Published in final edited form as:

Biochem J. ; 474(23): 3871–3886. doi:10.1042/BCJ20170642.

OleA Glu117 is key to condensation of two fatty-acyl coenzyme A substrates in long-chain olefin biosynthesis

Matthew R. Jensen^{*,†}, Brandon R. Goblirsch^{*,†}, James K. Christenson^{*,†}, Morgan A. Esler^{*,†}, Fatuma A. Mohamed^{*}, Lawrence P. Wackett^{*,†}, and Carrie M. Wilmot^{*,†,1}

^{*}Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, Minneapolis, Minnesota, USA

[†]The BioTechnology Institute, University of Minnesota, Saint Paul, Minnesota, USA

Abstract

In the interest of decreasing dependence on fossil fuels, microbial hydrocarbon biosynthesis pathways are being studied for renewable, tailored production of specialty chemicals and biofuels. One candidate is long-chain olefin biosynthesis, a widespread bacterial pathway that produces waxy hydrocarbons. Found in three- and four-gene clusters, *oleABCD* encode the enzymes necessary to produce *cis*-olefins that differ by alkyl chain length, degree of unsaturation, and alkyl chain branching. The first enzyme in the pathway, OleA, catalyzes the Claisen condensation of two fatty acyl-coenzyme A molecules to form a β -keto acid. In this report, the mechanistic role of *Xanthomonas campestris* OleA Glu117 is investigated through mutant enzymes. Crystal structures were determined for each mutant as well as their complex with the inhibitor cerulenin. Complemented by substrate modelling, these structures suggest that Glu117 aids in substrate positioning for productive carbon-carbon bond formation. Analysis of acyl-coenzyme A substrate hydrolysis shows diminished activity in all mutants. When the active site lacks an acidic residue in the 117 position, OleA cannot form condensed product, demonstrating Glu117 has a critical role upstream of the essential condensation reaction. Profiling of pH dependence shows that the apparent pK_a for Glu117 is impacted by mutagenesis. Taken together, we propose that Glu117 is the general base needed to prime condensation via deprotonation of the second, non-covalently bound substrate during turnover. This is the first example of a member of the thiolase superfamily of condensing enzymes to contain an active site base originating from the second monomer of the dimer.

Summary Statement—OleA is the first reported thiolase enzyme whose active site general base, a glutamic acid (Glu117), is derived from a homodimeric interface. Structural and functional characterization of glutamic acid mutants establishes its mechanistic role in initiating olefin biosynthesis.

¹To whom correspondence should be addressed (wilmo004@umn.edu).

[‡]Current address: Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, Virginia, USA

Author Contribution: M.R.J., B.R.G., and F.A.M made the site-directed mutants and purified the mutant proteins. M.R.J. and M.A.E. screened for protein crystal conditions. M.R.J performed the crystallography and model building. M.R.J., J.K.C., and M.A.E. performed activity assays. M.R.J. and C.M.W. drafted the manuscript, with all authors intellectually contributing to the final draft. C.M.W. and L.P.W. conceived and directed the project.

Keywords

OleA; Thiolase; Condensation; Cerulenin; Crystallography

Introduction

The transition from reliance on fossil fuels to developing renewable hydrocarbons for biofuels and commodity chemicals has been a major focus in recent biotechnology research [1-3]. Microbial production of specialty hydrocarbons represents a viable option for this transition [1]. Much research has been conducted to identify candidate organisms for heterologous expression and synthesis of these compounds; however, a great deal is yet unknown about how and why the native organisms produce these hydrocarbons.

One such platform of microbial hydrocarbon production is long-chain (C_{23} - C_{33}) olefin biosynthesis. These waxy hydrocarbons were first characterized structurally in 1969, and they have been described as arising from a head-to-head condensation of two fatty acids, leading to an odd-numbered alkene with a double bond at the site of condensation [4, 5]. The genes and enzymes responsible for olefin biosynthesis have been elucidated across deeply rooted branches of the prokaryotic tree of life [5, 6]. The enzymes encoded by four clustered genes in these bacteria – *oleABCD* – comprise the Ole pathway. The first enzyme in the pathway, OleA (thiolase superfamily), catalyzes the Claisen condensation of two fatty acyl-coenzyme A (CoA) molecules to form a β -keto acid [7, 8]. OleD (short-chain dehydrogenase/reductase superfamily) reduces the β -keto acid to a β -hydroxy acid in an NADPH-dependent reaction [9]. Recently, it was demonstrated that OleC (Acyl-CoA synthetase-Nonribosomal peptide synthetase adenylation domain-Luciferase superfamily) functions as a β -lactone synthetase, using ATP hydrolysis to form a β -lactone [10]. It is hypothesized that the final decarboxylation to form the *cis*-olefin product is catalyzed by OleB (α/β -hydrolase superfamily) [10]. OleBCD were proposed to form a multi-enzyme assembly that may function to efficiently move the hydrophobic pathway intermediates between enzyme active sites, preserve the stereochemistry, and sequester the highly reactive β -lactone from the cell [11]. As well as initiating the pathway, OleA has been proposed to act as a shuttle between fatty acid metabolism and the OleBCD assembly feeding its β -keto acid product to OleD [11].

Most thiolase superfamily enzymes exist as homodimers with a subset forming homotetrameric (dimer of dimers) structures [12, 13]. Thiolases can be classified into at least three distinct functional subcategories: β -ketoacyl-acyl-carrier protein synthases (KAS), polyketide synthases (PKS), and biosynthetic thiolases/3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) synthases [12]. All superfamily enzyme mechanisms proceed through a thioester acyl-enzyme intermediate with a conserved active site cysteine (acetyl transfer/transesterification) [12]. After binding a second substrate, a thioester-dependent Claisen condensation occurs through nucleophilic attack by a carbanion at the C2 position of the nucleophilic substrate. Thiolases such as FabH (KAS type III) generate a carbanion on C2 of the acyl carrier protein-bound second substrate by promoting decarboxylation [14]. On the other hand, biosynthetic thiolase and HMG-CoA synthase generate the carbanion

nucleophile in a non-decarboxylative fashion by deprotonation of one substrate by a catalytic base, a cysteine or glutamate, respectively [15, 16]. The final step involves product release (KAS, PKS, biosynthetic thiolase) or hydrolysis of a thioester-linked product (HMG-CoA synthase). A summary of thiolase superfamily enzyme mechanisms can be found in Supplementary Figure S1. It has been shown that OleA condenses two long-chain fatty acyl-CoA substrates using a non-decarboxylative mechanism [7]. Crystal structures of wild type (WT) OleA and Cys143 mutants co-crystallized with substrate revealed that OleA has a novel three-channel active site architecture (Supplementary Figure S2) [17]. The pantetheinate channel coordinates the phosphopantetheine moiety of the acyl-CoA substrates while two alkyl channels coordinate the acyl-enzyme intermediate and second acyl-CoA substrate in alkyl channels A and B, respectively [17]. Other structurally characterized thiolase superfamily enzymes, such as FabH, PKS type II, and HMG-CoA synthase condense smaller substrates and only require two (FabH and PKS type II) or one (HMG-CoA synthase) substrate binding channels, respectively [14, 16, 18, 19]. The only thiolase superfamily enzymes known to require three distinct channels are OleA and pyrone and pseudopyrone ketosynthases [17, 20]. It has been proposed that these enzymes also utilize an active site general base, a glutamate, that comes from the second subunit of the dimer to initiate Claisen condensation, but this has never been demonstrated [17, 21].

As the first committed step of olefin biosynthesis, OleA is an important target for protein engineering. *In vivo* studies have revealed that OleA controls the hydrocarbon produced by the Ole pathway [22]. *Xanthomonas campestris* produces a wide variety of hydrocarbons ranging from C₂₈ to C₃₁ with one, two, or three double bonds *in vivo* [5]. As a consequence, OleA from *X. campestris* is promiscuous and condenses fatty acyl-CoA substrates ranging from C₈ to C₁₆ *in vitro* [7]. Therefore, the *X. campestris* Ole pathway represents a good target for protein engineering. Understanding the mechanism necessary to coordinate, activate, and condense long-chain hydrophobic substrates may provide insights needed to tailor the pathway for non-natural olefins.

In this study, we have solved the crystal structures of *X. campestris* OleA E117A, E117Q, and E117D mutant enzymes and cerulenin-inhibited states. These structures and the characterization of these enzymes suggest that Glu117 is responsible for initiating the condensation of the fatty acyl substrates during turnover by acting as a general base. Modeling studies suggest non-canonical substrate binding that contributes to diminished activity. To our knowledge, this is the first example of a thiolase superfamily enzyme using a catalytic base donated by the second subunit of the dimer.

Experimental

Preparation of OleA Glu117 site mutants

The initial cloning of the *Xanthomonas campestris* pv. *campestris* str. ATCC 33913 (NP_635607.1) synthetic *oleA* gene into pET28b+ (Novagen, Madison, WI) vector was previously described [7]. Site-directed mutagenesis was performed using custom primers from Integrated DNA Technologies (IDT, Coralville, IA). Mutations to change the Glu117 residue in WT OleA to the corresponding E117A and E117Q were introduced via PCR using Phusion polymerase (New England Biolabs). Parental template DNA was removed by

DpnI digestion. The E117D variant was purchased as an *E. coli* codon-optimized gBlock (IDT) and was assembled into pET28b+ via Gibson assembly (New England Biolabs). Successful mutations were verified by DNA sequencing (ACGT, Inc.).

Expression and purification of OleA mutant enzymes

Mutant OleA was expressed in *Escherichia coli* BL21 (DE3) competent cells (Invitrogen). Flasks of 1 liter LB media containing 50 µg/mL kanamycin sulfate were inoculated with 5 mL saturated overnight culture carrying the appropriate *oleA* vector. Cells were grown at 37°C until an A₆₀₀ of 0.5-0.6 was obtained. Expression was induced by addition of 100 µM isopropyl β-D-1-thiogalactopyranoside. After 4 hr at 37°C, cells were harvested by centrifugation at 4000 × *g* for 10 min.

Cell were resuspended in buffer consisting of 200 mM sodium chloride and 20 mM Tris HCl pH 7.4 containing a single protease inhibitor cocktail tablet (Roche Applied Science). Lysis was carried out using a sonicator (QSonica Q700). The soluble fraction was separated by centrifugation at 27,000 × *g* for 1 hr and passed through a 0.2 µm syringe filter. Filtered lysate was purified using nickel affinity chromatography (5 mL HisTrap HP column) driven by an ÄKTA FPLC (GE Healthcare Life Sciences). Protein was eluted using a linear gradient consisting of resuspension buffer with the addition of 500 mM imidazole pH 7.4. The purity and size of WT and mutant OleA were analyzed by SDS-PAGE and SimplyBlue SafeStain (Invitrogen). Fractions containing only OleA were pooled and concentrated to below 2 mL. The sample was run over size exclusion chromatography using a HiLoad 16/600 Superdex 200 pg column (GE Healthcare Life Sciences). The single peak at 80 min was collected, analyzed by SDS-PAGE, and concentrated to 115 µM (9 mg/mL) for crystallization trials. OleA E117A, E117Q, and E117D mutations were confirmed by mass spectrometry (Bruker Autoflex Speed Matrix-assisted laser desorption/ionization time of flight mass spectrometer, University of Minnesota Center for Mass Spectrometry and Proteomics).

Detection of coenzyme A hydrolysis by OleA mutants

The hydrolytic displacement of CoA from myristoyl-CoA by WT and mutant OleA enzymes were detected using 5,5'-dithio-bis-(2-nitrobenzoic acid) (DTNB) as described previously [17]. OleA (20 µg) was incubated at room temperature with 65 µM myristoyl-CoA in 490 µL of 200 mM sodium chloride, 20 mM Tris HCl pH 7.4, and 4% v/v ethanol for incubation times ranging from 5 minutes to 24 hours. Reaction mixtures were quenched by denaturing the enzyme using rapid addition of 500 µL of 8.25 M guanidine HCl pH 7.4 before DTNB (300 µM) addition. After a ten minute incubation with DTNB, the free thiol, CoA-SH, was quantified spectrophotometrically at 412 nm ($\epsilon_{412}=13,600 \text{ M}^{-1}\text{cm}^{-1}$) [7, 23, 24]. Enzyme-dependent production of CoA was controlled for by subtracting non-enzymatic hydrolysis contributions to absorbance for each mutant enzyme.

Detection of mutant OleA condensation of myristoyl-CoA substrate

WT and mutant OleA turnover assays were conducted in glass vials at room temperature. OleA (10 µg) was mixed with 50 µM myristoyl-CoA in buffer consisting of 200 mM sodium chloride and 20 mM Tris HCl pH 7.4. After incubations of 1, 3, and 24 hours, reactions were

quenched with an equal volume of ethyl acetate. The organic phase was extracted and 5 μ L was analyzed by gas chromatography, splitting the sample between the flame ionization detector and the mass spectrometer as previously described [25]. Elution program conditions consisted of 250°C inlet temperature, 100°C start temperature, 10°C/min ramp rate, 320°C final temperature, and a 5 min hold time (total time of 27 min). The thermally-decarboxylated OleA product elution time was confirmed by the commercial standard, 14-heptacosanone (Sigma Aldrich).

pH dependence and apparent pK_a determination of WT and E117D OleA

The pH dependence of WT and E117D OleA activity was investigated by product condensation and myristoyl-CoA hydrolysis as described above using the following buffer conditions: 0.05 M sodium citrate (pH 5.5), 0.2 M sodium chloride; 0.05 M bis-tris propane (pH 6.0-9.5), 0.2 M sodium chloride; 0.05 M CAPS (pH 10.0-10.5), 0.2 M sodium chloride. WT OleA activity was measured in triplicate after 8 minutes, and E117D OleA activity was measured in triplicate after 24 hours. Nonlinear fitting analyses were performed using GraphPad Prism 7.03 (GraphPad Software, La Jolla California USA). WT OleA data exhibited sigmoidal behavior that rapidly diminished after pH 9.5 and were fit to the equation $y = y_{lim}/(1+10^{pK_a-pH})$ as described previously [26]. As such, only the increase in activity going from acidic to basic conditions was compared. E117D OleA curves were fit to the Gaussian function in GraphPad Prism 7.03. Apparent pK_a values for WT were calculated as the inflection point of the curve, and apparent pK_a values for E117D were calculated as one standard deviation below the mean.

Crystallization of OleA Glu117 mutant enzymes

Initial identification of crystallization conditions for OleA was previously described [8]. E117A, E117Q, and E117D protein crystals were grown under similar conditions utilizing the hanging-drop vapor diffusion method. Mutant-alone crystallizations produced crystals of poor diffraction quality, but similar morphology to WT OleA crystals. To improve diffraction quality and potentially trap catalytic intermediates, mutant OleA (230 μ M, 9 mg/mL) was incubated with 1 mM myristoyl-CoA for times ranging from 1 to 24 hours. A 1 μ L drop of substrate-incubated mutant OleA was combined with 1 μ L of well solution (0.1 M sodium citrate pH 4.2, 0.10-0.12 M potassium phosphate dibasic, and 14-17% PEG 8000) and was suspended over 500 μ L of well solution. Crystal trays were incubated at 19°C until crystals appeared (1-2 days). Both rod-shaped and bipyramidal crystals were harvested and cryoprotected in well solution containing 25% v/v glycerol before flash-freezing in liquid nitrogen.

Co-crystals of OleA with the fungal inhibitor cerulenin were previously reported [8]. OleA E117A, E117Q, and E117D (230 μ M, 9 mg/mL) were reacted with 1 mM cerulenin for 1 hr. Dialysis was used to remove excess cerulenin and avoid phase separation in the crystallization drop. Successful crystallization conditions were as follows: 0.07-0.10 M manganese (II) chloride, 0.1 M MES pH 6.0. 13-15% w/v PEG 4000. Long, grooved crystals were harvested and cryoprotected in well solution containing 25% v/v glycerol before flash-freezing in liquid nitrogen.

X-ray data collection, processing, and refinement

All E117A and E117Q datasets were collected at GM/CA-CAT beamlines 23-ID-B (MARmosaic 300 CCD) and 23-ID-D (Pilatus3 6M) at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL). Datasets for E117D crystals were collected at SBC-CAT beamline 19-BM (ADSC Quantum 210r) at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL). Crystals were exposed to either 1.02 or 1.03 Å wavelength X-rays at 100K with an attenuated beam (5-10-fold) set to match the size of each crystal. Strong electron density was observed at monomer crystal contacts for all mutant-cerulenin datasets, prompting the investigation of anomalous diffraction by a bound metal ion. As manganese (II) chloride was a major component of the crystallization well solution, anomalous diffraction data were collected for E117Q-cerulenin crystals above (1.85 Å) and below (1.94 Å) the manganese absorption edge. Recorded data were processed in either XDS or HKL2000 [27, 28]. The 1.85 Å resolution structure of the WT OleA structure (Protein Data Bank (PDB) ID 3ROW) was used as an initial phasing model to solve the isomorphous E117A, E117Q, and E117D myristoyl-CoA co-crystal structures using direct Fourier synthesis in CCP4 (REFMAC5) [29, 30]. Although the same crystal form as WT OleA bound with cerulenin, the E117A-cerulenin co-crystal dataset required phasing by molecular replacement (PhaserMR) using the original 1.7 Å resolution WT OleA-cerulenin crystal structure (PDB ID 3S21) as a search model [31]. This phasing method was used to assign the correct space group, $P3_221$, indicated by a strong increase in the translational Z score (TFZ) peak compared to $P3_121$ enantiomorphic space group (TFZ of 45.5 vs. 21.4, respectively). E117Q-cerulenin and E117D-cerulenin co-crystal structures were phased using the E117A-cerulenin structure by direct Fourier synthesis in REFMAC5. Reflections set aside for R_{free} were carried over from the WT OleA and WT OleA-cerulenin structures, respectively [8]. Subsequent rounds of model building in COOT and restrained refinement in REFMAC5 were carried out for all structural models until all interpretable $2mF_o-DF_c$ and mF_o-DF_c electron density was explained [32]. When necessary, the occupancy of each alternative conformer for Cys143 was adjusted in 10% increments until the B factor values for the sulfur atoms were similar to each other and surrounding residues after refinement. A final round of restrained refinement with translation-libration-screw refinement was performed before statistical validation (MolProbity) [33]. Root mean square deviation (rmsd) calculations were performed using the SUPER command in PyMOL, matching all alpha carbons between aligned structures [34].

Results

Mutagenesis strategy, expression, and purification of WT and mutant OleA enzymes

An amino acid sequence alignment of known OleA enzymes shows that Glu117 is invariant (Figure 1). In order to investigate the role of Glu117, three mutants were produced: Ala117, Gln117, and Asp117. The alanine mutation was chosen to probe the effect of no carboxylate group at the 117 position. Glutamine was selected to remove the acid/base properties but retain hydrogen bonding capabilities. Aspartate was introduced to reduce the length of the residue but retain the negative charge and the potential to act as a general base.

WT and mutant OleA were recombinantly expressed and purified yielding between 15-20 mg OleA per liter of culture. Size exclusion chromatography produced a single, monodisperse population, corresponding to the OleA homodimer. The purity of this sample was estimated to be >95% pure by SDS-PAGE. Mass spectrometry demonstrated that each OleA variant had the expected molecular mass.

Detection of coenzyme A hydrolysis by OleA mutants

The colorimetric DTNB assay measures free thiol groups produced either by hydrolysis through the transesterification and condensation steps of productive OleA catalysis or potentially futile off-pathway hydrolysis of the acyl-CoA substrates (Supplementary Figure S3). Table 1 summarizes the results of WT and mutant OleA hydrolysis of myristoyl-CoA. WT OleA completely hydrolyzes available substrate in under 15 minutes. Only after 24 hour incubations with substrate are E117A and E117Q OleA mutants able to hydrolyze 25% of the substrate present, while E117D OleA hydrolyzes 50%.

Detection of mutant OleA condensation of myristoyl-CoA substrate

The ability of OleA E117A, E117Q, and E117D mutants to form product was analyzed by gas chromatography/mass spectrometry (GC/MS). As the β -keto acid product is unstable in solution, the thermally-degraded ketone side product is observed in each chromatogram (Supplementary Figure S4A). The commercially available ketone (14- heptacosanone) standard elutes at 16 minutes. The WT OleA product elutes at the same retention time and has an identical mass spectrum as the ketone standard (Supplementary Figure S4B). After 15 minutes, WT OleA converts nearly 90% of substrate to product (Table 2). E117A and E117Q OleA exhibit no product formation discernible above baseline, even after 24 hours (Table 2) (Supplementary Figure S4D). Only E117D OleA is able to condense substrate and in 24 hours produce the expected 14-heptacosanone product, albeit only ~25% conversion (Table 2) (Supplementary Figure S4C).

pH dependence and apparent pK_a determination of WT and E117D OleA

WT and E117D OleA product condensation and myristoyl-CoA hydrolysis activities were tested at varying pH levels in order to determine an apparent pK_a for the ionized residue involved in each assay (Figure 2). General behavior towards pH was identical across both assays for each enzyme, but different between WT and E117D OleA. Although a bell shaped curve was obtained for E117D OleA, WT exhibited sigmoidal pH dependency as activity increased from acidic to basic pH, but a precipitous loss of activity was observed above pH 9.5. As such, we assigned the pK_a of residue 117 to the acidic pK_a of each assay, where behavior was similar for the two enzymes. For myristoyl-CoA hydrolysis, the pK_a for WT OleA (6.4) is ~1 pH unit lower than the E117D mutant enzyme (7.2) (Figure 2A). A similar trend is observed for β -keto acid production, showing pK_a values of 6.9 for WT and 7.5 for E117D OleA (Figure 2B). A general shift from acidic to more basic pK_a values from myristoyl-CoA hydrolysis to product formation experiments is present for both OleA enzymes.

Crystallization, data processing, and refinement of OleA Glu117 mutant structures

Data collection and refinement statistics for all mutants and cerulenin co-crystals are reported in Table 3. The mutant-alone E117A, E117Q, and E117D OleA structures were solved to 1.95 Å, 1.75 Å, and 1.84 Å resolution, respectively. All were solved in the $P2_12_12_1$ space group with one OleA homodimer in the asymmetric unit (Table 3). The rmsd over all alpha carbon atoms between WT OleA and E117A, E117Q, and E117D mutants are 0.23 Å, 0.17 Å, and 0.13 Å, respectively. The E117A-cerulenin, E117Q-cerulenin, and E117D-cerulenin structures were solved to 1.66 Å, 2.07 Å, and 2.08 Å resolution, respectively. As with WT-cerulenin, the mutant-cerulenin complexes were in the $P3_221$ space group with a single monomer in the asymmetric unit (Table 3). A crystallographic symmetry mate completes the dimeric biological assembly. The calculated rmsd values between OleA WT-cerulenin and E117A-cerulenin, E117Q-cerulenin, and E117D-cerulenin are 0.15 Å, 0.15 Å, and 0.17 Å, respectively. The $P3_221$ crystal form was only attained with cerulenin binding and could not be achieved with unbound or substrate reacted forms. Similar to WT OleA-cerulenin, each mutant-cerulenin complex contains a metal ion that is present at a crystal contact. This ion was confirmed to be a manganese ion from the crystallization by anomalous diffraction (Supplementary Figure S5A).

Although co-crystallized with myristoyl-CoA, the E117A, E117Q, and E117D OleA are presented with no substrate, product, or derivative bound and are reported as unbound, mutant-alone structures. Data from crystals grown in the absence of myristoyl-CoA were of inferior diffraction quality to crystallization in the presence of substrate and were not pursued. Additional attempts to co-crystallize mutants with C_8 - C_{12} and C_{16} acyl-CoA substrates, or to soak mutant crystals with substrate, led to empty active sites and inferior diffraction. Cerulenin co-crystallization experiments were performed to assess covalent binding capacity of each mutant enzyme and how each mutant positions covalent adducts.

Overall structures of OleA Glu117 mutants

Previous OleA studies have demonstrated that three substrate-binding channels emanate from the active site centered on the catalytic Cys143 (Supplementary Figure S2) [8, 17]. All three Glu117 mutant OleA structures show these channels with no major changes to their size or shape. Aligning the active site residues of all three mutant enzymes with WT OleA shows little perturbation of active site residue positions in both monomer active sites. However, multiple conformers and side chain rotamers are observed in the E117A and E117Q OleA structures compared to WT (Figure 3). Notably, in the E117A and E117Q structures, Cys143 adopts two conformers in rapid rotational equilibrium differing by $\sim 120^\circ$ (Figure 3B & C). The occupancy of each conformer is 60%:40% for the E117A structure and 30%:70% for the E117Q structure (toward residue 117:away from residue 117). The E117D mutant contains only the orientation of Cys143 observed in WT, which is positioned away from Asp117 (Figure 3D). This may suggest the orientation of Cys143 is influenced by negative charge repulsion between the thiolate of Cys143 and the carboxylate of Glu117. Each Glu117 mutant contains an active site water molecule bound in the second oxyanion hole formed by the main-chain amides of Cys143 and Ser347. The E117D structure contains a second water molecule held in place by hydrogen bonds between the side chain carboxyl of Asp117, the side chain hydroxyl of Ser347, and the water molecule bound in oxyanion

hole 2 (Figure 3D). Oxyanion hole 1, formed by the side chains of His285 and Asn315, indicates a lack of ordered water at this site in all three structures. Compared to WT, the sidechain hydroxyl group of Ser347 is rotated away from Cys143 in the E117A and E117Q OleA structures (Figure 3B & C). Only in the WT and E117D OleA structure does the side chain of Ser347 rotate toward the catalytic cysteine, forming a hydrogen bond with a water molecule in E117D (Figure 3A & D). However, the Asp side chain is too short to form an additional hydrogen bond to the side chain of Ser347, as seen in WT. All three structures exhibit increased disorder in monomer B of the homodimer, with monomer A showing more order due to extensive crystal contacts (Supplementary Figure S5B).

As these crystal forms were co-crystallized with myristoyl-CoA substrate, several possible ligands (substrate, product, and intermediates), orientations, and occupancy levels were modeled into both alkyl channels A and B in order to test for low occupancy binding that was suggested by spurious broken density. Modelling myristoyl-CoA, myristic acid, and the acyl-Cys143 intermediate within alkyl channel A all produced maps lacking ligand density ($2mF_o-DF_c$) and/or positive difference density (mF_o-DF_c) in this channel following refinement. These models also caused an increase in R_{free} during refinement. Only the modelling and refinement of myristoyl-CoA in alkyl channel B at 10% occupancy caused a significant increase in connected positive mF_o-DF_c map for the entire ligand and decreased R_{free} . This also revealed positive difference electron density for a second alkyl chain conformer within alkyl channel B, and vice versa when only this second conformer was modelled. Modeling and refining both conformers of myristoyl-CoA simultaneously into alkyl channel B at 10% occupancy each still led to significant positive mF_o-DF_c difference density for the full length of ligand (Supplementary Figure S6B). However, simulated annealing omit maps did not reveal evidence of bound ligand and further adjustments to occupancy were unsuccessful in fully explaining the map with no observable improvements to the $2mF_o-DF_c$ electron density. The indication of low level binding of substrate is only observed in alkyl channel B of monomer B in the E117A and E117Q OleA mutant structures, which is consistent with previous reports of substrate binding only in monomer B of the dimer due to less extensive crystal contacts [17]. These observations support that the electron density is not solely due to model bias, but ultimately no ligand was included during final refinement of the E117A and E117Q OleA structures.

Cerulenin-bound structures of OleA Glu117 mutants

OleA Glu117 mutants were covalently modified with cerulenin before crystallization. Cerulenin, a fungal inhibitor of fatty acid biosynthesis, acts to covalently modify active site cysteine (OleA Cys143) residues through reaction with its epoxide ring (Figure 4) [35, 36]. E117A, E117Q, and E117D OleA have refined occupancies of bound cerulenin at 70%, 70%, and 100%, respectively (Figure 5). In the E117A and E117Q OleA structures, a second, unbound conformer of Cys143 is present at 30% occupancy (Figure 5B & C). Cerulenin is bound exclusively in alkyl channel A as observed in the WT OleA-cerulenin structure [8]. The electron density indicates that the alkyl chain of cerulenin exhibits increased mobility as it extends deeper into the binding channel. The C3 hydroxyl of cerulenin occupies oxyanion hole 1 while the oxygen of the carboxamide moiety occupies oxyanion hole 2. The side chains of Ser347 and Gln117 form hydrogen bonds with the

amine of the cerulenin carboxamide. Similar to WT OleA, Asp117 forms a hydrogen bond with Ser347 (Figure 5A & D). As with WT, all three cerulenin-bound mutant structures exhibit increased disorder including a lack of clear electron density for the β -hairpin that closes alkyl channel B from solvent (Supplementary Figure S5C & D).

Discussion

In order to further understand the mechanism of OleA and its subcategory classification in the thiolase superfamily of enzymes, we solved the structures of the catalytically-impaired OleA E117A, E117Q, and E117D mutant enzymes. Glu117 is absolutely conserved across OleA sequences and is located on a loop that in all mutants and WT is inserted into the opposite monomer active site of the physiological homodimer (Figure 6). Due to the proximity of Glu117 to bound substrates trapped in Cys143 mutant enzyme crystal structures, it was proposed that Glu117 acts as the general base needed to prime condensation between two fatty acyl-CoA substrates [17]. These structures steered us to investigate the impact Glu117 mutants have on turnover, structure, and ability to bind substrates, leading us to present an updated mechanism for OleA (Figure 7).

The first step in the mechanism of OleA involves the transesterification of the acyl moiety of the first acyl-CoA substrate. The only proposed role for Glu117 had been to initiate Claisen condensation [17]. However, it is clear that all Glu117 mutants can catalyze CoA-SH release, as is evident by decreased rates of myristoyl-CoA hydrolysis and a shift in the pK_a of the pH dependence curve for E117D OleA (Table 1) (Figure 2A). E117A and E117Q OleA are only able to hydrolyze ~25% of the available myristoyl-CoA after a 24 hour incubation, whereas WT OleA has completely hydrolyzed this amount of myristoyl-CoA in under 15 minutes (Table 1). This demonstrates that these mutants are hydrolyzing multiple myristoyl-CoA, and if this residue is absolutely required to act as a general base for condensation, a process of catalyzed hydrolysis is still occurring and the enzyme is not trapped within the first turnover. The ability of E117D OleA to release CoA-SH is also impacted even though it retains a carboxylate group in the active site similar to WT and can form β -keto acid. The apparent pK_a of the E117D mutant compared to WT OleA is shifted to a more basic pK_a by 0.8 pH units consistent with the loss of the hydrogen bond to Ser347 (Figure 3A & D). The pK_a values for the WT and E117D myristoyl-CoA hydrolysis reactions are more acidic than the pK_a values for product formation, indicating a less hydrophobic environment, which is more consistent with this pK_a being associated with CoA-SH release when only a single myristoyl-CoA is bound (Figure 2). Release of CoA-SH from the second substrate following transesterification would occur in a more hydrophobic environment in which two alkyl chains would be present in the active site. The OleA Glu117 mutant levels of myristoyl-CoA hydrolysis are higher than the previously reported C143A and C143S OleA mutants that lack the ability to undergo transesterification [17]. This suggests OleA E117A and E117Q mutants may still be able to form low levels of acyl-enzyme intermediate, albeit at a diminished rate, and the reactivity with cerulenin confirms Cys143 can still act as a nucleophile (Figure 5B & C). Our attempts to trap the acyl-enzyme intermediate by co-crystallization with these mutants were unsuccessful indicating that this species, if present, does not accumulate. However, it should be noted that modelling of substrate into the electron density of these mutants indicated low level binding of substrate

in alkyl channel B, which is where the second substrate binds in WT OleA following transesterification. As such, the higher rate of hydrolysis could also be a consequence of an increased rate of futile hydrolysis when the first substrate incorrectly occupies alkyl channel B rather than alkyl channel A (Supplementary Figures S3 & S6). Based on active site architecture, hydrolysis is likely mediated by an activated water with a candidate being the water molecule bound in oxyanion hole 2 (Figure 3B, C, & D). As E117A and E117Q OleA cannot form product, the myristoyl-CoA hydrolysis is slow and futile, releasing free CoA-SH and fatty acid (Supplementary Figure S3A). In the case of E117D OleA, this must be able to catalyze transesterification as it can form β -keto acid. The environmental changes to active site pK_a and hydrogen bonding capabilities these mutations introduce likely play a role in decreasing the rate of myristoyl-CoA hydrolysis, whether it occurs through transesterification or direct hydrolysis by an activated water molecule.

The second step of the OleA mechanism is the Claisen condensation reaction (Figure 7). Production of β -keto acid only occurs if a new carbon-carbon bond forms between the two alkyl substrates. E117A and E117Q OleA are unable to condense substrate, indicating that Claisen condensation cannot proceed without an acidic residue at position 117 (Table 2). In line with Glu117 acting as the general base, E117D OleA is the only tested mutant able to produce β -keto acid. As with WT OleA, Asp117 is able to act as a base and is capable of forming a hydrogen bond with the hydroxyl group of Ser347 when cerulenin is bound in alkyl channel A (Figure 5D) [8]. This interaction has been suggested to be important for positioning and stabilizing Glu117 for proton abstraction, likely by helping define the pK_a of the carboxylate [17]. Since this hydrogen bond can still form in the OleA E117D-cerulenin mutant, the diminished rate of product formation likely comes from the shorter length of aspartate compared to the native glutamate leading to sub-optimal positioning of the side chain for proton abstraction. In addition to slowing the rate of β -keto acid formation, E117D OleA causes a shift in pK_a compared to WT (Figure 2B). As the pK_a s observed for both WT and E117D OleA (6.9 and 7.5, respectively) are different from those observed for myristoyl-CoA release (6.4 and 7.2, respectively), this is consistent with the pK_a s being associated with a different chemical step, likely the condensation step of the OleA mechanism. The E117D mutant causes a 0.6 pH unit increase in the pK_a (6.9 to 7.5), indicating that Glu117 is involved in this step (Figure 2B). Structural changes caused by this mutation cannot be ruled out as a cause for this shift in apparent pK_a . However, OleA WT and E117D structures show very few structural differences (0.13 Å rmsd), lending further evidence to the association of this pK_a with Claisen condensation. Compared to the pK_a values for myristoyl-CoA hydrolysis, the pK_a values for OleA WT and E117D substrate condensation reactions are more basic, indicating a more hydrophobic environment. This is consistent with two bound alkyl chains in the active site, as would occur prior to condensation. The increased hydrophobicity due to the second substrate binding and from which we propose Glu117 abstracts a proton, makes ionization in the active site less favorable, causing an increase in the apparent pK_a for Glu117.

The potential for non-canonical binding of substrate may affect the rates of transesterification and Claisen condensation in OleA Glu117 mutants. It has been proposed that during turnover, the first substrate binds in alkyl channel A for transesterification. This is based on cerulenin binding in alkyl channel A of OleA, as also observed in the Glu117

mutants, and the positioning of substrate trapped in crystals of C143A and C143S OleA (Figure 5) [8, 17]. Therefore, it is surprising that there is evidence of myristoyl-CoA in alkyl channel B in the OleA E117A and E117Q structures, but not in alkyl channel A (Supplementary Figure S6). If OleA E117A and E117Q mutants improperly bind substrate in alkyl channel B, reduction in hydrolysis and turnover activity compared to WT would be expected. Binding in alkyl channel B exposes substrate to a different set of residues than is found at the opening to alkyl channel A, including the mobile β -hairpin that sequesters alkyl channel B from solvent. The increased flexibility and different chemical environment in alkyl channel B may facilitate futile hydrolysis of substrate in these mutants. Hydrolysis in alkyl channel B would allow for release of myristic acid from the active site due to dynamics of the β -hairpin, causing an apparent increase in myristoyl-CoA hydrolysis compared to that of catalytically dead Cys143 mutant enzymes [17]. This theory is consistent with the trapping of hydrolyzed myristic and lauric acids in alkyl channel A of the C143S mutant enzyme, where they appear to be trapped [17]. Interestingly, non-canonical binding of myristoyl-CoA is not observed in the OleA E117D structure, consistent with its ability to produce β -keto acid. The E117D OleA retains an acidic residue within alkyl channel B while E117A and E117Q enzymes do not. This suggests that the presence of a negative charge within alkyl channel B upon initial substrate binding may be important for proper positioning of the first substrate alkyl chain into the correct active site channel (alkyl channel A) for productive turnover. We also note that crystals of the mutants obtained in the absence of substrate are less well ordered than those of WT, as demonstrated by the observed poor diffraction, and consistent with the absence of Glu117 being associated with increased dynamics of the polypeptides.

Together these results show a clear reduction in product formation, a decrease in myristoyl-CoA hydrolysis, and a significant shift in the apparent pK_a when Glu117 is replaced by aspartate. Due to its conservation and position in the active site, we propose that Glu117 is the general base needed for proton abstraction to initiate condensation (Figure 7). When incubated with cerulenin, Cys143 from each mutant is able to be deprotonated and act as a nucleophile, properly orienting the alkylated product in alkyl channel A. This is consistent with the presence of an acidic residue not being absolutely required until after the acyl-enzyme intermediate is formed. The inability of OleA E117A and E117Q mutants to form condensed product suggests that Glu117 is directly involved during the Claisen condensation step (Figure 7). The active site residues of OleA are equivalent to those of HMG-CoA synthase, which also employs a non-decarboxylative mechanism in which proton abstraction occurs using the general base, Glu79 (*Staphylococcus aureus* numbering) [37]. However, Glu79 in HMG-CoA synthase is located on the same protein chain as the other active site residues, whereas Glu117 in OleA is located on the other polypeptide of the homodimer and is a novel feature for thiolase superfamily enzymes. The occurrence of the active site base originating from the other monomer causes the spatial orientation of the glutamate to shift in OleA relative to HMG-CoA synthase by ~ 5 Å (Figure 8A). Based on co-crystals of mutant OleA and myristoyl-CoA substrate, it has been proposed that Glu117 abstracts a proton from C2 of the second non-acylated substrate (Figure 8B) [17]. In contrast, Glu79 of HMG-CoA synthase has been proposed to abstract a proton from C2 of the acetyl-Cys111 (Figure 8C) [16, 37]. This difference in substrate activation for condensation likely arises from the need

for two alkyl coordinating channels, necessitating the shift in position of these bases and the glutamate not being sequentially equivalent between the two enzymes.

In the context of the thiolase superfamily, OleA appears to borrow features and characteristics from several of the functional subcategories. First, KAS and PKS type thiolases utilize a specificity pocket or hydrophobic channel to accommodate the growing acyl chain during condensation that is equivalent to alkyl channel B [14, 19]. In these enzymes the inhibitor cerulenin also binds in channel B [38]. OleA utilizes not only this channel but a second alkyl channel as well for proper orientation of two long-chain acyl-CoA substrates, and cerulenin binds in alkyl channel A [17]. Second, HMG-CoA synthases and biosynthetic thiolases prime condensation through a non-decarboxylative proton abstraction step. This is accomplished by an active site base, a glutamate for HMG-CoA synthase and a cysteine for biosynthetic thiolase [15, 16]. OleA appears to prime for condensation through deprotonating the nucleophilic substrate using Glu117. Third, we propose that the nucleophilic substrate for OleA is the second substrate that binds in alkyl channel B. KAS, PKS, and biosynthetic thiolases all utilize this strategy of nucleophilic attack, where the second substrate C2 carbanion attacks the thioester of the acyl-enzyme intermediate (Supplementary Figure S1A & C). Even though OleA and HMG-CoA synthase share a great deal of active site residue orientation as well as both utilizing a glutamate for proton abstraction, structural arguments support OleA activating the second substrate for condensation rather than the acyl-enzyme intermediate as observed for HMG-CoA synthase.

In conclusion, we have solved the crystal structures for OleA E117A, E117Q, and E117D mutants and their complexes with cerulenin. Analyses of the myristoyl-CoA hydrolysis and product turnover rates, along with the apparent pK_a values for these reactions, show that these mutants have slowed myristoyl-CoA hydrolysis and are unable to form product if residue 117 is not an acid. Based on these data and spatial arguments from these and previous structures, we propose that Glu117 is the general base acting to deprotonate the second substrate, priming it for nucleophilic attack during the Claisen condensation step. Glu117 appears to also be important for orienting substrate within the appropriate alkyl binding channel for productive turnover. This is the first thiolase superfamily enzyme to utilize an active site base originating from the other monomer of the dimer. Compared to other thiolase enzymes, OleA seems to utilize mechanistic strategies from the three different subcategories. Thus, OleA represents a new hybrid subcategory of thiolase mechanisms.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Final X-ray diffraction data were collected at Argonne National Laboratory, GM/CA and the Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. GM/CA@APS has been funded in whole or in part with Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). We would like to acknowledge the staff at Sectors 19 and 23 at Argonne National Laboratory Advanced Photon Source, Argonne, IL for their technical support. Computational resources and software were made available by the Minnesota Supercomputing Institute.

Preliminary X-ray diffraction data were collected at the Kahlert Structural Biology Lab, University of Minnesota, and we thank Ed Hoeffner for his technical support.

Funding: This work was support by a grant from The BioTechnology Institute, University of Minnesota to C.M.W. and L.P.W (1000014885 10866 MNT11), National Institutes of Health (NIH) Chemistry-Biology Interface Training Grant T32GM008700 (M.R.J.), NIH Future Biotechnology Development Training Grant T32GM008347 (J.K.C.), and University of Minnesota Doctoral Dissertation Fellowship (B.R.G.).

References

- Wackett LP. Engineering microbes to produce biofuels. *Curr Opin Biotechnol.* 2011; 22:388–393. [PubMed: 21071201]
- Rude MA, Schirmer A. New microbial fuels: a biotech perspective. *Curr Opin Microbiol.* 2009; 12:274–281. [PubMed: 19447673]
- Steen EJ, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, Del Cardayre SB, Keasling JD. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature.* 2010; 463:559–562. [PubMed: 20111002]
- Albro PW, Dittmer JC. Biochemistry of long-chain, nonisoprenoid hydrocarbons III Metabolic relation of long-chain fatty acids and hydrocarbons and other aspects of hydrocarbon metabolism in *Sarcina lutea*. *Biochemistry.* 1969; 8:1913–1918. [PubMed: 5785213]
- 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; 76:3850–3862. [PubMed: 20418421]
- Friedman L, Da CB. Hydrocarbon-producing genes and methods of their use. Google Patents. 2010
- 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; 286:10930–10938. [PubMed: 21266575]
- 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; 51:4138–4146. [PubMed: 22524624]
- 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; 50:9633–9640. [PubMed: 21958090]
- Christenson JK, Richman JE, Jensen MR, Neufeld JY, Wilmot CM, Wackett LP. beta-Lactone Synthetase Found in the Olefin Biosynthesis Pathway. *Biochemistry.* 2017; 56:348–351. [PubMed: 28029240]
- Christenson JK, Jensen MR, Goblirsch BR, Mohamed F, Zhang W, Wilmot CM, Wackett LP. Active Multienzyme Assemblies for Long-Chain Olefinic Hydrocarbon Biosynthesis. *J Bacteriol.* 2017; 199
- Haapalainen AM, Merilainen G, Wierenga RK. The thiolase superfamily: condensing enzymes with diverse reaction specificities. *Trends Biochem Sci.* 2006; 31:64–71. [PubMed: 16356722]
- Modis Y, Wierenga RK. A biosynthetic thiolase in complex with a reaction intermediate: the crystal structure provides new insights into the catalytic mechanism. *Structure.* 1999; 7:1279–1290. [PubMed: 10545327]
- Musayev F, Sachdeva S, Scarsdale JN, Reynolds KA, Wright HT. Crystal structure of a substrate complex of *Mycobacterium tuberculosis* beta-ketoacyl-acyl carrier protein synthase III (FabH) with lauroyl-coenzyme A. *J Mol Biol.* 2005; 346:1313–1321. [PubMed: 15713483]
- Modis Y, Wierenga RK. Crystallographic analysis of the reaction pathway of *Zoogloea ramigera* biosynthetic thiolase. *J Mol Biol.* 2000; 297:1171–1182. [PubMed: 10764581]
- Chun KY, Vinarov DA, Zajicek J, Miziorko HM. 3-Hydroxy-3-methylglutaryl-CoA synthase A role for glutamate 95 in general acid/base catalysis of C-C bond formation. *J Biol Chem.* 2000; 275:17946–17953. [PubMed: 10748155]
- 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; 291:26698–26706. [PubMed: 27815501]

18. Gajiwala KS, Margosiak S, Lu J, Cortez J, Su Y, Nie Z, Appelt K. Crystal structures of bacterial FabH suggest a molecular basis for the substrate specificity of the enzyme. *FEBS Lett.* 2009; 583:2939–2946. [PubMed: 19665020]
19. Pan H, Tsai S, Meadows ES, Miercke LJ, Keatinge-Clay AT, O'Connell J, Khosla C, Stroud RM. Crystal structure of the priming beta-ketosynthase from the R1128 polyketide biosynthetic pathway. *Structure.* 2002; 10:1559–1568. [PubMed: 12429097]
20. Sucipto H, Sahner JH, Prusov E, Wenzel SC, Hartmann RW, Koehnke J, Muller R. In vitro reconstitution of alpha-pyrone ring formation in myxopyronin biosynthesis. *Chemical Science.* 2015; 6:5076–5085.
21. Kresovic D, Schempp F, Cheikh-Ali Z, Bode HB. A novel and widespread class of ketosynthase is responsible for the head-to-head condensation of two acyl moieties in bacterial pyrone biosynthesis. *Beilstein J Org Chem.* 2015; 11:1412–1417. [PubMed: 26425196]
22. 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; 76:3842–3849. [PubMed: 20418444]
23. Alexson SE, Nedergaard J. A novel type of short- and medium-chain acyl-CoA hydrolases in brown adipose tissue mitochondria. *J Biol Chem.* 1988; 263:13564–13571. [PubMed: 2901416]
24. Ellman GL. A colorimetric method for determining low concentrations of mercaptans. *Arch Biochem Biophys.* 1958; 74:443–450. [PubMed: 13534673]
25. Aukema KG, Kasinkas L, Aksan A, Wackett LP. Use of silica-encapsulated *Pseudomonas* sp. strain NCIB 9816-4 in biodegradation of novel hydrocarbon ring structures found in hydraulic fracturing waters. *Appl Environ Microbiol.* 2014; 80:4968–4976. [PubMed: 24907321]
26. Vararattanavech A, Ketterman AJ. A functionally conserved basic residue in glutathione transferases interacts with the glycine moiety of glutathione and is pivotal for enzyme catalysis. *Biochem J.* 2007; 406:247–256. [PubMed: 17523921]
27. Kabsch W. Xds. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:125–132. [PubMed: 20124692]
28. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol.* 1997; 276:307–326.
29. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr.* 2011; 67:235–242. [PubMed: 21460441]
30. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr.* 1997; 53:240–255. [PubMed: 15299926]
31. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr.* 2007; 40:658–674. [PubMed: 19461840]
32. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004; 60:2126–2132. [PubMed: 15572765]
33. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:12–21. [PubMed: 20057044]
34. Schrodinger LLC. The PyMOL Molecular Graphics System, Version 1.8.05. 2015
35. Omura S. The antibiotic cerulenin, a novel tool for biochemistry as an inhibitor of fatty acid synthesis. *Bacteriol Rev.* 1976; 40:681–697. [PubMed: 791237]
36. Ronnett GV, Kleman AM, Kim EK, Landree LE, Tu Y. Fatty acid metabolism, the central nervous system, and feeding. *Obesity (Silver Spring).* 2006; 14 Suppl 5:201S–207S. [PubMed: 17021367]
37. Theisen MJ, Misra I, Saadat D, Campobasso N, Miziorko HM, Harrison DH. 3-hydroxy-3-methylglutaryl-CoA synthase intermediate complex observed in “real-time”. *Proc Natl Acad Sci U S A.* 2004; 101:16442–16447. [PubMed: 15498869]
38. Price AC, Choi KH, Heath RJ, Li Z, White SW, Rock CO. Inhibition of beta-ketoacyl-acyl carrier protein synthases by thiolactomycin and cerulenin Structure and mechanism. *J Biol Chem.* 2001; 276:6551–6559. [PubMed: 11050088]

Abbreviations

CoA	coenzyme A
KAS	β -ketoacyl-acyl-carrier protein synthases
PKS	polyketide synthases
HMG-CoA synthase	3-hydroxy-3-methylglutaryl-CoA synthase
WT	wild type
DTNB	5,5'-dithio-bis-(2-nitrobenzoic acid)
LB	Lysogeny broth
MES	2-morpholin-4-ylethanesulfonic acid
PEG	polyethylene glycol
PDB	Protein Data Bank
GC/MS	gas chromatography–mass spectrometry
GC-FID	gas chromatography–flame ionization detector
rmsd	root mean square deviation

<i>X. campestris</i>	V	S	R	D	Y	L	117	E	P	S	T	A	S
<i>G. uraniireducens</i>	V	C	R	E	Q	F	97	E	P	A	T	A	C
<i>S. oneidensis</i>	V	C	R	D	Q	H	97	E	P	A	T	A	C
<i>S. ambofaciens</i>	T	S	Q	D	L	I	97	E	P	A	T	A	H
<i>A. aurescens</i>	V	T	R	R	N	L	103	E	P	S	V	A	V
<i>K. radiotolerans</i>	V	T	R	P	H	L	119	E	P	A	V	A	T
<i>M. luteus</i>	V	T	R	A	T	L	106	E	P	S	V	A	V
<i>S. maltophilia</i>	V	S	R	D	Y	L	97	E	P	S	T	A	S
<i>C. aggregans</i>	V	C	Q	D	Y	L	97	E	P	S	T	A	C

Figure 1.

Amino acid sequence alignment of a subset of bacterial OleA enzymes showing the invariant glutamate residue. Bacterial genus and species names are as follows: *Xanthomonas campestris* pv. *campestris* str. ATCC 33913, *Geobacter uraniireducens* Rf4, *Shewanella oneidensis* MR-1, *Streptomyces ambofaciens* ATCC 23877, *Arthrobacter aurescens* TC1, *Kineococcus radiotolerans* SRS30216, *Micrococcus luteus* NCTC 2665, *Stenotrophomonas maltophilia* K279a, *Chloroflexus aggregans* DSM 9485.

Figure 2.

OleA WT and E117D enzyme activity vs. pH curves. **A)** Myristoyl-CoA hydrolysis by WT (black circles) and E117D (white squares) OleA at indicated pH levels. Calculated pK_a values for each enzyme are shown in the inset table. **B)** Turnover of β-keto acid by WT (black circles) and E117D (white squares) OleA at indicated pH levels. Calculated pK_a values for each enzyme are shown in the inset table. All WT OleA activity was measured after eight minutes, and all E117D OleA activity was measured after 24 hours. Results are presented as the mean of triplicate data ± standard deviation. pK_a values were calculated from nonlinear regression of each curve as detailed in “Experimental” and are presented in each table ± standard error for that variable.

Figure 3.

OleA WT and mutant active site comparison. **A)** Active site of WT OleA (PDB ID 3ROW) showing the hydrogen bonding network and water molecules (gray) bound in both oxyanion holes (1 and 2, blue dashed circles) [8]. **B)** Active site of OleA E117A showing the hydrogen bonding network and a water molecule (blue) bound in the second oxyanion hole (2, blue dashed circle). Cys143 adopts two conformers at 60%:40% occupancy. **C)** Active site of OleA E117Q showing the hydrogen bonding network and a water molecule (green) bound in the second oxyanion hole (2, blue dashed circle). Cys143 adopts two conformers at 70%:30% occupancy. **D)** Active site of OleA E117D showing the hydrogen bonding network and water molecules (yellow). All hydrogen bonds are represented by black dashed lines. Monomer A was used to generate each figure.

Figure 4.

Covalent inhibition of OleA Cys143 by cerulenin. Upon binding, cerulenin occupies alkyl channel A in WT OleA (PDB ID 3ROW) and is stabilized by interactions with both oxyanion holes (OxH, blue) [8]. Oxyanion hole 1 is formed by the side chains of His285 and Asn315. Oxyanion hole 2 is formed by the main-chain amides of Cys143 and Ser347. Dashed black lines represent hydrogen bonds.

Figure 5.

Cerulenin-bound mutant OleA active site. **A)** Active site of WT OleA-cerulenin (gray, PDB ID 3S21) with ligand bound in alkyl channel A at 100% occupancy [8]. **B)** Active site of OleA E117A-cerulenin (blue) with ligand bound in alkyl channel A at 70% occupancy. **C)** Active site of OleA E117Q-cerulenin (green) with ligand bound in alkyl channel A at 70% occupancy. **D)** Active site of OleA E117D-cerulenin (yellow) with ligand bound in alkyl channel A at 100% occupancy. All hydrogen bonds are represented as black dashed lines. In all panels, oxyanion hole 1 (1, blue dashed circle) is occupied by the C3 hydroxyl of cerulenin, and oxyanion hole 2 (2, blue dashed circle) is occupied by carboxamide carbonyl oxygen of cerulenin. The simulated annealing ligand omit $mF_o - DF_c$ maps are contoured at 2σ .

Figure 6. OleA E117D showing the dimer interface. **A)** OleA E117D dimer highlighting the insertion of the monomer B loop (red) containing Asp117 into the active site of monomer A. **B)** Monomer A of OleA E117D showing the active site loop containing Asp117 donated by monomer B (red).

1. Transesterification

2. Claisen condensation

3. β-keto thioester hydrolysis

Figure 7.

Proposed three-step mechanism of β-keto acid formation by OleA. Glu117 acts as a general base that deprotonates C2 of the second acyl-CoA substrate to initiate condensation.

Figure 8.

Active site comparisons between OleA and HMG-CoA synthase. **A)** Active site overlay between WT OleA (gray, PDB 3ROW) and HMG-CoA synthase (green, PDB 1XPK) [8, 37]. HMG-CoA synthase is modelled with an acetylated Cys111. The active site glutamate base is shifted between the structures by an average of 5 Å due to the active site loop from which they originate. **B)** OleA C143S (purple, PDB 4KU3) active site substrate complex showing substrate coordination. The purple solid line represents the position of the covalent bond between the first substrate (bound in alkyl channel A, orange) and the native Cys143 [17]. The gray dashed line represents the proposed deprotonation of the second substrate (bound in alkyl channel B, green) by Glu117. The red dashed line shows the likely path and direction of nucleophilic attack. **C)** HMG-CoA synthase active site poised for condensation between the acetylated Cys111 and acetoacetyl-CoA. The gray dashed line represents the priming deprotonation step. The red dashed line shows the path and direction of nucleophilic attack.

Table 1

Hydrolysis of myristoyl-CoA by WT and mutant OleA enzymes.

Enzyme	Enzyme concentration (μM)	Rxn time (hrs)	CoA-SH product (nmol) ^a	Percent yield (%) ^b
WT	0.5	0.25	33 ± 0.5	100
E117A	0.5	0.25	0.8 ± 0.1	2.5
E117A	0.5	24	8.6 ± 0.3	27
E117Q	0.5	0.25	<0.1	<0.1
E117Q	0.5	24	7.8 ± 0.2	24
E117D	0.5	0.25	<0.1	<0.1
E117D	0.5	24	16 ± 0.9	50

Values shown are the average in triplicate with standard deviation.

^aFree coenzyme A detected as described under "Experimental"

^bStarting substrate was 32.5 nmoles; 32.5 nmoles product is 100% turnover

Table 2

Product turnover by WT and mutant OleA enzymes.

Enzyme	Enzyme concentration (μ M)	Rxn time (hrs)	Ketone peak area (a.u.) ^a	Normalized ketone peak area (%) ^b
WT	0.25	0.25	5599 $\pm$ 146	86.1
WT	0.25	24	6510 $\pm$ 155	100
E117A	0.25	24	N.D. ^c	N.D. ^c
E117Q	0.25	24	N.D. ^c	N.D. ^c
E117D	0.25	24	1610 $\pm$ 131	24.7

^aValues shown are the average in triplicate of the GC-FID peak area (arbitrary units) with standard deviation.^bGC-FID peak area as a percentage of WT 24 hour reaction peak area^cN.D: not detectable; peak area below 50 a.u. considered indistinguishable from baseline

Table 3

OleA Glu117 mutant data collection and refinement statistics^a.

	E117A OleA	E117Q OleA	E117D OleA	E117A OleA – Cerulenin cocrystal	E117Q OleA – Cerulenin cocrystal	E117D OleA – Cerulenin cocrystal
Data Collection						
Wavelength (Å)	1.03	1.03	1.02	1.03	1.03	1.02
Space group	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 3 ₂ 2 1	P 3 ₂ 2 1	P 3 ₂ 2 1
Unit cell (Å)	82.4 × 86.0 × 103.6	82.2 × 85.9 × 103.0	81.9 × 85.4 × 102.8	90.0 × 90.0 × 69.6	90.3 × 90.3 × 69.4	89.8 × 89.8 × 69.5
Unit cell (°)	α = β = γ = 90	α = β = γ = 90	α = β = γ = 90	α = β = 90; γ = 120	α = β = 90; γ = 120	α = β = 90; γ = 120
Resolution (Å)	50.0-1.96 (1.99-1.96)	50.0-1.75 (1.78-1.75)	50.0-1.84 (1.87-1.84)	50.0-1.66 (1.69-1.66)	29.6-2.07 (2.13-2.07)	50.0-2.08 (2.12-2.08)
Measured reflections	210319	544020	463599	217730	115512	206002
Unique reflections	52682	73521	63324	38595	20256	19937
Completeness (%)	99.9 (100)	99.9 (100)	100 (100)	99.3 (99.0)	99.7 (96.7)	98.8 (100)
R _{merge} (%) ^b	10.1 (71.4)	8.0 (82.0)	8.8 (97.5)	7.7 (86.5)	8.1 (80.4)	14.3 (72.0)
I/σ _I	11.8 (2.0)	10.8 (2.8)	11.7 (2.5)	10.7 (2.1)	17.6 (2.0)	11.8 (3.5)
CC [*]	0.971 (0.857)	0.987 (0.947)	0.986 (0.921)	0.982 (0.905)	0.999 (0.745)	0.988 (0.925)
Multiplicity	4.0 (4.0)	7.4 (7.4)	7.3 (6.9)	5.6 (5.6)	5.7 (4.9)	10.3 (10.0)
Wilson B factor (Å ²)	26.8	25.5	23.9	21.7	32.2	36.1
Source	APS 23 ID-B	APS 23 ID-B	APS19-BM	APS 23 ID-B	APS 23 ID-B	APS 19-BM
Processing software	HKL2000	HKL2000	HKL2000	HKL2000	XDS	HKL2000
Refinement						
Resolution (Å)	1.96	1.75	1.84	1.66	2.07	2.08
Monomers in ASU	2	2	2	1	1	1
R _{work} ^c	0.1688	0.1612	0.1582	0.1741	0.1768	0.1711
R _{free} ^d	0.2141	0.1974	0.1979	0.2096	0.2443	0.2289
Ramachandran statistics (%) ^e						
favored	96.18	96.0	96.2	96.3	96.0	96.7
outliers	0.44	0.60	0.30	0.00	0.31	0.30
RMS deviation						
bond lengths (Å)	0.0266	0.0292	0.0267	0.0314	0.0214	0.0246

	E117A OIeA	E117Q OIeA	E117D OIeA	E117A OIeA– Cerulenin cocystal	E117Q OIeA– Cerulenin cocystal	E117D OIeA– Cerulenin cocystal
bond angles (°)	2.410	2.511	2.448	2.757	2.098	2.306
Average B-factor (Å ²)	33.9	31.7	28.9	34.3	44.0	48.5
ligand B-factor (Å ²)	–	–	–	30.0	49.2	67.2
Clashscore (percentile) ^e	3.68 (99 th)	3.58 (98 th)	3.72 (98 th)	8.40 (79 th)	5.59 (97 th)	5.44 (97 th)
Protein Data Bank ID	5VXD	5VXF	5VXH	5VXE	5VXG	5VXI

^aData in parentheses are for the highest resolution shell.

^b $R_{\text{merge}} = \sum_i |I_{\text{hkl}}| / \sum_i I_{\text{hkl}}$, where $|I_{\text{hkl}}|$ is the observed intensity and $\{I_{\text{hkl}}\}$ is the average intensity of multiple measurements.

^c $R_{\text{work}} = \sum_i |F_o| - |F_c| / \sum_i |F_o|$, where $|F_o|$ is the observed structure factor amplitude and $|F_c|$ is the calculated structure factor amplitude.

^d R_{free} is the R factor based on 5% of the data excluded from refinement.

^eBased on values obtained from MolProbity (33).