BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

# Crystal Structures of *Xanthomonas campestris* OleA Reveal Features That Promote Head-to-Head Condensation of Two Long-Chain Fatty Acids

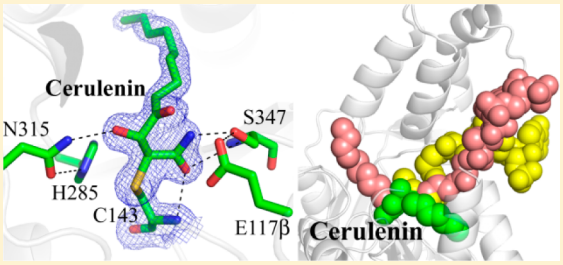
Brandon R. Goblirsch,[†] Janice A. Frias,[‡] Lawrence P. Wackett,[‡] and Carrie M. Wilmot*[,†]

[†]Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, Minneapolis, Minnesota 55455, United States

[‡]Department of Biochemistry, Molecular Biology, and Biophysics and BioTechnology Institute, University of Minnesota, St. Paul, Minnesota 55108, United States

Supporting Information

**ABSTRACT:** OleA is a thiolase superfamily enzyme that has been shown to catalyze the condensation of two long-chain fatty acyl-coenzyme A (CoA) substrates. The enzyme is part of a larger gene cluster responsible for generating long-chain olefin products, a potential biofuel precursor. In thiolase superfamily enzymes, catalysis is achieved via a ping-pong mechanism. The first substrate forms a covalent intermediate with an active site cysteine that is followed by reaction with the second substrate. For OleA, this conjugation proceeds by a nondecarboxylative Claisen condensation. The OleA from *Xanthomonas campestris* has been crystallized and its structure determined, along with inhibitor-bound and xenon-derivatized structures, to improve our understanding of substrate positioning in the context of enzyme turnover. OleA is the first characterized thiolase superfamily member that has two long-chain alkyl substrates that need to be bound simultaneously and therefore uniquely requires an additional alkyl binding channel. The location of the fatty acid biosynthesis inhibitor, cerulenin, that possesses an alkyl chain length in the range of known OleA substrates, in conjunction with a single xenon binding site, leads to the putative assignment of this novel alkyl binding channel. Structural overlays between the OleA homologues, 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) synthase and the fatty acid biosynthesis enzyme FabH, allow assignment of the two remaining channels: one for the thioester-containing pantetheinate arm and the second for the alkyl group of one substrate. A short $\beta$-hairpin region is ordered in only one of the crystal forms, and that may suggest open and closed states relevant for substrate binding. Cys143 is the conserved catalytic cysteine within the superfamily, and the site of alkylation by cerulenin. The alkylated structure suggests that a glutamic acid residue (Glu117$\beta$) likely promotes Claisen condensation by acting as the catalytic base. Unexpectedly, Glu117$\beta$ comes from the other monomer of the physiological dimer.




The impetus to discover microbial pathways capable of producing new, high-energy molecules has become stronger in response to a dwindling fossil fuel supply. Consequently, microbes that are capable of generating compounds such as aliphatic isoprenoid compounds and alkanes from fatty acid deformylation have been isolated.[1,2] The biocatalyzed generation of energy-rich long-chain olefins by microbes across multiple phyla produces hydrocarbons ranging in size from $C_{27}$ to $C_{31}$.[3] Future large-scale production of such compounds would provide a higher-energy biofuel in comparison to current technologies such as the ethanol fermentation process.[4]

The ability of microbes to generate long-chain hydrocarbons has been well documented.[5−9] Consistent among the microbial olefinic hydrocarbon products is a double bond located at the median carbon.[6] This transformation is initiated by a "head-to-head" condensation, so-called because the carboxyl carbon on one fatty acyl group reacts with the $\alpha$-carbon of another fatty acyl group to generate a new carbon−carbon bond.[10] Recent work has demonstrated that a dedicated gene ensemble (*oleABCD* for olefin biosynthesis) is required for the production of olefins formed through head-to-head condensation.[11,12] Sequence homology has identified the *oleBCD* gene products as an $\alpha/\beta$-hydrolase, an AMP-dependent ligase, and a short-chain dehydrogenase, respectively.[3] The *oleA* gene product was predicted to belong to the thiolase superfamily that contains members known to catalyze carbon−carbon bond condensation reactions.

Heterologous *oleA* gene expression and in vitro characterization of purified OleA have confirmed the enzyme is capable of initiating olefin biosynthesis.[3,13] Specifically, the enzyme can condense two coenzyme A (CoA) charged fatty acids to produce a long-chain $\beta$-ketoacid via a nondecarboxylative

**Received:** July 28, 2011
**Revised:** April 23, 2012
**Published:** April 23, 2012





dx.doi.org/10.1021/bi300386m | *Biochemistry* 2012, 51, 4138−4146

Claisen condensation reaction (Figure 1A).[13] The catalytic cycle begins with the transesterification of the first fatty acyl-

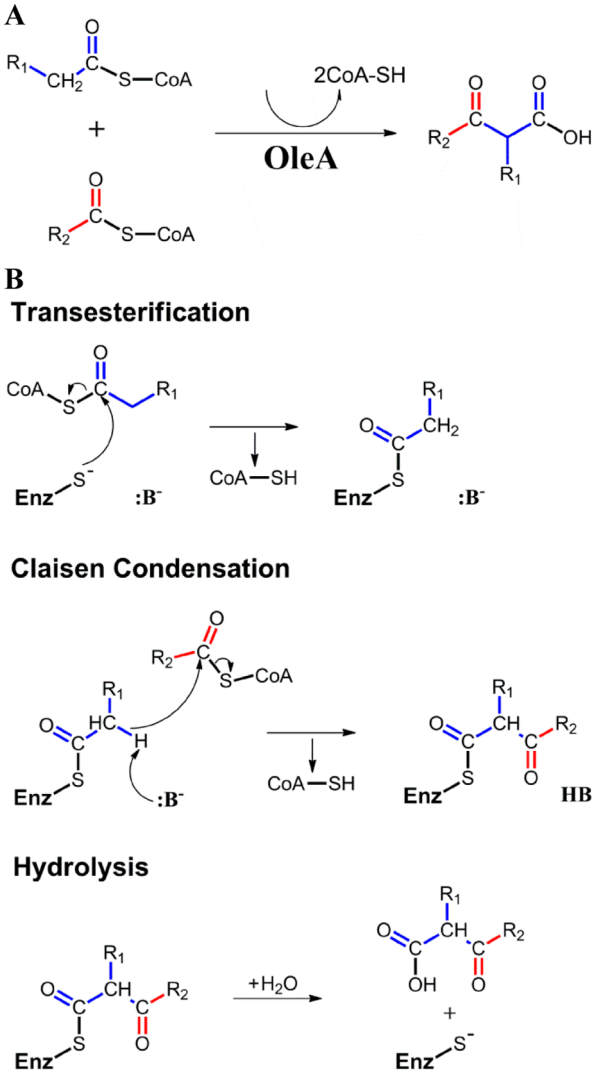



**Figure 1.** OleA-catalyzed condensation with CoA-charged substrates. (A) Overall reaction. (B) Three steps of the catalytic cycle.

CoA group to an active site cysteine (Figure 1B). The second fatty acyl-CoA then binds, and abstraction of a proton from the cysteine-tethered acyl group is thought to generate a $\beta$-carbanion capable of nucleophilic attack on the CoA thioester. The final step is hydrolysis of the cysteine acyl to free the $\beta$-ketoacid product. In vitro studies have shown that the $\beta$-ketoacid product can then be converted to an olefin by incubation with OleC and OleD.[13]

A decarboxylative Claisen condensation mechanism is utilized by other enzymes in the thiolase superfamily, such as the fatty acid biosynthetic (Fab) enzymes. The Fab enzymes condense fatty acid acyl-CoA substrates with malonic acid, which is typically charged with an acyl carrier protein (ACP) rather than CoA.[14] Crystal structures have been determined for members of the Fab enzyme family (FabB, FabP, and FabH), and the homologous 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) synthase that employs a catalytic mechanism similar to that of OleA.[15−21] These enzymes are functional homodimers, with an active site in each monomer containing the proposed reactive cysteine. Structures of FabH (also known as $\beta$-ketoacyl-acyl carrier protein synthase III) and HMG-CoA have been determined in complex with substrates, substrate analogues, and inhibitors that define substrate binding in these enzymes.[18−21] A frequently used inhibitor is the antibiotic cerulenin, a natural product that irreversibly inhibits fatty acid biosynthesis by forming a covalent adduct with the active site cysteine (Figure 2).[22,23] HMG-CoA synthase condenses two

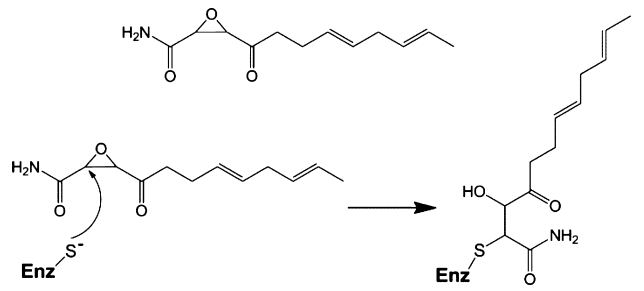



**Figure 2.** Inhibition of OleA through covalent modification of the active site cysteine by cerulenin.

acetyl-CoA molecules to yield acetoacyl-CoA, in a ping-pong mechanism that requires a single substrate binding channel to bring each CoA-activated substrate sequentially to the active site.[20,21] Fab enzymes, in particular FabH from *Mycobacterium tuberculosis* that elongates long-chain fatty acyl substrates ($C_{16}$−$C_{22}$) by two carbons, require a single binding channel for the alkyl chain prior to Claisen condensation.[18,19] However, OleA is unusual in that both its substrates possess long alkyl chains ($C_{10}$−$C_{16}$). In that context, an additional alkyl binding channel is required. In this study, we have determined the structure of OleA alone, in complex with cerulenin, and in complex with cerulenin and xenon to improve our understanding of how this atypical thiolase enzyme may accommodate both long-chain fatty acyl-CoA substrates during turnover.

## EXPERIMENTAL PROCEDURES

**Protein Expression, Purification, and Crystallization.** The expression and purification of *Xanthomonas campestris* OleA were conducted as described previously.[13] Protein stocks were concentrated between 220 and 300 $\mu$M in 500 mM NaCl and 20 mM Hepes (pH 7.4). To covalently bind cerulenin to OleA, the protein stock was reacted with 6 mM cerulenin for 1 h, resulting in the complete inhibition of the enzyme (Figure 2). Initial crystallization trials of ∼1600 conditions for both OleA and cerulenin-treated OleA were conducted at the Hauptman-Woodward Medical Research Institute.[24] Multiple conditions gave crystals for both enzyme forms, with the final condition selected for ease of growth, reproducibility, and diffraction quality. All crystals were grown via hanging drop vapor diffusion at 20 °C. OleA was crystallized in drops containing 1 $\mu$L of protein and 1 $\mu$L of mother liquor [16−21% (w/v) PEG 8000, 70−100 mM potassium phosphate dibasic, and 100 mM sodium citrate (pH 4.2)], while OleA−cerulenin cocrystals were grown in drops containing 1 $\mu$L of protein and 1 $\mu$L of mother liquor [13−18% (w/v) PEG 4000, 80−100 mM manganese chloride, and 100 mM MES (pH 6.0)]. OleA crystals were cryoprotected in a mother liquor solution containing 20% PEG 400, while the OleA−cerulenin cocrystals were cryoprotected in a mother liquor solution containing 25% glycerol, prior to being flash-frozen in liquid nitrogen.


dx.doi.org/10.1021/bi300386m | *Biochemistry* 2012, 51, 4138−4146

**Table 1. Data Collection and Refinement Statistics**[a]

| | OleA unbound | OleA–cerulenin cocrystal | Xe-pressurized OleA–cerulenin cocrystal | Xe-pressurized OleA–cerulenin cocrystal |
|---|---|---|---|---|
| | | Data Collection | | |
| wavelength (Å) | 1.03 | 1.03 | 1.03 | 2.29 |
| space group | $P2_12_12_1$ | $P3_221$ | $P3_221$ | $P3_221$ |
| unit cell (Å) | 82.2 × 85.4 × 102.7 | 90.5 × 90.5 × 69.8 | 90.0 × 90.0 × 69.5 | 90.0 × 90.0 × 69.5 |
| resolution (Å) | 50.0–1.85 (1.88–1.85) | 50.0–1.70 (1.73–1.70) | 50.0–1.95 (1.98–1.95) | 50.0–2.46 (2.50–2.46) |
| no. of measured reflections | 509308 | 270278 | 267054 | 113152 |
| no. of unique reflections | 61685 | 36319 | 24115 | 11836 |
| completeness (%) | 98.8 (97.9) | 99.0 (100) | 99.9 (99.8) | 97.5 (90.5) |
| $R_{merge}$ (%)[b] | 5.5 (38.5) | 4.4 (42.3) | 5.8 (40.8) | 6.7 (30.6) |
| $I/\sigma I$ | 40.2 (5.3) | 44.5 (5.0) | 55 (7.1) | 48.7 (9.8) |
| multiplicity | 8.2 (8.1) | 7.4 (7.4) | 11.1 (11.1) | 8.8 (9.6) |
| Wilson $B$ factor (Å$^2$) | 21.4 | 23.3 | 27.4 | |
| | | Refinement | | |
| resolution (Å) | 44.0–1.85 (1.89–1.85) | 27.6–1.70 (1.75–1.70) | 29.5–1.95 (2.03–1.95) | |
| no. of reflections (working set, test set) | 57593 (2364), 3059 (123) | 34442 (2646), 1852 (130) | 22513 (2344), 1224 (139) | |
| $R_{work}$ (%)[c] | 16.0 (19.4) | 17.8 (22.4) | 18.0 (19.9) | |
| $R_{free}$ (%)[d] | 19.1 (28.6) | 20.9 (25.2) | 21.82 (25.2) | |
| no. of protein atoms | 5244 | 2590 | 2614 | |
| no. of other atoms | 594 | 249 | 146 | |
| Ramachandran statistics (%) | | | | |
| allowed | 100 | 100 | 100 | |
| outliers | 0 | 0 | 0 | |
| rmsd | | | | |
| bond lengths (Å) | 0.007 | 0.007 | 0.008 | |
| bond angles (deg) | 1.038 | 1.013 | 1.027 | |
| average $B$ factor (Å$^2$) | 21.2 | 23.5 | 27.1 | |
| coordinate error (ML-based)[e] | 0.18 | 0.18 | 0.22 | |
| Protein Data Bank entry | 3ROW | 3S21 | 3S23 | |

[a]Data in parentheses are for the highest-resolution shell. [b]$R_{merge} = \sum_i|I_{hkl,i} - \langle I_{hkl}\rangle|/\sum_{hkl}\sum_i I_{hkl,i}$, where $I$ is the observed intensity and $\langle I_{hkl}\rangle$ is the average intensity of multiple measurements. [c]$R_{work} = \sum||F_o| - |F_c||/\sum|F_o|$, 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 keeping the same 5% between isomorphous data sets. [e]Estimated coordinate error based on maximum likelihood estimate in Phenix 1.7-650.[26]

**Xenon Derivitization of OleA–Cerulenin Cocrystals.** OleA–cerulenin crystals were harvested and washed in their respective cryoprotectant solutions. Crystals were then transferred into a pressure cell for derivatization (4DX Systems). Crystals were pressurized with xenon (Airgas Inc.) at ∼130 psi for 15 min. Crystals were then immediately flash-frozen in liquid nitrogen.

**X-ray Data Collection, Processing, and Refinement.** The data sets using 1.03 Å X-rays were collected at GM/CA-CAT beamlines 23-ID-D and 23-ID-B of the Advanced Photon Source (APS) (Argonne National Laboratory, Argonne, IL) at 100 K using a MARmosaic 300 CCD. The beam size was adjusted to match the crystal size and orientation and attenuated by 100–250-fold; 200° of diffraction data were collected on a single crystal, during which time no significant degradation of diffraction quality was observed. The data set at a wavelength of 2.29 Å was collected at the University of Minnesota using a MicroMax-007 chromium rotating anode X-ray generator with VariMaxCr optics and a R-AXIS IV++ detector with a helium cone (Rigaku) at 100 K. Data collection statistics are listed in Table 1.

OleA and OleA–cerulenin cocrystals crystallized in distinct space groups ($P2_12_12_1$ and $P3_221$, respectively). The orthorhombic OleA structure was determined by molecular replacement. The crystal structure of the annotated FabH from *Xanthomonas oryzae* without solvent [Protein Data Bank (PDB) entry 3FK5] was used as a search model (sequence 91% identical to that of *X. campestris* OleA). The *X. oryzae* enzyme has not been biochemically characterized, and the high level of sequence identity with *X. campestris* OleA, coupled with the gene being found within an *ole* gene cluster,[3] suggests that this gene is misannotated and actually encodes an OleA. Molecular replacement with PHASER placed a dimer within the asymmetric unit (AU). For all structures, cycles of model building with COOT[25] and restrained refinement using PHENIX 1.7-650[26] with TLS were performed until all interpretable regions of the $2F_o - F_c$ and $F_o - F_c$ electron density maps were explained (PDB entry 3ROW).

OleA–cerulenin cocrystals were in a different space group ($P3_221$), so the structure was determined by molecular replacement using PHASER and the OleA model without solvent as the search model.[27] The correct enantiomorph, $P3_221$, was confirmed by a strong translation function $Z$ score (TFZ) peak of 30.6 (TFZ, 8.5 for $P3_121$) with a single monomer in the AU. The physiological dimer was generated through the 2-fold crystallographic axis. The final structure was consistent with cerulenin bound at 100% occupancy (PDB entry 3S21). X-ray diffraction data from a xenon-pressured OleA–cerulenin cocrystal were collected at wavelengths of 1.03 and 2.29 Å, the latter leading to a strong anomalous signal due to the xenon ($\lambda$ = 2.29 Å; Xe $f''$ = 11.9 electrons). Difference Fourier was used to collect the 1.03 Å wavelength data, while


dx.doi.org/10.1021/bi300386m | *Biochemistry* 2012, 51, 4138–4146

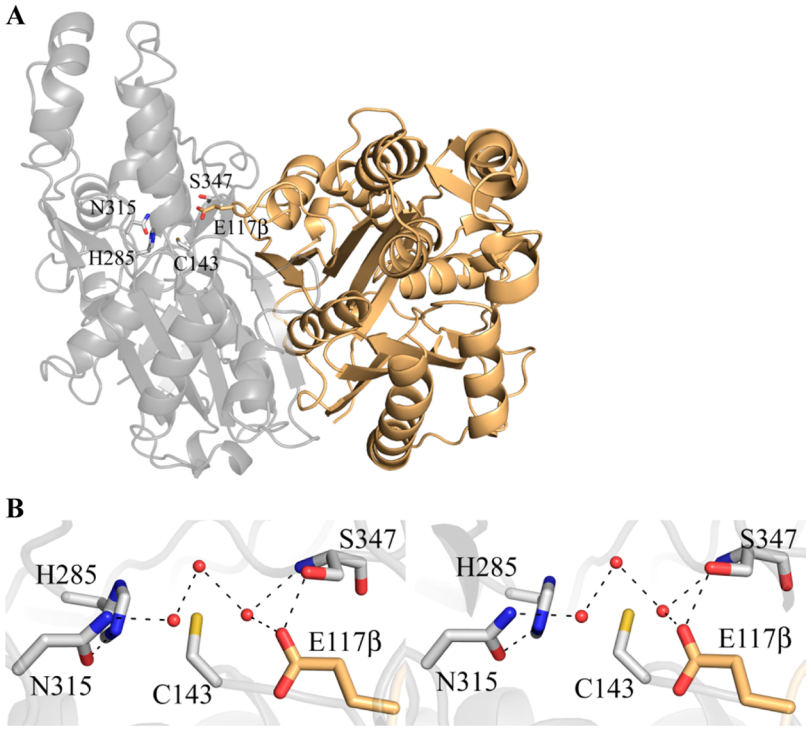



**Figure 3.** OleA dimer and active site. (A) Physiological dimer of OleA. The two monomers are shown as gray and tan cartoons. Each monomer contains one active site. The gray cartoon active site residues are drawn as sticks. Note that E117$\beta$ derives from the neighboring monomer. (B) Stereoview of the OleA active site. The $2F_o - F_c$ electron density map is contoured at 1.0$\sigma$. Ordered solvent molecules are represented as red spheres. This figure was produced using PyMOL (http://www.pymol.org/).

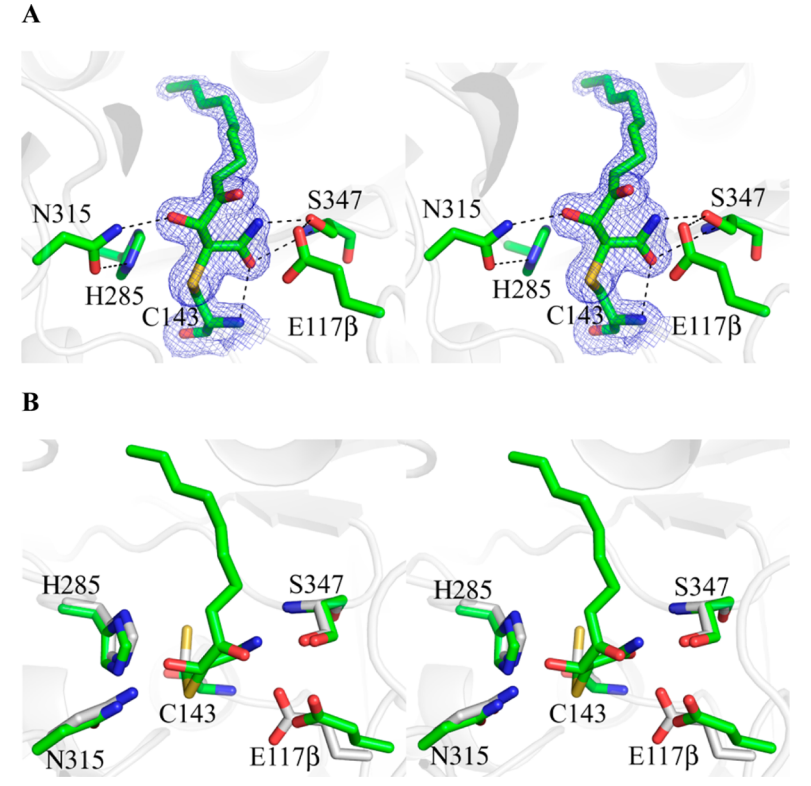



**Figure 4.** Stereoview of the OleA active site with cerulenin bound to Cys143. Active site residues and the inhibitor are drawn as sticks. Solvent molecules are represented as red spheres. Hydrogen bond contacts are represented by dashed lines. Blue mesh illustrates the $2F_o - F_c$ electron density maps contoured at 1.0$\sigma$. (A) Active site of OleA cocrystallized with cerulenin (green carbons). (B) Overlay of OleA active sites in the unbound (gray carbons) and cerulenin-bound states. The overlay was generated in PyMOL using SUPER. This figure was produced using PyMOL (http://www.pymol.org/).


dx.doi.org/10.1021/bi300386m | *Biochemistry* 2012, 51, 4138−4146

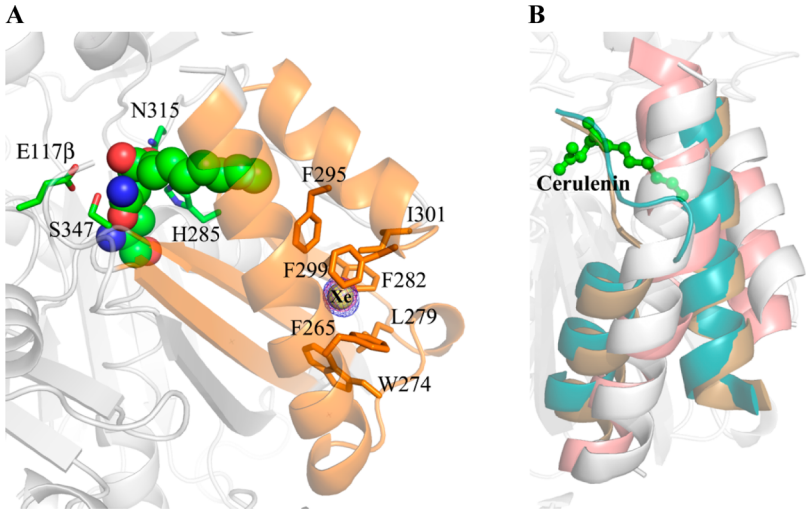



**Figure 5.** Proposed OleA channel A. (A) The positions of the cerulenin alkyl chain and xenon (Xe) atom are consistent with a continuous alkyl binding channel. The cerulenin bound to Cys143 and the xenon are drawn in space-filling mode and colored by atom (green carbons and pale yellow xenons). Electron density for the Xe is included ($2F_o - F_c$ map contoured at 1.0$\sigma$, blue mesh; anomalous contoured at 9.0$\sigma$, pink mesh). The OleA secondary structural elements that form the channel and contact the cerulenin and Xe are represented as orange cartoons. The bulky hydrophobic side chains that line the most distal part of the channel from Cys143 are shown as orange sticks. (B) Two helices that line channel A of OleA (gray cartoon) are more closely packed in FabF (PDB entry 1B3N, cyan carbons), FabB (PDB entry 1FJ8, brown carbons), and FabH (PDB entry 2QX1, pink carbons). This figure was produced using PyMOL (http://www.pymol.org/).

the PHENIX maps program suite was used to generate anomalous peak maps using the 2.29 Å wavelength data.[26] Both used the OleA–cerulenin cocrystal structure (without solvent) for phasing. Two xenon sites were identified (occupancies of 45 and 35%) (PDB entry 3S23). Final refinement statistics for all structures are listed in Table 1.

## RESULTS

**Overall Structures of the Unbound and Inhibited OleA Dimer.** The crystal structures of unbound OleA and OleA–cerulenin cocrystals were determined to 1.85 and 1.70 Å resolution, respectively (Table 1). The unbound structure of OleA is a dimer consistent with the oligomeric state of most characterized thiolase superfamily members, including FabH and HMG-CoA synthase (Figure 3A).[28] The monomer consists of two domains, with the dimer interface (2135 Å$^2$) being formed between a single equivalent domain from each monomer.[29] The OleA–cerulenin structure contains a single monomer in the AU with a crystallographic 2-fold axis running directly through the dimer. Nevertheless, the gross structures of the three OleA forms are similar, with the largest overall rmsd between any pair being ≤0.7 Å across all matching atoms (Figure S1 of the Supporting Information).

An 11-residue $\beta$-hairpin (residues 239–249) is ordered in only a single monomer of the $P2_12_12_1$ crystal form and is disordered in the $P3_221$ crystal form. The $\beta$-hairpin contains hydrophobic residues (Leu243, Met246, and Val247) that pack against OleA when ordered.

**Active Site of OleA.** The active site of each monomer contains a single catalytic cysteine residue (Cys143) resting at the base of a solvent accessible channel (Figure 3B). A novel feature of OleA is a glutamate (Glu117$\beta$) in the active site that originates from a loop in the neighboring monomer. Immediately above Cys143 are two pockets positioned to stabilize tetrahedral, oxyanionic intermediates generated during turnover. The first is formed by the side chains of His285 and Asn315 (oxyanion hole 1) and the second by the main-chain amides of Ser347 and the catalytic Cys143 (oxyanion hole 2). Ordered waters interlink these residues, and a hydrogen bond between the carboxylate side chain of Glu117$\beta$ and the hydroxyl of Ser347 (Figure 3B) may be important for positioning Glu117$\beta$ within the active site.

**Active Site Changes Induced by Reaction with Cerulenin.** Cerulenin treatment leads to the formation of a covalent adduct with Cys143 of OleA (Figures 2 and 4A). The entire cerulenin molecule is well resolved with clear positioning of the amide and hydroxyl moieties, and is at 100% occupancy. The amide of the bound cerulenin forms hydrogen bonds with the backbone amide nitrogen of Cys143 and Ser347, and the hydroxyl of Ser347. The adduct displaces the water network and causes movement and increased mobility in the side chain of Glu117$\beta$, which retains only a single hydrogen bond to the hydroxyl of Ser347 (Figure 4A). The cerulenin hydroxyl is hydrogen bonded to Asn315 of oxyanion hole 1. To accommodate the cerulenin, Cys143 rotates approximately 120° from its position in the unbound form of the enzyme (Figure 4B). The imidazole ring of His285 in the unbound and cerulenin-complexed structures is rotated by 180° in $\chi_2$ (flipped) compared to the common conformer observed in other thiolase superfamily members and hydrogen bonds directly to Asn315 of oxyanion hole 1.

**A New Long-Chain Alkyl Binding Pocket Defined in OleA.** The cerulenin complex experimentally identifies an undescribed alkyl channel that can be accessed in OleA. When compared to the substrates of *X. campestris* OleA ($C_{10}$, $C_{12}$, $C_{14}$, $C_{16}$, and $C_{16:1}$), cerulenin is equivalent to a $C_{11:2}$ fatty acid [the two double bonds are *trans* (Figure 2)]. To seek further evidence, we pressurized an OleA–cerulenin crystal with xenon gas to identify hydrophobic sites that might further delineate this new channel and suggest compatibility with longer *X. campestris* OleA substrates. The $2F_o - F_c$ maps indicate the presence of an electronically dense atom at the base of the channel originating from the alkyl chain of Cys143-tethered cerulenin (Figure 5A). The xenon site did not refine to full


dx.doi.org/10.1021/bi300386m | *Biochemistry* 2012, 51, 4138–4146

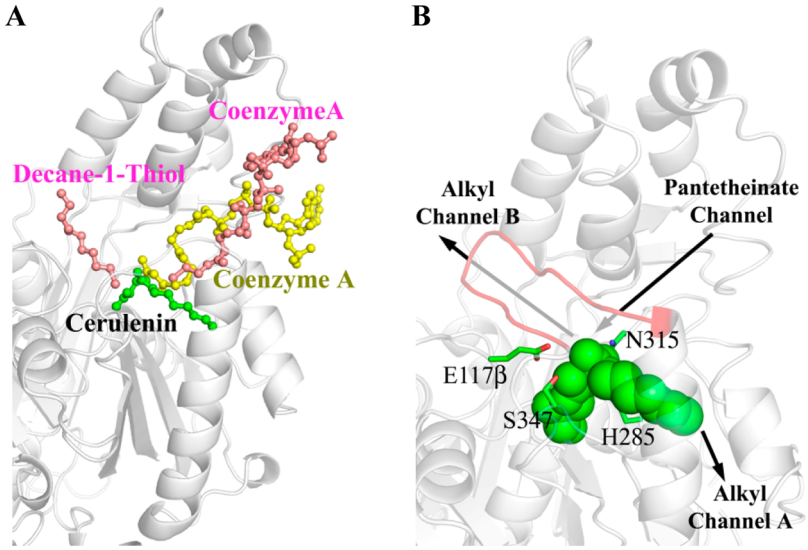



**Figure 6.** Illustration of OleA binding channels. (A) OleA bound with cerulenin overlaid with FabH (bound with decane-1-thiol and coenzyme A, PDB entry 2QX1) and HMG-CoA synthase (bound with coenzyme A, PDB entry 1TXT). OleA is drawn as a gray cartoon and cerulenin as a green ball-and-stick model. For the sake of clarity, the monomers of FabH and HMG-CoA synthase have been omitted and their bound ligands are drawn as balls and sticks (pink and yellow, respectively). Decane-1-thiol occupies alkyl channel B in FabH. Coenzyme A occupies the pantetheinate channel in both FabH and HMG-CoA synthase. (B) Putative assignment of OleA binding channels. Active site residues are drawn as sticks (green carbons), and bound cerulenin is drawn in green space-filling mode. Labeled arrows indicate the position of the alkyl and pantetheinate channels in the OleA monomer (gray cartoon). An 11-residue $\beta$-hairpin (residues 239−249) that could be modeled in only one monomer of the $P2_12_12_1$ crystal form is drawn as a red cartoon. This figure was produced using PyMOL (http://www.pymol.org/).

**Table 2. Structures Used in Overlays with OleA**

| organism | enzyme | PDB entry | rmsd[a] from OleA (Å) | ligands | channels bound |
|---|---|---|---|---|---|
| *M. tuberculosis* | FabH | 2QX1[18] | 2.01 | decane-1-thiol, CoA-SH | alkyl channel B, pantetheinate channel |
| *Staphylococcus aureus* | HMG-CoA synthase | 1TXT[20] | 2.62 | CoA-SH | pantetheinate channel |
| *E. coli* | FabB | 1FJ8[17] | 3.93 | cerulenin | alkyl channel B |
| *E. coli* | FabF | 1B3N[16] | 4.04 | lauric acid | alkyl channel B |

[a]Root-mean-square-deviation calculated over all matched atoms between structures by SUPER in PyMOL (http://www.pymol.org/).

occupancy (45%), but anomalous maps collected on a Cr anode X-ray source allowed for unambiguous assignment. Taken together, the position of the cerulenin $C_{11:2}$ alkyl chain (whose length is in the range of those of *X. campestris* OleA substrates) and this xenon site suggest a continuous channel beginning at the enzyme active site and terminating at a group of bulky hydrophobic residues (Figure 5A). Residues 252−272 and 289−301 are a pair of $\alpha$-helices that form the ceiling of the channel, while a $\beta$-sheet consisting of residues 280−284, 339−346, and 349−355 forms the base. The interior of the channel contains hydrophobic side chains creating a favorable environment for binding of a long alkyl chain. The channel (designated channel A) extends beyond the length of the cerulenin alkyl chain to the xenon site and is compatible with the $C_{10}$−$C_{16}$ fatty acyl substrates of OleA.[13]

## DISCUSSION

OleA initiates long-chain olefin biosynthesis by condensing two fatty acyl-CoA substrates at a cysteine-centered active site (Figures 1 and 3). The pantetheinate channel that accommodates CoA- or ACP-charged substrates has been well characterized and structurally visualized in the thiolase superfamily enzymes FabF, FabB, FabH, and HMG-CoA synthase, and the channel is conserved in OleA (Figure 6A and Table 2).[16,17,19,21] However, OleA is distinct in requiring two channels to simultaneously accommodate two long alkyl chains. HMG-CoA synthase substrates do not have long alkyl chains, while members of the Fab enzyme family need to accommodate a single long alkyl chain.[18,19] Hydrophobic channel A defined by cerulenin and xenon in the OleA structure has not been previously described (Figure 5A). In vitro data of *X. campestris* OleA demonstrate the enzyme can use $C_{10}$, $C_{12}$, $C_{14}$, $C_{16}$, and $C_{16:1}$ substrates.[13] Modeling an extended alkyl chain from the cerulenin is consistent with channel A being able to accommodate $C_{16}$ alkyl chains, although the exact length tolerance remains unclear.

While the alkyl chain of cerulenin lies in OleA channel A, the crystal structure of the *Escherichia coli* FabB−cerulenin complex shows the cerulenin alkyl chain bound in the opposing direction (Figure S2 of the Supporting Information and Table 2).[17] This channel (known as B) has been shown to bind alkyl chains in other Fab enzymes and thus appears to be the preferential binding site for the long-chain alkyl substrate utilized in fatty acid biosynthesis (Figure 6A and Table 2). The two helices of OleA channel A are spread apart when compared to the closely packed helices observed in the Fab enzymes (Figure 5B), consistent with the idea that alkyl chains can more easily thread between the $\beta$-sheet and two helices that define channel A.

Channel B in OleA appears to be very similar to channel B in the Fab enymes. A $\beta$-hairpin (residues 239−249) that defines the ceiling of channel B is ordered in only one of the monomers


dx.doi.org/10.1021/bi300386m | *Biochemistry* 2012, 51, 4138−4146

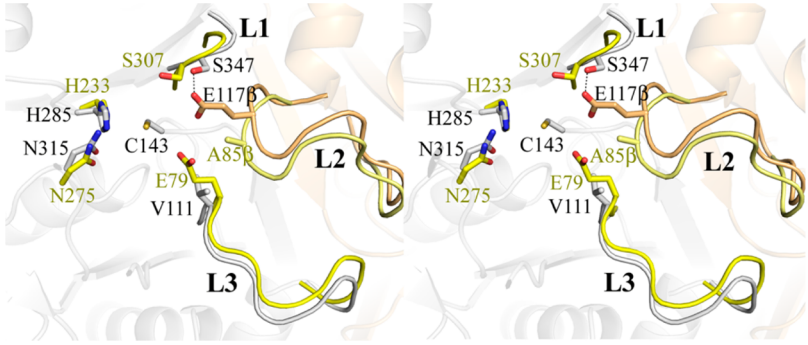



**Figure 7.** Overlay of HMG-CoA synthase (PDB entry 1TXT) onto OleA. The dimer of OleA is shown as a faded cartoon colored as in Figure 3. Three loop elements (L1−L3) conserved between OleA and HMG-CoA synthase are highlighted, with L2 originating from the neighboring monomer. Residues forming a single active site in each enzyme are drawn as sticks. For the sake of simplicity, a single active site cysteine (C143) from OleA is drawn. Residues from OleA are colored by monomer and labeled in black. E117$\beta$ (tan carbons) derives from the L2 loop element making hydrogen bonding contacts with S347 found in the L1 loop. Spatially equivalent residues from HMG-CoA are labeled in yellow and colored by monomer. The active site of HMG-CoA synthase is formed completely from one monomer (yellow carbons), including the E79 catalytic base that derives from the L3 loop. In HMG-CoA synthase, the position of E117$\beta$ aligns with that of A85$\beta$ (pale yellow carbons). The overlay was generated in PyMOL using SUPER. This figure was produced using PyMOL (http://www.pymol.org/).

within the $P2_12_12_1$ crystal form and disordered in the OleA−cerulenin structure. In *M. tuberculosis* FabH, this region is similarly disordered and ordered in different crystal structures and has been proposed to represent an open state allowing easy access for the long alkyl chain of a CoA-charged substrate and a closed state with an ordered $\beta$-hairpin optimized for productive catalytic turnover.[19] As cerulenin binds in the newly defined channel A, this is likely the site where the OleA first substrate binds. The OleA second long-chain fatty acyl-CoA substrate would then bind in channel B. This would favor ordering of the $\beta$-hairpin to attain the closed state, as suggested for FabH,[18] promoting the Claisen condensation and hydrolysis (Figure 1B). This model is also consistent with the observed release of fatty acids as aborted reaction products in OleA turnover.[13] Acylated Cys143 in the open conformation is likely more vulnerable to premature hydrolysis within the time frame required to bind the second fatty acyl-CoA substrate.

The OleA active site reveals a spatial arrangement of chemical groups conserved within the thiolase superfamily (Figure 3). Oxyanion holes are formed on either side of the centralized Cys143, and the active site is fully occupied by solvent (Figure 3B). In OleA, cerulenin binding reveals hydrogen bonding partners in each pocket that are likely important during turnover (Figure 4A). In HMG-CoA synthase, biochemical and structural studies have identified the histidine as the general base that deprotonates the cysteine to initiate the nucleophilic attack on the first substrate that leads to acetylation.[21,30] By analogy, His285 likely acts as the general base in this part of the OleA reaction.

The subsequent activation of the acetylated Cys to initiate condensation with the second substrate requires an additional base that in biosynthetic thiolases is another Cys and in HMG-CoA synthase is a Glu.[31,32] OleA does not have a second Cys in the active site but does contain a glutamate (Glu117$\beta$). The residue is in a loop element (L2) from the neighboring monomer, and this is the first superfamily member in which both monomers contribute to the active site. The glutamic acid base of HMG-CoA synthase (Glu79) originates from loop element L3 within the same monomer that contributes the other active site residues (Figure 7). In OleA, the sequential and structural equivalent of Glu79 of HMG-CoA synthase is Val111, whereas the equivalent of Glu117$\beta$ of OleA in HMG-CoA synthase is an alanine [Ala85$\beta$ (Figure 7)]. In the cerulenin complex, Glu117$\beta$ is relatively disordered and is not in contact with the cerulenin, being ∼4 Å distant at closest approach. Having Glu117$\beta$ in a loop that is not structurally part of the core protein fold containing the active site may provide a more flexible environment in which to spatially coordinate the chemistry between the two long-chain fatty acid acyl-CoA substrates during OleA turnover. Such plasticity may allow for the binding of large substrates while permitting the reactive acyl-thioester moieties of CoA-charged fatty acids to be correctly positioned within the active site.

Interestingly, most organisms with the *oleABCD* operon produce a single-length olefin product, while in *X. campestris*, multiple length olefins are produced.[3] The chain length appears primarily due to the action of OleA, as genetic experiments that switched *oleA* genes between species took on the olefin profile of the recombinant OleA protein.[3] The substrate promiscuity of *X. campestris* OleA is an unusual trait within the thiolase superfamily. Differences in the alkyl channels of OleA enzymes may explain the diverse product length profiles among different OleA proteins.

## CONCLUSIONS

Our results support the presence of three major substrate binding channels in OleA that promote the condensation of the two long-chain fatty acyl-CoA substrates. The location of the fatty acid biosynthesis inhibitor cerulenin, in conjunction with a xenon binding site, strongly supports channel A being the additional alkyl binding channel required for OleA catalysis. A glutamic acid residue (Glu117$\beta$) from the neighboring monomer likely promotes Claisen condensation by acting as the catalytic base. Future mutagenesis studies of Glu117$\beta$ and His285 should test their catalytic roles as bases in OleA turnover.

## ASSOCIATED CONTENT

### Supporting Information

$\alpha$-Carbon overlay of all reported OleA structures (Figure S1) and overlay of the OleA−cerulenin structure with the *E. coli* FabB−cerulenin structure (Figure S2). This material is available free of charge via the Internet at http://pubs.acs.org.


dx.doi.org/10.1021/bi300386m | *Biochemistry* 2012, 51, 4138−4146

### Accession Codes

Coordinates and structure factors have been deposited in the Protein Data Bank as entries 3ROW (OleA unbound), 3S21 (OleA−cerulenin cocrystal), and 3S23 (Xe-pressurized OleA−cerulenin cocrystal).

## AUTHOR INFORMATION


### Corresponding Author

*E-mail wilmo004@umn.edu. Phone: (612) 624-2406. Fax: (612) 624-5121.



### Funding

This work was supported by National Institutes of Health Grant GM-90260 (C.M.W.) and Chemistry Biology Interface Training Grant GM-008700 (J.A.F.), the Department of Energy ARPA-E via Grant DE-AR0000007 (L.P.W.), and University of Minnesota Doctoral Dissertation Fellowships (J.A.F. and B.R.G.).


### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS


We thank Jasmine Erickson for assistance with expression and purification of OleA used in these studies. We thank staff at sector 23, GM/CA-CAT, of the Advanced Photon Source for assistance. GM/CA-CAT is funded by National Cancer Institute Grant Y1-CO-1020 and National Institute of General Medical Sciences Grant Y1-GM-1104. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract DE-AC02-06CH11357. Computer resources were provided by the Basic Sciences Computing Laboratory of the University of Minnesota Supercomputing Institute, and we thank Can Ergenekan for support. We also thank Ed Hoeffner at the Kahlert Structural Biology Laboratory (KSBL) at the University of Minnesota.


## ABBREVIATIONS

CoA, coenzyme A; Ole, olefin biosynthesis enzyme; Fab, fatty acid biosynthesis enzyme; ACP, acyl carrier protein; HMG-CoA synthase, 3-hydroxy-3-methylglutaryl-CoA synthase; PDB, Protein Data Bank; AU, asymmetric unit; TFZ, translation function *Z* score; rmsd, root-mean-square deviation.

## REFERENCES


(1) Li, M., Nørgaard, H., Warui, D. M., Booker, S. J., Krebs, C., and Bollinger, J. M., Jr. (2011) Conversion of fatty aldehydes to alka(e)nes and formate by a cyanobacterial aldehyde decarbonylase: Cryptic redox by an unusual dimetal oxygenase. *J. Am. Chem. Soc. 133*, 6158−6161.

(2) Lee, S. K., Chou, H., Ham, T. S., Lee, T. S., and Keasling, J. D. (2008) Metabolic engineering of microorganisms for biofuels production: From bugs to synthetic biology to fuels. *Curr. Opin. Biotechnol. 19*, 556−563.

(3) Sukovich, D. J., Seffernick, J. L., Richman, J. E., Gralnick, J. A., and Wackett, L. P. (2010) Widespread head-to-head hydrocarbon biosynthesis in bacteria and role of OleA. *Appl. Environ. Microbiol. 76*, 3850−3862.

(4) Wackett, L. P. (2008) Microbial-based motor fuels: Science and technology. *Microb. Biotechnol. 1*, 211−225.

(5) Albro, P. W., and Dittmer, J. C. (1969) The biochemistry of long-chain, nonisoprenoid hydrocarbons. IV. Characteristics of synthesis by a cell-free preparation of *Sarcina lutea*. *Biochemistry 8*, 3317−3324.

(6) Albro, P. W., and Dittmer, J. C. (1969) The biochemistry of long-chain, nonisoprenoid hydrocarbons. 3. The metabolic relationship of long-chain fatty acids and hydrocarbons and other aspects of hydrocarbon metabolism in *Sarcina lutea*. *Biochemistry 8*, 1913−1918.

(7) Albro, P. W., and Dittmer, J. C. (1969) The biochemistry of long-chain, nonisoprenoid hydrocarbonss. II. The incorporation of acetate and the aliphatic chains of isoleucine and valine into fatty acids and hydrocarbon by *Sarcina lutea* in vivo. *Biochemistry 8*, 953−959.

(8) Albro, P. W., and Dittmer, J. C. (1969) The biochemistry of long-chain, nonisoprenoid hydrocarbons. I. Characterization of the hydrocarbons of *Sarcina lutea* and the isolation of possible intermediates of biosynthesis. *Biochemistry 8*, 394−404.

(9) Tornabene, T. G., and Oro, J. (1967) 14-C incorporation into the fatty acids and aliphatic hydrocarbons of *Sarcina lutea*. *J. Bacteriol. 94*, 349−358.

(10) Albro, P. W., Meehan, T. D., and Dittmer, J. C. (1970) Intermediate steps in the incorporation of fatty acids into long-chain, nonisoprenoid hydrocarbons by lysates of *Sarcina lutea*. *Biochemistry 9*, 1893−1898.

(11) Beller, H. R., Goh, E. B., and Keasling, J. D. (2010) Genes involved in long-chain alkene biosynthesis in *Micrococcus luteus*. *Appl. Environ. Microbiol. 76*, 1212−1223.

(12) Sukovich, D. J., Seffernick, J. L., Richman, J. E., Hunt, K. A., Gralnick, J. A., and Wackett, L. P. (2010) Structure, function, and insights into the biosynthesis of a head-to-head hydrocarbon in *Shewanella oneidensis* strain MR-1. *Appl. Environ. Microbiol. 76*, 3842−3849.

(13) Frias, J. A., Richman, J. E., Erickson, J. S., and Wackett, L. P. (2011) Purification and Characterization of OleA from *Xanthomonas campestris* and Demonstration of a Non-decarboxylative Claisen Condensation Reaction. *J. Biol. Chem. 286*, 10930−10938.

(14) Heath, R. J., and Rock, C. O. (2002) The Claisen condensation in biology. *Nat. Prod. Rep. 19*, 581−596.

(15) Miziorko, H. M., and Lane, M. D. (1977) 3-Hydroxy-3-methylglutaryl-CoA synthase. Participation of acetyl-S-enzyme and enzyme-S-hydroxymethylglutaryl-SCoA intermediates in the reaction. *J. Biol. Chem. 252*, 1414−1420.

(16) Moche, M., Schneider, G., Edwards, P., Dehesh, K., and Lindqvist, Y. (1999) Structure of the complex between the antibiotic cerulenin and its target, $\beta$-ketoacyl-acyl carrier protein synthase. *J. Biol. Chem. 274*, 6031−6034.

(17) Price, A. C., Choi, K. H., Heath, R. J., Li, Z., White, S. W., and Rock, C. O. (2001) Inhibition of $\beta$-ketoacyl-acyl carrier protein synthases by thiolactomycin and cerulenin. Structure and mechanism. *J. Biol. Chem. 276*, 6551−6559.

(18) Sachdeva, S., Musayev, F., Alhamadsheh, M. M., Neel Scarsdale, J., Tonie Wright, H., and Reynolds, K. A. (2008) Probing reactivity and substrate specificity of both subunits of the dimeric *Mycobacterium tuberculosis* FabH using alkyl-CoA disulfide inhibitors and acyl-CoA substrates. *Bioorg. Chem. 36*, 85−90.

(19) Sachdeva, S., Musayev, F. N., Alhamadsheh, M. M., Scarsdale, J. N., Wright, H. T., and Reynolds, K. A. (2008) Separate entrance and exit portals for ligand traffic in *Mycobacterium tuberculosis* FabH. *Chem. Biol. 15*, 402−412.

(20) Campobasso, N., Patel, M., Wilding, I. E., Kallender, H., Rosenberg, M., and Gwynn, M. N. (2004) *Staphylococcus aureus* 3-hydroxy-3-methylglutaryl-CoA synthase: Crystal structure and mechanism. *J. Biol. Chem. 279*, 44883−44888.

(21) Theisen, M. J., Misra, I., Saadat, D., Campobasso, N., Miziorko, H. M., and Harrison, D. H. (2004) 3-Hydroxy-3-methylglutaryl-CoA synthase intermediate complex observed in "real-time". *Proc. Natl. Acad. Sci. U.S.A. 101*, 16442−16447.

(22) Omura, S. (1976) The antibiotic cerulenin, a novel tool for biochemistry as an inhibitor of fatty acid synthesis. *Bacteriol. Rev. 40*, 681−697.

(23) Ronnett, G. V., Kleman, A. M., Kim, E. K., Landree, L. E., and Tu, Y. (2006) Fatty acid metabolism, the central nervous system, and feeding. *Obesity 14* (Suppl. 5), 201S−207S.

(24) Luft, J. R., Collins, R. J., Fehrman, N. A., Lauricella, A. M., Veatch, C. K., and DeTitta, G. T. (2003) A deliberate approach to

screening for initial crystallization conditions of biological macromolecules. *J. Struct. Biol. 142*, 170−179.
(25) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D66*, 486−501.
(26) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D66*, 213−221.
(27) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658−674.
(28) Haapalainen, A. M., Merilainen, G., and Wierenga, R. K. (2006) The thiolase superfamily: Condensing enzymes with diverse reaction specificities. *Trends Biochem. Sci. 31*, 64−71.
(29) Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol. 372*, 774−797.
(30) Misra, I., and Miziorko, H. M. (1996) Evidence for the interaction of avian 3-hydroxy-3-methylglutaryl-CoA synthase histidine 264 with acetoacetyl-CoA. *Biochemistry 35*, 9610−9616.
(31) Chun, K. Y., Vinarov, D. A., Zajicek, J., and Miziorko, H. M. (2000) 3-Hydroxy-3-methylglutaryl-CoA synthase. A role for glutamate 95 in general acid/base catalysis of C-C bond formation. *J. Biol. Chem. 275*, 17946−17953.
(32) Kursula, P., Ojala, J., Lambeir, A. M., and Wierenga, R. K. (2002) The catalytic cycle of biosynthetic thiolase: A conformational journey of an acetyl group through four binding modes and two oxyanion holes. *Biochemistry 41*, 15543−15556.
(33) Chen, V. B., Arendall, W. B., III, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D66*, 12−21.