Title

Structural Mechanism of Regioselectivity in an Unusual Bacterial Acyl-CoA Dehydrogenase



Journal

Journal of the American Chemical Society, 142(2)

ISSN

0002-7863

Authors

Blake-Hedges, Jacquelyn M

Pereira, Jose Henrique

Cruz-Morales, Pablo

et al.

Publication Date

2020-01-15

DOI

10.1021/jacs.9b09187

Peer reviewed

# Structural mechanism of regioselectivity in an unusual bacterial acyl-CoA dehydrogenase

Authors:
Jacquelyn M. Blake-Hedges,[1,2,3] Jose Henrique Pereira,[2,4] Pablo Cruz-Morales,[2,3] Mitchell G. Thompson,[2,3,5] Jesus F. Barajas,[2,3,6] Jeffrey Chen,[1] Rohith N. Krishna,[1] Leanne Jade G. Chan,[2,3] Danika Nimlos,[1] Catalina Alonso-Martinez,[1] Edward E.K. Baidoo,[2,3] Yan Chen,[2,3,6] Jennifer W. Gin,[2,3,6] Leonard Katz,[2,7] Christopher J. Petzold,[2,3,6] Paul D. Adams,[2,3,4] Jay D. Keasling[2,3,7,8,9,10]

[1]Department of Chemistry, University of California, Berkeley, Berkeley, CA 94720
[2]Joint BioEnergy Institute, Emeryville, CA, 94608
[3]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
[4]Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
[5]Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, CA 94720
[6]Department of Energy Agile BioFoundry, Emeryville, CA 94608, USA
[7]QB3 Institute, University of California, Berkeley, Emeryville, CA, 94608
[8]Department of Chemical & Biomolecular Engineering, Department of Bioengineering, University of California, Berkeley, Berkeley, CA 94720
[9]Novo Nordisk Foundation Center for Biosustainability, Technical University Denmark, DK2970-Horsholm, Denmark
[10]Center for Synthetic Biochemistry, Shenzhen Institutes for Advanced Technologies, Shenzhen, China

## Abstract

Terminal alkenes are easily derivatized, making them desirable functional group targets for polyketide synthase (PKS) engineering. However, they are rarely encountered in natural PKS systems. One mechanism for terminal alkene formation in PKSs is through the activity of an acyl-CoA dehydrogenase (ACAD). Herein, we use biochemical and structural analysis to understand the mechanism of terminal alkene formation catalyzed by an γ,δ-ACAD from the biosynthesis of the polyketide natural product FK506, TcsD. While TcsD is homologous to canonical α,β-ACADs, it acts regioselectively at the γ,δ-position and only on α,β-unsaturated substrates. Furthermore, this regioselectivity is controlled by a combination of bulky residues in the active site and a lateral shift in the positioning of the FAD cofactor within the enzyme. Substrate modeling suggests that TcsD utilizes a novel set of hydrogen bond donors for substrate activation and positioning, preventing dehydrogenation at the α,β position of substrates. From the structural and biochemical characterization of TcsD, key residues that contribute to regioselectivity and are unique to the protein family were determined and used to identify other putative γ,δ-ACADs that belong to diverse natural product biosynthetic gene clusters. These predictions are supported by the demonstration that a phylogenetically distant homolog of TcsD also regioselectively oxidizes α,β-

unsaturated substrates. This work exemplifies a powerful approach to understand unique enzymatic reactions and will facilitate future enzyme discovery, inform enzyme engineering, and aid natural product characterization efforts.

## Introduction

Natural products often have complex chemical structures which can confer potent biological activity. Evolution selects for the diversification of these secondary metabolites, making natural product biosynthetic pathways rich resources for the discovery of both lead compounds for drug discovery and also enzymes with unique functions.[1–3] Next-generation sequencing has led to the identification of numerous biosynthetic gene clusters (BGCs), but the pool of "orphan" BGCs (i.e. BGCs with no cognate natural product identified) remains largely untapped. Predicting natural product structures from primary DNA sequence is challenging, as the *in silico* functional annotation of enzymes within BGCs is limited. Amino acid sequence homology can suggest a general function, but without in depth structural and biochemical characterization of one or more members of an enzyme family, precise predictions of the final natural product structure are difficult. The identification of specificity-conferring motifs in polyketide synthase (PKS) acyltransferase (AT) and ketoreductase (KR) domains, for example, has allowed for more precise predictions of final polyketide natural product structure, including the identity of the alkyl substituents incorporated by (AT) and final stereochemical outcome (KR) of a given PKS module.[4] While much effort has been dedicated to elucidating signature motifs within PKS domains, many PKS-associated enzymes that generate less common functional groups (such as non-canonical starter and extender units) are not as well annotated or understood. Better characterization of the enzymes implicated in the biosynthesis of unique and reactive moieties would facilitate the engineering of novel polyketides with applications in medicine, in agriculture, or as commodity chemicals.[1,5,6]

One particular moiety of interest for PKS engineering is the alkene, as alkenes could easily be chemically derivatized to introduce a multitude of other desirable functionalities into a natural product.[7–19] Alkenes are often generated within a polyketide via the action of reductive domains, but they are sequestered within the polyketide backbone[20] and are therefore less sterically accessible for chemical modification than terminal alkenes.[8,10-16,21-29] Terminal alkenes, in addition to their innate reactivity, can also confer improved biological activity to polyketides that display biological activity, as is exemplified by the improved drug tolerability and efficacy of a synthetic epothilone analog sagopilone (anti-cancer activity).[30,31] There are few known examples of polyketides that contain terminal alkenes, including FK506, haliangicin, curacin A, and tautomycetin (Figure 1A).[32-33,34] In curacin A, a terminal alkene is generated

through a unique off-loading mechanism involving a sulfotransferase (ST) and thioesterase domain [25]. The tautomycetin terminal alkene is known to be formed after the chain has been released from the PKS but the process has yet to be fully characterized.[36,37]

A different mechanism for terminal alkene formation occurs via the action of an acyl-CoA dehydrogenase (ACAD) that oxidizes the γ,δ-position of a fatty acyl-CoA or acyl carrier protein (ACP), as observed in the haliangicin and FK506 pathways.[38,39] The identification of other terminal alkene-forming ACADs is difficult, though, because of the enzymes' homology to canonical ACADs. Thus, identifying important sequence motifs is critical to more accurate annotation. ACADs are oxidoreductase flavoenzymes well known for catalyzing the first oxidative step of fatty acid β-oxidation: the dehydrogenation of saturated fatty acyl-Coenzyme A (CoA) thioesters to form the corresponding α,β-unsaturated product (Figure 1B).[40] The ACADs from the haliangicin and FK506 biosynthetic pathways, however, oxidize the γ,δ-position of a substrate. In order to facilitate better annotation of γ,δ-ACADs and, in turn, the identification of polyketide natural products that potentially contain terminal alkenes, a better functional characterization of γ,δ-ACADs is necessary. Herein, we report on our studies of the basis for the shift in regiochemistry of one of these unusual ACADs, TcsD, which forms the terminal alkene of the allylmalonyl-CoA extender unit in the biosynthesis of the polyketide FK506.[39]

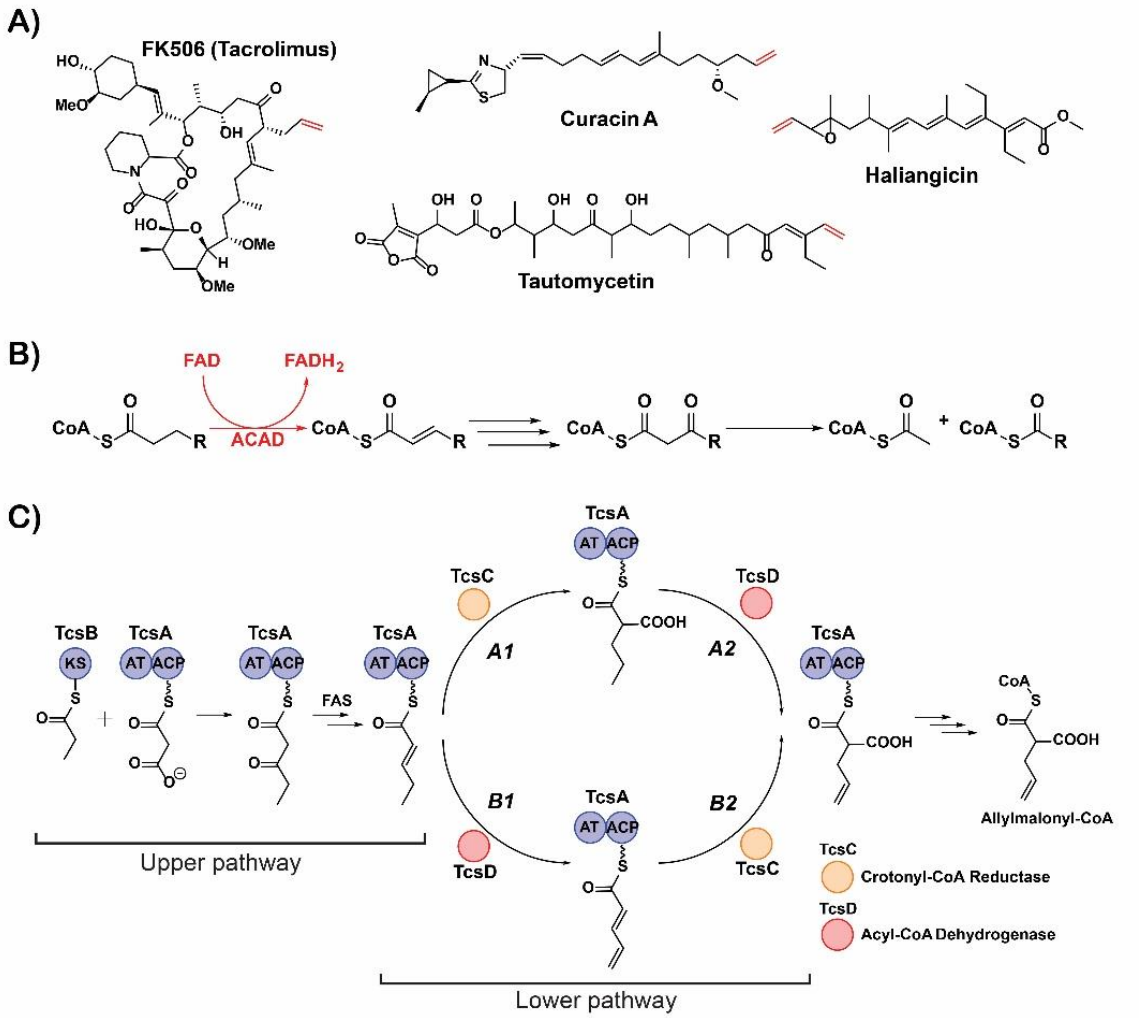



**Figure 1.** Terminal alkene-containing natural products and alkene formation by acyl-CoA dehydrogenases. **A)** Examples of terminal alkene-containing polyketide natural products. Terminal alkenes are highlighted in red. **B)** Process of fatty acid β-oxidation. The canonical activity of acyl-CoA dehydrogenases (ACADs) is the dehydrogenation of saturated fatty acyl-CoAs to form α,β-unsaturated products with concomitant reduction of FAD, highlighted in red. **C)** Proposed steps of allylmalonyl-CoA biosynthesis (adapted from reference 39). KS = ketosynthase domain, AT = acyltransferase domain, ACP = acyl carrier protein, FAS = fatty acid synthase.

The allylmalonyl-CoA biosynthetic pathway was initially proposed as a hybrid PKS-fatty acid synthase (FAS) pathway in which a free-standing ketosynthase, TcsB, first condenses a propionate group with a malonyl-CoA-derived extender unit selected by the AT domain of TcsA to form 3-oxo-pentanoyl-TcsA (Figure 1C, upper pathway).[39] After reduction by the producing organism's FAS, it was proposed that the final two enzymes TcsC (a crotonyl-CoA reductase) and TcsD (an ACAD) work interchangeably to convert 2-pentenoyl-TcsA to allylmalonyl-TcsA, with TcsD forming the unique terminal alkene moiety (lower pathway).

The sequence homology that TcsD shares with other ACADs suggests that it utilizes the same chemical mechanism to form the γ,δ-alkene of the allyl functional group. ACADs employ a catalytic base, typically a glutamate

residue, to deprotonate the acidic α-proton of a fatty acyl-CoA substrate.[40-43] The concomitant transfer of a hydride from the β-carbon of the substrate to FAD results in the formation of an α,β-unsaturated product and the reduced flavin, $FADH_2$.[44] The reaction is mediated by $pK_a$ perturbations of both the catalytic glutamate residue and the substrate α-protons that occur within the active site of the enzyme. The $pK_a$ of the glutamate is raised from ~6 to ~9 due to desolvation of the active site upon the binding of a hydrophobic substrate,[45-47] while the substrate protons are activated via hydrogen bonds of the substrate thioester carbonyl group with the amide backbone of the glutamate and the 2'-hydroxyl group of FAD.[48-51] While this mechanism of substrate activation is plausible for the γ-protons of a substrate such as 2-pentenoyl-TcsA (which contains an α,β-alkene that can propagate electronic effects from the thioester to the γ-carbon), the activation of propylmalonyl-TcsA and its conversion to allylmalonyl-TcsA (Figure 1C, pathway *A2*) is highly unlikely due to the aliphatic nature of the substrate.

Here we interrogate the activity of TcsD on both potential substrates and show that pathway *B1* (Figure 1C) is the only route of allylmalonyl-TcsA formation. Additionally, we show that TcsD oxidizes only α,β-unsaturated substrates and is regioselective for the γ,δ position of these substrates. Further, we present a high resolution TcsD crystal structure and propose a structural mechanism by which it exhibits precise regiocontrol over this transformation. Combined structural and biochemical analyses of this enzyme revealed signature residues that contribute to its unique regioselectivity and facilitate the identification of homologs that display the same biochemical activity. A better understanding of this unique enzyme's activity will not only inform the characterization of other homologs and their associated BGCs, but the insights gained herein can also aid future enzyme engineering efforts.

## Results and Discussion

### Biochemical activity of TcsD

In order to understand the mechanisms underlying TcsD activity, we initially sought to determine the native substrate(s) of the enzyme by biochemically interrogating both pathway *B*1 and *A*2 (Figure 1C). The substrates 2-pentenoyl-TcsA (pathway *B1*) and propylmalonyl-TcsA (pathway *A2*) were generated by loading the corresponding Coenzyme A esters onto Ser374 of the AT-ACP didomain protein TcsA. These substrates were then incubated with TcsD in the presence of the external electron acceptor ferrocenium hexafluorophosphate to facilitate enzyme turnover.[52] Assays were analyzed using targeted LC-MS/MS to monitor for the characteristic phosphopantetheine ejection transition (Figure 2A).[53] As expected, TcsD converted nearly all of the 2-pentenoyl substrate to the corresponding 2,4-pentadienoyl-TcsA

product, with no activity detected in boiled TcsD controls (Figure 2B). However, TcsD was unable to convert propylmalonyl-TcsA to allylmalonyl-TcsA under identical assay conditions (Figure S1). We therefore concluded that the biosynthesis of allylmalonyl-CoA can only proceed through route *B*, in which TcsD first oxidizes 2-pentenoyl-TcsA to 2,4-pentadienoyl-TcsA (*B*1) and subsequently TcsC performs a reductive carboxylation to form allylmalonyl-TcsA (*B*2).

Due to the noncanonical regiochemistry of the TcsD-mediated dehydrogenation of 2-pentenoyl-TcsA, we further investigated whether TcsD is regioselective for the γ,δ-position or if it simply oxidizes any appropriately activated substrate. It is known that some ACADs can abstract the acidic γ-proton of α,β-unsaturated substrates after dehydrogenation.[41,42,54] Accordingly, it is plausible that TcsD is a promiscuous enzyme that dehydrogenates any substrate present on the ACP domain of TcsA and that any specificity it exhibits *in vivo* arises from the sequestration of substrates on a protein (TcsA) instead of Coenzyme A. We therefore probed TcsD activity on a panel of α,β-unsaturated and fully saturated substrates. The enzyme was active on another α,β-unsaturated substrate, 2-hexenoyl-TcsA, but inactive on the seven carbon 2-heptenoyl-TcsA (Figure 2B). On the substrates butyryl-TcsA and pentanoyl-TcsA, where possible α,β-unsaturation could be expected, no oxidation activity was observed which indicated that TcsD acts regioselectively at the γ,δ-position (Figure 2B, S2 & S3).

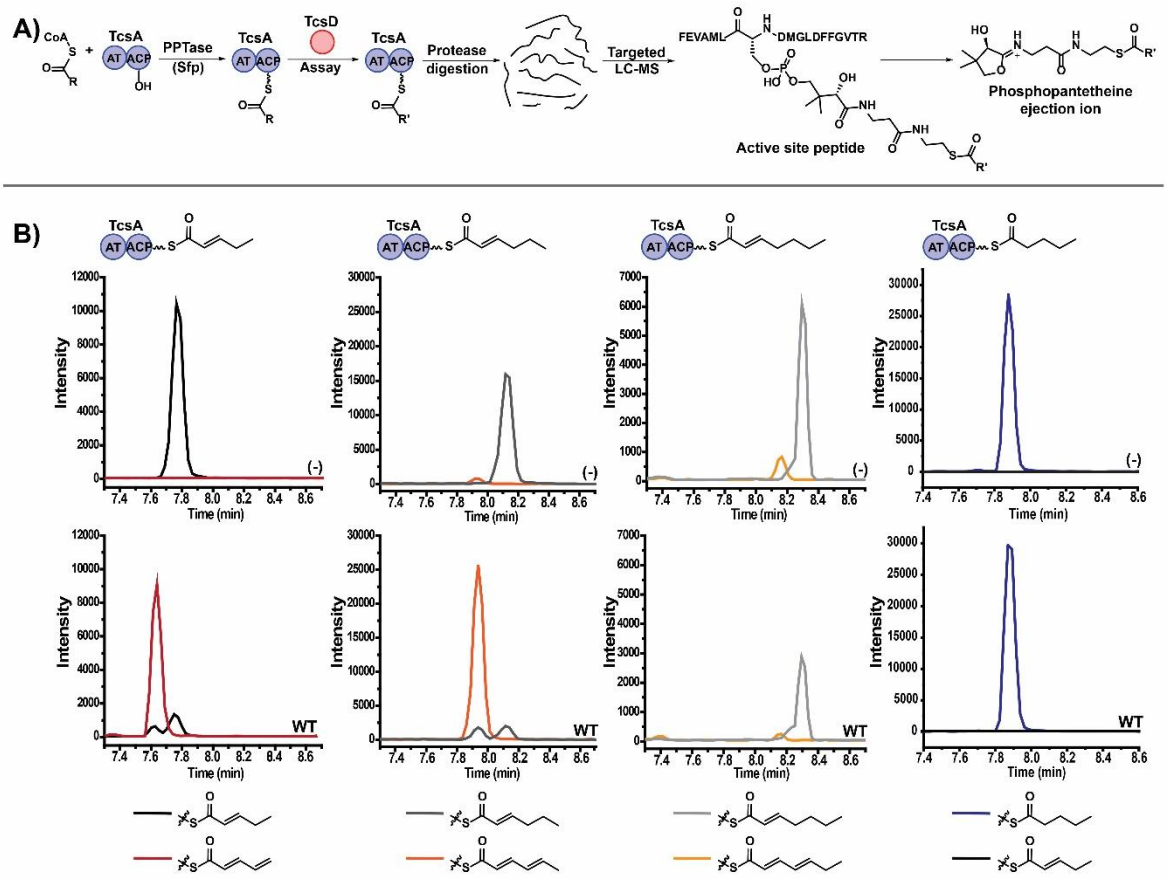



**Figure 2**. Biochemical activity of TcsD on ACP-bound substrates. **A)** Experimental design for phosphopantetheine ejection assays of TcsD activity on TcsA-bound substrates **B)** LC-MS/MS chromatograms of TcsD activity assays on various substrates. Substrates from left to right: 2-pentenoyl-TcsA, 2-hexenoyl-TcsA, 2-heptenoyl-TcsA, pentanoyl-TcsA. Top row: negative controls with boiled TcsD (-). Bottom row: assays with intact TcsD (WT).

## Crystal structure of TcsD

The strict substrate specificity of TcsD at the γ,δ-position suggested that there are structural elements within the enzyme's active site that control regioselectivity. To test this hypothesis, we obtained a crystal structure of TcsD, which was solved to 1.75Å resolution. The TcsD structure displays many similarities to those of other ACADs, such as the conserved ACAD fold consisting of a tetrameric quaternary structure, which is further organized into two sets of homodimers (Figure S4). Each subunit contains a single FAD cofactor and is composed of 3 subdomains consisting of a set of N- and C-terminal alpha helix domains that surround a middle beta sheet domain. The FAD cofactor adopts an extended conformation and is situated in a pocket formed by the C-terminal alpha helix domain, the middle beta sheet domain, and the C-terminal domain of the adjacent subunit (Figure S4).

The active site of TcsD also shares a similar overall organization with α,β-ACADs. Like the structures of the human, rat, pig, and bacterial ACADs, the catalytic base of TcsD, Glu364, sits poised at the top of the active site pocket immediately above the fatty acyl binding chamber.[40,43,55-57] The back side of the pocket that bounds the fatty acyl binding

region is shaped by the residue immediately upstream of the catalytic glutamate, Ile363, and three residues from helix 5: Phe79, Leu83, and Leu86. The FAD cofactor sits at the bottom of the active site, positioned below Glu364(Figure 3A).

The isoalloxazine ring of FAD is anchored via conserved hydrogen bonds with residues Thr116, Gly118, Ser119, and Thr151, while the adenosine pyrophosphate portion of the molecule extends into a cavity formed between the loop that follows β-sheet 1 and helices 10 and 11 of the adjacent subunit; it is positioned through polar interactions with Ser125 and Glu337 of the same subunit and Met338, Gly340, and Gly341 of the adjacent subunit (Figure S5). Similar to other ACAD structures, an aromatic residue, Phe149, is positioned on the si face of the isoalloxazine ring. The FAD is positioned adjacent to the substrate binding cavity which is bound on the opposite side by several conserved residues from helix 7 and the loop following sheet 1, including Gly118, Ser119, Glu120, Ser125, Leu127, Leu224, and Leu228 which form the phosphopantetheine binding region.

Although similar in overall structure, TcsD possesses many features that distinguish it from canonical ACADs. The residues

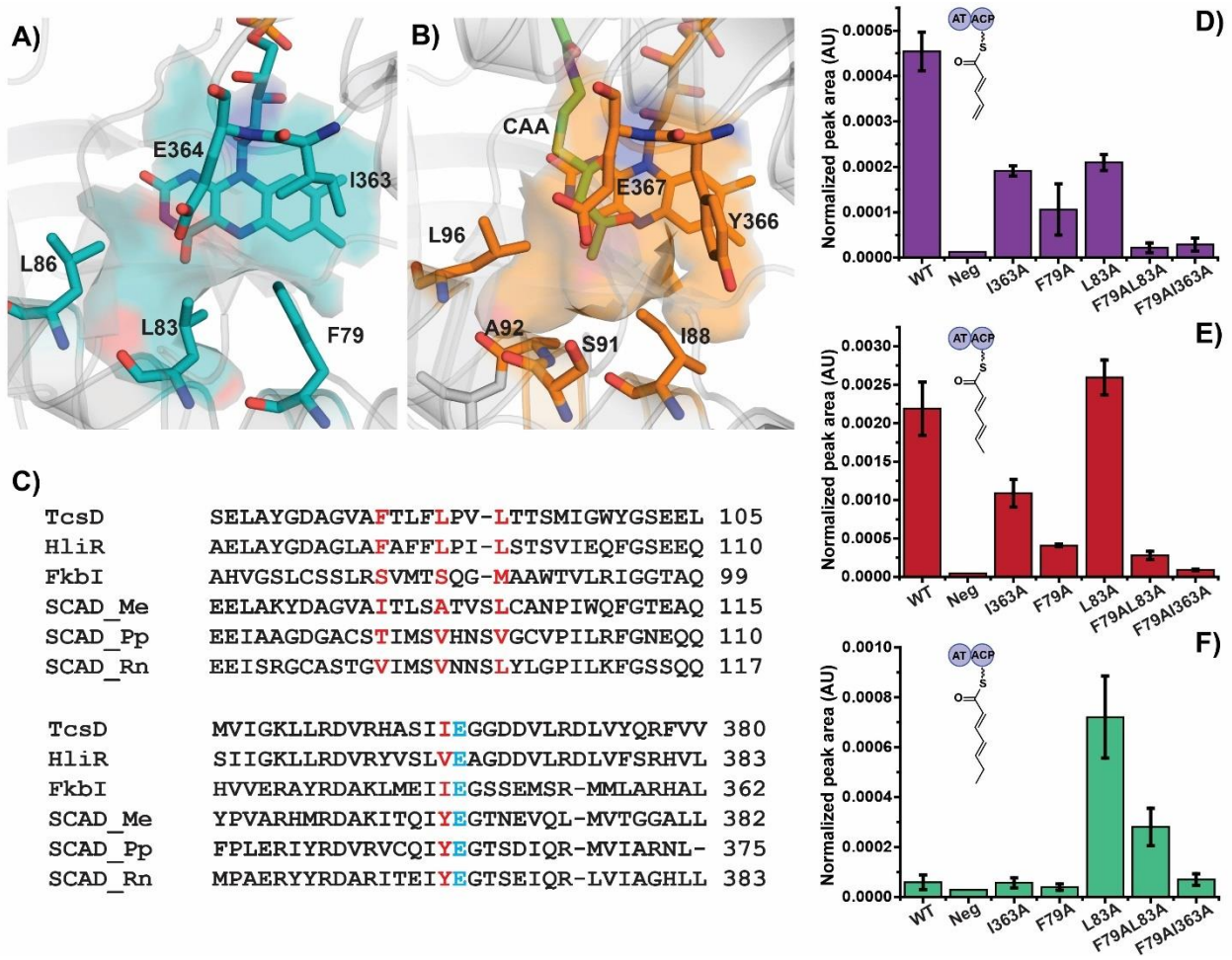



**Figure 3**. Unique active site features of TcsD and activity of active site mutants. **A)** Shape of the fatty acyl binding region of TcsD. **B)** Shape of the fatty acyl binding region of the *Megasphaera elsdenii* butyryl-CoA dehydrogenase (pdb entry 1buc), which was co-crystallized with acetoacetyl-CoA (CAA)[57] **C)** Sequence alignment of TcsD with HliR,[38] FkbI,[59] *M. elsdenii* butyryl-CoA dehydrogenase (SCAD_Me, 1buc),[57] *Pseudomonas putida* KT2440 short chain acyl-CoA dehydrogenase (SCAD_Pp),[58] and rat short chain acyl-CoA dehydrogenase (SCAD_Rn, 1jqi).[56] The residues in red correspond to residues that form the acyl-binding region of the enzyme active site (Phe79, Leu83, Leu86, and Ile363 in TcsD). The active site glutamate is highlighted in blue. **D)**, **E)**, **F)** TcsD mutant activity on 2-pentenoyl-TcsA* (**D**), 2-hexenoyl-TcsA* (**E**), and 2-heptenoyl-TcsA* (**F**) calculated as the LC-MS/MS peak area of the product normalized to a control peptide from TcsA (Normalized peak area, see Methods in SI). Data displayed are the average of three replicates. AU = arbitrary units *Note: TcsD activity on 5 carbon substrates cannot be compared to TcsD activity on 6 or 7 carbon substrates.

surrounding the TcsD active site entrance have diverged from the corresponding residues that are conserved in α,β-ACADs and contribute to the positioning of the adenosine portion of the coenzyme A substrate. TcsD does not possess the conserved Asp and Thr residues that form hydrogen bonds with the adenine base of Coenzyme A (Figure S6), nor the small helix between β-sheets 5 and 6 that often hydrogen bonds with the coenzyme A phosphate groups (Figure S7). The loop between β-sheets 5 and 6 in TcsD has instead been shortened. Additionally, a short helix, helix 9, is present in the TcsD structure in a position that would collide with the adenine base of Coenzyme A as it is positioned in α,β-ACAD

structures. While helix 9 in the TcsD structure seems to interfere with Coenzyme A binding, it is also present in the TcsD homolog from the haliangicin pathway, HliR, suggesting that this helix might contribute to the regiochemical shift observed in both enzymes (Figure S8). TcsD has unique structural features that form the fatty acyl binding region of the active site pocket. In particular, Ile363 and the residues of helix 5 that line the back side of the active site pocket protrude further into the active site than other ACADs, forming a much shallower binding pocket (Figure 3A & B). More specifically, Phe79 and Leu83 protrude directly into the active site pointing toward the fatty acyl binding position. In addition, these two residues are bulkier than the residues forming the active site of canonical ACAD homologs, contributing to the reduced active site pocket depth seen in the TcsD structure. The difference in active site shape and depth is further illustrated by comparing TcsD with the structures and sequences of α,β-ACADs containing co-crystallized substrates. Alignment and superposition of the *Megasphaera elsdenii* butyryl-CoA dehydrogenase (BCAD) in complex with acetoacetyl-CoA[57] with the TcsD active site shows that Phe79, Leu83, Leu86, Ile363 or a combination of these residues would sterically clash with substrates longer than acetoacetyl-CoA if they were to bind within the TcsD active site in a similar manner. The *M. elsdenii* BCAD and other short chain α,β-ACADs contain less bulky residues at the positions corresponding to Phe79, Leu83, and Ile363, but Leu86 is generally conserved in TcsD, HliR, and α,β-ACADs (Figure 3B & S8).[34,56,58-60] In addition, the bulky residues Phe79 and Leu83 distinguish TcsD from FkbI, the hydroxymalonate semialdehyde dehydrogenase from the methoxymalonyl-ACP biosynthetic pathway of FK506 which has serines at both of these positions.[59]Given the unique structural characteristics of TcsD, we hypothesized that the positioning of the amino acids Phe79, Leu83, and Ile363 could control the regioselectivity of TcsD by preventing either proton abstraction by Glu364 or hydride transfer to FAD through steric repulsion. More specifically, these large residues would prohibit a five carbon substrate from entering the active site far enough so that Glu364 can access the protons bound to the substrate α-carbon or for the β-carbon to transfer a hydride to N5 of FAD. Rather, the substrate would be pushed towards the active site entrance, positioning Glu364 above the γ-carbon and the δ-carbon within an appropriate distance of N5 of FAD. As the substrates pentanoyl-TcsA and 2-pentenoyl-TcsA are similarly chemically activated (the $pK_a$s of the α- and γ-protons are similar), the regioselectivity of TcsD would be controlled purely by steric interactions, not an electronic preference for one substrate.

**Biochemical activity of TcsD mutants**

In order to verify whether the regioselectivity of TcsD is sterically controlled, a mutant of TcsD with a larger active site pocket that can accommodate a longer substrate is required. Specifically, TcsD mutants that act upon a

substrate with two additional carbons should also be able to bind pentanoyl-TcsA in a manner that properly positions the α- and β-carbons within an accessible distance of Glu364 and FAD, respectively, as the α-carbon is two carbons removed from the γ-carbon. The residues that surround the active site of TcsD were therefore selectively mutated to alanines in order to accommodate longer substrates. Individual alanine mutants of Phe79, Leu83, and Ile363 in addition to double mutants of neighboring residues (e.g. L83A/L86A, F79A/L83A, and F79A/I363A) were generated, and the dehydrogenation activity of mutants was then probed on a panel of α,β-unsaturated substrates. The active site mutants were generally less active on 2-pentenoyl-TcsA than the wild type enzyme (Figure 3D). However, the I363A and L83A mutants displayed nearly equal or equivalent activity to the wild type on 2-hexenoyl-TcsA (Figure 3E). Proteins containing the L83A mutation were also active on the 2-heptenoyl-TcsA substrate (Figure 3F). This indicates that Leu83 controls the chain length of the substrate.

After demonstrating that the active site of the L83A mutant had been enlarged to accommodate two additional carbons, we next probed its activity on 4- and 5-carbon fully saturated substrates to test if the enzyme was now capable of dehydrogenating the α,β-position in the absence of the wild type steric interactions. We found that the mutant enzyme was still inactive on butyryl- and pentanoyl-TcsA (data not shown), suggesting that the observed regioselectivity is not controlled exclusively by steric interactions with the fatty acyl tail of a substrate.

**Substrate modeling into TcsD active site**

Given the retention of selectivity of the TcsD L83A mutant, we hypothesized that the regioselectivity of dehydrogenation could instead be controlled by intermolecular forces affecting the positioning of the thioester end of the substrate, which is determined by two hydrogen bonds (H-bonds) within the active site. We hypothesized that the canonical ACAD H-bonds should be conserved in TcsD as they are crucial to the chemical mechanism of the enzyme. To test this, we attempted to co-crystallize TcsD with 2-pentenoyl-CoA and 2-hexenoyl-CoA in order to show the substrate positioning within the active site; however, we were unable to observe density corresponding to either substrate in the TcsD active site pocket. It is possible that the presence of an unidentifiable density occupying the active site, possibly a polyethylene glycol fragment derived from the crystallization buffer, blocks substrates from entering the substrate-binding region (Figures S14, S15, & S16 & Supplementary Discussion).

Instead, we computationally modeled the native TcsD substrate into the active site to understand which structural features might affect substrate binding. The positioning of the thioesters in the structures of several ACADs (which were co-crystallized with substrate analogs) was used to guide the placement of

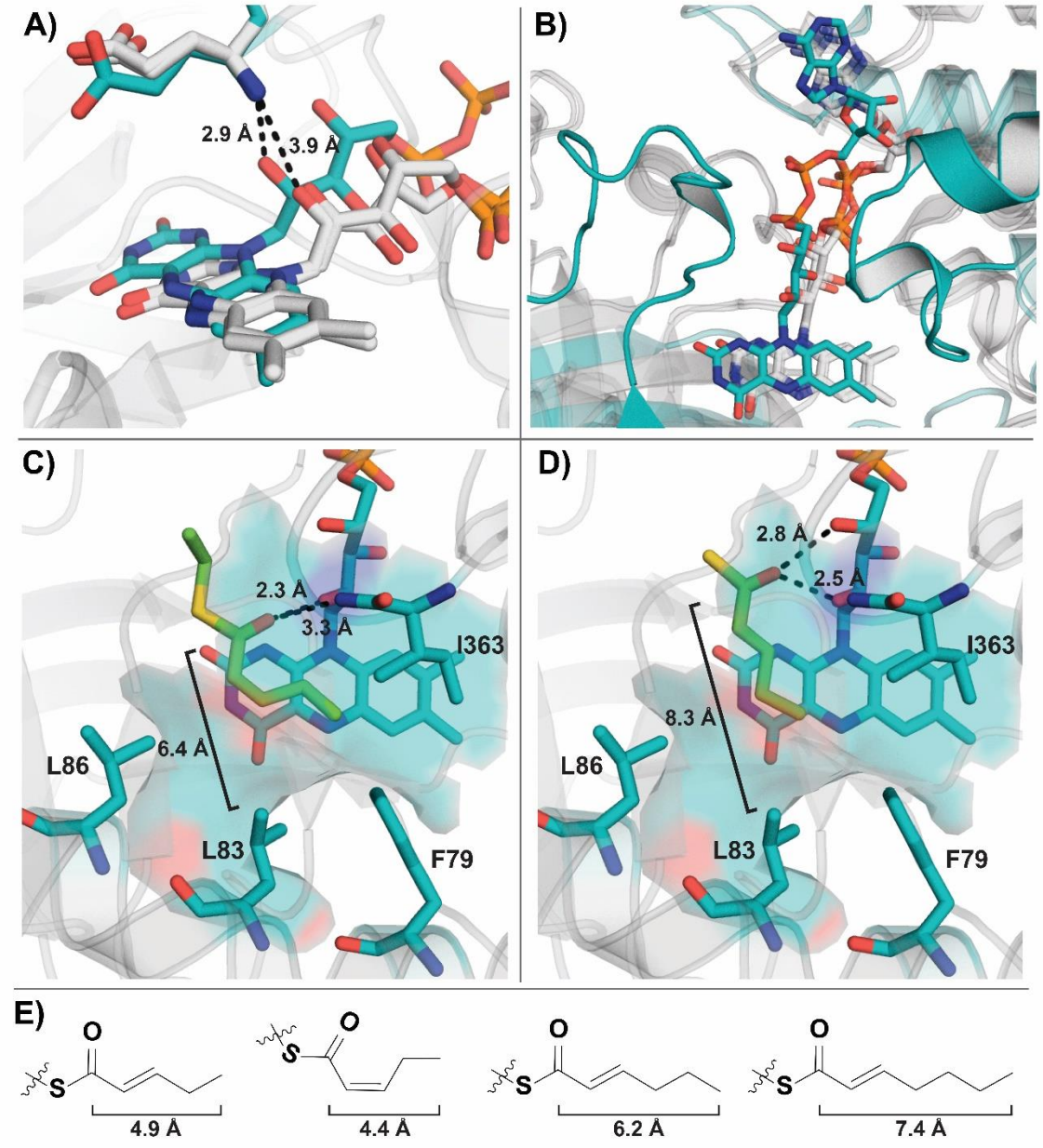



**Figure 4** FAD shift and substrate modeling in TcsD active site. **A)** Shift in positioning of the 2'-OH group of FAD in TcsD (teal) relative to the active site glutamate and compared to the pig (*Sus scrofa*) medium chain ACAD in the apo (pdb code 3mdd)[43] and substrate-bound (pdb code 1udy)[55] forms (both white). Dotted lines represent distances between 2'-OH groups of FAD and amide nitrogens of active site glutamates. **B)** Lateral shift of FAD in TcsD (teal) with respect to the active site glutamate and compared to the pig (*Sus scrofa*) medium chain ACAD in the apo (pdb code 3mdd) and substrate-bound (pdb code 1udy) forms (both white) **C)** TcsD active site* containing a modeled 2-pentenoyl thioester group in the *cis* conformation. Dotted lines = distances between the substrate thioester carbonyl group and the 2'-OH of FAD and amide nitrogen of Glu364 **D)** TcsD active site* containing a modeled 2-pentenoyl thioester group in the *trans* conformation. Dotted lines = distances between the substrate thioester carbonyl group and the 2'-OH and 4'-OH of FAD **E)** Distance from carbonyl carbon to the tail carbon of various α,β-unsaturated fatty acyl substrates. *Note: Glu364 was omitted from C) and D) in order to visualize the hydrogen bonding interactions of the substrate with FAD.

a substrate into the TcsD active site. More specifically, the structures of several homologs and TcsD were superimposed by aligning the peptide backbones of the three residue loop that contains the catalytic glutamate (e.g. Ile363, E364, and G365). This three residue loop was chosen as an anchor point because it is highly conserved and contains one of the H-

bond donors that contributes to thioester positioning, so the relative positioning of structural elements within the ACAD active sites can be used to infer relative substrate positioning. Not only was this type of alignment useful for appropriately positioning modeled substrates, but it also allowed for a more precise determination of slight structural differences that were not as obvious when analyzing global structural alignments. In particular, it highlighted stark differences in the positioning of the FAD cofactor in TcsD relative to several homologs (Figure 4A and 4B).

In relation to the active site glutamate (Glu364), the FAD molecule bound within the TcsD active site is shifted laterally, moving it in the direction of the pantetheine-binding region of the enzyme (Figure 4A and B). This movement of the ribityl side chain of FAD is accompanied by a shift in α-helices 10 and 11 and the loop that follows β-sheet 1, which interact with the FAD through various H-bonds. The repositioning of α-helices 10 and 11 results in a change in the conformation of the FAD ribose ring due to H-bonding interactions of the 2' and 3'-OH groups with the amide backbone of Met338 and of Glu337 and Gly339 of the adjacent subunit, respectively. The conformational change pushes the FAD phosphate groups towards the β-sheet domain of the enzyme. As a result of these structural changes, the positioning of the 2'-hydroxyl group of the FAD ribityl side chain is shifted closer to the nitrogen of the Glu364 amide bond (Figure 4A). We hypothesized that this would result in a concomitant shift in the positioning of the substrate thioester moiety, making TcsD act at the γ,δ-position instead ofat the canonical α,β-position of substrates.

In order to better understand how the change in FAD positioning affects substrate binding, we modeled a substrate analog consisting of a 2-pentenoyl thioester into the active site. Initially, we positioned the thioester carbonyl oxygen within hydrogen bonding distance of both the 2'-OH of the FAD and the nitrogen of the amide bond of Glu364 (Figure 4C). Anchoring the carbonyl group in this location

would force the substrate alkene into a *cis* conformation, as this is the only conformation that places the δ-carbon in close enough proximity to N5 of FAD. However, while this conformation enforces the proper positioning of the substrate with respect to FAD, it places the δ-carbon too close to residues Phe79 and Ile363 which would result in unfavorable steric repulsion between the substrate and the amino acid side chains lining the active site. Moreover, with this substrate conformation there is no additional space within the active site to accommodate a sixth carbon on the substrate which conflicts with our biochemical data.

After eliminating the possibility of substrate binding in the *cis* conformation, a *trans* substrate was next modeled into the TcsD active site. When the substrate alkene bond adopts a *trans* conformation, its positioning is more consistent with the biochemical activity of the enzyme as it places the fatty acyl tail turned toward Leu83 and the δ-carbon in

proximity to the N5 of FAD (Figure 4D). However, with the substrate in this extended position, H-bonding of the substrate with the canonical ACAD H-bond donors cannot occur without a steric clash between Leu83 and the fatty acyl tail of the substrate. We therefore modeled the carbonyl carbon of the substrate so that it H-bonds with the 2' and 4'-OH groups of FAD, not with the canonical amide nitrogen of Glu364. Given the size dimensions of the active site and the biochemical data presented herein, this appears to be the most probable positioning of the substrate within the active site, suggesting that TcsD utilizes a novel H-bond donor pair to achieve its unique regioselectivity.

**Genome mining reveals previously unidentified γ,δ-ACADs**

After identifying structural features that contribute to the unique regioselectivity of TcsD, we applied this mechanistic knowledge to find unidentified or misannotated γ,δ-ACADs among sequenced bacteria. In particular, we used the presence of bulky residues at positions corresponding to Phe79, Leu83, and Ile363 and helix 9 as requirements for the identification of γ,δ-ACADs. With these constraints, approximately 100 likely γ,δ-ACADs were identified using a combination of Hidden Markov Model (HMM), local sequence alignment (protein BLAST), and CORASON-BGC searches (Table S9 and S10). [61-63]

Notably, both TcsD and HliR, the only γ,δ-ACADs with associated natural products, were re-identified in our search. Of the newly-identified enzymes one feature in particular was strongly conserved: the presence of helix 9, which we had identified as a unique feature in the TcsD structure (Figure 5A and B). The exact purpose of this loop remains unclear and will be the subject of further investigation. Additionally, we found that the presence of bulky residues at positions in the substrate acyl binding region were highly conserved across the family, with 98% of homologs displaying a Phe residue (2% Leu) and 96% displaying a Leu or Ile residue (4% Val) at the positions corresponding to Phe79 and Leu83 in TcsD (Figure 5C and S9). The residues immediately preceding the catalytic glutamate (i.e. residue Ile363 in TcsD) were all bulky aliphatic residues as well (87% Val, 11% Ile, 2% Phe). However, bulky residues at this position are not unique to the γ,δ-ACAD family as they are also observed in hydroxymalonate semialdehyde dehydrogenases such as FkbI.

Most of the TcsD homologs were associated with identifiable BGCs, although some were not located near any obvious secondary metabolite genes. The majority of the BGC-associated genes were located within type I PKS (T1PKS) clusters, but several were also identified within type II PKS (T2PKS) and nonribosomal peptide synthetase (NRPS) clusters (Figure 5D and S10). None of the newly-identified BGCs containing the putative γ,δ-ACADs have been experimentally characterized, but many show significant homology to known clusters (Table S10). Of the clusters that

include putative γ,δ-ACADs, several contained proteins with homology to BGCs with known products, including the arsenopolyketides (T1PKS), oligomycin (T1PKS), E-837 (T1PKS), chlorothricin (T1PKS), butyrolactols (T1PKS), polyoxypeptin (T1PKS-NRPS), hedamycin (T2PKS), and erythrochelin (NRPS). Many of the homologous BGCs identified by antiSMASH have fatty acyl tails that could be potential substrates for γ,δ-ACADs in the uncharacterized clusters we identified.

Within the identified BGCs, analysis of the genomic contexts of the putative γ,δ-ACADs showed several patterns of syntenic genes (Figure 5D and S10). The syntenic genomic regions can be grouped based on the type of gene cluster (e.g., probable butyrolactol or arsenopolyketide BGC) within which the putative γ,δ-ACADs occur. The groups of genes that are found near γ,δ-ACADs can also be used to postulate the enzymes' native substrates and even certain aspects of the molecular structure of the final

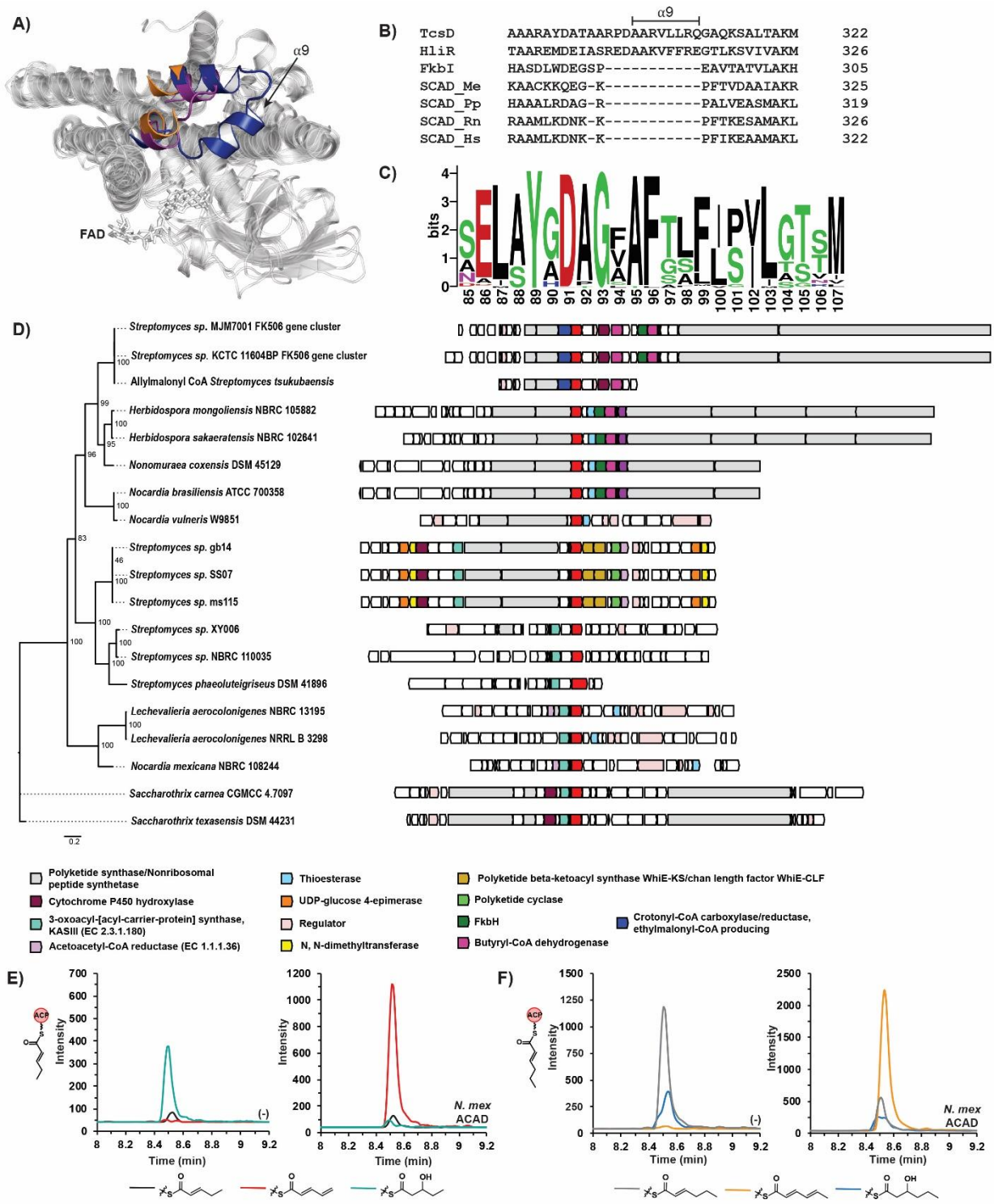



**Figure 5** Genome mining and characterization of γ,δ-ACADs. **A)** Global overlay of TcsD with *M. elsdenii* butyryl-CoA dehydrogenase (BCAD_Me, 1buc)[57] and *Streptomyces hygroscopicus* hydroxymalonate semialdehyde dehydrogenase, FkbI (1r2j)[59] highlighting the insertion of helix 9 into γ,δ-ACAD structures. The colored regions are helix 9 of TcsD (blue, labeled α9) and the corresponding regions in BCAD_Me (purple) and FkbI (orange). FAD is depicted as sticks. **B)** Sequence alignment of TcsD with HliR,[38] FkbI, *M. elsdenii* butyryl-CoA dehydrogenase (SCAD_Me, 1buc), *Pseudomonas putida* KT2440 short chain acyl-CoA dehydrogenase (SCAD_Pp),[59] and rat short chain acyl-CoA dehydrogenase (SCAD_Rn, 1jqi).[56] The sequence alignment highlights the presence of helix 9 in γ,δ-ACADs which is not encountered in other ACAD families. **C)** Amino acid position weight matrix showing the relative abundance of amino acids at each position within γ,δ-ACADs. Positions 96 and 100 correspond to the positions of Phe79 and Leu83 of TcsD, respectively. Logo was generated from a sequence alignment of all γ,δ-ACADs identified in this work. **D)** Genomic contexts and synteny of a representative set of γ,δ-ACADs identified in this work.TcsD homologs (putative γ,δ-ACADs) are shown in red and aligned. Other notable genes are annotated by

color. **E)** Targeted LC-MS/MS chromatograms showing the activity of the *Nocardia mexicana* NBRC 108244 putative γ,δ-ACAD on 2-pentenoyl-ACP **F)** Targeted LC-MS/MS chromatograms showing the activity of the *Nocardia mexicana* NBRC 108244 putative γ,δ-ACAD on 2-hexenoyl-ACP. (-) is used to mark negative controls with boiled *N. mexicana* ACAD, while assays are marked as *N. mex* ACAD. Chromatograms representing the 2,3-enoyl-, 2,4-dienoyl-, and 3-hydroxy transitions are color-coded below each set of chromatograms. The 3-hydroxyacyl groups are normal degradation products of 2,3-enoyl-thioesters in aqueous environments where hydration of the alkene can occur.

natural product. The γ,δ-ACADs found in the erythrochelin-like gene clusters in *Saccharothrix sp.*, for example, are consistently clustered with a free-standing ketosynthase (KS) and acyl carrier protein (ACP) pair (Figure 5D). It is plausible that the KS domain forms a 5 carbon ACP-bound substrate upon which the γ,δ-ACAD can act after it is reduced to a 2-pentenoyl form. The genomic context of this and many of the other putative γ,δ-ACADs can be used to infer function and will inform future biosynthetic studies on the pathways and enzymes.

The presence of conserved residues and helix 9 in the putative γ,δ-ACADs is highly suggestive that they would exhibit the same activity as TcsD. To interrogate this hypothesis, we biochemically characterized a TcsD homolog. We chose the putative γ,δ-ACAD from *Nocardia mexicana* NBRC 108244 because it is located on a distant branch of the γ,δ-ACAD phylogenetic tree (Figure 5D and S10) and has low sequence identity relative to TcsD (50.7%). Though sharing only 51% sequence identity, the *N. mexicana* homolog (Nmex-ACAD) possesses the conserved Phe, Leu, and Ile residues that line the TcsD active site (Figure S11). It is located within a gene cluster that is predicted to encode a ladderane/butyrolactone-like natural product BGC (Figure S12). There are several related biosynthetic genes located immediately next to the Nmex-ACAD, including an acyl carrier protein, a ketosynthase, and a putative 3-oxoacyl-ACP reductase. Based on its genomic context we hypothesized that, like TcsD, the Nmex-ACAD may act on ACP-bound substrates. We therefore expressed and purified both the Nmex-ACAD and its neighboring acyl carrier protein (Nmex-ACP) and assayed Nmex-ACAD activity on the same panel of substrates. The Nmex-ACAD showed the same activity profile as TcsD, converting 2-pentenoyl-Nmex-ACP (Figure 5E) and 2-hexenoyl-Nmex-ACP (Figure 5F) to the corresponding dienoyl-ACP products, but it showed no activity on butyryl-, pentanoyl-, or 2-heptenoyl-ACP (Figure S13). While this data does not confirm that all the putative enzymes we have identified are γ,δ-ACADs, it strongly supports the prediction that these enzymes have the same activity as they are more closely phylogenetically related to TcsD than the Nmex-ACAD and share the conserved bulky residues that form the enzyme fatty acyl binding pocket.

**Conclusion**

Terminal alkenes in polyketide natural products can be formed in several ways, including through the action of an acyl-CoA dehydrogenase. In this work we have described the biochemical and structural characterization of TcsD, the

terminal alkene-forming γ,δ-ACAD from the biosynthesis of the allylmalonyl-CoA extender unit implicated in the biosynthesis of the polyketide FK506. We showed that TcsD acts on 2-pentenoyl-TcsA but not on propylmalonyl-TcsA, suggesting that the bottom half of the allylmalonyl-CoA pathway proceeds only through pathways *B1* and *B2*. TcsD only acts on 5- and 6-carbon α,β-unsaturated substrates and is regioselective for the γ,δ-position.

A crystal structure of TcsD revealed the unique features of the active site of the enzyme. Residues Phe79, Leu83, and Ile363 form a bulky wall in the substrate binding region of the enzyme, preventing the entrance of long fatty acyl substrates. Leu83 controls the chain length of the substrate. A TcsD L83A mutant acts on 2-heptenoyl-TcsA, but even with a larger active site pocket the mutant remains regioselective for the γ,δ-position of substrates. Closer analysis of the protein structure revealed that the enzyme regioselectivity is likely due to a shift in the positioning of the FAD cofactor. We show through substrate modeling that, because of the FAD shift and the dimensions of the TcsD active site, TcsD most likely employs a novel hydrogen bond donor pair (the 2'-OH and 4'-OH groups of FAD) to position and activate substrates. While Leu83 does not exclusively control regioselectivity, it contributes to regioselectivity by reducing the size of the active site.

The structural and biochemical conclusions from biochemically and structurally characterizing TcsD allowed us to determine key residues that define γ,δ-ACADs. Through HMM and local alignment searches, approximately 100 putative γ,δ-ACADs were identified in sequenced bacterial genomes. Nearly all of the homologs contained a Phe-Leu/Ile pair at the positions corresponding to Phe79-Leu83 in TcsD, respectively. The identification of other homologs also highlighted the conservation of helix 9, which appears to be a feature that is unique to the γ,δ-ACAD family. Most of the homologous enzymes were encountered in identifiable secondary metabolite BGCs, but some were notably located near no other canonical specialized metabolic enzymes. The synteny of genes located near the γ,δ-ACADs and the type of BGC to which they belong correlates strongly with their phylogenetic clustering and can be used to group the enzymes into several subfamilies. Finally, we showed that one of the TcsD homolog from *Nocardia mexicana*, which is phylogenetically one of the most distant enzymes from TcsD, also performs a regioselective dehydrogenation of the γ,δ-carbon of 2-pentenoyl- or 2-hexenoyl-ACP substrates, suggesting that this activity is conserved across the entire family.

This work exhibits how selective pressure causes enzymes to diverge not only at the amino acid sequence level, but also how it can result in significant shifts in protein structure to generate enzymes with divergent functions. Furthermore, it exemplifies how an understanding of the mechanisms employed by unique enzymes can be used as a means to refine the definitions of enzyme families and identify uncharacterized homologs. It will inform future efforts to

characterize the identified homologs and the BGCs they reside within and can be used as a guide for the future discovery of natural products that contain terminal alkene handles.

## Acknowledgements

This work was performed as part of the DOE Joint BioEnergy Institute (https://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, supported by the U.S. Department of Energy, Energy Efficiency and Renewable Energy, Bioenergy Technologies Office. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan). In addition, this work was supported by the Agile BioFoundry (https://agilebiofoundry.org), through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. The Advanced Light Source is a Department of Energy Office of Science User Facility under Contract No. DE-AC02-05CH11231. The ALS-ENABLE beamlines are supported in part by the National Institutes of Health, National Institute of General Medical Sciences, grant P30 GM124169-01. J.M.B. was supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1106400.

## Author Information

*Corresponding author (keasling@berkeley.edu)

## Notes

J.D.K. has financial interests in Amyris, Lygos, Demetrix, Napigen, Maple Bio, Ansa Biotechnologies, Berkeley Brewing Sciences, and Apertor Labs.

## Associated Content

*Supporting Information:* Materials, methods, detailed experimental procedures, supplementary discussion, bioinformatic analysis, supplementary figures, crystallographic information.

*Structural data deposition:* The atomic coordinates and structural factors of TcsD have been deposited in the Protein Data Bank, PDB ID code 6U1V.

*LC-MS/MS data deposition:* All targeted LC-MS/MS data from TcsD assays has been uploaded to Panorama Public[64] and is publicly available at the following link: https://panoramaweb.org/Structural%20control%20of%20bacterial%20acyl-CoA%20dehydrogenase.url

**References**


(1) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **2016**, *79*, 629–661.

(2) Dias, D. A.; Urban, S.; Roessner, U. A Historical Overview of Natural Products in Drug Discovery. *Metabolites* **2012**, *2*, 303–336.

(3) Scott, T. A.; Piel, J. The Hidden Enzymology of Bacterial Natural Product Biosynthesis. *Nat. Rev. Chem.* **2019**, *3*, 404-425.

(4) Minowa, Y.; Araki, M.; Kanehisa, M. Comprehensive Analysis of Distinctive Polyketide and Nonribosomal Peptide Structural Motifs Encoded in Microbial Genomes. *J. Mol. Biol.* **2007**, *368*, 1500–1517.

(5) Waldron, C.; Matsushima, P.; Rosteck, P. R.; Broughton, M. C.; Turner, J.; Madduri, K.; Crawford, K. P.; Merlo, D. J.; Baltz, R. H. Cloning and Analysis of the Spinosad Biosynthetic Gene Cluster of Saccharopolyspora Spinosa. *Chem. Biol.* **2001**, *8*, 487–499.

(6) Hagen, A.; Poust, S.; Rond, T. de; Fortman, J. L.; Katz, L.; Petzold, C. J.; Keasling, J. D. Engineering a Polyketide Synthase for in Vitro Production of Adipic Acid. *ACS Synth. Biol.* **2016**, *5*, 21–27.

(7) Copéret, C.; Adolfsson, H.; Barry Sharpless, K. A Simple and Efficient Method for Epoxidation of Terminal Alkenes. *Chem. Commun.* **1997**, 1565–1566.

(8) Mlynarski, S. N.; Schuster, C. H.; Morken, J. P. Asymmetric Synthesis from Terminal Alkenes by Cascades of Diboration and Cross-Coupling. *Nature* **2014**, *505*, 386–390.

(9) Becker, H.; Sharpless, K. B. A New Ligand Class for the Asymmetric Dihydroxylation of Olefins. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 448–451.

(10) Xu, J.; Fu, Y.; Luo, D.-F.; Jiang, Y.-Y.; Xiao, B.; Liu, Z.-J.; Gong, T.-J.; Liu, L. Copper-Catalyzed Trifluoromethylation of Terminal Alkenes through Allylic C-H Bond Activation. *J. Am. Chem. Soc.* **2011**, *133*, 15300–15303.

(11) Breit, B.; Seiche, W. Hydrogen Bonding as a Construction Element for Bidentate Donor Ligands in Homogeneous Catalysis: Regioselective Hydroformylation of Terminal Alkenes. *J. Am. Chem. Soc.* **2003**, *125*, 6608–6609.

(12) Urkalan, K. B.; Sigman, M. S. Palladium-Catalyzed Oxidative Intermolecular Difunctionalization of Terminal Alkenes with Organostannanes and Molecular Oxygen. *Angew Chem Int Ed Engl* **2009**, *48*, 3146–3149.

(13) Rucker, R. P.; Whittaker, A. M.; Dang, H.; Lalic, G. Synthesis of Tertiary Alkyl Amines from Terminal Alkenes: Copper-Catalyzed Amination of Alkyl Boranes. *J. Am. Chem. Soc.* **2012**, *134*, 6571–6574.

(14) Coombs, J. R.; Morken, J. P. Catalytic Enantioselective Functionalization of Unactivated Terminal Alkenes. *Angew Chem Int Ed Engl* **2016**, *55*, 2636–2649.

(15) Nishimura, T.; Kakiuchi, N.; Onoue, T.; Ohe, K.; Uemura, S. Palladium(II)-Catalyzed Oxidation of Terminal Alkenes to Methyl Ketones Using Molecular Oxygen. *J. Chem. Soc., Perkin Trans. 1* **2000**, 1915–1918.

(16) Jensen, C. M.; Trogler, W. C. Catalytic Hydration of Terminal Alkenes to Primary Alcohols. *Science* **1986**, *233*, 1069–1071.

(17) Du, X.; Huang, Z. Advances in Base-Metal-Catalyzed Alkene Hydrosilylation. *ACS Catal.* **2017**, *7*, 1227–1243.

(18) Fürstner, A. Olefin Metathesis and Beyond. *Angew. Chem. Int. Ed.* **2000**, *39*, 3012–3043.

(19) Dong, Y.; Matson, J. B.; Edgar, K. J. Olefin Cross-Metathesis in Polymer and Polysaccharide Chemistry: A Review. *Biomacromolecules* **2017**, *18*, 1661–1676.

(20) Keatinge-Clay, A. T. The Structures of Type I Polyketide Synthases. *Nat. Prod. Rep.* **2012**, *29*, 1050–1073.

(21) Buslov, I.; Becouse, J.; Mazza, S.; Montandon-Clerc, M.; Hu, X. Chemoselective Alkene Hydrosilylation Catalyzed by Nickel Pincer Complexes. *Angew Chem Int Ed Engl* **2015**, *54*, 14523–14526.

(22) Cannon, J. S.; Grubbs, R. H. Alkene Chemoselectivity in Ruthenium-Catalyzed Z-Selective Olefin Metathesis. *Angew Chem Int Ed Engl* **2013**, *52*, 9001–9004.

(23) Cristina Silva Costa, D. Additions to Non-Activated Alkenes: Recent Advances. *Arabian Journal of Chemistry* **2017**.

(24) Chatterjee, A. K.; Choi, T.-L.; Sanders, D. P.; Grubbs, R. H. A General Model for Selectivity in Olefin Cross Metathesis. *J. Am. Chem. Soc.* **2003**, *125*, 11360–11370.
(25) O'Leary, D. J.; Blackwell, H. E.; Washenfelder, R. A.; Grubbs, R. H. A New Method for Cross-Metathesis of Terminal Olefins. *Tetrahedron Lett.* **1998**, *39*, 7427–7430.
(26) Andrus, M. B.; Lepore, S. D.; Sclafani, J. A. Selective Dihydroxylation of Non-Conjugated Dienes in Favor of the Terminal Olefin. *Tetrahedron Lett.* **1997**, *38*, 4043–4046.
(27) Campbell, A. N.; White, P. B.; Guzei, I. A.; Stahl, S. S. Allylic C-H Acetoxylation with a 4,5-Diazafluorenone-Ligated Palladium Catalyst: A Ligand-Based Strategy to Achieve Aerobic Catalytic Turnover. *J. Am. Chem. Soc.* **2010**, *132*, 15116–15119.
(28) Bigi, M. A.; White, M. C. Terminal Olefins to Linear α,β-Unsaturated Ketones: Pd(II)/Hypervalent Iodine Co-Catalyzed Wacker Oxidation-Dehydrogenation. *J. Am. Chem. Soc.* **2013**, *135*, 7831–7834.
(29) Dong, G.; Teo, P.; Wickens, Z. K.; Grubbs, R. H. Primary Alcohols from Terminal Olefins: Formal Anti-Markovnikov Hydration via Triple Relay Catalysis. *Science* **2011**, *333*, 1609–1612.
(30) Klar, U.; Buchmann, B.; Schwede, W.; Skuballa, W.; Hoffmann, J.; Lichtner, R. B. Total Synthesis and Antitumor Activity of ZK-EPO: The First Fully Synthetic Epothilone in Clinical Development. *Angew Chem Int Ed Engl* **2006**, *45*, 7942–7948.
(31) Klar, U.; Hoffmann, J.; Giurescu, M. Sagopilone (ZK-EPO): From a Natural Product to a Fully Synthetic Clinical Development Candidate. *Expert Opin. Investig. Drugs* **2008**, *17*, 1735–1748.
(32) Kino, T.; Hatanaka, H.; Hashimoto, M.; Nishiyama, M.; Goto, T.; Okuhara, M.; Kohsaka, M.; Aoki, H.; Imanaka, H. FK-506, a Novel Immunosuppressant Isolated from a Streptomyces. I. Fermentation, Isolation, and Physico-Chemical and Biological Characteristics. *J. Antibiot.* **1987**, *40*, 1249–1255.
(33) Mitsuhashi, S.; Matsuura, N.; Ubukata, M.; Oikawa, H.; Shima, H.; Kikuchi, K. Tautomycetin Is a Novel and Specific Inhibitor of Serine/Threonine Protein Phosphatase Type 1, PP1. *Biochem. Biophys. Res. Commun.* **2001**, *287*, 328–331.
(34) Fudou, R.; Iizuka, T.; Yamanaka, S. Haliangicin, a Novel Antifungal Metabolite Produced by a Marine Myxobacterium. 1. Fermentation and Biological Characteristics. *J. Antibiot.* **2001**, *54*, 149–152.
(35) Gu, L.; Wang, B.; Kulkarni, A.; Gehret, J. J.; Lloyd, K. R.; Gerwick, L.; Gerwick, W. H.; Wipf, P.; Håkansson, K.; Smith, J. L.; Sherman, D.H. Polyketide Decarboxylative Chain Termination Preceded by O-Sulfonation in Curacin a Biosynthesis. *J. Am. Chem. Soc.* **2009**, *131*, 16033–16035.
(36) Scaglione, J. B.; Akey, D. L.; Sullivan, R.; Kittendorf, J. D.; Rath, C. M.; Kim, E.-S.; Smith, J. L.; Sherman, D. H. Biochemical and Structural Characterization of the Tautomycetin Thioesterase: Analysis of a Stereoselective Polyketide Hydrolase. *Angew Chem Int Ed Engl* **2010**, *49*, 5726–5730.
(37) Li, W.; Luo, Y.; Ju, J.; Rajski, S. R.; Osada, H.; Shen, B. Characterization of the Tautomycetin Biosynthetic Gene Cluster from Streptomyces Griseochromogenes Provides New Insight into Dialkylmaleic Anhydride Biosynthesis. *J. Nat. Prod.* **2009**, *72*, 450–459.
(38) Sun, Y.; Feng, Z.; Tomura, T.; Suzuki, A.; Miyano, S.; Tsuge, T.; Mori, H.; Suh, J.-W.; Iizuka, T.; Fudou, R.; Ojika, M. Heterologous Production of the Marine Myxobacterial Antibiotic Haliangicin and Its Unnatural Analogues Generated by Engineering of the Biochemical Pathway. *Sci. Rep.* **2016**, *6*, 22091-22101.
(39) Mo, S.; Kim, D. H.; Lee, J. H.; Park, J. W.; Basnet, D. B.; Ban, Y. H.; Yoo, Y. J.; Chen, S.; Park, S. R.; Choi, E. A.; Kim, E.; Jin, Y-Y.; Lee, S-W.; Park, J.Y.; Liu, Y.; Lee, M.O.; Lee, K.S.; Kim, S.J.; Kim, D.; Park, B.C.; Lee, S.; Kwon, H.J.; Suh, J-W.; Moore, B.S.; Lim, S-K.; Yoon, Y.J. Biosynthesis of the Allylmalonyl-CoA Extender Unit for the FK506 Polyketide Synthase Proceeds through a Dedicated Polyketide Synthase and Facilitates the Mutasynthesis of Analogues. *J. Am. Chem. Soc.* **2011**, *133*, 976–985.
(40) Ghisla, S.; Thorpe, C. Acyl-CoA Dehydrogenases. A Mechanistic Overview. *Eur. J. Biochem.* **2004**, *271*, 494–508.
(41) Fendrich, G.; Abeles, R. H. Mechanism of Action of Butyryl-CoA Dehydrogenase: Reactions with Acetylenic, Olefinic, and Fluorinated Substrate Analogues. *Biochemistry* **1982**, *21*, 6685–6695.
(42) Powell, P. J.; Thorpe, C. 2-Octynoyl Coenzyme A Is a Mechanism-Based Inhibitor of Pig Kidney Medium-Chain Acyl Coenzyme A Dehydrogenase: Isolation of the Target Peptide. *Biochemistry* **1988**, *27*, 8022–8028.

(43) Kim, J. J.; Wang, M.; Paschke, R. Crystal Structures of Medium-Chain Acyl-CoA Dehydrogenase from Pig Liver Mitochondria with and without Substrate. *Proc Natl Acad Sci USA* **1993**, *90*, 7523–7527.
(44) Ghisla, S.; Thorpe, C.; Massey, V. Mechanistic Studies with General Acyl-CoA Dehydrogenase and Butyryl-CoA Dehydrogenase: Evidence for the Transfer of the Beta-Hydrogen to the Flavin N(5)-Position as a Hydride. *Biochemistry* **1984**, *23*, 3154–3161.
(45) Trievel, R. C.; Wang, R.; Anderson, V. E.; Thorpe, C. Role of the Carbonyl Group in Thioester Chain Length Recognition by the Medium Chain Acyl-CoA Dehydrogenase. *Biochemistry* **1995**, *34*, 8597–8605.
(46) Mancini-Samuelson, G. J.; Kieweg, V.; Sabaj, K. M.; Ghisla, S.; Stankovich, M. T. Redox Properties of Human Medium-Chain Acyl-CoA Dehydrogenase, Modulation by Charged Active-Site Amino Acid Residues. *Biochemistry* **1998**, *37*, 14605–14612.
(47) Becker, D. F.; Fuchs, J. A.; Stankovich, M. T. Product Binding Modulates the Thermodynamic Properties of a Megasphaera Elsdenii Short-Chain Acyl-CoA Dehydrogenase Active-Site Mutant. *Biochemistry* **1994**, *33*, 7082–7087.
(48) Vock, P.; Engst, S.; Eder, M.; Ghisla, S. Substrate Activation by Acyl-CoA Dehydrogenases: Transition-State Stabilization and PKs of Involved Functional Groups. *Biochemistry* **1998**, *37*, 1848–1860.
(49) Engst, S.; Vock, P.; Wang, M.; Kim, J. J.; Ghisla, S. Mechanism of Activation of Acyl-CoA Substrates by Medium Chain Acyl-CoA Dehydrogenase: Interaction of the Thioester Carbonyl with the Flavin Adenine Dinucleotide Ribityl Side Chain. *Biochemistry* **1999**, *38*, 257–267.
(50) Lau, S. M.; Brantley, R. K.; Thorpe, C. The Reductive Half-Reaction in Acyl-CoA Dehydrogenase from Pig Kidney: Studies with Thiaoctanoyl-CoA and Oxaoctanoyl-CoA Analogues. *Biochemistry* **1988**, *27*, 5089–5095.
(51) Stankovich, M. T.; Sabaj, K. M.; Tonge, P. J. Structure/Function of Medium Chain Acyl-CoA Dehydrogenase: The Importance of Substrate Polarization. *Arch. Biochem. Biophys.* **1999**, *370*, 16–21.
(52) Lehman, T. C.; Hale, D. E.; Bhala, A.; Thorpe, C. An Acyl-Coenzyme a Dehydrogenase Assay Utilizing the Ferricenium Ion. *Anal. Biochem.* **1990**, *186*, 280–284.
(53) Dorrestein, P. C.; Bumpus, S. B.; Calderone, C. T.; Garneau-Tsodikova, S.; Aron, Z. D.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Kelleher, N. L. Facile Detection of Acyl and Peptidyl Intermediates on Thiotemplate Carrier Domains via Phosphopantetheinyl Elimination Reactions during Tandem Mass Spectrometry. *Biochemistry* **2006**, *45*, 12756–12766.
(54) Cummings, J. G.; Thorpe, C. 3-Methyleneoctanoyl-CoA and 3-Methyl-Trans-2-Octenoyl-CoA: Two New Mechanism-Based Inhibitors of Medium Chain Acyl-CoA Dehydrogenase from Pig Kidney. *Biochemistry* **1994**, *33*, 788–797.
(55) Satoh, A.; Nakajima, Y.; Miyahara, I.; Hirotsu, K.; Tanaka, T.; Nishina, Y.; Shiga, K.; Tamaoki, H.; Setoyama, C.; Miura, R. Structure of the Transition State Analog of Medium-Chain Acyl-CoA Dehydrogenase. Crystallographic and Molecular Orbital Studies on the Charge-Transfer Complex of Medium-Chain Acyl-CoA Dehydrogenase with 3-Thiaoctanoyl-CoA. *J. Biochem.* **2003**, *134*, 297–304.
(56) Battaile, K. P.; Molin-Case, J.; Paschke, R.; Wang, M.; Bennett, D.; Vockley, J.; Kim, J.-J. P. Crystal Structure of Rat Short Chain Acyl-CoA Dehydrogenase Complexed with Acetoacetyl-CoA: Comparison with Other Acyl-CoA Dehydrogenases. *J. Biol. Chem.* **2002**, *277*, 12200–12207.
(57) Djordjevic, S.; Pace, C. P.; Stankovich, M. T.; Kim, J. J. Three-Dimensional Structure of Butyryl-CoA Dehydrogenase from Megasphaera Elsdenii. *Biochemistry* **1995**, *34*, 2163–2171.
(58) McMahon, B.; Gallagher, M. E.; Mayhew, S. G. The Protein Coded by the PP2216 Gene of Pseudomonas Putida KT2440 Is an Acyl-CoA Dehydrogenase That Oxidises Only Short-Chain Aliphatic Substrates. *FEMS Microbiol. Lett.* **2005**, *250*, 121–127.
(59) Watanabe, K.; Khosla, C.; Stroud, R. M.; Tsai, S. C. Crystal Structure of an Acyl-ACP Dehydrogenase from the FK520 Polyketide Biosynthetic Pathway: Insights into Extender Unit Biosynthesis. *J. Mol. Biol.* **2003**, *334*, 435–444.
(60) Chan, Y. A.; Boyne, M. T.; Podevels, A. M.; Klimowicz, A. K.; Handelsman, J.; Kelleher, N. L.; Thomas, M. G. Hydroxymalonyl-Acyl Carrier Protein (ACP) and Aminomalonyl-ACP Are Two Additional Type I Polyketide Synthase Extender Units. *Proc Natl Acad Sci USA* **2006**, *103*, 14349–14354.
(61) Cruz-Morales, P.; Ramos-Aboites, H. E.; Licona-Cassani, C.; Selem-Mójica, N.; Mejía-Ponce, P. M.; Souza-Saldívar, V.; Barona-Gómez, F. Actinobacteria Phylogenomics, Selective Isolation from an Iron Oligotrophic

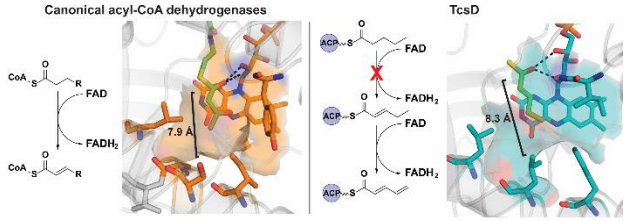
Environment and Siderophore Functional Characterization, Unveil New Desferrioxamine Traits. *FEMS Microbiol. Ecol.* **2017**, *93*.

(62) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410.

(63) Finn, R. D.; Clements, J.; Eddy, S. R. HMMER Web Server: Interactive Sequence Similarity Searching. *Nucleic Acids Res.* **2011**, *39*, W29-37.

(64) Sharma, V.; Eckels, J.; Schilling, B.; Ludwig, C.; Jaffe, J. D.; MacCoss, M. J.; MacLean, B. Panorama Public: A Public Repository for Quantitative Data Sets Processed in Skyline. *Mol. Cell. Proteomics* **2018**, *17*, 1239–1244.

For Table of Contents Only



# Supplementary Information for

## Structural mechanism of regioselectivity in an unusual bacterial acyl-CoA dehydrogenase

Jacquelyn M. Blake-Hedges,[1,2,3] Jose Henrique Pereira,[2,3,4] Pablo Cruz-Morales,[2,3] Mitchell G. Thompson,[2,3,5] Jesus F. Barajas,[2,3,6] Jeffrey Chen,[1] Rohith N. Krishna,[1] Leanne Jade G. Chan,[2,3] Danika Nimlos,[1] Catalina Alonso-Martinez,[1] Edward E. K. Baidoo,[2,3] Yan Chen,[2,3,6] Jennifer W. Gin,[2,3,6] Leonard Katz,[2,7] Christopher J. Petzold,[2,3,6] Paul D. Adams,[2,3,4] Jay D. Keasling[2,3,7,8,9,10]

[1]Department of Chemistry, University of California, Berkeley, Berkeley, CA 94720
[2]Joint BioEnergy Institute, Emeryville, CA, 94608
[3]Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
[4]Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
[5]Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, CA 94720
[6]Department of Energy Agile BioFoundry, Emeryville, CA 94608, USA
[7]QB3 Institute, University of California, Berkeley, Emeryville, CA, 94608
[8]Department of Chemical & Biomolecular Engineering, Department of Bioengineering, University of California, Berkeley, Berkeley, CA 94720
[9]Novo Nordisk Foundation Center for Biosustainability, Technical University Denmark, DK2970-Horsholm, Denmark
[10]Center for Synthetic Biochemistry, Shenzhen Institutes for Advanced Technologies, Shenzhen, China

**Table of Contents**

**Materials, Reagents, and Strains**

All strains used and generated in this work are listed in Table S2. Routine *E. coli* cultures were grown at 37°C in Luria-Bertani (LB) Miller medium (BD Biosciences, USA) and were supplemented with kanamycin (50 mg/L, Sigma Aldrich, USA), or carbenicillin (100mg/L, Sigma Aldrich, USA). *Pseudomonas putida* KT2440 was grown at 30°C in Luria-Bertani (LB) Miller medium (BD Biosciences, USA), and *Streptomyces tsukubaensis* NRRL 18488 was grown at 30°C in BD Bacto tryptic soy broth (TSB) medium (Fisher Scientific). All chemicals and reagents were purchased from Sigma Aldrich unless otherwise noted.

**DNA Manipulation**

All plasmids used and generated in this work are listed in Tables S1. The genomic DNA samples from *Pseudomonas putida* KT2440 were purified using a DNeasy Blood and Tissue Kit (Qiagen, USA). Plasmids and PCRs were routinely isolated using the Qiaprep Spin Miniprep kit (Qiagen, USA) and DNA Clean & Concentration Kit (Zymo Research, USA). All primers were purchased from Integrated DNA Technologies (IDT, Coralville, IA). Constructs were generated using restriction enzyme cloning (TcsD, TcsA wildtype coding sequences), Gibson assembly (TcsD mutants, PP2216), or Golden Gate assembly (*N. mex* ACAD and ACP) as previously described.[1,2]

**Construct Cloning: TcsD and TcsA**

Genomic DNA of *Streptomyces tsukubaensis* NRRL 18488 was prepared by resuspending cells grown in TSB medium in LC-grade DMSO. The coding sequences of TcsD and TcsA were amplified from this genomic DNA sample using Q5 Hotstart high-fidelity polymerase (New England Biolabs, USA) using primers which incorporated flanking NdeI and XhoI restriction sites. The genes were then ligated into a pET28a backbone which had been linearized with the FastDigest enzymes NdeI and XhoI (Thermo Scientific, USA), appending a C-terminal 6x histidine tag onto each. Targeted mutations were generated in TcsA and TcsD using Gibson cloning. Mutations were embedded in the overlap regions of the Gibson primers listed in Table S3**.**

**Construct Cloning: PP2216**

PP2216 was amplified from *Pseudomonas putida* KT2440 genomic DNA and cloned into a pBbE7a backbone with a C-terminal 6x histidine tag using Gibson assembly.

**Construct Cloning: *Nocardia mexicana* ACP and ACAD**

The coding sequences for the *Nocardia mexicana* NBRC 108244 putative γ,δ-acyl-CoA dehydrogenase (GenBank accession WP_068019852.1) and the neighboring acyl carrier protein (GenBank accession WP_068019854.1) were reverse translated and codon-optimized for *E. coli* using BOOST. [3] Genes were designed with flanking Golden Gate sites and purchased from Genscript (Piscataway, NJ, USA). The genes were then inserted into a pBbE7a backbone with a C-terminal 6x histidine tag using Golden Gate assembly. The sequences of the synthetic gene fragments are listed in Table S4.

**Protein Expression**

All proteins were expressed in *E. coli* BL21 (DE3) cells using the following protocol. An overnight culture of Luria Broth containing the appropriate antibiotic was used to inoculate 0.8L of Terrific Broth (EMD Millipore) in a 2L flask. Cells were grown at 37ºC to an OD of 0.6-0.8 then induced with 250 µM IPTG (Teknova, USA). Cells were grown for approximately 18 hours at 18ºC then pelleted and stored at -20°C until purification.

**Protein Purification for biochemical assays and substrate biosynthesis**

Frozen cell pellets were resuspended in lysis buffer that had been prechilled to 4°C. For TcsD, TcsD mutants, and the *Nocardia mexicana* ACAD, lysis buffer consisted of 50 mM Tris pH 9.0, 300 mM NaCl, 10 mM imidazole, 8% glycerol. For all other proteins (TcsAS98A, Sfp, PP2216, MatB T207G/M306I, *Nocardia mexicana* ACP), a lysis buffer composed of 50 mM sodium phosphate pH 7.2, 300 mM NaCl, 10 mM imidazole, 8% glycerol was used. After resuspension in lysis buffer, cells were sonicated on ice to lyse. Lysates were then centrifuged at 4°C at 40,000g for 30 minutes to pellet insolubles. The supernatants were then applied to Ni-NTA resin in an Econo-Pak column (Bio-Rad, USA) that had been pre-equilibrated with 10 column volumes of lysis buffer. The resin was washed with several column volumes of lysis buffer and subsequently with several column volumes of lysis buffer containing 50 mM imidazole. The proteins were then eluted using a stepwise gradient of lysis buffer containing 100 mM, 200 mM and 400 mM imidazole. Fractions were analyzed via SDS-PAGE, and fractions containing >90% pure protein were pooled and concentrated using Amicon Ultra 100 kDa (TcsD, TcsDmutants, and *Nocardia mexicana* ACAD), 3kDa (*Nocardia mexicana* ACP)10 kDa (all other proteins) Molecular Weight Cutoff (MWCO) centrifuge filters (Millipore). Proteins were

transferred to storage buffer (lysis buffer containing no imidazole, 100 mM NaCl, and 8-10% glycerol) by buffer exchanging in the centrifuge filters and frozen in liquid nitrogen. Proteins were stored at -80ºC until assays.

**Protein purification for crystallography**

A frozen cell pellet of *E. coli* BL21 (DE3) expressing TcsD was purified using Ni-NTA affinity chromatography as described above. After concentration to a minimal volume, the protein was further purified using size exclusion chromatography. It was loaded onto a HiPrep 26/60 Sephacryl S-300 High Resolution column that had been pre-equilibrated with size exclusion buffer (50 mM Tris pH 9, 500 mM NaCl, 10% glycerol) at 4°C. The protein was eluted and concentrated again then dialyzed overnight at 4°C into crystallization buffer (50 mM Tris pH 9, 100 mM NaCl).

**Protein Crystallization and Structure Determination**

An initial crystallization screen was set up using a Phoenix robot (Art Robbins Instruments, Sunnyvale, CA) using the sparse matrix screening method.[4] The purified TcsD protein sample was concentrated to 9 mg/mL and crystallized at 25°C using the sitting drop method in 0.4 µL drops containing a 1:1 ratio of protein to crystallization solution: 0.1 M Sodium Citrate tribasic dihydrate, pH 5.0, and 10% PEG 6,000. Crystals were transferred to crystallization solution containing 20% glycerol prior to flash freezing in liquid nitrogen. X-ray diffraction data was collected at the Berkeley Center for Structural Biology on beamline 5.0.2 of the Advanced Light Source at the Lawrence Berkeley National Laboratory. The TcsD structure was determined by the molecular-replacement method with the program *PHASER*[5] using medium-chain acyl-CoA dehydrogenase from *Thermus thermophilus* (PDB ID: 1UKW) as the search model, which showed 32% sequence identity. Structure refinement was performed by *phenix.refine* program.[6] Manual rebuilding using COOT and the addition of water molecules allowed for construction of the final model.[7] The final models of the TcsD structure showed an R-work of 14.6% and R-free of 17.9% (Table S5). Root-mean-square deviations from ideal geometries for bond lengths, angles, and dihedrals were calculated with Phenix.[8] The overall stereochemical quality of the final models for tcsD was assessed using the MolProbity program.[9] The structural analysis was performed in COOT and PyMOL.[10]

**Synthesis of Pentanoyl-Coenzyme A.**

Pentanoyl-Coenzyme A was prepared using the anhydride method.[11] Approximately 0.02 mmol of Coenzyme A was added to a solution of saturated sodium bicarbonate in water. The solution was chilled to

0°C, then 1 mmol of valeric anhydride (Sigma Aldrich) was added. The reaction was allowed to proceed at 0°C with constant mixing. After approximately 6 hours, HCl was added to bring the pH to ~2. The solution was extracted twice with an equal volume of ethyl acetate, then the remaining aqueous solution was frozen until further purification. Pentanoyl-CoA was purified using an Agilent 1260 series preparatory HPLC system equipped with a 900µL sample loop and a UV detector coupled to a fraction collector. The product was isolated over an Agilent Prep-C18 column (21.2 x 150 mm, 5 µm pore size) using a mobile phase composed of 10 mM ammonium formate, pH 4.5 (solvent A) and methanol (solvent B) using the following method:

| **Time (min)** | **%A** | **%B** | **Flow rate (mL/min)** |
|---|---|---|---|
| 0 | 95 | 5 | 10 |
| 1 | 95 | 5 | 10 |
| 8 | 5 | 95 | 10 |
| 10 | 5 | 95 | 10 |
| 11 | 95 | 5 | 10 |
| 16 | 95 | 5 | 10 |

The fraction collector was programmed to collect all peaks that absorbed at 280 nm. Fractions containing pentanoyl-CoA were identified by direct infusion into an Applied Biosystems 4000 QTRAP mass spectrometer with the following mass spectrometer parameters: Q1 MS mode, scan range 100-1000 m/z, declustering potential: 70, entrance potential: 10, curtain gas: 10, IonSpray Voltage: 4800, Temperature: 300, Ion Source Gases: 40. The fractions containing the correct m/z for pentanoyl-Coenzyme A ([M+H]) were then combined and lyophilized, leaving behind pentanoyl-CoA as a white powder. The product was resuspended in LC grade water and quantified using the absorbance at 280 nm compared to a standard curve of hexanoyl-CoA in water (Sigma-Aldrich) measured on a Nanodrop ND-1000 Spectrophotometer. The identity of the Coenzyme A ester was additionally verified by targeted LC-MS/MS after loading the purified substrate onto TcsA S98A (methods described below).

**Trans-2-pentenoyl-, trans-2-hexenoyl-, and trans-2-heptenoyl-Coenzyme A.**

The 2,3-unsaturated CoA substrates were generated *in vitro* from pentanoyl-CoA (synthesized in this work), hexanoyl-CoA (Sigma Aldrich), or heptanoyl-CoA (CoALA Biosciences, USA) using the short chain acyl-CoA dehydrogenase PP2216.[12] Reactions contained the following components: 100mM sodium phosphate (pH

7.2), 500 µM Acyl-Coenzyme A, 1 mM ferrocenium hexafluorophosphate (FeHFP, Sigma Aldrich), 225 µM PP2216 in a final volume of 1 mL. FeHFP was always prepared fresh by dissolving solid FeHFP in 10mM HCl and sonicating in a water bath to dissolve solid particles.[13] Reactions were incubated at 25°C for 2 hours (2-pentenoyl-CoA) or 6 hours (2-hexenoyl and 2-heptenoyl-CoA) then filtered through 3kDa MWCO Amicon microcentrifuge filters (Sigma Aldrich) to remove the remaining PP2216 protein. Samples were then aliquoted and lyophilized. After lyophilization, samples were stored at -20°C and resuspended in HPLC water (Honeywell) at a concentration of 4mM (assuming maximum theoretical yield) immediately prior to use.

**Allylmalonyl- and propylmalonyl-Coenzyme A.**

Allylmalonyl-CoA and propylmalonyl-CoA were biosynthesized from the corresponding malonic acids using an engineered mutant of the *Streptomyces coelicolor* malonyl-Coenzyme A transacylase MatB, MatB T207G/M306I.[14] Reactions contained the following components: 100 mM sodium phosphate (pH 7.2), 2 mM $MgCl_2$, 8 mM ATP, 4 mM Coenzyme A, 4 mM diacid, 100 µM MatB T207G/M306I in a final volume of 0.5 mL. Prior to addition to assays, allylmalonic acid (Sigma Aldrich) and propylmalonic acid (VWR, BeanTown Chemical) were dissolved in 50 mM sodium phosphate buffer, pH 7.2. Reactions were incubated for 3 hours the filtered through 3kDa MWCO Amicon microcentrifuge filters (Sigma Aldrich) to remove the remaining MatB protein. Mixtures were aliquoted and frozen at -20°C until use in TcsD assays.

**Enzyme Assays: TcsD and mutants**

Assays of TcsD activity on TcsAS98A-bound substrates were carried out as follows: First, purified TcsAS98A was loaded with acyl-Coenzyme A substrates via the action of the promiscuous phosphopantetheinyl transferase Sfp.[15] Loading reactions (150 µL) contained 60 µM TcsAS98A, 10 µM Sfp, 250 µM acyl-Coenzyme A, 20 mM $MgCl_2$, and 50 mM sodium phosphate, pH 7.2. Reactions were initiated by the addition of Sfp and allowed to proceed at 37°C for 30 minutes. The loaded protein was then immediately used for TcsD assays, each of which contained: 45 µL of loading reaction, 0.5 mM FeHFP, and 2 µM TcsD mutant (or boiled TcsD mutant). FeHFP was prepared as described above. Reactions were allowed to proceed for 16 hours then quenched with and equal volume of HPLC methanol (Sigma Aldrich) and either stored at -20°C or immediately processed for LC-MS analysis.

**Enzyme Assays: *Nocardia mexicana* ACAD**

The Nocardia mexicana ACP was loaded with butyrl-, pentanoyl-, 2-pentenoyl-, 2-hexenoyl, or 2-heptenoyl-CoA using Sfp.[15] Loading reactions (120 µL) consisted of 50 µM *N. mexicana* ACP, 10 µM Sfp, 250 µM acyl-Coenzyme A, 20 mM $MgCl_2$, and 40 mM Tris buffer, pH 8. Reactions were initiated by the addition of Sfp and allowed to proceed at 37°C for 1 hour. The loading reactions were then used for *N. mexicana* ACAD assays which consisted of: 40 µL of loading reaction, 0.5 mM FeHFP, and 2 µM *N. mexicana* ACAD. Reactions were allowed to proceed at 37°C for 16 hours then quenched with HPLC methanol and immediately processed for LC-MS analysis.

**Sample preparation for LC-MS/MS analysis of TcsD and Nmex-ACAD assays**

Samples were prepared for proteomics analysis by chloroform/methanol extraction as previously reported.[16] Briefly, 150 µL of LC-grade methanol and 50 µL of LC-grade chloroform were sequentially added to the methanol-quenched samples, and samples were vortexed after each addition. Then, 150 µL LC-grade water was added, samples were vortexed to mix, and they were subsequently centrifuged at 21,000g for 1 minute to promote phase separation. After removal and disposal of the top (aqueous) layer, another 150 µL of LC methanol was added, samples were vortexed, then they were centrifuged again at 21,000g for 2 minutes. After removal of the supernatant, samples were allowed to dry for 5 minutes in a fume hood. Then, they were resuspended in freshly-prepared 100 mM ammonium bicarbonate buffer (*N. mexicana* ACAD assays) or 100 mM ammonium bicarbonate containing 20% LC-grade DMSO (TcsD assays). Protein concentrations were analyzed via the DC Assay (Bio-Rad). After quantification, samples were digested with an appropriate protease. Samples digested with only trypsin (Sigma-Aldrich) or only chymotrypsin (Promega) were digested at a ratio of 1:50 w/w protease:protein sample at 37°C for at least 6 hours. For Asp-N/trypsin dual digestions (TcsD and TcsD mutant assays), proteins were digested with 1:25 w/w Asp-N (New England Biolabs) in the presence of 0.75 mM $ZnSO_4$ for 2-4 hours at 37°C then immediately digested again with 1:10 w/w trypsin for 1 hour at 37 $^{o}$C. The *N. mexicana* ACAD assays were digested with Glu-C (Promega) at a ratio of 1:50 protease:protein sample at 37°C for 6 hours. After digestion, the samples were either frozen for future analysis or directly analyzed via LC-MS/MS.[17]

**Shotgun LC-MS/MS Analysis of TcsA-bound substrates**

Samples prepared for shotgun proteomic analysis were analyzed using Agilent 6550 iFunnel Q-TOF mass spectrometer (Agilent Technologies, Santa Clara, CA) coupled to an Agilent 1290 UHPLC system. Twenty (20)

µg of peptides were separated on a Sigma–Aldrich Ascentis Peptides ES-C18 column (2.1 mm × 100 mm, 2.7 µm particle size, operated at 60°C) at a 0.400 mL/min flow rate and eluted with the following gradient: initial condition was 95% solvent A (0.1% formic acid) and 5% solvent B (99.9% acetonitrile, 0.1% formic acid). Solvent B was increased to 35% over 5.5 min, and then increased to 80% over 1min, and held for 3.5 min at a flow rate of 0.6 mL/min, followed by a ramp back down to 5% B over 0.5 min where it was held for 2 min to re-equilibrate the column to original conditions. Peptides were introduced to the mass spectrometer from the LC by using a Jet Stream source (Agilent Technologies) operating in positive-ion mode (3,500 V). Source parameters employed gas temp (250°C), drying gas (14 L/min), nebulizer (35 psig), sheath gas temp (250°C), sheath gas flow (11 L/min), VCap (3,500 V), fragmentor (180 V), OCT 1 RF Vpp (750 V). The data were acquired with Agilent MassHunter Workstation Software, LC/MS Data Acquisition B.06.01 operating in Auto MS/MS mode whereby the 20 most intense ions (charge states, 2–5) within 300–1,400 m/z mass range above a threshold of 1,500 counts were selected for MS/MS analysis. MS/MS spectra (100–1,700 m/z) were collected with the quadrupole set to “Medium” resolution and were acquired until 45,000 total counts were collected or for a maximum accumulation time of 333 ms. Former parent ions were excluded for 0.1 min following MS/MS acquisition. The acquired data were exported as mgf files and searched against a customized protein database with Mascot search engine version 2.3.02 (Matrix Science).

**Targeted LC-MS/MS Analysis**

Samples were analyzed using an Agilent 1290 Infinity II liquid chromatography system coupled to an Agilent 6460 QQQ mass spectrometer (Agilent Technologies, Santa Clara, CA). Peptide samples (10 µg) were separated on an Ascentis Express Peptide ES-C18 column (2.7 µm particle size, 160 Å pore size, 50 x 2.1mm) fitted with a guard column (5 mm x 2.1 mm, Sigma Aldrich). The column was heated to 60°C. The mobile phase consisted of 0.1% formic acid in $H_2O$ (A) and 0.1% formic acid in acetonitrile (B). Peptides were eluted via the following gradient method:

| Time (min) | %A | %B | Flow rate (mL/min) |
|---|---|---|---|
| 0 | 95 | 5 | 0.4 |
| 2.2 | 95 | 5 | 0.4 |
| 7.7 | 65 | 35 | 0.4 |
| 8 | 20 | 80 | 0.4 |
| 10 | 20 | 80 | 0.4 |
| 10.5 | 95 | 5 | 0.4 |
| 12 | 95 | 5 | 0.4 |
| 15 | 95 | 5 | 0.4 |

Peptides were ionized using an Agilent Jet Stream ESI source operating in positive-ion mode with the following source parameters: Gas Temperature = 250°C, Gas Flow = 13 L/min, Nebulizer Pressure = 35 psi, Sheath Gas Temperature = 250°C, Sheath Gas Flow = 11 L/min, and Capillary Voltage = 3,500 V. Dwell times were set to 10ms. Data were acquired using Agilent MassHunter Data Acquisition (Version B.08.02, Build 8.2.8260.0). The mass spectrometry method was built using the Skyline targeted mass spectrometry environment (version 4.2.0 19072).[18] Collision energies were predicted by Skyline with the exception of phosphopantetheine-bearing peptides for which the collision energy was set to 30eV. The specific transitions monitored and the collision energies for each are listed in Tables S7 and S8.

**Targeted LC-MS/MS Data Availability**

Raw data were imported into Skyline and are available on the LC-MS data sharing platform Panorama Public at the following link: https://panoramaweb.org/Structural%20control%20of%20bacterial%20acyl-CoA%20dehydrogenase.url.[19] Sample names for TcsD assays are labeled as “protein-substrate-replicate.” For example, replicate one of the assay of wild type TcsD on pentanoyl-ACP is labeled “WT-val-01.” Substrate abbreviations are as follows: val = pentanoyl-ACP, but = butyryl-ACP, 2pent = 2-pentenoyl-ACP, 2hex = 2-hexenoyl-ACP, 2hept = 2-heptenoyl-ACP, allylmal = allylmalonyl-ACP, propmal = propylmalonyl-ACP. The data for TcsD wild type assays on 2-pentenoyl-ACP, 2-hexenoyl-ACP, and 2-heptenoyl-ACP is located in the file named “WT assays_set 1_pub data.skyd.” The data for TcsD wild type assays on propylmalonyl-ACP is located in the file named “Malonate assays_pub data.skyd.” The data for TcsD wild type assays on butyryl-

ACP and pentanoyl-ACP and the data for all mutant TcsD proteins on all substrates is located in the file named "Mutants assays_WT-but-val_pub data.skyd."

Sample names for *N. mexicana* ACAD assays are labeled as "substrate_replicate," and substrate abbreviations are the same as for TcsD assays. The data for butyryl-, pentanoyl-, and 2-pentenoyl-ACP are located in the file named "NmexACP_C4_C5_pub data.skyd," and data for 2-hexenoyl- and 2-heptenoyl-ACP are located in the file named "NmexACP_C6_C7_pub data.skyd."**Targeted LC-MS/MS Data Analysis: TcsD assays**

After importing data, chromatograms were manually curated. Correct peaks were identified by comparison of their retention times to control peaks (i.e. the 2-pentenoyl-ACP active site peptide should elute before the pentanoyl-ACP active site peptide). For phosphopantetheine ejection peaks, only the transition derived from the peptide "DLGVDSLAMTELQAHALQR" was used for analysis and quantification. Peaks that were not correctly automatically selected were manually selected in Skyline and integrated using the built in Skyline integration function. For negative controls, an area corresponding to the same retention time as the peak of interest was integrated. After integrating, all phosphopantetheine ejection ion peak areas were normalized to an internal control peptide from the within the TcsA protein (a peptide that does not contain catalytic residues) to account for differences in the amount of protein injected. The reported values in Figure 3 were averages of three normalized replicate peak areas. Error bars represent standard deviations of each set of replicates.

**Targeted LC-MS/MS Data Analysis: *N. mexicana* ACAD assays**

After importing data into Skyline, chromatograms were manually curated (including manual integration where necessary) as described above for TcsD assays. Plots in Figures 5E and 5F and Figure S13 correspond to the LC-MS/MS chromatograms of the phosphopantetheine ejection transition from the parent peptide "MDSLNLMDFLVYE" in the +2 charge state.

**Preparation of denatured TcsD supernatants for untargeted LC-MS analysis**

TcsD was expressed and purified via nickel affinity chromatography as described above. The highest purity nickel affinity elution fractions were pooled, and protein was concentrated to approximately 17 mg/mL or 375 µM (determined by absorbance using a Nanodrop and the molar absorbance of TcsD). After concentration, 150 µL of protein was denatured by either boiling for 10 minutes or the addition of 300 µL of LC grade acetonitrile. The denatured protein was pelleted by centrifuging at maximum speed in a benchtop centrifuge for

2 minutes. Supernatants were removed and lyophilized overnight. After lyophilization, supernatants were resuspended in 75 µL of LC grade water for an effective metabolite concentration of 750 µM. Samples were diluted to an effective metabolite concentration of 50 µM in 50:50 (v/v) LC grade water:methanol prior to untargeted LC-MS analysis. Standards of flavin adenine dinucleotide (FAD), Coenzyme A, pantetheine, pantothenate, and butyryl-CoA were prepared in 50:50 (v/v) LC grade methanol:water to a concentration of 20 µM.

**High resolution untargeted LC-MS analysis of denatured TcsD supernatants**

Supernatants of denatured TcsD samples were analyzed using an Agilent 6545 LC-QTOF system as previously described. [20] Analytes were separated using a SeQuant ZIC hydrophilic interaction chromatography (HILIC) column (150 mm length, 4.6 mm internal diameter, 5 µm particle size) coupled to a SeQuant ZIC-pHILIC guard column (20 mm length, 2.1 mm internal diameter, 5 µm particle size). The mobile phase consisted of 10 mM ammonium carbonate and 118.4 mM ammonium hydroxide in acetonitrile-water (60.2:39.8 v/v). Analytes were separated isocratically using the following pump parameters:

| **Time (min)** | **Mobile phase (%)** | **Flow rate (mL/min)** |
|---|---|---|
| 0 | 100 | 0.45 |
| 6 | 100 | 0.45 |
| 6.5 | 100 | 0.605 |
| 12.5 | 100 | 0.605 |

The mass spectrometer source was operated with the following parameters: Gas Temperature = 300°C, Drying Gas = 10 L/min, Nebulizer Pressure = 20 psi, Sheath Gas Temperature = 350°C, Sheath Gas Flow = 12 L/min, and Capillary Voltage = 3500 V, Nozzle Voltage = 2000 V. The TOF mass spectrometer was programmed to operated in scan mode with the following parameters: Fragmentor Voltage = 100 V, Skimmer Voltage = 50 V, Oct 1 RF Vpp = 400 V, Mass Range = 70 m/z – 1100 m/z, Acquisition Rate = 0.86 spectra/s, Acquisition Time = 1162.8 ms/spectrum, and Transients/spectrum = 9532 (negative ion mode) or 9485 (positive ion mode). All samples were analyzed in both negative ion mode and subsequently in positive ion mode. Data analysis was performed in Agilent MassHunter Qualitative Analysis (v B.05.00).

**Substrate modeling into TcsD active site**

Substrates for modeling were drawn using ChemDraw, and ligand restraints were generated using eLBOW in the Phenix software suite. Substrates were manually placed into the TcsD active site using Coot. Real Space Refinements of the substrates were performed using the unknown density in the substrate-binding region of the enzyme active site. The refinements resulted in the generation of both *cis* and *trans* isomers of the substrate with planar fatty acyl tails. After substrate generation, the TcsD structure was aligned with the structures of ACADs which had been co-crystallized with thioester substrates. Alignments were performed using the least squares fit LSQ superpose function of Coot. Proteins were aligned using the 3-residue loop containing the catalytic glutamate and the residues immediately upstream and downstream of the glutamate (Table S6). These alignments were used for analysis of the relative positioning of FAD cofactors and for substrate modeling. After alignment, the 2-pentenoyl substrates were manually translated and rotated into a position within the TcsD active site that corresponds to approximately the equivalent position of substrates in the aligned homolog structures. The feasibility of hydrogen bonding interactions was confirmed using the Measure tool in Coot considering a maximum distance of 3.5Å between heteroatoms. Substrate lengths were also determined using the Measure tool to calculate the distance from the carbonyl carbon of a given molecule to the distal carbon of the fatty acyl tail. The length of the 2-hexenoyl substrate was calculated using a substrate molecule that was generated using Ligand Builder in Coot. The approximate length of the 2-heptenoyl substrate was calculated by adding half the distance between the fourth and sixth (saturated) carbons of the 2-hexenoyl substrate to the total length of the 2-hexenoyl substrate.

**Bioinformatic identification of TcsD homologs**

First, a Hidden Markov Model (HMM) that describes γ,δ-ACADs was generated using the sequences of several TcsD enzymes from various FK506 producers and HliR.[21] The HMM was then used to search for homologs among all proteins in the Uniprot database, after which the genomes of organisms that contain potential γ,δ-ACADs were analyzed using antiSMASH.[22] An additional search was performed by querying the results of the HMM search against GenBank using protein BLAST.[23] Finally, CORASON-BGC was used for the identification of the syntenic genomic contexts of the identified genes.[24] The results from each of these searches were then manually curated based on the presence of key motifs identified through the structural and biochemical characterization of TcsD.

**Sequence alignments of TcsD and homologs**

A sequence alignment of TcsD, HliR, and several α,β-ACADs was generated using Clustal Omega.[25] The sequence alignment was visualized with respect to the TcsD structure using ESPript 3.0.[26] To calculate the positional occupancy of amino acids within the γ,δ-ACADs, a sequence alignment of the identified proteins was generated using MAFFT.[27] Positional occupancies were calculated using the bioinformatics tool UGene.[28]

**Position weight matrix**

The Shannon entropy of the aligned γ,δ-ACADs was calculated at each position within a sequence alignment. All of the identified γ,δ-ACAD sequences were aligned using MAFFT.[27] The MSA was then refined so that all positions maintained at least 10% occupancy using the ProDy python library. The Shannon entropy was then calculated using ProDy.[29,30] The sequence logo was generated using Weblogo (https://weblogo.berkeley.edu/logo.cgi).

## Supplementary Discussion

### LC-MS/MS Phosphopantetheine ejection assay method development

In order to assay TcsD activity on TcsA-bound substrates, we first developed a targeted proteomics assay for the identification of substrates tethered to the phosphopantetheinyl arm of the acyl carrier proteins (ACP) domain of TcsA. In a typical "phosphopantetheine ejection" experiment,[17] a protein of interest is digested with trypsin and the digested peptides are analyzed via LC-MS/MS (main text Figure 2A). However, after digesting TcsA with trypsin we were unable to detect an ion representing the ACP active site peptide, most likely because the large size (~40 residues) of the tryptic peptide placed it outside of the detectable range of our mass spectrometer (Figure S17). Similarly, the active site peptide could not be detected after digestion with another commonly-used protease, chymotrypsin. In order to obtain an appropriately-sized peptide, we performed a sequential digestion of TcsA with two proteases, Asp-N then trypsin, and identified the resulting "holo" and "acyl" active site peptides using high resolution untargeted LC-MS (Figure S17 and S18). Surprisingly, Asp-N did not consistently cleave at the predicted residue D373 of the conserved "DSL" motif of the acyl carrier protein (ACP) of TcsA but instead at the preceding D369, possibly due to steric effects arising from the presence of the phosphopantetheine arm appended to S374.

### Unidentified density in active site of TcsD crystal structure

An unknown metabolite consistently co-crystallized in the active site, occupying the space where substrates are predicted to bind (Figure S14). We attempted to identify this unknown density by performing untargeted high resolution LC-MS of denatured protein samples (Figures S15 and S16), but were unable to identify any unique masses in the supernatants of denatured TcsD samples. We also compared the denatured protein supernatant LC-MS chromatograms to several standards, including flavin adenine dinucleotide (FAD), Coenzyme A, pantetheine, pantothenate, and butyryl-CoA, but only the m/z corresponding to FAD could be detected in extracted ion chromatograms of either sample. We additionally used the Ligand Identification function of the Phenix software suite,[31,32,33] attempting to model 180 of the most common ligands from the Protein Data Bank into the density, but none of the ligands fit appropriately.

## Supplementary Data Tables

**Table S1**: Plasmids used in this study

| ICE Entry Number | Plasmid name | Description | Source |
|---|---|---|---|
| JPUB_013738 | pET28a-TcsA | TcsA in pET28a | This study |
| JPUB_013740 | pET28a-TcsAS98A | TcsA S98A in pET28a | This study |
| JPUB_013728 | pET28a-TcsD | TcsD in pET28a | This study |
| JPUB_013730 | pET28a-TcsDL83A | TcsD L83A in pET28a | This study |
| JPUB_013732 | pET28a-TcsDF79A | TcsD F79A in pET28a | This study |
| JPUB_013734 | pET28a-TcsDI363A | TcsD I363A in pET28a | This study |
| JPUB_013736 | pET28a-TcsDF79AL83A | TcsD F79AL83A in pET28a | This study |
| JPUB_013744 | pET28a-TcsDF79AI363A | TcsD F79AI363A in pET28a | This study |
| JPUB_013742 | pBbE7a-PP2216-6xHis | PP2216 in pBbE7a with a C-terminal 6xHis tag | This study |
| JPUB_013842 | pBbE7a-NmexACP-6xHis | Acyl carrier protein from *Nocardia mexicana* NBRC 108244 (*E. coli* codon-optimized) in pBbE7a with a C-terminal 6xHis tag | This study |
| JPUB_013844 | pBbE7a-NmexACAD-6xHis | Acyl-CoA dehydrogenase from *Nocardia mexicana* NBRC 108244 (*E. coli* codon-optimized) in pBbE7a with a C-terminal 6xHis tag | This study |
| N/A | pET-Sfp | Sfp in a pET vector | Reference 34 |
| N/A | pLK54 | MatB mutant | Reference 14 |
| N/A | pBbE7a-RFP | RFP behind a T7 promoter in a BglBrick plasmid | Reference 35 |

**Table S2**: Strains used in this study

| ICE Entry Number | Strain name | Description/Genotype | Source |
|---|---|---|---|
| N/A | *E. coli* DH5α | F⁻ *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20* φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17($r_K^-m_K^+$), λ⁻ | QB3 Macrolab (http://qb3.berkeley.edu/macrolab/) |
| N/A | *E. coli* NEB Turbo | F' *proA⁺B⁺ lacI^q* Δ*lacZM15 / fhuA2* Δ*(lac-proAB) glnV galK16 galE15 R(zgb-210::Tn10)*Tet^S *endA1 thi-1* Δ*(hsdS-mcrB)5* | New England Biolabs |
| N/A | *E. coli* BL21 (DE3) | F⁻ *ompT gal dcm lon hsdS_B*($r_B^-m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB⁺*]$_{K-12}$(λ^S) | New England Biolabs |
| N/A | *Streptomyces tsukubensis* NRRL 18488 | Wild type strain | US Department of Agriculture, Agricultural Research Service Culture Collection (NRRL), Patent Collection |
| N/A | *Pseudomonas putida* KT2440 (ATCC 47054) | Wild type strain | American Type Culture Collection (ATCC) |
| JPUB_013737 | JBEI-107002 | *E. coli* DH5α containing pet28a-TcsA | This study |
| JPUB_013739 | JBEI-107003 | *E. coli* DH5α containing pet28a-TcsA S98A | This study |
| JPUB_013727 | JBEI-106997 | *E. coli* NEB Turbo containing pet28a-TcsD | This study |
| JPUB_013729 | JBEI-106998 | *E. coli* DH5α containing pet28a-TcsD L83A | This study |
| JPUB_013731 | JBEI-106999 | *E. coli* DH5α containing pet28a-TcsD F79A | This study |
| JPUB_013733 | JBEI-107000 | *E. coli* DH5α containing pet28a-TcsD I363A | This study |
| JPUB_013735 | JBEI-107001 | *E. coli* DH5α containing pet28a-TcsD F79AL83A | This study |
| JPUB_013743 | JBEI-107005 | *E. coli* DH5α containing pet28a-TcsD F79AI363A | This study |
| JPUB_013741 | JBEI-107004 | *E. coli* DH5α containing pBbE7a-PP2216-6xHis | This study |
| JPUB_013841 | JBEI-133594 | *E. coli* DH5α containing pBbE7a-NmexACP-6xHis | This study |
| JPUB_013843 | JBEI-133595 | *E. coli* DH5α containing pBbE7a-NmexACAD-6xHis | This study |

**Table S3**: Primers used in this work

| **Primer name** | **Sequence (5'-->3')** |
|---|---|
| TcsA-NdeI-F | TATACATATGGCGTTCCTCTTCCCCGGC |
| TcsA-XhoI-R | TATACTCGAGCGCCGCCCGGAAACGGAA |
| TcsD-NdeI-F | AAACATATGAGCGAATCCGAACGCCTCGGTA |
| TcsD-XhoI-R | TTTCTCGAGGGTACGTTTCGCGGTGGGGACG |
| TcsA-AT0-F-1 | TCGCCGGGCACGCTCTCGGGGA |
| TcsA-AT0-R-1 | CCACTGGTAACAGGATTAGCAGAGCGAGGTATGT |
| TcsA-AT0-F-2 | CACCGCCTACATACCTCGCTCTGCTAATCC |
| TcsA-AT0-R-2 | GCTCCGAACTCCCCGAGAGCGTGC |
| TcsD-mut-F | CTCTGTAGCACCGCCTACATACCTCGC |
| TcsD-mut-R | ACTGGTAACAGGATTAGCAGAGCGAGG |
| TcsD-F79A-F | GGTGGCGGCCACCCTGTTTCTGC |
| TcsD-F79A-R | AGGACGGGCAGAAACAGGGTGGCCGCC |
| TcsD-L83A-F | ACCCTGTTTGCGCCCGTCCTGAC |
| TcsD-L83A-R | GCTGGTCGTCAGGACGGGCGCAAACAGGGT |
| TcsD-I363-bb-F | GACCCAGAGCGCTGCCGGCACCT |
| TcsD-I363-bb-R | CTCGTAGGACAGGTGCCGGCAGCGC |
| TcsD-I363A-F | CACGCTTCGATCGCCGAGGGCGGCGACGAC |
| TcsD-I363A-R | GTCGTCGCCGCCCTCGGCGATCGAAGCGTG |
| PP2216-E7a-F | CAAAAGATCTTTTAAGAAGGAGATATACATATGCTGGTAAATGACGAGCAACAAC |
| PP2216-E7a-R | CTCGAGCCGCCGCCCGGAAACGGAACCGGTAAGATTGCGCGCAATGACCATGC |
| E7a-link-His-F | GGATCTGGCACTGGTAGT |
| E7a-link-His-R | CATATGTATATCTCCTTCTTAAAAGATCT |
| E7a-link-GG-F | CAGGTCTCAGGATCTGGCACTGGTAGT |
| BioB-GG-R | CAGGTCTCACATATGTATATCTCCTTCTTAAAAGATCTT |

**Table S4**: Synthetic DNA fragments used in this work

| **DNA fragment name** | **Sequence (5'-->3')** |
|---|---|
| *Nocardia mexicana* ACP | caggtctcatATGAACGCGATAACTACAACCGAGATCGCCGAGGGCCTGCGGAGCATCGCCGAT<br>CGTCTGGACCTGGAACTCGAGAATGTCGATATTTCCGCGACGTCGTCGCTGGAAGACGACC<br>TCGAGATGGACTCGCTGAACCTGATGGACTTCCTGGTGTATCTCGAGAAGAGATATCACGT<br>GCAGGTCACCGACGAGCGCCTGCACGAGGTCGACACCATCGGCGATGTCGTGAACCTGCT<br>CAACGACCTGCTGGCATCGTCGCGCGACACCCCGCAGCCGTCCCGAGTCGGAAGTCGTgg<br>atatgagacctg |
| *Nocardia mexicana* ACAD | caggtctcatATGGGCGACATTGTCGAGCAGGCCCGCCATTTTGGACGTACAGTTCTAGCTGCT<br>GCCCCAGCACCTGATGTAGAAACCGTCTACGACGACCGGCACCCACTGTGGGAGCAGTTC<br>CGATCGGCCGACCTCGCCGACTGGTGGGTGCCCGCCGAATACGGCGGACGGGGTGTCG<br>GACTGTGCGAGTCGGTGAACGTCGTCTCCGAACTCTCCTACCACGACGCCGGATTCGCGT<br>TCGCCGCGTTTCTGCCGATCCTCGCGTCACGAATGCTCGAGCTCTACGGTCCCGAGGAGT<br>TGGCGCGCCGCTACTTGGCGGAGATGGCCACCCATGGATCCTTTGCAGCAGCGTTGGGA<br>AGTGAAGCTGAAGCTGGAAGTGAATTAGCTAGAACGCAGACCACGTTCCGCCGCGACGGC<br>GATGTGCTGCACATCAACGGCGATAAGCAGTTCTCGACCAACTTGGCGTTCGCGCGGTTC<br>TGCCTCGTGCTGGCCCGCGACGTCGACAACCCCCGAGATTTCGCCCTGATCCTGGTCCCG<br>TCGGACAGTCCGGGATTTGTTGTCGGCCAACGGTGGCCGATGTCGGGCCTGCACGGTAC<br>TGCCACGTACCCGGCAACTTTCACCGATTGCGTCGTCCCGGCAGCAAATCGGTTGTCGGG<br>CAACGGTATTCGGATCCTGGAGGTCGGGCTCGACGCGAGCCGGATCCTGATGGCCTCGA<br>TCGCGATCGGTCTGACCCGGCGGATTCGGGATCTGAGCATGGACTACGCGGCGAGCAAG<br>CGGCTCGGTGGGCAACCGCTGAACCGCAACGCCGTCTTCGGCGCCCGCATGGGTCAGCT<br>CGAGATGGAACTCGAAACGATGAAGGCCCAGTGCCGGTGCGCCGCGGCCGAATACGACG<br>ACATCTACCAGCGATCCGATCGGGCGGCGGTGTTCTATGCCGACGGCGTGCTGAAGTCG<br>GCGATTGTGGCCAAGATGCATTGTGGTCAGGTCGGCTGGCGGGTCGCGAGTCGGGCGTC<br>GGAGGCGTTCGGCGGTCTGGGCTACACCGCGGGCCACGACATCCAGCGGTGCCTGCGG<br>GACATGCGGCACATCGCGATCGTCGAGGGCGGCGACGACGTGCTGCGTGAACTGCTCTA<br>CGGACGCTACGTCAAGCGCGCATCGCGACGAGGAggatatgagacctg |

**Table S5:** - Summary of crystal parameters, data collection, and refinement statistics. Values in parentheses are for the highest resolution shell.

| **Crystal parameters** | |
|---|---|
| Space group | P 1 21 1 |
| Unit cell | 69.52 100.06 148.51 90 100.31 90 |
| **Data collection statistics** | |
| Wavelength (Å) | 1.00000 |
| Resolution range (Å) | 68.4 - 1.75 (1.813 - 1.75) |
| Total reflections | 835542 (85389) |
| Unique reflections | 196484 (19281) |
| Multiplicity | 4.3 (4.4) |
| Completeness (%) | 97.65 (96.41) |
| Mean I/sigma(I) | 8.99 (1.39) |
| Wilson B-factor | 19.7 |
| R-merge | 0.1119 (1.063) |
| R-meas | 0.1277 (1.207) |
| R-pim | 0.06086 (0.5664) |
| CC1/2 | 0.997 (0.524) |
| CC* | 0.999 (0.829) |
| **Refinement and model statistics** | |
| Reflections used in refinement | 196470 (19281) |
| Reflections used for R-free | 9864 (1003) |
| R-work | 0.1462 (0.2810) |
| R-free | 0.1787 (0.3138) |
| CC(work) | 0.974 (0.807) |
| CC(free) | 0.959 (0.765) |
| Number of non-hydrogen atoms | 13452 |
| macromolecules | 11781 |
| ligands | 212 |
| solvent | 1459 |
| Protein residues | 1546 |
| RMS(bonds) | 0.012 |
| RMS(angles) | 0.98 |
| Ramachandran favored (%) | 99.28 |
| Ramachandran allowed (%) | 0.72 |
| Ramachandran outliers (%) | 0 |
| Rotamer outliers (%) | 0.08 |
| Clashscore | 3.04 |
| Average B-factor | 27.08 |
| macromolecules | 26.03 |
| ligands | 17.5 |
| solvent | 36.93 |
| Number of TLS groups | 31 |

**Table S6:** ACAD structures used in alignment for substrate modeling

| PDB code | Host organism | Protein description | Amino acids used for LSQ alignment |
|---|---|---|---|
| 6U1V | *Streptomyces tsukubensis* NRRL 18488 | TcsD | Ile 363 |
| | | | Glu 364 |
| | | | Gly 365 |
| 1jqi | *Rattus norvegicus* | Rat short chain ACAD | Tyr 367 |
| | | | Glu 368 |
| | | | Gly 369 |
| 2vig | *Homo sapiens* | Human short chain ACAD | Tyr 391 |
| | | | Glu 392 |
| | | | Gly 393 |
| 1udy | *Sus scrofa* | Pig medium chain ACAD | Tyr 375 |
| | | | Glu 376 |
| | | | Gly 377 |
| 1buc | *Megasphaera elsdenii* | Bacterial butyryl-CoA ACAD | Tyr 366 |
| | | | Glu 367 |
| | | | Gly 368 |
| 3mdd | *Sus scrofa* | Pig medium chain ACAD | Tyr 375 |
| | | | Glu 376 |
| | | | Gly 377 |

**Table S7**: Transitions used for the targeted LC-MS/MS-based detection of protein-bound intermediates

| Protein | Peptide | Precursor m/z | Precursor Charge | Product m/z | Product Charge | Collision Energy (eV) |
|---|---|---|---|---|---|---|
| tr\|E9KTG6\|E9KTG6_9ACTN_TcsAS98A | APPPTGGMLAVK | 569.818084 | 2 | 970.539014 | 1 | 18.7 |
| tr\|E9KTG6\|E9KTG6_9ACTN_TcsAS98A | APPPTGGMLAVK | 569.818084 | 2 | 873.48625 | 1 | 18.7 |
| tr\|E9KTG6\|E9KTG6_9ACTN_TcsAS98A | APPPTGGMLAVK | 569.818084 | 2 | 430.302396 | 1 | 18.7 |
| tr\|E9KTG6\|E9KTG6_9ACTN_TcsAS98A | AGELIAAAR | 436.253434 | 2 | 501.314357 | 1 | 14.5 |
| tr\|E9KTG6\|E9KTG6_9ACTN_TcsAS98A | AGELIAAAR | 436.253434 | 2 | 388.230293 | 1 | 14.5 |
| tr\|E9KTG6\|E9KTG6_9ACTN_TcsAS98A | AGELIAAAR | 436.253434 | 2 | 317.193179 | 1 | 14.5 |
| tr\|E9KTG6\|E9KTG6_9ACTN_TcsAS98A | DASALYATTMR | 600.289887 | 2 | 742.355236 | 1 | 19.6 |
| tr\|E9KTG6\|E9KTG6_9ACTN_TcsAS98A | DASALYATTMR | 600.289887 | 2 | 579.291907 | 1 | 19.6 |
| tr\|E9KTG6\|E9KTG6_9ACTN_TcsAS98A | DASALYATTMR | 600.289887 | 2 | 508.254793 | 1 | 19.6 |
| sp\|P39135\|SFP_BACSU | TKPISLEIAK | 550.339706 | 2 | 870.529495 | 1 | 18.1 |
| sp\|P39135\|SFP_BACSU | TKPISLEIAK | 550.339706 | 2 | 573.360639 | 1 | 18.1 |
| sp\|P39135\|SFP_BACSU | TKPISLEIAK | 550.339706 | 2 | 460.276575 | 1 | 18.1 |
| sp\|P39135\|SFP_BACSU | TKPISLEIAK | 550.339706 | 2 | 331.233982 | 1 | 18.1 |
| Holo-ACP | DLGVDSLAMTELQAHALQR[+340.085794] | 1204.571117 | 2 | 261.126739 | 1 | 30 |
| Holo-ACP | DLGVDSLAMTELQAHALQR[+340.085794] | 803.38317 | 3 | 261.126739 | 1 | 30 |
| Holo-ACP | DSLAMTELQAHALQR[+340.085794] | 1012.470675 | 2 | 261.126739 | 1 | 30 |
| Holo-ACP | DSLAMTELQAHALQR[+340.085794] | 675.316209 | 3 | 261.126739 | 1 | 30 |
| Apo-ACP | DLGVDSLAMTELQAHALQR | 1034.52822 | 2 | 1297.668131 | 1 | 33.1 |
| Apo-ACP | DLGVDSLAMTELQAHALQR | 1034.52822 | 2 | 1166.627646 | 1 | 33.1 |
| Apo-ACP | DLGVDSLAMTELQAHALQR | 1034.52822 | 2 | 1065.579967 | 1 | 33.1 |
| Apo-ACP | DLGVDSLAMTELQAHALQR | 1034.52822 | 2 | 936.537374 | 1 | 33.1 |
| Apo-ACP | DLGVDSLAMTELQAHALQR | 1034.52822 | 2 | 823.45331 | 1 | 33.1 |
| Apo-ACP | DLGVDSLAMTELQAHALQR | 690.021239 | 3 | 936.537374 | 1 | 20 |
| Apo-ACP | DLGVDSLAMTELQAHALQR | 690.021239 | 3 | 823.45331 | 1 | 20 |
| Apo-ACP | DLGVDSLAMTELQAHALQR | 690.021239 | 3 | 695.394733 | 1 | 20 |
| Apo-ACP | DLGVDSLAMTELQAHALQR | 690.021239 | 3 | 624.357619 | 1 | 20 |
| Apo-ACP | DLGVDSLAMTELQAHALQR | 690.021239 | 3 | 487.298707 | 1 | 20 |
| Apo-ACP | DSLAMTELQAHALQR | 842.427778 | 2 | 1166.627646 | 1 | 27.1 |
| Apo-ACP | DSLAMTELQAHALQR | 842.427778 | 2 | 1065.579967 | 1 | 27.1 |
| Apo-ACP | DSLAMTELQAHALQR | 842.427778 | 2 | 936.537374 | 1 | 27.1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Apo-ACP | DSLAMTELQAHALQR | 842.427778 | 2 | 823.45331 | 1 | 27.1 |
| Apo-ACP | DSLAMTELQAHALQR | 842.427778 | 2 | 695.394733 | 1 | 27.1 |
| Apo-ACP | DSLAMTELQAHALQR | 561.954277 | 3 | 823.45331 | 1 | 15.4 |
| Apo-ACP | DSLAMTELQAHALQR | 561.954277 | 3 | 695.394733 | 1 | 15.4 |
| Apo-ACP | DSLAMTELQAHALQR | 561.954277 | 3 | 624.357619 | 1 | 15.4 |
| Apo-ACP | DSLAMTELQAHALQR | 561.954277 | 3 | 487.298707 | 1 | 15.4 |
| Apo-ACP | DSLAMTELQAHALQR | 561.954277 | 3 | 416.261593 | 1 | 15.4 |
| 2-pentenoyl-ACP | DLGVDSLAMTELQAHALQR[+422.127659] | 1245.59205 | 2 | 343.168604 | 1 | 30 |
| 2-pentenoyl-ACP | DLGVDSLAMTELQAHALQR[+422.127659] | 830.730459 | 3 | 343.168604 | 1 | 30 |
| 2-pentenoyl-ACP | DSLAMTELQAHALQR[+422.127659] | 1053.491607 | 2 | 343.168604 | 1 | 30 |
| 2-pentenoyl-ACP | DSLAMTELQAHALQR[+422.127659] | 702.663497 | 3 | 343.168604 | 1 | 30 |
| 2_4-pentadienoyl-ACP | DLGVDSLAMTELQAHALQR[+420.112009] | 1244.584225 | 2 | 341.152954 | 1 | 30 |
| 2_4-pentadienoyl-ACP | DLGVDSLAMTELQAHALQR[+420.112009] | 830.058575 | 3 | 341.152954 | 1 | 30 |
| 2_4-pentadienoyl-ACP | DSLAMTELQAHALQR[+420.112009] | 1052.483782 | 2 | 341.152954 | 1 | 30 |
| 2_4-pentadienoyl-ACP | DSLAMTELQAHALQR[+420.112009] | 701.991614 | 3 | 341.152954 | 1 | 30 |
| Hydroxypentanoyl-ACP | DLGVDSLAMTELQAHALQR[+440.138224] | 1254.597332 | 2 | 361.179169 | 1 | 30 |
| Hydroxypentanoyl-ACP | DLGVDSLAMTELQAHALQR[+440.138224] | 836.73398 | 3 | 361.179169 | 1 | 30 |
| Hydroxypentanoyl-ACP | DSLAMTELQAHALQR[+440.138224] | 1062.49689 | 2 | 361.179169 | 1 | 30 |
| Hydroxypentanoyl-ACP | DSLAMTELQAHALQR[+440.138224] | 708.667019 | 3 | 361.179169 | 1 | 30 |
| Dihydroxypentanoyl-ACP | DLGVDSLAMTELQAHALQR[+456.133138] | 1262.594789 | 2 | 377.174083 | 1 | 30 |
| Dihydroxypentanoyl-ACP | DLGVDSLAMTELQAHALQR[+456.133138] | 842.065618 | 3 | 377.174083 | 1 | 30 |
| Dihydroxypentanoyl-ACP | DSLAMTELQAHALQR[+456.133138] | 1070.494347 | 2 | 377.174083 | 1 | 30 |
| Dihydroxypentanoyl-ACP | DSLAMTELQAHALQR[+456.133138] | 713.998657 | 3 | 377.174083 | 1 | 30 |
| 3-hydroxy-45-pentenoyl-ACP | DLGVDSLAMTELQAHALQR[+438.122574] | 1253.589507 | 2 | 359.163519 | 1 | 30 |
| 3-hydroxy-45-pentenoyl-ACP | DLGVDSLAMTELQAHALQR[+438.122574] | 836.062097 | 3 | 359.163519 | 1 | 30 |
| 3-hydroxy-45-pentenoyl-ACP | DSLAMTELQAHALQR[+438.122574] | 1061.489065 | 2 | 359.163519 | 1 | 30 |
| 3-hydroxy-45-pentenoyl-ACP | DSLAMTELQAHALQR[+438.122574] | 707.995135 | 3 | 359.163519 | 1 | 30 |
| 2-hexenoyl-ACP | DLGVDSLAMTELQAHALQR[+436.143309] | 1252.599875 | 2 | 357.184254 | 1 | 30 |
| 2-hexenoyl-ACP | DLGVDSLAMTELQAHALQR[+436.143309] | 835.402342 | 3 | 357.184254 | 1 | 30 |
| 2-hexenoyl-ACP | DSLAMTELQAHALQR[+436.143309] | 1060.499432 | 2 | 357.184254 | 1 | 30 |
| 2-hexenoyl-ACP | DSLAMTELQAHALQR[+436.143309] | 707.33538 | 3 | 357.184254 | 1 | 30 |
| Hydroxyhexanoyl-ACP | DLGVDSLAMTELQAHALQR[+454.153874] | 1261.605157 | 2 | 375.194819 | 1 | 30 |
| Hydroxyhexanoyl-ACP | DLGVDSLAMTELQAHALQR[+454.153874] | 841.405864 | 3 | 375.194819 | 1 | 30 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Hydroxyhexanoyl-ACP | DSLAMTELQAHALQR[+454.153874] | 1069.504715 | 2 | 375.194819 | 1 | 30 |
| Hydroxyhexanoyl-ACP | DSLAMTELQAHALQR[+454.153874] | 713.338902 | 3 | 375.194819 | 1 | 30 |
| 3-hydroxy-45-hexenoyl-ACP | DLGVDSLAMTELQAHALQR[+452.138224] | 1260.597332 | 2 | 373.179169 | 1 | 30 |
| 3-hydroxy-45-hexenoyl-ACP | DLGVDSLAMTELQAHALQR[+452.138224] | 840.73398 | 3 | 373.179169 | 1 | 30 |
| 3-hydroxy-45-hexenoyl-ACP | DSLAMTELQAHALQR[+452.138224] | 1068.49689 | 2 | 373.179169 | 1 | 30 |
| 3-hydroxy-45-hexenoyl-ACP | DSLAMTELQAHALQR[+452.138224] | 712.667019 | 3 | 373.179169 | 1 | 30 |
| Dihydroxyhexanoyl-ACP | DLGVDSLAMTELQAHALQR[+470.148788] | 1269.602614 | 2 | 391.189733 | 1 | 30 |
| Dihydroxyhexanoyl-ACP | DLGVDSLAMTELQAHALQR[+470.148788] | 846.737502 | 3 | 391.189733 | 1 | 30 |
| Dihydroxyhexanoyl-ACP | DSLAMTELQAHALQR[+470.148788] | 1077.502172 | 2 | 391.189733 | 1 | 30 |
| Dihydroxyhexanoyl-ACP | DSLAMTELQAHALQR[+470.148788] | 718.67054 | 3 | 391.189733 | 1 | 30 |
| 2_4-hexadienoyl-ACP | DLGVDSLAMTELQAHALQR[+434.127659] | 1251.59205 | 2 | 355.168604 | 1 | 30 |
| 2_4-hexadienoyl-ACP | DLGVDSLAMTELQAHALQR[+434.127659] | 834.730459 | 3 | 355.168604 | 1 | 30 |
| 2_4-hexadienoyl-ACP | DSLAMTELQAHALQR[+434.127659] | 1059.491607 | 2 | 355.168604 | 1 | 30 |
| 2_4-hexadienoyl-ACP | DSLAMTELQAHALQR[+434.127659] | 706.663497 | 3 | 355.168604 | 1 | 30 |
| tr\|E9KTG9\|E9KTG9_9ACTN_TcsD | DAPLALYER | 524.277106 | 2 | 861.482879 | 1 | 17.3 |
| tr\|E9KTG9\|E9KTG9_9ACTN_TcsD | DAPLALYER | 524.277106 | 2 | 580.308937 | 1 | 17.3 |
| tr\|E9KTG9\|E9KTG9_9ACTN_TcsD | DAPLALYER | 524.277106 | 2 | 467.224873 | 1 | 17.3 |
| tr\|E9KTG9\|E9KTG9_9ACTN_TcsD | DAPLALYER | 524.277106 | 2 | 304.161545 | 1 | 17.3 |
| tr\|E9KTG9\|E9KTG9_9ACTN_TcsD | YTAVTVPR | 453.755809 | 2 | 472.287808 | 1 | 15.1 |
| tr\|E9KTG9\|E9KTG9_9ACTN_TcsD | YTAVTVPR | 453.755809 | 2 | 371.24013 | 1 | 15.1 |
| tr\|E9KTG9\|E9KTG9_9ACTN_TcsD | YTAVTVPR | 453.755809 | 2 | 272.171716 | 1 | 15.1 |
| 2-heptenoyl-ACP | DLGVDSLAMTELQAHALQR[+450.158959] | 1259.6077 | 2 | 371.199904 | 1 | 30 |
| 2-heptenoyl-ACP | DLGVDSLAMTELQAHALQR[+450.158959] | 840.074225 | 3 | 371.199904 | 1 | 30 |
| 2-heptenoyl-ACP | DSLAMTELQAHALQR[+450.158959] | 1067.507257 | 2 | 371.199904 | 1 | 30 |
| 2-heptenoyl-ACP | DSLAMTELQAHALQR[+450.158959] | 712.007264 | 3 | 371.199904 | 1 | 30 |
| 2_4-heptadienoyl-ACP | DLGVDSLAMTELQAHALQR[+448.143309] | 1258.599875 | 2 | 369.184254 | 1 | 30 |
| 2_4-heptadienoyl-ACP | DLGVDSLAMTELQAHALQR[+448.143309] | 839.402342 | 3 | 369.184254 | 1 | 30 |
| 2_4-heptadienoyl-ACP | DSLAMTELQAHALQR[+448.143309] | 1066.499432 | 2 | 369.184254 | 1 | 30 |
| 2_4-heptadienoyl-ACP | DSLAMTELQAHALQR[+448.143309] | 711.33538 | 3 | 369.184254 | 1 | 30 |
| Hydroxyheptanoyl-ACP | DLGVDSLAMTELQAHALQR[+468.169524] | 1268.612982 | 2 | 389.210469 | 1 | 30 |
| Hydroxyheptanoyl-ACP | DLGVDSLAMTELQAHALQR[+468.169524] | 846.077747 | 3 | 389.210469 | 1 | 30 |
| Hydroxyheptanoyl-ACP | DSLAMTELQAHALQR[+468.169524] | 1076.51254 | 2 | 389.210469 | 1 | 30 |
| Hydroxyheptanoyl-ACP | DSLAMTELQAHALQR[+468.169524] | 718.010785 | 3 | 389.210469 | 1 | 30 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Dihydroxyheptanoyl-ACP | DLGVDSLAMTELQAHALQR[+484.164439] | 1276.61044 | 2 | 405.205384 | 1 | 30 |
| Dihydroxyheptanoyl-ACP | DLGVDSLAMTELQAHALQR[+484.164439] | 851.409385 | 3 | 405.205384 | 1 | 30 |
| Dihydroxyheptanoyl-ACP | DSLAMTELQAHALQR[+484.164439] | 1084.509997 | 2 | 405.205384 | 1 | 30 |
| Dihydroxyheptanoyl-ACP | DSLAMTELQAHALQR[+484.164439] | 723.342424 | 3 | 405.205384 | 1 | 30 |
| 3-hydroxy-45-heptenoyl-ACP | DLGVDSLAMTELQAHALQR[+466.153874] | 1267.605157 | 2 | 387.194819 | 1 | 30 |
| 3-hydroxy-45-heptenoyl-ACP | DLGVDSLAMTELQAHALQR[+466.153874] | 845.405864 | 3 | 387.194819 | 1 | 30 |
| 3-hydroxy-45-heptenoyl-ACP | DSLAMTELQAHALQR[+466.153874] | 1075.504715 | 2 | 387.194819 | 1 | 30 |
| 3-hydroxy-45-heptenoyl-ACP | DSLAMTELQAHALQR[+466.153874] | 717.338902 | 3 | 387.194819 | 1 | 30 |
| Pentanoyl-ACP | DLGVDSLAMTELQAHALQR[+424.1] | 1246.599875 | 2 | 345.184254 | 1 | 30 |
| Pentanoyl-ACP | DLGVDSLAMTELQAHALQR[+424.1] | 831.402342 | 3 | 345.184254 | 1 | 30 |
| Pentanoyl-ACP | DSLAMTELQAHALQR[+424.1] | 1054.499432 | 2 | 345.184254 | 1 | 30 |
| Pentanoyl-ACP | DSLAMTELQAHALQR[+424.1] | 703.33538 | 3 | 345.184254 | 1 | 30 |
| Butyryl-ACP | DLGVDSLAMTELQAHALQR[+410.1] | 1239.59205 | 2 | 331.168604 | 1 | 30 |
| Butyryl-ACP | DLGVDSLAMTELQAHALQR[+410.1] | 826.730459 | 3 | 331.168604 | 1 | 30 |
| Butyryl-ACP | DSLAMTELQAHALQR[+410.1] | 1047.491607 | 2 | 331.168604 | 1 | 30 |
| Butyryl-ACP | DSLAMTELQAHALQR[+410.1] | 698.663497 | 3 | 331.168604 | 1 | 30 |
| Crotonyl-ACP | DLGVDSLAMTELQAHALQR[+408.1] | 1238.584225 | 2 | 329.152954 | 1 | 30 |
| Crotonyl-ACP | DLGVDSLAMTELQAHALQR[+408.1] | 826.058575 | 3 | 329.152954 | 1 | 30 |
| Crotonyl-ACP | DSLAMTELQAHALQR[+408.1] | 1046.483782 | 2 | 329.152954 | 1 | 30 |
| Crotonyl-ACP | DSLAMTELQAHALQR[+408.1] | 697.991614 | 3 | 329.152954 | 1 | 30 |
| Hydroxybutyryl-ACP | DLGVDSLAMTELQAHALQR[+426.1] | 1247.589507 | 2 | 347.163519 | 1 | 30 |
| Hydroxybutyryl-ACP | DLGVDSLAMTELQAHALQR[+426.1] | 832.062097 | 3 | 347.163519 | 1 | 30 |
| Hydroxybutyryl-ACP | DSLAMTELQAHALQR[+426.1] | 1055.489065 | 2 | 347.163519 | 1 | 30 |
| Hydroxybutyryl-ACP | DSLAMTELQAHALQR[+426.1] | 703.995135 | 3 | 347.163519 | 1 | 30 |
| Propylmalonyl-ACP | DLGVDSLAMTELQAHALQR[+468.1] | 1268.594789 | 2 | 389.174083 | 1 | 30 |
| Propylmalonyl-ACP | DLGVDSLAMTELQAHALQR[+468.1] | 846.065618 | 3 | 389.174083 | 1 | 30 |
| Propylmalonyl-ACP | DSLAMTELQAHALQR[+468.1] | 1076.494347 | 2 | 389.174083 | 1 | 30 |
| Propylmalonyl-ACP | DSLAMTELQAHALQR[+468.1] | 717.998657 | 3 | 389.174083 | 1 | 30 |
| Allymalonyl-ACP | DLGVDSLAMTELQAHALQR[+466.1] | 1267.586964 | 2 | 387.158433 | 1 | 30 |
| Allymalonyl-ACP | DLGVDSLAMTELQAHALQR[+466.1] | 845.393735 | 3 | 387.158433 | 1 | 30 |
| Allymalonyl-ACP | DSLAMTELQAHALQR[+466.1] | 1075.486522 | 2 | 387.158433 | 1 | 30 |
| Allymalonyl-ACP | DSLAMTELQAHALQR[+466.1] | 717.326773 | 3 | 387.158433 | 1 | 30 |

Note: peptides with added masses in brackets (e.g. “DSLAMTELQAHALQR[+466.1]”) depict modified peptides with phosphopantetheine arms/substrates.

**Table S8**: Transitions used for the targeted LC-MS/MS-based detection of protein-bound intermediates

| Protein | Peptide | Precursor Ion | Precursor charge | Product Ion | Product charge | Collision Energy (eV) |
|---|---|---|---|---|---|---|
| Holo-Nmex-ACP | MDSLNLMDFLVYLE[+340.085794].light | 1021.94986 | 2 | 261.12674 | 1 | 30 |
| Holo-Nmex-ACP | MDSLNLMDFLVYLE[+340.085794].light | 681.635668 | 3 | 261.12674 | 1 | 30 |
| Nmex-ACP | GLRSIADRLDLE.light | 679.37534 | 2 | 760.38356 | 1 | 22.1 |
| Nmex-ACP | GLRSIADRLDLE.light | 679.37534 | 2 | 645.35662 | 1 | 22.1 |
| Nmex-ACP | GLRSIADRLDLE.light | 679.37534 | 2 | 489.25551 | 1 | 22.1 |
| Nmex-ACP | GLRSIADRLDLE.light | 453.252652 | 3 | 489.25551 | 1 | 11.5 |
| Nmex-ACP | GLRSIADRLDLE.light | 453.252652 | 3 | 376.17144 | 1 | 11.5 |
| Nmex-ACP | GLRSIADRLDLE.light | 453.252652 | 3 | 261.1445 | 1 | 11.5 |
| Nmex-ACP | NVDISATSSLEDDLE.light | 804.367772 | 2 | 820.35707 | 1 | 25.9 |
| Nmex-ACP | NVDISATSSLEDDLE.light | 804.367772 | 2 | 733.32504 | 1 | 25.9 |
| Nmex-ACP | NVDISATSSLEDDLE.light | 804.367772 | 2 | 620.24098 | 1 | 25.9 |
| Pentanoyl-Nmex-ACP | MDSLNLMDFLVYLE[+424.143309].light | 1063.97862 | 2 | 345.18425 | 1 | 30 |
| Pentanoyl-Nmex-ACP | MDSLNLMDFLVYLE[+424.143309].light | 709.65484 | 3 | 345.18425 | 1 | 30 |
| Pentenoyl-Nmex-ACP | MDSLNLMDFLVYLE[+422.127659].light | 1062.9708 | 2 | 343.1686 | 1 | 30 |
| Pentenoyl-Nmex-ACP | MDSLNLMDFLVYLE[+422.127659].light | 708.982956 | 3 | 343.1686 | 1 | 30 |
| Pentadienoyl-Nmex-ACP | MDSLNLMDFLVYLE[+420.112009].light | 1061.96297 | 2 | 341.15295 | 1 | 30 |
| Pentadienoyl-Nmex-ACP | MDSLNLMDFLVYLE[+420.112009].light | 708.311073 | 3 | 341.15295 | 1 | 30 |
| Hydroxypentanoyl-Nmex-ACP | MDSLNLMDFLVYLE[+440.138224].light | 1071.97608 | 2 | 361.17917 | 1 | 30 |
| Hydroxypentanoyl-Nmex-ACP | MDSLNLMDFLVYLE[+440.138224].light | 714.986478 | 3 | 361.17917 | 1 | 30 |
| Butyryl-Nmex-ACP | MDSLNLMDFLVYLE[+410.127659].light | 1056.9708 | 2 | 331.1686 | 1 | 30 |
| Butyryl-Nmex-ACP | MDSLNLMDFLVYLE[+410.127659].light | 704.982956 | 3 | 331.1686 | 1 | 30 |
| Crotonyl-Nmex-ACP | MDSLNLMDFLVYLE[+408.112009].light | 1055.96297 | 2 | 329.15295 | 1 | 30 |
| Crotonyl-Nmex-ACP | MDSLNLMDFLVYLE[+408.112009].light | 704.311073 | 3 | 329.15295 | 1 | 30 |
| Hydroxybutyryl-Nmex-ACP | MDSLNLMDFLVYLE[+426.122574].light | 1064.96825 | 2 | 347.16352 | 1 | 30 |
| Hydroxybutyryl-Nmex-ACP | MDSLNLMDFLVYLE[+426.122574].light | 710.314595 | 3 | 347.16352 | 1 | 30 |
| Hexanoyl-Nmex-ACP | MDSLNLMDFLVYLE[+438.158959].light | 1070.98645 | 2 | 359.1999 | 1 | 30 |
| Hexanoyl-Nmex-ACP | MDSLNLMDFLVYLE[+438.158959].light | 714.326723 | 3 | 359.1999 | 1 | 30 |
| Hexenoyl-Nmex-ACP | MDSLNLMDFLVYLE[+436.143309].light | 1069.97862 | 2 | 357.18425 | 1 | 30 |
| Hexenoyl-Nmex-ACP | MDSLNLMDFLVYLE[+436.143309].light | 713.65484 | 3 | 357.18425 | 1 | 30 |
| Hexadienoyl-Nmex-ACP | MDSLNLMDFLVYLE[+434.127659].light | 1068.9708 | 2 | 355.1686 | 1 | 30 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Hexadienoyl-Nmex-ACP | MDSLNLMDFLVYLE[+434.127659].light | 712.982956 | 3 | 355.1686 | 1 | 30 |
| Hydroxyhexanoyl-Nmex-ACP | MDSLNLMDFLVYLE[+454.153874].light | 1078.9839 | 2 | 375.19482 | 1 | 30 |
| Hydroxyhexanoyl-Nmex-ACP | MDSLNLMDFLVYLE[+454.153874].light | 719.658361 | 3 | 375.19482 | 1 | 30 |
| Heptanoyl-ACP | MDSLNLMDFLVYLE[+452.174609].light | 1077.99427 | 2 | 373.21555 | 1 | 30 |
| Heptanoyl-ACP | MDSLNLMDFLVYLE[+452.174609].light | 718.998606 | 3 | 373.21555 | 1 | 30 |
| Hydroxyheptanoyl-Nmex-ACP | MDSLNLMDFLVYLE[+468.169524].light | 1085.99173 | 2 | 389.21047 | 1 | 30 |
| Hydroxyheptanoyl-Nmex-ACP | MDSLNLMDFLVYLE[+468.169524].light | 724.330245 | 3 | 389.21047 | 1 | 30 |
| Heptenoyl-Nmex-ACP | MDSLNLMDFLVYLE[+450.158959].light | 1076.98645 | 2 | 371.1999 | 1 | 30 |
| Heptenoyl-Nmex-ACP | MDSLNLMDFLVYLE[+450.158959].light | 718.326723 | 3 | 371.1999 | 1 | 30 |
| Heptadienoyl-Nmex-ACP | MDSLNLMDFLVYLE[+448.143309].light | 1075.97862 | 2 | 369.18425 | 1 | 30 |
| Heptadienoyl-Nmex-ACP | MDSLNLMDFLVYLE[+448.143309].light | 717.65484 | 3 | 369.18425 | 1 | 30 |
| sp\|P39135\|SFP_BACSU | YSDLLAKDKDE.light | 648.819532 | 2 | 705.34136 | 1 | 21.1 |
| sp\|P39135\|SFP_BACSU | YSDLLAKDKDE.light | 648.819532 | 2 | 634.30425 | 1 | 21.1 |
| sp\|P39135\|SFP_BACSU | YSDLLAKDKDE.light | 648.819532 | 2 | 506.20928 | 1 | 21.1 |

**Table S9: Protein sequences used to build γ,δ-ACAD HMM used in genome mining**

| GenBank Accession Number | Host organism | Sequence |
|---|---|---|
| ADU56309.1 | *Streptomyces sp*. KCTC 11604BP | MSESERLGIVRDFVAREILGREGILDSLADAPLALYERFAETGLMNWWV PKEHGGLGLGLEESVRIVSELAYGDAGVAFTLFLPVLTTSMIGWYGSEEL KERFLGPLVARRGFCATLGSEHEAGSELARISTTVRRDGDTLVLDGTKA FSTSTDFARFLVVIARSADDPARYTAVTVPRDAPGLRVDKRWDVIGMRA SATYQVSFSDCRVPGDNALNGNGLRLLEIGLNASRILIAASALGVARRIR DVCMEYGKTKSLKGAPLVKDGVFAGRLGQFEMQIDVMANQCLAAARAY DATAARPDAARVLLRQGAQKSALTAKMFCGQTAWQIASTASEMFGGIG YTHDMVIGKLLRDVRHASIIEGGDDVLRDLVYQRFVVPTAKRT |
| WP_055551085.1 | *Streptomyces kanamyceticus* | MSEPEHLDTVRKFVAQEVLGRETHLDSLADAPLALYERFAETGLMNWW VPEEHGGLGLGLEDSVRIVSELAYGDAGVAFTLFLPILTTSMVSWYGSA ELKEKLLDPLVAHRGFCATLGSEHEAGSELAKISTVVRRDGEGLVLDGT KAFSTSTDFAQFLVVIARSAENPTRYLAVAVERDAPGLRIDKRWDVIGLR ASATYQVSFSDCHVPAGNALDGHGLRLLEIGLNASRILIAATALGVARRIR DLCMEYAKTKSLKGAPLVNDAVFAGRLGQFEMQIEVMANQCLAAARTY DATAARPDAARTLLRQGAQKSALTAKMFCGQTAWQIASTASEMFGGIG YTHDVPIGKLLRDVRHASIIEGGDDVLRDLVFHRFVVPTAKRT |
| WP_006350839.1 | *Streptomyces tsukubensis* | MSESERLGIVRDFVAREILGREGILDSLADAPLALYERFAETGLMNWWV PKEHGGLGLGLEESVRIVSELAYGDAGVAFTLFLPVLTTSMIGWYGSEEL KERFLGPLVARRGFCATLGSEHEAGSELARISTTVRRDGDTLVLDGTKA FSTSTDFARFLVVIARSADDPARYTAVTVPRDAPGLRVDKRWDVIGMRA SATYQVSFSDCRVPGDNALNGNGLRLLEIGLNASRILIAASALGVARRIR DVCMEYGKTKSLKGAPLVKDGVFAGRLGQFEMQIDVMANQCLAAARAY DATAARPDAARVLLRQGAQKSALTAKMFCGQTAWQIASTASEMFGGIG YTHDMVIGKLLRDVRHASIIEGGDDVLRDLVYQRFVVPTAKRT |
| AMM72019.1 | *Haliangium ochraceum* DSM 14365 | MSADTTKKNPLIEPIRGFVREHVLGREQQLDAGGELPLDIYEAFRKAGLA NWWLPESYGGHGLSLEQSVDIVAELAYGDAGLAFAFFLPILSTSVIEQFG SEEQKQRYLPALAKSGGSCATMASEEKAGSELIRTEALARGSAEEGFKL SGDKYFSTNADTAELLIVYARIAGPTPAYGAFLVPRSADGIRIVRRWEMN GLRASGTYELELRDCPAESQLAGNGLRILEVGLNSSRTLMAACAVGIAR RVRDVCLGYARNKEIKNTKLFNNHVFGAKLGQMEAELDGMMAVCRTAA REMDEIASREDAAKVFFREGTLKSVIVAKMLCGQLGWKIASVGSESLGG LGYTHDSIIGKLLRDVRYVSLVEAGDDVLRDLVFSRHVLPRFMSEIE |

**Table S10: Uncharacterized homologs of TcsD that contain characteristic γ,δ-ACAD motifs**

| Species | Protein GenBank Accession | DNA GenBank Accession | Gene cluster type (Antismash annotation) | Closest homologous cluster (Antismash) |
|---|---|---|---|---|
| Streptomyces kanamyceticus | ADU56239.1 | HM116536.1 | N/A | Allylmalonyl-ACP/FK506 (T1PKS) |
| Streptomyces tacrolimicus | ADU56353.1 | HM116538.1 | N/A | Allylmalonyl-ACP/FK506 (T1PKS) |
| Streptomyces tsukubensis VKM Ac-2618D | WP_006350839.1 | NZ_SGFG01000008.1 | N/A | Allylmalonyl-ACP/FK506 (T1PKS) |
| Streptomyces sp. KCTC 11604BP | ADU56309.1 | HM116537.1 | N/A | Allylmalonyl-ACP/FK506 (T1PKS) |
| Streptomyces tsukubensis NRRL 18488 | WP_006350839.1 | NZ_AJSZ01000908.1 | N/A | Allylmalonyl-ACP/FK506 (T1PKS) |
| Streptomyces tsukubensis F601 | WP_077974278.1 | NZ_MVFC01000056.1 | N/A | Allylmalonyl-ACP/FK506 (T1PKS) |
| Streptomyces sp. MJM7001 | WP_006350839.1 | HQ696504.1 | N/A | Allylmalonyl-ACP/FK506 (T1PKS) |
| Haliangium ochraceum DSM 14365 | AMM72019.1 | KU523555.1 | N/A | Haliangicin (T1PKS) |
| Nocardia brasiliensis ATCC 700358 | WP_014984668.1 | NC_018681.1 | T1PKS | Akaeolide (polyketide) (16%) |
| Nocardia brasiliensis HUJEG-1 isolate P-200 | WP_014984668.1 | NZ_LRRM01000006.1 | T1PKS | Akaeolide (polyketide) (16%) |
| Saccharomonospora saliphila YIM 90502 | WP_019815635.1 | NZ_KB912660.1 | PKS-like,T1PKS | Arsenopolyketides (45%) |
| Streptomyces sp. ADI96-15 | WP_023416081.1 | NZ_ML123109.1 | PKS-like,T1PKS | Arsenopolyketides (54%) |
| Streptomyces sp. PVA 94-07 | WP_023416081.1/ESQ07377.1 | NZ_CM002273.1 | PKS-like,T1PKS | Arsenopolyketides (58%) |
| Streptomyces sp. Root63 | WP_023416081.1 | NZ_LMGX01000018.1 | PKS-like,T1PKS | Arsenopolyketides (58%) |

| | | | | |
|---|---|---|---|---|
| Streptomyces sp. Root1295 | WP_023416081.1 | NZ_LMEL01000021.1 | PKS-like,T1PKS | Arsenopolyketides (58%) |
| Streptomyces sp. GBA 94-10 | WP_023416081.1 | NZ_CM002271.1 | PKS-like,T1PKS | Arsenopolyketides (58%) |
| Streptomyces sp. CB00072 | WP_073868671.1 | NZ_LIPB01000003.1 | PKS-like,T1PKS | Arsenopolyketides (58%) |
| Millisia brevis | WP_066907456.1 | NZ_BCRN01000007.1 | NRPS,T1PKS | Aurantimycin (18%) |
| Pseudonocardia endophytica | WP_132431031.1 | NZ_SMFZ01000002.1 | NRPS,PKS-like,T1PKS | Butyrolactol A (46%) |
| Herbidospora sp. NEAU-GS14 | WP_137246715.1 | NZ_SZQA01000007.1 | T1PKS | Butyrolactol A (33%) |
| Herbidospora yilanensis strain NBRC 106371 | WP_062352189.1 | NZ_BBXE01000030.1 | NRPS,T1PKS | Butyrolactol A (40%) |
| Streptomyces sp. AmelKG-E11A | WP_099283133.1 | NZ_AQRJ01000070.1 | T1PKS/NRPS/lassopeptide | Butyrolactol A (40%) |
| Streptomyces uncialis | WP_073788609.1 | NZ_LFBV01000001.1 | T1PKS/NRPS/lassopeptide | Butyrolactol A (40%) |
| Herbidospora cretacea strain NRRL B-16917 | WP_030450088.1 | NZ_JODQ01000001.1 | NRPS,T1PKS,arylpolyene | Butyrolactol A (46%) |
| Herbidospora daliensis strain NBRC 106372 | WP_062432695.1 | NZ_BBXF01000003.1 | NRPS,T1PKS,arylpolyene | Butyrolactol A (46%) |
| Herbidospora sakaeratensis strain NBRC 102641 | WP_062330499.1 | NZ_BBXC01000007.1 | NRPS,T1PKS,ladderane | Butyrolactol A (46%) |
| Herbidospora mongoliensis strain NBRC 105882 | WP_066370148.1 | NZ_BBXD01000017.1 | NRPS,T1PKS,ladderane | Butyrolactol A (46%) |
| Alloactinosynnema album | WP_091377637.1 | | T1PKS/NRPS | Butyrolactol A (46%) |
| Nonomuraea sp. PA1-10 | WP_139634331 | NZ_VDLX01000013.1 | NRPS,T1PKS | Butyrolactol A (53%) |
| Nonomuraea coxensis DSM 45129 | WP_020541001.1 | NZ_KB903944.1 | T1PKS | Butyrolactol A (66%) |
| Labedaea rhizosphaerae DSM 45361 | WP_133849167.1 | NZ_SNXZ01000002.1 | NRPS,T1PKS,betalactone,transAT-PKS-like | Butyrolactol A (40%) |
| Nocardia vulneris W9851 | WP_052281359.1 | NZ_JNFP00000000.1 | T1PKS | Chalcomycin (9%) |

| | | | | |
|---|---|---|---|---|
| Nocardia vulneris NBRC 108936 | WP_052281359.1 | NZ_BDCI01000001.1 | T1PKS | Chalcomycin (9%) |
| Amycolatopsis coloradensis | WP_076160419.1 | NZ_MQUQ01000006.1 | butyrolactone | Chlorothricin (4%) |
| Amycolatopsis coloradensis | WP_076160419.1 | NZ_MQUQ00000000 (NZ_MQUQ01000006.1) | butyrolactone | Chlorothricin (4%) |
| Actinocrispum wychmicini | WP_132116054.1 | NZ_SLWS01000003.1 | PKS-like,T1PKS | Chlorothricin (48%) |
| Nocardia suismassiliense S-137 | WP_107655957.1 | NZ_LT985361.1 | NRPS,T1PKS,ladderane | Coelimycin (29%) |
| Nocardia mexicana | WP_068019852.1 | NZ_QQAZ01000001 | butyrolactone/ladderane | Colabomycin (4%) |
| Nocardia sp. BMG51109 | WP_024802972.1 | NZ_JAFQ01000004.1 | T1PKS | Stambomycin (40%) |
| Streptomyces peucetius subsp. caesius ATCC 27952 | WP_017584471.1 | NZ_CP022438.1 | T1PKS,T2PKS | Oligomycin (61%) |
| Actinosynnema mirum ATCC 29888 | WP_015803413.1 | NC_013093.1 | T1PKS | Cyclizidine (41%) |
| Streptomyces puniciscabiei | WP_069777509.1 | NZ_CP017248.1 | T1PKS | E-837 (100%) |
| Lechevalieria aerocolonigenes | WP_051784425.1 | NZ_BBOJ01000031.1 | arylpolyene,butyrolactone | Enduracidin (4%) |
| Saccharothrix sp. NRRL B-16348 | WP_053716783.1 | NZ_LGED01000125.1 | NRPS | Erythrochelin (85%) |
| Saccharothrix carnea | WP_106615758.1 | NZ_PYAX01000004.1 | NRPS | Erythrochelin (85%) |
| Saccharothrix texasensis | WP_123742396.1 | NZ_RJKM01000001.1 | NRPS | Erythrochelin (85%) |
| Streptomyces sp. P3 | WP_107448820.1 | NZ_CP028369.1 | LAP,PKS-like,T1PKS,T2PKS | Hedamycin (43%) |
| Kitasatospora griseola | WP_043910458.1 | NZ_JXZB01000002.1 | NRPS,PKS-like,T1PKS,T2PKS | Hedamycin (31%) |
| Kitasatospora sp. CB02891 | WP_100586111.1 | NZ_NNBO01000004.1 | NRPS,PKS-like,T1PKS,T2PKS | Hedamycin (34%) |
| Streptomyces sp. FBKL.4005 | WP_059247715.1 | NZ_NPKF01000001.1 | LAP,PKS-like,T1PKS,T2PKS | Hedamycin (43%) |
| Streptomyces reticuli | WP_059247715.1 | NZ_LN997842.1 | LAP,PKS-like,T2PKS | Hedamycin (46%) |
| Actinocorallia populi strain A251 | WP_106396532.1 | NZ_PVZV01000001.1 | PKS-like,T1PKS,T2PKS | Hedamycin (56%) |

| | | | | |
|---|---|---|---|---|
| Streptomyces sp. t99 | WP_030720926.1 | NZ_NTGQ01000069.1 | T1PKS,T2PKS | Hedamycin (59%) |
| Streptomyces sp. st140 | WP_030720926.1 | NZ_NTGS01000037.1 | T2PKS | Hedamycin (59%) |
| Streptomyces sp. f51 | WP_030720926.1 | NZ_NTHH01000040.1 | T1PKS,T2PKS | Hedamycin (62%) |
| Thermomonospora curvata ATCC 19995 | WP_012853108.1 | NC_013510.1 | PKS-like,T1PKS,T2PKS | Hedamycin (81%) |
| Streptomyces sp. ms115 | WP_097947834.1 | NZ_NTHD01000048.1 | PKS-like,T1PKS,T2PKS | Hedamycin (87%) |
| Streptomyces sp. b62 | WP_030720926.1 | NZ_NTHK01000012.1 | PKS-like,T1PKS,T2PKS | Hedamycin (87%) |
| Streptomyces sp. f150 | WP_030720926.1 | NZ_NTHG01000036.1 | PKS-like,T1PKS,T2PKS | Hedamycin (87%) |
| Streptomyces griseus subsp. griseus NRRL F-5144 | WP_030720926.1 | NZ_JOGA01000026.1 | PKS-like,T1PKS,T2PKS | Hedamycin (87%) |
| Streptomyces sp. gb14 | WP_030720926.1 | NZ_NTHF01000021.1 | PKS-like,T1PKS,T2PKS | Hedamycin (87%) |
| Streptomyces sp. SS07 | WP_030720926.1 | NZ_KZ268499.1 | PKS-like,T1PKS,T2PKS,terpene | Hedamycin (87%) |
| Streptomyces rubellomurinus subsp. indigoferus | KJS54098.1 | JZKG00000000.1 | T1PKS | Hedamycin (9%) |
| Streptomyces sp. NRRL S-146 | WP_031110854.1 | NZ_JOAW01000493.1 | T2PKS | Hedamycin (9%) |
| Streptomyces sp. Ag82_G6-1 | WP_097222862.1 | NZ_OCNA01000001.1 | PKS-like, T1PKS, T2PKS | Hedamycin (90%) |
| Nocardia alba DSM 44684 | TCJ89880.1 | SMFR01000008.1 | PKS-like,T1PKS,T2PKS | Hedamycin (25%) |
| Nocardia alba NBRC 108234 | WP_067458355.1 | NZ_BDAX01000033.1 | PKS-like,T1PKS,T2PKS | Hedamycin (25%) |
| Streptomyces sp. NP10 | WP_126932507.1 | NZ_PDIQ01000024.1 | PKS-like,T1PKS,T2PKS | Hedamycin (50%) |
| Streptomyces sp. JS01 | WP_030720926.1 | NZ_JPWW01000020.1 | PKS-like,T1PKS,T2PKS | Hedamycin (87%) |
| Streptomyces sp. Root264 | KRD19112.1 | NZ_LMIZ01000001 | LAP,PKS-like,T1PKS,T2PKS | Hedamycin (43%) |
| Actinophytocola oryzae DSM 45499 | WP_133905254.1 | NZ_SOCP01000009.1 | T1PKS | Incednine (17%) |
| Streptomyces sp. WAC 01529 | WP_125514820.1 | CP029617.1 | T1PKS/NRPS | Lorneic acid A (23%) |

| | | | | |
|---|---|---|---|---|
| uncultured bacterium | ASV46999.1 | KY560362.1 | T1PKS | Lorneic acid A (23%) |
| Streptomyces oceani | WP_070197325.1 | NZ_LJGU00000000.1 | T1PKS | Lorneic acid A (23%) |
| Streptomyces puniciscabiei | WP_069776466.1 | NZ_CP017248.1 | T1PKS,terpene | Lorneic acid A (28%) |
| Streptomyces sp. DS1-2 | WP_120696069.1 | NZ_RBDY01000003.1 | T1PKS | Methymycin / pikromycin (57%) |
| Sorangium cellulosum | KYF78568.1 | JEMB01002769.1 | T1PKS | Micromonolactam (100%) |
| Actinosynnema pretiosum strain X47 | WP_096495686.1 | NZ_CP023445.1 | T1PKS | Microtermolide A (33%) |
| Streptomyces sp. CNQ-509 | WP_047018908.1 | NZ_CP011492.1 | NRPS,PKS-like,T1PKS | Microtermolide A (53%) |
| Streptomyces sp. AZ1-7 | WP_120696069.1 | NZ_RBDX01000002.1 | T1PKS | Nocardiopsin (21%) |
| Catenulispora acidiphila DSM 44928 | WP_015795553.1 | NC_013131.1 | T1PKS/NRPS | Octacosamicin (29%) |
| Streptomyces sp. NRRL S-118 | WP_031080613.1 | NZ_KL591043.1 | T1PKS | Piericidin A1 (58%) |
| Streptomyces griseus subsp. griseus NRRL F-5144 | WP_030723159.1 | NZ_JOGA01000040.1 | NRPS | Polyoxypeptin (40%) |
| Streptomyces phaeoluteigriseus strain DSM 41896 | OJT46852.1 | MPOH00000000.2 | NRPS,T1PKS,other | Polyoxypeptin (75%) |
| Streptomyces sp. XY006 | WP_094055126.1 | NZ_NOKT01000017.1 | NRPS,other | Polyoxypeptin (40%) |
| Streptomyces sp. E5N91 s-91 | WP_121712942.1 | NZ_RAIE01000715.1 | other | Polyoxypeptin (40%) |
| Pseudonocardiaceae bacterium YIM PH 21723 | WP_120088448.1 | NZ_QZFT00000000.1 | T1PKS | Pyrronazol B (9%) |
| Nocardia altamirensis NBRC 108246 | WP_069164184.1 | NZ_BDAY01000046.1 | T1PKS | Stambomycin (40%) |
| Streptomyces sp. NRRL WC-3742 | WP_031071135.1 | NZ_JOCF01000031.1 | NRPS,T1PKS | Lydicamycin (40%) |
| Kitasatospora azatica KCTC 9699 | WP_083976688.1 | NZ_JQMO01000003.1 | PKS-like,T1PKS | Zincophorin (61%) |

| | | | | |
|---|---|---|---|---|
| Streptomyces rubellomurinus ATCC 31215 | WP_017584471.1 | NZ_JZKH01000090.1 | N/A | N/A |
| Streptomyces albidoflavus | WP_128462586.1 | NZ_QHCQ00000000.1 (NZ_SCDQ01000064.1) | N/A | N/A |
| Streptomyces sp. PRh5 | WP_051573751.1 | NZ_JABQ01000071.1 | N/A | N/A |
| Streptomyces sp. FXJ1.172 | WP_067044802.1 | NZ_LWRP01000068.1 | N/A | N/A |
| Streptomyces sp. AVP053U2 | WP_062189972.1 | NZ_LMTQ02000008.1 | N/A | N/A |
| Enhygromyxa salina | KIG11693.1 | JMCC00000000.2 | N/A | N/A |
| Streptomyces tricolor NRRL B-16925 | WP_086702637.1 | NZ_MUMF00000000.1 | N/A | N/A |
| Streptomyces sp. 8K308 | WP_132929808.1 | NZ_SMKC01000043.1 | N/A | N/A |
| Mycobacteroides abscessus | WP_079869619.1/SHS51726.1 | NZ_FSAT01000004.1 | N/A | N/A |
| Streptomyces atriruber | WP_055564739.1 | NZ_LIPN01000084.1 (NZ_SMKI00000000.1) | N/A | N/A |
| Streptomyces hainanensis DSM 41900 | WP_132818863.1 | NZ_SMKI01000162.1 | N/A | N/A |
| Nocardiopsis valliformis DSM 45023 | WP_017584471.1 | NZ_ANAZ01000058.1 | N/A | N/A |
| Streptomyces sp. E2N166 | WP_121750360.1 | NZ_RAIF01000032.1 | N/A | N/A |
| Streptomyces sp. NBRC 110035 | WP_042171193.1 | NZ_BBNN01000027.1 | N/A | N/A |
| Plesiocystis pacifica SIR-1 | WP_006969759.1/EDM81215.1 | NZ_ABCS01000005.1 | transAT-PKS | N/A |
| Nonomuraea sp. SBT364 | WP_049575489.1 | NZ_LAVL01000205.1 | T1PKS | N/A |

Abbreviations: Polyketide synthase (PKS), type I polyketide synthase (T1PKS), type II polyketide synthase (T2PKS), nonribosomal peptide synthetase (NRPS), lassopeptide (LAP

## Supplementary Figures

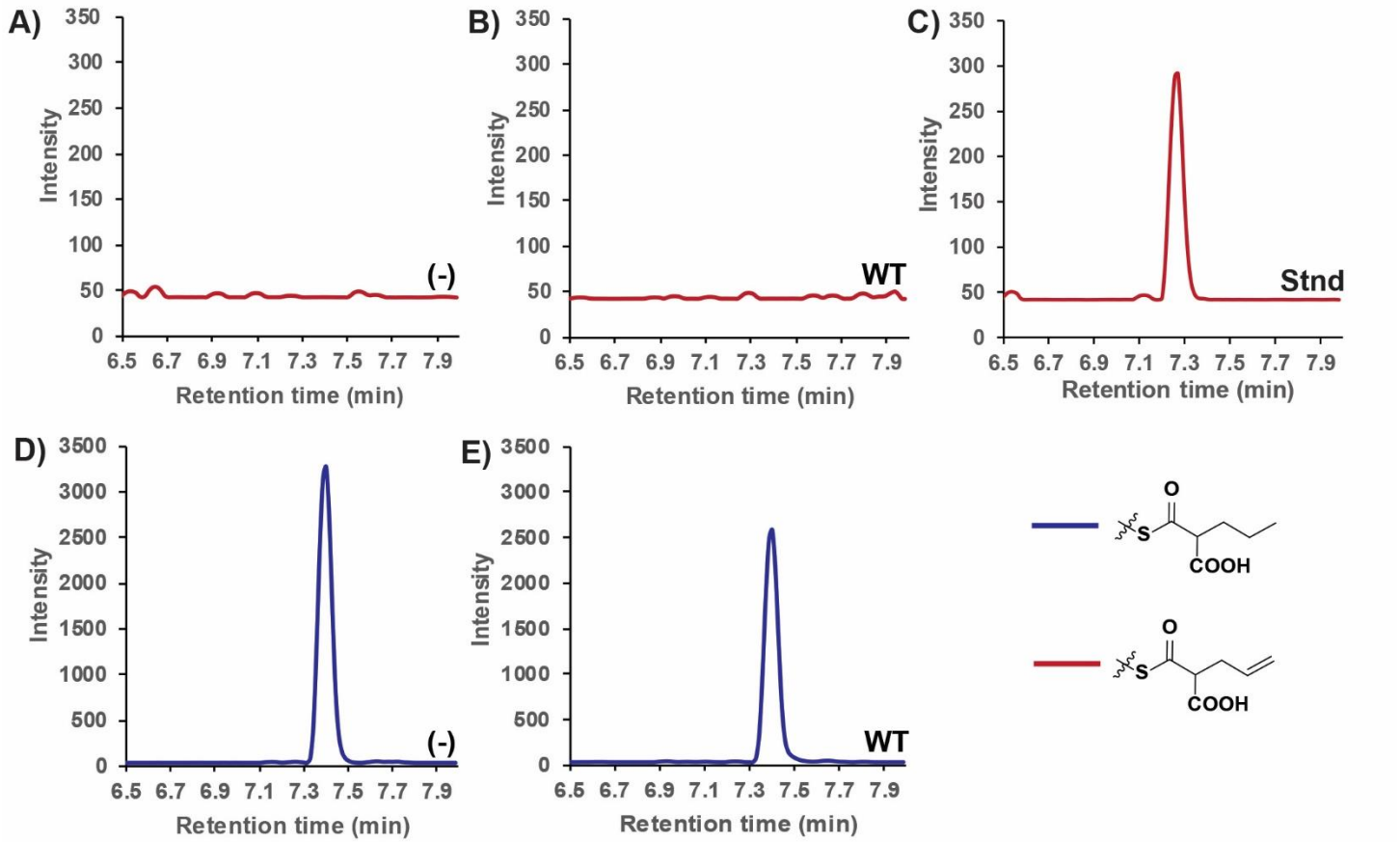



**Figure S1**. TcsD activity on propylmalonyl-ACP (pathway *A2* from Figure 1) analyzed via targeted LC-MS/MS. WT = wild type TcsD, (-) = negative control, Stnd = standard. The chromatograms shown depict: **A)** the allylmalonyl-ACP transition for the negative control reaction **B)** the allylmalonyl-ACP transition for the wild type TcsD reaction **C) a**n allylmalonyl-ACP standard **D)** the propylmalonyl-ACP transition for the negative control reaction **E)** the propylmalonyl-ACP transition for the TcsD wild type reaction

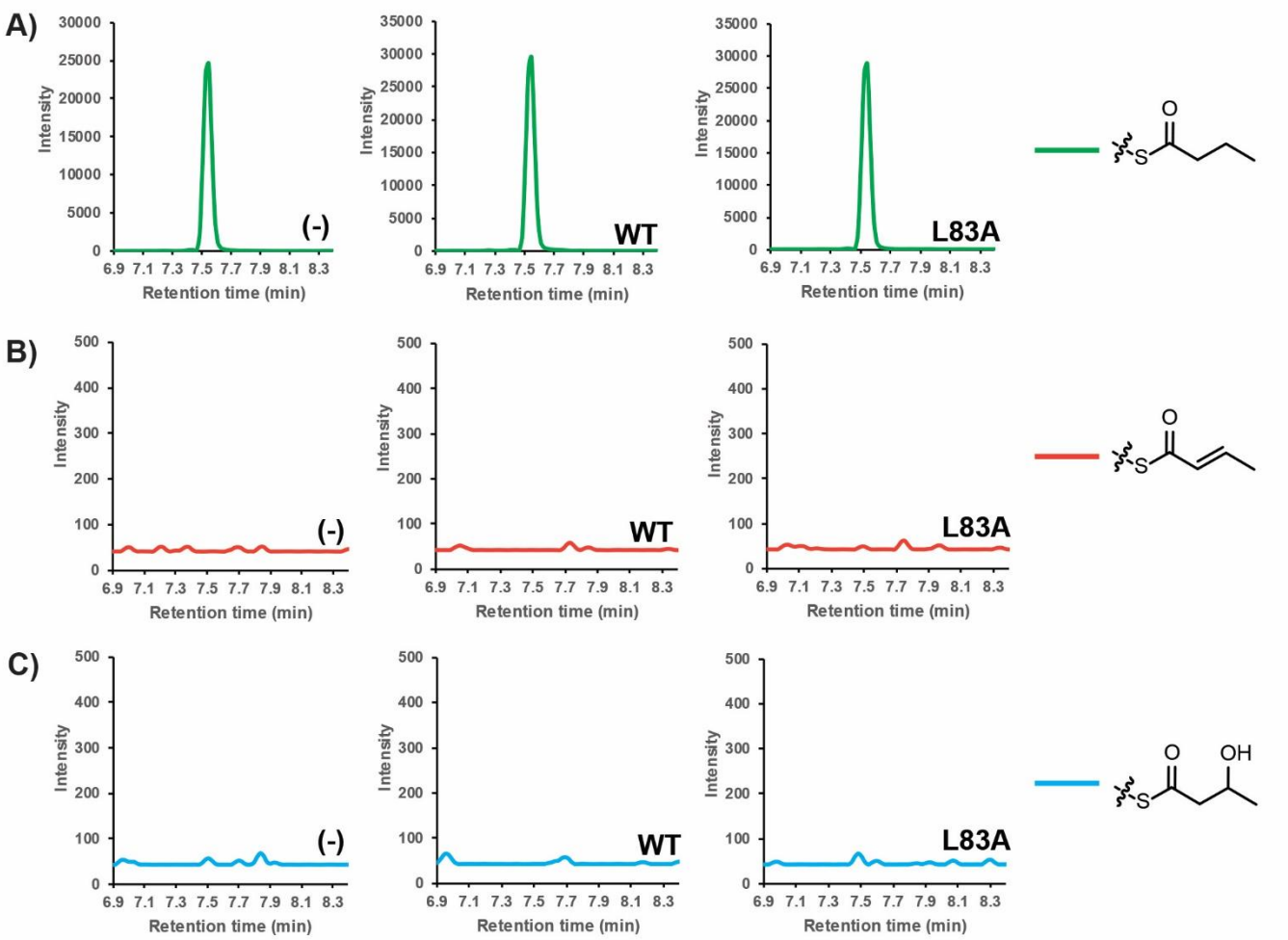



**Figure S2**. Activity of TcsD wild type and L83A mutant on butyryl-ACP analyzed by targeted LC-MS/MS. WT = wild type TcsD, (-) = negative control, L83A = TcsD L83A mutant. Chromatograms representing different transitions are color coded (key is shown on right side of figure). The chromatograms shown depict: **A)** butyryl-ACP, **B)** crotonyl-ACP, or **C)** 3-hydroxybutyryl-ACP for each assay (negative control, TcsD wild type, or TcsD L83A).

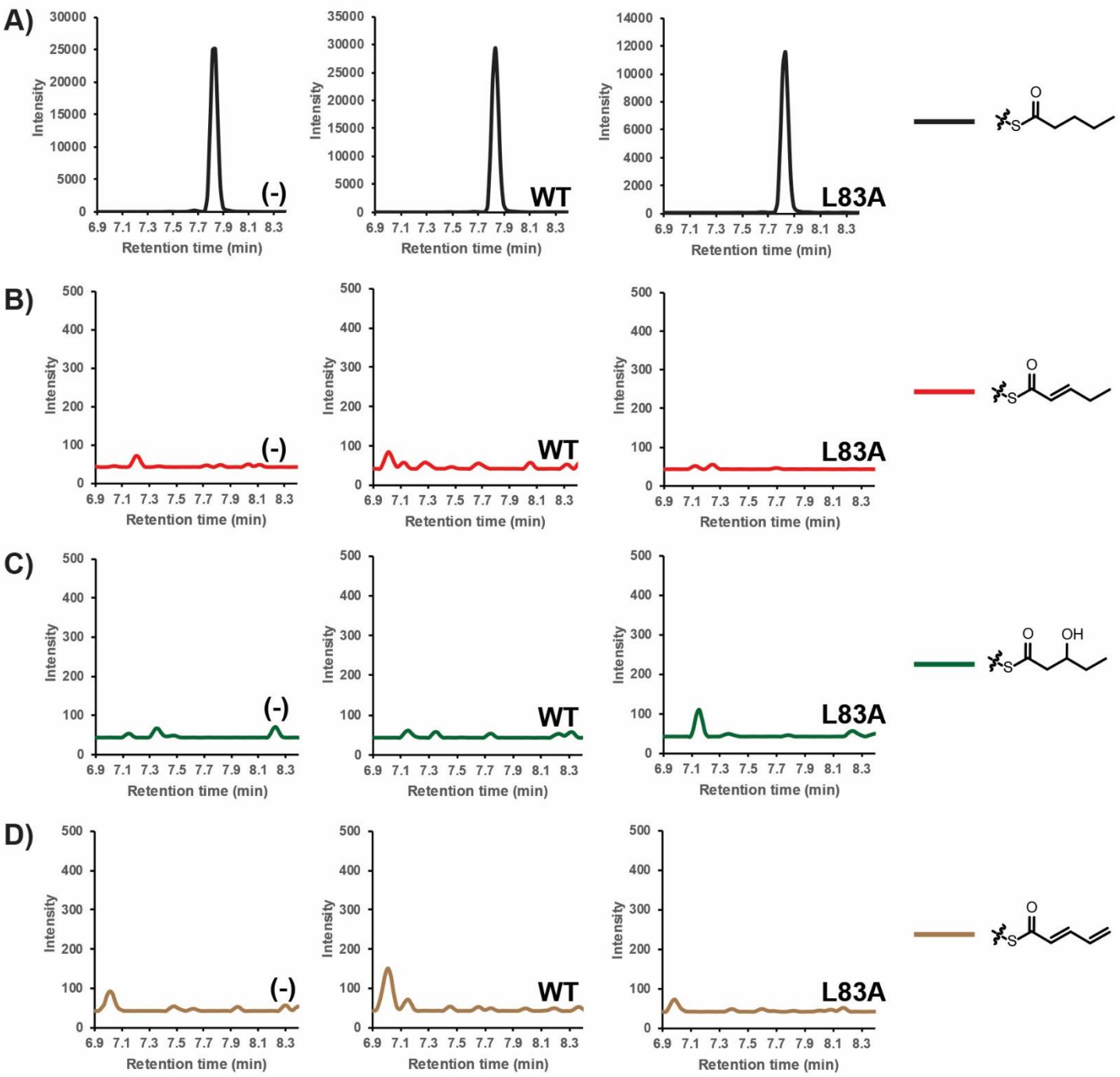



**Figure S3**. Activity of TcsD wild type and L83A mutant on pentanoyl-ACP analyzed by targeted LC-MS/MS. WT = wild type TcsD, (-) = negative control, L83A = TcsD L83A mutant. Chromatograms representing different transitions are color coded (key is shown on right side of figure). The chromatograms shown depict: **A)** pentanoyl-ACP, **B)** 2-pentenoyl-ACP, **C)** the 3-hydroxypentanoyl-ACP, or **D)** 2,4-pentadienoyl-ACP for each assay (negative control, TcsD wild type, or TcsD L83A).

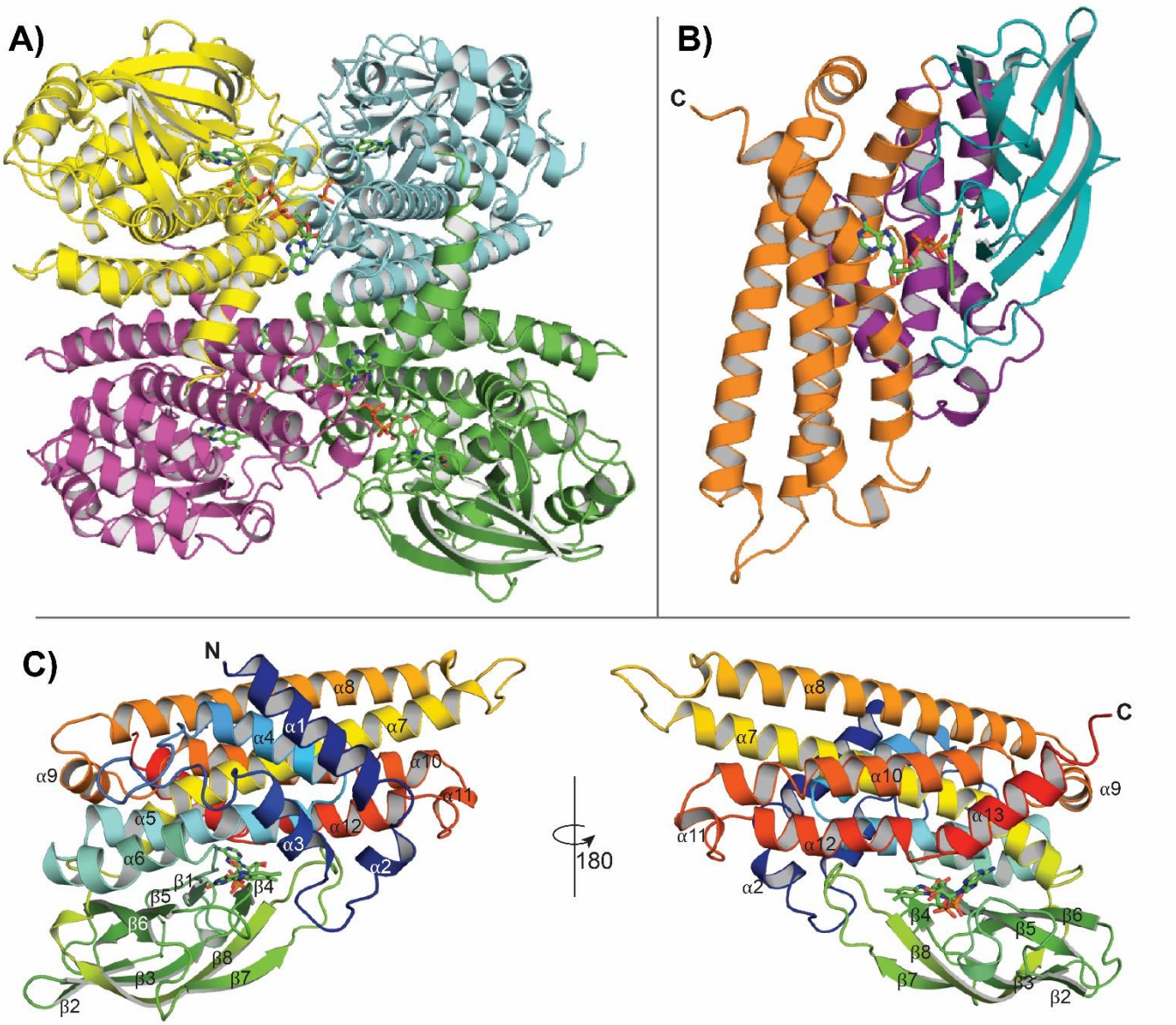



**Figure S4**. Overall structure of TcsD. **A)** Tetrameric structure of TcsD. Each subunit is shown in a different color. FAD cofactors are shown as sticks. **B)** A single TcsD subunit. The N-terminal α-helix domain (purple), middle β-sheet domain (teal), and C-terminal α-helix domain (orange) are highlighted. C) A single TcsD monomer colored in a progressive rainbow from its N- (blue) to C-terminus (red). Helices and sheets are numbered.

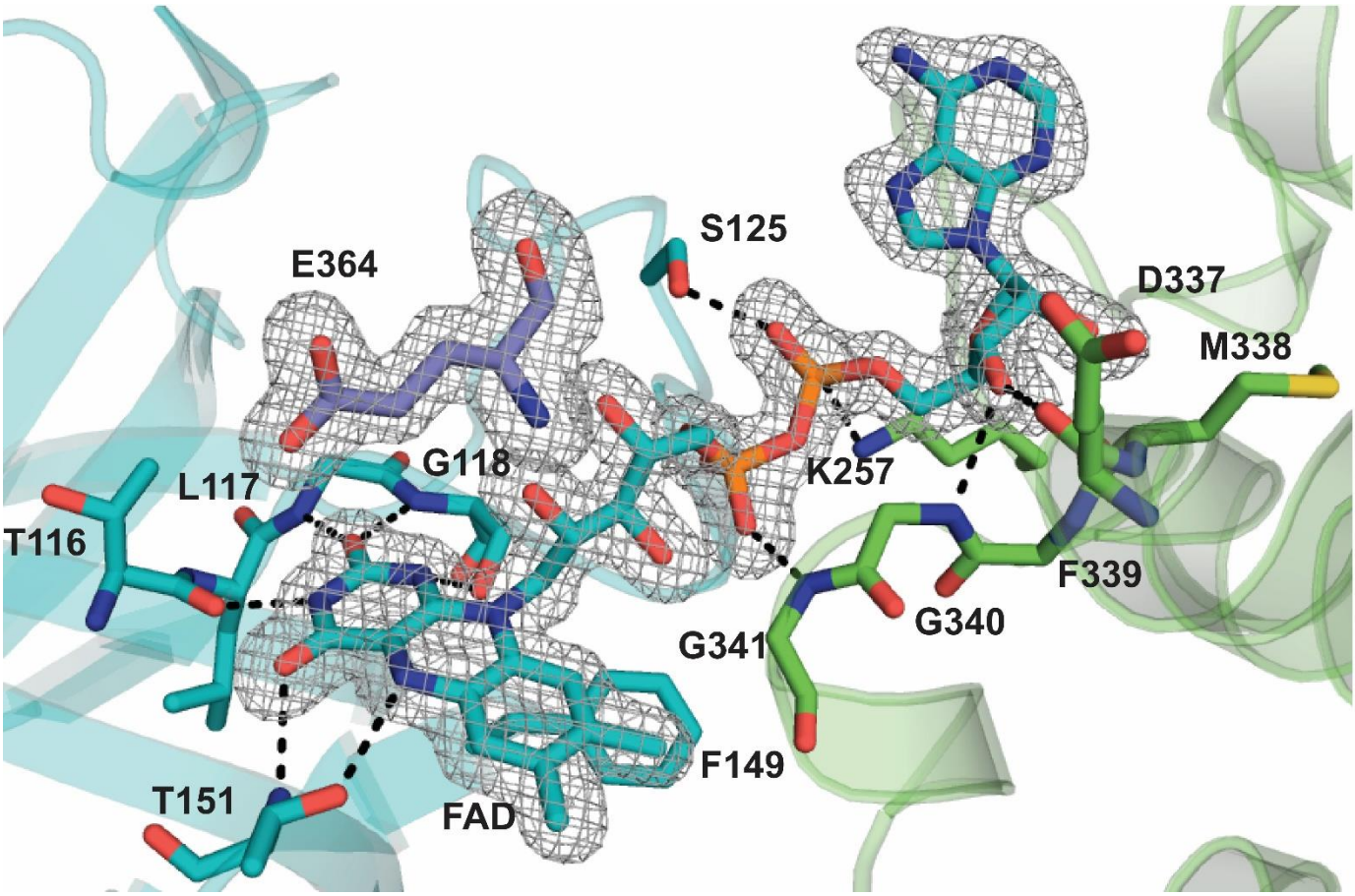



**Figure S5**.FAD-interacting residues of TcsD. The mFo-DFc omit electron density map of FAD and the catalytic glutamate is shown with gray mesh and is contoured at 3σ. Residues that interact with FAD via hydrogen bonds are shown as sticks, with hydrogen bonds depicted as black dashed lines. Residues pertaining to the same subunit as FAD (T116, L117, G118, S119, S125, F149, T151) are shown in blue, while residues pertaining to the adjacent subunit (K257, D337, M338, F339, G340, G341) are shown in green. The catalytic glutamate, E364, is shown in purple.

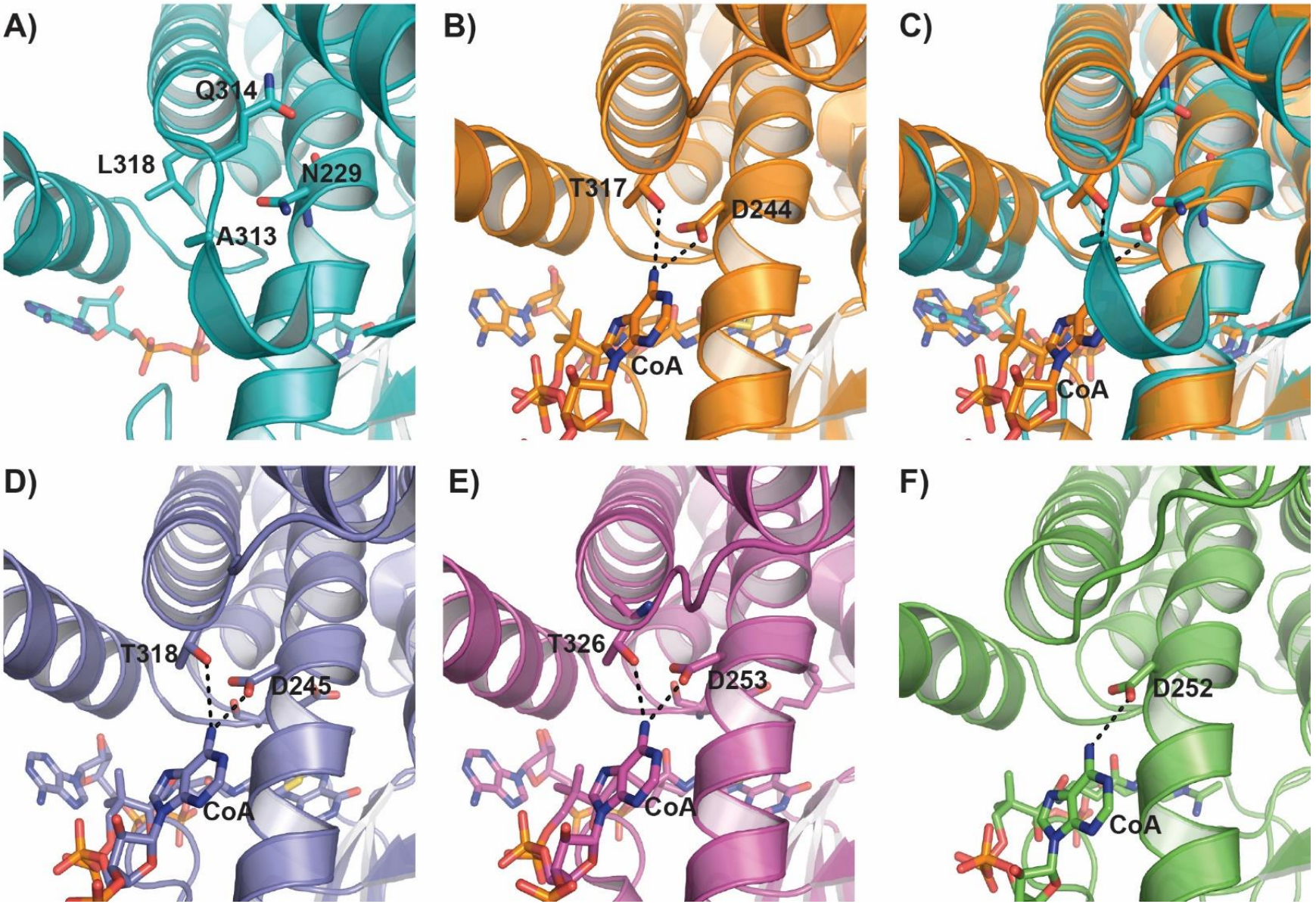



**Figure S6**. Residues of ACADs that hydrogen bond with the nucleotide portion of CoA in **A)** TcsD, **B)** *M. esldenii* butyryl-CoA dehydrogenase (1buc), **C)** *M. esldenii* butyryl-CoA dehydrogenase overlaid with TcsD, **D)** *Sus scrofa* medium chain ACAD (1udy), **E)** rat short chain ACAD (1jqi), **F)** human isovaleroyl-CoA dehydrogenase (1ivh). TcsD has an Asn residue in the place of the conserved Asp of α,β-ACADs. Helix 9 occupies some of the space where the nucleotide portion of CoA would normally bind. The conserved Thr residue is replaced with a Leu residue.

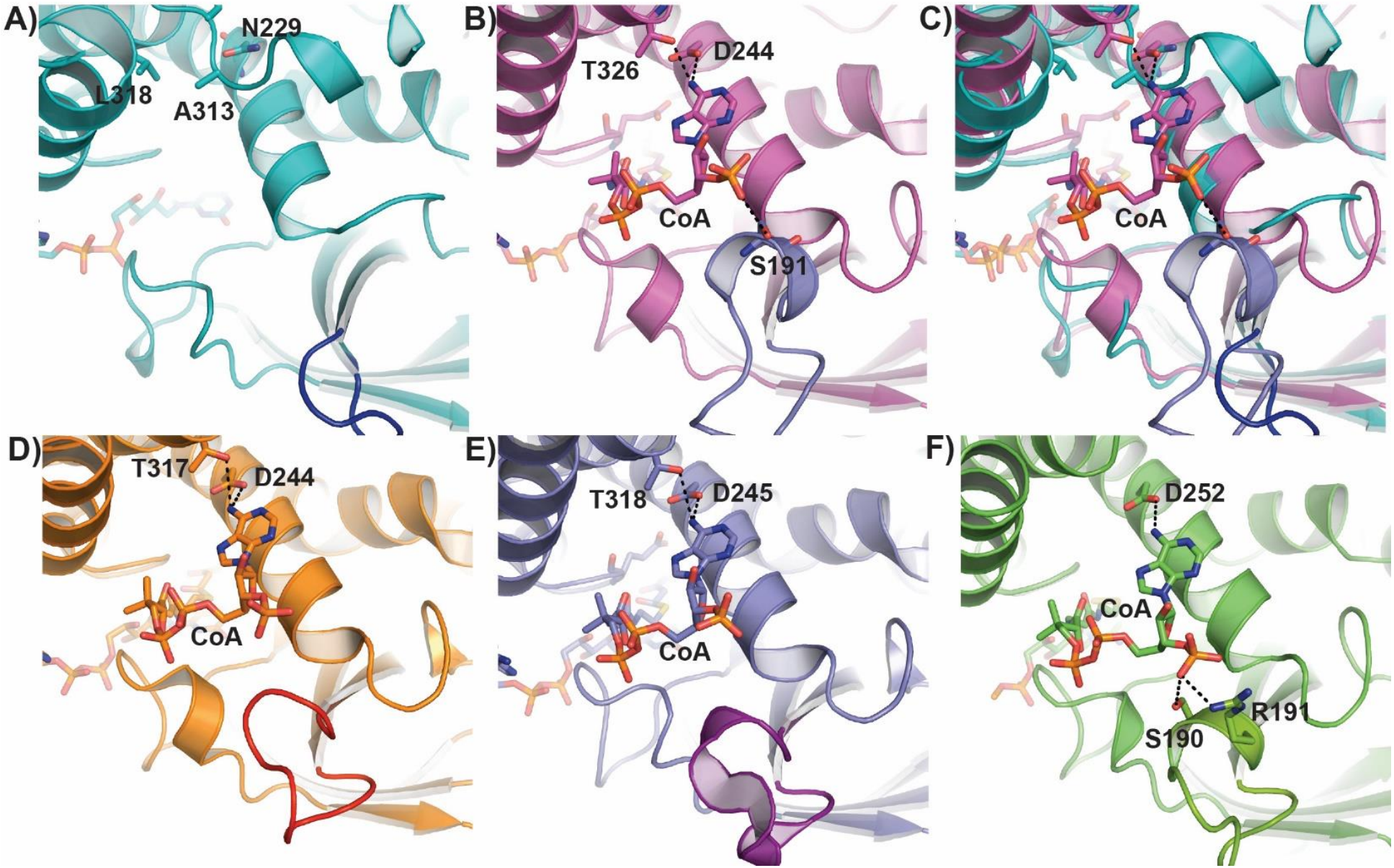



**Figure S7**.CoA phosphate-binding region of **A)** TcsD, **B)** rat short chain ACAD (1jqi), **C)** rat short-chain ACAD overlaid with TcsD, **D)** *M. esldenii* butyryl-CoA dehydrogenase (1buc), **E)** *Sus scrofa* medium chain ACAD (1udy), **F)** human isovaleroyl-CoA dehydrogenase (1ivh). The loop that approaches the phosphate groups is colored in a different shade in each panel.

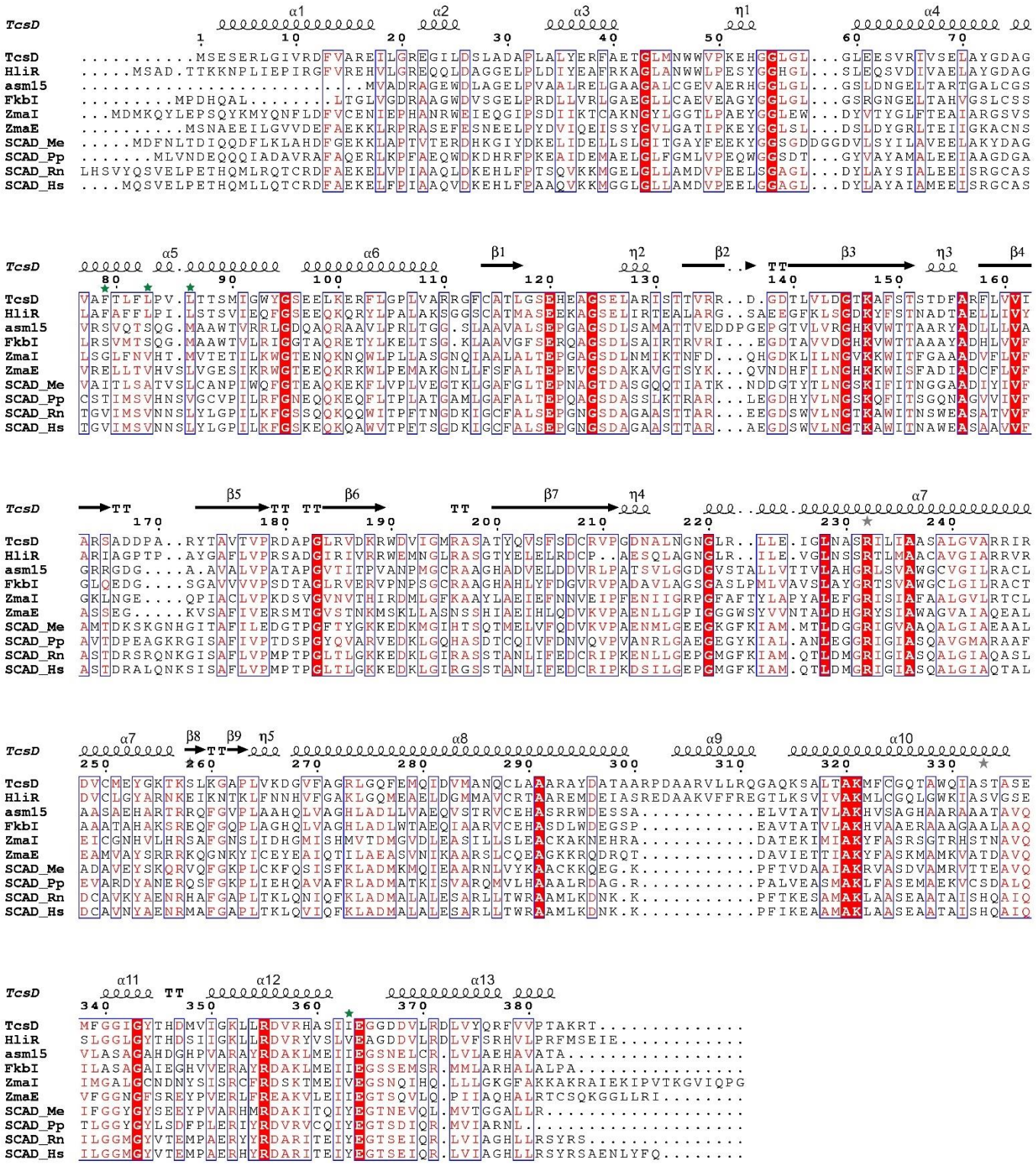



**Figure S8**. Sequence alignment of TcsD with homologs. Structural elements noted above the alignment are based on the TcsD structure. Proteins included in the alignment and their Genbank accession numbers are: γ,δ-ACADs TcsD (ADU56309.1) and HliR (AMM72019.1), hydroxymalonate semialdehyde dehydrogenases asm15 (AAM54093.1), FkbI (TAI41666.1), and ZmaE (AAD40109.1), aminomalonate semialdehyde dehydrogenase ZmaI (AAR87758.1), and α,β-ACADs SCAD_Me (pdb 1buc), SCAD_Pp (NP_744365.1), SCAD_Rn (pdb 1jqi), and SCAD_Hs (pdb 2vig). Green stars denote residues that form the active site pocket.

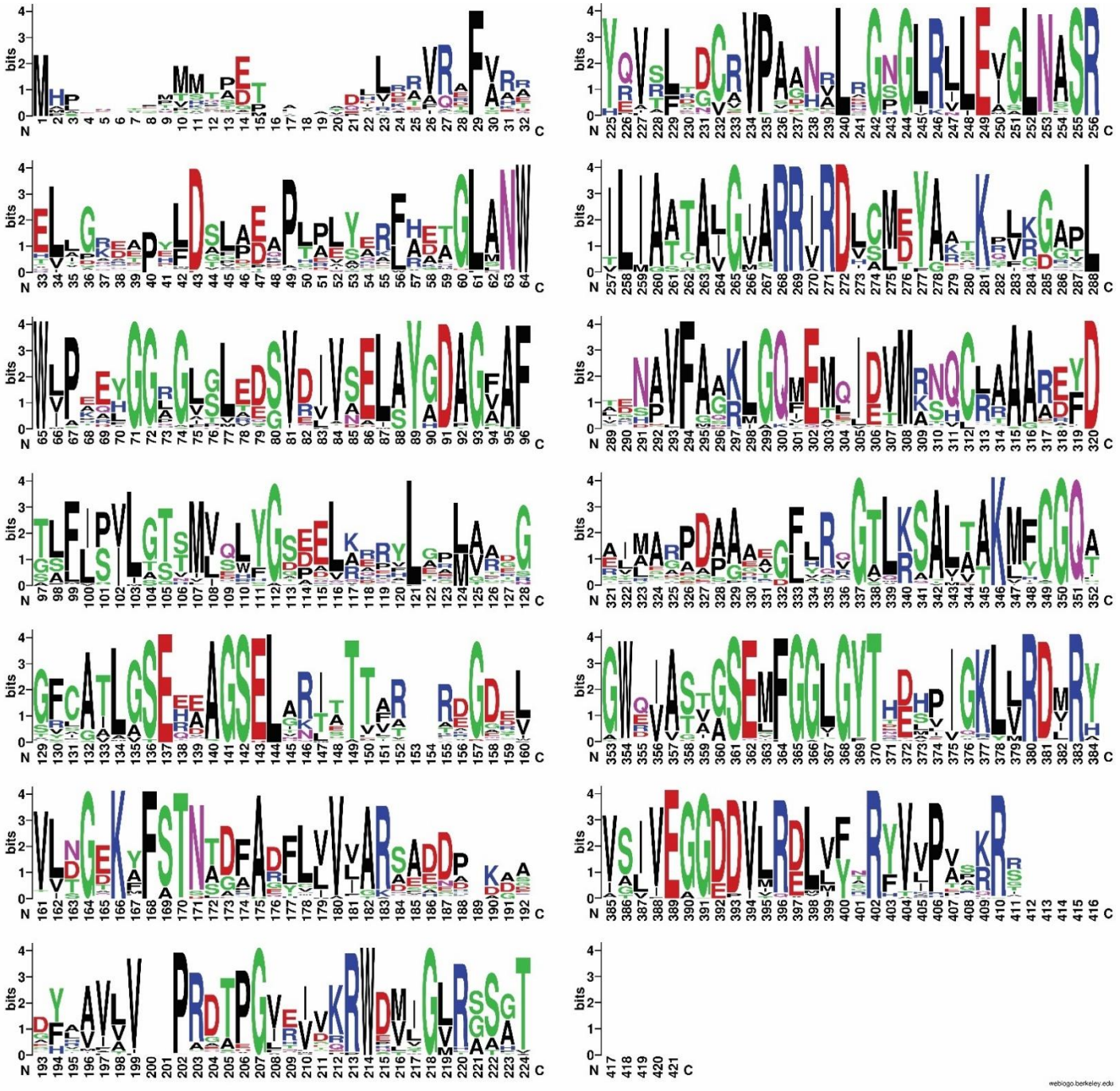



**Figure S9**. Web logo depicting the conservation of amino acids across a multiple sequence alignment of all γ,δ-ACADs identified in this work. Larger-sized letters represent higher conservation, while smaller letters represent less conserved amino acids.

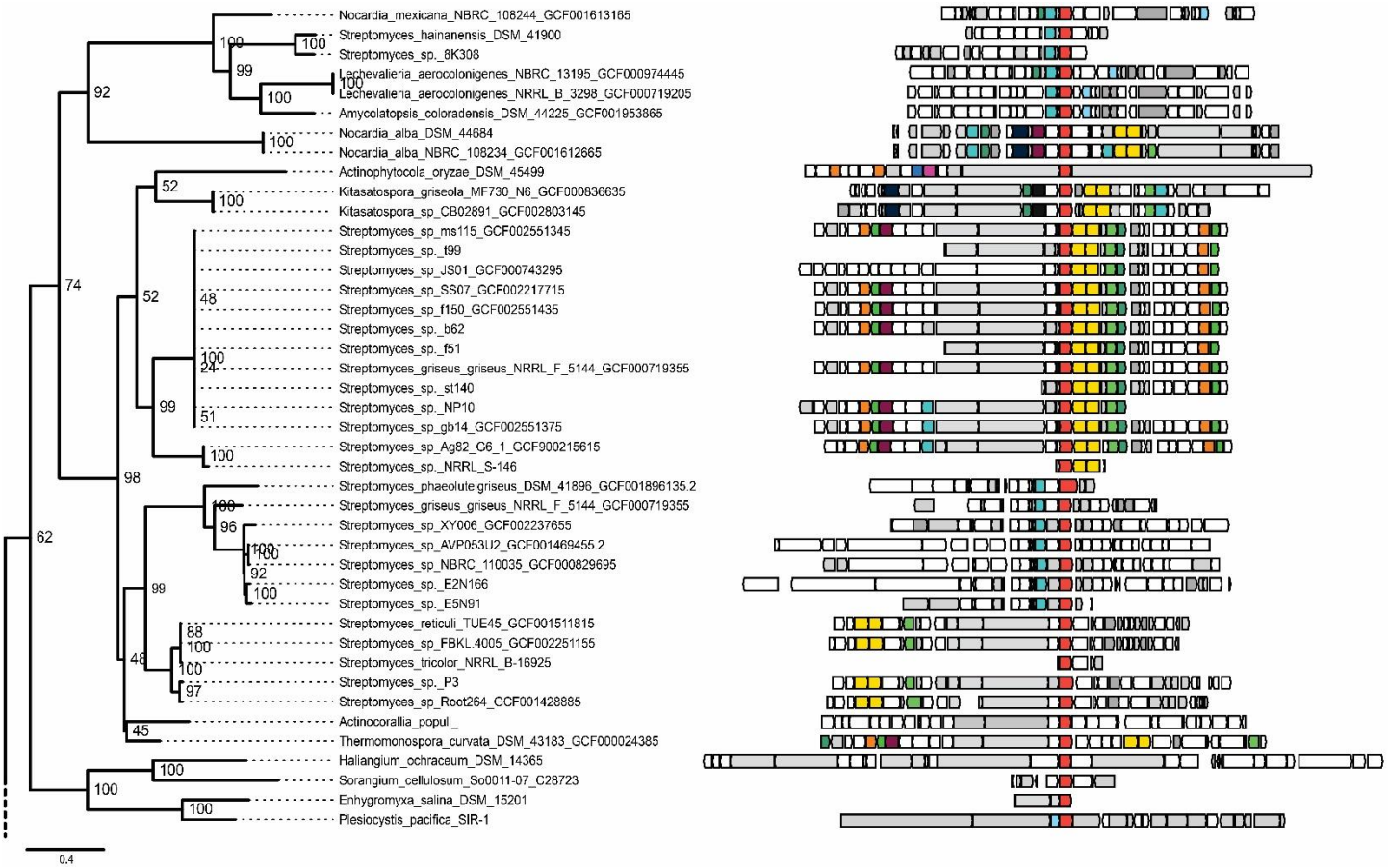



**Figure S10A**. Phylogenetic tree and genomic contexts of γ,δ-ACADs identified in this work. This figure represents the top portion of the tree and connects to the bottom portion (Figure S10B).

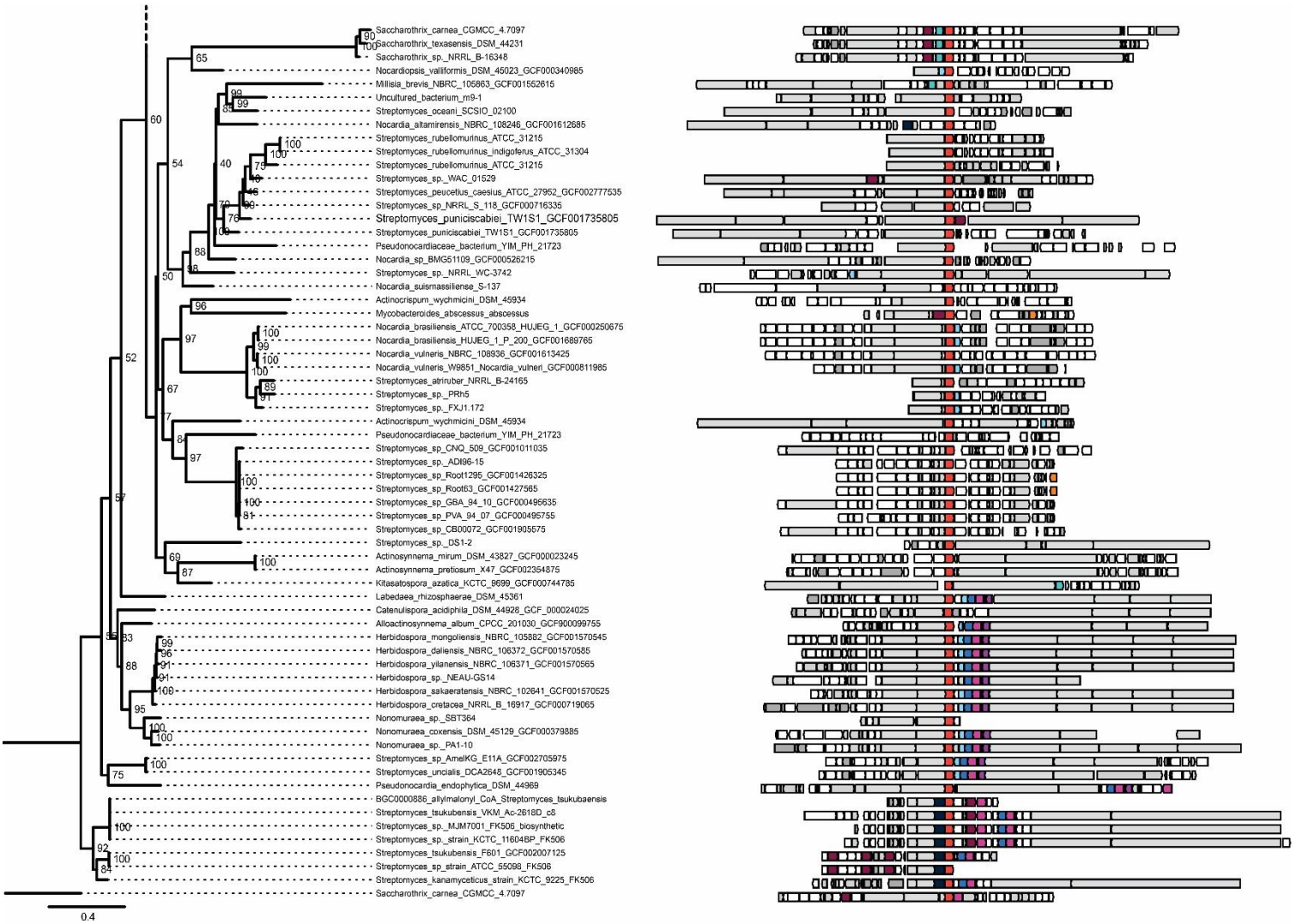



**Figure S10B**. Phylogenetic tree and genomic contexts of γ,δ-ACADs identified in this work. This figure represents the bottom portion of the tree and connects to the top portion (Figure S10A).

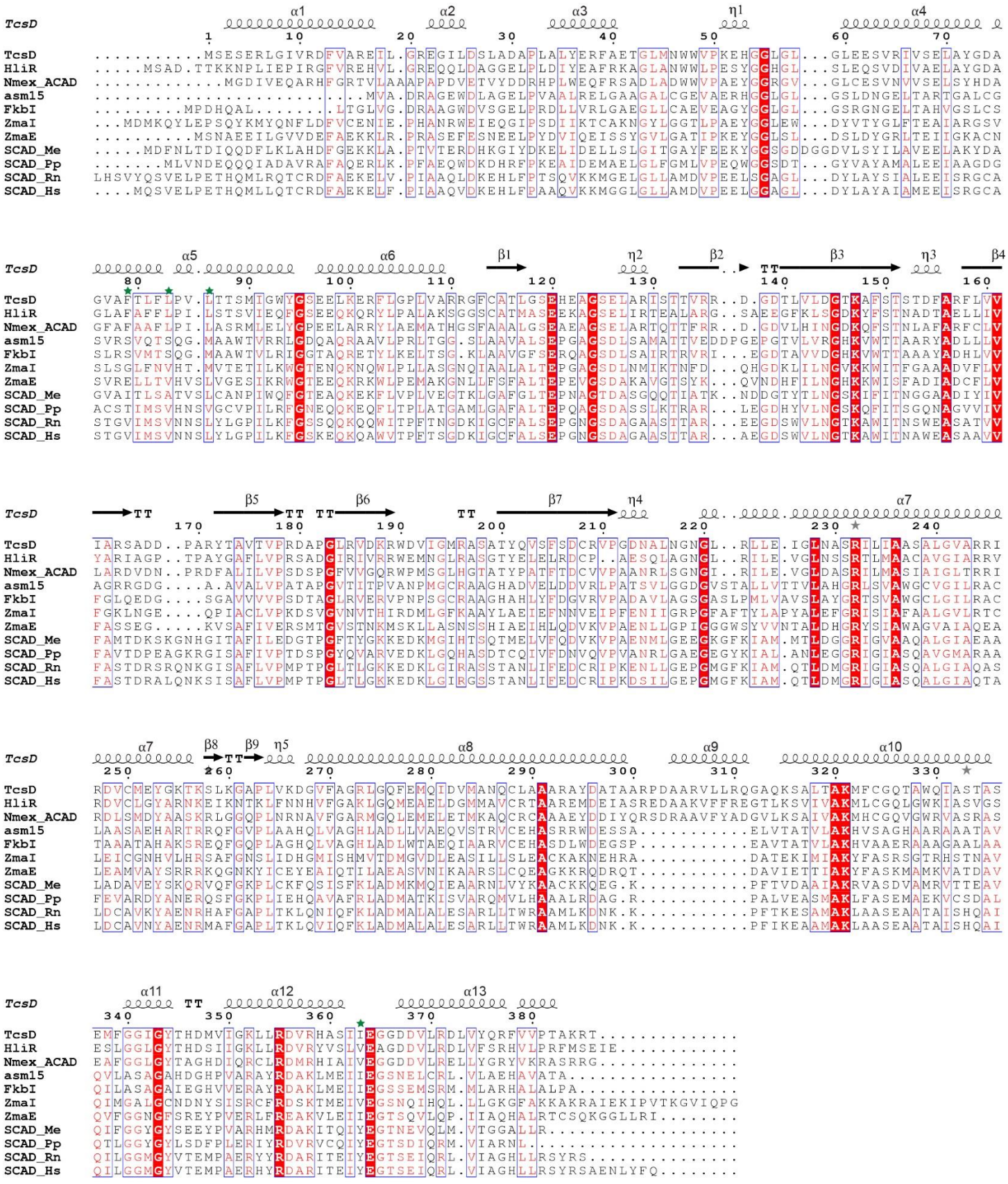
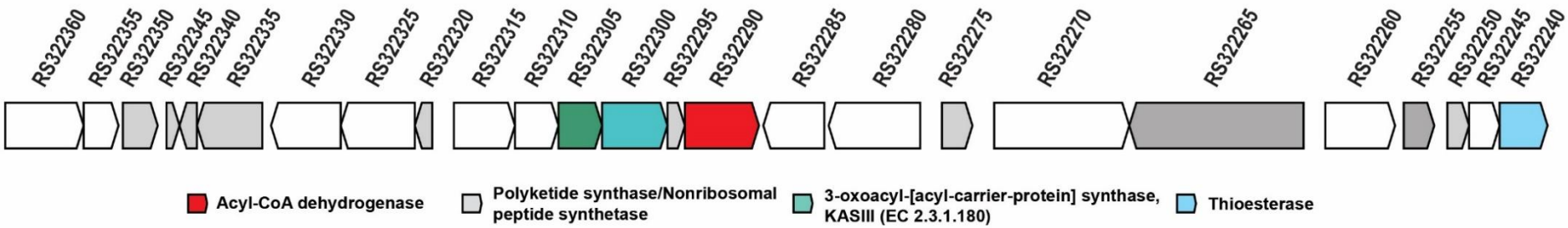



**Figure S11**. Sequence alignment of TcsD with homologs, including Nmex-ACAD. Structural elements are based on the TcsD structure, and green stars denote residues that line the active site pocket. Proteins included in the alignment and their Genbank accession numbers are listed in the description for Figure S8.




| Locus tag | Gene product | GenBank Protein ID |
|---|---|---|
| DFR68_RS32240 | thioesterase | WP_068019833.1 |
| DFR68_RS32245 | hypothetical protein | WP_068019835.1 |
| DFR68_RS32250 | hypothetical protein | WP_068019837.1 |
| DFR68_RS32255 | response regulator transcription factor | WP_068019839.1 |
| DFR68_RS32260 | histidine kinase | WP_068019841.1 |
| DFR68_RS32265 | AAA family ATPase | WP_068019843.1 |
| DFR68_RS32270 | MMPL family transporter | WP_068019846.1 |
| DFR68_RS32275 | ester cyclase | WP_084519643.1 |
| DFR68_RS32280 | MFS transporter | WP_068019850.1 |
| DFR68_RS32285 | hypothetical protein | WP_114699758.1 |
| DFR68_RS32290 | acyl-CoA dehydrogenase | WP_068019852.1 |
| DFR68_RS32295 | acyl carrier protein | WP_068019854.1 |
| DFR68_RS32300 | ketoacyl-ACP synthase III | WP_084519644.1 |
| DFR68_RS32305 | SDR family oxidoreductase | WP_068019857.1 |
| DFR68_RS32310 | HAD family phosphatase | WP_068019859.1 |
| DFR68_RS32315 | hypothetical protein | WP_084519645.1 |
| DFR68_RS32320 | acyl carrier protein | WP_068019863.1 |
| DFR68_RS32325 | beta-ketoacyl-[acyl-carrier-protein] synthase family protein | WP_068019865.1 |
| DFR68_RS32330 | hypothetical protein | WP_068019867.1 |
| DFR68_RS32335 | DUF2236 domain-containing protein | WP_068019869.1 |
| DFR68_RS32340 | hypothetical protein | WP_068019872.1 |
| DFR68_RS32345 | hypothetical protein | WP_114699759.1 |
| DFR68_RS32350 | SRPBCC family protein | WP_068019876.1 |
| DFR68_RS32355 | hypothetical protein | WP_068019878.1 |
| DFR68_RS32360 | MCE family protein | WP_084519646.1 |

**Figure S12**. Genomic context of the *Nocardia mexicana* γ,δ-acyl-CoA dehydrogenase (Nmex-ACAD) and predicted gene products for each coding sequence.

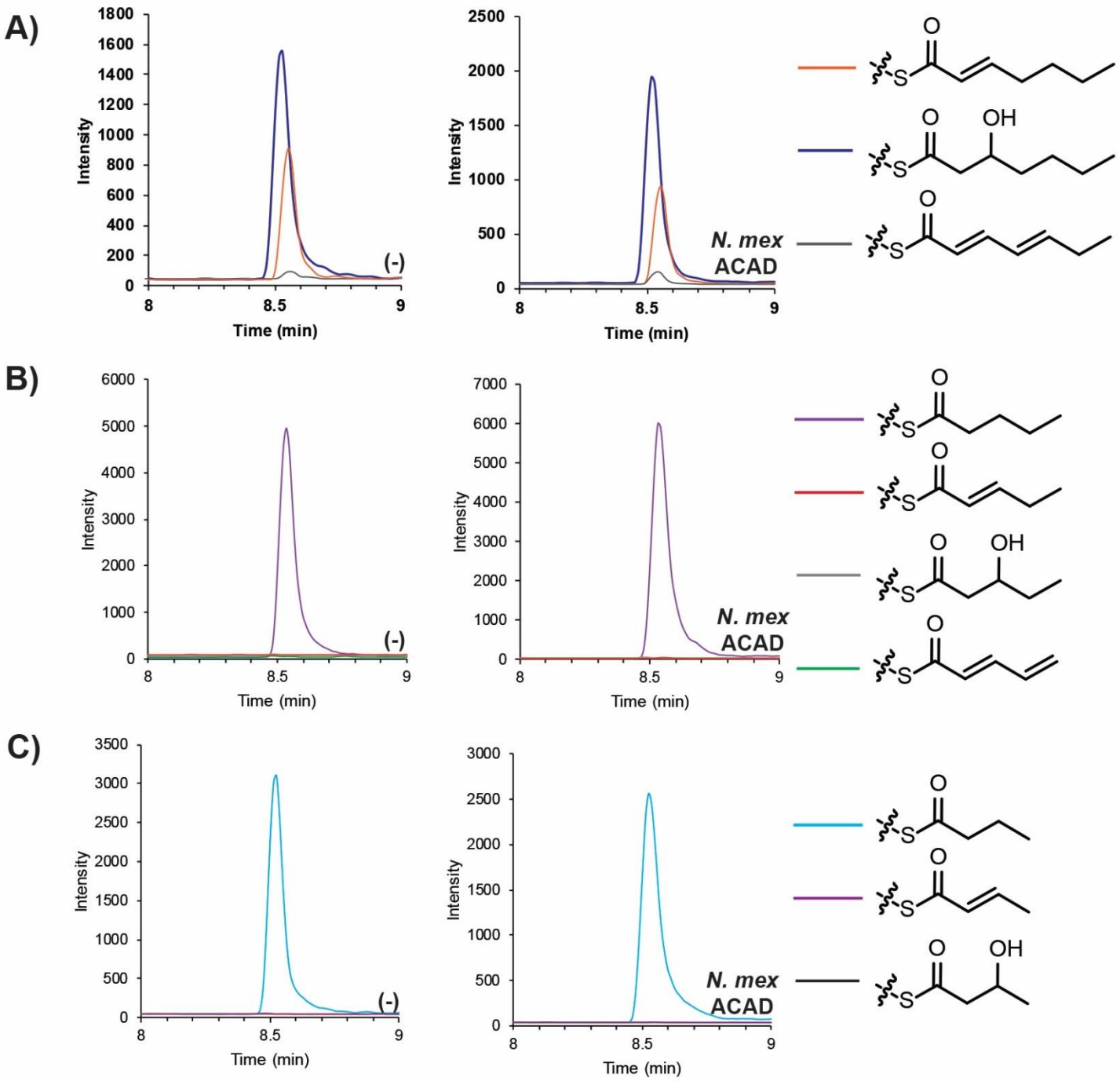



**Figure S13**. Activity of Nmex-ACAD analyzed by targeted LC-MS/MS. *N.mex* ACAD = wild type Nmex-ACAD, (-) = negative control. Chromatograms representing different transitions are color coded (key is shown on right side of figure). The chromatograms shown depict assays on the following substrates: **A)** 2-heptenoyl-ACP, **B)** pentanoyl-ACP, or **C)** butyryl-ACP. For all samples, only the substrate (no product) was observed.

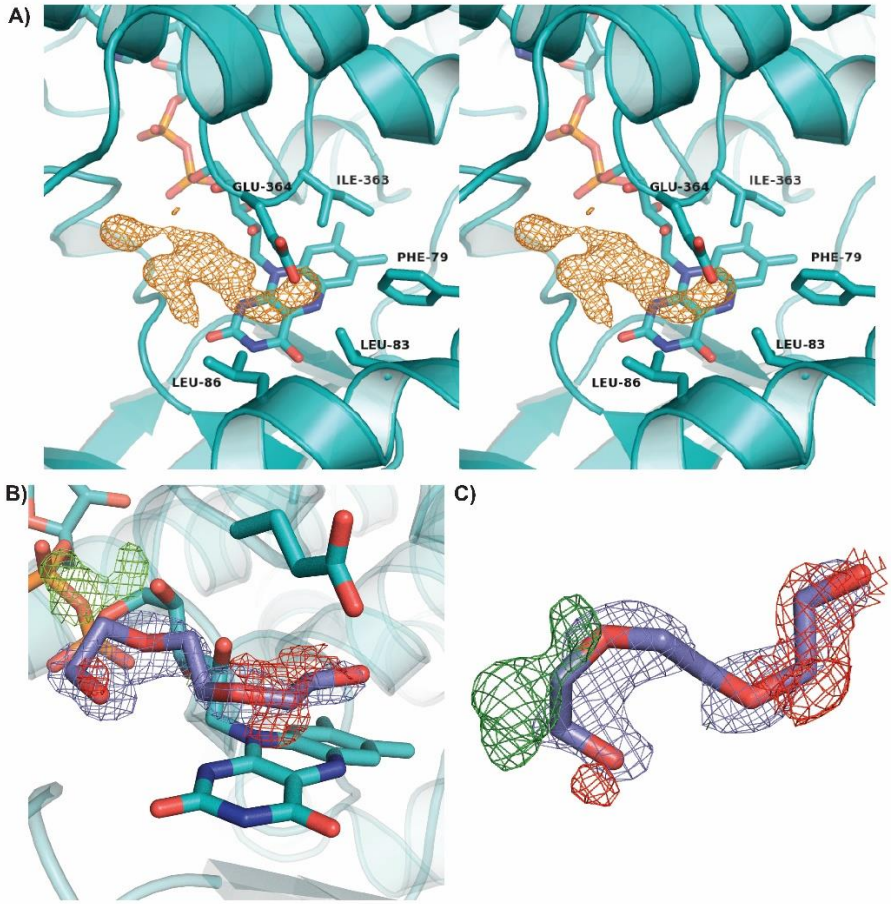



**Figure S14**. Unidentified density in active site of TcsD and modeling of PEG trimer into density. **A)** Shape of the unidentified density (in orange mesh) in the TcsD active site shown in stereo (mFo-DFc map at 3σ) **B) and C)** Model of TcsD after refinement with a PEG trimer (purple sticks) occupying the unknown density, showing that this PEG fragment does not fit properly within the density. The 2mFo-DFc map is shown in blue mesh and contoured at 1σ. The mFo-DFc difference map is shown in green (contoured at 3σ) and red (contoured at -3σ).

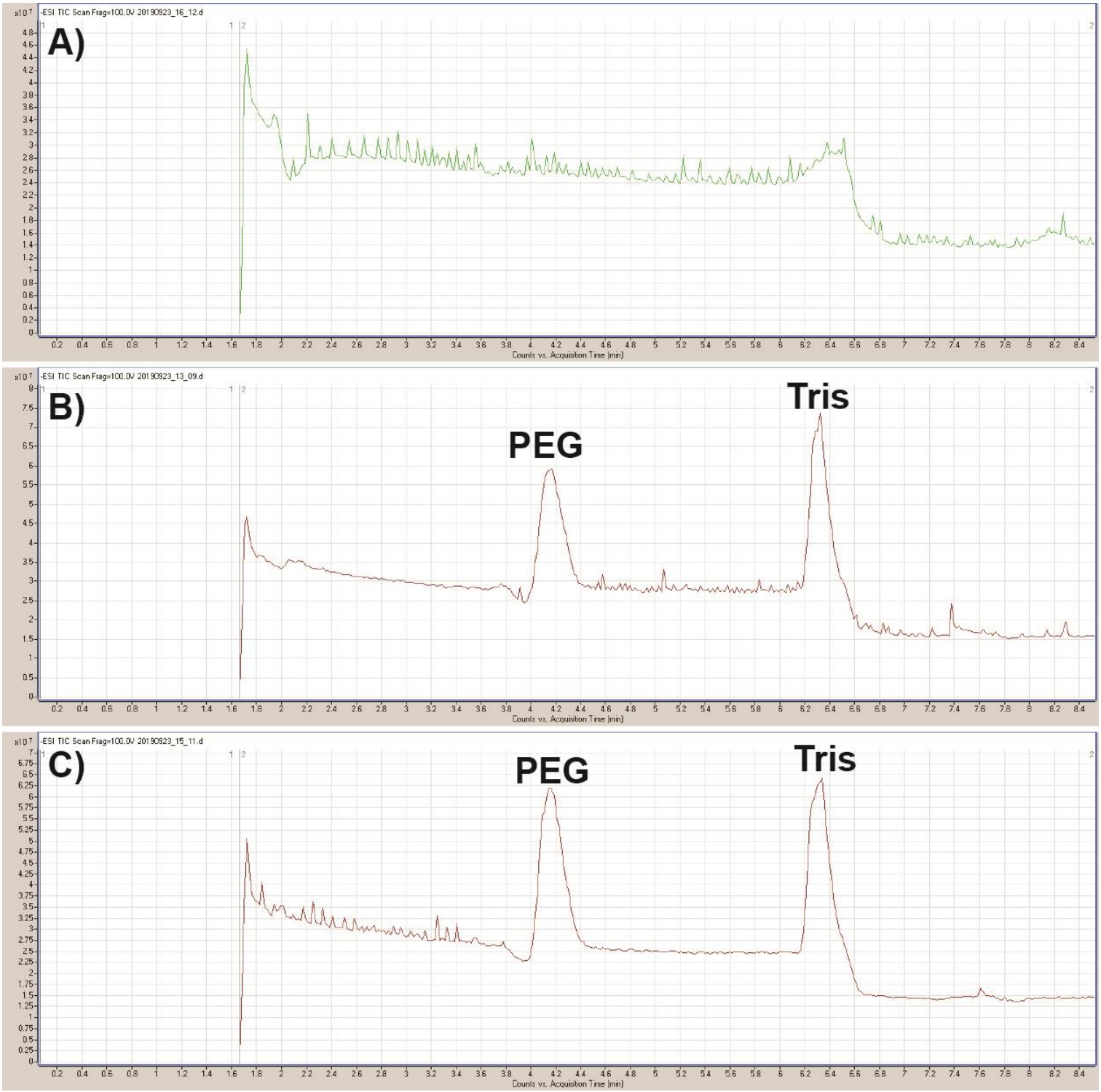



**Figure S15**. High resolution untargeted LC-TOF analysis of denatured purified TcsD samples in negative ion mode. Samples are as follows: **A)** Crystallization buffer **B)** Supernatant from boiled TcsD **C)** Supernatant from acetonitrile-denatured TcsD. Peaks pertaining to PEG (a common LCMS contaminant) and Tris are labeled.

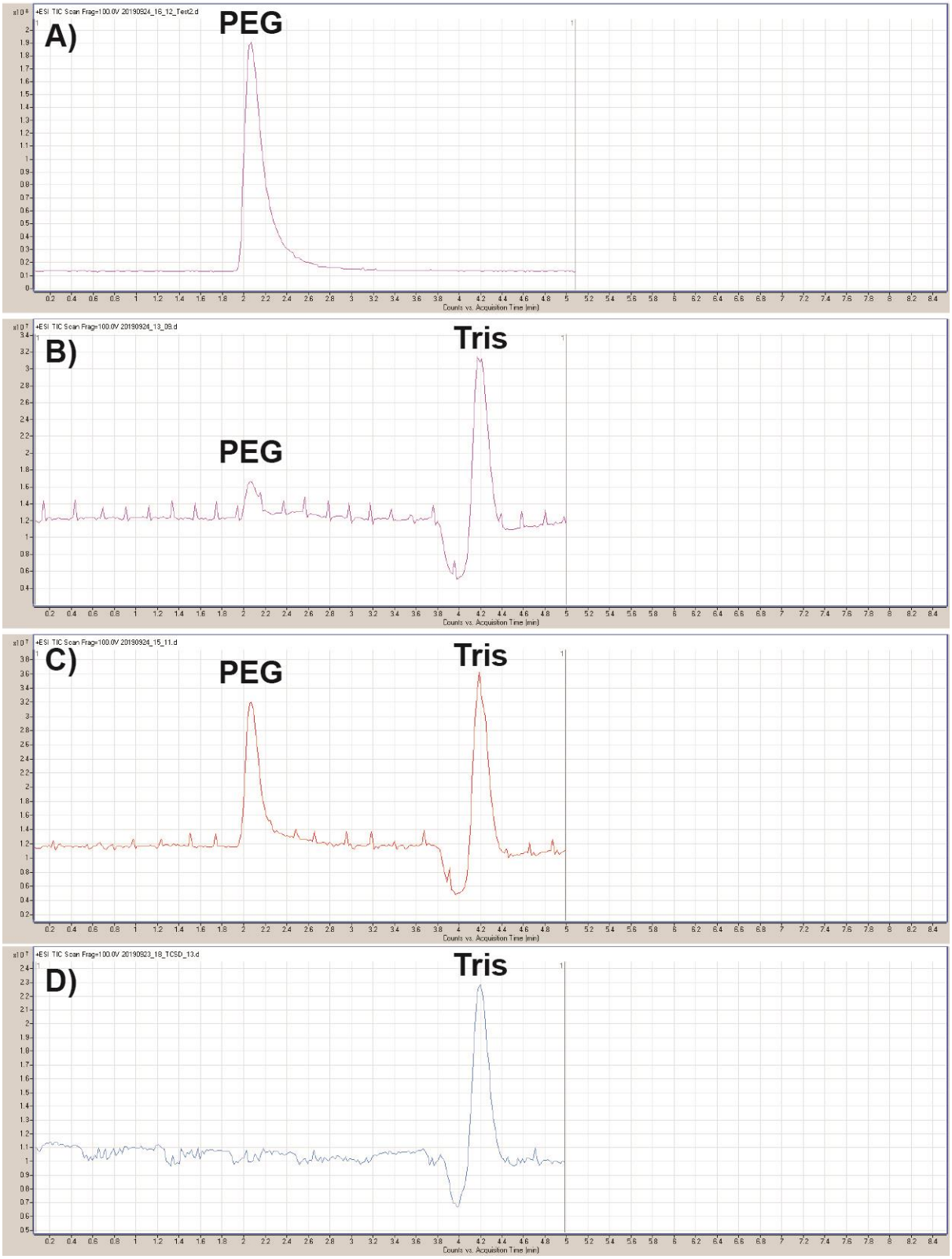



**Figure S16**. High resolution untargeted LC-TOF analysis of denatured purified TcsD samples in positive ion mode. Samples are as follows: **A)** Crystallization buffer **B)** Supernatant from boiled TcsD **C)** Supernatant from acetonitrile-denatured TcsD, **D)** lysis buffer. Peaks pertaining to PEG (a common LCMS contaminant) and Tris are labeled.

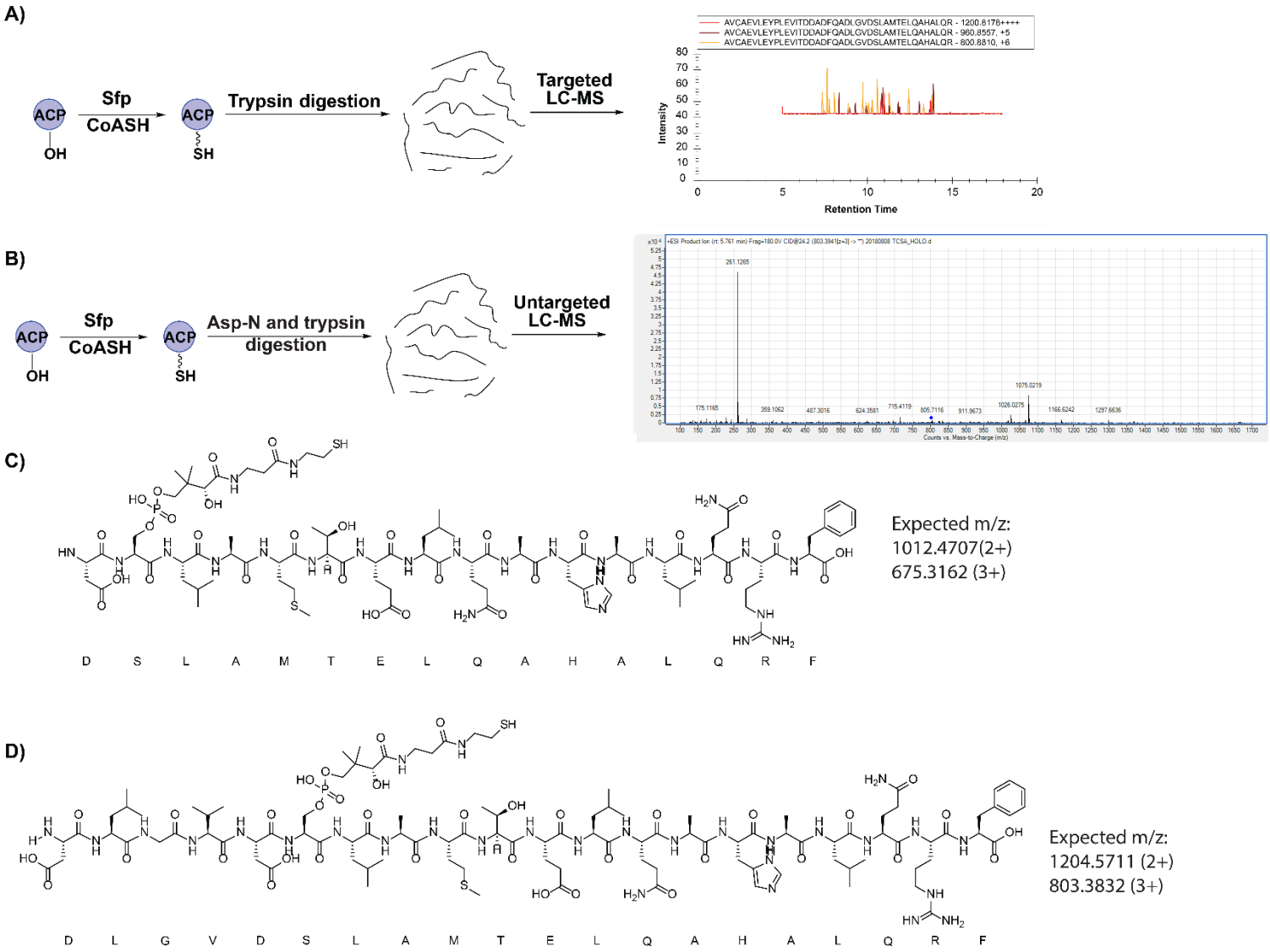



**Figure S17.** Development of a phosphopantetheine ejection method for studying intermediates bound to TcsA-ACP. **A)** Digestion of TcsA with trypsin results in an active site peptide 40 amino acids in length which is not detectable using targeted LC-MS/MS. **B)** Dual digestion of TcsA with Asp-N and trypsin results in a detectable peptide from which phosphopantetheine ejection can be observed. **C)** Predicted TcsA active site peptide resulting from TcsA digestion with Asp-N and trypsin. **D)** Observed TcsA active site peptide resulting from TcsA digestion with Asp-N and trypsin (verified with data in Figure S18).

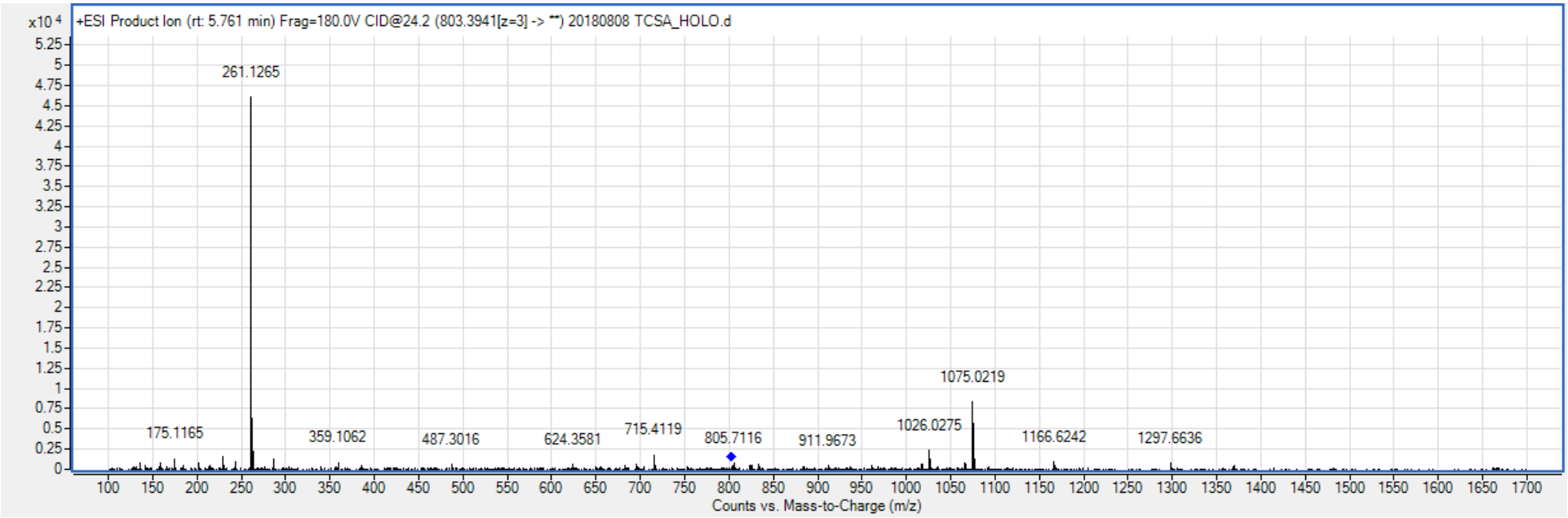



**Figure S18.** Product ion mass spectrum showing phosphopantetheine ejection ion resulting from holo-TcsA digested with Asp-N and trypsin. The product ion is derived from a parent ion with m/z = 803.394 [z=3], representing the peptide “DLGVDSLAMTELQAHALQR” in which Asp-N misses one cleavage, not the expected “DSLAMTELQAHALQR” active site peptide (see Figure S17).

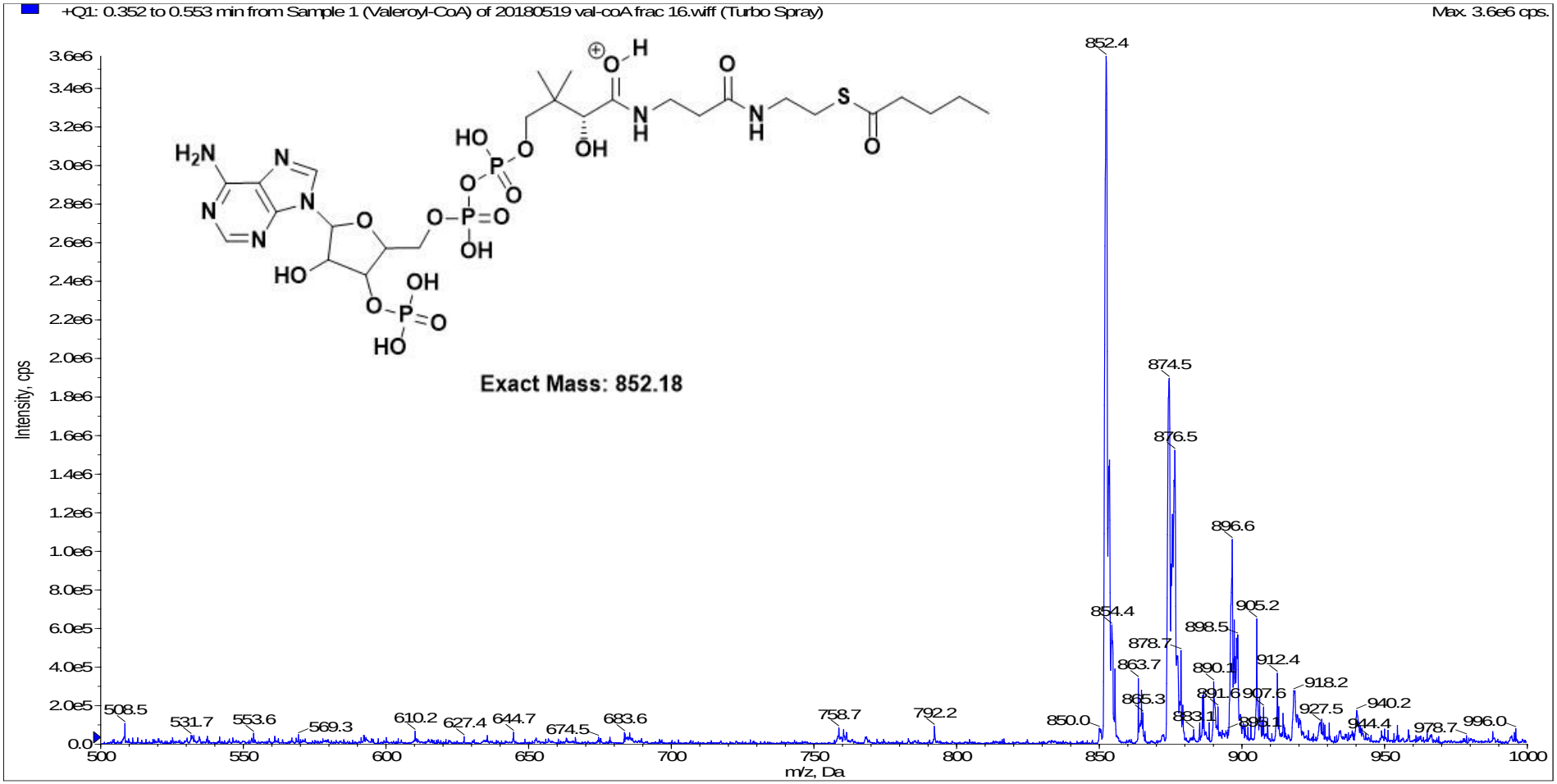



**Figure S19.** Mass spectrum of synthesized pentanoyl-CoA. An m/z of 852.18 [z=1] represents the [M+H] ion, while the m/z of 874.5 [z=1] represents the [M+Na] ion.

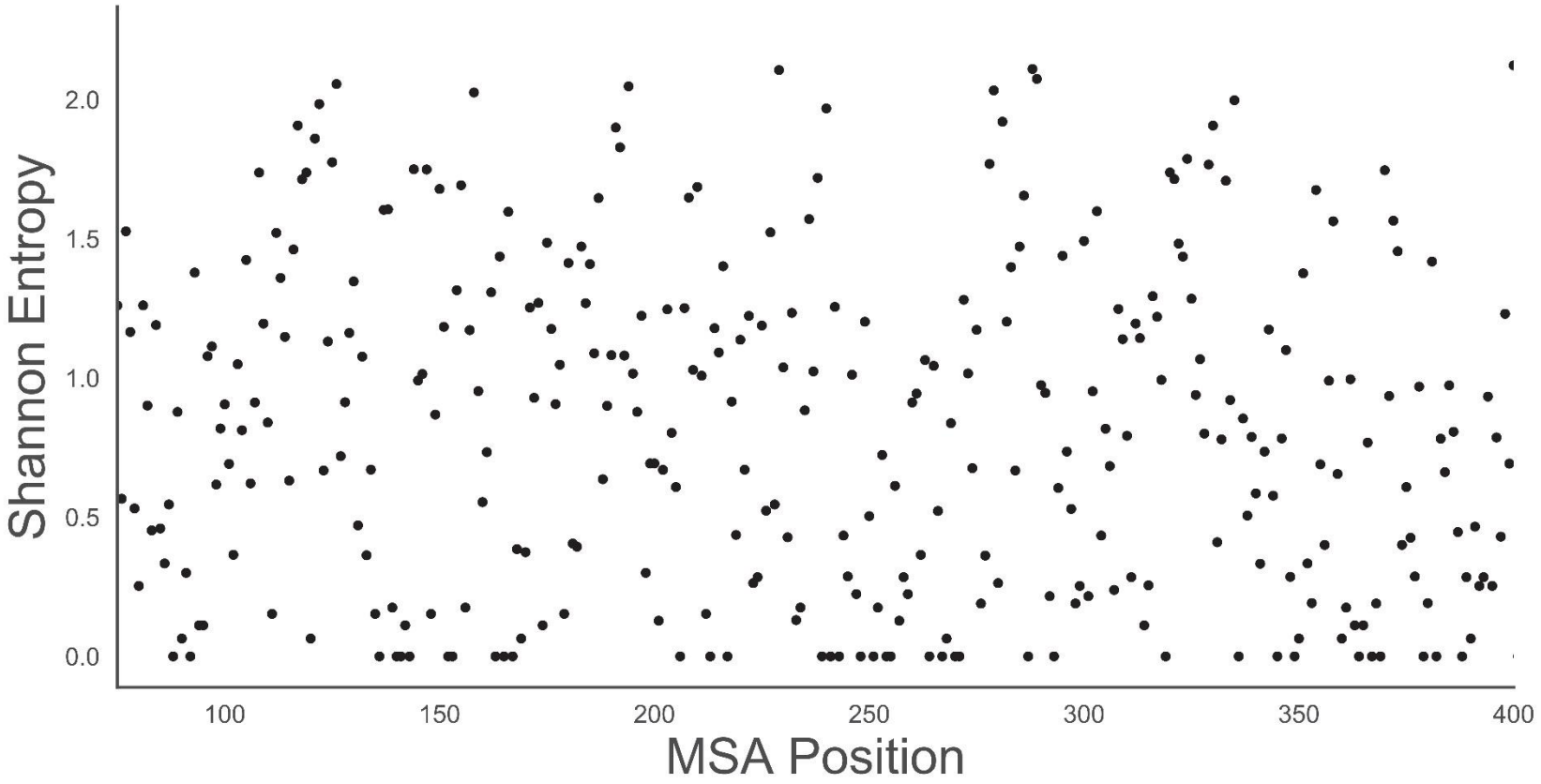



**Figure S20**. Shannon entropy plot depicting amino acid conservation at each position of a multiple sequence alignment of all γ,δ-ACADs identified in this work. An entropy value of 0 represents 100% conservation whereas larger entropy values represent less conservation. This plot was used to generate a positional conservation web logo (Figure S6).

## Data Availability

***Structural data deposition***: The atomic coordinates and structural factors of TcsD have been deposited in the Protein Data Bank, PDB ID code 6U1V.

***LC-MS/MS data deposition***: All targeted LC-MS/MS data from TcsD assays has been uploaded to Panorama Public[16] and is publicly available at the following link: https://panoramaweb.org/Structural%20control%20of%20bacterial%20acyl-CoA%20dehydrogenase.url .

***Strain availability***: All plasmids and strains generated in this work are available to the public through the Joint Bioenergy Institute's Inventory of Composable Elements: https://public-registry.jbei.org/

## Supplementary Information References


(1) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **2009**, *6*, 343–345.

(2) Engler, C.; Gruetzner, R.; Kandzia, R.; Marillonnet, S. Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. *PLoS ONE* **2009**, *4*, e5553.

(3) Oberortner, E.; Cheng, J.-F.; Hillson, N. J.; Deutsch, S. Streamlining the Design-to-Build Transition with Build-Optimization Software Tools. *ACS Synth. Biol.* **2017**, *6*, 485–496.

(4) Jancarik, J.; Kim, S. H. Sparse Matrix Sampling: A Screening Method for Crystallization of Proteins. *J. Appl. Crystallogr.* **1991**, *24*, 409–411.

(5) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.

(6) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards Automated Crystallographic Structure Refinement with Phenix.Refine. *Acta Crystallogr. D Biol. Crystallogr.* **2012**, *68*, 352–367.

(7) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.

(8) Adams, P. D.; Afonine, P. V.; Bunkóczi, 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.; Zwart, P.H.. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213–221.

(9) Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B.; Snoeyink, J.; Richardson, J. S.; Richardson, D.C. MolProbity: All-Atom Contacts and Structure Validation for Proteins and Nucleic Acids. *Nucleic Acids Res.* **2007**, *35*, W375-83.

(10) The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. https://pymol.org/2/ (accessed Jul 16, 2019).

(11) Mishra, P. K.; Drueckhammer, D. G. Coenzyme A Analogues and Derivatives: Synthesis and Applications as Mechanistic Probes of Coenzyme A Ester-Utilizing Enzymes. *Chem. Rev.* **2000**, *100*, 3283–3310.

(12) McMahon, B.; Gallagher, M. E.; Mayhew, S. G. The Protein Coded by the PP2216 Gene of Pseudomonas Putida KT2440 Is an Acyl-CoA Dehydrogenase That Oxidises Only Short-Chain Aliphatic Substrates. *FEMS Microbiol. Lett.* **2005**, *250*, 121–127.

(13) Lehman, T. C.; Hale, D. E.; Bhala, A.; Thorpe, C. An Acyl-Coenzyme a Dehydrogenase Assay Utilizing the Ferricenium Ion. *Anal. Biochem.* **1990**, *186*, 280–284.

(14) Koryakina, I.; McArthur, J.; Randall, S.; Draelos, M. M.; Musiol, E. M.; Muddiman, D. C.; Weber, T.; Williams, G. J. Poly Specific Trans-Acyltransferase Machinery Revealed via Engineered Acyl-CoA Synthetases. *ACS Chem. Biol.* **2013**, *8*, 200–208.

(15) Quadri, L. E.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P.; Walsh, C. T. Characterization of Sfp, a Bacillus Subtilis Phosphopantetheinyl Transferase for Peptidyl Carrier Protein Domains in Peptide Synthetases. *Biochemistry* **1998**, *37*, 1585–1595.

(16) Wessel, D.; Flügge, U. I. A Method for the Quantitative Recovery of Protein in Dilute Solution in the Presence of Detergents and Lipids. *Anal. Biochem.* **1984**, *138*, 141–143.

(17) Dorrestein, P. C.; Bumpus, S. B.; Calderone, C. T.; Garneau-Tsodikova, S.; Aron, Z. D.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Kelleher, N. L. Facile Detection of Acyl and Peptidyl Intermediates on Thiotemplate Carrier Domains via Phosphopantetheinyl Elimination Reactions during Tandem Mass Spectrometry. *Biochemistry* **2006**, *45*, 12756–12766.

(18) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: An Open Source Document Editor for Creating and Analyzing Targeted Proteomics Experiments. *Bioinformatics* **2010**, *26*, 966–968.

(19) Sharma, V.; Eckels, J.; Schilling, B.; Ludwig, C.; Jaffe, J. D.; MacCoss, M. J.; MacLean, B. Panorama Public: A Public Repository for Quantitative Data Sets Processed in Skyline. *Mol. Cell. Proteomics* **2018**, *17*, 1239–1244.

(20) Baidoo, E. E. K.; Wang, G.; Joshua, C. J.; Benites, V. T.; Keasling, J. D. Liquid Chromatography and Mass Spectrometry Analysis of Isoprenoid Intermediates in Escherichia Coli. *Methods Mol. Biol.* **2019**, *1859*, 209–224.

(21) Finn, R. D.; Clements, J.; Eddy, S. R. HMMER Web Server: Interactive Sequence Similarity Searching. *Nucleic Acids Res.* **2011**, *39*, W29-37.
(22) Medema, M. H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R. AntiSMASH: Rapid Identification, Annotation and Analysis of Secondary Metabolite Biosynthesis Gene Clusters in Bacterial and Fungal Genome Sequences. *Nucleic Acids Res.* **2011**, *39*, W339-46.
(23) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410.
(24) Cruz-Morales, P.; Ramos-Aboites, H. E.; Licona-Cassani, C.; Selem-Mójica, N.; Mejía-Ponce, P. M.; Souza-Saldívar, V.; Barona-Gómez, F. Actinobacteria Phylogenomics, Selective Isolation from an Iron Oligotrophic Environment and Siderophore Functional Characterization, Unveil New Desferrioxamine Traits. *FEMS Microbiol. Ecol.* **2017**, *93*.
(25) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J.D.; Higgins, D.G. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539.
(26) Robert, X.; Gouet, P. Deciphering Key Features in Protein Structures with the New ENDscript Server. *Nucleic Acids Res.* **2014**, *42*, W320-4.
(27) Katoh, K.; Kuma, K.; Toh, H.; Miyata, T. MAFFT Version 5: Improvement in Accuracy of Multiple Sequence Alignment. *Nucleic Acids Res.* **2005**, *33*, 511–518.
(28) Okonechnikov, K.; Golosova, O.; Fursov, M.; UGENE team. Unipro UGENE: A Unified Bioinformatics Toolkit. *Bioinformatics* **2012**, *28*, 1166–1167.
(29) Bakan, A.; Dutta, A.; Mao, W.; Liu, Y.; Chennubhotla, C.; Lezon, T. R.; Bahar, I. Evol and ProDy for Bridging Protein Sequence Evolution and Structural Dynamics. *Bioinformatics* **2014**, *30*, 2681–2683.
(31) Bakan, A.; Dutta, A.; Mao, W.; Liu, Y.; Chennubhotla, C.; Lezon, T. R.; Bahar, I. Evol and ProDy for Bridging Protein Sequence Evolution and Structural Dynamics. *Bioinformatics* **2014**, *30*, 2681–2683.
(32) Terwilliger, T. C.; Klei, H.; Adams, P. D.; Moriarty, N. W.; Cohn, J. D. Automated Ligand Fitting by Core-Fragment Fitting and Extension into Density. *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 915–922.
(33) Terwilliger, T. C.; Adams, P. D.; Moriarty, N. W.; Cohn, J. D. Ligand Identification Using Electron-Density Map Correlations. *Acta Crystallogr. D Biol. Crystallogr.* **2007**, *63*, 101–107.
(34) Curran, S. C.; Hagen, A.; Poust, S.; Chan, L. J. G.; Garabedian, B. M.; de Rond, T.; Baluyot, M.-J.; Vu, J. T.; Lau, A. K.; Yuzawa, S.; Petzold, C.J.; Katz, L.; Keasling, J.D. Probing the Flexibility of an Iterative Modular Polyketide Synthase with Non-Native Substrates *in Vitro*. *ACS Chem. Biol.* **2018**, *13*, 2261–2268.
(35) Lee, T. S.; Krupa, R. A.; Zhang, F.; Hajimorad, M.; Holtz, W. J.; Prasad, N.; Lee, S. K.; Keasling, J. D. BglBrick Vectors and Datasheets: A Synthetic Biology Platform for Gene Expression. *J. Biol. Eng.* **2011**, *5*, 12.