Exploring Beta-Lactones as an Unexpected Link Between Hydrocarbon and Natural Product Biosynthesis in Bacteria

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE UNIVERSITY OF MINNESOTA
BY

James Kurt Christenson

IN PARTIAL FULFILLMENT OF THE REQUIERMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Dr. Lawrence P. Wackett, Advisor

May 2017

Acknowledgements

I am wholly indebted to my adviser, Dr. Larry Wackett, for his sacrificial service to me through all of graduate school. I dare not count the number of hours that he has poured into my life, not only to develop me as a researcher, but as a person as well. I marvel at the freedom and trust he gave me to explore and learn even when I was so inexperienced. I wish him many successful years to come and hope many others benefit from his mentorship.

I would also like to thank Dr. Carrie Wilmot for her scientific expertise and patient guidance on my project. Thank you for always finding the time to read drafts and give advice even in the busiest of times. My doctoral work is far better because of your faithful support. Thank you also to all the other staff and faculty at the University of Minnesota, Dr. Michael Freeman, Dr. Romas Kazlauskas, Dr. Mikael Elias, and many others, who have given of their time to support my project and career development.

I also want to thank my lab mates for such a caring and collaborative environment. Thank you, Dr. Jack Richman, for your seemingly inexhaustible supply of organic chemistry knowledge. I doubt the discoveries contained in this work would have happened without his guidance. Thank you Dr. Kelly Aukema, for your endless answering of questions, both personal and lab related, and your servant's heart. Thank you, Dr. Tony Dodge, for always pointing me in the right direction when I had questions. Thank you, Matt Jensen, collaborator, friend, and partner, in crime, for your constant input, scientific acumen, and tireless patience during my barrages of questions and ideas. Thank you, Serina Robinson, for your expertise in bioinformatics and exuberant spirit. I look forward to all the discoveries you will make on the project in the years to come. Thank you, Tiffany Engel and Jennifer Neufeld, for helping me so faithfully in lab. My

graduate work would have taken much longer without your incredible support. I also want to thank all the past and present members of the Wackett lab for their support and kindness during my time here. Finally, I want to thank all those who have been part of our lunch group that has made life as a graduate student so much richer: Bob, Diego, Matt, Kelly, Adi, and so many others.

To my dear wife, Genna: thank you for your support and faithfulness during these last five years. There were seasons when you sacrificed many personal comforts for this work, and I will ever be grateful. Thank you, mom and dad, for making me who I am today. This work would not have been possible without loving, supportive parents who sacrificially gave to impart a love of earthly and heavenly knowledge into my life. I will always be in your debt. Thank you, Jonathan and Peter, for their efforts to keep me sane and social during my graduate work. Finally, thanks also to my Bible study group; their regular presence in my life was a constant reminder that there is more to living than a dissertation.

Dedication

I would like to dedicate this work to my wife, Genna; my family; and my friends; you all have so enriched and influenced my life during graduate school.

ABSTRACT

Recent decades have seen a great interest in using biological systems as sources for the diverse chemicals that are used in modern society. Many of these processes, such as the production of ethanol from corn starch, rely on microbes to consume a low value feedstock and produce a higher-value product. Making these microbial processes efficient and economical depends on a fundamental understanding of the complex network of biochemical pathways and the individual chemical reactions occurring in a given microbe.

The production of long-chain olefins (alkenes) from two fatty acids by diverse bacterial species has been well reported in literature and patented for the production of hydrocarbon compounds. A four-gene cluster, *oleABCD*, encodes the enzymes responsible for this conversion; however, only two of the four enzymatic steps had been characterized previously. Here, the third enzyme of the pathway, OleC, was characterized and found to convert β-hydroxy acids to β-lactones in an ATP-dependent reaction. OleC is the first β-lactone synthetase ever characterized, and sequence analysis suggests that the biosynthesis of other β-lactone natural products is carried out by OleC homologues. β-Lactone-containing natural products show great potential as antibiotic and anticancer compounds and are known for their potent inhibition of proteases and esterases. For example, the β-lactone salinosporamide from *Salinospora pacifica* is in phase II clinical trials for glioblastoma and multiple myeloma, and lipstatin (Orlistat) from *Streptomyces toxytricini* is the only over-the-counter, FDA-approved anti-obesity drug. Despite the β-lactone's use in modern medicine, no enzyme capable of synthesizing the β-lactone ring had previously ever been identified.

The final enzymatic step of olefin biosynthesis was found to be the decarboxylation of the β-lactone to an olefin by OleB. This enzymatic reaction was not known to exist in nature, making OleB the first β-lactone decarboxylase. Sequence alignments of OleB demonstrate it is a part of the α/β-hydrolase enzyme superfamily, and many OleB sequences have been erroneously

annotated as haloalkane dehalogenases. Studies on the mechanism of OleB reveal a similar mechanism to haloalkane dehalogenases that relies on the same Asp-His-Asp catalytic triad and proceeds *via* an acyl-enzyme intermediate.

Finally, the hydrophobic nature of the ole-pathway metabolites and the toxicity of the β-lactone intermediate were hypothesized to require a direct transfer between the active sites of Ole enzymes. As such, the four pathway enzymes, OleABCD, were recombinantly expressed in *Escherichia coli*, and OleBCD were found to form an enzyme complex after two affinity chromatography steps. Immunoblots of OleC from the native *X. campestris* on non-denaturing gels were found to migrate similarly to the recombinantly purified OleBCD complex and much slower than OleC by itself, suggesting formation of a complex in the native *X. campestris* host. Electron microscopy images showed distinct, circular assemblies of OleBCD between 24 and 40 nm in size with an average of 32 nm. OleA is thought to act as a soluble shuttle between the pathway's fatty acid precursors and the OleBCD complex. The complex then functions to efficiently traffic the hydrophobic metabolites and sequester the toxic β-lactone intermediate.

Table of Contents

List of Tables

List of Figures

CHAPTER 1 - Introductions

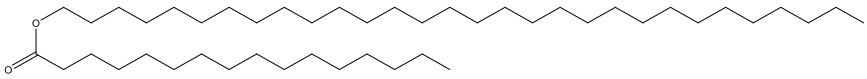
"It's a dangerous business, Frodo, going out your door. You step onto the road, and if you don't keep your feet, there's no knowing where you might be swept off to." – J.R.R. Tolkien (Lord of the Rings)

Hydrocarbons in Nature

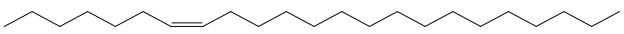
Every living organism requires a staggering diversity of chemicals to flourish. A cell takes advantage of the unique physical and chemical properties of every distinct molecule within its grasp to survive and adapt as a distinct entity in an environment. One class of compounds that is universally found in every living cell is the hydrocarbon substructure. Often covalently attached to other compounds, the hydrocarbon moiety is the primary component of cellular membranes and performs the critical role of delineating a cell from its environment. Hydrocarbons are composed solely of carbon and hydrogen and are generally unreactive, energy dense, and extremely hydrophobic. Because of these and other properties, hydrocarbons are suited for a variety of functions in nature from energy storage to cell to cell communication (Figure 1-1).

Many biochemical pathways dispersed throughout the tree of life utilize the reduced carbon created during membrane biosynthesis to generate a diversity of hydrocarbon-containing compounds for a wide variety of applications. For example, most plants will coat their leaves with hydrocarbons as a defense against moisture loss and protection against ultraviolet radiation.[1] Various insects such as fruit flies release hydrocarbons as pheromones *cis*-7-tricosene.[2] Bacteria are also known to produce hydrocarbons in addition to their membrane components, some of which are incorporated into bio-active natural products that are used in the medical field.[3]

Plant Wax Ester: triacontanyl palmitate



Insect Pheromone: *cis*-7-tricosene



Bacterial Antibiotic: lipstatin

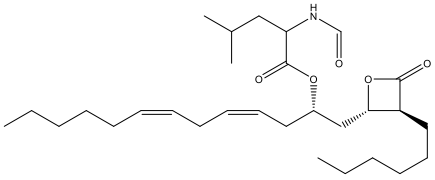


Figure 1-1. *The hydrocarbon moiety is found throughout the tree of life. Triacontanyl palmitate and lipstatin contain only a hydrocarbon substructure, while cis-7-tricosene is a true hydrocarbon.*

Hydrocarbons in Modern Society

Modern society values hydrocarbons primarily for their use as fuels and chemical building blocks. Currently, the greatest source of hydrocarbons comes from non-renewable crude oil which is pumped from the ground. Hydrocarbon content in a barrel of crude oil varies greatly by location, between 50-98%.[4] The most significant portion of the hydrocarbons, >85%, are used as fuels for cars, planes, or industrial plants, while most of the remaining percentage is dedicated to chemical building blocks for many plastics.[5] The small remaining hydrocarbons are used for a wide array of smaller volume applications from cosmetics to candles.

Global crude oil consumption has risen steadily over the past century and reached an estimated 35 billion barrels per year in 2015.[6] For comparison, this is approximately 5 cubic kilometers, or the volume of water the Mississippi rivers discharges in 3 days.[7] This level of consumption has raised significant environmental concerns on both local and global scales and have been well detailed in many publications.[8–10] Additionally, many consider the use of non-

renewable hydrocarbons for fuels and chemicals as unsustainable because of the finite quantity of crude oil reserves. However, the quantity of global crude oil reserves remains a subject of active debate.

In response to these alarming facts, recent decades have seen increased action by governments, private organizations, and industries to find alternative sources of energy and crude oil-derived chemicals. Significant investments have been made to harness living systems such as plants and microbes as renewable options for fuels, commodity chemicals, and various small-volume compounds for niche applications. There are now many success stories from these efforts such as the production of bio-diesel from rapeseed, the fermentation of 1,4-butanediol as a chemical precursor from a bacterium, and the fermentation of squalene by yeast for the cosmetic industry.[11–13] The drive to reduce society's dependence on crude oil propelled our lab to study the bacterial biosynthesis of hydrocarbons from fatty acids.

Placing this Dissertation within a Larger Context

The work of this dissertation describes our efforts to promote the sustainable development and production of hydrocarbons from bacterial fatty acids. Fatty acids are excellent precursors for fuels and various niche chemical compounds because of their high natural abundance in every cell. In addition to hydrocarbon biosynthesis, fatty acids feed a wide variety of biochemical pathways that incorporate the hydrocarbon moiety into compounds for quorum sensing, antibiotic production, surfactants for solubilizing hydrophobic food sources, and others (Figure 1-2).[3,14,15] The work described in this dissertation originally set out to study only the production of hydrocarbons from fatty acids, but unexpectedly discovered a strong connection between the biosynthesis of hydrocarbons and that of natural products with antibiotic and anti-cancer properties. As such, the ensuing text will first provide the seemingly disparate but necessary background on bacterial natural products with medicinal value then focus on bacterial hydrocarbon biosynthesis.

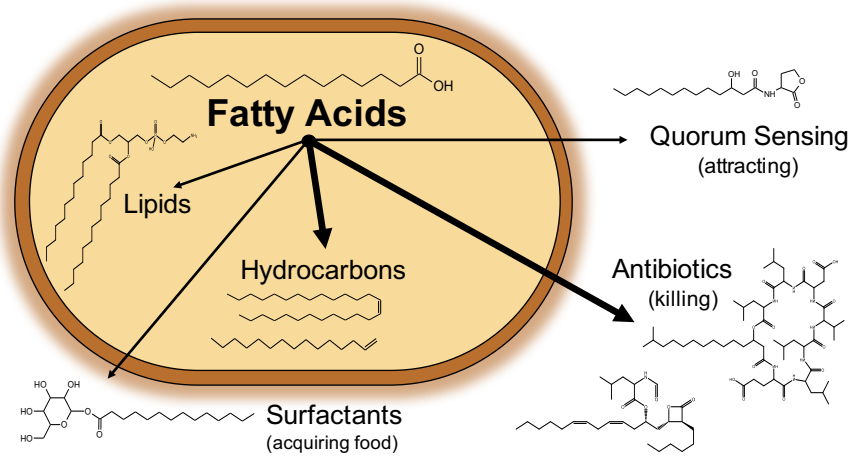


Figure 1-2. *In bacteria, the reduced carbon of fatty acids is utilized by various biochemical pathways to create compounds for diverse functions. The thick arrows leading to "Hydrocarbons" and "Antibiotics" indicate the focus of this work, and do not represent metabolic flux.*

β-Lactone Antibiotics

Life's endless struggle for survival has generated a vast array of natural compounds with potent toxicity to both prokaryotic and eukaryotic cells that can be harnessed by modern medicine. Microbes access the metabolites within their reach and combine them using standard enzymatic reactions to generate staggering chemical diversity. Many of these nascent biosynthesis pathways to new compounds confer no evolutionary advantage to the host and are quickly lost, while others that provide a competitive edge are passed on to offspring and optimized into a streamlined enzymatic pathway.

β-Lactone natural products are compounds isolated from bacteria that are known to have potent anti-cancer, antibiotic, and anti-obesity properties.[16] The four-membered β-lactone ring is very reactive and can alkylate active site nucleophiles of proteases, lipases, and esterases.[3,16,17] For example, the fatty acid-derived β-lactone natural product lipstatin (Orlistat, Xenical, Alli) from *Streptomyces* is the only FDA-approved over-the-counter anti-obesity drug on the market. Lipstatin acts by inhibiting human pancreatic lipase thereby

preventing the proper assimilation of fats from the diet.[3] Salinosporamide is a bicyclic β-lactone produced by *Salinispora tropica* that is known to inhibit human 20S protease function.[18] Salinosporamide is now in phase II clinical trials for glioblastoma and multiple myeloma and acts by inhibiting the tumor cells ability to degrade pro-apoptotic proteins.[19] Synthetic β-lactones such as 3-benzyl, 4-propyl oxetanone are known to inhibit the ClpP protease of *Mycobacterium tuberculosis*.[17,20] These results are especially exciting as proteases represent novel targets for antibiotics, suggesting β-lactones could provide an option for treating resistant organisms. Many other β-lactones from both natural and synthetic sources show potent anti-biotic and anti-cancer abilities.[21,22]

Despite the current usage and widespread potential of β-lactone natural products, no enzyme capable of synthesizing the critical β-lactone ring has been characterized. The biosynthetic gene clusters for many β-lactones have been identified, but characterization has proved difficult.[3,16,23] The biosynthesis of the fatty-acid derived lipstatin is probably the best characterized β-lactone natural product and its putative biosynthesis pathway has been described.[3] In lipstatin, the critical β-lactone ring is thought to form spontaneously from the cyclization of the β-hydroxy thioester moiety.[3] However, this has never been directly demonstrated.

Bacterial Hydrocarbon Production

Hydrocarbons derived from fatty acids are dispersed throughout the bacteria kingdom, but are most frequently found in cyanobacteria.[24] Studies reporting the cellular benefit of hydrocarbons are surprisingly rare, but they are generally understood to help maintain proper membrane fluidity and function.[25] There are three known biochemical pathways that utilize fatty acid precursors to generate hydrocarbons, the ADO, OleT, and Ole pathways, named for the enzymes that produce them (Figure 1-3).[24,26] The Ole pathway is the focus of this work, but an overview of the first two pathways will be given.

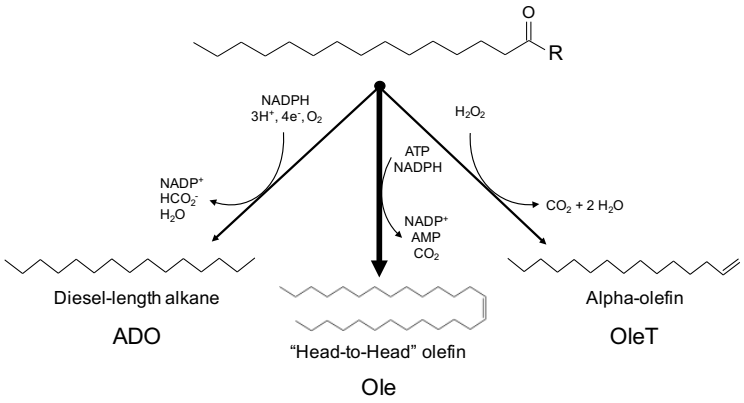


Figure 1-3. *Summary of the three known bacterial pathways. The enzymatic functions of the OleT and ADO pathways have been described, but the chemical steps of the Ole pathway are incomplete. The "R" group can be an OH, CoenzymeA, or Acyl Carrier Protein (ACP).*

Overview of ADO and OleT pathways

The ADO pathway was first identified in cyanobacteria by Schirmer *et al.* in 2010 and generates alkanes and alkenes from fatty acid-acyl carrier proteins (acyl-ACP) precursors.[27] The pathway is a two-step process; the first enzyme is an NADPH-dependent acyl-ACP reductase (AAR) which reduces ACP-linked fatty acids to fatty aldehydes, releasing ACP. The second step is an aldehyde deformylating oxygenase (ADO) that generates formate and the final alkane or alkene product that is one carbon shorter than the starting fatty acid.[24] The hydrocarbon tail of most cellular fatty acids is similar to the hydrocarbon profile of diesel fuel, and efforts have been made to implement this pathway in the production of fuels.[14] However, to my knowledge, no commercial process is yet in existence.

The OleT pathway is composed of a single enzyme that carries out the decarboxylation of fatty acids to terminal alkenes (olefins).[28] Found in *Firmicutes*, OleT is a P450 monooxygenase that can accept a wide range of fatty acids for decarboxylation. OleT requires exogenous source of electrons for the oxidation of the carboxylic acid head group, and H_2O_2 has been demonstrated to

work *in vitro*.[29] Again, there has been interest in implementing this enzyme for the production of biofuels, but to my knowledge, no such process exists.[30]

Olefin (alkene) Biosynthesis Pathway (Ole)

The synthesis of long-chain olefins represents the third pathway from fatty acids to hydrocarbons in bacteria.[31] The hydrophobic products of this pathway were first identified in the late 1960's by Albro and Dittmer while analyzing the membrane components of *Micrococcus luteus*. They reported long-chain olefinic hydrocarbons composed of between 27 and 29 carbons, and correctly surmised that these olefins arose from combining two fatty acids. Over the next decades, long-chain olefins were identified in multiple different bacteria, but it was not until 2008 that the genes responsible for olefin biosynthesis, *oleABCD*, were identified by LS9 Inc.[32] The genes were patented by LS9 Inc. for their use in the production of hydrocarbons and a ketone intermediate.[32]

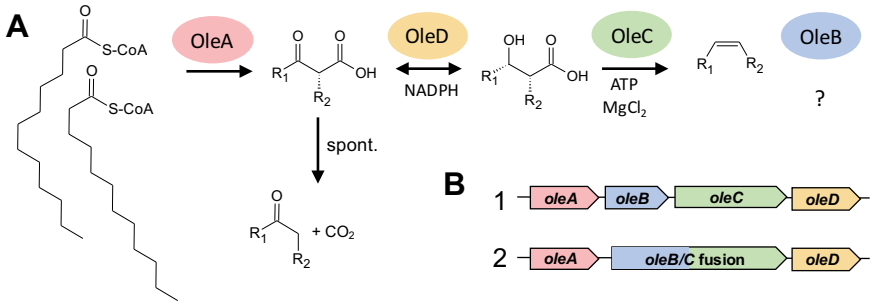


Figure 1-4. *(A) Understanding of the enzymatic steps of the long-chain olefin biosynthesis prior to this work. The function of OleC as shown was demonstrated to be incorrect. The function of OleB was unknown. (B)Two typical ole-gene clusters. OleB and OleC are found as a natural fusion in all gram-positive bacteria containing the ole genes.*

Reports characterizing olefin biosynthesis were published rapidly over the next few years. Frias *et al.* studied long-chain olefin from various *Arthrobacter* species and concretely established the *cis*-olefin as the sole stereochemical configuration produced by this pathway (Figure 1-4A).[33] Bioinformatics work by Sukovich *et al*. in 2010 revealed *oleABCD* gene clusters in 69 species spread

throughout the bacterial kingdom.[34] All gram-positive bacteria containing this pathway were found to contain a fusion of the *oleB and oleC* genes, but were still found to produce olefins (Figure 1-4B). Fourteen of the 69 were grown, and membrane composition analysis revealed diverse olefin profiles.[34] Some species, such as the plant pathogen *Xanthomonas campestris*, were found to synthesize over a dozen different long-chain olefins with various lengths and up to three double bonds. Others however, such as the Antarctic bacterium *Shewanella frigidmarina*, were found to produce a single polyolefin, 3,6,9,12,15,19,22,25,28-hentriacontanonaene. In cold temperature growth studies wild type *Shewanella* onidensis displayed a much shorter lag phase when growth was shifted from 26°C to 4°C when compared to *oleABCD* deletion strains.[25] A temperature related phenotype for polyolefin producing strains is consistent with its frequent occurrence in cold-water bacteria. However, it is unclear if this phenotype is specific to the polyolefin producing bacteria only or is broadly applicable to all bacteria containing the olefin biosynthesis pathway.

OleA – The first enzyme of olefin biosynthesis

OleA is the first enzyme of the olefin biosynthesis pathway, and it is by far the most well characterized. Work by Beller, *et al*. on recombinantly expressed and purified OleA from *Micrococcus luteus* reported that this enzyme accepts two coenzyme A-linked fatty acids and combines them head-to-head to form a ketone (Figure 1-4) *via* a decarboxyative Claisen condensation reaction.[35] More recent work on OleA from *Xanthomonas campestris* by Frias *et al*. supported some of Beller's findings, but revealed that the true product of OleA was a β-keto acid that can spontaneously decarboxylate to form the ketone observed in the earlier studies.[36] Frias *et al*. conducted LC/MS analysis of the OleA product with synthetic standards rather than GC/MS analysis which causes the β-ketoacid to spontaneously decarboxylate in the high temperatures of the GC inlet port.[36] Additionally, only the β-ketoacid, not the ketone, was able to be processed by the second pathway enzyme OleD. A schematic of the olefin biosynthesis pathway prior to the work of this dissertation is shown in Figure 1-4.

Crystal structures of OleA have helped determine much of the mechanism of this enzyme. Goblirsch *et al*. crystallized *Xc.* OleA and found that it was an obligate homodimer with a key active site glutamic acid originating from the opposite monomer.[37] Mutant work with myristoyl-CoA substrates trapped the two acyl chains of the fatty acids *in crystallo* and allowed the active site and substrate binding channels to be mapped.[38] Subsequent work on OleA has defined more of the critical catalytic residues and established an HMG CoA-synthase-like mechanism. Briefly, the first acyl-CoA substrate binds in alkyl channel A to allow the transfer of the fatty acid to Cys_{143} by nucleophilic attack of the thioester bond. The second acyl-CoA binds into alkyl channel B before Glu_{117} from the opposite OleA subunit abstracts an alpha proton from this second substrate to generate a carbanion. The electrons of the carbanion then attack the carbonyl carbon of the enzyme-linked fatty acid to release the nucleophilic Cys and form a carbon-carbon bond, thereby generating a CoA-linked dialkyl β-keto thioester. The CoA is then hydrolyzed to release the final β-keto acid product.[38] It is unclear how this hydrophobic β-keto acid product with two alkyl tails is transferred to the next pathway enzyme.

OleD – An NADPH depended dehydrogenase

The seminal study on OleD was published in 2011 by Bonnett and Kancharla *et al*. which establish it as the second enzyme of olefin biosynthesis.[39] In this work, OleD from *Stenotrophomonas maltophilia* was recombinantly expressed and purified as an active dimer. OleD was found to promote the reversible reduction of the β-keto acid to a β-hydroxy acid in an NADPH-dependent reaction.[39] To determine the stereochemical preference of OleD, the four diastereomers of the β-hydroxy acid were synthesized and reacted with OleD in the oxidative direction. Reactions were followed by the appearance of NADPH absorbance, and it was found that the 2R,3S-configuration was preferred 50-fold over the next closest diastereomer. Enzyme kinetic analysis found that both alkyl chains are required for processing of the β-hydroxy acids. The optimal substrate was found to be *syn*-2-decyl-3-hydroxytetradecanoic acid; however,

substrates with longer alkyl groups that are more representative of *in vivo* substrates were not able to be tested because of their extreme hydrophobicity. Again, it is unclear how the transfer of this compound to the downstream enzymes occurs *in vivo*.

OleC – A controversial ending to olefin biosynthesis

A subsequent report by Kancharla *et al*. in 2016 described OleC as the final enzyme of the olefin biosynthesis pathway.[40] Their findings will be summarized here; however, our research (reported in Chapter 2) directly contradicts their functional assignment of OleC. Building on their earlier OleD work, all four β-hydroxy acid diastereomers were reacted with MgATP and OleC from *Stenotrophomonas maltophilia*; all four were found to exclusively produce *cis*-olefin when analyzed by GC and ^{1}H NMR.[40] The discovery that only the OleACD proteins are necessary for the completion of olefin biosynthesis is surprising given the ubiquitous presence of the fourth gene, *oleB,* in *oleABCD* gene clusters. OleC has a broad substrate specificity as all four diastereomers were acted upon with similar k_{cat}/Km values as monitored by PPi release from ATP.[40] The mass of a proposed β-hydroxy acid-AMP intermediate was identified by mass spectrometry as the basis for their purported mechanism. Briefly, OleC is reported to use ATP to activate the alcohol of the β-hydroxy acid substrate with an AMP. The substrate then undergoes a decarboxylative elimination to release carbon dioxide, AMP, and the final olefin product. This hydroxyl activation mechanism is similar to diphosphomevalonate decarboxylase of the galacto-, homoserine, mevalonate and phosphomevalonate kinases (GHMP) superfamily. However, OleC is a clearly member of the acyl-CoA, non-ribosomal peptide synthetase, and luciferase superfamily (ANL).[41] Every member of the ANL superfamily characterized to date uses ATP to activate the carboxylic acid group of a substrate with an AMP leaving group rather than the hydroxyl group proposed. It is unclear why OleC would deviate from other members of the superfamily.

OleB – The forgotten member of the Ole pathway

Prior to my dissertation research, nothing was known about the function of OleB in olefin biosynthesis; if fact, literature reports that OleACD are both necessary and sufficient to complete olefin biosynthesis from acyl-CoAs.[36,40] However, its ubiquitous presence in all 69 *ole* gene clusters identified by Sukovich *et al*. strongly suggests an important function.[34] OleB sequence alignments identify this protein as an α/β-hydrolase, and OleB is often annotated in genomes as a haloalkane dehalogenase. It has been hypothesized that OleB may function as a non-essential isomerase to promote the correct stereochemistry of pathway intermediates. However, there is no data to support this theory as the recombinant purification of OleB has proved difficult. The state of our knowledge of OleB can be summarized by a previous dissertation from our lab, "no mechanism or defined role [has] been identified. No experimentation can be done. . . until the purification yield and OleB stability is improved."[42]

Goals of Research

The overall goal of this project is to elucidate the function and mechanism of the enzymes in the long-chain olefin biosynthesis pathway to facilitate its implementation in the production of specialty chemicals. Of the three bacterial pathways to hydrocarbons, only the Ole pathway enzymatic steps remained undiscovered. The first two enzymes, OleA and OleD, were comparatively well characterized, and were not the focus of this dissertation. We had doubts about the literature assigned function of OleC, necessitating a closer investigation. Also, purification and functional identification of OleB is a significant goal of this work. Finally, as the intermediates of olefin biosynthesis are extremely hydrophobic, we sought to understand how the intermediates are transferred between the enzymes of this pathway.

CHAPTER 2 - OleC

One sometimes finds what one is not looking for.
— Sir Alexander Fleming

Biochemistry, **2017**, *56* (2), pp 348–351
DOI: 10.1021/acs.biochem.6b01199

β-Lactone Synthetase Found in Olefin Biosynthesis Pathway

James K. Christenson, Jack E. Richman, Matthew R. Jensen, Jennifer Y. Neufeld, Carrie M. Wilmot, and Lawrence P. Wackett

OleC Summary: The first β-lactone synthetase enzyme is reported, creating an unexpected link between the biosynthesis of olefinic hydrocarbons and highly functionalized natural products. The enzyme OleC, involved in the microbial biosynthesis of long-chain olefinic hydrocarbons, reacts with *syn*- and *anti*-β-hydroxy acid substrates to yield *cis*- and *trans*-β-lactones, respectively. Protein sequence comparisons reveal that enzymes homologous to OleC are encoded in natural product gene clusters that generate β-lactone rings, suggesting a common mechanism of biosynthesis.

Introduction

The β-lactone (2-oxetanone) substructure is well known in organic synthesis and microbial natural products, some of which are presently being investigated for anti-obesity, anti-cancer, and antibiotic properties.[3,16,43–45] Although multiple organic synthesis routes exist for β-lactones, no specific enzyme that catalyzes the formation of this functional group had been identified.[46] While defining the chemistry of a well-known olefinic hydrocarbon biosynthesis pathway, we identified a β-lactone synthetase whose presence extends into natural product biosynthesis.

The olefin biosynthesis pathway is encoded by a four-gene cluster, *oleABCD*, and is found in over 250 divergent bacteria.[34] Ole enzymes produce long-chain hydrocarbon *cis*-alkenes from activated fatty acids through the reported chemistry shown in **Figure 2-1**.[31,33] OleA, the first enzyme of the pathway, has been studied in *Xanthomonas campestris* (*Xc*) and found to catalyze the "head-to-head" coupling of CoA-activated fatty acids (**1**) to unstable β-keto acids (**2**).[36] The second enzyme, OleD, couples the reduction of **2** with NADPH oxidation to yield stable β-hydroxy acids (**3**) as defined in *Stenotrophomonas maltophilia* (*Sm*).[39] Finally, using GC detection methods, we have observed and others have reported that *Sm* OleC catalyzes an apparent decarboxylative dehydration reaction to generate the final *cis*-olefin product.[40] Together, these findings indicate that there is no defined purpose for the ever-present fourth gene in the cluster, *oleB*.

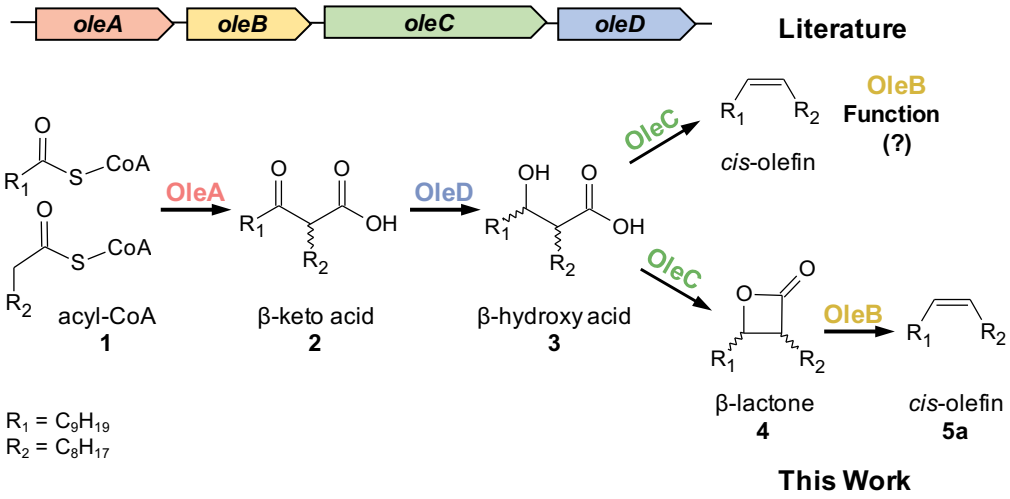


Figure 2-1. *ole gene cluster and enzymatic pathway.*

We now report that the observation of *cis*-alkenes as products of OleC is an artifact, a misleading observation caused by GC-analytical methods used to identify the high-boiling point products of OleC. Rather, using ^{1}H NMR, we demonstrate that OleC from four different bacteria produce thermally-labile β-lactones from β-hydroxy acids in an ATP-dependent reaction; no alkenes are observed. Further analyses of gene clusters for β-lactone-containing natural products reveal *oleC* homologues that likely perform this previously unknown biological β-lactone ring closure reaction.

The first suggestions of β-lactone synthetase activity arose when monitoring reactions of *Xc* OleC with ATP, $MgCl_2$, and a synthetic, diastereomeric mixture of **3** by GC. Full ^{1}H NMR spectra of all synthesized compounds from this study are shown in Supplemental Figures S1-S6. Two peaks were observed by GC, coupled to both a mass spectrometer (MS) and flame ionization detector (FID), with mass spectra and retention times identical to synthetic *cis*- and *trans*-olefin standards (**5a** and **5b**, respectively). However, the GC/FID peak area of the enzymatically-generated olefin varied significantly with GC inlet temperature and inlet liner purity, while synthetic standards were unaffected, suggesting that the observed olefin may be a thermal decomposition product of the actual OleC product.

Results and Discussion

To test this hypothesis, *Xc* OleC reactions with **3** were scaled to generate sufficient quantities for ^{1}H NMR. No resonances consistent with the prepared olefin standards (Figures S1 and S2) (**5a**, **5b**) were observed; rather, four distinct multiplets, appearing as doublet of doublets of doublets, arose (Figures S7 and S8). These resonances were consistent with the two hydrogens of *cis*- and *trans*-β-lactone rings and perfectly matched our authentic standards of *cis*- and *trans*-3-octyl-4-nonyloxetan-2-one (**4a**, **4b**) (Figures S3 and S4). Furthermore, when **4a** and **4b** were analyzed by GC, retention times and mass spectra matched the olefin standards **5a** and **5b**, with sensitivity to inlet conditions being observed. The thermal decarboxylation of *cis*- and *trans*-β-lactone to *cis*- and *trans*-olefin, respectively, is well known.[47,48] We believe that thermal decomposition during GC/MS analysis caused the product of OleC catalysis to be misidentified. Additionally, when reviewing supplemental NMR data from the literature report of *Sm* OleC characterization, resonances of the *cis*- and *trans*-β-lactones, consistent with those synthesized here, are visible.[40]

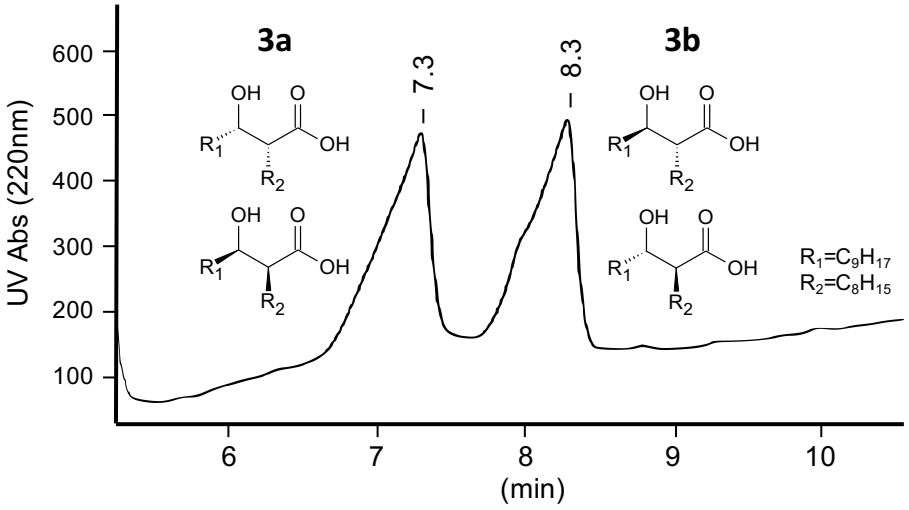


Figure 2-2. *HPLC separation of syn- and anti-β-hydroxy acids (**3a** and **3b**, respectively) for use as Xc OleC substrates. The rise in baseline is caused by impurities from the MTBE gradient that absorb at 220 nm.*

The stereochemical origins of **4a** and **4b** were then investigated by reacting *Xc* OleC with *syn*- and *anti*-β-hydroxy acids (Figures S9 and S10). HPLC was used to separate **3** into its *syn*- and *anti*-diastereomeric pairs (**3a** and **3b**) (**Figure 2-2**). Examining **3a** and **3b** by ^{1}H NMR and GC/MS, post-methylation, demonstrated each contained <10% of the opposite racemic diastereomer (Figures S5 and S6). When reacting with *Xc* OleC, **3a** produced **4a**, while **3b** generated **4b** (**Figure 3-2**). GC/MS analysis supported this conclusion, as OleC reactions with **3a** and **3b** yielded the β-lactone breakdown products, **5a** and **5b**, respectively. OleC consumed greater than 90% of substrates **3a** and **3b** by GC/MS, supporting the conclusions of Kancharla *et al.,* that all four β-hydroxy acid isomers are utilized by OleC.[40] Taken together, *Xc* OleC represents the first reported β-lactone synthetase, converting β-hydroxy acid substrates to β-lactones in the presence of ATP and $MgCl_2$. MgATP is likely required to activate the hydroxyl or carboxyl group and promote β-lactone ring formation.

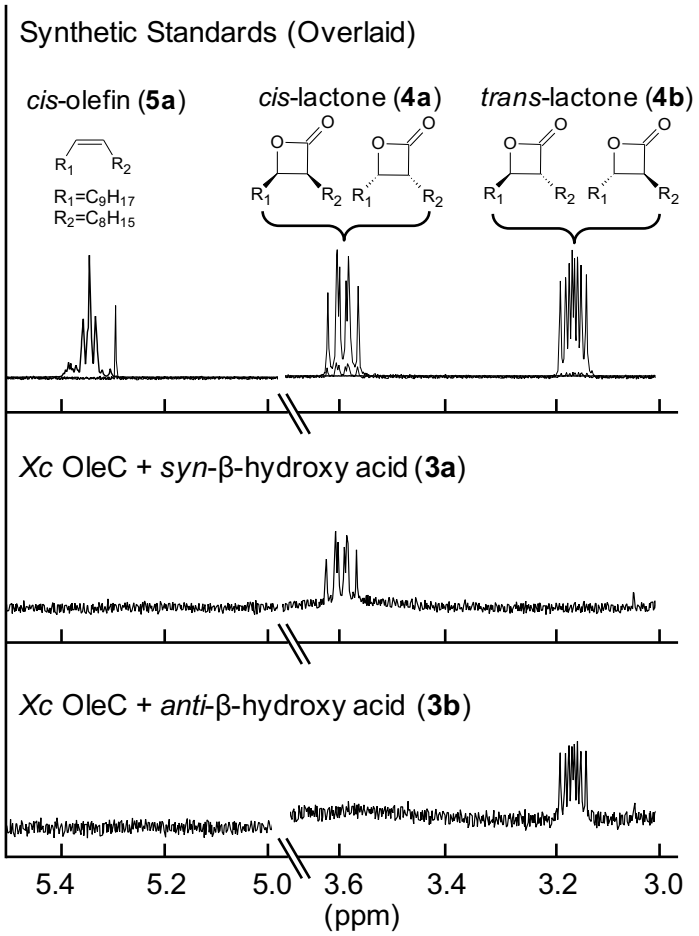


Figure 2-3. *^{1}H NMR (400 MHz; $CDCl_3$) analyses of Xc OleC products with β-hydroxy acid substrates.*

To determine if β-lactone synthetase activity is a common enzymatic step in long-chain olefin biosynthesis, we obtained four *oleC* genes from *oleABCD* gene clusters in divergent microorganisms (**Table 2-1**). Purified OleC enzyme from the four organisms was reacted overnight with ATP and **3,** then analyzed by ^{1}H NMR and GC/MS. The products of OleC proteins from the bacteria *Stenotrophomonas maltophilia*, *Arenimonas malthae*, and *Lysobacter dokdonensis* were both **4a** and **4b** β-lactones (Figures S11-S13), with no **5a** or **5b** olefins being observed, indicating that OleC enzymes from diverse sources are β-lactone synthetases. The Gram-positive bacterium *Micrococcus luteus* (*Ml*) was specifically chosen because its sequence is very divergent from *Xc* OleC,

and it contains a natural fusion of the *oleB* and *oleC* genes. This natural *oleBC* fusion is found in all actinobacteria, which comprise ~30% of the microorganisms that contain identifiable *oleABCD* genes. Reaction of purified *Ml* OleBC fusion with $MgCl_2$, ATP, and **3** produced β-lactones **4a** and **4b** as well as small amounts of *cis*-olefin, **5a** (Figure S14). No trace of *trans*-olefin, **5b**, was detected. Further characterization is ongoing, but we believe that OleB performs a *syn*-elimination of *cis*-β-lactone to form the final *cis*-olefin product. This is consistent with previous studies of microorganisms expressing *ole* genes that contain olefins with a *cis*-relative configuration exclusively.[31,33,34] These data also demonstrate that an enzyme domain with only 35% amino acid identity to the *Xc* OleC generates β-lactones, indicating that this activity is likely common among all olefinic hydrocarbon biosynthesis OleC homologues.

Table 2-1. Other OleC enzymes make β-lactones

Organism	Ascension #	% ID[a]
Xanthomonas campestris	WP_011035474.1	100
Stenotrophomonas maltophilia	AFC01244.1	77
Arenimonas malthae	WP_043804215.1	73
Lysobacter dokdonensis	WP_036166093.1	70
Micrococcus luteus[b]	WP_010078536.1	35

[a] % identity based on amino acid sequence
[b] OleC and OleB are a natural fusion in *Ml*.

Establishing the widespread nature of lactone synthetase activity within olefinic hydrocarbon biosynthesis led us to search sequence databases for OleC homologues in other biosynthetic pathways. OleC is a member of the ubiquitous AMP-dependent ligase/synthetase enzyme superfamily; as such, homologues are found in all organisms.[49] As of November 2016, a BLAST search of NCBI's non-redundant protein sequence database identified over 900 sequences with greater than 35% sequence identity and over 16,000 with greater than 25% sequence identity to *Xc* OleC.

Of the sequences examined, two *Xc* OleC homologues were clearly encoded in gene clusters known to produce β-lactone natural products (**Figure 2-**

4). The first, LstC, is an uncharacterized enzyme found in the lipstatin biosynthesis pathway from *Streptomyces toxytricini*. Lipstatin is the precursor to Orlistat, the only over-the-counter, FDA-approved anti-obesity drug. LstC is a member of the AMP-dependent ligase/synthetase superfamily, and its protein sequence is 38% similar to *Xc* OleC, more similar than the sequence of the β-lactone synthetase domain of *Ml* OleBC (35%). Surprisingly, further investigation revealed homologues to OleA and OleD encoded by the lipstatin gene cluster, suggesting that the two gene clusters have a common ancestry. The synthesis of both lipstatin and olefinic hydrocarbons are initiated by the condensation of two fatty acyl-CoAs to form a β-keto acid. In the case of lipstatin, the two fatty acids are 3-hydroxy-linoleic and octanoic acid.[3] The hydroxyl group of linoleic acid is later functionalized by LstE and LstF with a modified valine.[3] LstD and OleD likely perform the same NADPH-dependent reduction of the β-keto group to a hydroxyl group. Formation of the *trans*-β-lactone is likely accomplished by the OleC homologue LstC, to generate the final product in lipstatin biosynthesis. Olefin biosynthesis is completed by the putative OleB-dependent elimination of CO_2 to generate the final olefin product. The lipstatin gene cluster lacks any gene product that shows homology to OleB, consistent with the accumulation of the β-lactone natural product and further supporting our hypothesis that OleB performs the final step in the biosynthetic pathway to olefins.

The gene cluster responsible for the biosynthesis of ebelactone A, a commercially available esterase inhibitor, in *Streptomyces aburaviensis* shows a gene, *orf1*, with 46% amino acid sequence identity to *Xc* OleC and is directly adjacent to *ebeA-G* (**Figure 2-4**). Unlike lipstatin, ebelactone A is formed partly by a polyketide synthase multi-domain protein rather than fatty acid condensation; as such, OleA and OleD homologues are not encoded in the surrounding gene cluster. Literature reports suggest that the β-lactone ring of ebelactone A is formed spontaneously from the final, enzyme-linked, β-hydroxy-thioester intermediate.[16] While a spontaneous β- hydroxy-thioester cyclization is mechanistically plausible, β-lactone formation from β-hydroxy-thioesters in

ubiquitous pathways such as fatty acid oxidation or synthesis has not been reported to our knowledge.[50] Additionally, β-hydroxy-thioester intermediates are extremely common in polyketide synthesis pathways, while β- lactone formation is comparatively rare. An Orf1-independent cyclization would require a unique property of ebelactone A precursors or a novel polyketide domain architecture to promote β-lactone ring cyclization. However, in favor of an Orf1-independent mechanism is that no thioesterase domain exists in the final polyketide synthesis domain, suggesting that no free β-hydroxy acid is released for the putative ATP-dependent Orf1 to act on. Other polyketide-type β-lactone gene clusters, such as those for salinosporamide A, cinnabaramide, and oxazolomycin, do not encode an OleC homologue with high sequence similarity (>35%). Polyketide-derived β-lactones are thought to form by the cyclization of the final thioester, enzyme-linked intermediate, but this has never been characterized.[23,43,51,52] It is reasonable to hypothesize that specialized polyketide synthase domains represent a second mechanism of β-lactone formation. Regardless, the discovery of a stand-alone β-lactone synthetase here creates new opportunities for the natural product field. Preliminary screening of *Streptomyces* and *Nocardia* genomes suggests that β-lactone natural products may be more widespread than currently realized.

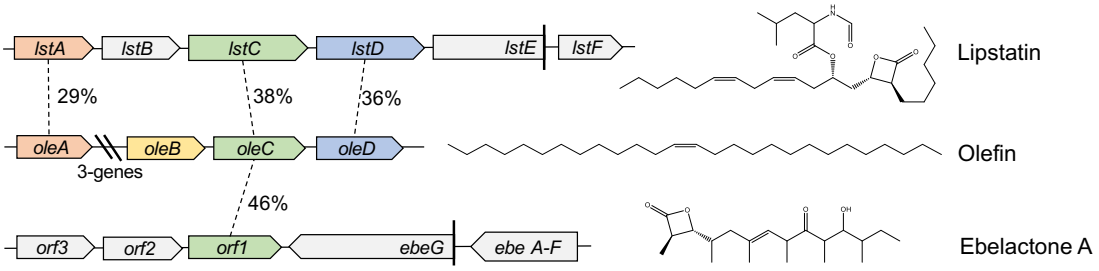


Figure 2-4. *Homologous OleC enzymes encoded in β-lactone biosynthesis gene clusters. Percent identity is based on amino acid sequences. The E values for OleC to LstC and Orf1 are 2e-72 and 1e-143, respectively. Lipstatin is the precursor to the anti-obesity drug Orlistat. Ebelactone A is a commercially available general esterase inhibitor.*

ACKNOWLEDGMENTS

We thank Anna Kim, Tiffany Engel, and Shintaro Nakayama for their practical support in the lab.

Once you eliminate the impossible, whatever remains, no matter how improbable, must be the truth. – Sir Arthur Conan Doyle

This chapter has been prepared for submission to ACS *Biochemistry.*

OleB from bacterial hydrocarbon biosynthesis is a β-lactone decarboxylase sharing key features with haloalkane dehalogenases

James K Christenson, Serina L Robinson, Tiffany A. Engel, Jack E. Richman, An N. Kim, Larry P. Wackett.

OleB Summary: OleB is an α/β-hydrolase ubiquitous in bacteria that biosynthesize long-chain olefinic hydrocarbons, but its function and mechanism has remained obscure. We report that OleB from the gram-negative bacterium *Xanthomonas campestris* performs an unprecedented β-lactone decarboxylation reaction, to complete *cis*-olefin biosynthesis. ^{1}H NMR analysis of OleB reactions revealed a selectivity for decarboxylating *cis*-β-lactones and no detectable activity with *trans*-β-lactones consistent with the stereochemistry of pathway intermediates. Protein sequence analyses showed OleB proteins are most related to haloalkane dehalogenases (HLDs). Unexpectedly, it was determined that a subfamily of HLDs, denoted as class III, is comprised mostly of OleB proteins suggesting a mis-annotation. The OleB proteins contain the canonical Asp-His-Asp catalytic triad characteristic of HLDs, but *X. campestris* OleB showed no discernible dehalogenase activity against typical haloalkane substrates. Rather, a haloalkane substrate mimic alkylated wild type *Xc* OleB but

not $OleB_{D114A}$ implicating this residue as the active site nucleophile. To further confirm our findings, a sequence-divergent OleB, found as part of a natural OleBC fusion, from the gram-positive bacterium *Micrococcus luteus* was demonstrated to have the same activity, stereochemical preference, and dependence on the proposed Asp nucleophile. An analogous mechanism to haloalkane dehalogenases is proposed in which Asp_{114} attacks the carbonyl carbon of the β-lactone ring, producing the olefin and leaving carbonate bound to the Asp_{114}. His_{277} activates water for attack to release bicarbonate and regenerate the resting form of the enzyme. These findings define a previously unrecognized reaction in nature and represent a novel cellular route for producing hydrocarbons.

Introduction

The α/β-hydrolase enzyme scaffold is a very common fold, used to catalyze a wide array of chemical reactions.[53,54] The vast majority of α/β-hydrolases that have been studied initiate catalysis *via* attack of a catalytic nucleophile to form an acyl-enzyme intermediate that is hydrolyzed by a water molecule that is activated by a conserved histidine residue, with subsequent release of the product and a return of resting enzyme.[55,56] Despite their prevalence in biology, approximately 35% of enzymes annotated as α/β-hydrolases do not have a known substrate, thus their cellular function remains unknown.[54]

One such α/β-hydrolase is encoded by the gene denoted as *oleB*, and resides in the *ole* (olefin) gene cluster responsible for the biosynthesis of long-chain olefins.[34,25] Early studies demonstrated that long-chain olefins are generated following the head-to-head condensation of two fatty acyl-CoA molecules.[36] The olefins can be 19-31 carbons in length and contain a central double bond at the site of C-C bond formation.[34,31,36] Genetic work in *Shewanella oneidensis* concretely linked the four-gene cluster, *oleABCD*, to hydrocarbon production, and the *ole*-genes have now been identified in over 300 divergent bacteria.[25,34] The α/β-hydrolase, OleB, is encoded from stand-alone gene or as part of a gene fusion with *oleC*. Recently, the OleB, OleC, and OleD proteins from *Xanthomonas campestris* were found to associate *in vivo* to form an active, multi-enzyme complex when recombinantly expressed and purified from *Escherichia coli*, further suggesting an important function for OleB.[57]

Until recently, only OleACD were thought to be required for the generation of long-chain olefins, leaving no apparent function for OleB.[40] The roles of OleA and OleD as the first two pathway steps had been previously established, with OleA preforming the Claisen condensation of two acyl-CoAs to form a β-ketoacid and OleD catalyzing the NADPH-dependent reduction of the keto acid to produce a β-hydroxy acid (Scheme 1).[36,38,39] The third enzyme, OleC, was initially thought to react with the β-hydroxy acid in the presence of ATP to produce the long-chain

olefin that is the endpoint of the metabolic pathway. More recently, it was discovered that OleC forms a stable β-lactone under physiological conditions.[58] In the earlier work, the OleC reaction product, the β-lactone, had been analyzed using gas chromatography at high temperature, resulting in a spontaneous decarboxylation reaction to make the observed olefin.[58] This most recent report reserves a possible role for OleB, but the decarboxylation of a β-lactone to an olefin has never been reported for any enzyme.

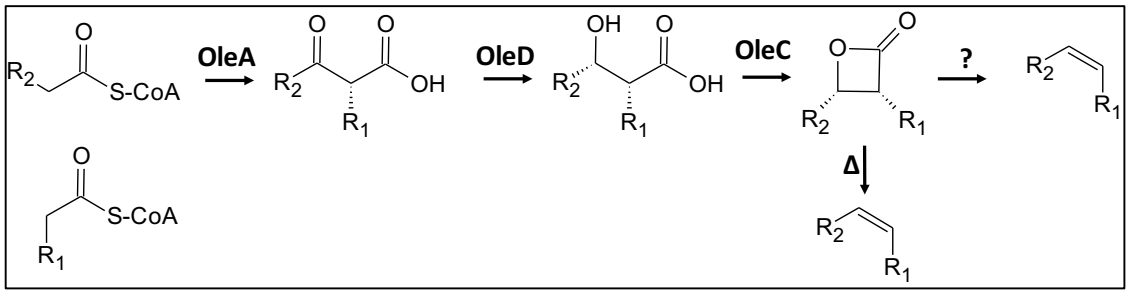


Figure 3-1. *Reactions previously established for the enzymes OleA, OleD, and OleC. Preliminary evidence by Christenson et al. suggests that OleB may perform the final β-lactone decarboxylation reaction, but this had never been demonstrated.*[58] *The R groups are typically linear alkanes between 12-14 carbons, but can contain unsaturation or methyl branching.*

Here, we report an optimized purification scheme and new information that the α/β-hydrolase, OleB, from *Xanthomonas campestris* selectively decarboxylates *cis*-β-lactones to *cis*-alkenes, the final step in bacterial long-chain olefin biosynthesis. This is a previously unreported enzymatic activity and represents a novel cellular route to hydrocarbons. Sequence comparisons to other α/β-hydrolases revealed OleB proteins and OleBC fusion proteins cluster within haloakane dehaolgenases (HLDs). Indeed, a previously designated group of haloalkane dehalogenases, HLD-III, are indicated to contain proteins that function physiologically as OleB enzymes.[59] This connection to HLD proteins also provided testable hypotheses regarding a possible reaction mechanism for OleB

and specific residues were shown by site directed mutagenesis to be critical for catalysis.

Methods

Chemical synthesis of β-hydroxy acids, β-lactone, and olefin.

All compounds, *cis*- and *trans*-3-octyl-4-nonyloxetan-2-one (β-lactones), 3-hydroxy-2-octyldodecanoic acid (β-hydroxy acids), *cis*- and *trans*-9-nonadecene (olefins) where chemically synthesized as described previously.[58] Briefly, β-hydroxy acids were synthesized from decanoic acid and decanal and recrystallization yielded a 1:1:1:1 ratio of isomers.[60] The *cis*-β-lactone was synthesized from decanoic acid *via* a ketene dimer that was subsequently hydrogenated to yield a *cis*-β-lactone.[44] *Trans*-β-lactone was separated from a *cis*- and *trans*-β-lactone mixture generated from the precursor β-hydroxy acid with sulfonyl chloride.[61] The *cis*-olefin was generated from the coupling of 1-decyne with 1-bromononane precursors followed by hydrogenation with Lindlar catalyst.[62,63] Photoisomerization of the *cis*-olefin generated the *trans*-olefin standard.[64]

Generating mutants of OleB and OleBC.

Site-directed mutations of OleB derived from the wild-type protein sequences from *Xanthomonas campestris* ATCC 33913 (WP_012437021.1) and *Micrococcus luteus* OleBC (WP_010078536.1) were made with New England Biolabs Q5 quick change site directed mutagenesis kit following manufacturer's instructions. All primers were ordered from Integrated DNA Technologies (IDT). To confirm each mutant, single colonies were grown in 5 mL cultures at 37 °C overnight under kanamycin selection. Plasmids were isolated using a QIAGEN Miniprep kit and sent to ACGT Inc for sequencing.

Purification of OleB and OleBC fusion.

The buffer for OleB purification contained 200 mM NaCl, 20mM $NaPO_4$, 10% glycerol, and 0.5% PEG 400 (Hampton Research) at pH 7.4. *E. coli* BL-21 DE3 cells containing OleB with a 6x Histidine tag on the N-terminal were, grown, sonicated, and crude protein was purified using a Ni^{2+} column. Protein

concentrations of purified OleB solutions were measured by Bradford assay. Purified OleB solutions were routinely stored at -80°C. OleBC and $OleB_{D163A}C$ fusion proteins from *Ml* were generated with a 6x Histidine tag and purified as described previously.[58]

OleB reactions with β-lactone followed by ^1H NMR.

For enzyme reactions, the appropriate substrate *cis*- or *trans*-β-lactone was first dissolved in ethanol at 0.17 mg/ml. Reactions were carried out in separatory funnels containing 1.0 mg *X. campestris* OleB (or *M. luteus* OleBC fusion), 3.0 ml of the β-lactone substrate, 10 µl of 10% 1-bromonaphthalene as an internal standard, and 100 ml buffer (200 mM NaCl, 20 mM NaPO4, pH 7.4), and incubated at room temperature overnight. Reactions were extracted twice, with 10 ml and 5 ml methylene chloride, consecutively. The organic extracts were pooled and back-extracted with 15 ml double-distilled H_2O. The organic fraction was dried, dissolved in $CDCl_3$, and placed in 5 mm NMR tubes with tetramethylsilane (TMS) as a reference. A Varian Inova 400 MHz NMR spectrometer using a 5 mm Auto-X Dual Broadband probe at 20°C was used for all spectral acquisitions. Spectra were typically acquired using 1,024 pulses with a 3 second pulse delay.

OleB reactions with haloalkanes.

The following haloalkane substrates were dissolved in ethanol to a concentration of 5 mM for testing with *Xc* OleB: 1-iodobutane, 1,3-iodobutane, 1-chlorobutane, 1-bromopentane, 1-chlorohexane, 1-bromooctane, 1-iodoundecane, and 7-(bromomethyl)pentadecane. Reactions were carried out in glass GC vials contained purified OleB (40 μg) and 10 μL of substrate in 500 μL of 200 mM NaCl, 20 mM $NaPO_4$ at pH 7.4. Reactions were incubated at room temperature overnight, followed by extraction with *tert*-butyl methylether (MTBE). The MTBE extract was transferred to a clean GC vial and analyzed by gas chromatography/mass spectrometry (GC/MS Agilent 7890a & 5975c with an Agilent J&W bd-ms1 column 30 m length, 0.25 mm diameter, 0.25 μm film).

Bioinformatic analysis of OleB and haloalkane dehalogenases.

Sequences for representative members of the α/β-hydrolase protein superfamily were retrieved from the Protein Data Bank (PDB) on 3/13/2017 using the SCOP classification for α/β-hydrolases and filtered to include only representative bacterial sequences for each protein family (Figure S1). To obtain a higher resolution phylogeny for the relationship of OleB/OleBC sequences with haloalkane dehalogenases (Figure 3-3), accession numbers for characterized HLD-I, -II, and -III accession numbers were pulled from Nagata *et al*. (2015). Five experimentally characterized OleB and OleBC sequences were obtained from Sukovich *et al*.[34] Protein sequences were aligned and curated using the DECIPHER package in R.[65] Due to the length of OleBC fusion sequences interfering with proper alignment, the last 550 residues were trimmed to eliminate the OleC region. Maximum-likelihood phylogenies with 100 bootstrap replicates were inferred from alignment using the JTT method using the phangorn R package.[66] A structural homology model for the Xanthomonas campestris OleB sequence (WP_012437021.1) was built using default parameters in Phyre2.[67]

Mass spectroscopy of acyl-enzyme intermediate with haloalkanes.

To identify an acyl-enzyme intermediate, Matrix Assisted Laser Desorption Ionization (MALDI) was carried out on wild type OleB and $OleB_{D114A}$ proteins that had been reacted with 7-(bromomethyl)pentadecane (TCI). These two substrates contain the reactive chemical group in the middle of long di-alkyl tails, thereby mimicking the β-lactone substrate of OleB. Reactions contained 500 μM substrate and 40 μg of OleB in 100 μL of buffer (20 mM NaCl, 5 mM $NaPO_4$, pH 7.4). Reaction were prepared for MALDI using standard C_4 ZipTip (Millipore) procedures and spotted on a plate with sinapinic acid. Samples were analyzed on a Bruker Autoflex Speed MALDI-TOF.

Results

Purification of monomeric OleB without detergents.

The OleB protein from *X. campestris* had been purified previously in a study showing that OleB, OleC, and OleD combine to form large enzyme assemblies on the order of 2 MDa molecular weight.[57] The individual activity of the protein was not demonstrated in that study. Moreover, in that previous report, OleB purification required the presence of 0.05% Triton X-100 to maintain the protein in a soluble form. Despite that, the purified OleB protein formed large, non-homogeneous aggregates when not in admixture with OleC and OleD. In the present study, it was discovered that the addition polyethylene glycol (PEG 400) to purification buffers stabilized OleB, making it more amenable to purification and concentration. Purification yields increased to 8 mg/L of culture compared to 2 mg/L previously reported.[57] Unlike the OleB purified in Triton X-100, the protein purified with PEG 400 migrated largely as a monomer by gel filtration. This monomeric protein form was used in these studies, although the Triton-purified OleB was shown to catalyze the same reaction as described below.

OleB utilizes only *cis*-β-lactones.

It was necessary to synthesize β-lactones containing two hydrocarbons tails in the range of C_8-C_{14} to determine if OleB catalyzes the terminal reaction of long-chain olefin biosynthesis. Those chain lengths were previously shown to be in the biologically relevant range. Both *cis*- and *trans*-3-octyl-4-nonyl-2-oxetanone (*cis*- and *trans*-β-lactone) were chemically synthesized and used to determine if OleB catalyzes a decarboxylation reaction. ^{1}H NMR demonstrated that both the *cis*- and *trans*-β-lactones enantiomeric pairs contained <10% of the opposite configuration. ^{1}H NMR analysis of OleB reactions (Figure 3-2) showed that 47% of the *cis*-β-lactone underwent decarboxylation to the *cis*-olefin in an overnight reaction that appears to have gone to completion. This suggested that OleB selectively acts on only one of the *cis*-β-lactone enantiomers. An OleB reaction mixture with *trans*-β-lactone showed only 4% underwent decarboxylation to a *cis*-olefin. This 4% product is likely caused by the small contamination of *cis*-β-lactone in the *trans*- β-lactone sample. Olefin was undetectable in control reactions lacking enzyme. A synthetic *trans*-olefin standard was prepared to aid in analytical methods, but there was no evidence for this compound being produced in OleB reaction mixtures. This observation agrees with multiple literature reports that bacteria exclusively produce *cis*-olefins.[31,33] Taken together, these data supported the idea that OleB acts physiologically to catalyze decarboxylation of *cis*-β-lactones to yield *cis*-olefins that complete the olefin biosynthetic pathway.

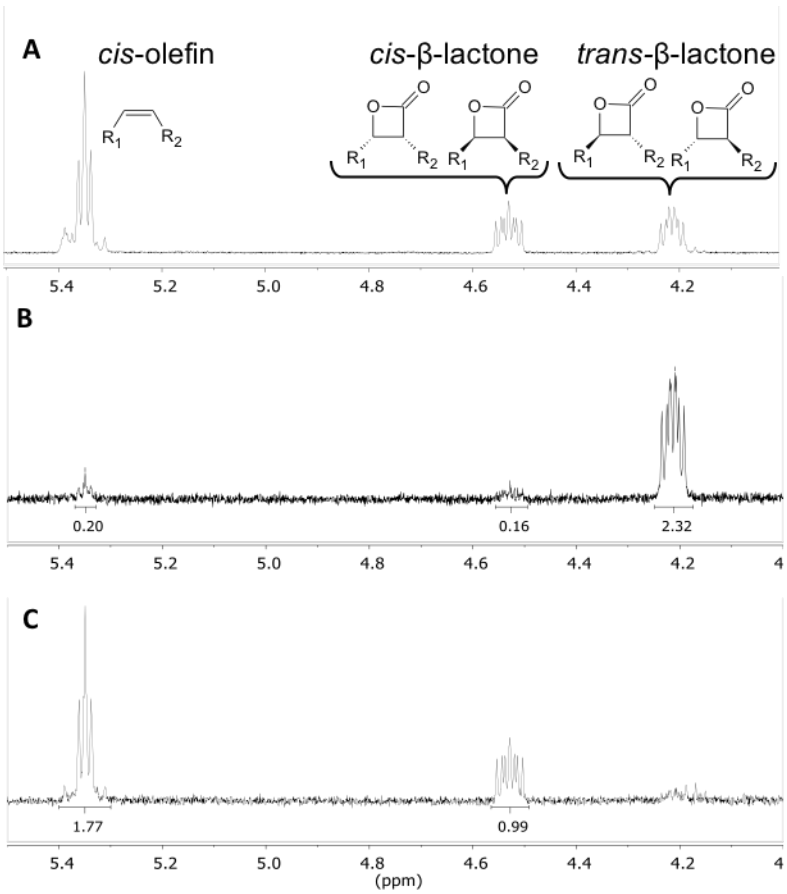


Figure 3-2. *OleB decarboxylation of cis-β-lactones to cis-olefins followed by* ^{1}H NMR*. (A)* ^{1}H NMR *showing synthetic standards of cis- and trans-3-octyl-4-nonyloxetan-2-one and cis-9-nonadecene. (B)* ^{1}H NMR *for reaction of OleB with synthetic trans-β-lactone (minor cis- β-lactone centered at 4.55 shown) showing no reaction towards this enantiomeric pair. The small peak for cis-olefin is believed to originate from the minor cis-β-lactone contaminant. (C) OleB + cis-β-lactone showing approximately half of the starting material has been converted to cis-olefin suggesting that only one of the cis-enantiomers reacted with OleB.*

OleB clusters with type-III haloalkane dehalogenases.

The sequences of *X. campestris* and other OleB proteins were aligned with the broad class of bacterial α/β-hydrolases to identify nearest evolutionary relationships. A sequence was considered to be OleB if it derived from organisms shown to produce olefins or when the genes for *oleACD* could be identified within 3 genes of *oleB*. OleB protein sequences clustered most closely with haloalkane dehalogenases (HLDs) (Figure S1). Within the OleB sequences, there was also a clear distinction between the independent OleB proteins, typically found in gram-negative bacteria, and the natural OleBC fusion proteins produced in gram-positive *Actinobacteria*.

The α/β-hydrolase superfamily is highly divergent necessitating a finer scale clustering of only HLD and OleB sequences. A small subset of HLDs and OleB proteins is shown in Figure 3-3A. Our analysis recovered the classification of HLDs into three subgroups: HLD-I, -II, and -III.[59] Unexpectedly, OleB sequences were not a separate cluster, but were interspersed with the HLD-III subgroup. Closer analysis of the genomic regions for those 36 putative HLD-III sequences suggested in Chovancova's original work found that 72% were part of an *oleABCD* gene cluster (Figure S2), indicating that those proteins likely function as β-lactone decarboxylases.[59] Additionally, the only two HLD-III proteins characterized to date are reported to have "very low activities with typical substrates of haloalkane dehalogenases."[68] In light of this, we believe at least a portion of the HLD of subgroup III has been misannotated, and should instead be considered β-lactone decarboxylases. This connection is explored in greater depth below and in the discussion.

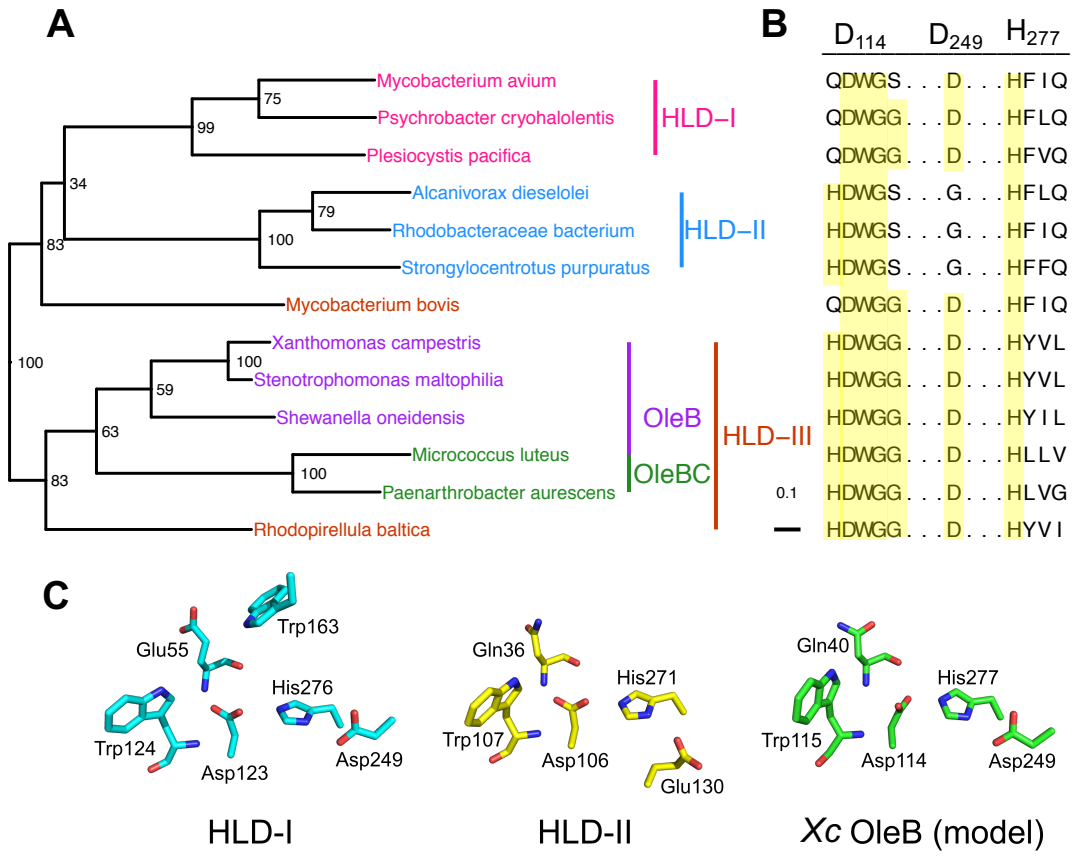


Figure 3-3. *Sequence analysis and structural modeling of OleB and haloalkane dehalogenase proteins. (a) Phylogenetic tree of OleB and OleBC sequences aligned with characterized members of the HLD family separated into classes I, II and III (Chovancova et al.).*[59] *OleB and OleBC sequences cluster with the HLD class III Rhodopirellula baltica sequence* (CAM90600.1)*. Only 13 sequences were included in the final alignment for ease of viewing. Unrooted maximum-likelihood tree was estimated using the Jones Taylor Thornton model of amino acid evolution. Bootstrap values are displayed at each node (100 data resamplings). The low bootstrap value between HLD-I and HLD-II is consistent with previous work.*[59] *Scale bar represents 0.1 changes per amino-acid position. (b) Multiple sequence alignment revealed a putative catalytic triad of amino acids. Numbering of the proposed catalytic triad at the top is based on the amino acid position in the Xanthomonas campestris OleB sequence (WP_012437021.1) studied here. (c) Crystal structures of the HLD-I and HLD-II (2XT0 - Plesiocystis*

pacifica and 4BRZ - Rhodobacteraceae bacterium) were compared with a homology model of Xc OleB generated by Phyre2.[67,69,70]

OleB and HLD alignments and site-directed mutagenesis suggest catalytic residues.

Sequence alignments of OleB proteins and well-studied HLDs were performed in order to identify residues that might be directly involved in catalysis (Figure 3-3B). X-ray crystal structures and mutagenesis studies have delineated the catalytic residues and mechanism of HLD Is and HLD-IIs, but little is known about HLD-III class proteins and no structures are available. OleB from *X. campestris* showed significant sequence identity with HLDs and allowed the putative assignment of the catalytic triad in OleB as Asp_{114}, His_{277} and Asp_{249}. These residues are completely conserved in all OleB sequences and align perfectly with HLD-I. The HLD-II enzymes are known to utilize a glutamate derived from the end of the beta-sheet 6 in place of D249. Despite that, a comparison of X-ray structures from the HLD-I and -II classes with a homology model of the *X. campestris* OleB protein suggests that the catalytic triad of D114, D249, and H277 may be isostructural and isofunctional between OleB and all HLD proteins (Figure 3-3C). The comparable D114 residue has been identified to serve as a nucleophile for halide displacement in HLD reactions. The backbone nitrogen of Trp124 and Glu55 from HLD-I and the equivalent Trp107 and Gln36 in HLD-II are known stabilize the oxyanion intermediate of haloalkane dehalogenation.[69,70] The sidechain nitrogens of Trp124 and Trp163 in HLD-I and Trp107 and Gln26 in HLD-II are known to stabilize the displaced halide atom during the catalytic cycle.[69,70] The equivalent residues are completely conserved in OleB proteins.

Site-directed mutagenesis of the *X. campestris* OleB protein was conducted to test the hypothesis that the residues identified by alignments and structural modeling might comprise a catalytic triad. Three mutants were made and tested for activity: D114A, H277A, and D249A. $OleB_{D114A}$ and $OleB_{H277A}$

showed no detectable activity towards *cis*-β-lactones when monitored by ^{1}H NMR. $OleB_{D249A}$, however, showed decarboxylation activity lower than wild-type OleB. The ^{1}H NMR assay used here did not allow us to measuring steady-state kinetic parameters. Moreover, the OleB substrates lack UV/Vis absorbance, have very poor solubility in water, and are thermally unstable making assays difficult. However, alternate assay methods are currently under investigation and may allow the determination of kinetic parameters in future studies.

OleB shows no detectable dehalogenase activity towards haloalkane substrates.

OleB from *Xc* was tested with haloalkane substrates to assess dehalogenase activity. Jesenska *et al*. tested 30 haloalkane substrates with two purified HLD-III proteins and found very limited activity with a select number of compounds.[68] Substrates showing the highest activity in those studies and substrates containing long alkyl tails similar to native OleB substrates were tested here. No detectable activity was observed against the following haloalkane substrates: 1-iodobutane, 1-iodoundecane, 1-cholohexane, 1-bromohexane, 1-chlorobutane, 1-bromobutane, and 7-(methylbromo)pentadecane when monitored by GC-FID/MS. The level of activity of activity is less than 0.2 h^{-1}.

Haloalkane substrate mimics form stable acyl-enzyme intermediates to inhibit OleB.

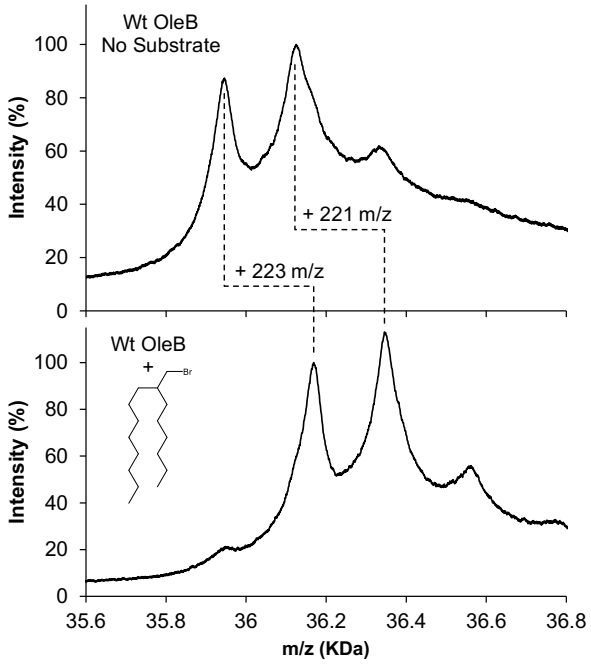


Figure 3-4. *OleB forms a stable acyl-enzyme intermediate when reacted with 7-(bromomethyl)pentadecane. OleB show a mass shift ~222 m/z consistent the nucleophilic attack and displacement of bromine with the substrate. The mass shift expected with the loss of bromide is 225 m/z. $OleB_{D114A}$ did not show any mass shift when reacted with the bromo-alkane substrate. This data is consistent with a haloalkane dehalogenase like mechanism. The two major peaks and one minor peak appear in all OleB and OleB mutant MALDI-TOF experiments and are presumed to be the result of an ion with a m/z of ~180.*

HLD proteins react to form an acyl enzyme intermediate between the nucleophilic Asp and the substrate. To investigate whether OleB might form a covalent enzyme intermediate, OleB was reacted with 7-(bromomethyl)pentadecane and a shift in protein mass was examined by MALDI-TOF mass spectrometry. A mass shift in OleB corresponding to the mass of the

debrominated alky chain was identified (Figure 3-4). A parallel incubation without the brominated substrate analog served as a control and showed the expected mass of the wild-type OleB. Another experiment was conducted with the putative nucleophilic residue, D114, mutated to an unreactive residue alanine. The D114A mutant OleB enzyme did not show a mass shift when incubated with 7-(methylbromo)pentadecane, suggested of the role D144 is similarly to its function in haloalkane dehalogenases. The modified wild-type protein appeared to be stable, suggesting that if an acyl-intermediate is being formed, OleB is unable to remove the alkyl group.

OleBC fusion from *Micrococcus luteus* functions as Xc OleB.

To ensure that previous findings are not confined to the single *Xc* OleB protein, the *M. luteus* OleBC fusion protein was purified and assayed here. The OleB domain of *M. luteus* is only 32% identical to *Xc* OleB and the OleC domain is known to have β-lactone synthetase activity.[58] When the nucleophilic Asp of *M. luteus* OleBC fusion (Asp163) was mutated to Ala and reacted with OleC substrate, a mixture of *syn*- and *anti*-β-hydroxy acids, only *trans*- and *cis*-β-lactones were observed. However, under the same conditions, the wild-type *Ml* OleBC fusion protein formed less *cis*-β-lactone, and resonances consistent with *cis*-olefin appeared (Figure 3-5). These findings are consistent with OleB acting on a single *cis*-β-lactone enantiomer to yield a *cis*-olefin in a reaction dependent on D163 as a nucleophile. Additionally, these data demonstrate that divergent sequences within the HLD-III subgroup have β-lactone decarboxylase activity.

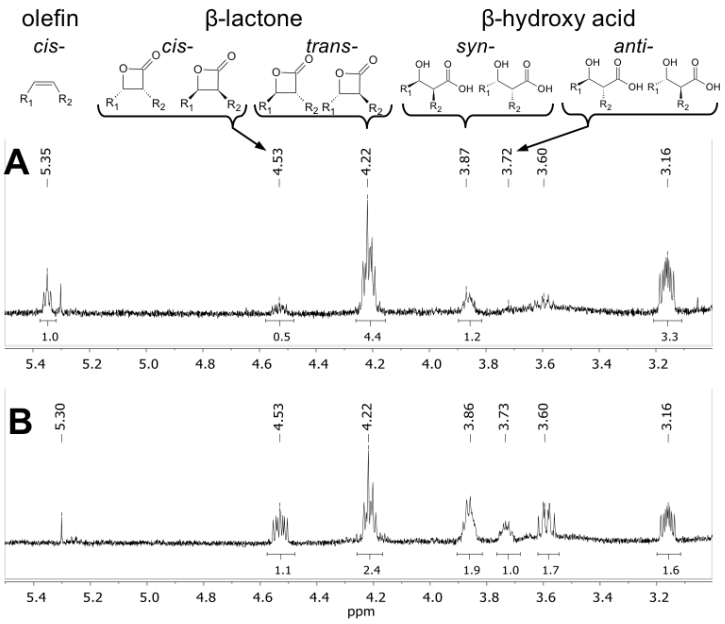


Figure 3-5. *Natural OleBC fusion from MI when reacted with β-hydroxy acids. (A) Wild type OleBC fusion accumulates trans-β-lactone as well as equivalent amounts of cis-olefin and cis-β-lactone consistent with only one of the cis-β-lactone enantiomers reacted with by the OleB domain. (B) The functional OleC domain of $OleB_{D114A}C$ can convert syn- and anti-β-hydroxy acids to cis- and trans-β-lactones respectively, but the mutant OleB domain does not generate cis-olefin.*

Discussion

To our knowledge, OleB is the first enzyme reported to decarboxylate a β-lactone to form a *cis*-olefin. There are other known α/β-hydrolase superfamily members from plants that perform decarboxylation reactions, such as MKS1 from *Solanum habrochaites* (wild tomato), that decarboxylate β-keto acids to methylketones.[71] However, these show only ~12% sequence identity to *Xc* OleB and are reported to rely on a completely different mechanism. Additionally, α/β-hydrolase superfamily members, such as AidH, from *Ochrobactrum sp*. are known to hydrolyze five-membered γ-lactone rings of quorum sensing molecules to 4-hydroxy acids.[72] Again, these lactonases show little sequence identity to

OleB (~19%) and contain a serine at the catalytic nucleophile suggesting a different mechanism.

OleB appears to react preferentially with only one enantiomer of the synthetic *cis*-β-lactone pair. The preceding pathway enzymes, OleA and OleD, are known to generate the 2R,3S-configuration in the β-hydroxy acid. OleC is believed to retain this stereochemistry during its ring closure reaction to the β-lactone. As such, it is likely that OleB acts on the 2R,3S-β-lactone to produce a *cis*-olefin, but studies to identify the chirality of the remaining lactone must be conducted. No *trans*-olefin was ever observed, consistent with multiple literature reports that the Ole pathway exclusively produces *cis*-olefin.

Sequence alignments and homology modeling reveal OleB is closely related to HLDs. Chovoncova *et al.* described three subfamilies of HLDs (I, II, and III).[59] However, 72% of the HLD-III subfamily from this original work were found to be encoded in *oleABCD* gene clusters. Both subfamilies I and II have multiple crystal structures and the mechanisms of these enzymes are well understood (Figure 3-3C), but no structures are available for HLD-IIIs. The two previously characterized subfamily III members have poor dehalogenase activity and the HLD-III OleB has no detectable dehalogenase activity.[68] However, the annotated HLD-IIIs, *Xc* OleB and *Ml* OleBC fusion, were found to have β-lactone decarboxylase activity, indicating at least part of this HLD-III subgroup is misannotated. Further bioinformatics work, coupled with biochemical data, is needed to distinguish between these two enzyme functions. Additionally, the enzymatic function of sequences that cluster with HLD-IIIs (OleBs), but are not part of *oleABCD* gene clusters, must be explored.

We find no reason to significantly deviate from the well-studied mechanism of haloalkane dehalogenases when considering a mechanism for the decarboxylation of β-lactones. In both sequence and structural alignments, the conserved Asp114 from *Xc* OleB aligns perfectly with the nucleophilic aspartic acid of haloalkane dehalogenases. Additionally, MALDI-MS of OleB and $OleB_{D114A}$ implicates this Asp as the critical nucleophile to generate the canonical

acyl enzyme intermediate in the HLD mechanism. The function of *Xc* OleB is dependent on His277 consistent with its complete conservation within the entire α/β-hydrolase superfamily. The role of the second acidic residue (Asp249 in HLD Is or Glu130 in HLD-IIs) in maintaining the correct protonation state of His for the activation of water agrees with our data that *Xc* OleB is slower when Asp249 is mutated to an Ala. Considering the aforementioned data, we propose the following β-lactone decarboxylation mechanism for OleB.

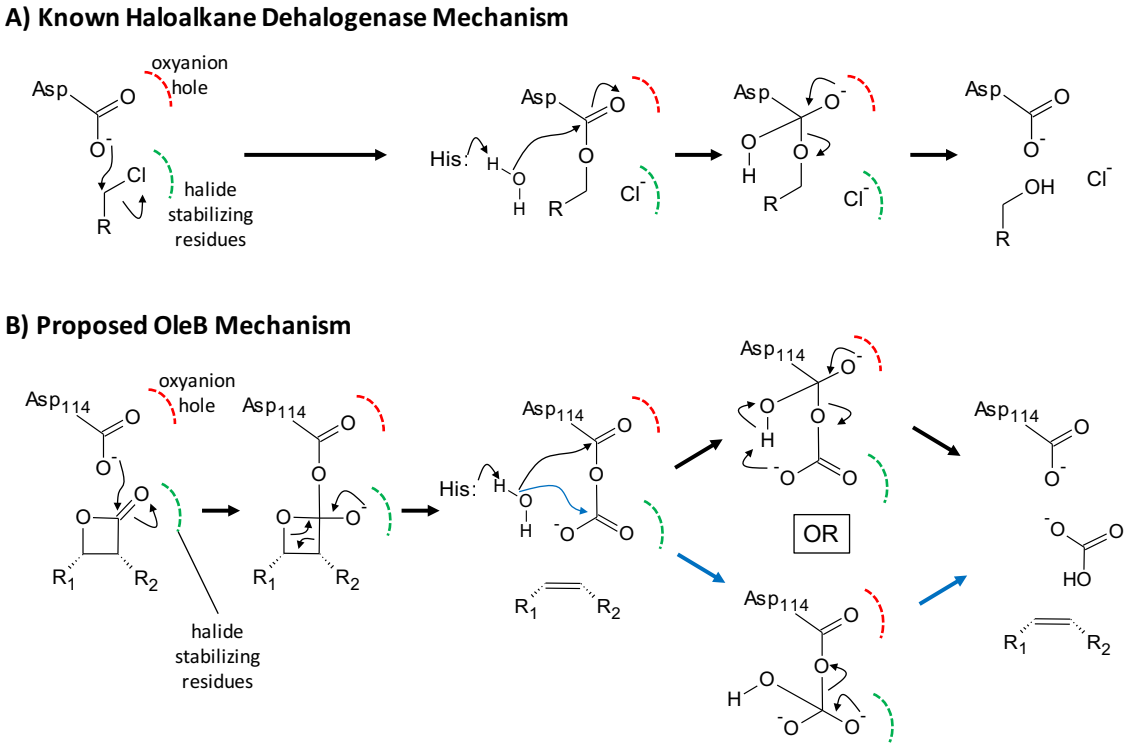


Figure 3-6. *Proposed mechanism of OleB as a β-lactone decarboxylase. (A) General mechanism of HLDs as described by multiple reviews.*[56,73] *(B) Analogous mechanism using β-lactone as a substrate with Xc OleB. The putative anhydride intermediate allows two possible mechanisms of resolution for OleB, and it is currently unclear which occurs.*

Figure 3-6 shows the canonical mechanism for HLDs and proposes an analogous mechanism for OleB. The nucleophilic Asp114 of OleB attacks the carbonyl carbon of the β-lactone ring to generate a tetrahedral intermediate. The side chains of Trp115 and Gln40 are in equivalent spatial and sequence positions to act as halide stabilizing residues, but no halide is present in the lactone moiety. Instead, these residues could act to stabilize the oxyanion in first tetrahedral intermediate. This first tetrahedral intermediate resolves to expel the olefin product and generate an anhydride as the equivalent to the acyl enzyme intermediate of HLDs.

There are now two possible centers for the attack of water activated by His277, the carbonyl of aspartic acid, or the carbonyl originating from the β-lactone. In favor of the Asp (upper pathway, Figure 3-6B) is the fact that this is the canonical pathway for HLDs and presumably contains the optimal bond angles and distances for attack. Additionally, the backbone nitrogens that create the oxyanion hole in HLDs (X of the HGXP motif and Trp adjacent to the nucleophile) are in the same spatial position in our model and are 100% conserved across all OleB sequences. However, in favor of the lower pathway (Figure 3-6B) is the biochemical evidence that no haloalkane substrates turn over with OleB. Hydroxide attack of the lower carbonyl nicely explains the trapping of the acyl-enzyme intermediate when OleB is reacted with 7(bromo-methyl)pentadecane. Additionally, the resulting second tetrahedral intermediate would be at the same site as the first proposed in step two. This mechanism is simpler, as OleB would only need to have the necessary residues to stabilize an oxyanion in one location rather than two. Regardless of the pathway, resulting products are identical: alkene, bicarbonate, and the regenerated enzyme.

In summary, OleB is concretely defined as the final step of the long-chain olefin biosynthesis pathway by decarboxylating the β-lactone product of OleC. OleB shows many similarities to haloalkane dehalogenases and comprises most of the sequences reported in the HLD subgroup III suggesting a misannotation of

this group of enzymes. OleB proteins contain the conserved Asp-His-Asp/Glu catalytic triad of HLDs, and current evidence supports an analogous mechanism.

CHAPTER 4 OleBCD Assemblies

"We are only as strong as we are united, as weak as we are divided."
— J.K. Rowling (Harry Potter and the Goblet of Fire)

J. Bacteriol. **2017**, 199 (9), pp e00890-16.
doi: 10.1128/JB.00890-16.

Active Multi-Enzyme Assemblies for Long-Chain Olefinic Hydrocarbon Biosynthesis

James K. Christenson, Matthew R. Jensen, Brandon R. Goblirsch, Fatuma Mohamed, Wei Zhang, Carrie M. Wilmot, and Lawrence P. Wackett

OleBCD Summary: Bacteria from different Phyla produce long-chain olefinic hydrocarbons derived from an OleA-catalyzed Claisen condensation of two fatty acyl-CoA substrates, followed by reduction and oxygen elimination reactions catalyzed by the proteins OleB, OleC, and OleD. In this report, OleA, OleB, OleC, and OleD were individually purified as soluble proteins, and all were found to be essential for reconstituting hydrocarbon biosynthesis. Recombinant co-expression of tagged OleABCD proteins from *Xanthomonas campestris* in *Escherichia coli* and purification over His_{6x}- and FLAG-columns resulted in OleA separating whilst OleBCD purified together, irrespective of which of the four Ole proteins were tagged. Hydrocarbon biosynthetic activity of co-purified OleBCD assemblies could be reconstituted by adding separately-purified OleA.

Immunoblots of non-denaturing gels using anti-OleC reacted with *X. campestris* crude protein lysate indicated the presence of a large protein assembly containing OleC in the native host. Negative stain electron microscopy of recombinant OleBCD revealed distinct, large structures with diameters primarily between 24-40 nm. Assembling OleB, OleC, and OleD into a complex may be important to maintain stereo-chemical integrity of intermediates, facilitate the movement of hydrophobic metabolites between enzyme active sites, and protect the cell against the highly reactive β-lactone intermediate produced by the OleC-catalyzed reaction.

Importance

Bacteria biosynthesize hydrophobic molecules to maintain a membrane, for carbon storage, and for antibiotics that help them survive in their niche. The hydrophobic compounds are often synthesized by a multi-domain protein or by large multi-enzyme assemblies. The present study reports on the discovery that long-chain olefinic hydrocarbons, made by bacteria from different Phyla, are produced by multi-enzyme assemblies in *X. campestris*. The OleBCD multi-enzyme assemblies are thought to compartmentalize and sequester olefin biosynthesis from the rest of the cell. This system provides additional insights into how bacteria control specific biosynthetic pathways.

Introduction

Long-chain olefinic hydrocarbons are abundant in nature. First described in 1929 in cabbage leaf extracts, plant C_{29} ketones and hydrocarbons were proposed to arise from a head-to-head condensation of fatty acids, with a loss of one of the "head" carboxylic acid groups as carbon dioxide.[74] Bacterial long-chain olefinic hydrocarbons were characterized structurally in 1969 during studies on the lipid components of *Micrococcus luteus*.[31]

The genes and enzymes responsible for the head-to-head mechanism of hydrocarbon biosynthesis in bacteria were elucidated some forty years later by the company LS9, Inc. (now Renewable Energy Group) and patented.[32] A subsequent study in 2010 reported the presence of the four hydrocarbon biosynthetic genes, *oleABCD*, in sixty-nine prokaryotes from many deeply rooted branches of the prokaryotic tree of life (Figure. 4-1).[34] The cellular function(s) of the *ole* gene cluster has only been studied in *Shewanella onidensis* MR-1. The presence of the *oleABCD* genes were found to promote more rapid cell growth during a shift to colder temperatures, consistent with those genes being commonly found in cold temperature bacteria from polar and marine environments.[34,25,75] However, bacteria containing *ole* gene clusters inhabit a diversity of ecological niches, suggesting there may be a range of functions for long-chain olefinic hydrocarbons.

Only recently have the activities of each Ole enzyme been assigned (Figure. 4-1), as initial work was hampered by the instability and hydrophobicity of the pathway intermediates. OleA is a soluble protein that condenses C_{10}-C_{16} fatty acyl groups *via* a non-decarboxylative Claisen condensation reaction to form a β-keto acid.[36] If the β-keto acid is not quickly reduced by OleD, it spontaneously decarboxylates in aqueous solution to a ketone, as observed in earlier studies.[35,36] Physiologically, OleD catalyzes an NADPH-dependent reduction of the β-keto acid intermediate to produce a β-hydroxy acid that is significantly more stable than the OleD substrate.[39] Until recently, OleC was thought to catalyze the final reaction in olefin biosynthesis. However, OleC is now

known to react with the product of the OleD reaction to generate a thermally labile β-lactone.[58] The β-lactone undergoes spontaneous, and non-biological, decarboxylation to an olefin when monitored by gas chromatography (GC), which led to the initial incorrect assignment of OleC function.[40,58] Recent data suggest that OleB acts as an unprecedented β-lactone decarboxylase to yield the final olefin product.[58] Homologous gene clusters lacking an *oleB* gene have been identified in *Streptomyces*, and these produce β-lactone natural products rather than olefins.[58] Microbial secondary metabolites containing β-lactones often serve as antibiotics and are known as general esterase and protease inhibitors.[3,16]

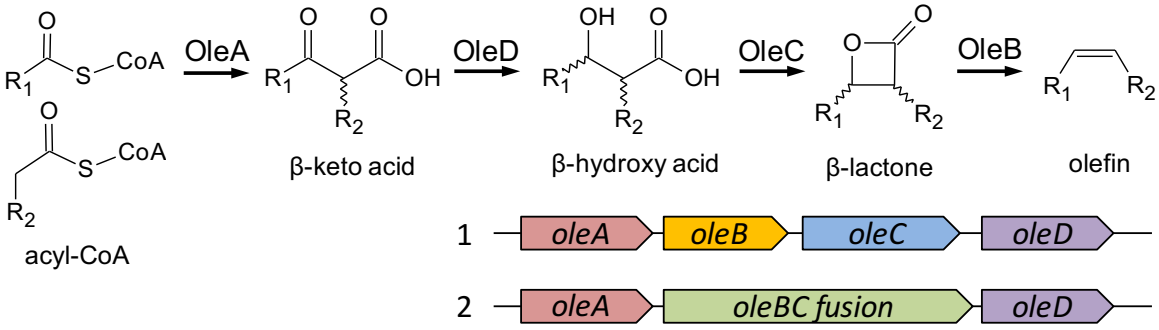


Figure 4-1 *Olefin biosynthesis enzymatic pathway and gene clusters. Gene clusters 1 and 2 are the two most common gene arrangements found in oleABCD containing organisms. The two fatty acyl CoAs are typically linear alkyl chains between 10 and 16 carbons in length.*

It was considered here that Ole enzymes might form a multi-enzyme complex to effectively process the hydrophobic, unstable, and potentially toxic intermediates of olefin biosynthesis. The presence of an *oleBC* gene fusion in actinobacterial *ole* gene clusters further suggested a physical interaction amongst the enzymes of olefin biosynthesis. In the present study, we report the purification of active protein assemblies consisting of OleB, OleC, and OleD proteins. Assemblies of OleBCD were obtained by recombinantly co-expressing all four *ole* genes from *Xanthomonas campestris pv. campestris* ATCC 33913 (*X. campestris*) in *Escherichia coli* with various affinity tag combinations. Non-

denaturing gels immunoblotted with anti-OleC showed OleC from *X. campestris* cell lysate migrated as a high molecular band, similarly to purified OleBCD assemblies. Results from size-exclusion chromatography and electron microscopy were consistent with ordered assemblies of OleBCD with an average molecular weight of ~1.9 MDa.

Results

Individual purification and physical properties of the four OleABCD enzymes.

The purification of individually-expressed OleA, OleC, and OleD proteins derived from different bacteria had been reported previously, whilst the purification of an OleB protein is reported for the first time here. In this study, all four *X. campestris* Ole proteins were purified in a recombinant form from separate *E. coli* expression host cell-lines.[25,35,36,39,40,58] In contrast to conclusions reached in a previous study, we found that OleB is required for reconstituting hydrocarbon biosynthesis, consistent with the ubiquitous occurrence of *oleB* in the olefin gene cluster.[40] The OleA and OleD reactions had been previously demonstrated to produce a β-keto acid and β-hydroxy acid, respectively (Figure 4-1).[36,39] OleC, purified from *Stenotrophomonas maltophilia,* was reported to produce an olefin, but we recently demonstrated that OleC from that organism produces a β-lactone, and that OleB is required to transform that to the final olefinic hydrocarbon.[40,58]

Relevant physical properties of the Ole proteins are reported in Table 1. The oligomeric state of each protein was determined by gel filtration chromatography. However, detergents were required in the purifications of OleB and OleD, which could affect the observed oligomeric states. Moreover, all attempts to concentrate either protein resulted in immediate precipitation, even in the presence of detergents. Glycerol increased purification yields of OleB, OleC, and OleD to the level shown in Table 1, but did not boost the yields of OleA. Since the proteins all appeared to be in the soluble fraction of the cell, the

insoluble nature of OleB and OleD in the absence of detergents suggested to us that Ole proteins might associate with each other *in vivo*.

Table 4-1

	X. campestris Ole Proteins			
Properties	OleA	OleB	OleC	OleD
Subunit MW (KDa)	36.6	34.1	58.5	36.1
Oligomeric State	Dimer	Tetramer	Monomer	Aggregate[1]
Required Detergent	-	Triton X-100	-	Tween 20
Yield (mg/L culture)	15	1.9	5.8	1.1
Precipitated on concentration?[2]	No	Yes	No	Yes

[1] *Xc* OleD runs >600 KDa on SEC, but remains active and soluble. *Sm* OleD is a dimer (9).

[2] The sample precipitated after spin concentration in 30,000 MWCO Amicon spin filters

OleBCD forms active multi-enzyme assemblies without OleA.

To examine our hypothesis that Ole enzymes form a multi-enzyme complex, we simultaneously expressed *oleABCD* in *E. coli* with different affinity tag arrangements (#1-3) shown in Figure 4-2A. Tag arrangement #1 contained an N-terminal His_{6x} tag on OleB and a C-terminal FLAG tag on OleC (His_{6x}-BADC-FLAG). Even when no detergent was used, three clear SDS-PAGE bands corresponding to the size of OleC, OleA and/or OleD, and OleB eluted from the Ni^{2+} column when the imidazole concentration reached ~140 mM. OleA and OleD are nearly indistinguishable by SDS-PAGE as they differ by only 508 Daltons. Subsequent purification over an anti-FLAG column further purified these three bands. However, the addition of myristoyl-CoA substrate and cofactors generated no olefin product unless separately purified OleA was added back to the reaction mix. Reaction products of OleA with co-purified OleBCD were found to be identical with our synthetic *cis*-olefin standard (Figure S1). These data suggested that the association between OleBCD survives the purification, whereas OleA is lost during chromatography. For comparison, the individually expressed OleA, OleB, OleC, and OleD were reconstituted and also

shown to produce olefin from CoA-charged fatty acids. In both samples, the absolute configuration of the major product was *cis*-olefin. No appreciable difference in the rate of olefin production was observed, possibly because the rate-limiting step for *in vitro* assays is release of the hydrophobic olefin from OleB.

To examine if the tag placement altered OleA's binding to the putative complex of OleBCD, expression of tag arrangements 2 and 3 were undertaken. Tag arrangement #2 (His_{6x}-OleABDC-FLAG) resulted in the co-elution of three bands at ~100 mM imidazole from the Ni^{2+} column that were identified by mass spectrometry (MS) as OleB, OleC-FLAG, and OleD, while His_{6x}-OleA (MS identified) eluted separately at 200 mM imidazole. The observation that some OleBCD was retained on the Ni^{2+} column until imidazole was added, suggested an interaction with His_{6x}-OleA, which was disrupted prior to His_{6x}-OleA being competitively eluted off the column. Tag arrangement #3 (His_{6x}-BACD) resulted in the co-elution of OleBCD with no OleA activity. Taken together, these data suggest that OleBCD form active, stable multi-enzyme assemblies to which OleA weakly associates.

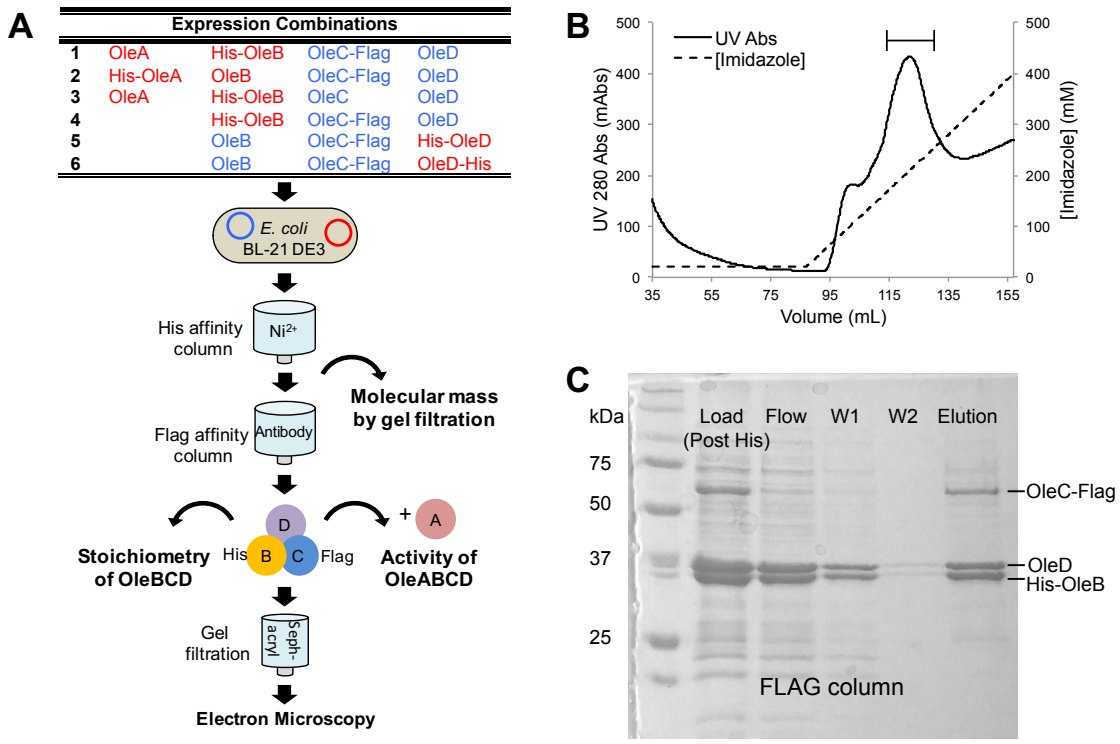


Figure 4-2 *Plasmid expression, purification, and SDS-PAGE analysis of Ole protein co-expressions. (A) The top table represents all tag combinations constructed. The placement of the His_{6x} and FLAG tags on the left and right of the protein correspond to N- and C-terminal tags respectively. Colors in the table represent the grouping of genes on a single vector for expression. (B) Purification of OleBCD (#4 tag arrangement) over His-affinity column with increasing amounts of imidazole. Peak fractions indicated by the bar were collected, concentrated, and loaded onto the anti-FLAG column (lane labelled "Load" on gel). (C) SDS-PAGE showing the post His-affinity column concentrate and the anti-FLAG purification of OleBCD #4. Bound samples were twice washed (W1 and W2) with buffer before elution with FLAG peptide. Note that OleC, the weakest intensity band, contained the FLAG tag.*

Optimizing the co-purification of the OleBCD.

To obtain pure assemblies of OleBCD routinely, OleA was no longer co-expressed with OleBCD; only *oleBCD* genes were expressed in the *E. coli* host. Of the three additional tag arrangements tested (#4-6, Figure 4-2A), only the purification of #4, His_{6x}-BDC-FLAG, yielded active assemblies after Ni^{2+} and anti-FLAG columns. In contrast to observations made with OleB and OleD proteins purified independently, the co-purified OleBCD did not require detergents and could be concentrated without precipitation. A His_{6x}-tag on OleD (either N- or C-terminal, tag arrangements #5 and #6) appeared to prevent formation of the enzyme assemblies. The purification of tag arrangement #4 is shown in Figure 4-2B,4- 2C. Glycerol was found to increase purification yields, but was not required to obtain active assemblies. The individual components of this OleBCD complex were confirmed by MS, and this sample was used in further characterization studies.

OleBCD complex identification in native X. campestris.

The level of expression of olefinic hydrocarbons in wild-type bacteria is modest, and so we used sensitive immunoblotting methods and non-denaturing gels to identify the anticipated multi-enzyme assembly within the native *X. campestris* host.[25] A polyclonal antibody against *X. campestris* OleC was raised against a specific peptide predicted to be on the protein surface, as described in the Materials and Methods. The polyclonal antibody was then validated by showing binding activity in immunoblots against purified OleC, co-purified OleBCD, and lysates of *E. coli* expressing OleBCD. Figure 4-3 shows a non-denaturing gel that had been immunoblotted in which a lysate fraction derived from *X. campestris* cells is compared with co-purified OleBCD and OleC alone. The band from the supernatant of *X. campestris* lysate migrates similarly to the OleBCD assemblies, and more slowly than OleC alone. These data are consistent with the idea that the assemblies of OleBCD that form in *E. coli* likely assemble in a similar manner in the native host, *X. campestris*. It was also found in the soluble fraction, similar to recombinant expression in *E. coli*.

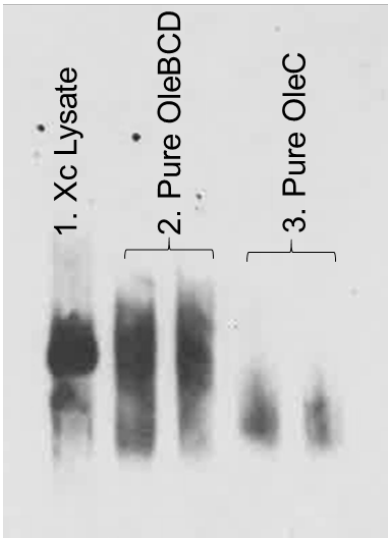


Figure 4-3 *Anti-OleC immunoblot on a non-denaturing protein gel. Native OleC from the supernatant of lysed X. campestris cells migrates closely with recombinantly expressed and purified X. campestris OleBCD rather than individually-purified OleC.*

Size estimation of OleBCD assemblies by gel filtration chromatography.

Gel filtration chromatography was used to obtain insight into the molecular size of co-purified OleBCD. Ni^{2+}column-purified OleBCD (#4) eluted in the void of a GE 16/600 Superdex 200 gel filtration column, indicating a size larger than 600 KDa. The void fraction was collected and found to be active upon addition of OleA, cofactors, and myristoyl-CoA. We subsequently attempted to purify OleBCD assemblies with detergents to see if a minimal functional complex could be isolated. However, the OleBCD still ran in the void with the addition of Tween 20 or Triton X-100, while CHAPS, a zwitterionic detergent, led to complete dissociation of OleBCD at all concentrations.

Subsequently, OleBCD was analyzed on a Sephacryl gel filtration column that is able to separate globular proteins up to 8 MDa. OleBCD (tag arrangement #4) eluted significantly after the void (45 mL) at a volume of 65.0 mL, consistent with a high molecular mass (Figure 4-4). The observed peak lies outside the

range of commercially available protein standards (up to 600 KDa), but a molecular mass centered at ~1.9 MDa was determined by extrapolation. The activity of each fraction was assayed in the presence of excess OleA, and activity tracked precisely with the A_{280} absorbance, consistent with stable, active assemblies of OleBCD (Figure 4-4). SDS-PAGE analysis of fractions showed peak OleBCD protein-staining intensity consistent with the UV absorbance and activity data.

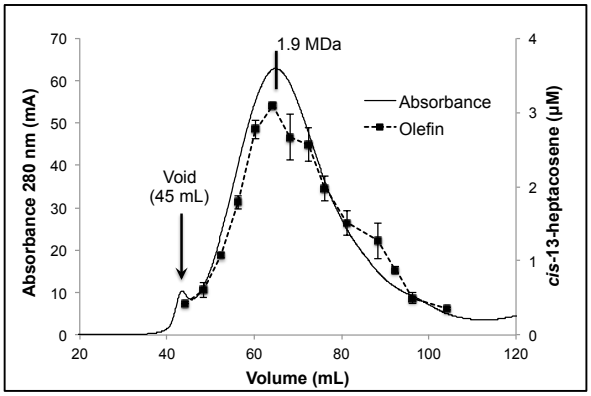


Figure 4-4 *Size-exclusion chromatography of Ni^{2+} column purified OleBCD (#4). The peak absorbance at 65.0 mL corresponded to an estimated size of 1.9 MDa when extrapolated from protein standards. 200 µL of each 4 mL fraction was tested for OleBCD activity by the addition of OleA, myristoyl-CoA and cofactors.*

Electron microscopy imaging of the OleBCD assemblies.

The suggested size based on gel filtration prompted analysis of OleBCD by transmission electron microscopy (TEM) using negative staining. This required purification of OleBCD (tag arrangement #4) without glycerol using Ni^{2+}, anti-FLAG, and gel filtration columns to obtain a highly purified sample. Glycerol increases the purification yield of OleBCD assemblies, but prevents sample dehydration for effective TEM imaging. Multiple purifications of OleBCD produced very similar electron micrographs, one of which is shown in Figure 5-5.

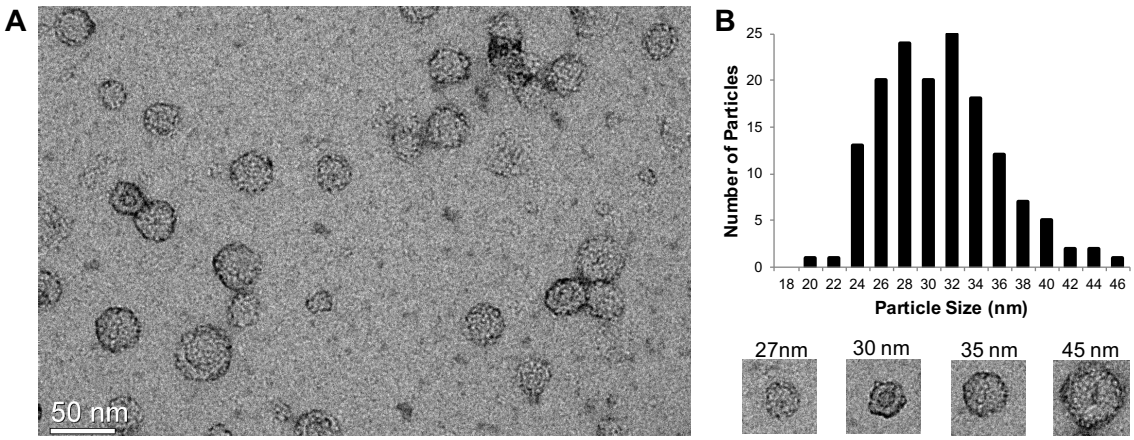


Figure 4-5 *Electron micrograph of OleBCD assemblies with particle size analysis. (A) OleBCD assemblies by TEM. (B) Histogram analysis of >150 particles from four different frames with representative micrographs shown below. Measurement was of the greatest particle diameter. Average assembly size is 30.3 nm and the median is 29.6 nm.*

TEM analysis of the negatively stained OleBCD revealed distinct assemblies for OleBCD, primarily ranging in size from 24-40 nm in diameter. The annular, hexagonal assemblies with an average diameter of ~27 nm are particularly striking for their organized structure (Figure 4-5 and Figure S2). Some of the sample heterogeneity may be different orientations of the same structure on the TEM grid, different size complexes, assembly intermediates, or the result of the dehydration and negative staining protocol required for TEM. We note that size-exclusion chromatography (SEC) does show that OleBCD activity is associated with a broad peak (Figure 4-4). The relationship between the assemblies pictured in Figure 4-5 remains unclear, but further work is underway to obtain more high-resolution structural information by cryo-electron microscopic (cryo-EM) methods.

Estimated stoichiometry of the OleBCD assemblies.

To estimate the stoichiometry of each of the OleB, OleC, and OleD proteins within the combined assemblies, gel band intensities of the purified OleBCD, shown in Figure 4-2C, were compared to band intensities of a gel

standard curve made with each purified protein. Figure S3 shows equimolar amounts of separately purified and quantified OleB, OleC, and OleD by SDS-PAGE as determined by Bradford assay and UV_{280}. Using ImageJ, the intensities of the bands were plotted against protein concentration to generate standard curves for each of the Ole proteins (Figure S4). Comparing the band intensities of the co-purified OleBCD assemblies (Figure 4-2C) with these standard curves revealed a stoichiometry of nearly 4:1:4 of OleB:OleC:OleD within the combined assemblies. Analysis of two other gels gave similar results. Taking this result with the ~1.9 MDa size as assigned by SEC, this would be most consistent with a single complex composed of 24, 6, and 24 subunits of OleB, OleC, and OleD, respectively. However, the broadness of the SEC peak, and the variable sizes of the particles observed by cryo-EM, could also be consistent with other multiples of the 4:1:4 stoichiometry. It is also possible that a mixture of different component combinations (perhaps assembly intermediates) lead to the overall observed 4:1:4 stoichiometry. However, olefin production upon addition of OleA, cofactors, and substrate occurs in all fractions following SEC, and activity tracks with the amount of protein in the fraction (Figure 4-4). This indicates that OleB, OleC, and OleD are present in all fractions, although they may not be in the same proportions. We note that the observed correlation between activity and total protein indicates only that the rate-determining step of olefin biosynthesis correlates with total protein.

Discussion

Many biosynthetic pathways are assembled as a single large protein with multiple domains; for example, polyketide synthases and type I fatty acid synthases.[76,77] In other cases, individually-expressed proteins self-assemble in the cytosol, for example the lumazine synthase/riboflavin synthase protein complex.[78,79] The structure and function of a number of polyketide synthase and fatty acid synthase systems have been studied in significant detail.[80–82] By contrast, the higher-order structure of hydrocarbon biosynthetic machinery has been far less characterized.

In the present study, the OleABCD enzymes from *X. campestris* were expressed and purified individually, and shown to reconstitute long-chain olefin biosynthesis. Co-expression studies demonstrated that OleBCD forms active, stable, assemblies. While hundreds of bacteria are thought to produce long-chain, head-to-head hydrocarbons, *X. campestris* has been used as a model system because the enzymes are relatively stable and have been shown to possess broad substrate, specificity making them suitable for bioengineering (4). OleA purifies as a stable, soluble dimer, and the structure and mechanism of this component has been the subject of several studies.[36–38] OleA catalyzes the Claisen condensation of fatty acyl-CoA substrates, producing a β-keto acid that feeds into the OleBCD enzymes.[36]

There are several reasons that multi-enzyme OleBCD assemblies might be functionally superior to individual OleB, OleC, and OleD proteins operating independently within a bacterial cell. Firstly, a multi-enzyme assembly of Ole proteins could protect the cell by sequestering the reactive Ole pathway β-lactone intermediate and preventing non-specific reaction with cytosolic proteins. The β-lactone moiety is common in microbial natural products, and can function as an antibiotic by reacting with active site residues in essential esterase and protease enzymes.[3,16] The β-lactone tetrahydrolipstatin from *Streptomyces toxytricini* is very similar in structure to the β-lactone produced by OleC, and its biosynthesis is encoded by an *ole*-like gene cluster that lacks the *oleB* decarboxylase gene.[3,58] Tetrahydrolipstatin is a potent human pancreatic lipase inhibitor, and is the only FDA approved, over-the-counter, anti-obesity drug. Since the β-lactone of olefin biosynthesis is only an intermediate, the sequestration of a potentially harmful compound by an enzyme complex likely affords sufficient protection to the bacterium. It is also relevant that *oleB* and *oleC* are a gene fusion in all actinobacteria, which comprise approximately 30% of known *ole*-gene-containing organisms. It is unknown if assemblies of OleBCD form in these bacteria, as the enzyme fusion should allow the direct transfer of the β-lactone from the OleC to OleB domains for decarboxylation. Fusion creates

a 1:1 stoichiometry between OleB and OleC, but our protein quantitation for the *X. campestris* multi-enzyme assemblies suggests a 4:1 relationship, respectively, suggesting there may be different configurations between species.

An additional benefit of an OleBCD complex would be to promote efficient trafficking of the highly hydrophobic intermediates that comprise the Ole pathway from one active site to another. It is common for the multi-step biosynthesis or degradation of hydrophobic compounds to be accomplished by either membrane-bound enzymes or enzymes complexes.[76,77] However, sequence analysis of Ole proteins did not predict any transmembrane domains or lipid anchor sequences. Moreover, the enzymes have been observed in the soluble protein fraction of the native host *X. campestris* and in the recombinant host *E. coli*. An example of a well-characterized bacterial complex producing and transferring hydrophobic intermediates is the fatty acid synthesis (FAS I) machinery from *Mycobacterium tuberculosis*.[83] Cryo-EM studies of FAS I from *M. tuberculosis* revealed a 2.0 MDa homohexameric enzyme complex with a diameter of 25 nm and the same hexagonal symmetry observed in our annular assemblies by TEM of the OleBCD assemblies. However, the FAS I complex uses an acyl carrier protein within a central cavity to ferry enzyme-tethered intermediates between active sites for iterative fatty acid elongation reactions. While a central cavity or porous, channeled structure is possible based on the size and molecular mass of the OleBCD assemblies, there is no evidence for any enzyme-tethered intermediate shuttle in the Ole biosynthetic pathway at this time.

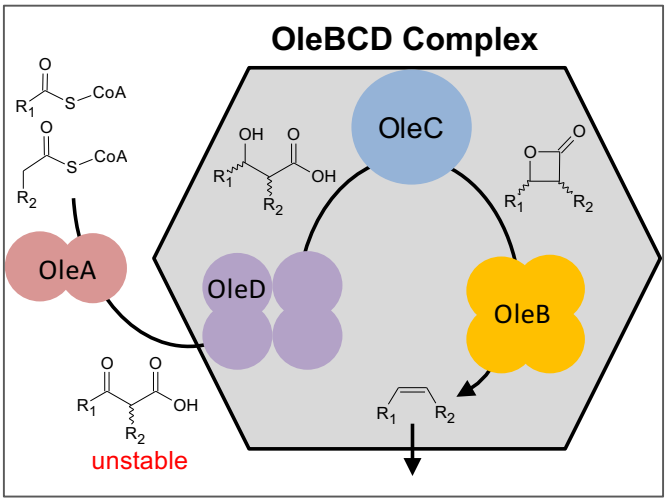


Figure 4-6. *Model for the biosynthesis of long-chain, head-to-head olefins in X. campestris showing OleA acts predominantly as a soluble cytosolic dimer, condensing acyl-CoAs, docking, and providing its reaction product to the OleBCD complex, that ultimately releases a long-chain olefin.*

In summary, we propose the following model for long-chain olefin biosynthesis in *X. campestris* (Figure 4-6). OleA receives fatty acyl-CoA products from the fatty acid biosynthetic machinery, acting as a shuttle between that complex and an OleBCD complex. It is likely that OleA has some binding affinity for an OleBCD complex, as suggested by assemblies of OleBCD showing weaker affinity for His_{6x}-OleA on the Ni^{2+} column. However, this interaction is not sufficiently strong to survive passage through several chromatographic steps. The current data support the idea that the consecutively-catalyzed OleD-OleC-OleB reactions occur sequestered within multi-enzyme structures.

The product olefin is ultimately found in the membrane fraction, and the mechanism of transfer from the soluble OleBCD assemblies to the membrane remains to be elucidated. Further studies by cryo-EM will help reveal more details of the OleBCD assembly architecture, and the relationship between the component enzymes. Long-chain hydrocarbon production has been confirmed experimentally in at least 20 different bacteria (2,4,6-9), and sequence analysis

reveals *oleABCD* gene clusters in hundreds of other bacteria, suggesting multi-enzyme OleBCD structures may be wide-spread amongst microorganisms.

Materials and Methods

Chemicals.

All reagents were purchased from Sigma-Aldrich, St. Louis, MO with the following exceptions. NADPH was purchased from EMD Millipore, Billerica, MA. Glycerol was purchased from IBI Scientific, Peosta, IA. Synthetic *cis*- and *trans*-olefins were prepared as described.[58]

Cloning, expression, and purification of Ole proteins.

All protein sequences for *ole* genes used in this study were from *X. campestris pv. campestris* ATCC 33913 (OleA: NP_635607.1, OleB: NP_635611.1, OleC: NP_635613.2, OleD: NP_635614.1). *E. coli* codon-optimized genes were ordered from DNA2.0, Inc. Polymerase, restriction enzymes, ligase, and super-competent cells were purchased from New England Biolabs and used according to manufacturer instructions. The four vector backbones, pET28b$^+$, pET30b$^+$, pCOLA, and pCDF, were obtained from Novogen. A complete list of constructs, gene-insertion sites, and antibiotic selectable markers from this study can be found in Supplemental Table S1.

The following general expression and purification protocols were developed for all Ole enzyme configurations. Plasmids were transformed into *E. coli* BL21 (DE3) cells (Invitrogen) and plated on Luria-Bertani (LB) agar with appropriate antibiotic selection (50 µg/µl). Single colonies were selected and grown at 37°C in 5 mL LB overnight as starter cultures. LB media (1.0 L) containing antibiotic was inoculated with a 5 mL starter cultures and grown at 37°C until an OD_{600} of 0.50 was reached. Protein expression was then induced with 100 µM isopropyl β-D-1-thiogalactopyranoside (IPTG), and growth was shifted to 16°C overnight. Cells were harvested by centrifugation and resuspended in 5 mL of buffer containing 200 mM NaCl, 20 mM HNa_2PO_4 pH 7.4, and 10% glycerol (buffer) per liter of LB media. Sonication or French pressure cell lysis was used to disrupt the cells, followed by centrifugation at

35,000 x *g*. The supernatant liquid was filtered through 0.45 µm and 0.22 µm filters (Millipore), and loaded onto a GE HisTrap HP 5 mL column (His column). Gradient elution in buffer with 500 mM imidazole (pH 7.4) was achieved using an ÄKTA fast protein liquid chromatography instrument (General Electric Company). Ole proteins typically eluted around 200 mM imidazole. Co-expressed protein mixtures containing an DYKDDDDK (FLAG) tag were then manually loaded onto a column containing 2.5 mL of Anti-FLAG M2 resin (Sigma-Aldrich) (anti-FLAG column). The anti-FLAG column was washed with 8 column volumes of buffer before eluting with 10 ml of 100 µg/mL FLAG peptide. The protein in the solution was concentrated using 30,000 MWCO spin filters (Amicon) and frozen at -80°C for later use.

The following protein-specific modifications were made to the above protocol. Whenever buffer additives were required as described below they were included for all purification steps. **OleBCD Assemblies**: Vector combinations, shown in Figure 2A, were transformed into *E. coli* under both kanamycin and streptomycin selection. No detergents were needed during purification of Ole assemblies, but 10 v/v% glycerol was found to increase protein yield. For electron microscopy experiments, glycerol was omitted from the purification buffer. **OleA**: After IPTG induction, cells containing pET28b$^+$ *oleA* with an N-terminal His_{6x} tag were grown at 37°C for 4 hours before purification in buffer. **OleB**: Buffer required 0.05 w/v% Triton X-100 (critical micelle concentration (cmc) ~0.02 w/v%) for purification of pET28b$^+$ expressed OleB with a His_{6x} N-terminal tag. Triton X-100 interfered with A_{280} measurements, necessitating that protein elution be followed by Bradford assay and protein gels.[84] Protein could not be concentrated without precipitation. **OleC**: An N-terminal His_{6x} tag on OleC does not purify, necessitating the C-terminal His_{6x} tag of pET30b$^+$ for expression and purification. **OleD**: Buffer containing 0.02 % w/v Tween-20 (cmc ~0.007 % w/v) was required to purify OleD with an N-terminal His_{6x} tag expressed from a pET28b$^+$ vector. Protein could not be concentrated without precipitation.

Olefin Activity Assays. Ole enzyme activity assays were conducted in glass vials at room temperature with 500 µL of buffer (200 mM NaCl, 20 mM $NaPO_4$, pH 7.4). Reactions were conducted in duplicate and contained 10 µg of OleA and 30 µg of co-purified OleBCD, as measured by Bradford assay. Assays involving separately purified OleA, OleB, OleC, and OleD contained: 10 µg of OleA, 12.1 µg of OleB, 5.1 µg of OleC and 12.8 µg of OleD to mimic the putative 4:1:4 B:C:D stoichiometry of 30 µg of co-purified OleBCD. ATP, NADPH, and $MgCl_2$ were added to final concentrations of 1 mM, 500 µM, and 1 mM, respectively in all assays. The slight difference in buffer additives to Ole proteins was ignored as they were found not to significantly affect reactions. Reactions were initiated by the addition of myristoyl coenzyme A (myristoyl-CoA) to a concentration of 50 µM and quenched with 500 µL of ethyl acetate at appropriate time points. Five microliters of organic phase was removed and analyzed by GC, using both a flame ionization detector (FID) and a mass spectrometry (MS) detector, as previously described.[58] Program conditions were as follows: start temp, 100°C; ramp rate, 10°C/min; final temp, 320°C; hold time, 5 min; total time, 27 min. A synthetic Olefin standard, *cis*-13-heptacosene, was synthesized as described previously.[58] The retention time of *cis*-13-heptacosene was 18.26 min.

Mass spectrometry protein identification.

Ole proteins were separated using 12.5% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Bands were excised, trypsin-digested, and submitted to the University of Minnesota Center for Mass Spectrometry and Proteomics for peptide identification using liquid chromatography/mass spectrometry.[85] Peptides were initially separated by a Paradigm Platinum Peptide Nanotrap (Michrom Bioresources, Inc.) pre-column followed by an analytical capillary column (100 µm×12 cm) packed with C18 resin (5 µm, 200 Å MagicC18AG, Michrom Bioresources, Inc.). Mass spectrometry was performed on an LTQ (Thermo Electron Corp., San Jose, CA).

Immunoblots for OleC in native *X. campestris* lysate.

Polyclonal antibodies against an *X. campestris* OleC peptide (AIDDAAIPEWSGVR) were raised by GenScript in rabbit. A secondary horse radish peroxidase-conjugated donkey anti-rabbit IgG antibody was purchased from Jackson Immunoresearch Laboratories. *X. campestris* (ATCC 33913) for immunoblots was grown in 250 mL LB media at 30°C before harvesting and sonication generated the *X. campestris* lysate. OleC and OleBCD (Figure 4-2A, tag arrangement #4) were purified as described above. Native gels were made as described by and contained 1.9% polyacrylamide and 0.75% agarose in 40 mM tris-acetate buffer at pH 8.4.[86] Protein was mixed with loading buffer (15% sucrose in tris-acetate) and run at 100V. Gels were transferred to a nitrocellulose membrane and washed thoroughly with 200mM NaCl, 20 mM $NaPO_4$ (pH 7.4), and 0.1 % Tween 20 as a blocking agent. Membranes were incubated with the primary antibody overnight, washed, and incubated with the secondary antibody for 1 hr. ECL Plus Western Blotting Substrate (Pierce) and Classic Blue Autoradiology Film BX (MidSci) were used to visualize the membrane.

Gel-filtration chromatography.

A HiLoad 16/600 Superdex 200 (GE, normal column) and a HiPrep 16/60 Sephacryl S-400 HR (GE, ultra-high MW column) were used to separate the proteins and determine molecular mass. The normal small size column can separate globular proteins up to 600 kDa, while the ultra-high MW column can separate globular proteins up to 8.0 MDa. A Gel Filtration HMW Calibration Kit was purchased from GE to calibrate the gel filtration columns (up to 669 kDa), with larger molecular weights being extrapolated. Typically, 2.0 mg of protein were loaded onto columns and run according to the manufacturer's specifications.

Electron microscopy of OleBCD assemblies.

Following His-column, anti-FLAG, and gel-filtration purification, OleBCD assemblies were characterized using transmission electron microscopy (TEM).

Grids (400 mesh Cu carbon grids) were glow discharged at 15 mA and 0.39 mBar for 1 minute using a Pelco easiGlow glow discharge cleaning system (Ted Pella, Inc. Redding, CA). Samples containing 0.08 mg/mL OleBCD were drop cast onto these grids and blotted with filter paper to remove excess sample. Grids were washed twice with deionized water and blotted dry. Uranyl formate (0.05%) stain was added to each grid and blotted a final time. Negative stain micrographs were collected using a Field emission gun (FEG) Tecnai G2 F30 TEM (FEI, Hillsboro, OR). Images were recorded at room temperature on a Gatan 4k x 4k CCD camera (Gatan, Inc. Pleasanton, CA). ImageJ was used to analyze the sizes of greater than 150 particles in four TEM micrographs. Size was recorded as the greatest diameter of the particle.

SDS-PAGE standard curve.

Quantification of staining intensities of the Ole proteins on SDS-PAGE was carried out using individually purified Ole proteins as described above. Concentrations of OleB, OleC, and OleD were determined by the Bradford Assay and supported by UV absorbance at 280 nm (A_{280}). OleB required Triton X-100 to prevent protein precipitation, and so could not be verified by A_{280}. A 4.0 μM master mix of the three proteins was created and loaded in different volumes into wells of a 12.5% SDS-PAGE gel. Gels were stained with Simple Blue Safe Stain (Life Technologies), and images were analyzed using ImageJ.

Acknowledgements

We thank Isaac Hamilton for his help with gas chromatography assays and standard curves, as well as Iwaki Shigehiro for his help in the purification of Ole proteins. We also thank the University of Minnesota Center for Mass Spectrometry and Proteomics for their help in peptide fragment identification. JKC was supported in part by the NIH Training for Future Biotechnology Development (T32GM008347). MRJ was supported in part by the NIH Chemistry-Biology Interface Training Grant (T32GM008700). We acknowledge the support of the Biotechnology Institute from the University of Minnesota, and the MnDRIVE Initiative from the Office of the Vice President for Research of the University of Minnesota (to CMW and LPW). The electron micrographs were collected using a Tecnai TF30 TEM maintained by the Characterization Facility, College of Science and Engineering, University of Minnesota.

CHAPTER 5 - Conclusions

It always seems impossible until it's done. - Nelson Mandela

The work contained in this dissertation has significantly contributed to both our understanding of long-chain hydrocarbon biosynthesis as well as the bacterial biosynthesis of β-lactone natural products by identifying β-lactones as an unexpected common biosynthetic intermediate (Figure 5-1).

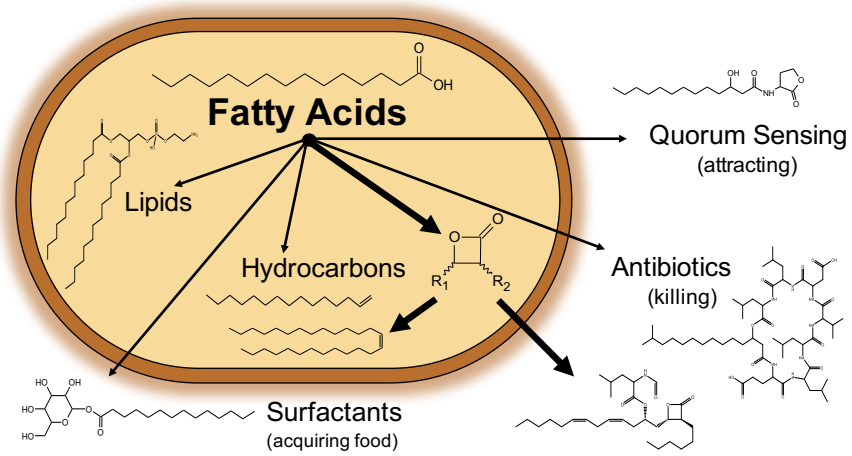


Figure 5-1. *The β-lactone moiety is an unexpected common intermediate of both long-chain olefin biosynthesis and medicinal β-lactone natural products.*

Concerning the Ole pathway, the final two steps of olefin biosynthesis have now been concretely defined. OleC has been demonstrated to be a β-lactone synthetase, the first described in literature. This discovery also corrected an error in the literature that mis-assigned the function of OleC based on GC assays that caused the thermal breakdown of the true β-lactone product of OleC to an olefin. Additionally, the final enzyme of the pathway, OleB, has now been purified for the first time and its function identified as a β-lactone decarboxylase. OleB is a member of the α/β-hydrolase superfamily that is most similar to the well-studied haloalkane dehalogenases (HLDs). The decarboxylation of a β-lactone to an olefin is an unprecedented enzyme reaction in nature and,

therefore, represents a novel cellular route to hydrocarbons. The spatial coordination of the Ole proteins into an enzyme complex was also identified in both a native organism and recombinant work. This complex is believed to facilitate the transfer of the hydrophobic intermediates of the Ole pathway from one enzyme active site to the next as well as promote the rapid decarboxylation of the reactive β-lactone intermediate. It is hoped that these discoveries on the Ole pathway will aid future researchers in both academic and industrial settings to efficiently and sustainably provide the materials used in modern society.

Concerning the β-lactone, the first enzyme to perform the critical β-lactone ring-closure reaction has now been identified as OleC. Bacterial β-lactone natural products are currently used in clinical settings and others are being explored for their potent anti-cancer and antibiotic properties. The β-lactone synthetase OleC, was found to have sequence homology to proteins encoded by known β-lactone gene clusters, suggesting a common mechanism of formation. However, some β-lactone gene clusters do not appear to contain an OleC homolog, suggesting the existence of at least one additional mechanism of formation. It is hoped that these findings will facilitate the production of β-lactone natural products as drugs and expedite the discovery of novel β-lactone compounds for clinical trials.

Future Directions

OleBCD Complex

The identification of an OleBCD complex in hydrocarbon biosynthesis offers a glimpse into the cellular processing of hydrophobic and toxic intermediates. Recent advances in cryo-EM has made this technique amenable for structural characterization of enzyme complexes with sub 5 angstroms resolution; this method could be applied to an OleBCD complex to attain a defined structure. The OleBCD complex has only been demonstrated in *Xanthomonas campestris*, and it would be beneficial to investigate it in other organisms. Actinobacteria contain a natural fusion of the OleB and OleC proteins, presumably to aid in the decarboxylation of the lactone. It is unclear

what form an Ole complex would take in one of these organisms. It is also possible that the biosynthetic enzymes of lipstatin, which bears striking resemblance to the Ole pathway, could form an enzyme complex. To our knowledge, it remains unknown how *Streptomyces* species protect themselves against the toxic β-lactones. A complex would allow for the specific localization of the β-lactone and possibly help contain the product until it can be appropriately processed.

OleC

There is still work to be done on OleC, the β-lactone synthetase, in relation to its mechanism, substrate specificity, and placement within the ANL-superfamily. OleC uses ATP and likely forms an AMP intermediate on either the carboxylic acid or the hydroxyl group before undergoing an intramolecular ring closure reaction to release AMP. Substrate mimic and O^{18} labeling studies should clarify the mechanism of OleC. Additionally, the substrate specificity of OleC may prove important for its implementation as a biocatalyst in the production of medically-relevant β-lactones. A crystal structure of OleC would be very helpful in these endeavors to map the substrate channels and key binding residues.

OleB

As previously discussed, OleB is typically mis-annotated in genomes as a haloalkane dehalogenase because of its close sequence and mechanistic relationship to HLDs. Future bioinformatics work backed up by proper biochemical experiments are needed to help define the similarities and differences between these two classes of enzymes. It is tantalizing to speculate that the decarboxylation of a β-lactone may be a resistance or self-protection mechanism for β-lactone antibiotics in nature. Once the amino acid sequence space of β-lactone decarboxylase has been defined, bioinformatics can likely identify non-Ole pathway β-lactone decarboxylases that may play a role in detoxification.

Identification of novel β-lactone

A key component to discovering novel β-lactone natural products with medicinal value is defining sequences motifs that separate β-lactone synthetases from other ANL-superfamily members. Again, a combination of bioinformatics and biochemical assays should help define the β-lactone synthetase sequence space. Once identified, genomic mining can be conducted to identify putative β-lactone natural product gene clusters. The characterization of hypothetical β-lactone biosynthesis gene clusters and the compounds they produce will likely take some time but could yield valuable medical compounds.

"And further, my son, be admonished by these. Of making many books there is no end, and much study is wearisome to the flesh. Let us hear the conclusion of the whole matter: fear God and keep His commandments, for this is man's all. For God will bring every work into judgment, including every secret thing, whether good or evil." – King Solomon (Ecclesiastes 12:12-14)

References

1. Jetter R, Kunst L. Plant surface lipid biosynthetic pathways and their utility for metabolic engineering of waxes and hydrocarbon biofuels. *Plant J*. 2008;54(4):670-683. doi:10.1111/j.1365-313X.2008.03467.x.
2. Chung H, Carroll SB. Wax, sex and the origin of species: Dual roles of insect cuticular hydrocarbons in adaptation and mating. *BioEssays*. 2015;37(7):822-830. doi:10.1002/bies.201500014.
3. Bai T, Zhang D, Lin S, *et al.* Operon for biosynthesis of lipstatin, the beta-lactone inhibitor of human pancreatic lipase. *Appl Environ Microbiol*. 2014;80(24):7473-7483. doi:10.1128/AEM.01765-14.
4. National Research Council Staff. *Oil in the Sea: Inputs, Fates, and Effects.* Washington: National Academies Press; 1984. http://public.eblib.com/choice/publicfullrecord.aspx?p=3377105. Accessed May 29, 2017.
5. US Energy Information Administration. *Petroleum Supply Annual*. Washington, DC 20585: Department of Energy; 2016.
6. British Petroleum. *BP Statistical Review of World Energy June 2016*. 65th Editi. Pureprint Group Limited, UK; 2016.
7. National Park Service. Missippi River Facts. https://www.nps.gov/miss/riverfacts.htm. Published 2017. Accessed May 30, 2017.
8. Schleussner CF, Rogelj J, Schaeffer M, *et al.* Science and policy characteristics of the Paris Agreement temperature goal. *Nat Clim Chang*. 2016;6(9):827-835. doi:10.1038/nclimate3096.
9. Amundson R, Berhe AA, Hopmans JW, Olson C, Sztein AE, Sparks DL. Soil and human security in the 21st century. *Science*. 2015;348(6235):1261071-1261071. doi:10.1126/science.1261071.
10. Clark PU, Shakun JD, Marcott SA, *et al.* Consequences of twenty-first-century policy for multi-millennial climate and sea-level change. *Nat Clim*

Chang. 2016;6(4):360-369. doi:10.1038/nclimate2923.

11. Kluts IN, Brinkman MLJ, de Jong SA, Junginger HM. Biomass Resources: Agriculture. *Adv Biochem Eng Biotechnol*. April 2017. doi:10.1007/10_2016_66.
12. Leavell MD, McPhee DJ, Paddon CJ. Developing fermentative terpenoid production for commercial usage. *Curr Opin Biotechnol*. 2016;37:114-119. doi:10.1016/j.copbio.2015.10.007.
13. Yim H, Haselbeck R, Niu W, *et al.* Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat Chem Biol*. 2011;7(7):445-452. doi:10.1038/nchembio.580.
14. Zhou YJ, Buijs NA, Zhu Z, Qin J, Siewers V, Nielsen J. Production of fatty acid-derived oleochemicals and biofuels by synthetic yeast cell factories. *Nat Commun*. 2016;7:11709. doi:10.1038/ncomms11709.
15. Dorobantu LS, Yeung AKC, Foght JM, Gray MR. Stabilization of oil-water emulsions by hydrophobic bacteria. *Appl Environ Microbiol*. 2004;70(10):6333-6336. doi:10.1128/AEM.70.10.6333-6336.2004.
16. Wyatt MA, Ahilan Y, Argyropoulos P, Boddy CN, Magarvey NA, Harrison PH. Biosynthesis of ebelactone A: isotopic tracer, advanced precursor and genetic studies reveal a thioesterase-independent cyclization to give a polyketide β-lactone. *J Antibiot (Tokyo)*. 2013;66(7):421-430. doi:10.1038/ja.2013.48.
17. Compton CL, Schmitz KR, Sauer RT, Sello JK. Antibacterial activity of and resistance to small molecule inhibitors of the ClpP peptidase. *ACS Chem Biol*. 2013;8(12):2669-2677. doi:10.1021/cb400577b.

18. Eustaquio AS, McGlinchey RP, Liu Y, *et al.* Biosynthesis of the salinosporamide A polyketide synthase substrate chloroethylmalonyl-coenzyme A from S-adenosyl-L-methionine. *Proc Natl Acad Sci*. 2009;106(30):12295-12300. doi:10.1073/pnas.0901237106.
19. Miller CP, Ban K, Dujka ME, *et al.* NPI-0052, a novel proteasome inhibitor, induces caspase-8 and ROS-dependent apoptosis alone and in combination with HDAC inhibitors in leukemia cells. *Blood*. 2007;110(1):267-277. doi:10.1182/blood-2006-03-013128.
20. Sello JK, Compton CL. Antibacterial beta-lactones, and methods of identification, manufacture and use. 2015;(WO2015031300 A1). http://www.google.com/patents/WO2015031300A1. Accessed May 29, 2017.
21. Smith JW, Axelrod F, Kridel SJ, Romo D, Purohit V, Ma G. Method for the asymmetric synthesis of beta-lactone compounds. 2012;(US8124794 B2). http://www.google.com/patents/US8124794. Accessed May 29, 2017.
22. Wolf F, Bauer JS, Bendel TM, *et al.* Biosynthesis of the β-Lactone Proteasome Inhibitors Belactosin and Cystargolide. *Angew Chemie Int Ed*. 2017;56(23):6665-6668. doi:10.1002/anie.201612076.
23. Rachid S, Huo L, Herrmann J, *et al.* Mining the Cinnabaramide Biosynthetic Pathway to Generate Novel Proteasome Inhibitors. *ChemBioChem*. 2011;12(6):922-931. doi:10.1002/cbic.201100024.
24. Beller HR, Lee TS, Katz L. Natural products as biofuels and bio-based chemicals: fatty acids and isoprenoids. *Nat Prod Rep*. 2015;32(10):1508-1526. doi:10.1039/C5NP00068H.
25. Sukovich DJ, Seffernick JL, Richman JE, Hunt KA, Gralnick JA, Wackett LP. Structure, function, and insights into the biosynthesis of a head-to-head hydrocarbon in *Shewanella oneidensis* strain MR-1. *Appl Environ Microbiol*. 2010;76(12):3842-3849. doi:10.1128/AEM.00433-10.
26. Wang W, Lu X. Microbial Synthesis of Alka(e)nes. *Front Bioeng Biotechnol*. 2013;1. doi:10.3389/fbioe.2013.00010.

27. Schirmer A, Rude MA, Li X, Popova E, Del Cardayre SB. Microbial Biosynthesis of Alkanes. *Science (80-)*. 2010;329(5991):559-562. doi:10.1126/science.1187936.
28. Rude MA, Baron TS, Brubaker S, Alibhai M, Del Cardayre SB, Schirmer A. Terminal olefin (1-alkene) biosynthesis by a novel P450 fatty acid decarboxylase from *Jeotgalicoccus* species. *Appl Environ Microbiol*. 2011;77(5):1718-1727. doi:10.1128/AEM.02580-10.
29. Grant JL, Hsieh CH, Makris TM. Decarboxylation of fatty acids to terminal alkenes by cytochrome P450 compound I. *J Am Chem Soc*. 2015;137(15):4940-4943. doi:10.1021/jacs.5b01965.
30. Liu Y, Wang C, Yan J, et al. Hydrogen peroxide-independent production of α-alkenes by $OleT_{JE}$ P450 fatty acid decarboxylase. *Biotechnol Biofuels*. 2014;7(1):28. doi:10.1186/1754-6834-7-28.
31. Albro PW, Dittmer JC. The biochemistry of long-chain, nonisoprenoid hydrocarbons. I. Characterization of the hydrocarbons of *Sarcina lutea* and the isolation of possible intermediates of biosynthesis. *Biochemistry*. 1969;8(1):394-404.
32. Friedman L, da Costa B. Hydrocarbon-producing genes and methods of their use. 2007;(WO2008147781 A2). http://www.google.com/patents/WO2008147781A2. Accessed May 29, 2017.
33. Frias JA, Riehman JE, Wackett LP. C 29 olefinic hydrocarbons biosynthesized by *Arthrobacter* species. *Appl Environ Microbiol*. 2009;75(6):1774-1777. doi:10.1128/AEM.02547-08.

34. Sukovich DJ, Seffernick JL, Richman JE, Gralnick JA, Wackett LP. Widespread head-to-head hydrocarbon biosynthesis in bacteria and role of OleA. *Appl Environ Microbiol*. 2010;76(12):3850-3862. doi:10.1128/AEM.00436-10.
35. Beller HR, Goh EB, Keasling JD. Genes involved in long-chain alkene biosynthesis in *Micrococcus luteus*. *Appl Environ Microbiol*. 2010;76(4):1212-1223. doi:10.1128/AEM.02312-09.
36. Frias JA, Richman JE, Erickson JS, Wackett LP. Purification and characterization of OleA from *Xanthomonas campestris* and demonstration of a non-decarboxylative Claisen condensation reaction. *J Biol Chem*. 2011;286(13):10930-10938. doi:10.1074/jbc.M110.216127.
37. Goblirsch BR, Frias JA, Wackett LP, Wilmot CM. Crystal structures of *Xanthomonas campestris* OleA reveal features that promote head-to-head condensation of two long-chain fatty acids. *Biochemistry*. 2012;51(20):4138-4146. doi:10.1021/bi300386m.
38. Goblirsch BR, Jensen MR, Mohamed FA, Wackett LP, Wilmot CM. Substrate trapping in crystals of the thiolase OleA identifies three channels that enable long chain olefin biosynthesis. *J Biol Chem*. 2016;291(52):26698-26706. doi:10.1074/jbc.M116.760892.
39. Bonnett SA, Papireddy K, Higgins S, Del Cardayre S, Reynolds KA. Functional characterization of an NADPH dependent 2-alkyl-3-ketoalkanoic acid reductase involved in olefin biosynthesis in *Stenotrophomonas maltophilia*. *Biochemistry*. 2011;50(44):9633-9640. doi:10.1021/bi201096w.
40. Kancharla P, Bonnett SA, Reynolds KA. *Stenotrophomonas maltophilia* OleC-Catalyzed ATP-Dependent Formation of Long-Chain Z-Olefins from 2-Alkyl-3-hydroxyalkanoic Acids. *ChemBioChem*. 2016;17(15):1426-1429. doi:10.1002/cbic.201600063.
41. Gulick AM. Conformational dynamics in the acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly

luciferase. *ACS Chem Biol*. 2009;4(10):811-827. doi:10.1021/cb900156h.

42. Frias JA. Microbial synthesis of fuel hydrocarbons: enzymes and metabolic pathways. *University of Minnesota Dissertation*, 2011.
43. Feling RH, Buchanan GO, Mincer TJ, Kauffman CA, Jensen PR, Fenical W. Salinosporamide A: A highly cytotoxic proteasome inhibitor from a novel microbial source, a marine bacterium of the new genus *Salinospora*. *Angew Chemie - Int Ed*. 2003;42(3):355-357. doi:10.1002/anie.200390115.
44. Lee MJ, Gwak HS, Park BD, Lee ST. Synthesis of mycolic acid biosurfactants and their physical and surface-active properties. *J Am Oil Chem Soc*. 2005;82(3):181-188. doi:10.1007/s11746-005-5170-8.
45. Masamune S, Hayase Y, Chan WK, Sobczak RL. Tylonolide hemiacetal, the aglycone of tylosin and its partial synthesis. *J Am Chem Soc*. 1976;98(24):7874-7875. doi:10.1021/ja00440a096.
46. Wang Y, Tennyson RL, Romo D. β-lactones: Intermediates for natural product total synthesis and new transformations. *Heterocycles*. 2004;64(1):605-658. doi:10.3987/REV-04-SR(P)3.
47. Noyce DS, Elden EH. The Stereochemistry of the Decarboxylation of beta-Lactones to Form Olefins. *J Org Chem*. 1966;31(December):4043-4047. doi:10.1021/jo01350a037.
48. Mulzer J, Zippel M, Brüntrup G. Thermische Decarboxylierung von β-Lactonen: Sterische Mesomeriehinderung als Hinweis auf eine zwitterionische Zwischenstufe. *Angew Chemie*. 1980;92(6):469-470. doi:10.1002/ange.19800920612.
49. Conti E, Franks NP, Brick P. Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure*. 1996;4(3):287-298. doi:10.1016/S0969-2126(96)00033-0.
50. Dick LR, Cruikshank AA, Melandri FD, *et al.* Mechanistic Studies on the Inactivation of the Proteasome by Lactacystin. *J Biol Chem*. 1996;272(1):182-188. doi:10.1074/jbc.271.13.7273.
51. Zhao C, Coughlin JM, Ju J, et al. Oxazolomycin biosynthesis in

Streptomyces albus JA3453 featuring an “acyltransferase-less” type I polyketide synthase that incorporates two distinct extender units. *J Biol Chem*. 2010;285(26):20097-20108. doi:10.1074/jbc.M109.090092.

52. Hemmerling F, Hahn F. Biosynthesis of oxygen and nitrogen-containing heterocycles in polyketides. *Beilstein J Org Chem*. 2016;12:1512-1550. doi:10.3762/bjoc.12.148.
53. Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P, Chatonnet A. ESTHER, the database of the α/β-hydrolase fold superfamily of proteins: Tools to explore diversity of functions. *Nucleic Acids Res*. 2013;41(D1):D423--D429. doi:10.1093/nar/gks1154.
54. Rauwerdink A, Kazlauskas RJ. How the Same Core Catalytic Machinery Catalyzes 17 Different Reactions: The Serine-Histidine-Aspartate Catalytic Triad of α/β-Hydrolase Fold Enzymes. *ACS Catal*. 2015;5(10):6153-6176. doi:10.1021/acscatal.5b01539.
55. Nardini M, Dijkstra BW. alpha/beta hydrolase fold enzymes: The family keeps growing. *Curr Opin Struct Biol*. 1999;9(6):732-737. doi:10.1016/S0959-440X(99)00037-8.
56. Nagata Y, Ohtsubo Y, Tsuda M. Properties and biotechnological applications of natural and engineered haloalkane dehalogenases. *Appl Microbiol Biotechnol*. 2015;99(23):9865-9881. doi:10.1007/s00253-015-6954-x.
57. Christenson JK, Jensen MR, Goblirsch BR, *et al.* Active multienzyme assemblies for long-chain olefinic hydrocarbon biosynthesis. *J Bacteriol*. 2017;199(9):e00890--16. doi:10.1128/JB.00890-16.
58. Christenson JK, Richman JE, Jensen MR, Neufeld JY, Wilmot CM, Wackett LP. β-Lactone Synthetase Found in the Olefin Biosynthesis Pathway. *Biochemistry*. 2017;56(2):348-351. doi:10.1021/acs.biochem.6b01199.
59. Chovancova E, Kosinski J, Bujnicki JM, Damborsky J. Phylogenetic analysis of haloalkane dehalogenases. *Proteins Struct Funct Genet*.

2007;67(2):305-316. doi:10.1002/prot.21313.

60. Mulzer J, Brüntrup G, Hartz G, Kühl U, Blaschek U, Böhrer G. Additionen von Carbonsäure-Dianionen an α,β-ungesättigte Carbonylverbindungen - Steuerung der 1,2-/1,4-Regioselektivität durch sterische Substituenteneffekte. *Chem Ber*. 1981;114(11):3701-3724. doi:10.1002/cber.19811141123.
61. Crossland RK, Servis KL. Facile synthesis of methanesulfonate esters. *J Org Chem*. 1970;35(3):3195-3196. doi:10.1021/jo00834a087.
62. Lindlar H. Ein neuer Katalysator für selektive Hydrierungen. *Helv Chim Acta*. 1952;35(2):446-450. doi:10.1002/hlca.19520350205.
63. Buck M, Chong JM. Alkylation of 1-alkynes in THF. *Tetrahedron Lett*. 2001;42(34):5825-5827. doi:10.1016/S0040-4039(01)01131-5.
64. Thalmann A, Oertle K, Gerlach H. Ricinelaidic acid lactone. *Org Synth*. 1985;63:192. doi:10.15227/orgsyn.063.0192.
65. Wright ES. DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics*. 2015;16(1):322. doi:10.1186/s12859-015-0749-z.
66. Schliep KP. Phangorn: Phylogenetic analysis in R. *Bioinformatics*. 2011;27(4):592-593. doi:10.1093/bioinformatics/btq706.
67. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. 2015;10(6):845-858. doi:10.1038/nprot.2015.053.
68. Jesenská A, Monincová M, Koudeláková T, *et al.* Biochemical characterization of haloalkane dehalogenases DrbA and DmbC, representatives of a novel subfamily. *Appl Environ Microbiol*. 2009;75(15):5157-5160. doi:10.1128/AEM.00199-09.
69. Hesseler M, Bogdanović X, Hidalgo A, *et al.* Cloning, functional expression, biochemical characterization, and structural analysis of a haloalkane dehalogenase from *Plesiocystis pacifica* SIR-1. *Appl Microbiol Biotechnol*. 2011;91(4):1049-1060. doi:10.1007/s00253-011-3328-x.

70. Novak HR, Sayer C, Isupov MN, Gotz D, Spragg AM, Littlechild JA. Biochemical and structural characterisation of a haloalkane dehalogenase from a marine *Rhodobacteraceae*. *FEBS Lett*. 2014;588(9):1616-1622. doi:10.1016/j.febslet.2014.02.056.
71. Auldridge ME, Guo Y, Austin MB, *et al*. Emergent Decarboxylase Activity and Attenuation of α/β-Hydrolase Activity during the Evolution of Methylketone Biosynthesis in Tomato. *Plant Cell*. 2012;24(4):1596-1607. doi:10.1105/tpc.111.093997.
72. Gao A, Mei GY, Liu S, *et al.* High-resolution structures of AidH complexes provide insights into a novel catalytic mechanism for N-acyl homoserine lactonase. *Acta Crystallogr Sect D Biol Crystallogr*. 2013;69(1):82-91. doi:10.1107/S0907444912042369.
73. Koudelakova T, Chovancova E, Brezovsky J, *et al.* Substrate specificity of haloalkane dehalogenases. *Biochem J*. 2011;435(2):345-354. doi:10.1042/BJ20101405.
74. Channon HJ, Chibnall AC. The ether-soluble substances of cabbage leaf cytoplasm: The isolation of n-nonacosane and di-n-tetradecyl ketone. *Biochem J*. 1929;23(2):168-175.

75. Nichols DS, Nichols PD, McMeekin TA. A new n-C31:9 polyene hydrocarbon from Antarctic bacteria. *FEMS Microbiol Lett*. 1995;125(2-3):281-285. doi:10.1016/0378-1097(94)00509-P.
76. Blom T, Somerharju P, Ikonen E. Synthesis and biosynthetic trafficking of membrane lipids. *Cold Spring Harb Perspect Biol*. 2011;3(8):1-16. doi:10.1101/cshperspect.a004713.
77. Pawar S, Schulz H. The structure of the multienzyme complex of fatty acid oxidation from *Escherichia coli. J Biol Chem*. 1981;256(8):3894-3899.
78. Haase I, Gräwert T, Illarionov B, Bacher A, Fischer M. Recent Advances in Riboflavin Biosynthesis. *Methods Mol Biol*. 2014;1146:15-40. doi:10.1007/978-1-4939-0452-5_2.
79. Ladenstein R, Fischer M, Bacher A. The lumazine synthase/riboflavin synthase complex: Shapes and functions of a highly variable enzyme system. *FEBS J*. 2013;280(11):2537-2563. doi:10.1111/febs.12255.
80. Smith JL, Skiniotis G, Sherman DH. Architecture of the polyketide synthase module: Surprises from electron cryo-microscopy. *Curr Opin Struct Biol*. 2015;31:9-19. doi:10.1016/j.sbi.2015.02.014.
81. Robbins T, Liu Y-C, Cane DE, Khosla C. Structure and mechanism of assembly line polyketide synthases. *Curr Opin Struct Biol*. 2016;41:10-18. doi:10.1016/j.sbi.2016.05.009.
82. Enderle M, Mccarthy A, Paithankar KS, Grininger M. Crystallization and X-ray diffraction studies of a complete bacterial fatty-acid synthase type I. *Acta Crystallogr Sect Struct Biol Commun*. 2015;71(11):1401-1407. doi:10.1107/S2053230X15018336.
83. Boehringer D, Ban N, Leibundgut M. 7.5-Å cryo-EM structure of the mycobacterial fatty acid synthase. *J Mol Biol*. 2013;425(5):841-849. doi:10.1016/j.jmb.2012.12.021.
84. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.

Anal Biochem. 1976;72(1-2):248-254. doi:10.1016/0003-2697(76)90527-3.

85. Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal Chem*. 1996;68(5):850-858. doi:10.1021/ac950914h.
86. Nadano D, Aoki C, Yoshinaka T, Irie S, Sato TA. Electrophoretic characterization of ribosomal subunits and proteins in apoptosis: Specific downregulation of S11 in staurosporine-treated human breast carcinoma cells. *Biochemistry*. 2001;40(50):15184-15193. doi:10.1021/bi0108397.

Appendices

Specific Contributions

Many people contributed to both the published and unpublished works contained in this dissertation. Dr. Jack Richman generated the vast majority of synthetic compounds used in Chapters 2, 3, and 4. An Kim supported this synthesis work for the experiments contained in Chapter 3. Dr. Jack Richman also provided much of the NMR characterization and helped with data analyses. The structural biology work of this dissertation was supported by Matthew Jensen and his adviser, Dr. Carrie Wilmot. Specifically, Matthew Jensen generated the homology model of OleB used in Chapter 3 and the EM images used in Chapter 4 with the aid of Dr. Wei Zhang. Serina Robinson compiled and analyzed the bioinformatics data contained in Chapter 3. Dr. Brandon Goblirsch and Fatuma Mohamed did much of the groundwork for the OleBCD complex work of Chapter 4, including cloning and initial purifications. Jennifer Neufeld and Tiffany Engel provided significant practical support in lab for the work in Chapters 2 and 3, respectively.

Chapter 2 - OleC

Synthesis and characterization of 3-hydroxy-2-octyldodecanoic acid (3).

Synthesis of **3** was prepared from reaction of 1.0 g of decanoic acid with 2.2 equivalents of 1 M LDA in THF at -30° C followed by warming and solvent exchange to remove hexane as described by Mulzer (1). Addition of decanal (2.2 equivalents in THF) produced 95% pure **3** by ^{1}H NMR with the major contaminant identified as decanoic acid. The (+/-) *syn*-β-hydroxy acid diastereomers (2R,3S and 2S, 3R) were produced in 1:1.3 ratio with the (+/-) *anti*-β-hydroxy acid diastereomers (2R,3R and 2S,3S). Recrystalization yielded crystals of an equal molar ratio of all four possible diastereomeric isomers by ^{1}H NMR and GC/MS analyses, leaving the *anti*-diastereomer highly enriched in the supernatant. **Figure S15** shows the general synthesis and highlights the four diastereomeric products. All ^{1}H NMR (400 MHz, $CDCl_3$) peaks match literature values well (2) with the single β-hydrogens of the *syn*- and *anti*-β-hydroxy acids showing shifts of 3.86 and 3.72 ppm, respectively. The single α-hydrogens of the two diastereomeric pairs were not resolved, both falling near 2.48 ppm.

HPLC separation of *syn*- and *anti*-diastereomer pairs.

The synthesized mixture of *syn*- and *anti*-3-hydroxy-2-octyldodecanoic acids (**3**) was separated into the *syn*- and *anti* diastereomeric pairs (**3a** and **3b**) by High Pressure Liquid Chromatography (HPLC, Hewlett Packard) using a reverse phase C18 column (Agilent eclipse plus 4.6x250 mm) and flow rates of 1.0 mL/min. The LC column was prewashed with 100% acetonitrile (ACN) with 4 mM HCl and 40% methyl tert-butyl ether (MTBE) in ACN prior to injection of 100 µL samples of 2.0 mg/ml of (**3**) in the acidic ACN. The elution program at room temperature was as follows: 100% ACN hold for 2 min; linearly ramp MTBE to 40% in ACN over 8 min; hold 40% MTBE for 5 min; reverse to 100% ACN in 3 minutes. Total program was 18 min, with a detection wavelength of 220 nm. Fractions were collected manually from 10 runs to accumulate approximately 1.0 mg of each racemic diastereomer. The elution of *syn*- before *anti*-**3** in reverse phase HPLC is consistent with literature (2). GC/MS and ^{1}H NMR analyses

confirmed both **3a** and **3b** at 90% purity. ^{1}H NMR (400 MHz, $CDCl_3$) values for the single α- and β-hydrogens of the (+/-) *syn*-β-hydroxy acids were 2.50 and 3.85 ppm, respectively, while the same hydrogens in the (+/-) *anti*-β-hydroxy acid show shifts of 2.47 and 3.72 ppm, respectively (**Fig S5** and **S6**).

Synthesis and characterization of *cis*- and *trans*-4-nonyl-3-octyloxtan-2-one (4a, 4b).

Starting with decanoic acid, the method of Lee (3) produced octylketene dimer in good yield. Only one isomer (presumably *cis*) was observed by ^{1}H NMR. Prior to hydrogenation the ketene dimer was purified by silica gel chromatography. Hydrogenation with 10% Pd/C catalyst in EtOAc/EtOH, as described by Lee (3), produced (+/-) *cis*-4-nonyl-3-octyloxtan-2-one (**4a**) and low levels (~3%) of the *trans* isomer, as observed by ^{1}H NMR/$CDCl_3$. The methine protons of the *cis*-isomer at carbons 4 and 3 are sharp multiplets (d,d,d) at 4.53 and 3.59 ppm, respectively (**Fig S3**). The corresponding multiplets for the *trans* isomer were at 4.21 and 3.16 ppm (**Fig S4**). *trans*-4-nonyl-3-octyloxtan-2-one (**4b**) was isolated and purified from a complex mixture of products containing about equal amounts of the *cis* and *trans*-isomers of the oxtan-2-one. The complex mixture was prepared by the reaction of methane sulfonyl chloride (2 equivalents Et_3N) in dichloromethane with 3-hydroxy-2-octyldodecanoic acid (**3**) (~1:1 *syn*/*anti*) according to Crossland (4). Column chromatography (SiO_2 gel, hexane stepwise to dichloromethane) gave fractions of (+/-) *trans*-4-nonyl-3-octyloxtan-2-one eluting before the (+/-) *cis*-isomer. This separation was monitored by ^{1}H NMR, and the final *trans*-β-lactone (**4b**) spectrum is shown in **Figure S4**.

Synthesis and characterization of *cis*- and *trans*-9-nonadecene (5a, 5b).

The alkyne precursor, 9-nonadecyne, was produced in ~80% yield by the reaction of lithiated (n-butyl lithium/THF, -78° C) 1-decyne under reflux with 1-bromononane and catalytic amounts of sodium iodide (5). The product was distilled at 112° C/0.06 mm Hg. 9-Nonadecyne was hydrogenated to *cis*-9-nonadecene under Lindlar conditions (6). A glass-lined, 50 mL, Parr pressure

reactor (50mL) was charged with commercial Lindlar catalyst (5% $Pd/CaCO_3$ poisoned with Pb), quinoline, and 9-nonadecyne and stirred under 500 psi H_2. The reaction was followed by ^1H NMR and various batches produced ~90% yield of *cis*-9-nonadecene along with low levels (<10%) of the *trans*- isomer and nonadecane overnight. The quinoline, catalyst, and hydroquinolines were removed by filtration through silica gel. Distillation (127-129° C/0.15 mmHg) gave 90% pure *cis*-9-nonadecene (**5a**) (10% trans by ^1H NMR and GC/MS) (**Fig S1**). *trans*-9-Nonadecene (**5b**) was generated by the reduction of 9-nonadecyne by Na/NH_3 using standard conditions (7). Analysis by ^1H NMR showed a mixture of products with 57% starting alkyne, 36% *trans*-9-nonadecene, and 7% *cis*-9-nonadecene. The vinyl proton multiplets for the *trans*- and *cis*-isomers are nearly completely separated, falling at 5.39 and 5.36 ppm, respectively. The mutiplet for the four propargylic methylene protons of the starting alkyne at 2.15 ppm is clearly separated from the allylic multiplets of the *trans*- and *cis*-alkenes at 1.97 and 2.01 ppm, respectively. GC/MS also supports the conclusions of the ^1H NMR analysis. To further confirm our assignments of *cis*- and *trans* olefin, photoisomerization of *cis*-9-nonadecene to *trans* was done according to literature (8). The overnight irradiation (254nm) of a solution of 11mg of 90:10 mixture of *cis*- and *trans*-9-nonadecene and 4 mg diphenyl disulfide in hexane (1.65 g) in a quartz UV cell produced a 21:79 mixture of the *cis*- and *trans*- alkenes respectively by ^1H NMR (**Fig S2**).

^1H NMR analyses of the Ole-pathway metabolites.

All NMR spectra were acquired on a Varian Inova 400 MHz NMR spectrometer using a 5 mm Auto-X Dual Broadband probe at 20°C. Typically 512 pulses were used with a one second pulse delay. Solvent was always $CDCl_3$ with TMS as a reference. Enzyme reactions contained 1-bromo-napthalene as an integration reference. Common known contaminant resonances in our spectra are the $CDCl_3$ (singlet, 7.26 ppm), dichloromethane (singlet, 5.30 ppm), and the plasticizer bis(2-ethylhexyl) phthalate (dioctyl phthalate) (multiplet 4.22 ppm). The origin of the quartet at 3.98 ppm in several figures is unknown.

GC/MS analysis of Ole pathway metabolites.

Separation and identification of metabolites was accomplished by gas chromatography/mass spectrometry (GC/MS Agilent 7890a & 5975c) equipped with an Agilent J&W bd-ms1 column (30 m, 0.25 mm diameter, 0.25 μm film). Olefin (from either β-lactone breakdown or standards) was detected without derivatization, but the β-hydroxy acid required methylation of the carboxylic acid group by diazomethane. Preparation of ethereal alcoholic solutions of diazomethane from N-methyl-N-nitroso-*p*-toluenesulfonamide was carried out using a commercial kit from Aldrich (Technical information Bulletin No. AI-113). All samples were extracted with ethyl acetate (EtOAc) and mixed with 100 μL of diazomethane solution when needed. Five microliters of each sample were injected into the injection port (230° C). The 20 min program was as follows: hold 100° C for 3 min; ramp linearly to 280° C for 16 min; hold 280° C for 2 min.

Protein expression and purification.

All genes were *Escherichia coli (E. coli)* codon optimized and obtained from either IDT or DNA 2.0. Constructs were created using standard cloning techniques with restriction digest, ligation, and transformation reagents all purchased from New England Biolabs (NEB) and used according to manufacturer's protocols. Constructs with pET28b$^+$ contained an N-terminal 6X His tag, while pET30b$^+$ contained a C-terminal tag. All vector constructs were sequence verified and are listed in **Table S1**.

Table S1

Protein	Organism	Ascension #	Vector	Buffer Additive
OleA	*Xanthomonas campestris*	WP_011035468.1	pET28b$^+$	5 mM BME
OleC	*Xanthomonas campestris*	WP_011035474.1	pET30b$^+$	-
OleD	*Xanthomonas campestris*	WP_011035475.1	pET28b$^+$	0.025% Tween 20
OleC	*Stenotrophomonas maltophilia*	AFC01244.1	pET30b$^+$	-
OleC	*Arenimonas malthae*	WP_043804215.1	pET30b$^+$	-
OleC	*Lysobacter Dokdonensis*	WP_036166093.1	pET30b$^+$	-
OleB-C	*Micrococcus luteus*	WP_010078536.1	pET30b$^+$	-

OleB-C is a natural fusion of OleB and OleC in *Ml*

All vectors were transformed into BL21 (DE3) cells (NEB). Starter cultures (5 mL) were grown in LB media overnight at 37° C with kanamycin selection. Cells were transferred to 2 L flasks containing 1 L of LB media with kanamycin and grown to an optical density of 0.5 at 37° C. Growth was transitioned to 16° C, and expression was induced by isopropyl β-D-1-thiogalactopyranoside (IPTG, 100 μM). Cells were pelleted and resuspended in 10 mL of buffer containing 200 mM NaCl, 20 mM $NaPO_4$, and 10% glycerol at pH 7.4. Protein-specific buffer additives are indicated in **Table S1**. Cells were sonicated, spun at 33,000xg for 45 min, and the supernatant was filtered through 0.45 and 0.22 μm filters (Millipore), successively. The supernatant was then loaded onto a GE HisTrip HP 1 mL column using an Äkta fast protein liquid chromatography instrument (GE). The column was first washed with 125 mM imidazole buffer before protein elution with 400 mM imidazole buffer. Protein concentration was determined using the Bradford protein assay before samples were aliquoted and frozen at -80° C for future use.

Formation of β-lactones by OleC from β-hydroxy acids.

Reactions to generate β-lactones contained 1.0 mg OleC, 20 mg of ATP, and 30 mg of $MgCl^2 \bullet 6H_2O$ in 100 mL of buffer without glycerol (200mM NaCl and 20 mM $NaPO_4$ at pH 7.4). The appropriate β-hydroxy acids were dissolved in DMSO to 1.0 mg/mL, and 1-bromo-napthalene at 0.5 mg/mL was included as an internal reference. Replacing DMSO with ethanol for substrate delivery had no effect on the results. Reactions were initiated with 1.5 mL of the DMSO-substrate solution and allowed to run for 24 hours before three successive extractions were performed with 10 mL, 5 mL, and 5 mL of dichloromethane. Samples were allowed to evaporate at room temperature before dissolving in $CDCl_3$ for NMR. Extraction with ethylacetate instead of dichloromethane produced the same results. **Figures S8-S14** show the ^1H NMR results of OleC experiments.

Computational analyses of OleC homologues. A BLAST search with the non-redundant protein sequence database from the National Center for

Biotechnology Information (NCBI) was used to identify homologs of OleC (WP_011035474.1) from *Xanthomonas campestris pv. campestris ATCC 33913*. BLAST parameters were as follows: matrix = BLOSUM62; gap cost = Existence:11 Extentension:1; word size = 6; max matches in a query range = 0; expect threshold = 10. The most distantly related sequences that maintained 25% identity had E values of 2e-14 or better.

Supplemental Figures

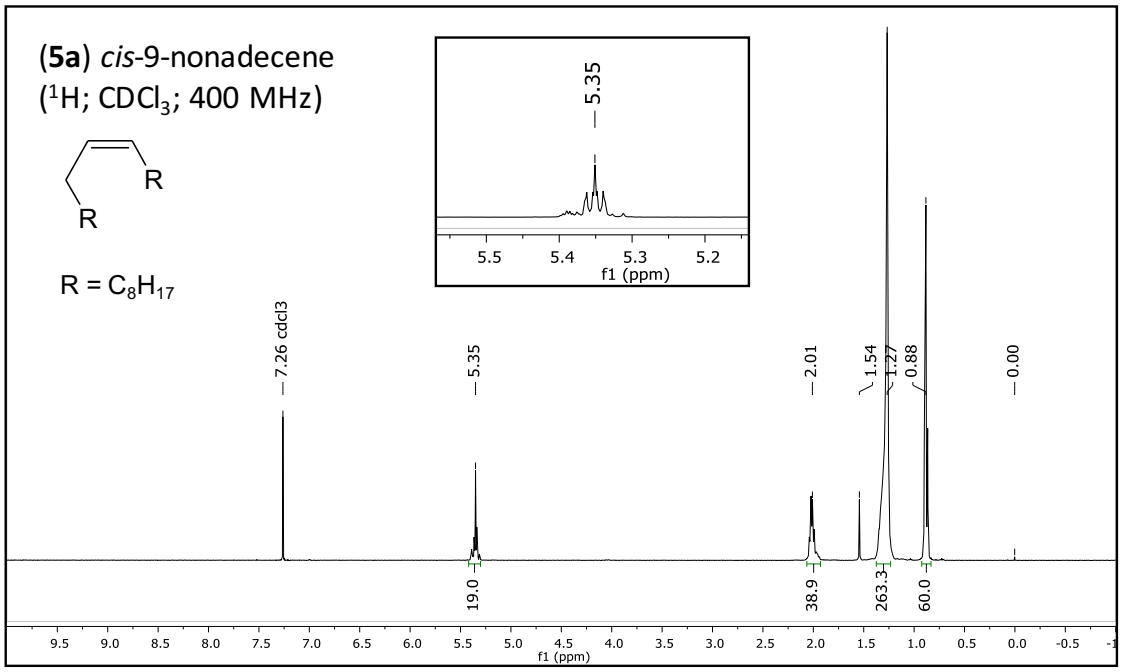


Figure S1. *cis*-Olefin standard.

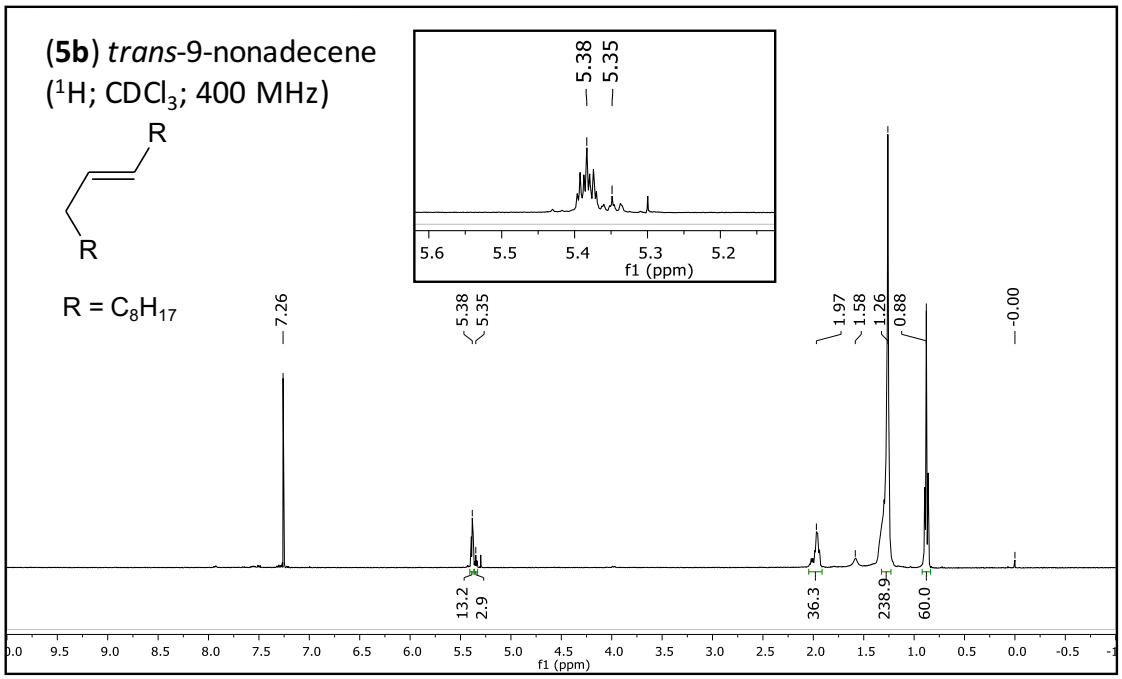


Figure S2. *trans*-Olefin standard. Approximately 20% *cis*-olefin can be seen.

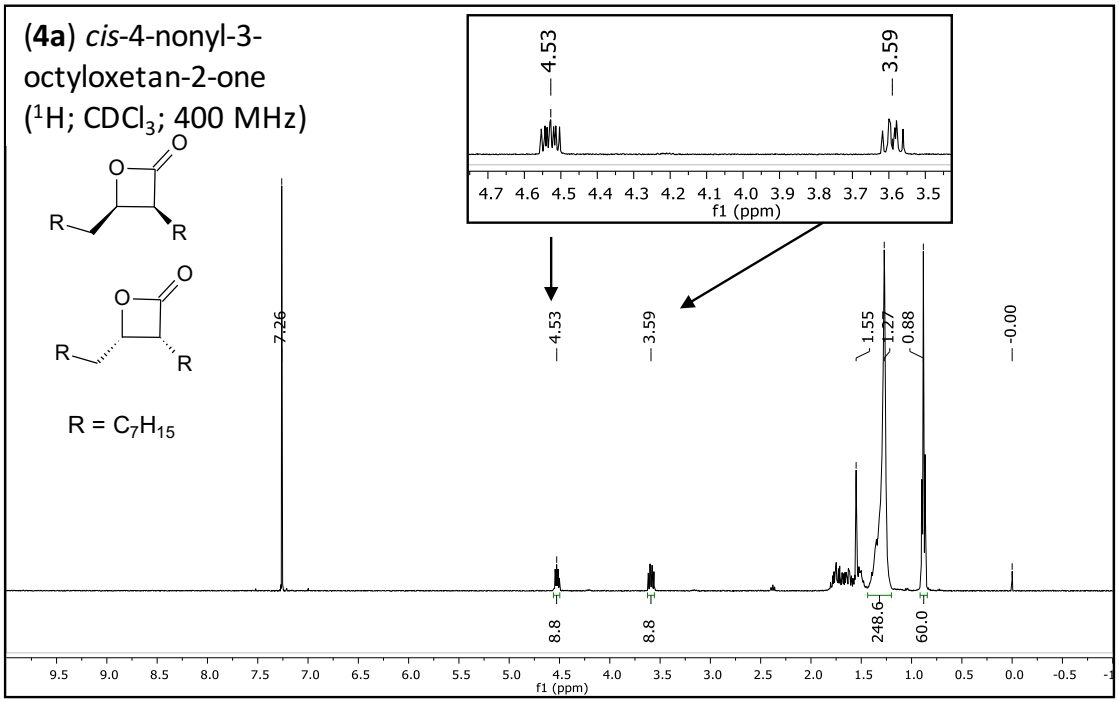


Figure S3. Synthetic *cis*-β-lactone.

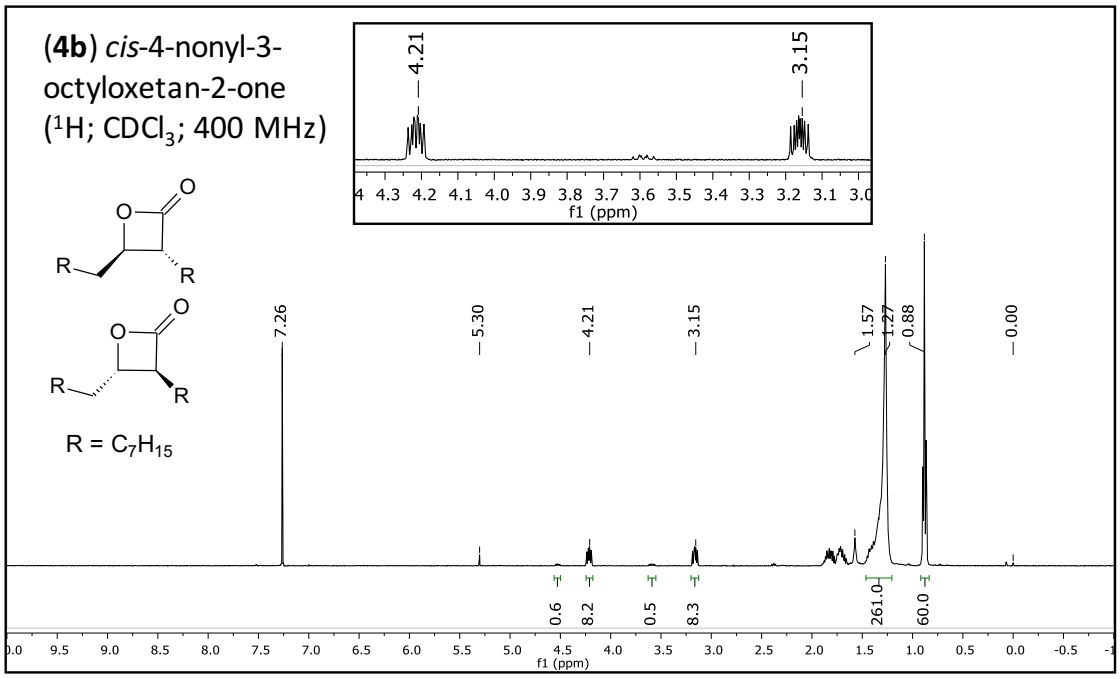


Figure S4. Synthetic *trans*-β-lactone. Singlet at 5.30 ppm is dichloromethane. A very small amount *cis*-β-lactone (<10%) can be seen at 3.59 ppm

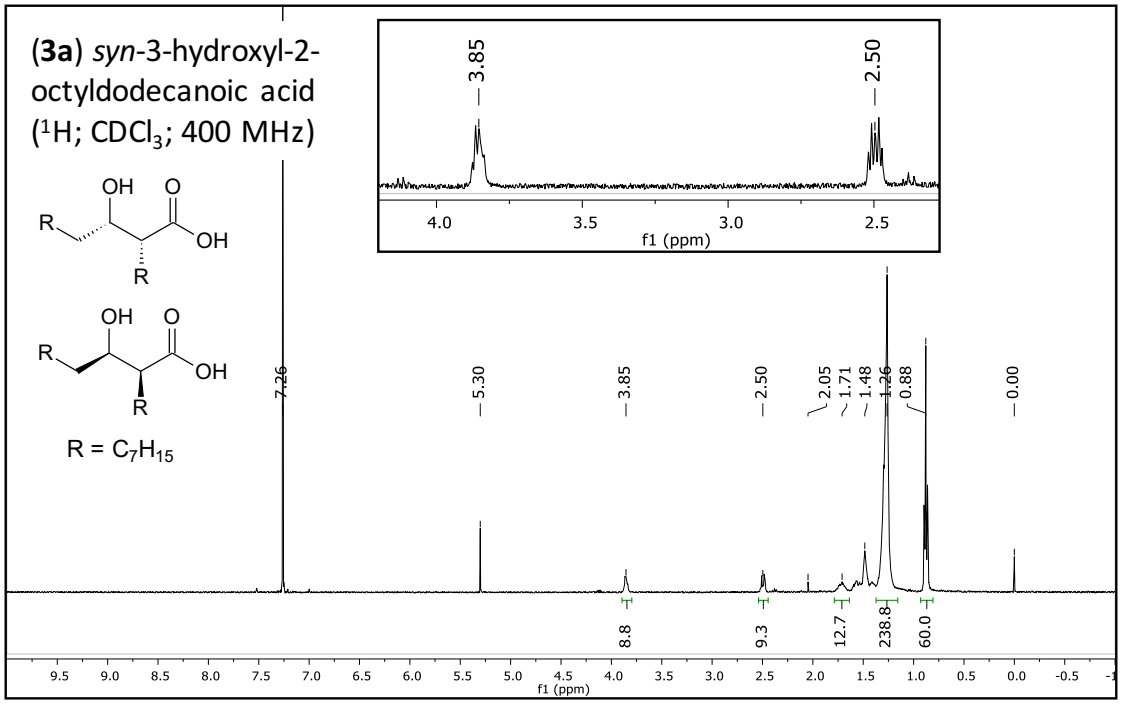


Figure S5. HPLC separated (+/-) *syn*-diasteromeric pairs (2R,3S and 2S,3R). Singlet at 5.30 is dichloromethane.

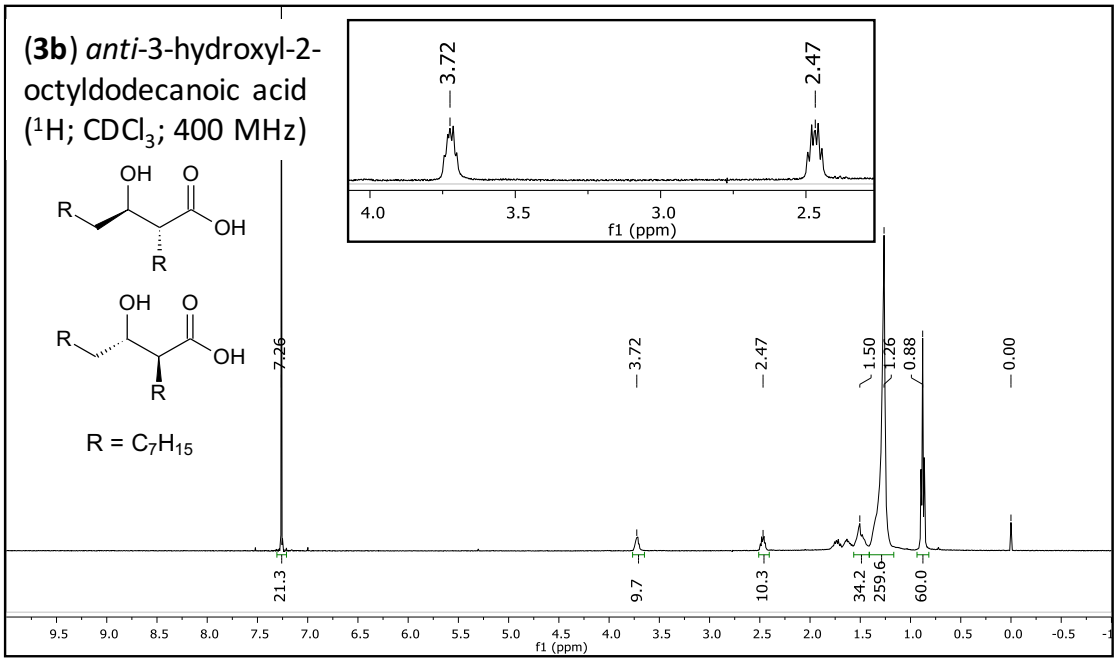


Figure S6. HPLC separated (+/-) *anti*-diasteromeric pairs (2R,3R and 2S,3S).

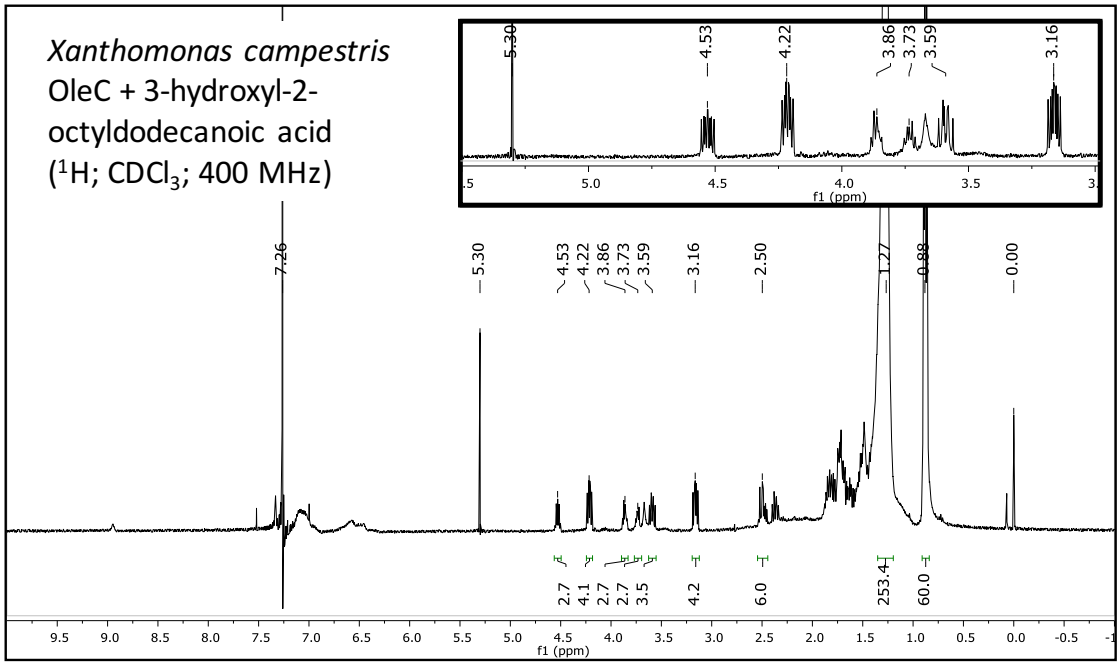


Figure S7. All four isomers of **3** with *Xc* OleC overnight. The singlet of dichloromethane is visible at 5.30 ppm.

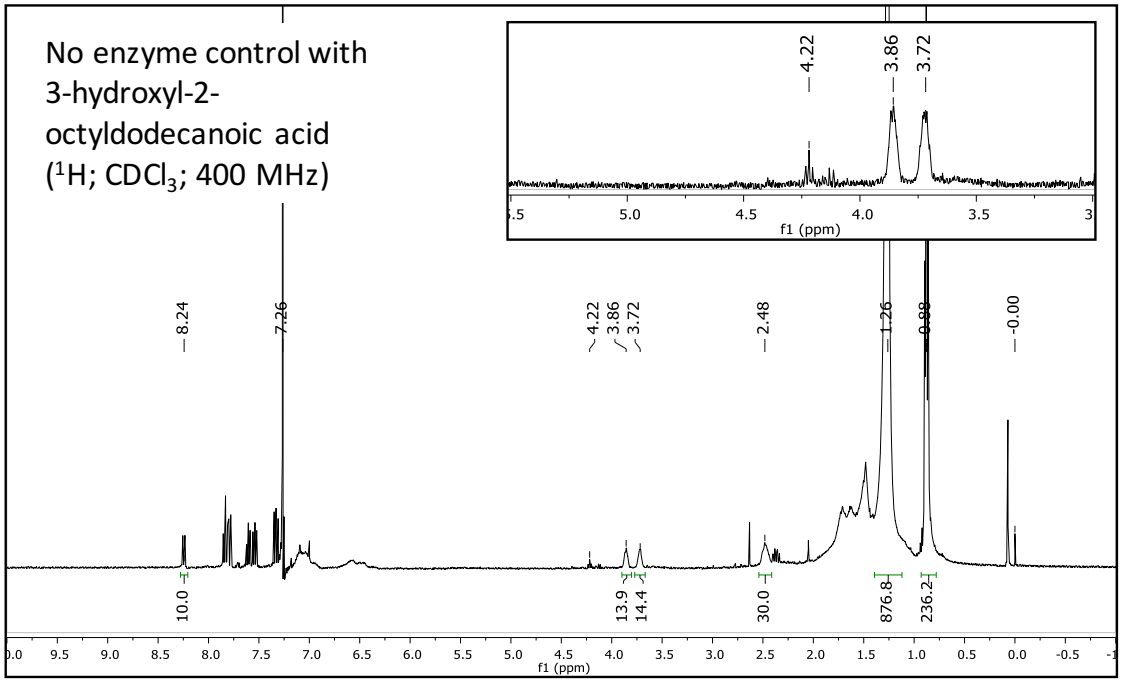


Figure S8. All four isomers of **3** with no enzyme. Some dioctyl phthalate contamination is visible at 4.22 ppm.

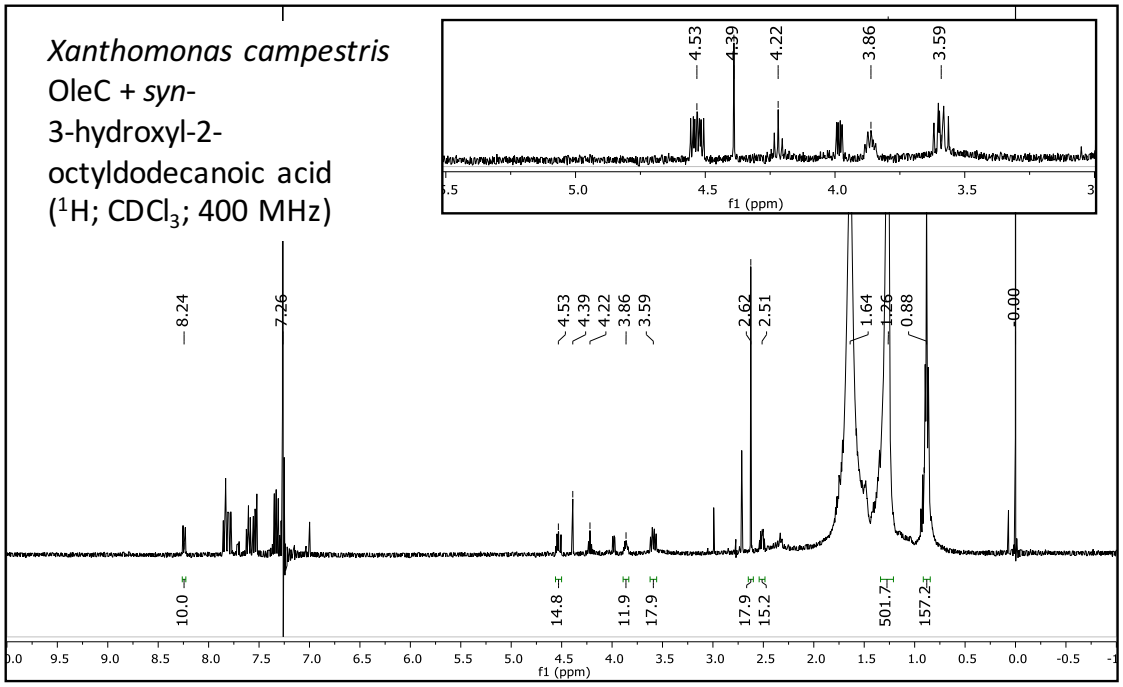


Figure S9. *Xc* OleC with **3a** overnight. A multiplet from the plasticizer dioctyl phthalate is visible at 4.22 ppm. Only (+/-) *cis*-β-lactone is observed.

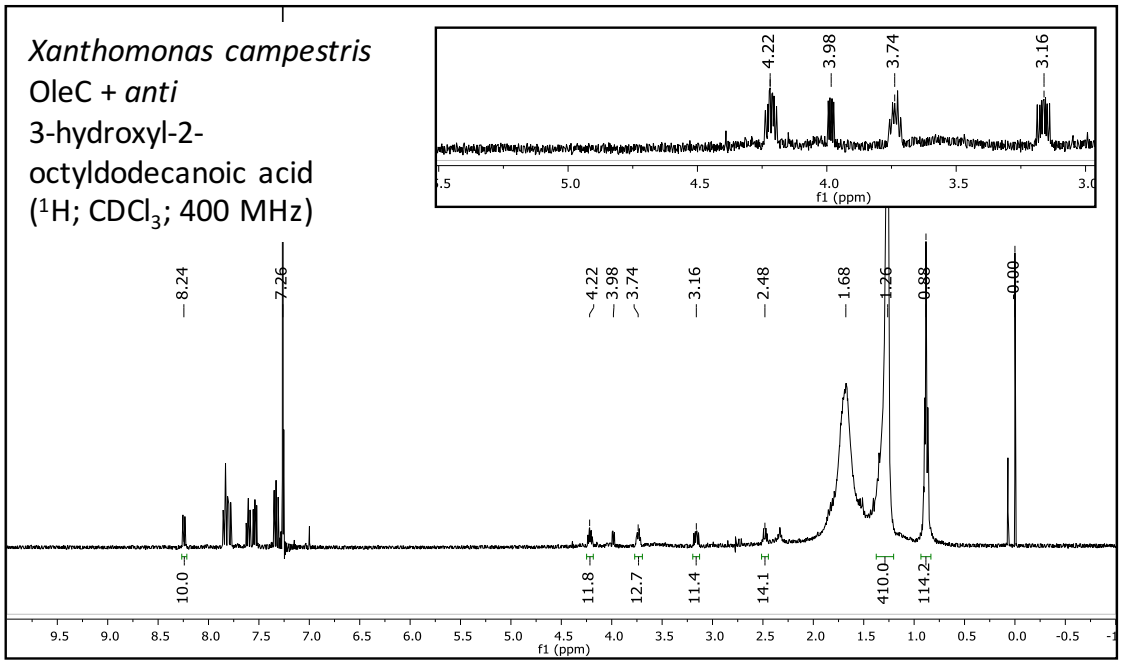


Figure S10. *Xc* OleC with **3b** overnight. Only (+/-) *trans*-β-lactone is observed.

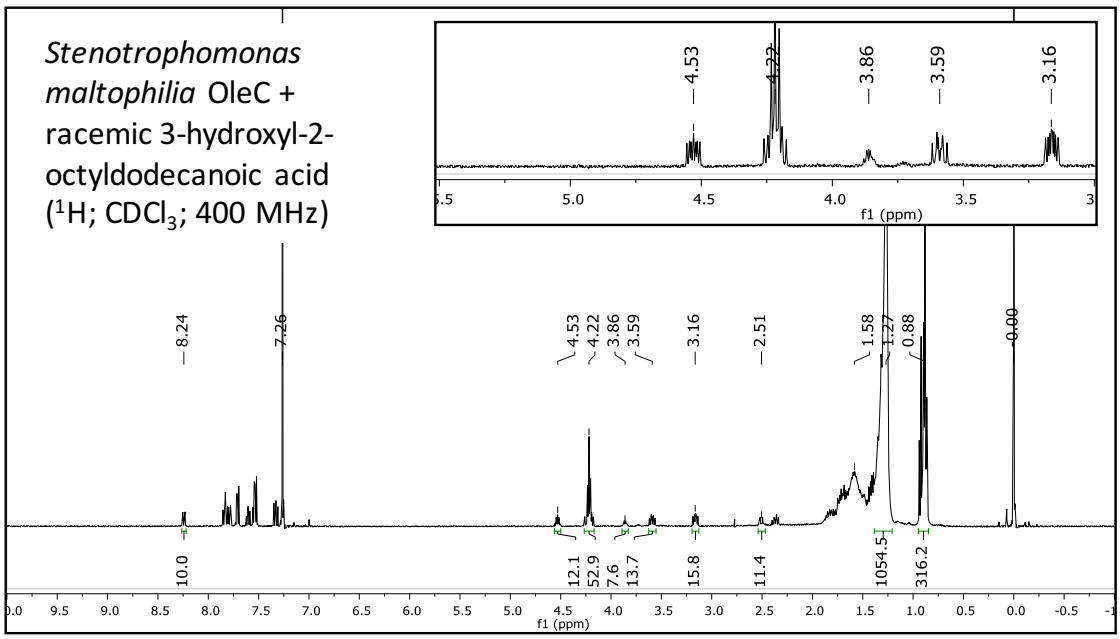


Figure S11. *Sm* OleC with **3** overnight. Some dioctyl phthalate contamination masks the *trans*-β-lactone at 4.22 ppm.

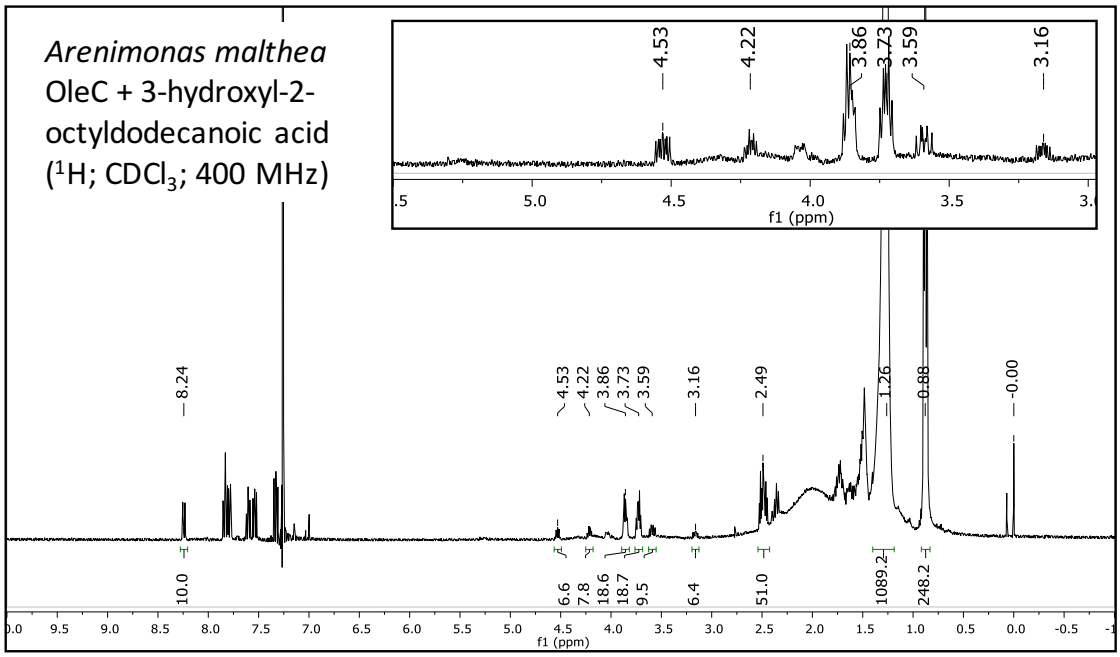


Figure S12. *Am* OleC with **3** overnight.

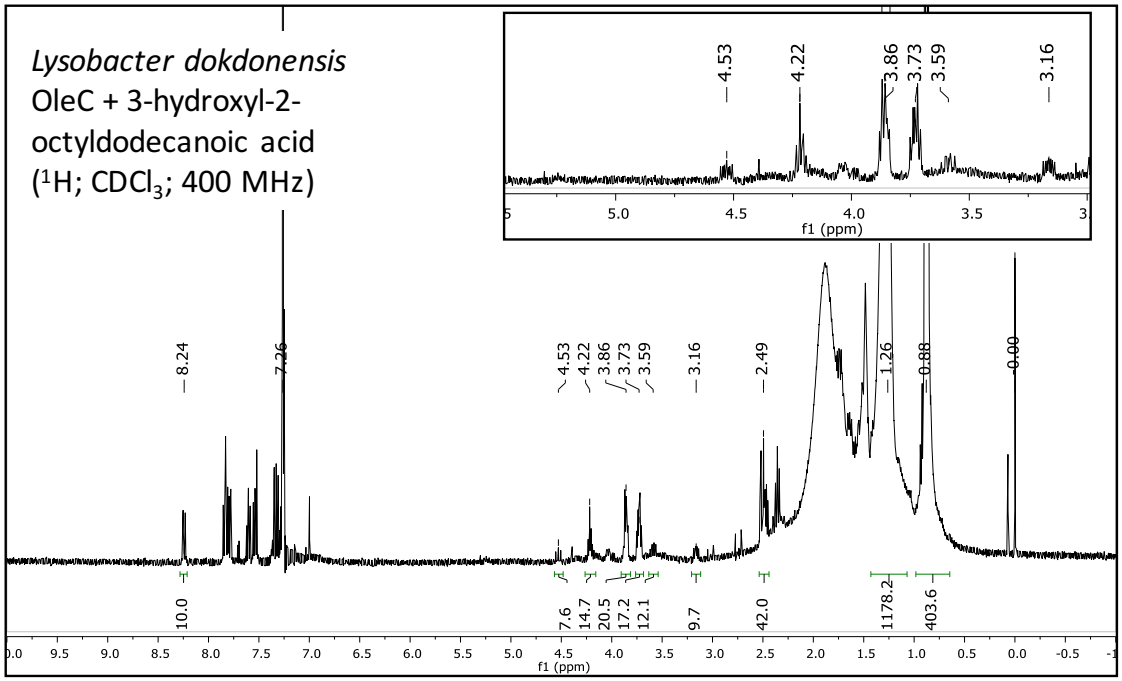


Figure S13. *Ld* OleC with **3** overnight. Some dioctyl phthalate contamination masks the *trans*-β-lactone at 4.22 ppm.

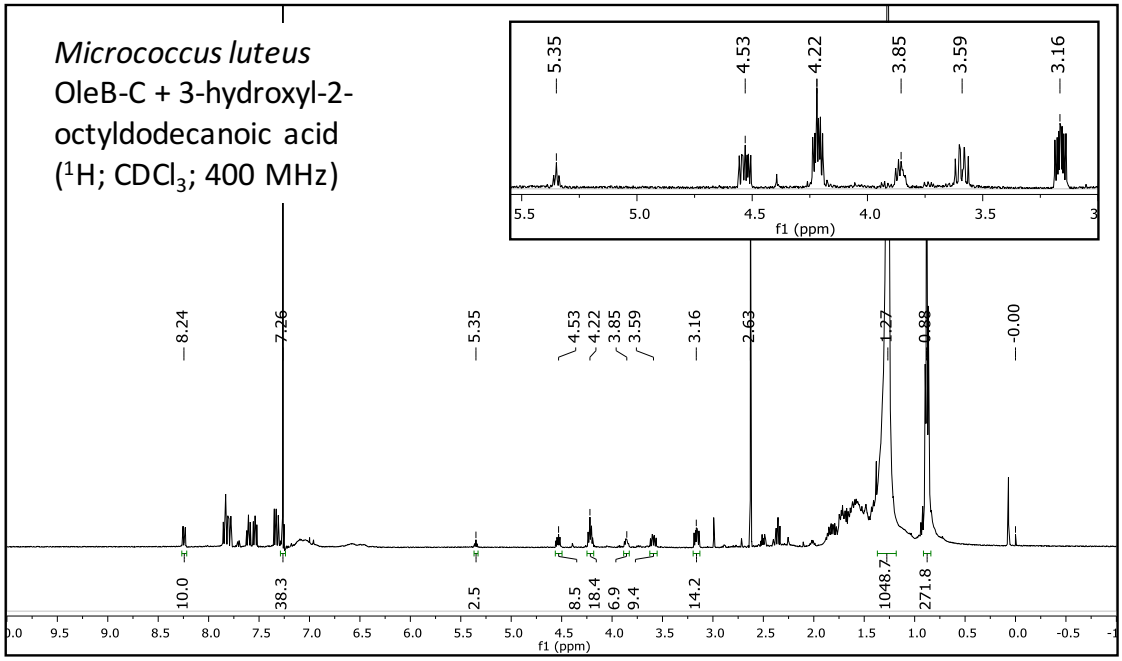


Figure S14. *Ml* OleB-C fusion with **3** overnight. Note the presence of only *cis*-olefin at 3.35 ppm.

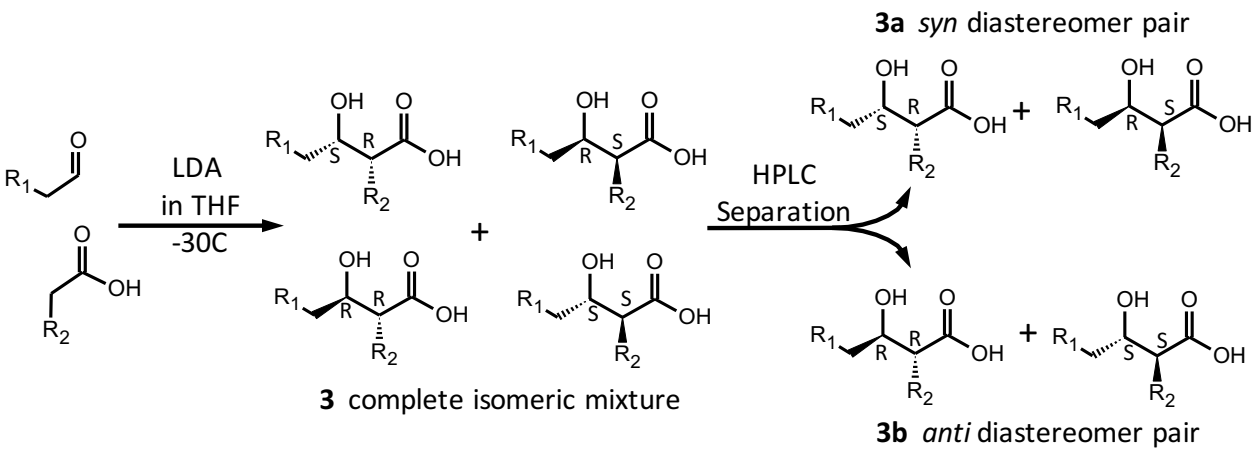


Figure S15. Preparation of -3-hydroxy-2-octyldodecanoic acid as substrates for OleC

Chapter 2- Supplemental References

(1) Mulzer, J.; Bruentrup, G.; Hartz, G.; Kuhl, U.; Blascheck, U.; Boehrer, G. *Chem. Ber.* **1981**, 3701.

(2) Bonnett, S.A.; Papireddy, K.; Higgins, S.; del Cardavre, S.; Renyolds, K.A. *Biochemistry*, **2011**, 50, 9633.

(3) Lee, M.; Gwak, H.S.; Park, B. D.; Lee, S.-T. J. *Am. Oil. Chem. Soc*. **2005**, 82 (3), 181.

(4) Crossland, R. K.; Servis, K. L.; *J. Org. Chem*., **1970**, 35 (9), 3195

(5) Buck, M.; Chong, M.J. *Tet. Lett*. **2001**, 42, 5825.

(6) Lindlar, H.; Dubuis, R. *Org. Syn.*, Coll.Vol. .5, **1973**, 880-883.

(7) House, H. O.; Kinloch, E. F. *J. Org. Chem*. **1974**, 39 (6) 747.

(8) Thalmann, A.; Oertle, K.; Gerlach, H. *Org. Syn*. **1990**, Coll.Vol. 7, 470.

Supplemental Figures

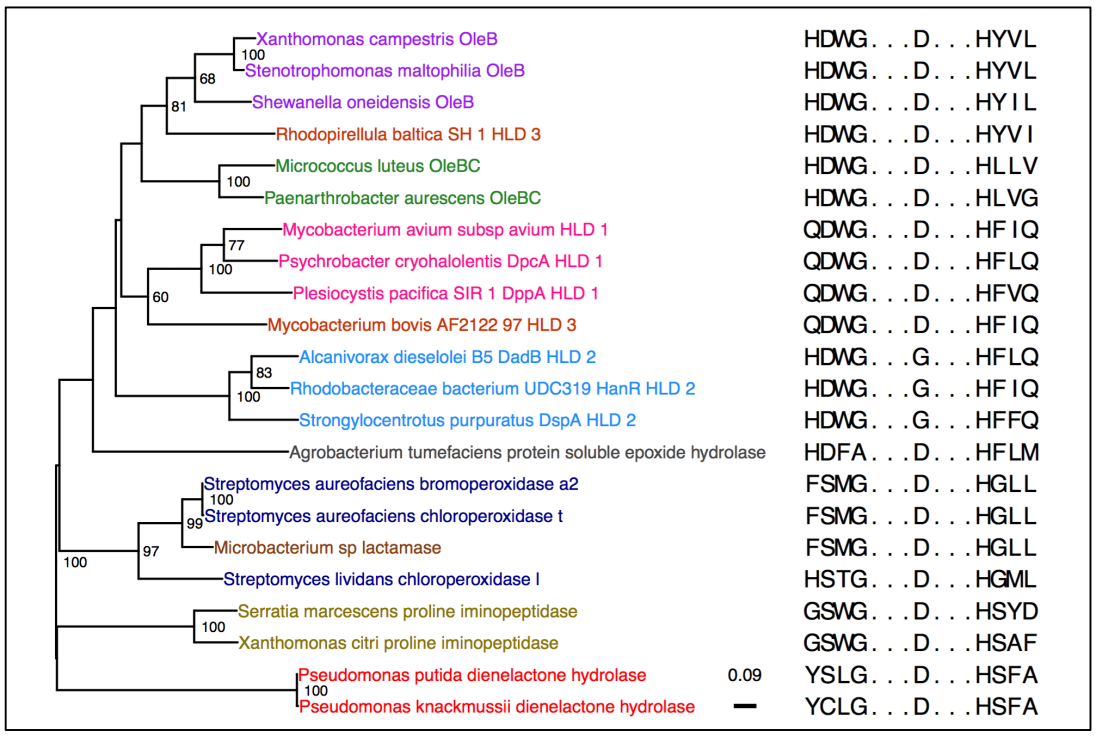


Figure S1. *Alignment of OleB and OleBC fusion proteins within bacterial α/β-hydrolase enzymes.*

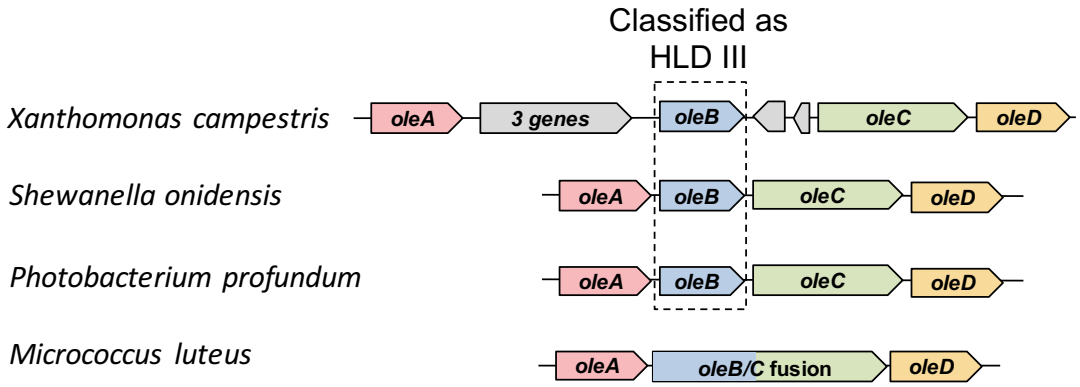


Figure S2. *OleB proteins are encoded in oleABCD gene clusters; however, many were annotated haloalkane dehalogenase (HLD) subfamily III in work by Chovancova et al. The OleB domain of OleBC fusion proteins like Micrococcus luteus was not included in Chovancova's alignments, but clusters within the HLD III subgroup.*

Supplemental Table S1

Plasmids	Description	Source
pCOLA Duet (kan)		
oleA/*oleB*	Inducible expression of a N-terminal His_{6x} tagged *oleA* with *oleB*. SacI:PstI(*oleA*)/BglII:XhoI (*oleB*)	This Study
oleB/*oleA*	Inducible expression of a N-terminal His_{6x} tagged *oleB* with *oleA*. BamHI:SacI (*oleB*)/NceI:XhoI (*oleA*)	This Study
pCDF Duet (sm)		
oleD/*oleC*	Inducible expression of *oleD* with *oleC*. NcoI:BamHI (*oleD*)/NdeI:XhoI (*oleC*)	This Study
oleD/*oleC*	Inducible expression of *oleD* with *oleC* with a C-terminal Flag tag. NcoI:BamHI (*oleD*)/NdeI:XhoI (*oleC*)	This Study
oleB/*oleC*	Inducible expression of *oleB* and *oleC* with a C-terminal Flag tag. NcoI:SacI (*oleB*)/NdeI:XhoI (*oleC*)	This Study
pET28b+ (kan)		
oleA	Inducible expression of *oleA* with a N-terminal His_{6x} tag. Cloned into NdeI:XhoI site	(1)
oleB	Inducible expression of *oleB* with a N-terminal His_{6x} tag. Cloned into NcoI:BamHI site	This Study
oleD	Inducible expression of *oleD* with a N-terminal His_{6x} tag. Cloned into NdeI:XhoI site	(1)
pET30b+ (kan)		
oleC	Inducible expression of *oleC* with a C-terminal His_{6x} tag. Cloned into NdeI:HindIII site.	(1)
oleD	Inducible expression of *oleD* with a C-terminal His_{6x} tag. Cloned into NdeI:XhoI site.	This Study

kanamycin (kan), streptomycin (sm)

Supplemental Figures

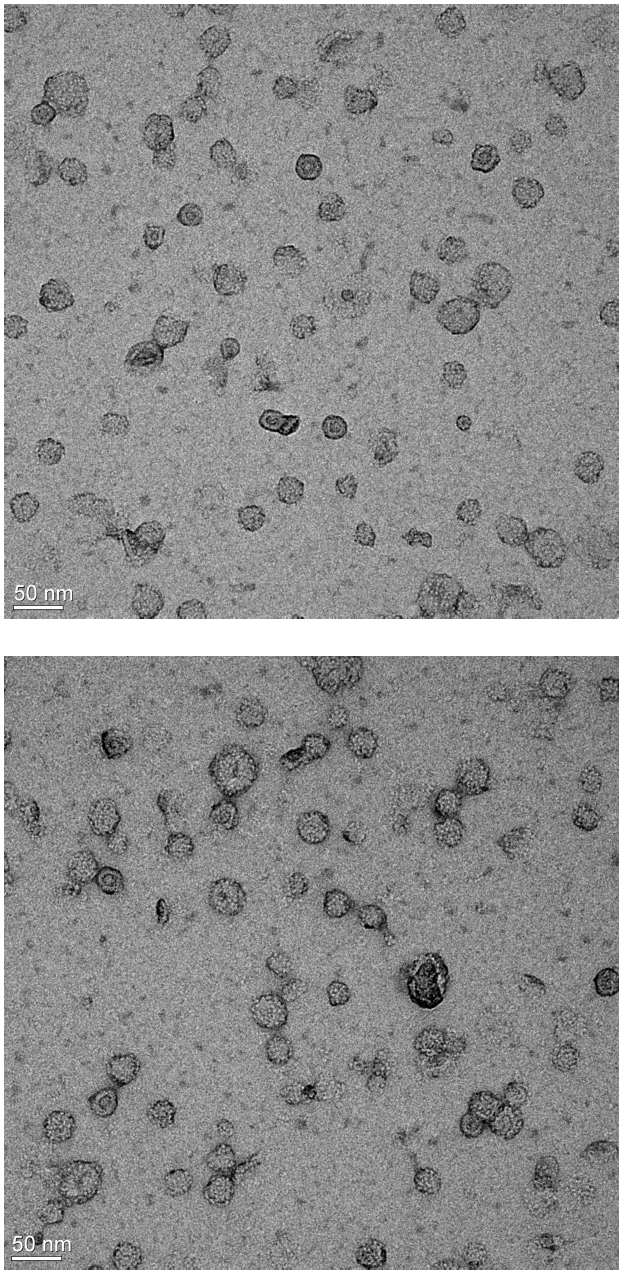


Figure S1. Additional TEM micrographs showing OleBCD (tag arrangement #4) purified over His, Gel Filtration, and FLAG columns.

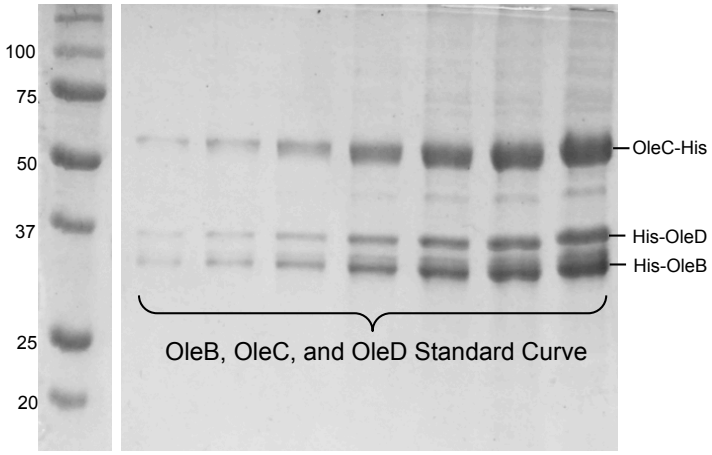


Figure S2. SDS-PAGE standard curve of Ole proteins to determine band staining densities. Separately-expressed OleB, OleC, and OleD were purified over a His affinity column. A master mix containing 4μM of each Ole protein was loaded onto an SDS-PAGE to create this gel standard curve (1 μl, 2 μl, 4 μl, 8 μl, 12 μl, 18 μl, 24 μl). The stoichiometry of the OleBCD complex was estimated by comparing band intensities to the generated standard curves (Fig. S3).

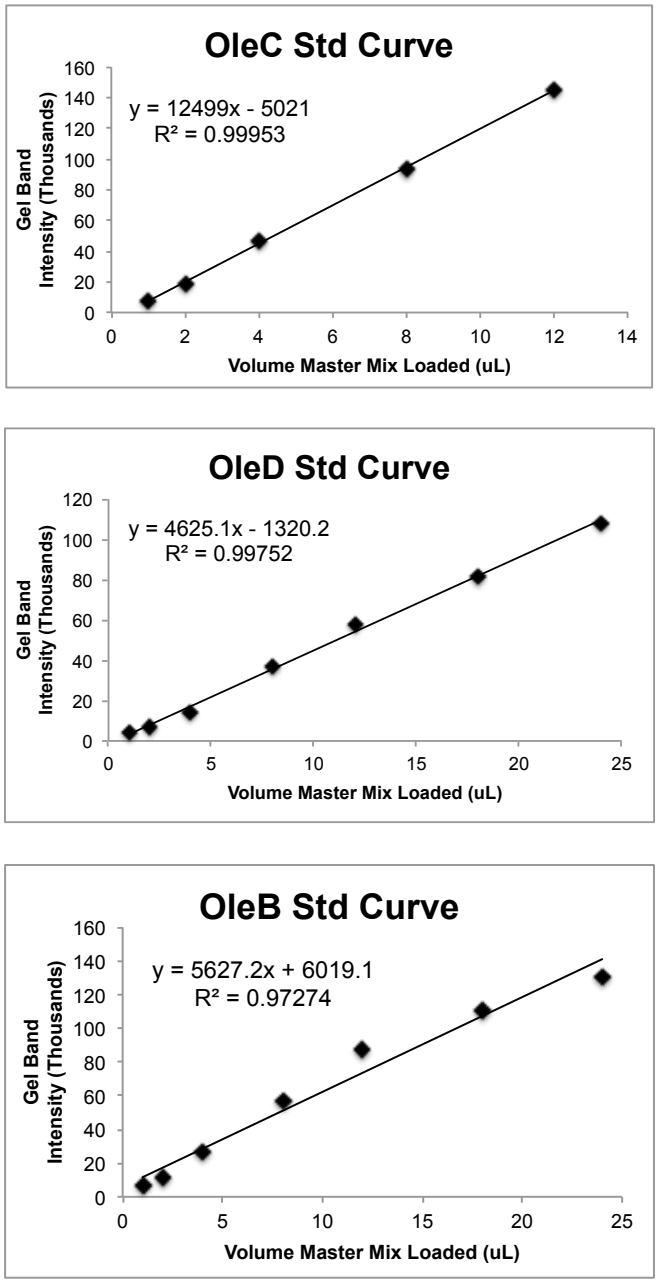


Figure S3. Densitometry analysis of Fig. S2. ImageJ was used for the analysis of gel band intensity. Band intensities of the OleBCD complex in Figure 5B were compared to these standard curves to determine a stoichiometric ratio of 4:1:4 OleB:OleC:OleD for the OleBCD complex.

Chapter 4 Supplemental References

1. Frias JA, Richman JE, Erickson JS, and Wackett LP. 2011. Purification and characterization of OleA from *Xanthomonas campestris* and demonstration of a non-decarboxylative Claisen condensation reaction. J Biol Chem 286**:**10930-10938. doi: 10.1074/jbc.M110.216127.