligand. The C1 and C2 functional groups of OMA are situated deeper within the active site pocket where backbone interactions occur.

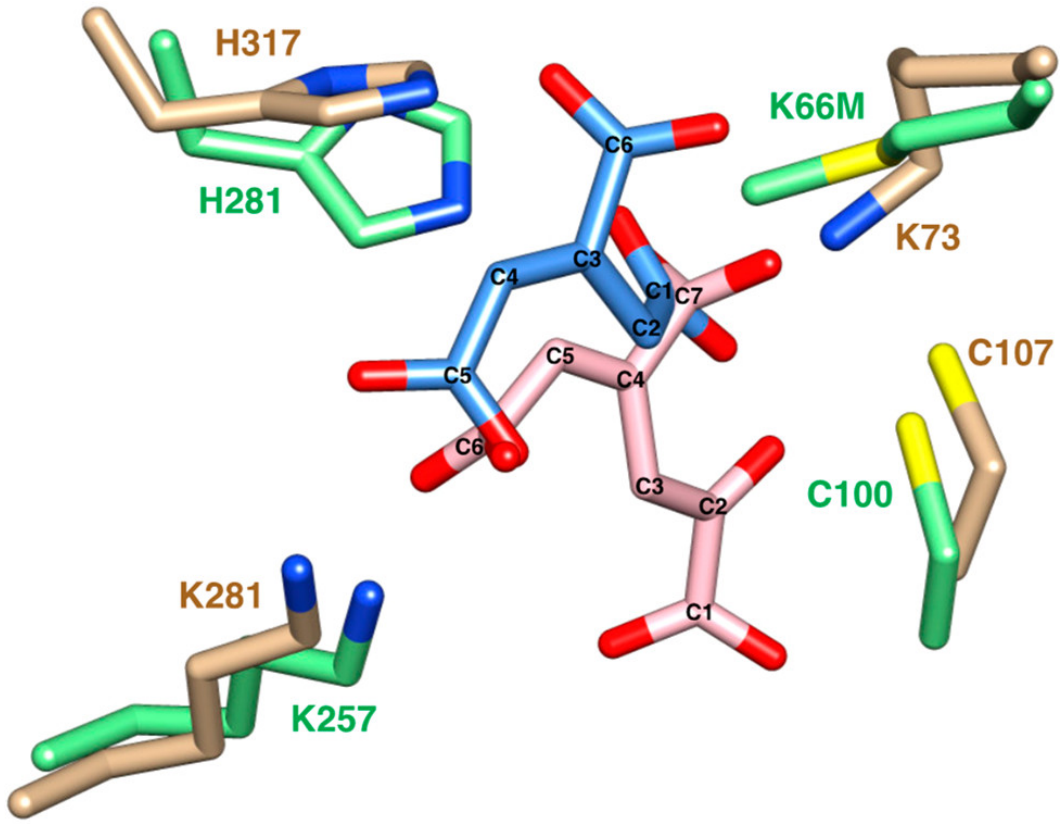


Figure 48. Superimposed models of PrpF (tan) complexed with *trans*-aconitate (blue) (PDB entry 2PW0, chain a) and LigU (green) complexed with ligand (OMA or KCH) (pink) (PDB entry 6P3H, chain a).

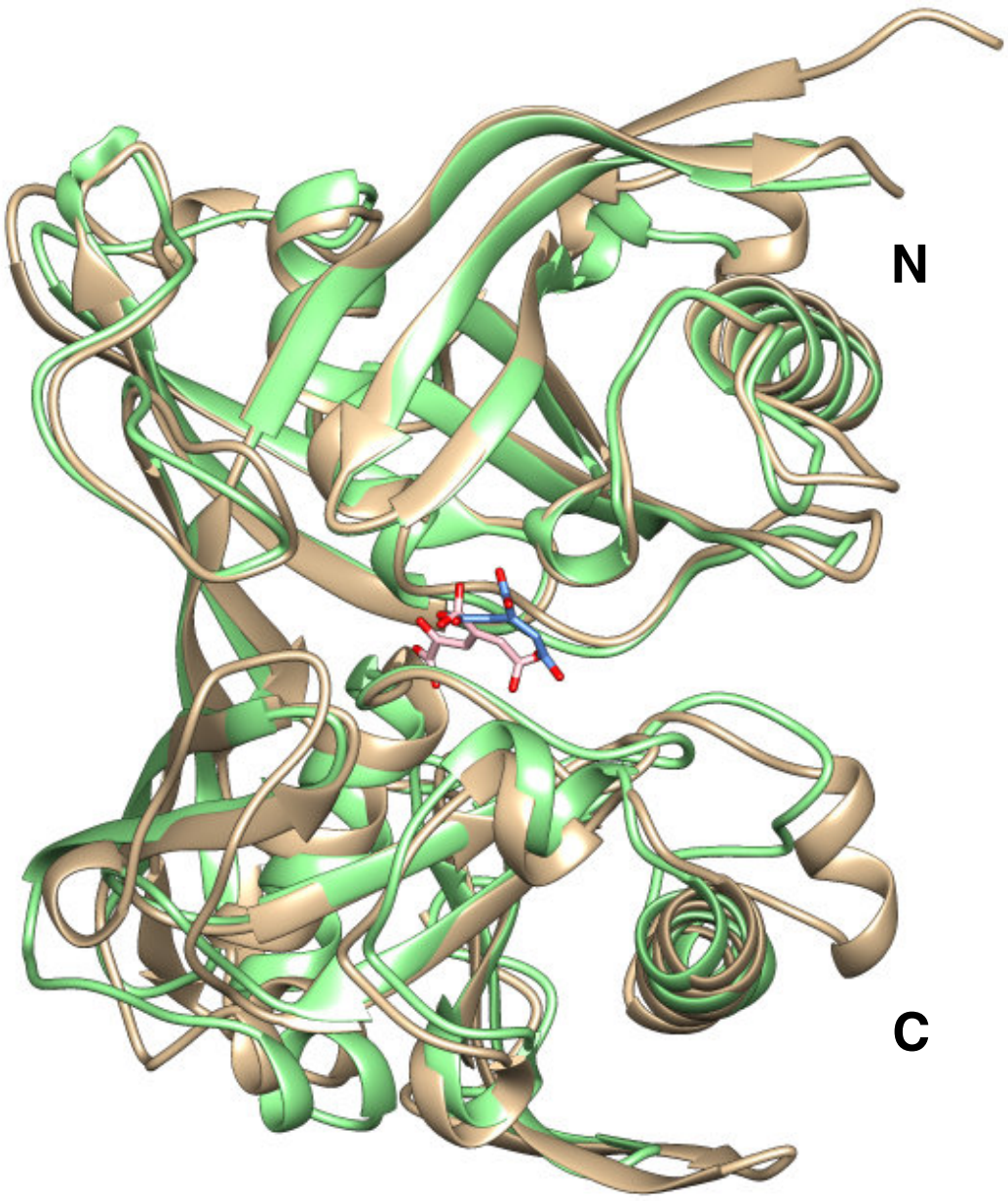


Figure 49. Superimposed ribbon model of the monomers of LigU K66M and PrpF with ligand bound. LigU K66M (PDB entry: 6P3H, chain A) is depicted in green with OMA colored pink, and PrpF (PDB entry: 2PW0, chain A) (6) is depicted in tan with trans- aconitate colored blue.

4.5. Conclusions

We have determined the three-dimensional structure of LigU in the presence and absence of a bound substrate. Solvent−deuterium exchange experiments and kinetic analysis of site-directed mutants have demonstrated that LigU catalyzes the isomerization of OMA (2) to KCH (3) via the abstraction of the pro-*S* hydrogen from C3 by Cys-100 and donation of a proton to C5 to the same face of the enolate intermediate by Lys-66 (**Scheme 15**).

4.6. References

1. Becker, J., and Wittmann, C. (2019) A field of dreams: Lignin valorization into chemicals, materials, fuels, and health-care products. *Biotechnol. Adv.,* doi.org/10.1016/j.biotechadv.2019.02.016
2. Chaves, J. E., Presley, G. N., and Michener, J. K. (2019) Modular Engineering of Biomass Degradation Pathways. *Processes, 7,* 230.
3. Cragg, S. M., Beckham, G. T., Bruce, N. C., Bugg, T. D. H., Distel, D. L., Dupree, P., Etxabe, A. G., Goodell, B. S., Jellison, J., McGeehan, J. E., McQueen-Mason, S. J., Schnorr, K., Walton, P. H., Watts, J. E. M., and Zimmer, M. (2015) Lignocellulose degradation mechanism across the Tree of Life. *Curr. Opin. Chem. Biol. 29,* 108-119.
4. Lubbers, R. J. M., Dilokpimol, A., Visser, J., Mäkelä M. R., Hildén, K. S., and Vries, R. P. (2019) A comparison between the homocyclic aromatic metabolic pathways from plant-derived compounds by bacteria and fungi. *Biotech Adv.,* doi.org/10.1016/j.biotechadv.2019.05.002
5. Noda, Y., Nishikawa, S., Shiozuka, K. I., Kadokura, H., Nakajima, H., Yoda, K., Katayama, Y., Morohoshi, N., Haraguchi, T., and Yamasaki, M. (1990) Molecular Cloning of the Protocatechuate 4,5- Dioxygenase Genes of *Pseudomonas paucimobilis*. *J. Bacteriol. 172,* 2704−2709.
6. Masai, E., Momose, K., Hara, H., Nishikawa, S., Katayama, Y., and Fukuda, M. (2000) Genetic and Biochemical Characterization of 4-Carboxy-2-

Hydroxymuconate-6-Semialdehyde Dehydrogenase and Its Role in the Protocatechuate 4,5-Cleavage Pathway in *Sphingomonas paucimobilis* SYK-6. *J. Bacteriol. 182,* 6651−6658.

7. Hobbs, M. E., Malashkevich, V., Williams, H. J., Xu, C., Sauder, J. M., Burley, S. K., Almo, S. C., and Raushel, F. M. (2012) Structure and Catalytic Mechanism of LigI: Insight into the Amidohydrolase Enzymes of cog3618 and Lignin Degradation. *Biochemistry 51,* 3497− 3507.
8. Hogancamp, T. N., and Raushel, F. M. (2018) Functional Annotation of LigU as a 1,3-Allylic Isomerase during the Degradation of Lignin in the Protocatechuate 4,5-Cleavage Pathway from the Soil Bacterium *Sphingobium* sp. SYK-6. *Biochemistry 57,* 2837−2845.
9. Hogancamp, T. N., Mabanglo, M. F., and Raushel, F. M. (2018) Structure and Reaction Mechanism of the LigJ Hydratase: An Enzyme Critical for the Bacterial Degradation of Lignin in the Protocatechuate 4,5-Cleavage Pathway. *Biochemistry 57,* 5841-5850.
10. Hara, H., Masai, E., Miyauchi, K., Katayama, Y., and Fukuda, M. (2003) Characterization of the 4-Carboxy-4-Hydroxy-2-Oxoadipate Aldolase Gene and Operon Structure of the Protocatechuate 4,5-Cleavage Pathway Genes in *Sphingomonas paucimobilis* SYK-6. *J. Bacteriol. 185,* 41-50.
11. Wang, W., Mazurkewich, S., Kimber, M. S., and Seah, S. Y. K. (2010) Structural and Kinetic Characterization of 4-Hydroxy-4-methyl-2-oxoglutarate/4-Carboxy-4-hydroxy-2-oxoadipate Aldolase, a Protocatechuate Degradation Enzyme

Evolutionarily Convergent with the HpaI and DmpG Pyruvate Aldolases. *J. Biol. Chem. 47,* 36608-36615.

12. Garvey, G. S., Rocco, C. J., Escalante-Semerena, J. C., and Rayment, I. (2007) The three-dimensional crystal structure of the PrpF protein of *Shewanella oneidensis* complexed with *trans*-aconitate: Insights into its biological function. *Protein Sci. 16,* 1274−1284.
13. Rocco, C. J., Wetterhorn, K. M., Garvey, G. S., Rayment, I., and Escalante-Semerena, J. C. (2017) The PrpF protein of *Shewanella oneidensis* MR-1 catalyzes the isomerization of 2-methyl-*cis*-aconitate during the catabolism of propionate via the AcnD-dependent 2- methylcitric acid cycle. *PLoS One 12,* e0188130.
14. Cirilli, M., Zheng, R., Scapin, G., and Blanchard, J. S. (1998) Structural Symmetry: The Three-Dimensional Structure of *Haemophilus Influenzae* Diaminopimelate Epimerase. *Biochemistry 37,* 16452-16458.
15. Liu, L., Iwata, K., Yohda, M., and Miki, K. (2002) Structural insight into gene duplication, gene fusion, and domain swapping in the evolution of PLP-independent amino acid racemases. *FEBS Letters 528,* 114-118.
16. Llyod, A. J., Huyton, T., Turkenburg, J., and Roper, D. I. (2004) Refinement of *Haemophilus influenzae* diaminopimelic acid epimerase (DapF) at 1.75 Å resolution suggests a mechanism for stereocontrol during catalysis. *Acta Cryst. D60,* 397-400.
17. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: A Software Environment for

Integrated Models of Biomolecular Interaction Networks. *Genome Res. 13*, 2498– 2504.

18. Gerlt, J. A., Allen, K. N., Almo, S. C., Armstrong, R. N., Babbitt, P. C., Cronan, J. E., Dunaway-Mariano, D., Imker, H. J., Jacobson, M. P., Minor, W., Poulter, C. D., Raushel, F. M., Sali, A., Shoichet, B. K., and Sweedler, J. V. (2011) The Enzyme Function Initiative. *Biochemistry 50*, 9950– 9962.
19. Gerlt, J. A., Bouvier, J. T., Davidson, D. B., Imker, H. J., Sadkhin, B., Slater, D. R., and Whalen, K. L. (2015) Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta, Proteins Proteomics 1854*, 1019– 1037.
20. The UniProt Consortium. (2018) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research 47,* D506-D515.
21. Kersten, P. J., Dagley, S., Whittaker, J. W., Arciero, D. M., and Lipscomb, J. D. (1982) 2-Pyrone-4,6-Dicarboxylic Acid, a Catabolite of Gallic Acids in *Pseudomonas* Species. *J. Bacteriol. 152,* 1154−1162.
22. Liu, M., Mao, X., Ye, C., Huang, H., Nicholson, J. K., and Lindon, J. C. (1998) Improved WATERGATE Pulse Sequence for Solvent Suppression in NMR Spectroscopy. *J. Magn. Reson. 132,* 125−129.
23. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol 276,* 307-326.

24. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr 67,* 271-281.
25. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution?. *Acta Cryst. D69,* 1204-1214.
26. Adams, P. D., Afonine, P. V., Bunkcózi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst. D66,* 213-221.
27. Velarde, M., Macieira, S., Hilberg, M., Bröker, G., Tu, S., Golding, B. T., Pierik, A. J., Buckel, W., and Messerschmidt, A. (2009) Crystal Structure and Putative Mechanism of 3-Methylitaconate-Δ-isomerase from *Eubacterium barkeri. J. Mol. Biol. 391,* 609−620.
28. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Cryst. 40,* 658-674.
29. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Cryst. D66,* 486-501.
30. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P.

D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Cryst. D68,* 352-367.

31. Lebedev, A. A., Young, P., Isupov, M. N., Moroz, O. V., Vagin, A. A., and Murshudov, G. N. (2012) JLigand: a graphical tool for the CCP4 template-restraint library. *Acta Cryst. D68,* 431-440.
32. Moriarty, N. W., Draizen, E. J., and Adams, P. D. (2017) An editor for the generation and customization of geometry restraints. *Acta Cryst. D73,* 123-130.
33. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Cryst. D66,* 12-21.
34. Meng, E. C., Pettersen, E. F., Couch, G. S., Huang, C. C., and Ferrin, T. E. (2006) Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinformatics 7,* 339.
35. Alberty, R. A., Miller, W. G., and Fisher, H. F. (1957) Studies of the Enzyme Fumarase. VI. Study of the Incorporation of Deuterium into L-Malate during the Reaction in Deuterium Oxide. *J. Am. Chem. Soc. 15,* 3973-3977.
36. Blanchard, J. S., and Cleland, W. W. (1980) Use of Isotope Effects To Deduce the Chemical Mechanism of Fumarase. *Biochemistry 19,* 4506-4513.
37. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem. 13,* 1605-12.

38. Klinman, J. P. and Rose, I. A. (1971) Purification and Kinetic Properties of Aconitate Isomerase from *Pseudomonas putida*. *Biochemistry 10,* 2253-2258.

39. Klinman, J. P. and Rose, I. A. (1971) Mechanism of the Aconitate Isomerase Reaction. *Biochemistry 10,* 2259-2266.

5. CONCLUSIONS

5.1. LigU Isomerase Function

Sphingobium sp. SYK-6 is a Gram-negative soil bacterium that contributes to the degradation of lignin. Lignin provides structural support and protection to plants as a complex aromatic heteropolymer. The lignin degradation pathway of guaiacyl moieties leads to the intermediate, protocatechuate (PCA), which is further degraded via the 4,5-cleavage pathway in which PCA is ultimately metabolized to pyruvate and oxaloacetate. In this pathway, LigI has been shown to catalyze the hydrolysis of 2-pyrone-4,6-dicarboxylate to (4*E*)-oxalomesaconate (OMA). Here we have demonstrated, using ^{1}H and ^{13}C nuclear magnetic resonance spectroscopy, that LigU from *Novosphingobium* sp. strain KA1 catalyzes the isomerization of the double bond between C4 and C5 in (4*E*)-OMA to (3*Z*)-2-keto-4-carboxy-3-hexenedioate (KCH), where the double bond has migrated to be between C3 and C4 via a 1,3-allylic isomerization. LigU is most closely related in amino acid sequence to methylaconitate isomerase (PrpF) from *Shewanella oneidensis* and methylitaconate-Δ-isomerase (Mii) from *Eubacterium barkeri*. The kinetic constants for the isomerization of OMA to KCH by LigU at pH 8.0 were determined to be 1300 ± 120 s^{-1} and (7.7 ± 1.5) × 10^{6} M^{-1} s^{-1} for k_{cat} and k_{cat}/K_{m}, respectively. We have also shown that the product of the LigU-catalyzed reaction is the preferred substrate for the LigJ hydratase. In this reaction, LigJ catalyzes the hydration of KCH to 4-carboxy-4-hydroxy-2-oxoadipate (CHA).

5.2. LigJ Hydratase Proposed Mechanism

LigJ from the soil bacterium *Sphingobium* sp. SYK-6 catalyzes the reversible hydration of KCH to CHA in the degradation of lignin in the PCA 4,5-cleavage pathway. LigJ is a member of the amidohydrolase superfamily and an enzyme in cog2159. The three-dimensional crystal structure of wild-type LigJ was determined in the presence [Protein Data Bank (PDB) entry 6DXQ] and absence of the product CHA (PDB entry 6DWV). The protein folds as a distorted $(\beta/\alpha)_8$-barrel, and a single zinc ion is bound in the active site at the C-terminal end of the central β-barrel. The product CHA is ligated to the zinc ion in the active site via the displacement of a single water molecule from the coordination shell of the metal center in LigJ. The product-bound structure reveals that the enzyme catalyzes the hydration of KCH with the formation of a chiral center at C4 with *S* stereochemistry. The E284Q mutant was unable to catalyze the hydration of KCH to CHA, and the structure of this mutant was determined in the presence of the substrate KCH (PDB entry 6DXS). On the basis of the structure of LigJ in the presence of KCH and CHA, it is proposed that the side chain carboxylate of Glu-284 functions as a general base in the abstraction of a proton from a bound water molecule for nucleophilic attack at C4 of the substrate. The reaction is facilitated by the delocalization of the negative charge to the metal center via the carbonyl group at C2 of the substrate. C3 of the substrate is subsequently protonated by Glu-284 functioning as a general acid. The overall reaction occurs by the *syn*-addition of water to the double bond between C4 and C3 of the substrate KCH. The kinetic constants for the hydration of KCH to CHA by LigJ at pH 8.0 are 25 s^{-1} (k_{cat}) and 2.6×10^6 M^{-1} s^{-1} (k_{cat}/K_m).

5.3. LigU Isomerase Proposed Mechanism

LigU from *Novosphingobium* sp. strain KA1 catalyzes the isomerization of OMA to KCH as part of the PCA 4,5-cleavage pathway during the degradation of lignin. The three-dimensional structure of the apo form of the wild-type enzyme was determined by X-ray crystallography, and the structure of the K66M mutant enzyme was determined in the presence of the substrate OMA. LigU is a homodimer requiring no cofactors or metal ions with a diaminopimelate epimerase structural fold, consisting of two domains with similar topologies. Each domain has a central α-helix surrounded by a β-barrel composed of antiparallel β-strands. The active site is at the cleft of the two domains. ^{1}H nuclear magnetic resonance spectroscopy demonstrated that the enzyme catalyzes the exchange of the pro-*S* hydrogen at C5 of KCH with D_2O during the isomerization reaction. Solvent–deuterium exchange experiments demonstrated that mutation of Lys-66 eliminated the isotope exchange at C5 and that mutation of C100 abolished exchange at C3. The positioning of these two residues in the active site of LigU is consistent with a reaction mechanism that is initiated by the abstraction of the pro-*S* hydrogen at C3 of OMA by the thiolate anion of Cys-100 and the donation of a proton at C5 of the proposed enolate anion intermediate by the side chain of Lys-66 to form the product KCH. The 1,3-proton transfer is suprafacial.

5.4. LigK Aldolase Future Work

Future work is to elucidate the mechanism of action of LigK from *Sphingobium* sp. SYK-6 that catalyzes the retro-aldol cleavage of CHA to oxaloacetate and pyruvate.

LigK belongs to cog0684, termed as the Regulator of ribonuclease activity A (RraA) protein family. Structurally, these proteins have a αββα sandwich fold. Based on the proposed mechanism of action, it is characterized as a Class II pyruvate aldolase in where a divalent metal ion is utilized to stabilize an enolate intermediate. Recent studies have shown pyruvate aldolases have a significant rate enhancement in the presence of inorganic phosphate. Chemically how the phosphate anion is enhancing the rate is unclear; ongoing efforts are being made to co-crystallize phosphate inside the active site of LigK. In addition to phosphate, substrate-bound crystals are also being pursued by site-directed mutagenesis, substrate analogs, and X-ray crystallography. The following mutants are being constructed based on studies from homologous proteins: E197Q, D100N, D122N, R121N, T143A, T143S, K145Q, and K145M. By deciphering the mechanism of LigK, the biochemical characterization of the last four enzymes of the PCA 4,5-cleavage pathway, specifically LigI hydrolase, LigU isomerase, LigJ hydratase, and LigK aldolase will be known.