

Insights into Substrate Modification by Dehydratases from Type I Polyketide Synthases

Alexandre Faille, Sabine Gavalda, Nawel Slama, Christian Lherbet, Laurent Maveyraud, Valérie Guillet, Françoise Laval, Annaïk Quémard, Lionel Mourey, Jean-Denis Pedelacq

▶ **To cite this version:**

Alexandre Faille, Sabine Gavalda, Nawel Slama, Christian Lherbet, Laurent Maveyraud, et al.. Insights into Substrate Modification by Dehydratases from Type I Polyketide Synthases. Journal of Molecular Biology, 2017, 429 (47), pp.1554 - 1569. 10.1016/j.jmb.2017.03.026 . hal-03002100

HAL Id: hal-03002100
https://hal.science/hal-03002100

Submitted on 20 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Insights into Substrate Modification by Dehydratases from Type I Polyketide Synthases

Alexandre Faille[1,†], Sabine Gavalda[1,†], Nawel Slama[1], Christian Lherbet[2], Laurent Maveyraud[1], Valérie Guillet[1], Françoise Laval[1], Annaïk Quémard[1], Lionel Mourey[1] and Jean-Denis Pedelacq[1]

1 - Institut de Pharmacologie et de Biologie Structurale, *Université de Toulouse, CNRS, UPS, 31077 Toulouse Cedex 04, France*
2 - ITAV USR 3505, *31106 Toulouse Cedex 1, France*

Correspondence to Lionel Mourey and Jean-Denis Pedelacq:. *Lionel.Mourey@ipbs.fr; Jean-Denis.Pedelacq@ipbs.fr.*
http://dx.doi.org/10.1016/j.jmb.2017.03.026
Edited by Georg Schulz

Abstract

Dehydration reactions play a crucial role in the *de novo* biosynthesis of fatty acids and a wide range of pharmacologically active polyketide natural products with strong emphasis on human medicine. The type I polyketide synthase PpsC from *Mycobacterium tuberculosis* catalyzes key biosynthetic steps of lipid virulence factors phthiocerol dimycocerosates and phenolic glycolipids. Given the insolubility of the natural C_{28}–C_{30} fatty acyl substrate of the PpsC dehydratase (DH) domain, we investigated its structure–function relationships in the presence of shorter surrogate substrates. Since most enzymes belonging to the (R)-specific enoyl hydratase/hydroxyacyl dehydratase family conduct the reverse hydration reaction *in vitro*, we have determined the X-ray structures of the PpsC DH domain, both unliganded (apo) and in complex with *trans*-but-2-enoyl-CoA or *trans*-dodec-2-enoyl-CoA derivatives. This study provides for the first time a snapshot of dehydratase–ligand interactions following a hydration reaction. Our structural analysis allowed us to identify residues essential for substrate binding and activity. The structural comparison of the two complexes also sheds light on the need for long acyl chains for this dehydratase to carry out its function, consistent with both its *in vitro* catalytic behavior and the physiological role of the PpsC enzyme.
© 2017 Published by Elsevier Ltd.

Introduction

Polyketides represent the most important class of naturally synthesized metabolites in human health with proven antibacterial, anticholesterolemic, immunosuppressant, antifungal, anticancer, antiparasitic, and insecticide activities [1]. They are biosynthesized by complex machineries, named polyketide synthases (PKSs), that human pathogens such as *Mycobacterium tuberculosis* use in combination with fatty acid synthases (FASs) to produce lipid virulence factors. These polyketide-derived lipids form a permeability barrier within the mycobacterial cell envelope, which constitutes the first line of defense against toxic molecules from the host and confers intrinsic resistance against treatments [2,3]. Among the virulence factors, phthiocerol dimycocerosates (DIMs) participate directly into the early steps of microbial infection. Although the underlying mechanism is still a matter of debate, two hypotheses have emerged: a host plasma membrane rearrangement that encourages phagocytosis and inhibits phagosome acidification [4], or a masking role of pathogen-associated molecular patterns that could afford some protection against multiple host effectors [5].

The *M. tuberculosis* genes involved in the biosynthesis of DIMs and closely related phenolic glycolipids (PGLs) are clustered in a 70-kb region of the chromosome [6]. Biosynthesis of these components implies the concerted action of an ensemble of six multidomain (type I) PKSs. Five monomodular PKSs (PpsA/B/C/D/E) are responsible for the elongation of long-chain fatty acid backbones that form the phthiocerol (or phenolphthiocerol in the case of

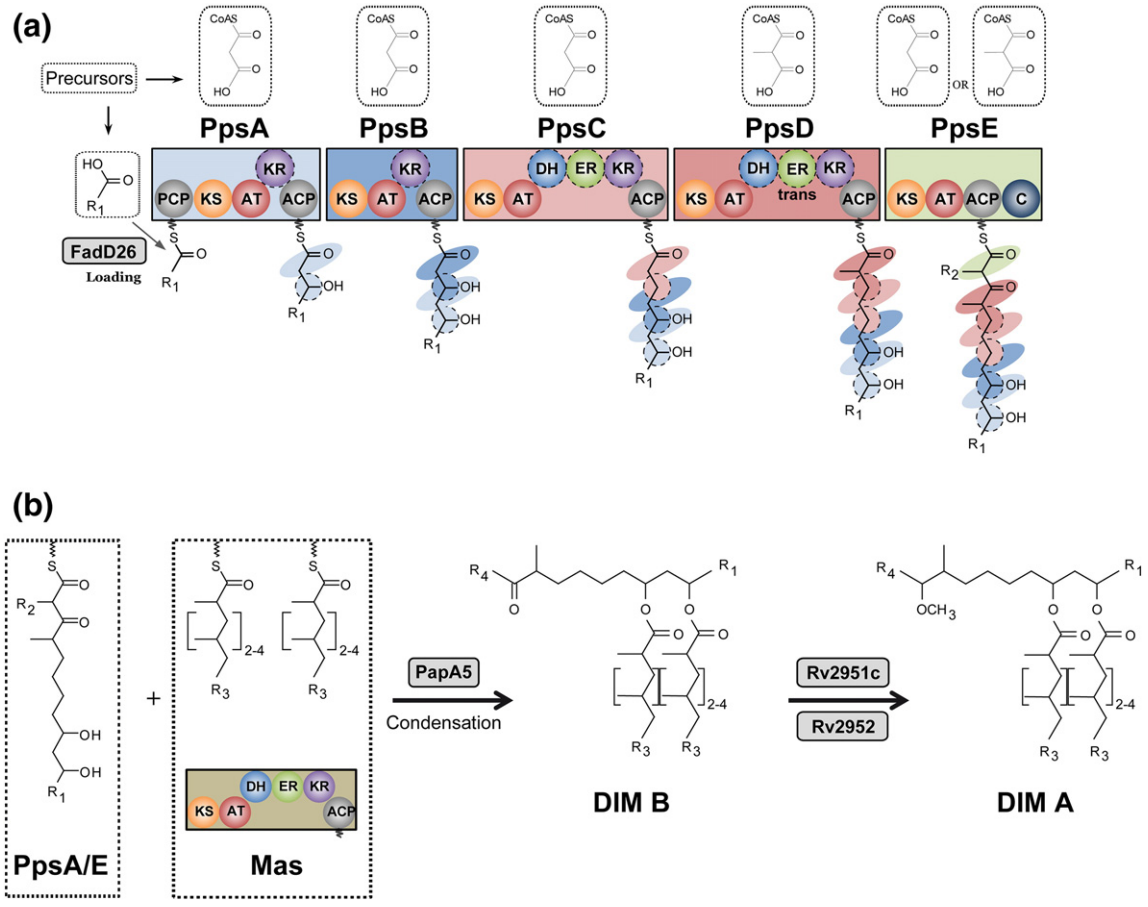


Fig. 1. Schematic representation of enzymes and products involved in phthiocerol dimycocerosate biosynthesis. (a) PpsA, PpsB, PpsC, PpsD, and PpsE are type I monomodular PKSs, which catalyze the formation of the C_{32}-C_{34} long β-diol chain called phthiocerol. Precursors are embedded in dashed rectangles. Within each Pps, essential (bottom) and optional (top) domains participate to the decarboxylative Claisen-type condensation and subsequent modifications at position C3, respectively. In the case of PpsD, there is no ER domain. Instead, the *trans*-acting enoylreductase Rv2953 catalyzes the reduction of the carbon–carbon double bond [57]. (b) Mas is a type-I iterative PKS, which catalyzes the formation of C_{24}-C_{30} long methyl-branched mycocerosic acids. The condensation reaction is achieved by the PapA5 acyltransferase to produce DIM B [10]. Further reduction and methylation reactions to form DIM A are achieved by Rv2951c [11] and Rv2952 [12], respectively; adapted from Trivedi *et al.* [9]. $R_1 = CH_3$-$(CH_2)_{20–22}$; R_2 = H, CH_3; $R_3 = (CH_2)_{16–18}$-CH_3; $R_4 = CH_3$, C_2H_5.

PGL biosynthesis) moiety. The iterative mycocerosic acid synthase (Mas) is involved in the production of multi-methyl branched mycocerosates, which esterify the hydroxy groups of the (phenol)phthiocerol backbone (Fig. 1). Phthiocerol biosynthesis starts with the activation and transfer of a C_{22}–C_{24} fatty acyl chain onto the peptidyl carrier protein (PCP) domain of PpsA [7] by the FadD26 fatty acyl-AMP ligase (FAAL) [8]. The acyltransferase (AT) domain, often referred to as the “gatekeeper”, catalyzes the transfer of malonyl-CoA to its cognate acyl carrier protein (ACP) domain. The ketosynthase (KS) domain of PpsA catalyzes the Claisen condensation of the precursor with the malonyl-CoA elongation unit to give an ACP-bound β-keto ester longer by two methylene units, followed by a reduction of the β-ketone with the help of the ketoreductase (KR) domain. The growing polyketide is then transferred to the ACP domain of PpsB to undergo the same treatment. Next, PpsC and PpsD elongate the precursor using malonyl- and methyl malonyl-CoA, respectively. PpsC and PpsD also catalyze the reduction of the previously incorporated β-ketones using a full set of β-carbon processing domains. First, the β-keto ester undergoes a reduction by the KR domain to form a β-hydroxyl subsequently modified by the dehydratase (DH) domain, thus resulting in a carbon–carbon double bond formation. The growing chain is then reduced by the enoylreductase (ER) domain to give a saturated carbon chain. Finally, PpsE catalyzes the condensation of the C_{30}–C_{32} acyl chain with either a malonyl- or a

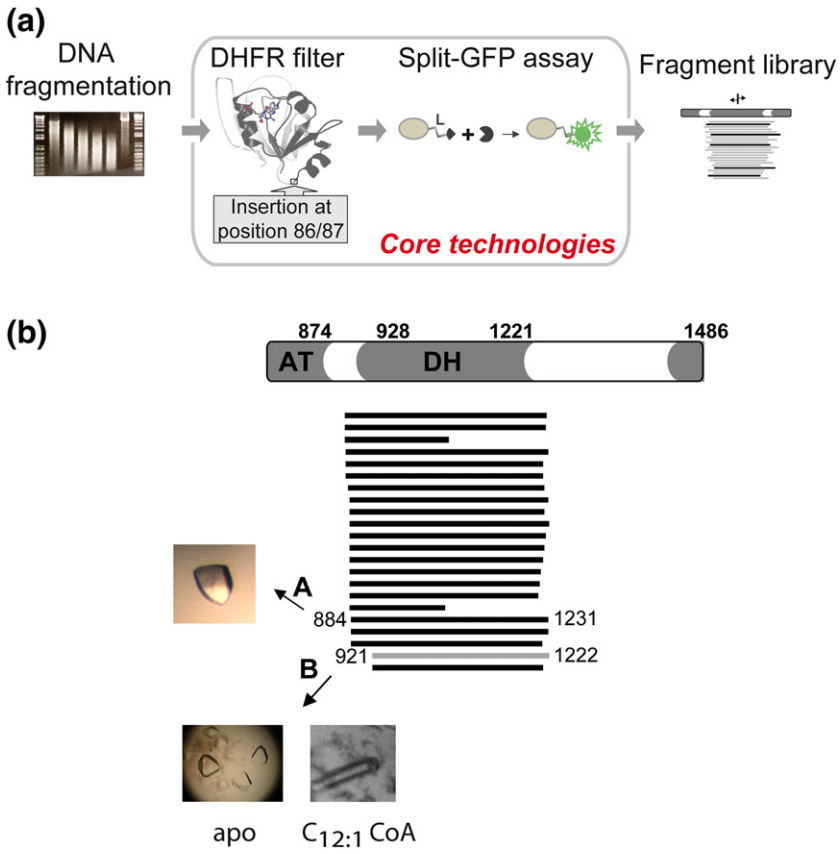


Fig. 2. The domain trapping strategy. (a) Our strategy integrates a DHFR filter to eliminate the incorrect reading frames and a split-GFP solubility screen to identify the soluble constructs. A 850–1500 bp library of fragments was inserted at position 86/87 of the DHFR. Inverse-PCR focused sublibraries of fragments helped in identifying soluble fragments centered onto a specific region of the protein of interest. (b) The top 25% most soluble clones centered onto the DH of the *Mycobacterium tuberculosis* polyketide synthase PpsC are represented. The fragment that permitted the X-ray structure determination is in light gray.

methyl malonyl-CoA to give the phthiocerol product (Fig. 1a). In parallel, mycocerosic acid biosynthesis starts with the transfer of a C_{18}–C_{20} by FadD28 FAAL [8] onto the ACP domain of Mas, which then iteratively elongates and processes the polyketide, with methyl malonyl-CoA as elongation unit [9]. Resulting mycocerosic acids are then transferred onto the phthiocerol moiety by the acyltransferase PapA5 to produce DIM B [10]. DIM B can be further reduced by the phthiodiolone KR encoded by the *Rv2951c* gene [11] and then methylated by the methyltransferase encoded by *Rv2952* [12] to give DIM A. Both DIM A and DIM B contribute to the virulence of *M. tuberculosis* [13].

To date, none of the accessible structures of DHs from modular PKSs has been determined with a metabolizable ligand in its active site. Only complexes involving the type II FAS DH FabA [14,15] and the β-oxidation hydratase Mfe2p-H2 [16] have been reported so far. Knowing that some dehydratases catalyze the reverse hydration reaction more efficiently than the forward dehydration reaction *in vitro* [17–20] and considering the insolubility of natural C_{28}–C_{30} fatty acyl chains, we investigated the structure–function relationships of the PpsC DH domain (DH_{PpsC}) in the presence of three shorter surrogate substrates: *trans*-but-2-enoyl-CoA ($C_{4:1}$-CoA), *trans*-dec-2-enoyl-CoA ($C_{10:1}$-CoA), and *trans*-dodec-2-enoyl-CoA ($C_{12:1}$-CoA). This study points out the critical role of longer acyl chains for the DH to carry out its function and provides a structural explanation as to why it is active in the presence of $C_{12:1}$-CoA but not $C_{4:1}$-CoA.

Results and Discussion

The PpsC dehydratase domain fold

A library of fragments centered onto the DH domain of PpsC from *M. tuberculosis* was first created to select for well-expressed and soluble constructs using the domain trapping method [21] (Fig. 2a). DH

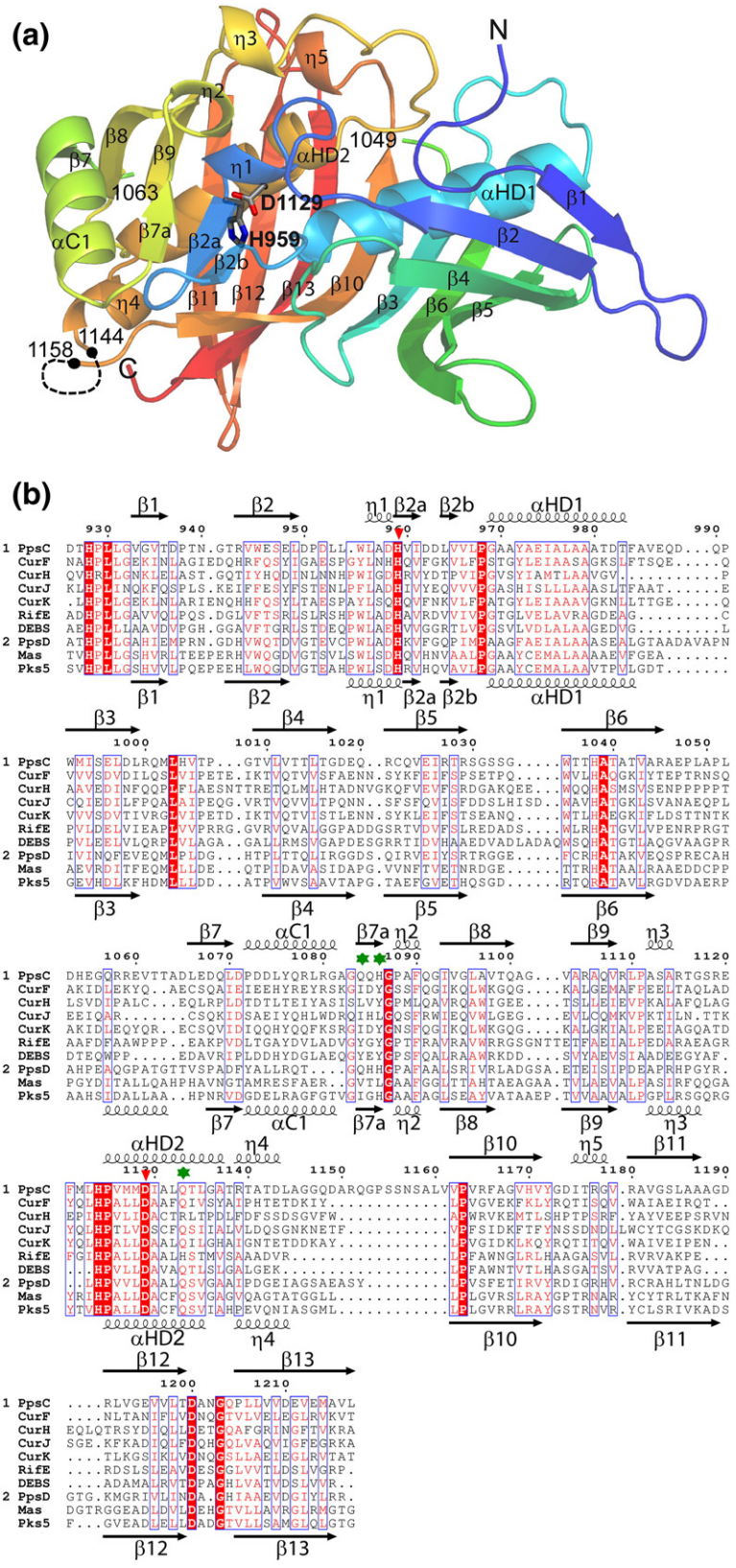


Fig. 3. *(legend on the next page)*

Table I. Structural neighbors of DH_{PpsC} accessible in the PDB

Name	Origin	Pathway	PDB IDs	% identity	rmsd[a]	Z–scores	Ligand[b]	Refs
DEHYDRATASE								
PpsC	*Mycobacterium tuberculosis*	Modular PKS	5nji	100	0		$C_{12:1}$–CoA	*
			4p7p	100	0.7	42.9	$C_{4:1}$–CoA	*
Pks5	*Mycobacterium smegmatis*		5bp2 / 5bp3	26	1.9	32.8	none	[22]
CurF	*Lyngbya majuscula*		3kg6	21	2.1	32.8	none	[23]
CurK			3kg7	19	2.2	31.7	none	[23]
CurJ			3kg8	24	2.1	31.1	none	
CurH			3kg9	24	2.2	29.5	none	
RifDH10	*Amycolatopsis mediterranei*		4ln9	24	2.5	28.4	none	[24]
DEBS	*Saccharopolyspora erythraea*		3el6	27	2.6	27.9	none	[25]
mFAS	*Sus scrofa*	FAS	2vz8 / 2vz9	19	2.5	22.0	none	[52]
FabZ	*Plasmodium falciparum*		3az8	12	2.6	14.4	S21	[53]
			3az9	12	2.6	14.4	K91	
			3aza	12	2.6	14.4	KM0	
			3azb	12	2.6	14.4	KM1	
	Helicobacter pylori		4zjb	13	2.8	12.6	ACP	[54]
FabA	*Pseudomonas aeruginosa*		4b0i	11	2.8	11.4	KBP	[15]
	Escherichia coli		4keh	11	2.8	11,3	ACP	[55]
			1mka	11	2.9	10.9	DAC	[14]
HYDRATASE								
Mfe–2 hydratase 2	*Candida tropicalis*	β–Oxidation	1pn4	7	3.5	10.6	HDC	[16]

*This study. The percent identities, rmsd values, and Z-scores have been calculated with DALI [56] using the $DH^{PpsC}/C^{12:1}$-CoA complex as a reference (PDB code 5nji). DALI gives higher Z-scores for highly similar structures. [a,b] Abbreviations used: rmsd, root mean square deviation for superposition of Cα atoms; S21, 4,4,4-trifluoro-1-(4-nitrophenyl)butane-1,3-dione; K91, 4-chloro-2-[(5-chloroquinolin-8-yl)oxy]phenol; KM0, 8-(benzyloxy)-5-chloroquinoline; KM1, 5-chloro-8-[(3-chlorobenzyl)oxy] quinolone; ACP, acyl carrier protein; KBP, S-[2-(acetylamino)ethyl] (3R)-3-hydroxydecanethioate; DAC, 2-decenoyl N-acetyl cysteamine; HDC, 3 R-hydroxydecanoyl-coenzyme A. Compounds S21, K91, KM0, and KM1 are a series of competitive inhibitors designed for FabZ.

fragments A (884–1231) and B (921–1222) could be purified and crystallized (Fig. 2b). Small crystals of fragment A never diffracted beyond 8–10 Å despite the implementation of extensive optimization screens. On the other hand, bipyramidal crystals of fragment B could grow overnight and diffracted to a maximum resolution of 2.7 Å (Table S1). DH_{PpsC} belongs to the hot dog fold enzyme superfamily whose first identified member was the *Escherichia coli* type II dehydratase–isomerase FabA [14]. The hot dog fold is characterized by a curved antiparallel β-sheet resembling a bun that wraps around an α-helical sausage (Fig. 3a). Unlike FabA, which functions as a homodimer with a single hot dog from each monomer associating to create an extended β-sheet [14], the PpsC DH double hot dog occurs in a single polypeptide chain. Landmarks include a 13-stranded β-sheet (topology β7-β8-β9-β11-β12-β13-β10-β3-β6-β5-β4-β2-β1), hot dog helices αHD1 and αHD2, a cap motif comprising helices η1, η2, and η3, a three-stranded β-sheet (β2a, β2b, and β7a), and helix αC1 (Fig. 3a). Residues in loop regions (amino acids 1049–1063 and 1144–1158) that connect strands β6 to β7 and helix η4 to strand β10 are missing from the refined X-ray structure, as a result of poorly defined electron density. The search for structural homologs in the

Fig. 3. Three-dimensional structure and structural alignment of DH domains from modular PKSs. (a) Ribbon representation of the wild-type DH_{PpsC} in its apo form with rainbow color coding from blue (N terminus) to red (C terminus). N- and C-terminal residues from missing regions (β6//β7 and η4//β10) are also indicated. (b) Sequence–based sequence alignment of DH_{PpsC} with homologs from modular PKSs for which the 3D structure are known (group 1). Group 2 includes DH domains from PKSs involved in the DIM biosynthesis pathway. The recently determined structure of the Mas-like Pks5 is also included. β strands and α helices of DH_{PpsC} (top) and DH_{Pks5} (bottom) are shown as arrows and coils, respectively. Sequence homologies are highlighted in red; sequence identities are shown as white letters on a red background. Strictly conserved active site residues involved in the catalytic mechanism are shown with red triangles. Green stars indicate the polar residues lining the active site cavity shown in Fig. S4. CurF, CurH, CurJ, and CurK are part of the curacin A biosynthetic pathway; RifE, rifamycin polyketide synthase; DEBS, 6-deoxyerythronolide B synthase; Mas, mycocerosic acid synthase. PDB references: CurF, 3kg6; CurH, 3kg7; CurJ, 3kg8; CurK, 3kg9; RifE, 4ln9; DEBS, 3el6; Pks5, 5bp2. UnitprotKB accession numbers: PpsC, P96202; CurF, Q6DNE7; CurH, Q6DNE5; CurJ, Q6DNE3; CurK, Q6DNE2; RifE, O54593; DEBS, Q03132; PpsD, P9WQE3; Mas, I6Y231; Pks5, I7FI52.

Protein Data Bank[‡] highlighted some strong similarities with double hot dog DH domains from the Mas-like PKS [22], the curacin A biosynthetic pathway [23], the rifamycin [24] and erythromycin [25] PKSs (Fig. 3b and Table 1). Strong similarities also exist with in the DH domain from mammalian type I FASs and, to a lesser extent, with in bacterial type II FASs (Table 1).

Long acyl chains are preferred for DH activity

The natural substrate of DH_{PpsC} is a C_{28}-C_{30} long acyl chain covalently attached to an ACP domain via a 4′-phosphopantetheinyl (P-pant) arm. After condensation with malonyl-CoA by the KS domain and subsequent reduction of the β-keto function by the KR domain, the growing acyl chain carries three chiral centers with hydroxyl groups at positions C3, C5, and C7 (Fig. 4a, compound 4). The *5S,7R* configurations have been deduced from the phthiocerol structure [26] and in the context of an esterification with two mycocerosic acids [27,28]. The *R* configuration of C3 could be assigned from the presence of conserved amino acid residues in the upstream KR domain [29]. The LDD motif is present in the KR domain from PpsC (starting at position 1892). The first aspartate is not strictly conserved in B-type KRs. This is also true for PpsC where it is replaced by an alanine (A1893). In addition to the LDD motif, the presence of a proline at position 1944 and the absence of tryptophan at position 1941 correlate well with the formation of B-type KR with *3R* stereochemistry [29]. The resulting KR product with *3R,5S,7R* configuration (Fig. 4a, compound 4) is predicted to be dehydrated to a *trans*-α,β-unsaturated bond [30] (Fig. 4a, compound 5). Although the presence of the downstream ER domain must favor the dehydration reaction in the context of an entire PpsC enzyme, it has been shown that isolated DH domains can catalyze the reverse reaction *in vitro* [17–20]. Therefore, we tested the enzymatic activity of DH_{PpsC} in the presence of soluble *trans*-2-unsaturated CoA derivatives with acyl chain length C4, C10, and C12 (Fig. 4a, compounds 1, 2, and 3). A specific activity of 58.1 ± 1.8 nmol.min^{-1}.mg^{-1} was measured for the DH domain in the presence of $C_{12:1}$-CoA (Fig. 4b). In identical experimental conditions, the specific activity was decreased 12-fold when the acyl chain was shortened by two carbons, while no activity was detected in the presence of $C_{4:1}$-CoA. The steady-state kinetic constants could only be determined for $C_{12:1}$-CoA with K_m and V_{max} values of 51 ± 10 μM and 47 ± 4 nM/s, respectively (Fig. S1). These results strongly suggest that the length of the acyl chain is essential for catalysis, as the absence of hydroxyl groups at positions 5 and 7 does not abolish enzyme activity. The formed product was analyzed by MALDI-TOF mass spectrometry. Analysis of the spectra

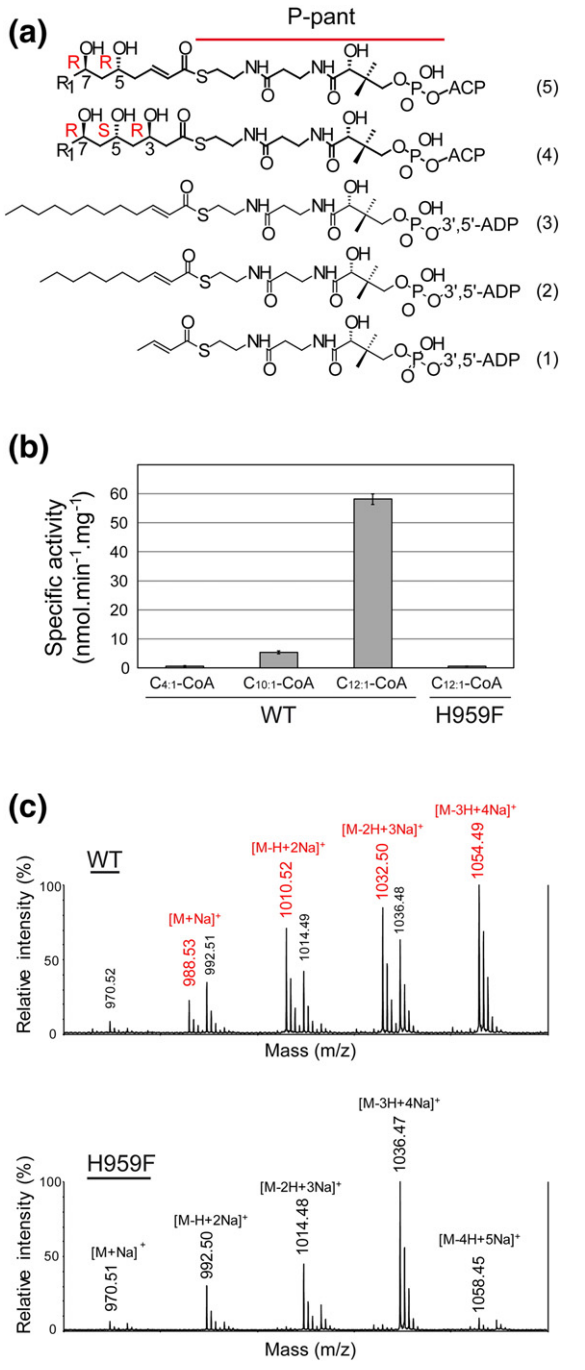


Fig. 4. Activity assays. Enzyme assays were performed on the wild-type DH_{PpsC} and H959F mutant using (1) *trans*-but-2-enoyl-CoA ($C_{4:1}$-CoA), (2) *trans*-dec-2-enoyl-CoA ($C_{10:1}$-CoA), and (3) *trans*-dodec-2-enoyl-CoA ($C_{12:1}$-CoA). The natural substrate (5) and dehydration product (4) are also depicted. $R_1 = CH_3$-$(CH_2)_{20–22}$. (b) Disappearance of the C=C double bond at 263 nm was detected in the presence of $C_{10:1}$-CoA and $C_{12:1}$-CoA. (c) The hydrated product was confirmed from MALDI-TOF MS experiments with the wild-type enzyme (top). No product was detected for the H959F variant (bottom). The substrate ion peaks are labeled in black and the hydration product ion peaks are labeled in red.

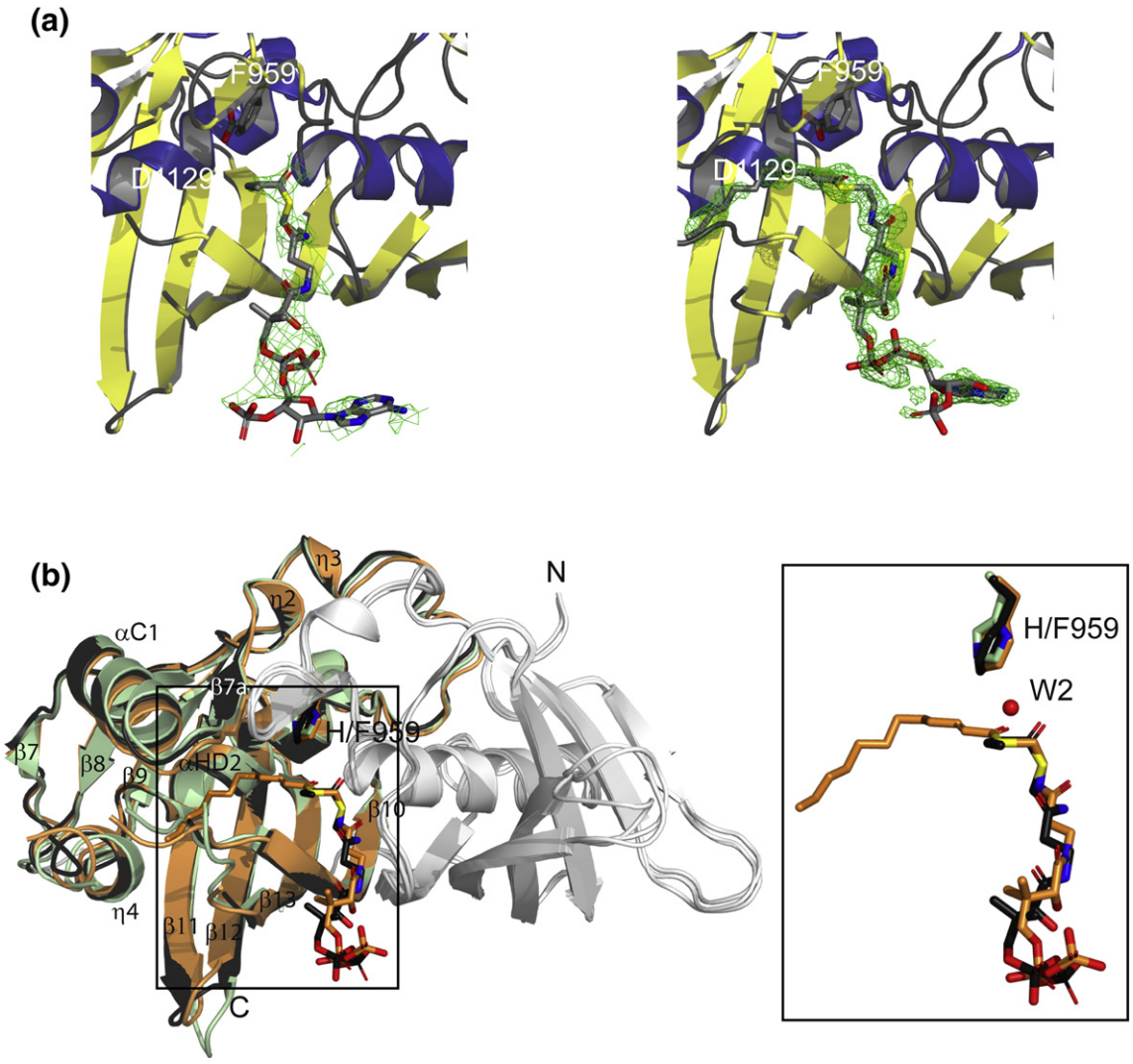


Fig. 5. Structural alignment of the DH_{PpsC} structures. (a) The active site histidine (H959) has been mutated into a phenylalanine (F959). Simulated-annealing 2Fo-Fc composite omit electron density maps contoured at 1.0σ around the CoA derivatives are shown in green for DH_{PpsC} in complex with $C_{4:1}$-CoA (left) and $C_{12:1}$-CoA (right). For clarity, only one alternate conformation is shown for the 3′,5′-ADP moiety of $C_{12:1}$-CoA outside the cavity. (b) The C-terminal hot dog domain of wild-type apo-DH_{PpsC} (green) is superimposed onto the $DH_{PpsC}/C_{4:1}$-CoA (black) and $DH_{PpsC}/C_{12:1}$-CoA (orange) structures. The same color code is used for the carbon atoms of CoA derivatives and the active site residues H959 and F959. The N-terminal hotdog is in white for all three structures. Secondary structure elements delineating the tunnel cavity are indicated. Closeup view of the superimposition between the $C_{4:1}$-CoA (black) and $C_{12:1}$-CoA (orange).

indicates that 3-hydroxy-dodecanoyl-CoA, the product of the reverse hydration reaction, was observed in the presence of the wild-type enzyme (Fig. 4c and S2). On the other hand, no product was detected when the catalytic histidine residue (H959) is mutated into a phenylalanine (Fig. 4c), as previously shown for DH domains of other modular PKSs [31–33].

Structural determinants for substrate specificity

In an attempt to elucidate how the length of the acyl chain can affect the activity of the enzyme, we focused our efforts on determining the X-ray structures of DH_{PpsC} in complex with $C_{4:1}$-CoA and $C_{12:1}$-CoA. Neither co-crystallization nor soaking experiments of wild-type DH crystals could give rise to the formation of a stable complex, a non-surprising result considering the transient nature of interactions that govern fatty acid and polyketide biosynthesis. To circumvent this obstacle, we used the DH_{PpsC} H959F inactive mutant that could sequester $C_{4:1}$-CoA and $C_{12:1}$-CoA. Using this strategy, both ligands could be successfully trapped into the active site with markedly improved diffraction limit to 1.5 Å for the $DH_{PpsC}/C_{12:1}$-CoA

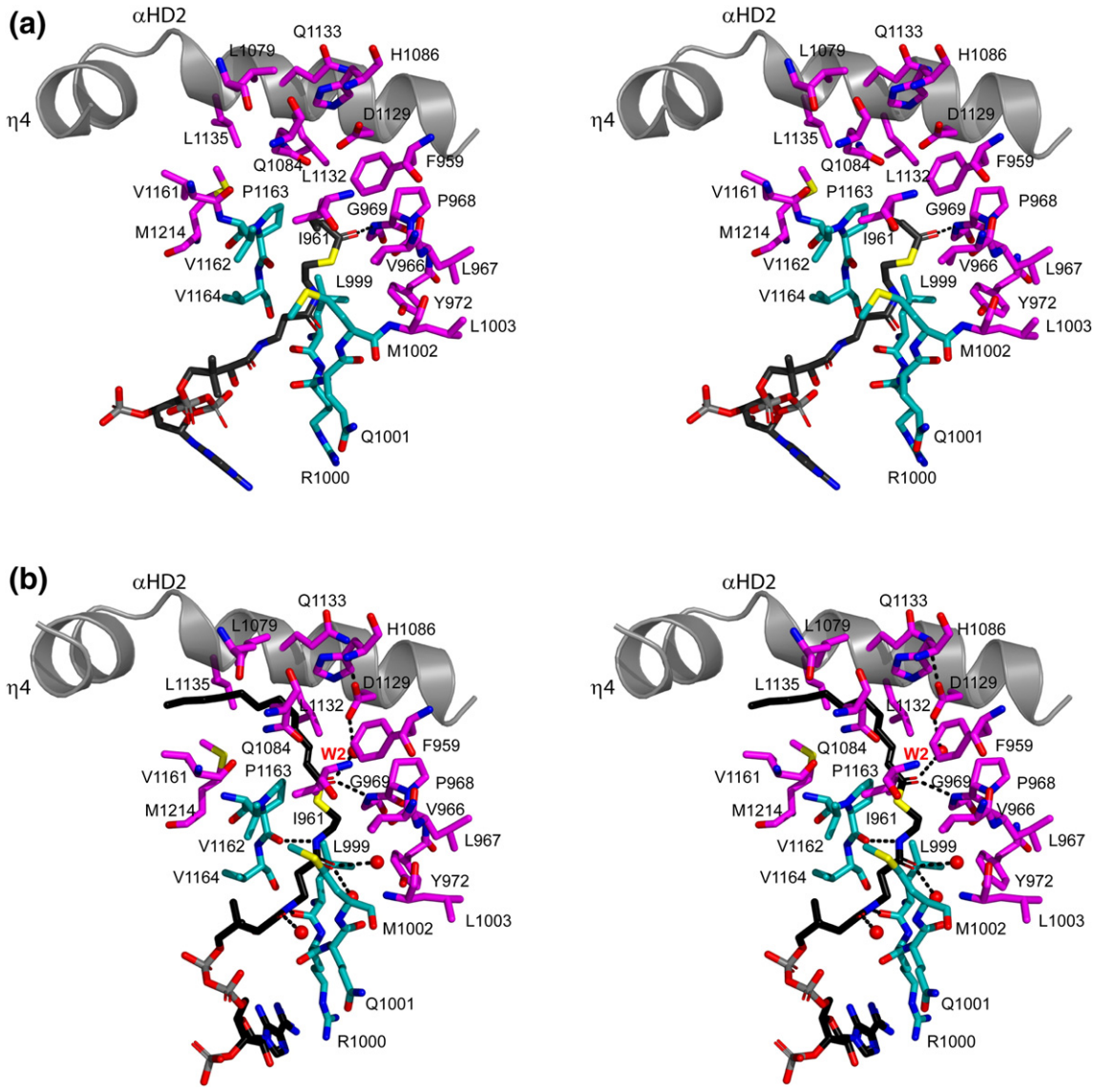


Fig. 6. Stereo view of the substrate-binding cavity. Residues lining the tunnel are shown as sticks for the DH_{PpsC} H959F variant in complex with (a) $C_{4:1}$-CoA and (b) $C_{12:1}$-CoA. The cavity entrance is delineated by residues 999–1002 and 1162–1164 with carbon atoms colored in cyan. The carbon atoms of residues lining the thioester group and acyl chain are colored in magenta. The carbon atoms of the ligand are in black. The presumed catalytic water molecule (W2), as found in the complex with $C_{12:1}$-CoA, is indicated.

complex (Table S1). Calculation of simulated annealing omit maps confirmed the presence of the acyl chain moieties inside a ~20 Å long tunnel for both substrates, with strong electron density pointing outside the tunnel for the P-pant arm carrying the acyl chain (Fig. 5a). The 3′,5′-ADP moiety of the coenzyme A has much less defined electron density. This might be justified by the absence of stabilizing interactions outside the cavity and/or with neighboring molecules in the crystal packing.

The tunnel entrance at the junction between the two hot dogs is delineated by two stretches of residues between strands β3 and β4 (positions 999 to 1002) and within strand β10 (positions 1162 to 1164) (Fig. 6 and S3). In the $DH_{PpsC}/C_{12:1}$-CoA complex, the backbone Cα of V1162 has shifted ~1 Å closer to the space occupied by $C_{4:1}$-CoA in the $DH_{PpsC}/C_{4:1}$-CoA complex and ~1.5 Å with respect to the apo enzyme. In addition, the side chain of M1002 has moved toward V1162, thus causing a narrowing of the entrance cavity in the $DH_{PpsC}/C_{12:1}$-CoA complex, nearly circular in shape. From this result, we predict that such structural movements, potentially exacerbated by the presence of a covalently linked ACP, may help

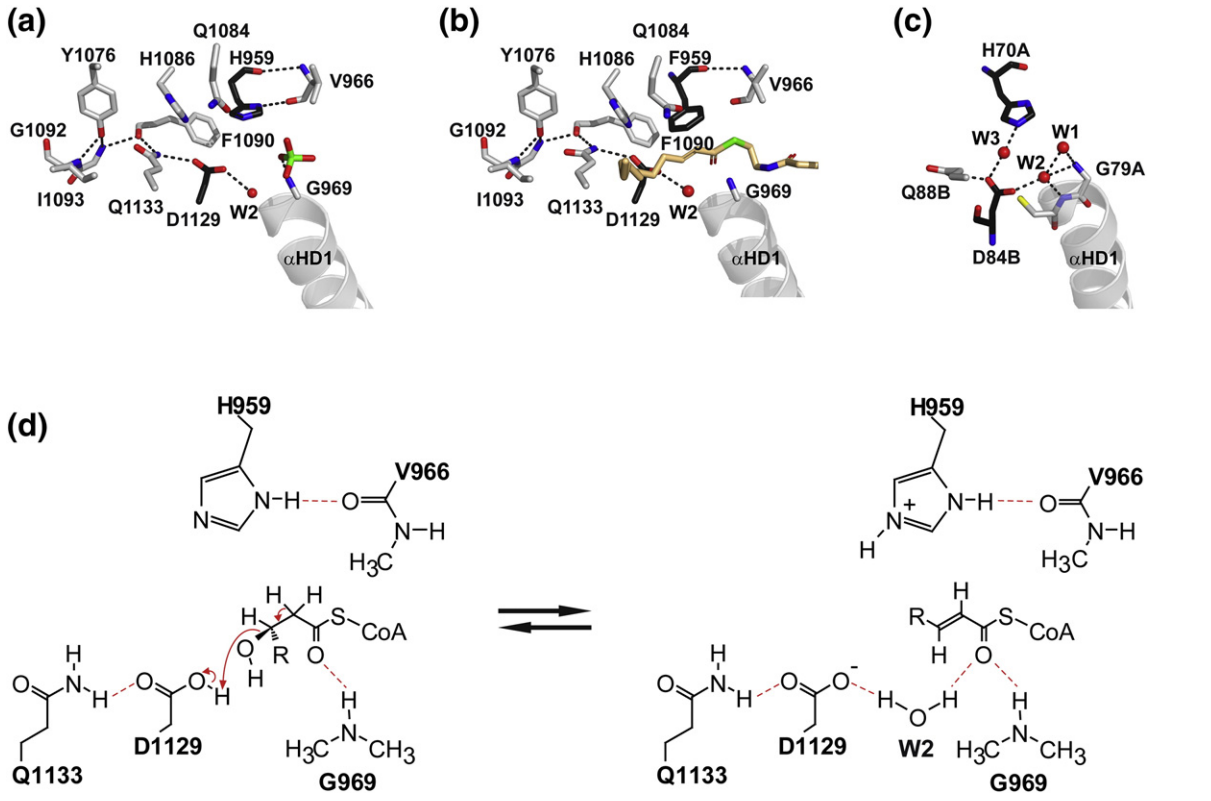


Fig. 7. Proposed mechanism of the dehydration reaction. Ball-and-stick representation of residues involved in the dehydration reaction as seen in the structure of (a) wild-type apo DH_{PpsC}, (b) DH_{PpsC} H959F mutant in complex with $C_{12:1}$-CoA, and (c) the dimeric *Escherichia coli* FabA (PDB code 1mkb). Cα atoms of the catalytic histidine and aspartate residues are in black. (d) Proposed dehydratase mechanism based on the X-ray structure of DH_{PpsC} in complex with an α,β-double bond substrate in a *trans* configuration. R = CH_3-$(CH_2)_8$.

stabilize the substrate within the tunnel cavity. Despite the presence of multiple polar interactions between the P-pant arm of CoA derivatives and DH_{PpsC}, superimposition of the structures of the two complexes reveals a shift in the position of the P-pant moiety (Fig. 5b) accompanied by significant changes in interaction patterns (Fig. 6a and b). This observation supports the idea that long acyl chains are key players in positioning correctly the P-pant and catalytic center during the (de)hydration reaction.

Among the six PKSs involved in DIM synthesis, only PpsC, PpsD, and Mas harbor a DH domain. Residues lining the acyl chain cavity in DH_{PpsC} are mostly hydrophobic, with the exception of Q1084, H1086, and Q1133, whose polar side chains might be involved in hydrogen bond interactions with hydroxyl groups at positions C5 and C7 in the $C_{12:1}$-CoA substrate (Fig. S4a). The side chains of Q1084 and Q1133 are hydrogen bonded to the catalytic residues H959 and D1129, respectively (Fig. 7a). All three residues are strictly conserved in PpsD, whereas Q1084 and H1086 have been replaced with hydrophobic residues V1048 and L1050 in the Mas sequence, respectively (Fig. 3b). This observation is consistent with a potential role of Q1084 and H1086 in stabilizing hydroxyl groups that are absent in Mas substrates. The three-dimensional structure of the DH domain from Mas is unknown, but the closely related structure of the DH domain from *Mycobacterium smegmatis* Pks5 (DH_{Pks5}), which shares 54% sequence identity, has recently been published [22]. Pks5 synthesizes lipid components of lipooligosaccharides characterized by the presence of methyl-branched fatty acids similar to Mas products and is referred to as a Mas-like PKS [34]. Both PpsC and Pks5 DH X-ray structures can be superimposed with a 1.9 Å rmsd between the aligned Cα atoms. Similarity in the shape of the cavity also characterizes the two structures. As shown in Fig. S4b, the presence of H1055 in DH_{Pks5}, equivalent to H1086 of PpsC, in the inner part of the cavity is somewhat puzzling, given the absence of polar groups on the acyl chain of the natural product of Pks5 [34]. In an attempt to evaluate the potential role of hydroxyls in stabilizing the natural substrate within the DH_{PpsC} cavity, molecular dynamics (MD) simulations were

performed on complexes with the $C_{12:1}$-CoA compound and with di-hydroxylated analogs bearing hydroxyl groups at positions 5 and 7 in an *S* or *R* configuration (Table S2). Binding energies, estimated using the MM/GBSA approximation, were found to be similar for all evaluated compounds, with differences smaller than 6.7 kJ.mol^{-1}. Additionally, the persistence of hydrogen bonds involving the hydroxyl groups in the course of the 10-ns simulation was investigated. Such hydrogen bonds were found to exist only transiently for the different di-hydroxylated $C_{12:1}$-CoA, with a fraction of existence smaller than 20% in most cases.

Active site residues and crystallographic water molecules

To date, none of the accessible structures of DHs from modular PKSs has been determined with a metabolizable ligand in its active site (Table 1). As a consequence, the proposed mechanism of dehydration described for the DH domain from DEBS module 4 [35], the curacin A [23], and rifamycin [24] biosynthetic pathways relies exclusively on three-dimensional structures of the apo form. Only complexes involving the stand-alone type II FAS DH FabA [14,15] and the β-oxidation hydratase Mfe2p-H2 [16] have been determined so far (Table 1). The suicide inhibitor 2-decenoyl-*N*-acetylcysteamine in FabA offers a distorted vision of the catalytic site [14] that makes difficult the interpretation of a dehydration mechanism based on the three-dimensional structure. The scenario of the catalytic histidine abstracting the α-proton and the aspartic acid donating a proton to the β-hydroxyl group was proposed and generalized to all the modular PKSs and type II FASs DH described in Table 1. Here, the structure of the DH_{PpsC}/$C_{12:1}$-CoA complex provides a snapshot of an unsaturated product after the dehydration reaction. At least three crystallographic water molecules (W1, W2, and W3) have been assigned a role in the mechanism of hydration and dehydration [14,36] (Fig. 7c). W1 and W3 are absent in all our DH_{PpsC} structures. In the wild-type apo-form, W1 is replaced by a phosphate group coming from the crystallization buffer, which is hydrogen bonded to G969 (Fig. 7a). W2 is the only resolved water molecule present in the high-resolution structure of the DH_{PpsC}/$C_{12:1}$-CoA complex, within hydrogen bond distance to the carboxylate oxygen of the catalytic residue D1129 and the thioester group of the substrate (Fig. 7b). W2 is located in the vicinity of the C3 atom of $C_{12:1}$-CoA (3.7 Å), where the hydroxylation reaction takes place. This suggests that W2 is the released water during the dehydration reaction, as originally proposed by Leesong and colleagues [14].

Furthermore, in the apo-DH_{PpsC} structure, residues V966 and Q1084 make hydrogen bonds with the catalytic residue H959, which maintains its side chain in a position that allows the hydration reaction to take place (Fig. 7b). Y1076 is part of a polar interaction network that probably increases the pKa value of the carboxylic acid group of D1129, thus allowing it to be protonated and to catalyze dehydration at physiological pH (Fig. 7b and d).

Conclusion

Dehydratases from type I FASs and PKSs are key players in the production of a wide range of pharmacologically active natural products with strong emphasis on human medicine. To date, structural information on dehydration reactions has been relying on modeling studies with limited knowledge of determinants for substrate specificity. In our studies, the catalytic activity of the DH domain from the *M. tuberculosis* PKS PpsC has been assayed by measuring the reverse hydration reaction in the presence of unsaturated CoA derivatives with different acyl chain length (C4, C10, and C12). We showed that substrates with long acyl chains are essential for the DH to carry out its function. Structures of the DH domain in complex with two unsaturated CoA derivatives, that is, $C_{4:1}$-CoA and $C_{12:1}$-CoA, constitute an important piece in the intricate puzzle of the DH reaction. Identification of key residues essential for interactions with the natural acyl phosphopantetheinyl moiety normally tethered to the ACP should enable the engineering of modified PKSs that can process non-natural substrates with improved or novel pharmacological properties.

Materials and Methods

PpsC fragment library, ORF selection, and cloning

The open reading frame (ORF) for the *M. tuberculosis ppsC* gene was amplified by PCR using forward 5′-GATATACATATGACCGCAGCGACACCAGATCG-3′ and reverse 5′-AATTCACTAGTTGACTCGCCTCGCGTCGCAGC-3′ primers. The underlined bases represent the NdeI and SpeI restriction sites, respectively. A large-size DNA library (850–1650 bp) was created from cleaned PCR product using a HydroShear device from Genomics Solutions (Ann Arbor, MI). Extremities of the fragments were polished using Vent polymerase (New England Biolabs, Beverly, MA) at 72 °C for 20 min. A slab of gel containing DNA fragments (from 850 to 1650 bp) was excised and recovered with a QIAquick® Gel Extraction Kit (Qiagen Inc. USA, Valencia, CA, USA). Library screening for ORFs into a permissive site of the *E. coli* dihydrofolate reductase (DHFR) was conducted as previously described [21]. Recovered DHFR insertion library was diluted for plasmid

preparation. Gel extracted and cleaned inserts from NdeI/SpeI restriction digests were ligated into the pTET ColE1 GFP11 vector and transformed into chemically competent BL21 (DE3), pET GFP 1–10 cells. *In vivo* solubility screenings were performed as previously described [37].

Protein production and purification of wild-type DH domain

Frozen cells from constructs A and B were used to grow 3-mL LB-kanamycin (35 μg/mL) cultures overnight at 32 °C prior to inoculation in baffled flasks containing 500 mL of the same media. Cells were allowed to grow for approximately 2 h at 37 °C before temperature was dropped from 37 °C to 30 °C. When OD_{600} reached 0.5–0.7, cells were induced with IPTG at a final concentration of 0.5 mM and grown for an additional 4 h prior to harvesting by centrifugation at 4000*g* for 20 min. Selenomethionine-labeled fragment B was produced following the protocol described by Guerrero and colleagues [38]. Cell pellets from both wild-type and selenomethionylated protein fragments were resuspended in 25 mL buffer A [100 mM Tris (pH 8.1) and 150 mM NaCl] and lysed by sonication on ice. We then added 3 mL of washed cobalt Talon Superflow resin (GE Healthcare 28-9575-02) to the supernatant, and the resulting mixture was shaken for a few minutes, allowing the 6His-tagged proteins to bind the cobalt resin. Two cycles of washing were then realized before the elution of non-specific contaminants with buffer A containing 10 mM imidazole and finally the elution of the bound protein with buffer A containing 250 mM imidazole. We dialyzed twice 6 mL eluted solution against 1 L of buffer B [25 mM Tris (pH 8.1) and 25 mM NaCl]. Protein purity was assessed on 10% SDS-PAGE along with PageRuler Prestained Protein Ladder (Thermo Scientific #26616) followed by Coomassie blue detection.

Crystallization and structure determination of wild-type DH domain

Highly purified proteins were used for crystallization using commercially available kits from Qiagen (Venlo, Netherlands) in Greiner 3 round plates (Stonehouse, UK) and the hanging drop vapor diffusion technique. The plates were filled using a Nanodrop ExtY crystallization robot (Innovadyne Technologies, Santa Rosa, CA, USA) prior to storage and imaging using a Rock Imager 1000 (Formulatrix, Bedford, MA, USA). Optimization of crystallization conditions was realized manually in Linbro plates (Hampton Research, Aliso Viejo, USA) containing 500 μL reservoir solution. For crystallization of the DH domain, 1 μL of protein (30 mg/mL) was mixed with 1 μL of 1.8 M Na/K phosphate (pH 5.8–6.0). Crystals of the DH fragment B belong to space group $P4_32_12$ with cell parameters $a = b = 89.5$ Å, $c = 159.3$ Å. Crystals were soaked in paratone for 5 s prior to flash freezing in liquid nitrogen. Datasets for the wild-type DH domain were collected at the ESRF beamlines ID29 and ID14 to a maximum resolution of 2.7 Å (Table S1). Peak and inflection datasets for selenomethionylated crystals of the DH domain were collected at the selenium absorption edge (ID29, ESRF). The DH structure was solved by the multiwavelength anomalous dispersion method using anomalous scattering from the six selenium-substituted methionine residues. Datasets were indexed using MOSFLM [39] and scaled with SCALA [40]. Phase calculation and density modification were realized with the SHELX software suite [41]. Structures were then built and refined by iterative cycles of manual model building in COOT [42] and refinement using REFMAC5 [43] included in the CCP4 software suite [44].

Synthesis of *trans*-dodec-2-enoyl-CoA ($C_{12:1}$-CoA)

Synthesis of methyl trans-dodec-2-enoate

Sodium hydroxide (482 mg, 12.0 mmol, 0.5 eq) was added to a solution of (methoxycarbonylmethyl) triphenyl phosphonium bromide (10 g, 24.1 mmol, 1.0 eq) in water (160 mL). The reaction mixture was stirred for 15 min. Then, the ylide was extracted with dichloromethane (3 × 50 mL). The organic phases were combined, dried over sodium sulfate, and filtrated. This solution (~150 mL) was added directly to a solution of decanal (3.17 mL, 16.8 mmol, 0.7 eq) in dichloromethane (50 mL). The reaction mixture was stirred for 2 h at room temperature. Then, the mixture was washed with water and brine, dried over magnesium sulfate, and concentrated under reduced pressure. To remove phosphine oxide, we added tetrahydrofuran to the crude mixture and filtrated it. The filtrate was concentrated to obtain a white solid (4.6 g). The product was used in the next step without further purification.

Synthesis of trans-dodec-2-enoic acid

A 0.3 M aqueous solution of potassium hydroxide (67.1 mmol, 4.0 eq) was added to a solution of ester (3.56 g, 16.8 mmol, 1.0 eq) in $EtOH/H_2O$. The mixture was refluxed overnight. After cooling, the aqueous phase was washed with dichloromethane and was acidified with HCl (6 N) to pH ~1. The carboxylic acid was extracted with dichloromethane (3 × 50 mL), and the organic phases were combined, dried over magnesium sulfate, filtrated, and concentrated under reduced pressure to afford an oily product (2.66 g). The residue was purified by flash chromatography (gradient 100% CH_2Cl_2 to 95/5 CH_2Cl_2/THF) for 20 min to afford a colorless oil.

TLC Rf 0.35 (dichloromethane) ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 7.12 (td, J = 7.2 Hz, 15.6 Hz, 1H); 5.85 (td, J = 1.5 Hz, 15.6 Hz, 1H); 3.73–3.83 (m, 1H); 2.26 (qd, J = 1.5 Hz, 7.5 Hz, 2H); 1.84–1.92 (m, 1H); 1.42–1.53 (m, 2H); 1.22–1.42 (m, 10H); 0.92 (t, J = 6.9 Hz, 3H).

Synthesis of $C_{12:1}$-CoA

$C_{12:1}$-CoA was synthesized using a reported method [45]. Briefly, in a dry round-bottom flask, the acid (55 mg, 0.27 mmol, 1.0 eq) was added in dry diethyl ether (5 mL) and cooled to 4 °C. Then, ethyl chloroformate (37 mg, 33 µL, 1.25 eq) and dry triethylamine (34 mg, 46 mL, 0.34 mmol, 1.2 eq) were added, and the reaction mixture was stirred overnight at room temperature. Water was added to the solution, and the mixed anhydride was extracted with diethyl ether (3 × 20 mL), dried over magnesium sulfate, filtered, and concentrated under reduced pressure to give a light-brown oil (60 mg). In a hemolysis tube, coenzyme A trilithium salt (100 mg, 0.13 mmol, 0.5 eq) was solubilized in water (1.5 mL), and the solution was basified to pH >8.0; a light-pink color was observed. The coenzyme A salt solution was added dropwise to the mixed anhydride solubilized in THF (2–3 mL). The reaction mixture was stirred at room temperature and THF was removed under reduced pressure. Water was added and the crude mixture was washed with diethyl ether (3 × 10 mL) and freeze-dried to afford a flocculent product (~120 mg).

The purification of $C_{12:1}$-CoA was carried out with the equipment of the technical platform at the Institute of Chemistry of Toulouse (ICT–FR CNRS 2599). The purification of the product was performed by semi-preparative HPLC (Autopurification Waters with a Photodiode Array Detector) on a C18 column (X bridge C18 5 µm, 150 mm × 19 mm ID). We used 22.5 mM ammonium acetate (pH 6.7) and acetonitrile with a flow rate of 20.5 mL/min. The elution was followed by UV detection at 259 nm. The starting conditions of the linear gradient were 80% of ammonium acetate buffer/20% acetonitrile (vol/vol). The equilibration time was 0.5 min in these conditions. Then, the percentage of acetonitrile varied from 20% to 50% over 10 min and kept at 50% for 2 min. The return to initial conditions is done in 0.5 min. The total run time of the separation was 20 min. The product was dissolved in the buffer (36 mg/mL) and filtered with 0.2 µm regenerated cellulose membrane. The injection volume was 1 mL. The product was eluted with a retention time of 8.97 min. The collected fractions were pooled and analyzed by analytical UPLC (UPLC Acquity Waters, column BEH C18 1.7 µm, 2.1 mm × 50 mm). The fractions were freeze-dried and furnished a white solid (~30 mg). The NMR data are in agreement with those reported in the literature [46].

Synthesis of trans-dec-2-enoyl-coA ($C_{10:1}$-CoA)

$C_{10:1}$-CoA was synthesized and purified as previously described [45].

Construction and structure determination of the H959F DH mutant

The H959F DH mutant was amplified by inverse PCR using complementary primers bearing two point mutations shifting the catalytic histidine codon CAC to a phenylalanine codon TTC. Forward (5′-GCTGTGG CTCGCCGATTTCGTCATCGACGATC-3′) and reverse (5′-GATCGTCGATGACGAAATCGGCGAGC-CACAGC-3′) primers were from Sigma-Aldrich. Amplified DNA was digested using DpnI and transformed into *E. coli* BL21 (DE3). Transformants were picked and the mutation was verified by sequencing the plasmid (GATC Biotech, Konstanz, Germany). The resulting H959F variant could be produced, purified, and crystallized in the exact same conditions as the wild-type DH fragment. $C_{4:1}$-CoA (Sigma-Aldrich) was solubilized in water and added to the drop containing DH crystals to a final concentration of 1 mM for 1 h. The DH variant was also purified and concentrated at 367 µM (11.7 mg/mL) in the presence of 50 mM Hepes (pH 7.5) and 50 mM NaCl prior to mixing with $C_{12:1}$-CoA in a 1:2 molar ratio. The same conditions of crystallization were used [1.8 M Na/K phosphate (pH 6)] that gave rod-shaped crystals with improved diffraction limit to 1.5 Å on beamline XALOC at ALBA (Barcelona, Spain). Crystals of the DH mutant in complex with the $C_{4:1}$-CoA belong to space group $P4_32_12$ with cell parameters $a = b$ = 83.9 Å, c = 166.8 Å. Crystals of the DH/$C_{12:1}$-CoA complex belong to space group $C121$ with cell parameters a = 117.4 Å, b = 70.8 Å, c = 42.6 Å α = 90°, β = 104.9°, γ = 90°. Data were processed using the XDS program package [47]. Structures were solved by molecular replacement using the wild-type DH structure as template in PHASER [48]. CoA derivatives and water molecules (3 sigma level in the Fo-Fc map) were then added, and the corresponding structures were refined by iterative cycles of manual model building in COOT [42] and PHENIX [49]. The DH_{PpsC}/$C_{4:1}$-CoA structure was then further refined using the high-resolution structure of the DH domain from DH_{PpsC}/$C_{12:1}$-CoA as a reference model in PHENIX [49]. Water molecules were built into the electron density using default parameters in PHENIX [49] and were inspected manually using COOT [42]. Simulated-annealing 2Fo-Fc composite omit electron density maps contoured at 1.0σ level were calculated with PHENIX [49].

Enzyme assays and analysis of hydrated products

Hydratase activity was monitored at 263 nm using a thermostated Uvikon 923 spectrophotometer

(Kontron Instruments) in the presence of *trans*-2-enoyl-CoA (ΔA of 0.67 for a variation of concentration of 100 μM). Kinetic assays were performed in a quartz cuvette for 1.5 min at room temperature, in 100 mM sodium phosphate buffer (pH 7.0), and in the presence of 10 μM $C_{12:1}$-CoA, $C_{10:1}$-CoA or $C_{4:1}$-CoA. After equilibration of the baseline, reactions were started by adding purified 0.250–1.275 μM DH domain (fragment B) WT or H959F, as specified. K_m and V_{max} were measured in the presence of variable concentrations (0–200 μM) of $C_{12:1}$-CoA, and data were fitted to the Michaelis–Menten equation by least-squares fits to a hyperbola using the program GraphPad Prism version 5.04. For MALDI-TOF MS analyses, the reaction media were first diluted 10-fold in water. Then, 1 μL samples were spotted onto the target plate, mixed with 1 μL of matrix [10 mg/mL of 2,5-dihydroxybenzoic acid in water:acetonitrile, 8:2 (vol/vol)], and allowed to crystallize at room temperature. MS analyses were performed using a 5800 MALDI-TOF/TOF Analyzer (Applied Biosystems/ABsciex). Ionization was achieved by irradiation with a Nd:YAG laser (349 nm). Mass spectra were acquired in reflectron-positive mode and were the sum of 2500 laser shots. Lipid standards were used to calibrate the mass spectrometer.

MD calculations

Coordinates for the four diastereoisomers of hydroxylated $C_{12:1}$-CoA at positions 5 and 7 were generated using VIDA 4.3.0 (OpenEye Scientific, Santa Fe, USA). The tleap module of AMBER 12 [50] was used to generate a periodic octahedral box containing the complex, 11,177 water molecules represented with the TIP3P model, and 23 sodium cations. Energy minimization and MD simulation were performed with the parallel version of the PMEMD module of AMBER12. After an initial energy minimization with progressively reduced restraints on the position of protein atoms, 40-ps MD simulations were run with temperature varying linearly from 0 to 300 K at constant volume, followed by a 400-ps MD at 300 K and a constant pressure of 1 bar, with the atomic coordinates saved every 2 ps. A production MD simulation of 10 ns was then performed at a constant pressure of 1 bar, with atomic coordinates saved every 10 ps. The MMPBSA.py procedure [51] was used to evaluate binding energies, without considering the entropic term since all evaluated compounds are highly similar (Table S2).

Accession codes

The coordinates for the crystal structures of WT and H959F mutant of DH_{PpsC} and of the DH_{PpsC}/$C_{4:1}$-CoA and DH_{PpsC}/$C_{12:1}$-CoA complexes have been deposited in the Protein Data Bank with PDB ID codes 4ooc, 5l84, 5i0k, and 5nji, respectively.

Acknowledgments

This work was supported by the Agence Nationale de la Recherche (XPKS-MYCO, grant 09-BLAN-0298-01). The equipment used for crystallization experiments and X-ray crystallography are part of the Integrated Screening Platform of Toulouse (PICT, IBiSA). We thank the staff of synchrotron beamlines ID14 and ID29 at the European Synchrotron Radiation Facility (Grenoble, France). Experiments were also performed on the XALOC beamline at the ALBA Synchrotron (Barcelona, Spain) with the invaluable assistance of Jordi Juanhuix.

Appendix A. Supplementary Data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2017.03.026.

Received 10 January 2017;
Received in revised form 21 March 2017;
Accepted 27 March 2017
Available online 1 April 2017

Keywords:
Mycobacterium tuberculosis;
polyketide synthase;
dehydratase;
metabolizable ligand;
X-ray structure

Present address: A. Faille, Cambridge Institute for Medical Research, University of Cambridge, Cambridge CB2 0QH, UK.

†A.F. and S.G. contributed equally to this work.
‡http://www.rcsb.org.

Abbreviations used:
PKS, polyketide synthase; FAS, fatty acid synthase; DIM, phthiocerol dimycocerosate; PGL, phenolic glycolipid; Mas, mycocerosic acid synthase; ACP, acyl carrier protein; PCP, peptidyl carrier protein; AT, acyltransferase; KS, ketosynthase; KR, ketoreductase; DH, dehydratase; ER, enoylreductase; DH_{PpsC}, PpsC DH domain; $C_{4:1}$-CoA, *trans*-but-2-enoyl-CoA; $C_{10:1}$-CoA, *trans*-dec-2-enoyl-CoA; $C_{12:1}$-CoA, *trans*-dodec-2-enoyl-CoA; P-pant, 4′-phosphopantetheinyl; MD, molecular dynamics; ORF, open reading frame; DHFR, dihydrofolate reductase; MALDI-TOF, matrix-assisted laser desorption-ionization time-of-flight; THF, tetrahydrofuran.

References

[1] K.J. Weissman, P.F. Leadlay, Combinatorial biosynthesis of reduced polyketides, Nat. Rev. Microbiol. 3 (2005) 925–936, http://dx.doi.org/10.1038/nrmicro1287.

[2] L. Nguyen, S. Chinnapapagari, J. Charles, C.J. Thompson, FbpA-dependent biosynthesis of trehalose dimycolate is required for the intrinsic multidrug resistance, cell wall structure, and colonial morphology of *Mycobacterium smegmatis* FbpA-dependent biosynthesis of trehalose dimycolate is required for the intri, J. Bacteriol. 187 (2005) 6603–6611, http://dx.doi.org/10.1128/JB.187.19.6603.

[3] V. Jarlier, H. Nikaido, Permeability barrier to hydrophilic solutes in *Mycobacterium chelonei*, J. Bacteriol. 172 (1990) 1418–1423.

[4] C. Astarie-Dequeker, L. Le Guyader, W. Malaga, F.K. Seaphanh, C. Chalut, A. Lopez, C. Guilhot, Phthiocerol dimycocerosates of *M. tuberculosis* participate in macrophage invasion by inducing changes in the organization of plasma membrane lipids, PLoS Pathog. 5 (2009) e1000289, http://dx.doi.org/10.1371/journal.ppat.1000289.

[5] C.J. Cambier, K.K. Takaki, R.P. Larson, R.E. Hernandez, D.M. Tobin, K.B. Urdahl, C.L. Cosma, L. Ramakrishnan, Mycobacteria manipulate macrophage recruitment through coordinated use of membrane lipids, Nature 505 (2014) 218–222, http://dx.doi.org/10.1038/nature12799.

[6] L.R. Camacho, P. Constant, C. Raynaud, M.A. Lanéelle, J.A. Triccas, B. Gicquel, M. Daffé, C. Guilhot, Analysis of the phthiocerol dimycocerosate locus of *Mycobacterium tuberculosis*. Evidence that this lipid is involved in the cell wall permeability barrier, J. Biol. Chem. 276 (2001) 19,845–19,854, http://dx.doi.org/10.1074/jbc.M100662200.

[7] A.K. Azad, T.D. Sirakova, D. Norvin, P.E. Kolattukudy, Gene knockout reveals a novel gene cluster for the synthesis of a class of cell wall lipids unique to pathogenic mycobacteria, J. Biol. Chem. 272 (1997) 16,741–16,745, http://dx.doi.org/10.1074/jbc.272.27.16741.

[8] O.A. Trivedi, P. Arora, V. Sridharan, R. Tickoo, D. Mohanty, R.S. Gokhale, Enzymic activation and transfer of fatty acids as acyl-adenylates in mycobacteria, Nature 428 (2004) 441–445, http://dx.doi.org/10.1038/nature02406.1.

[9] O.A. Trivedi, P. Arora, A. Vats, M.Z. Ansari, R. Tickoo, V. Sridharan, D. Mohanty, R.S. Gokhale, Dissecting the mechanism and assembly of a complex virulence mycobacterial lipid, Mol. Cell 17 (2005) 631–643, http://dx.doi.org/10.1016/j.molcel.2005.02.009.

[10] K.C. Onwueme, J.A. Ferreras, J. Buglino, C.D. Lima, L.E.N. Quadri, Mycobacterial polyketide-associated proteins are acyltransferases: proof of principle with *Mycobacterium tuberculosis* PapA5, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 4608–4613, http://dx.doi.org/10.1073/pnas.0306928101.

[11] K.C. Onwueme, C.J. Vos, J. Zurita, C.E. Soll, L.E.N. Quadri, Identification of phthiodiolone ketoreductase, an enzyme required for production of mycobacterial diacyl phthiocerol virulence factors, J. Bacteriol. 187 (2005) 4760–4766, http://dx.doi.org/10.1128/JB.187.14.4760-4766.2005.

[12] E. Pérez, P. Constant, F. Laval, A. Lemassu, M.A. Lanéelle, M. Daffé, C. Guilhot, Molecular dissection of the role of two methyltransferases in the biosynthesis of phenolglycolipids and phthiocerol dimycoserosate in the *Mycobacterium tuberculosis* complex, J. Biol. Chem. 279 (2004) 42,584–42,592, http://dx.doi.org/10.1074/jbc.M406134200.

[13] R. Siméone, P. Constant, W. Malaga, C. Guilhot, M. Daffé, C. Chalut, Molecular dissection of the biosynthetic relationship between phthiocerol and phthiodiolone dimycocerosates and their critical role in the virulence and permeability of *Mycobacterium tuberculosis*, FEBS J. 274 (2007) 1957–1969, http://dx.doi.org/10.1111/j.1742-4658.2007.05740.x.

[14] M. Leesong, B.S. Henderson, J.R. Gillig, J.M. Schwab, J.L. Smith, Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: two catalytic activities in one active site, Structure 4 (1996) 253–264.

[15] L. Moynié, S.M. Leckie, S.A. McMahon, F.G. Duthie, A. Koehnke, J.W. Taylor, M.S. Alphey, R. Brenk, A.D. Smith, J.H. Naismith, Structural insights into the mechanism and inhibition of the β-hydroxydecanoyl-acyl carrier protein dehydratase from *Pseudomonas aeruginosa*, J. Mol. Biol. 425 (2013) 365–377, http://dx.doi.org/10.1016/j.jmb.2012.11.017.

[16] M.K. Koski, A.M. Haapalainen, J.K. Hiltunen, T. Glumoff, A two-domain structure of one subunit explains unique features of eukaryotic hydratase 2, J. Biol. Chem. 279 (2004) 24,666–24,672, http://dx.doi.org/10.1074/jbc.M400293200.

[17] A. Witkowski, A.K. Joshi, S. Smith, Characterization of the beta-carbon processing reactions of the mammalian cytosolic fatty acid synthase: role of the central core, Biochemistry 43 (2004) 10,458–10,466, http://dx.doi.org/10.1021/bi048988n.

[18] R.J. Heath, C.O. Rock, Enoyl-acyl carrier protein reductase (fabI) plays a determinant role in completing cycles of fatty acid elongation in *Escherichia coli*, J. Biol. Chem. 270 (1995) 26,538–26,542, http://dx.doi.org/10.1074/jbc.270.44.26538.

[19] E. Sacco, A.S. Covarrubias, H.M. O'Hare, P. Carroll, N. Eynard, T.A. Jones, T. Parish, M. Daffé, K. Bäckbro, A. Quémard, The missing piece of the type II fatty acid synthase system from *Mycobacterium tuberculosis*, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 14,628–14,633, http://dx.doi.org/10.1073/pnas.0704132104.

[20] O. Vergnolle, F. Hahn, A. Baerga-Ortiz, P.F. Leadlay, J.N. Andexer, Stereoselectivity of isolated dehydratase domains of the borrelidin polyketide synthase: implications for *cis* double bond formation, Chembiochem 12 (2011) 1011–1014, http://dx.doi.org/10.1002/cbic.201100011.

[21] J.-D. Pedelacq, H.B. Nguyen, S. Cabantous, B.L. Mark, P. Listwan, C. Bell, N. Friedland, M. Lockard, A. Faille, L. Mourey, T.C. Terwilliger, G.S. Waldo, Experimental mapping of soluble protein domains using a hierarchical approach, Nucleic Acids Res. 39 (2011) e125, http://dx.doi.org/10.1093/nar/gkr548.

[22] D.A. Herbst, R.P. Jakob, F. Zähringer, T. Maier, Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases, Nature 531 (2016) 533–537, http://dx.doi.org/10.1038/nature16993.

[23] D. Akey, J. Razelun, J. Tehranisa, Crystal structures of dehydratase domains from the curacin polyketide biosynthetic pathway, Structure 18 (2010) 94–105, http://dx.doi.org/10.1016/j.str.2009.10.018.Crystal.

[24] D. Gay, Y.-O. You, A. Keatinge-Clay, D.E. Cane, Structure and stereospecificity of the dehydratase domain from the terminal module of the rifamycin polyketide synthase, Biochemistry 52 (2013) 8916–8928, http://dx.doi.org/10.1021/bi400988t.

[25] A. Keatinge-Clay, Crystal structure of the erythromycin polyketide synthase dehydratase, J. Mol. Biol. 384 (2008) 941–953, http://dx.doi.org/10.1016/j.jmb.2008.09.084.Crystal.

[26] M. Welby-Gieusse, J.F. Tocanne, Configuration absolue du phtiocerol A, du phtiotriol A et de la phtiodiolone A, Tetrahedron 26 (1970) 2875–2882.
[27] E. Casas-Arce, B. Ter Horst, B.L. Feringa, A.J. Minnaard, Asymmetric total synthesis of PDIM A: a virulence factor of *Mycobacterium tuberculosis*, Chem. Eur. J. 14 (2008) 4157–4159, http://dx.doi.org/10.1002/chem.200800243.
[28] S. Barroso, R. Castelli, M.P. Baggelaar, D. Geerdink, B. Terhorst, E. Casas-Arce, H.S. Overkleeft, G.A. Van Der Marel, J.D.C. Codée, A.J. Minnaard, Total synthesis of the triglycosyl phenolic glycolipid PGL-tb1 from *Mycobacterium tuberculosis*, Angew. Chem. Int. Ed. 51 (2012) 11,774–11,777, http://dx.doi.org/10.1002/anie.201206221.
[29] P. Caffrey, Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases, Chembiochem 4 (2003) 654–657, http://dx.doi.org/10.1002/cbic.200300581.
[30] A.T. Keatinge-Clay, A tylosin ketoreductase reveals how chirality is determined in polyketides, Chem. Biol. 14 (2007) 898–908, http://dx.doi.org/10.1016/j.chembiol.2007.07.009.
[31] D.J. Bevitt, J. Staunton, P.F. Leadlay, Mutagenesis of the dehydratase active site in the erythromycin-producing polyketide synthase, Biochem. Soc. Trans. 21 (1993) 30S.
[32] J. Wu, T.J. Zaleski, C. Valenzano, C. Khosla, D.E. Cane, Polyketide double bond biosynthesis. Mechanistic analysis of the dehydratase-containing module 2 of the picromycin/methymycin polyketide synthase, J. Am. Chem. Soc. 127 (2005) 17,393–17,404, http://dx.doi.org/10.1021/ja055672+.
[33] W.D. Fiers, G.J. Dodge, D.H. Sherman, J.L. Smith, C.C. Aldrich, Vinylogous dehydration by a polyketide dehydratase domain in curacin biosynthesis, J. Am. Chem. Soc. 138 (2016) 16024–16036, http://dx.doi.org/10.1021/jacs.6b09748 (jacs.6b09748).
[34] G. Etienne, W. Malaga, F. Laval, A. Lemassu, C. Guilhot, M. Daffé, Identification of the polyketide synthase involved in the biosynthesis of the surface-exposed lipooligosaccharides in mycobacteria, J. Bacteriol. 191 (2009) 2613–2621, http://dx.doi.org/10.1128/JB.01235-08.
[35] A.T. Keatinge-Clay, Stereocontrol within polyketide assembly lines, Nat. Prod. Rep. 33 (2016) 141–149, http://dx.doi.org/10.1039/C5NP00092K.
[36] J.W. Labonte, C.A. Townsend, Active site comparisons and catalytic mechanisms of the hot dog superfamily, Chem. Rev. 113 (2013) 2182–2204, http://dx.doi.org/10.1021/cr300169a.
[37] S. Cabantous, T. Terwilliger, G. Waldo, Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein, Nat. Biotechnol. 23 (2005) 102–107, http://dx.doi.org/10.1038/nbt1044.
[38] S.A. Guerrero, H.J. Hecht, B. Hofmann, H. Biebl, M. Singh, Production of selenomethionine-labelled proteins using simplified culture conditions and generally applicable host/vector systems, Appl. Microbiol. Biotechnol. 56 (2001) 718–723, http://dx.doi.org/10.1007/s002530100690.
[39] A.G.W. Leslie, H.R. Powell, Processing diffraction data with mosflm, in: Randy J. Read, Joel L. Sussman (Eds.), Evol. Methods Macromol. Crystallogr, 1, 245, Springer, Netherlands 2007, pp. 41–51, http://dx.doi.org/10.1074/jbc.275.13.9468.
[40] P.R. Evans, An introduction to data reduction: space-group determination, scaling and intensity statistics, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 282–292, http://dx.doi.org/10.1107/S090744491003982X.
[41] G.M. Sheldrick, Experimental phasing with SHELXC/D/E: combining chain tracing with density modification, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 479–485, http://dx.doi.org/10.1107/S0907444909038360.
[42] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 486–501, http://dx.doi.org/10.1107/S0907444910007493.
[43] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, Acta Crystallogr. D Biol. Crystallogr. 53 (1997) 240–255, http://dx.doi.org/10.1107/S0907444996012255.
[44] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G.W. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and current developments, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 235–242, http://dx.doi.org/10.1107/S0907444910045749.
[45] A. Quemard, J.C. Sacchettini, A. Dessen, C. Vilcheze, R. Bittman, W.R. Jacobs, J.S. Blanchard, Enzymatic characterization of the target for isoniazid in *Mycobacterium tuberculosis*, Biochemistry 34 (1995) 8235–8241.
[46] S. Parikh, D.P. Moynihan, G. Xiao, P.J. Tonge, Roles of tyrosine 158 and lysine 165 in the catalytic mechanism of InhA, the enoyl-ACP reductase from *Mycobacterium tuberculosis*, Biochemistry 38 (1999) 13,623–13,634, http://dx.doi.org/10.1021/bi990529c.
[47] W. Kabsch, XDS, Acta Crystallogr. D Biol. Crystallogr. D66 (2010) 125–132, http://dx.doi.org/10.1107/S0907444909047337.
[48] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, J. Appl. Crystallogr. 40 (2007) 658–674, http://dx.doi.org/10.1107/S0021889807021206.
[49] P.D. Adams, P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.-W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive python-based system for macromolecular structure solution, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 213–221, http://dx.doi.org/10.1107/S0907444909052925.
[50] D.A. Case, J.T. Berryman, R.M. Betz, D.S. Cerutti, T.E. Cheatham III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K.M. Merz, G. Monard, P. Needham, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, R. Salomon-Ferrer, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, D.M. York, P.A. Kollman, AMBER 2015, Univ. California, San Fr., 2015
[51] B.R. Miller, T.D. McGee, J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, MMPBSA.py: an efficient program for end-state free energy calculations, J. Chem. Theory Comput. 8 (2012) 3314–3321, http://dx.doi.org/10.1021/ct300418h.
[52] T. Maier, M. Leibundgut, N. Ban, The crystal structure of a mammalian fatty acid synthase, Science 321 (2008) 1315–1322, http://dx.doi.org/10.1126/science.1161269.
[53] K. Maity, B.S. Venkata, N. Kapoor, N. Surolia, A. Surolia, K. Suguna, Structural basis for the functional and inhibitory mechanisms of β-hydroxyacyl-acyl carrier protein

dehydratase (FabZ) of *Plasmodium falciparum*, J. Struct. Biol. 176 (2011) 238–249, http://dx.doi.org/10.1016/j.jsb.2011.07.018.

[54] L. Zhang, J. Xiao, J. Xu, T. Fu, Z. Cao, L. Zhu, H.-Z Chen, X. Shen, H. Jiang, L. Zhang, Crystal structure of FabZ-ACP complex reveals a dynamic seesaw-like catalytic mechanism of dehydratase in fatty acid biosynthesis, Cell Res. 26 (2016) 1330–1344, http://dx.doi.org/10.1038/cr.2016.136.

[55] C. Nguyen, R.W. Haushalter, D.J. Lee, P.R.L. Markwick, J. Bruegger, G. Caldara-Festin, K. Finzel, D.R. Jackson, F. Ishikawa, B. O'Dowd, J.A. McCammon, S.J. Opella, S.C. Tsai, M.D. Burkart, Trapping the dynamic acyl carrier protein in fatty acid biosynthesis, Nature 505 (2014) 427–431, http://dx.doi.org/10.1038/nature12810.

[56] L. Holm, P. Rosenström, Dali server: Conservation mapping in 3D, Nucleic Acids Res. 38 (2010) 545–549, http://dx.doi.org/10.1093/nar/gkq366.

[57] R. Siméone, P. Constant, C. Guilhot, M. Daffé, C. Chalut, Identification of the missing trans-acting enoyl reductase required for phthiocerol dimycocerosate and phenolglycolipid biosynthesis in Mycobacterium tuberculosis, J. Bacteriol. 189 (2007) 4597–4602, http://dx.doi.org/10.1128/JB.00169-07.

Supporting information

Insights into Substrate Modification by Dehydratases from Type I Polyketide Synthases

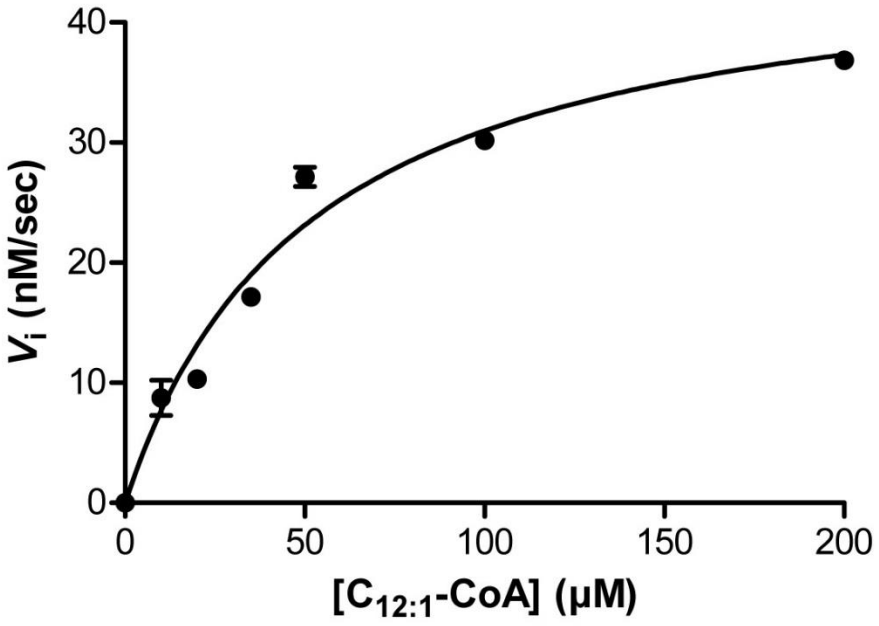


Figure S1. Determination of the steady-state kinetic constants of DH_{PpsC} for dodecenoyl-CoA. The hydratase activities were measured in a quartz cuvette at 263 nm in the presence of 250 nM enzyme, variable concentrations of $C_{12:1}$-CoA (0-200 μM), in 100 mM sodium phosphate buffer pH 7.0. Data were then fitted to the Michaelis-Menten equation by least-squares fitting to a hyperbola using the program GraphPad Prism. Values are means of replicates ± SEM. Non visible error bars correspond to very small SEM values.

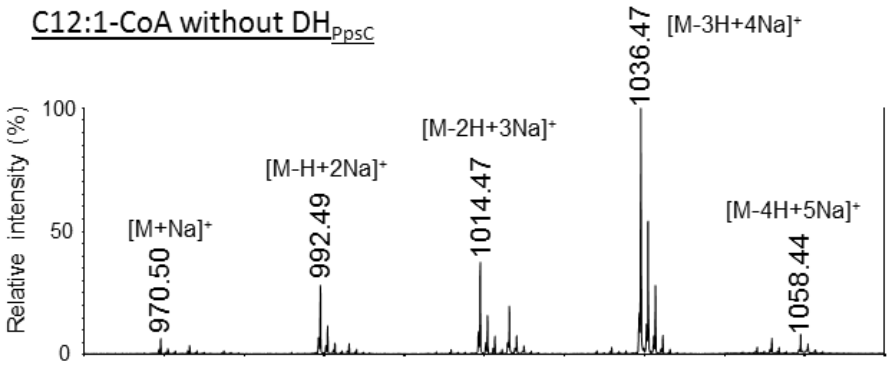


Figure S2. MALDI-TOF MS negative control. The substrate ion peaks are labeled in black.

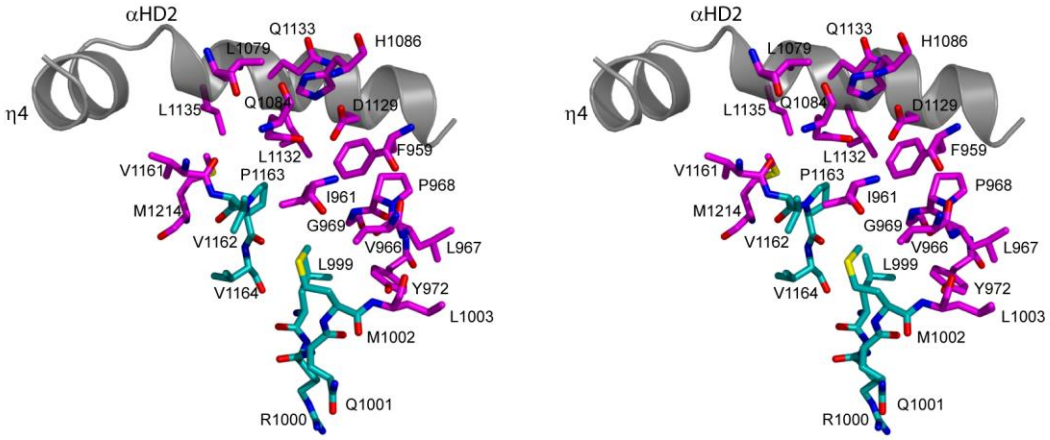


Figure S3. Stereoview of the substrate-binding cavity for the DH$_{PpsC}$ H959F variant in its apo form. Residues lining the tunnel are shown as sticks. The cavity entrance is delineated by residues 999-1002 and 1162-1164 with carbon atoms colored in cyan. The carbon atoms of residues lining the thioester group and acyl chain are colored in magenta.

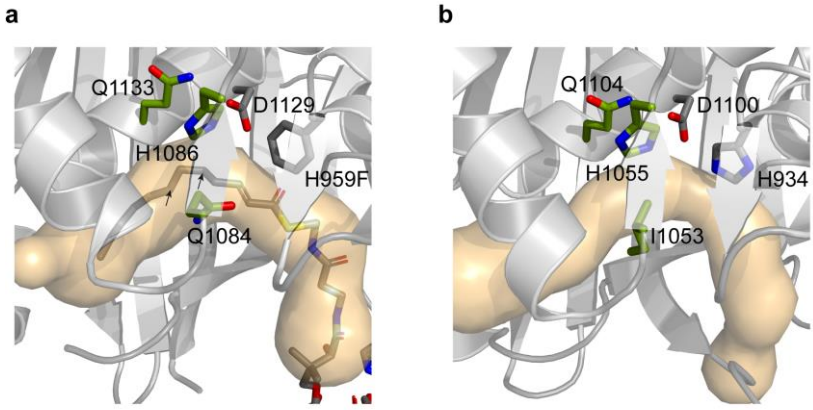


Figure S4. The substrate binding tunnel. Ribbon diagram of the dehydratase domain from (a) PpsC and (b) Pks5 (PDB code 5BP2, 1.9 Å root mean square deviation between superimposed Cα atoms). The V-shaped cavity is represented as a surface colored in light orange. The active site histidine and aspartate residues are shown as sticks (Cα atoms in gray, nitrogen atoms in blue, oxygen atoms in red). Polar residues pointing inside the cavity are indicated with green Cα atoms. The tunnel surface was calculated using CAVER [1]. Black arrows indicate the positions of the C5 and C7 atoms.

Table S1. Data collection, phasing, and refinement statistics of the DH_{PpsC}.

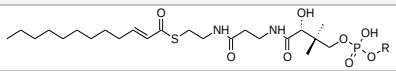
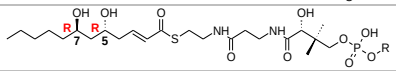
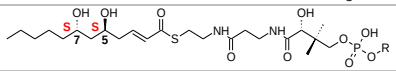
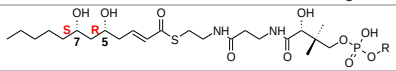
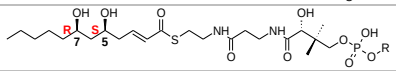
	WT – SeMet			H959F/apo	H959F/$C_{4:1}$-CoA	H959F/$C_{12:1}$-CoA
	Peak	*Inflection*	*Low remote*			
Data collection	ID29 (ESRF)			ID29 (ESRF)	ID29 (ESRF)	XALOC (ALBA)
Wavelength	0.979138	0.976250	0.979500	1.07200	0.97625	0.97922
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$C121$
Cell parameters - a, b, c (Å) - α, β, γ (°)	89.5, 89.5, 159.3 90, 90, 90			83.8, 83.8, 165.8 90, 90, 90	83.9, 83.9, 166.8 90, 90, 90	117.4, 70.8, 42.6 90, 104.9, 90
Resolution (Å)[a]	46.0 – 3.5 (3.71 – 3.50)	41.9 – 3.1 (3.29 – 3.10)	36.4 – 2.7 (2.80 – 2.70)	41.9 – 2.9 (3.00 – 2.90)	46.4 – 3.2 (3.39 – 3.20)	41.1 – 1.5 (1.59 – 1.50)
R_{meas} (%)[b]	7.8 (28.5)	18.6 (137.7)	7.3 (116.8)	11.2 (13.9)	17.0 (106.3)	4.7 (84.6)
CC(1/2)* (%)	99.9 (99.8)	99.7 (88.6)	99.9 (91.9)	100 (96.6)	99.4 (74.6)	99.9 (71.1)
Anomal Corr.	72 (11)					
I/σ(I)	24.0 (7.8)	7.8 (0.9)	15.8 (1.4)	18.7 (1.9)	9.0 (1.2)	15.9 (1.6)
Completeness (%)	99.8 (99.3)	95.3 (96.1)	99.8 (99. 7)	99.0 (100)	95.2 (95.1)	98.6 (97.2)
Reflections, total	179465 (29567)	51727 (4832)	116065 (15692)	144412 (13793)	57676 (8751)	244719 (38302)
Reflections, unique	15640 (2534)	10867 (1064)	18423 (1796)	13660 (1334)	9975 (1565)	53493 (8481)
Redundancy	11.5 (11.7)	4.8 (4.5)	6.3 (5.7)	10.6 (10.4)	5.8 (5.6)	4.6 (4.5)
Refinement						
Resolution (Å)			36.4 – 2.7	41.9 – 2.9	46.4 – 3.2	41.1 – 1.6
R_{work}/R_{free} (%/%)[c]			0.208/0.251	0.182/0.215	0.251/0.289	0.163/0.194
No. of non-H atoms - Protein - Water - Ligand			 1978 5 5	 2029 - -	 1980 - 53	 2041 293 122
Mean B value (Å²) - Protein - Water - Ligand			 90.1 81.5 138.9	 82.6 - -	 86.7 - 147.8	 33.2 42.2 39.1
Rmsd bond lengths (Å)			0.010	0.003	0.004	0.010
Rmsd bond angles (°)			1.71	0.55	0.84	1.47
Ramachandran - favored region - outliers			 96.18 0.38	 90.98 0.75	 87.60 0.78	 97.72 0.00

[a] Highest resolution shell is shown in parentheses.

[b] $Rmeas = (\sum_{hkl}\sqrt{\frac{n}{n-1}}\sum_{i=1}^{n}|I_i(hkl) - \bar{I}(hkl)|)/\sum_{hkl}\sum_{i=1}^{n}I_i(hkl)$

[c] $Rwork = \sum_{hkl}(|F_{obs}(hkl)| - |F_{calc}(hkl)|)/\sum_{hkl}|F_{obs}(hkl)|$ and Rfree is the R value for a test set of reflections consisting of a random 5% of the diffraction data not used in refinement.

Table S2. MM/GBSA Estimated binding energies.

Ligand	Structure	Estimated binding energy (kJ.mol^{-1})
$C_{12:1}$-CoA		-63.1 +/- 4.9
5R, 7R $C_{12:1}$-CoA		-61.1 +/- 8.2
5S, 7S $C_{12:1}$-CoA		-60.9 +/- 7.9
5R, 7S $C_{12:1}$-CoA		-56.4 +/- 5.0
5S, 7R $C_{12:1}$-CoA		-63.0 +/- 5.6

Reference

1. Chovancova E, Pavelka A, Benes P, Strnad O, Brezovsky J, Kozlikova B, Gora A, Sustr V, Klvana M, Medek P, et al. (2012) CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput Biol* **8**: 23–30.