Journal of Structural Biology 213 (2021) 107776

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Substrate specificity and conformational flexibility properties of the *Mycobacterium tuberculosis* β-oxidation trifunctional enzyme

Subhadra Dalwani [a], Outi Lampela [a,b], Pierre Leprovost [a,b], Werner Schmitz [c], André H. Juffer [a,b], Rik K. Wierenga [a,b], Rajaram Venkatesan [a,*]

[a] Faculty of Biochemistry and Molecular Medicine, University of Oulu, Oulu, Finland
[b] Biocenter Oulu, University of Oulu, Oulu, Finland
[c] Theoder-Boveri-Institut für Biowissenschaften der Universität Würzburg, Würzburg, Germany

ARTICLE INFO

Keywords:
Trifunctional enzyme
Substrate specificity
Fatty acid β-oxidation
M. tuberculosis

ABSTRACT

The *Mycobacterium tuberculosis* trifunctional enzyme (MtTFE) is an $\alpha_2\beta_2$ tetrameric enzyme. The α-chain harbors the *2E*-enoyl-CoA hydratase (ECH) and 3S-hydroxyacyl-CoA dehydrogenase (HAD) activities and the β-chain provides the 3-ketoacyl-CoA thiolase (KAT) activity. Enzyme kinetic data reported here show that medium and long chain enoyl-CoA molecules are preferred substrates for MtTFE. Modelling studies indicate how the linear medium and long acyl chains of these substrates can bind to each of the active sites. In addition, crystallographic binding studies have identified three new CoA binding sites which are different from the previously known CoA binding sites of the three TFE active sites. Structure comparisons provide new insights into the properties of ECH, HAD and KAT active sites of MtTFE. The interactions of the adenine moiety of CoA with loop-2 of the ECH active site cause a conformational change of this loop by which a competent ECH active site is formed. The NAD^+ binding domain (domain C) of the HAD part of MtTFE has only a few interactions with the rest of the complex and adopts a range of open conformations, whereas the A-domain of the ECH part is rigidly fixed with respect to the HAD part. Two loops, the CB1-CA1 region and the catalytic CB4-CB5 loop, near the thiolase active site and the thiolase dimer interface, have high B-factors. Structure comparisons suggest that a competent and stable thiolase dimer is formed only when complexed with the α-chains, highlighting the importance of the assembly for the proper functioning of the complex.

1. Introduction

For *Mycobacterium tuberculosis* (*Mtb*), lipid metabolism is important for its survival in the mammalian host and a large portion of its genome is accounted for by genes encoding proteins involved in lipid metabolism. Of its genome ~250 genes are known to code for enzymes involved in fatty acid metabolism as compared to only ~50 in *Escherichia coli* (Cole et al., 1998). *Mtb* is metabolically versatile and capable of switching its metabolic state during the intracellular stage to a state in which it can grow and survive using lipids (cholesterol and fatty acids) as a carbon source rather than glucose and glycerol (Lee et al., 2013; Nazarova et al., 2019; Sassetti and Rubin, 2003; Schnappinger et al., 2003; Wilburn et al., 2018). It has been reported that the *Mtb* genes required for the β-oxidation of fatty acids, a central pathway in lipid degradation, are upregulated in the intraphagosomal environment (Schnappinger et al., 2003). This adaptive mechanism could especially be advantageous during the latent stage of infection where *Mtb* utilizes fatty acids that are available from the host organism to persist and survive.

The β-oxidation cycle consists of four steps (Fig. S1), being (i) an acyl-CoA dehydrogenation (ACD), (ii) an enoyl-CoA hydration (ECH), (iii) a 3S-hydroxyacyl-CoA dehydrogenation (HAD) and (iv) a 3-ketoacyl-CoA thiolytic cleavage (KAT) that produces acetyl-CoA and acyl-CoA molecules that are two carbon atoms shorter. For example, the complete degradation pathway of palmitoyl-CoA into 8 molecules of acetyl-CoA requires 7 cycles consisting of 28 chemical steps. The enzymes catalyzing these reactions can be monofunctional as well as multifunctional enzymes. The latter enzymes can be bifunctional, like multifunctional enzyme type-1 (MFE1, catalyzing the ECH and HAD reactions), or trifunctional (like the trifunctional enzyme, TFE, catalyzing the ECH, HAD and KAT reactions). *Mtb* has multiple copies of genes that putatively code for the enzymes of each of these steps. For

* Corresponding author.
E-mail address: rajaram.venkatesan@oulu.fi (R. Venkatesan).

https://doi.org/10.1016/j.jsb.2021.107776
Received 14 April 2021; Received in revised form 30 June 2021; Accepted 4 August 2021
Available online 8 August 2021

example, it has 35 putative ACDs (*fadE1-E35*), 21 putative ECHs (*echA1-21*), 4 putative HADs (*fadB2-fadB5*) and 5 putative KATs (*fadA1-fadA5*) (Cole et al., 1998). While the function of several of these enzymes remains unknown, some have been better characterized (Bonds et al., 2020; Cox et al., 2019; Srivastava et al., 2015; Wipperman et al., 2014). Although there are multiple monofunctional β-oxidation enzymes, *Mtb* has only one trifunctional β-oxidation enzyme complex (MtTFE), which is encoded by the genes *fadA* and *fadB* (Venkatesan and Wierenga, 2013). The TFE complex is an $\alpha_2\beta_2$-tetramer (Fig. 1), which is capable of catalyzing the ECH (at the N-terminus of the α-chain), the HAD (at the C-terminus of the α-chain) and the KAT reactions (β-chain). In the $\alpha_2\beta_2$-tetramer the two α-chains are assembled with the β_2-dimer by extensive protein–protein interfaces (Fig. 1). The folds of both chains consist of several domains, as shown in Fig. S2. The MtTFE α-chain is 720 residues (Fig. S3) and it is homologous to the monomeric MFE1. The best characterized MFE1 is the rat peroxisomal MFE1 (RnMFE1). The sequence identity between these two sequences is 27%. The sequence alignment of MtTFE-α and RnMFE1 is provided in Fig. S3. The β-chain is 403 residues and the β_2-dimer is homologous to the thiolase dimer. The sequence alignment of MtTFE-β with a close homologue, the *Mtb* FadA5 thiolase, is given in Fig. S4. A comparison of 130 thiolase sequences has suggested that the FadA5 and MtTFE thiolase belong to the same thiolase subfamily (Anbazhagan et al., 2014). The sequence identity between these two sequences is 40% (Fig. S4). Figs. S3 and S4 provide the secondary structure nomenclature of the MtTFE α- and MtTFEβ-chains.

Bioinformatics studies have identified four TFE subfamilies (Venkatesan and Wierenga, 2013): (i) the mammalian mitochondrial TFE, like the human TFE (HsTFE), (ii) the mycobacterial TFE, like MtTFE, (iii) the bacterial aerobic TFE, like *E. coli* TFE (EcTFE) and its close homologue *Pseudomonas fragi* TFE (PfTFE) and (iv) the bacterial anaerobic TFE (such as the *E. coli* anEcTFE). Structural studies (Ishikawa et al., 2004; Liang et al., 2018; Sah-Teli et al., 2020; Xia et al., 2019) have shown that each of these TFEs are $\alpha_2\beta_2$ heterotetramers with the mycobacterial TFE, bacterial aerobic TFE and mitochondrial TFE showing very different assemblies. anEcTFE is proposed to have an assembly that resembles the mitochondrial TFE (like HsTFE) in its tetrameric assembly (Sah-Teli et al., 2019). anEcTFE and HsTFE are membrane associated, whereas EcTFE and MtTFE are soluble. The β_2/thiolase dimer forms the core of the tetramer in each of these subfamilies.

Enzyme kinetic studies have shown that EcTFE can degrade short chain, medium chain and long chain linear enoyl-CoA substrates, such that the best substrate is short chain enoyl-CoA substrate (Sah-Teli et al., 2019). The enzymological properties of PfTFE have not been studied in such great detail, but in any case the linear enoyl-CoA molecules are substrates (Imamura et al., 1990), and long chain enoyl-CoAs are more efficiently degraded than short chain enoyl-CoA molecules (Sato et al., 1992). The systematic substrate specificity studies of HsTFE and anEcTFE show that these TFEs degrade medium and long chain linear enoyl-CoA molecules much better than short chain enoyl-CoA (Sah-Teli et al., 2019). In addition, it has been shown that EcTFE (Yang et al., 1986) and HsTFE (Nada et al., 1995) have substrate channeling properties, like RnMFE1 (Yang et al., 1986). The function of EcTFE, anEcTFE and HsTFE concerns the degradation of linear *2E*-enoyl-CoA substrate molecules.

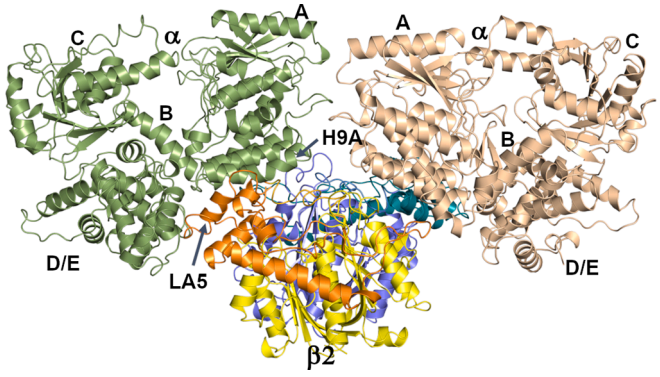

Fig. 1. The overall assembly of the MtTFE $\alpha_2\beta_2$ tetramer. The α subunits are shown in green (left) and wheat (right), the β subunits of the thiolase dimer are shown in yellow (front) and blue (back). The thiolase loop domains are in orange and turquoise of the yellow and blue thiolase subunits, respectively. The twofold axis runs vertically.

The three active sites of these TFEs, being the ECH, the HAD and the KAT active sites, have a very different shape and the mode of binding of CoA to each of these active sites is very different (Zhang et al., 2010). The ECH active site is formed by the crotonase fold of 2*E*-enoyl-CoA hydratases (also referred to as crotonases) and the CoA is bound in a bent conformation in the ECH active site of these crotonase fold enzymes, such that the adenine moiety is hydrogen bonded to the active site loop (loop-2) that forms the oxyanion hole and its phosphate moieties point into bulk solvent. The acyl tail is bound in a tunnel that connects the catalytic site with the bulk solvent (Kasaragod et al., 2013). In the HAD enzymes the pantetheine part of the CoA moiety is bound in an extended conformation in the pantetheine binding tunnel (Sridhar et al., 2020; Barycki et al., 2000), whereas the acyl tail is bound in a groove shaped by domain D. In the thiolase active site, the CoA moiety of the substrate is bound in a bulk solvent exposed groove and the acyl tail is bound either in a deep pocket (like in the *Zoogloea ramigera* biosynthetic thiolase (Haapalainen et al., 2006) or in a solvent exposed groove, like in the peroxisomal SCP2-thiolase (Kiema et al., 2019) and in the *Mtb* FadA5 thiolase (Schaefer et al., 2015).

The substrate specificity of MtTFE has not been studied systematically. It has been reported that *2E*-decenoyl-CoA is efficiently degraded (Venkatesan and Wierenga, 2013) and also it has been shown that the ECH active site catalyzes the isomerization of *3E*-octenoyl-CoA into *2E*-octenoyl-CoA, but *3Z*-octenoyl-CoA is not a substrate for this isomerization reaction (Srivastava et al., 2015). Here, the substrate specificity of MtTFE is studied systematically by enzyme kinetic measurements with linear enoyl-CoA derivatives of varying tail lengths, as well as having a 2-methyl branched tail, with or without a steroid moiety at its ω-end. Protein crystallographic studies of unliganded and ligand bound structures and the comparison with the structures of well-studied homologues reveal the importance of conformational flexibility in each of the three active sites.

2. Material and methods

2.1. Cloning

Seven active site variants of MtTFE were generated with either one, two or three active site point mutations. The previously reported active site residues of each of the active sites namely αGlu141 of the ECH part, αHis462 of the HAD part and βCys92 of the KAT active site were mutated to alanine using the QuickChange Lightning Multi Site-Directed Mutagenesis Kit from Agilent Technologies either as single mutations (MtTFE-αE141A, MtTFE-αH462A, MtTFE-βC92A) or in a combination to give the desired double (MtTFE-αE141AαH462A, MtTFE-αE141AβC92A, MtTFE-αH462AβC92A) or triple (MtTFE-αE141AαH462AβC92A) active site mutated variant. A previously cloned pET-Duet construct containing wild type MtTFE-α (*Mtb fadB*) and MtTFE-β (*Mtb fadA*) at MCSII and MCSI, respectively (Venkatesan and Wierenga, 2013), was used as the template to generate these variants. A list of the primers, that were used to generate the mutated variants, is given in Table S1. For the expression and purification of MtTFE-α only (MtTFE-alpha), a previously cloned construct (Venkatesan and Wierenga, 2013) of pACYC-Duet, containing *Mtb fadB* at MCSI, was used.

2.2. Protein expression and purification

MtTFE, each of its mutated variants and the separate MtTFE-alpha chain (not assembled with thiolase) were expressed in *E. coli* BL21

(DE3) cells and purified as previously described (Venkatesan and Wierenga, 2013). Briefly, the cells were grown at 37 °C in M9ZB media to an optical density of 0.6 at 600 nm and were then induced with 0.2 mM of IPTG at 22 °C for 15 h. The cells were then resuspended in 20 mM Tris buffer containing 500 mM NaCl and 5 mM imidazole at pH 8.0 and lysed by sonication. Subsequently, the proteins were purified by Ni-NTA affinity chromatography by gradient elution using 20 mM Tris buffer pH 8.0, containing 500 mM NaCl and increasing the imidazole concentration from 5 mM to 100 mM. The eluted peak fractions were pooled and excess of salt and imidazole removed by a buffer exchange protocol using a 100 kDa Centricon concentrator into 20 mM Tris, pH 7.9, containing 50 mM NaCl. The protein (in 20 mM Tris buffer, pH 7.9, 50 mM NaCl) was then subjected to anion exchange chromatography on a 6 mL Resource-Q column, using a gradient from the protein buffer to the elution buffer (20 mM Tris, 0.5 M NaCl, pH 7.9). Fractions containing MtTFE or its variants were then finally passed through a HiLoad 16/60 Superdex 200 size exclusion chromatography (SEC) column previously equilibrated with a buffer containing 20 mM HEPES, 120 mM KCl, 1 mM EDTA and 1 mM NaN_3 at pH 7.2. For some wild type MtTFE purifications another buffer (20 mM MOPS buffer pH 7.0) was used at the SEC step, resulting in a different crystal form in the crystallization step (having space group $C222_1$, instead of C2). The peak fractions were pooled, concentrated, frozen in 30–100 μL aliquots and stored at −70 °C for further studies.

2.3. Synthesis of the acyl-CoA substrates

Acetoacetyl-CoA and *2E*-butenoyl-CoA were purchased from Sigma–Aldrich (Missouri, U.S.A.). *2E*-decenoic acid and other chemicals used for substrate synthesis were from Merck (Darmstadt, Germany). 2-Methyl-*2E*-decenoic acid, *2E*-hexadecenoic acid, *24E*-THCe acid (Fig. S5) and *24E*-25DM-THCe acid (Fig. S5) were synthesized and conjugated with CoA as described previously (Sah-Teli et al., 2019). The identity of the acyl-CoA esters was confirmed by TLC, [^{1}H] NMR and high resolution mass spectrometry (HR-MS) analysis. 3,22-dioxo-chol-4-ene-24-oyl-CoA (Schaefer et al., 2015) (Fig. S5) was synthesized and provided by the laboratory of Professor Nicole S. Sampson, Stony Brook, New York.

2.4. Enzyme activity assays

Enzyme activities were measured by spectrophotometric activity assays using a Jasco V660 spectrophotometer as described earlier (Venkatesan and Wierenga, 2013). All assays were performed at 25 °C using 50 mM Tris (pH 8.5), 50 mM KCl, and 50 μg mL^{-1} bovine serum albumin as reaction buffer with a final volume of 500 μL in quartz cuvettes and were monitored for 3 min. To measure the specific activity of the ECH active site, 60 μM of each substrate was mixed into the reaction buffer and the reaction was initiated by adding 5 ng of MtTFE. ECH activity was measured by monitoring the disappearance of the C–C double bond of the substrate at 263 nm. A molar extinction coefficient of 6700 $M^{-1}cm^{-1}$ was used for calculating the specific activity. The specific activity is expressed as the number of μmole substrate converted per mg of enzyme, per minute. To measure the specific activity of the HAD active site, 60 μM of substrate, 1 mM CoA and 1 mM NAD^+ were mixed into the reaction buffer and the reaction was initiated by adding 100 ng of MtTFE to the reaction mixture. HAD activity was measured by monitoring the formation of NADH at 340 nm and by using an absorption coefficient of 6220 $M^{-1}cm^{-1}$ for NADH for calculating the specific activity. To determine the k_{cat} and K_m values with *2E*-hexadecenoyl-CoA for the ECH and HAD activities, the above reactions were repeated by varying the substrate concentration from 1 μM to 120 μM. GraphPad Prism version 8.4.2 was used to plot and analyze the data and the linear part of the progress curve was used to determine the initial rate of the reaction.

KAT activity with *2E*-decenoyl-CoA and *2E*-hexadecenoyl-CoA was qualitatively determined. First, a Mg^{2+}/3-ketoacyl-CoA complex was generated by incubating 60 μM of the respective *2E*-enoyl-CoA substrate with 100 ng of MtTFE in the reaction buffer containing 1 mM NAD^+ and 5 mM $MgCl_2$ at room temperature. Formation of the complex was monitored at 303 nm for 5 min or until a plateau of the curve at 303 nm was observed. KAT activity was initiated by adding 0.5 mM of CoA to the reaction mixture and the reaction was monitored by following the disappearance of the Mg^{2+}/3-ketoacyl-CoA complex at 303 nm. An absorption coefficient of 13900 $M^{-1}cm^{-1}$ for the Mg^{2+}/3-ketoacyl-CoA complex was used for KAT activity calculations. In addition, the KAT activity with 60 μM 3,22-dioxo-chol-4-ene-24-oyl-CoA was monitored both spectroscopically (in the presence of Mg^{2+}) as well as by the direct mass spectrometry assay. For this, 300 μM of 3,22-dioxo-chol-4-ene-24-oyl-CoA, 100 μM CoA and 0.2 μM of enzyme were mixed into reaction buffer in a 10 μL reaction volume and left to react at room temperature for 20 min. ZipTip pipette tips from Merck Millipore were then used to desalt and concentrate the reaction mixture into 2 μL of 10 mg mL^{-1} DHB matrix in 50% acetonitrile, that was directly applied onto the MALDI TOFF plate. Samples were run in the positive reflectron mode and formation of 3-oxo-pregn-4-ene-20-carboxyl-CoA (3-OPC-CoA) was detected. The ECH, HAD and KAT specific activities for the point mutated variants, as well as for MtTFE-alpha, were determined using 60 μM of *2E*-decenoyl-CoA, as described above (Table S2).

2.5. Stability measurements

Protein stability was measured by monitoring protein unfolding by increasing the temperature. 1 mg mL^{-1} of each protein in 20 mM HEPES, 120 mM KCl, 1 mM EDTA and 1 mM NaN_3 at pH 7.2 was filled into 10 μL standard capillary tubes (NanoTemper Technologies) and loaded onto the Prometheus NT.48 instrument (NanoTemper Technologies). The temperature was increased from 20 °C to 95 °C at a rate of 1 °C per min. An excitation intensity of 20% at 280 nm was applied so that the measured emission intensities at 330 nm and 350 nm was between 5000 and 15,000 fluorescent counts. The melting temperature was obtained by recording emission intensities at 350 nm and 330 nm. The emission fluorescent intensity ratio (350 nm/330 nm) and the first derivative were calculated using the PR.ThermControl software (NanoTemper Technologies).

2.6. Crystallization and crystal soaking experiments of wild type MtTFE

Two crystallization protocols have been used in the crystallization experiments of wild type MtTFE. In the first protocol MtTFE was crystallized as described previously (Venkatesan and Wierenga, 2013). Briefly, 0.5 μL of MtTFE solution (6 mg mL^{-1} in 20 mM HEPES pH 7.2, 120 mM KCl, 1 mM EDTA and 1 mM NaN_3), was mixed with the crystallization well solution (2 M $(NH_4)_2SO_4$, 100 mM Tris pH 8.5) in a 1:1 ratio and crystallized by vapor diffusion in hanging drops at room temperature. The formation of the crystals was monitored using IceBear (Daniel et al., 2021) and the crystals were frozen in liquid nitrogen for diffraction studies.

In the second protocol the protein buffer and the well solution buffer was a MOPS buffer. 0.5 μL of MtTFE solution (6 mg mL^{-1} in 20 mM MOPS pH 7.0, 120 mM KCl, 1 mM EDTA and 1 mM NaN_3) was mixed with a well solution of 2 M $(NH_4)_2SO_4$, 100 mM MOPS pH 6.5, 10% ethanol in 1:1 ratio and crystallized by vapor diffusion in hanging drop plates at room temperature. The crystals were harvested and frozen in liquid nitrogen.

For the crystallographic binding studies unliganded crystals, grown at pH 8.5, were transferred into mother liquor that contained also CoA, or *2E*-butenoyl-CoA, or NAD^+ using solutions containing 2 M $(NH_4)_2SO_4$, 100 mM Tris pH 8.5 supplemented with either 10 mM CoA, or 2.7 mM *2E*-butenoyl-CoA or 5 mM NAD^+, respectively. The crystals soaked with NAD^+ and CoA were allowed to equilibrate for 6–7 min before cryocooling in liquid nitrogen and crystals soaked with *2E*-

butenoyl-CoA were soaked for 30 min before cryocooling in liquid nitrogen.

2.7. Crystallization of MtTFE-alpha

An optimization screen was setup, using hanging drop vapor diffusion method in a 24 well plate around an initial crystallization hit in which the PEG4000 concentration was varied from 20 to 22% and the concentration of isopropyl alcohol was varied from 20 to 22%. The hanging drops were made with 3 μL of 4 mg mL^{-1} protein (in storage buffer 20 mM HEPES pH 7.2, 120 mM KCl, 1 mM EDTA and 1 mM NaN_3) and 1 μL of crystallization solution. This resulted in large crystals, but these diffracted weakly (~3.5 Å). Therefore, crystals from 2 drops of this optimization screen were harvested and a seed stock was prepared by vigorous vortexing with a bead for 5 mins. 4 mg mL^{-1} of MtTFE-alpha, containing also freshly added 2 mM NAD^+ was incubated on ice for 30

Table 1
Data collection, data processing and structure refinement statistics of the wild type MtTFE and MtTFE-alpha structures.

Data set	MtTFE (αβ-unliganded)	MtTFE-CoA-A	MtTFE-CoA-B	MtTFE-CoA-C	MtTFE(NAD^+)	MtTFE-alpha (NAD^+)
Data collection						
Beamline	BESSY II (14–1)	ESRF-ID29	DLS-I24	Petra-P13	DLS-I04	DLS-I03
Detector	Pilatus 6M	Pilatus 6M	Pilatus3 6M	Pilatus 6M	Eiger2 XE 16M	Eiger2 XE 16M
Wavelength (Å)	0.9184	1.0722	0.9686	0.9763	0.9795	0.9762
Temperature (K)	100	100	100	100	100	100
Data processing						
Space group	$C222_1$	C2	C2	C2	C2	P1
Unit cell parameters						
a,b,c (Å)	104.09, 204.67, 132.79	248.25, 135.75, 119.30	245.41, 136.60, 117.91	247.5, 136.3, 117.7	243.75, 133.24, 115.60	49.81, 86.81, 88.96
α,β,γ (^{O})	90, 90, 90	90, 110.30, 90	90, 110.45, 90	90, 110.50, 90	90, 110.14, 90	88.02, 89.91, 75.19
Processing software	xdsapp, AIMLESS	xds, STARANISO	fast_dp	edna_proc	xia2_3dii	xia2_DIALS, AIMLESS
Resolution range (Å)	48.45–2.1 (2.22–2.10)	48.5–2.79 (2.90–2.79)	29.76–2.79 (2.83–2.79)	58.74–2.1 (2.14–2.10)	86.84–2.42 (2.46–2.42)	88.9–2.70 (2.82–2.70)
$R_{p.i.m.}$ (%)	0.041 (0.296)	0.065 (0.554)	0.047 (0.442)	0.058 (1.423)	0.046 (1.176)	0.062 (0.296)
$CC_{1/2}$	0.997 (0.780)	0.995 (0.473)	0.998 (0.581)	0.998 (0.410)	0.999 (0.289)	0.988 (0.867)
I/sigmaI	10.6 (2.0)	10.5 (1.5)	14 (1.3)	8.9 (0.8)	10.2 (0.7)	7.7 (2.0)
Completeness (%)	99.6 (98.3)	95.1 (83.0, ellipsoid)	97.6 (52.9)	99.7 (99.7)	99.4 (99.8)	97.6 (97.8)
Multiplicity	6.8 (6.6)	3.4 (3.3)	6.9 (3.3)	5.1 (4.8)	3.5 (3.4)	3.1 (3.3)
No. of reflections	1,094,346	279,549	608,741	1,070,948	455,984	119,102
No. of unique reflections	160,514	81,071	88,737	212,119	131,141	38,685
V_M ($Å^3$/Da)	2.9	3.9	3.9	3.9	3.7	2.3
Wilson *B*-factor ($Å^2$)	39.3	44.4	62.6	46.8	54.2	33.4
Refinement statistics						
Resolution range (Å)	48.45–2.10	48.5–2.79	29.76–2.79	58.74–2.10	72.92–2.42	83.97–2.70
No. of used reflections	83,073	80,078	88,718	212,013	131,052	38,651
R_{work} (%)	19.49	17.82	18.71	20.54	21.39	21.28
R_{free} (%)	21.93	21.57	23.11	23.36	24.45	26.99
Total No. of atoms	8885	17,229	17,341	17,556	17,045	10,631
No. of waters	343	229	111	265	89	14
Average *B* factors ($Å^2$)						
Protein	45.0, 38.7 (Chains B, C)	55.6, 49.2, 39.2, 39.8 (Chains A, B, C, D)	74.7, 70.2, 52.9, 56.5 (Chains A, B, C, D)	72.8, 70.4, 56.6, 59.2 (Chains A, B, C, D)	96.2, 91.3, 72.1, 68.2 (Chains A, B, C, D)	71.1, 66.5 (Chains A, B)
Ligands						
CoA(ECH) (A, B)	–	–	72.0, 76.5	106.7, 107.4	–	–
CoA(KAT) (C, D)	–	97.2, 45.5	82.5, 98.0	96.5, 106.1	–	–
CoA(HAD/KAT) (C, D)	–	104.2, 113.2	–	–	–	–
CoA(ECH2) (A, B)	–	–	116.6, 115.3	–	–	–
CoA(ECH/HAD) (A, B)	–	–	–	104.1, 100.3	–	–
NAD^+ (A, B)	–	–	–	–	140.0, 140.3	98.6, 90.1
Waters	40.7	35.6	48.5	56.1	63.9	36.3
Rms deviations						
rms bond length (Å)	0.002	0.001	0.002	0.002	0.002	0.002
rms bond angle (^{O})	0.50	0.42	0.51	0.54	0.49	0.48
Ramachandran plot (%)[1]						
Favored	97.2	96.9	96.9	97.5	96.5	96.0
Allowed	2.6	3.0	3.1	2.4	3.3	3.8
Outliers	0.2	0.1	0.0	0.1	0.2	0.2
PDB ID	7O4Q	7O4R	7O4S	7O4T	7O4V	7O4U

[1] As determined with MolProbity (Chen et al., 2010).

min. Crystals were set up using the same optimization screen by mixing 3 μL of MtTFE-alpha NAD^+ mixture, 0.5 μL seed stock and 1 μL well solution. The best crystals were grown in 100 mM TEA pH 7.5, 22% PEG 4000, 20% isopropyl alcohol at room temperature. The crystals were harvested two months after the crystallization plates were set up, and flash frozen in liquid nitrogen.

2.8. Cocrystallization of the point mutated variants

An optimization screen around the crystallization conditions for wild type MtTFE of 2 M $(NH_4)_2SO_4$, 100 mM Tris pH 8.5 (Venkatesan and Wierenga, 2013) was designed in which 2 M $(NH_4)_2SO_4$ was kept constant but the pH varied from 5.5 to 8.5. Each variant (6 mg mL^{-1}) was incubated with either 2 mM CoA, or 2 mM *2E*-decenoyl-CoA or 500 μM *2E*-hexadecenoyl-CoA for 30 min on ice and crystallization drops were setup using protein–ligand mixture (0.5 μL) and crystallization well solution (0.5 μL) in sitting drop plates using the Mosquito Nano dispensing robot. The plates were left at room temperature, for up to two months, before the crystals were cryocooled in liquid nitrogen and checked for diffraction. Crystals of MtTFE-αE141A and MtTFE-αE141AαH462AβC92A were harvested from wells at pH 5.5, crystals of MtTFE-αH462A from wells at pH 6.5 and crystals of MtTFE-βC92A, MtTFE-αE141AβC92A and MtTFE-αH462AβC92A from wells at pH 7.5.

2.9. Data collection, structure determination and structure refinement

X-ray diffraction data were collected at different beamlines at BESSY II, Diamond, ESRF and EMBL (Hamburg) (Table 1, Table S3). Data reduction and scaling was either performed manually using XDS (Kabsch, 2010), AIMLESS (Evans and Murshudov, 2013) and STARANISO (Global Phasing Limited) (http://www.globalphasing.com) or by using the various data processing pipelines, available at the different synchrotron facilities, including xdsapp (Krug et al., 2012; Sparta et al., 2016), xia2 (Winter, 2010), xia2_3dii (Winter, 2010), fast_dp (Winter and McAuley, 2011), edna_proc (Incardona et al., 2009) and xia2_DIALS (Winter et al., 2018) (Table 1, Table S3). Initial phases were calculated by molecular replacement with coordinates derived from PDB ID 4B3H as the search model using either PHASER (McCoy et al., 2007) or MOLREP (Vagin and Teplyakov, 2010). Positioned coordinates after molecular replacement were used as the initial model to perform iterative rounds of model building with COOT (Emsley et al., 2010) and refinement, using Phenix (Afonine et al., 2012; Liebschner et al., 2019) and REFMAC5 of CCP4 (Murshudov et al., 2011; Potterton et al., 2018; Winn et al., 2011). The structure quality was assessed using MolProbity (Chen et al., 2010). Ligands were included in the later stages of the refinement whenever these ligands were clearly defined by the density maps (Table S4 and Table S5). In some structures the catalytic cysteines of the KAT active sites are oxidized (Table S4). NCS restraints were used in the initial refinement of some structures. The final structures were obtained after refinement with Phenix, not using NCS restraints. The cores of NCS related subunits superimpose on each other with a RMS value of less than 0.4 Å for the corresponding Cα-atoms. The refinement statistics and the PDB entry codes of the refined structures are reported in Table 1 and Table S3.

2.10. Structure analysis

The SSM protocol (Krissinel and Henrick, 2004) of COOT (Emsley et al., 2010) was used routinely for the superposition of the homologous structures and the MtTFE structures. For the analysis of the binding pocket of the ECH active site, the structures of rat ECH in complex with octanoyl-CoA (PDB ID 2DUB, chain A) (Engel et al., 1998), and of RnMFE1 in complex with 3S-hydroxydecanoyl-CoA (PDB ID 3ZWC, molecule A) (Kasaragod et al., 2013) were used. For the analysis of the CoA and NAD^+ binding pockets of the HAD active site, the structure of human 3S-hydroxyacyl-CoA dehydrogenase (HsHAD) complexed with acetoacetyl-CoA and NAD^+ (PDB ID 1F0Y, chain A) (Barycki et al., 2000) as well as the RnMFE1 structure of its complex with 3-ketodecanoyl-CoA (PDB ID 5OMO, molecule A) (Sridhar et al., 2020) were used. For analyzing the KAT active site, the structures of acetylated biosynthetic thiolase of *Z. ramigera* complexed with acetyl-CoA (PDB ID 1DM3, chain B) (Modis and Wierenga, 2000), the C89A-variant of the thiolase of *Z. ramigera* complexed with acetoacetyl-CoA (PDB ID 1M1O, chain B) (Kursula et al., 2002) and the *Mtb* FadA5 thiolase complexed with 3-oxo-pregn-4-ene-20-carboxylic acid (PDB ID 4UBT, chain A) (Schaefer et al., 2015) were used.

The images of the structures were generated using PyMOL (the PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). The molecular surface, color coded according to electrostatic potential, has been calculated with CCP4MG (McNicholas et al., 2011). These calculations were done with the MtTFE-CoA-B structure in which the protein part of the model covers all residues from αMet1 to αSer720 of both α-chains and from βSer2 to βVal403 of both β-chains and in which all side chains were completely built, but in which the ligands and waters were removed. The schematic drawings were made with MarvinSketch (ChemAxon). Sequence alignments were made with CLUSTAL Omega (Madeira et al., 2019) and the alignment figures were made using the web based program ESPRIPT 3.0 (Robert and Gouet, 2014). Secondary structure elements were assigned according to DSSP (Kabsch and Sander, 1983).

The normalized B-factors were calculated using a bash script for extracting the individual B-factors of the Cα atoms of each residue, which were then imported into Microsoft EXCEL, allowing for the calculation of the normalized B-factor, using the formula (Yaseen et al., 2016):

$$\text{B-factor(normalized)} = (\text{B-factor} - \mu)/\sigma$$

where μ is the mean B-factor of the given structure and σ is the standard deviation. To assess which loops and domains of MtTFE are the most flexible regions of the α and β-chains, the computed normalized B values were plotted as a function of residue number for the various structures.

2.11. Molecular dynamics simulation calculations to relax and energy minimize the modelled substrates

The mode of binding of CoA in the ECH and KAT active sites is known from the structural studies and for the HAD active site it can be inferred from the structures of complexes of the RnMFE1 homologue. This knowledge has been used for model building the mode of binding of the acyl tail of acyl-CoA in each of the active sites. The starting model has been obtained through model building of the acyl-CoA molecule into the active sites of the refined $\alpha_2\beta_2$ MtTFE-CoA-B structure in COOT, after removing from this structure the hexahistidine tag of the α-chain as well as all ligands and all waters. Both α-chains of the MtTFE-CoA-B structure include residues Met1 to Ser720 and both β-chains include residues Ser2 to Val403.

The starting model of *2E*-decenoyl-CoA bound in the ECH active site was obtained from the mode of binding of 3S-hydroxydecanoyl-CoA to the ECH active site of RnMFE1 (PDB ID 3ZWC; molecule A (Kasaragod et al., 2013)). The CoA part of the substrate was modelled with COOT as seen in the MtTFE-CoA-B structure and the thioester oxygen was built in its oxyanion hole. The protonation state of the two active site glutamates was chosen to be consistent with a competent active site for catalysis of the hydration reaction of the bound *2E*-decenoyl-CoA, being that αGlu119 is deprotonated and αGlu141 is protonated (Zhang et al., 2010; Kasaragod et al., 2013).

The starting model of the HAD active site complexed with 3-ketodecanoyl-CoA was obtained from the mode of binding of 3-ketodecanoyl-CoA in the HAD active site of RnMFE1 (PDB ID 5OMO, molecule A, in which the HAD active site is complexed with 3-ketodecanoyl-CoA (Sridhar et al., 2020)). The protonation state of the catalytic histidine

was chosen to be consistent with the active site complexed with the product 3-ketodecanoyl-CoA, such that αHis462 is doubly protonated and its ND1 atom is hydrogen bonded to the side chain of αGlu474, which is deprotonated (Barycki et al., 2001). The superimposed ligand was modelled in COOT in the MtTFE-CoA-B structure, such that the hydrogen bond interactions of the pantetheine moiety, the thioester oxygen atom and the 3-keto oxygen atom were preserved, as described previously (Sridhar et al., 2020).

The starting model of the KAT active site complexed with 3-ketodecanoyl-CoA was obtained from the mode of binding of acetoacetyl-CoA in the *Z. ramigera* thiolase active site. βCys92 was mutated into an alanine, because the reference structure is the structure of the C89A variant complexed with acetoacetyl-CoA (PDB ID 1M1O, chain B (Kursula et al., 2002)). The catalytic histidine, βHis359 was singly protonated (at NE2) and its ND1 atom was hydrogen bonded to OG1(βThr364). The CoA part of the starting model was modelled in COOT from the mode of binding of CoA in the MtTFE-CoA-B structure, and the hydrogen bond interactions of the thioester oxygen atom and the 3-keto atom were modelled as described previously (Kursula et al., 2002).

The molecular dynamics (MD) calculations were done with the complete tetramer. The force field parameters for the ligands were obtained using a fragment-based approach. In each case, a fragment for CoA was taken from the R.E.D database (Dupradeau et al., 2008) and partial charges for the acyl tails were derived using the RESP method (Bayly et al., 1993) based on QM calculations with Gaussian *09* (Frisch et al., 2009) and the HF 6-31G* basis set, using the R.E.D server (Vanquelef et al., 2011). Parameters for the bonds, angles and dihedrals were taken from the Lipid 17 and GAFF2 Amber libraries (Wang et al., 2004).

The protein models for all simulations were prepared using the LEaP module of AmberTools (Salomon-Ferrer et al., 2013). The TIP3PBOX water model was used to solvate a rectangular box extending 10 Å around each model. All systems were neutralized with counter ions. The Amber ff14SB force field (Maier et al., 2015) was used for the protein and ions. The generated files were converted with the parmed tool of the AmberTools to be compatible for use with the GROMACS software suite (Van Der Spoel et al., 2005). Steepest decent energy minimization was first performed in 50,000 steps followed by equilibration steps in NVT and in NPT ensembles for 100 ps each. The protein and ligand heavy atoms were restrained during equilibration. Temperature was controlled using Velocity Rescaling at 300 K and pressure was maintained at 1 bar. The equilibrated systems were then used for 1 ns MD simulation production runs with 2 fs time steps. Coordinates and energies were saved every 2 ps. The structures, captured in the last frame of each of the MD runs, were energy minimized and the obtained structures were used to inspect and visualize the mode of binding of the ligand.

3. Results and discussion

3.1. Medium and long chain enoyl-CoAs are the preferred substrates for MtTFE

Specific activities of several possible substrates (Fig. S5) have been measured. Of the tested substrates, only the unbranched, medium and long chain *2E*-enoyl-CoA compounds are good substrates of the hydratase and subsequent dehydrogenase active sites (Table 2). No HAD activity was observed for *2E*-butenoyl-CoA. Likewise, no thiolase activity is detected with the substrate acetoacetyl-CoA. For 3-ketodecanoyl-CoA and 3-ketohexadecanoyl-CoA, the thiolase activities could only be determined qualitatively, following the protocol as described in the Material and methods section. This qualitative thiolase assay showed activity when testing for degradation of 3-ketodecanoyl-CoA and 3-ketohexadecanoyl-CoA. Only residual thiolase activity was found for 3,22-dioxo-chol-4-ene-24-oyl-CoA, which is the substrate of the FadA5 thiolase (Schaefer et al., 2015), by probing the assay mixture using mass spectrometry, by which the product could be detected. 3,22-dioxo-chol-4-ene-24-oyl-CoA (Fig. S5) is a derivative of cholesterol, which is an important metabolite for *Mtb*. This is a more hydrophobic compound than the two derivatives of THC-CoA, *24E*-THCe-CoA and *24E*-25DM-THCe-CoA (Fig. S5), which were found not to be converted by the hydratase active site (Table 2). THC-CoA is an intermediate of the mammalian bile acid synthesis pathway (Ferdinandusse et al., 2005).

Table 2
Specific activity of wild type MtTFE with different substrates[1].

Substrate	Specific activity (μmole mg^{-1} min^{-1})	
	ECH	HAD
2E-butenoyl-CoA	6.7 ± 1.3	ND
2E-decenoyl-CoA	349.8 ± 80.6	5.0 ± 1.4
2-methyl-*2E*-decenoyl-CoA	7.5 ± 4.4	ND
2E-hexadecenoyl-CoA	125.2 ± 15.9	5.3 ± 0.2
24E-THCe-CoA	ND	ND
24E-25DM-THCe-CoA	ND	ND

ND, not detectable.
[1] These measurements were done at 60 μM substrate concentration.

The results of the activity assay with the substrate of FadA5 (3,22-dioxo-chol-4-ene-24-oyl-CoA, Fig. S5) indicate that this is a poor substrate for the MtTFE KAT active site. The substrate specificity of MtTFE is also notably different from the substrate specificity of EcTFE, the other bacterial soluble TFE, which degrades best the short chain enoyl-CoA substrates (Sah-Teli et al., 2019). In the thiolase phylogenetic tree calculations EcTFE-thiolase (together with the PfTFE-thiolase) is in a different cluster than the FadA5 and MtTFE-thiolases (Anbazhagan et al., 2014).

These kinetic experiments show that medium and long chain, unbranched enoyl-CoA compounds are good substrates for both ECH and HAD active sites. For *2E*-decenoyl-CoA and *2E*-hexadecenoyl-CoA the Michaelis-Menten kinetic values were measured, as listed in Table 3 for the hydratase and dehydrogenase activities. These kinetic data show that these compounds are efficiently degraded by MtTFE. It can be noted that the k_{cat} values for the hydratase activity for these substrates are approximately 20 fold larger than for the corresponding dehydrogenase activity, whereas the K_m values show that the affinities for the respective substrates is highest for the dehydrogenase active sites. The Michaelis-Menten kinetic constants for the ECH and HAD activities of MtTFE for the medium and long chain, linear, acyl-CoA substrates show that MtTFE degrades these substrates with higher efficiency than HsTFE (Sah-Teli et al., 2019), whereas both MtTFE and HsTFE show no catalytic conversion when assayed with the short chain acyl-CoA molecules. The kinetic data therefore suggest that linear medium and long chain *2E*-enoyl-CoA molecules are the preferred substrates of MtTFE.

3.2. Structural and modelling studies identify the acyl tail binding pockets of the ECH, HAD and KAT active sites

In order to facilitate crystallographic binding studies with substrates, point mutation variants were made, in which each of the respective active sites were inactivated by point mutations. In the hydratase active site, it concerns the MtTFE-αE141A variant, in the HAD active site it is

Table 3
Michaelis-Menten constants of the ECH and HAD active sites of wild type MtTFE with *2E*-decenoyl-CoA and *2E*-hexadecenoyl-CoA substrates.

Substrate	ECH	HAD
2E-decenoyl-CoA (Venkatesan and Wierenga, 2013)		
k_{cat} (s^{-1})	380.7 ± 39.0	10.3 ± 0.2
K_m (μM)	15.3 ± 3.5	1.7 ± 0.1
k_{cat}/K_m ($M^{-1}s^{-1}$)	25.7 ± 5.4 10^6	6.2 ± 0.4 10^6
2E-hexadecenoyl-CoA		
k_{cat} (s^{-1})	192.6 ± 35.2	9.0 ± 0.59
K_m (μM)	4.3 ± 1.5	1.0 ± 0.25
k_{cat}/K_m ($M^{-1}s^{-1}$)	49.5 ± 17.2 10^6	9.9 ± 1.9 10^6

the MtTFE-αH462A variant and in the KAT active site it is the MtTFE-βC92A variant. Also, the double and triple mutations were made. These variants were purified using the same protocols as for the wild type MtTFE and have similar stability as wild type (Table S2), as measured from the nanoDSF melting curves. The specific activities of these variants for *2E*-decenoyl-CoA are also listed in Table S2. Residual ECH and HAD activities could not be detected when assaying for the activity of the point mutated active site. The ECH specific activities of the HAD and KAT active site point mutated variants are similar to wild type, whereas the HAD specific activity of the MtTFE-βC92A variant is lower.

The crystallographic binding studies with these point mutation variants were done with short chain (*2E*-butenoyl-CoA), medium chain (*2E*-decenoyl-CoA) and long chain (*2E*-hexadecenoyl-CoA) substrates, using the cocrystallization protocol. The point mutated variants of MtTFE crystallize preferably in the space group $C222_1$ with one αβ-dimer in the asymmetric unit (Table S3) with the twofold axis of the thiolase dimer coinciding with a crystallographic twofold axis, in contrast to wild type MtTFE which preferably crystallizes in the space group C2 with the $\alpha_2\beta_2$-tetramer in the asymmetric unit (Venkatesan and Wierenga, 2013). In the C2 crystal form, the crystal packing is very loose ($V_m = 3.9$ Å^3/Da) whereas in the $C222_1$ crystal form the packing is much tighter ($V_m = 2.9$ Å^3/Da). The obtained structures of the point mutated variants (Table S3 and Table S4) show minimal structural changes in each of the catalytic sites. However, the C-domain is disordered to various extents in several of these structures. Analysis of the electron density maps show that the ECH and HAD active sites are unliganded in each of these structures,

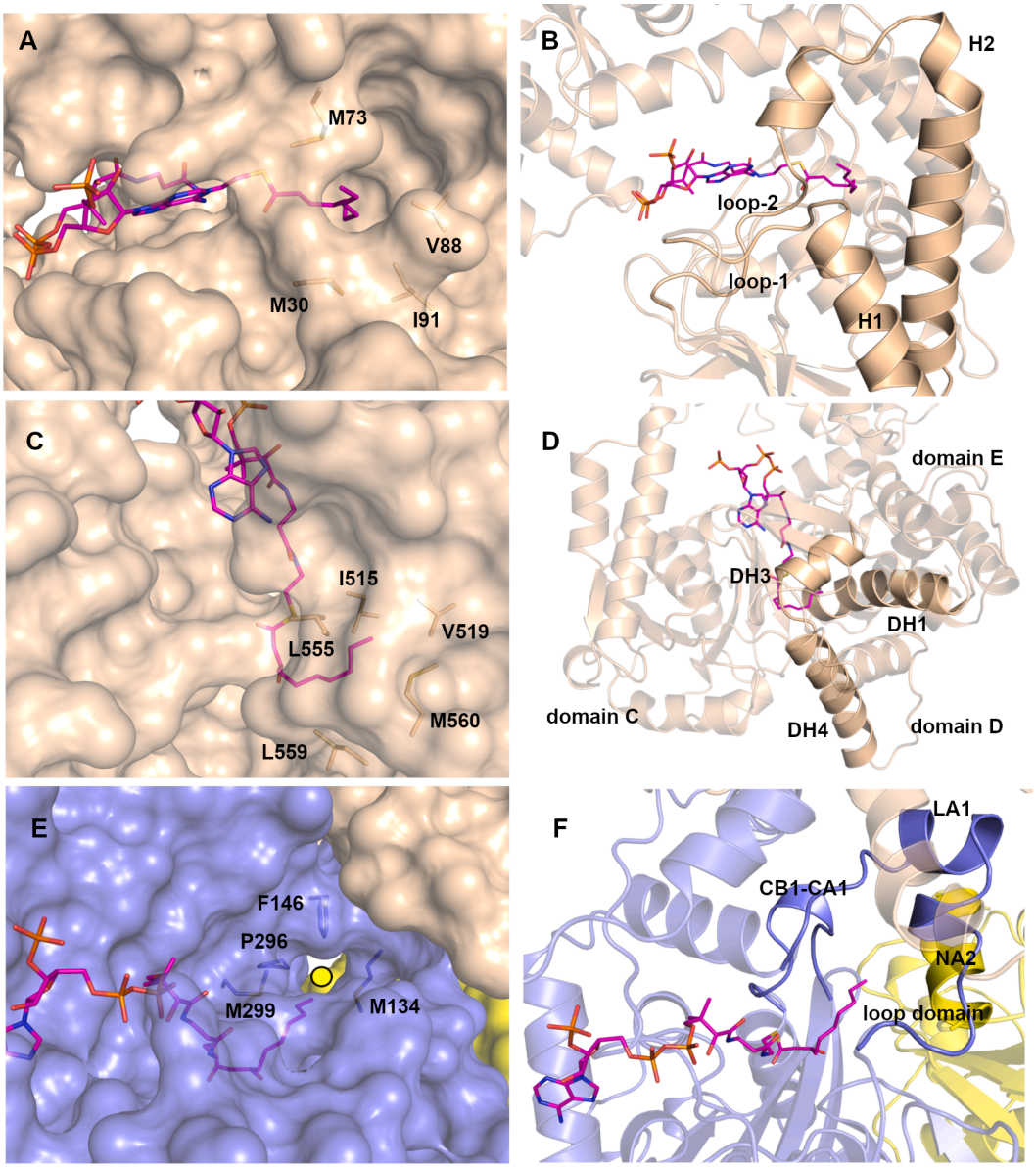

Fig. 2. The acyl binding pockets of the ECH, HAD and KAT active sites. (A, B) The ECH active site. (A) In the ECH active site the acyl tail of the modelled substrate fits in a tunnel that connects the catalytic site to the bulk solvent. (B) This tunnel is shaped by loop-1 and helix H1 (αMet30) as well as loop-2 (αMet73) and helix H2 (αVal88, αIle91 of the A-domain. (C, D) The HAD active site. (C) The acyl tail binding groove of the HAD active site is shaped by residues of the helices DH1 (αIle515, αVal519), DH3 (αLeu555) and DH4 (αLeu559, αMet560). (D) The mode of binding of the modelled substrate to the HAD part of the α-chain with respect to the C-domain. The pantetheine moiety binds in a tunnel shaped by the C and D/E-domains. (E, F) The KAT active site. (E) The acyl tail binding tunnel of the KAT active site. The catalytic site is shaped by the purple thiolase subunit. The tunnel that interacts with the acyl tail of the modelled substrate, is between the N-terminal region of the loop domain (purple, βMet134, βPhe146), the CB1-CA1 loop (βPro296, βMet299) and the NA2 helix of the other thiolase subunit (yellow, "o" labels βLeu75). The nearby H9A helix of the α-chain is in wheat color. (F) The N-terminal region of the loop domain and the CB1-CA1 loop are highlighted in purple and helix NA2 of the other thiolase subunit is in yellow.

whereas in some of the KAT active sites a CoA molecule is present (Table S4), despite the presence of enoyl-CoA substrate in the cocrystallization mother liquor. CoA can be present in the mother liquor as a contaminant but more likely it is present because of hydrolysis of the enoyl-CoA substrate molecules. The crystals were harvested about two months after the drops had been setup, probably allowing for the enoyl-CoA substrate to be hydrolyzed.

Subsequently, modelling studies, as described in the Material and methods section, have been carried out, to study the possible mode of binding of the acyl tail in each of the three active sites. In the ECH active site, the CoA part of the modelled substrate is bound as known from the experimental structures and its thioester oxygen atom binds in the oxyanion hole and the tail docks nicely in a tunnel shaped by loop-1/helix H1 and loop-2/helix H2 (Fig. 2A, B). Only a few side chain rearrangements of residues lining this tunnel are required to allow for this mode of binding. In the HAD active site the interactions of the thioester oxygen and the 3-keto oxygen atoms are preserved in the docked conformation of the substrate and its tail fits in a groove between helices DH1 (αIle515, αVal519), DH3 (αLeu555) and DH4 (αLeu559, αMet560) of domain D (Fig. 2C, D). In the KAT active site, the interactions of the thioester oxygen and the 3-keto oxygen of the modelled substrate are also preserved and the tail points into a tunnel that is also shaped by residues of the other thiolase subunit. This tunnel is lined by hydrophobic residues such as βMet134, βVal144, βPhe146 (of the N-terminal part of the loop domain), βLeu75 (of NA2 of the neighboring subunit), and βPro296, βMet299 (of the CB1-CA1 loop) (Fig. 2E, F). Despite the different geometries of the three active sites, it is clear that each active site can accommodate the extended acyl tails, which is in good agreement with the enzyme kinetic data.

3.3. Crystallographic studies with wild type MtTFE reveal three additional CoA binding sites

For the structural studies with wild type MtTFE five new crystal structures have been refined and analyzed (Table 1, Table S4). These include three new structures of MtTFE referred to as MtTFE-CoA-A, MtTFE-CoA-B and MtTFE-CoA-C, in which new CoA binding sites different from the ECH, HAD and KAT active site CoA binding sites have been identified. Each of these are described in the subsequent paragraphs of this section. Another MtTFE structure complexed with NAD^+, has also been determined (MtTFE(NAD^+). These four MtTFE structures are from crystals which were crystallized in space group C2 with one $\alpha_2\beta_2$-tetramer per asymmetric unit as observed before (Venkatesan and Wierenga, 2013). Interestingly, when using another crystallization protocol as described in the Material and methods section, the wild type MtTFE crystallized in the space group $C222_1$, the same crystal form in which the point mutated variants preferably crystallized with one αβ-dimer in the asymmetric unit. This structure is referred to as the MtTFE (αβ-unliganded) crystal form. In addition, the structure of the MtTFE-alfa-chain only, complexed with NAD^+ is also reported and referred to as the MtTFE-alfa(NAD^+) structure.

The CoAs bound in the three active sites are referred as CoA(ECH), CoA(HAD) and CoA(KAT), respectively. In the cocrystallization experiment of wild type MtTFE in the presence of CoA, a CoA binding site between the HAD part and the thiolase subunit has been identified in the MtTFE-CoA-A structure. This site is referred to as the CoA(HAD/KAT) site (Fig. S6, Table S5). This site is not near crystal contacts and CoA is bound at this site in both copies of the asymmetric unit. The key interaction at the CoA(HAD/KAT) site is the stacking interaction between the adenine ring of CoA and the side chains of αGln629 (of the HAD part) and βTrp244 of helix LA5 of the thiolase subunit. There is also a salt bridge interaction between the pyrophosphate moiety of CoA and the βLys336 side chain of the thiolase subunit. The mode of binding of the adenine moiety is very well defined by the electron density map (Fig. S6), whereas the mode of binding of the pantetheine part is less well defined. This is also seen in the other two structures, MtTFE-CoA-B and MtTFE-CoA-C where the pantetheine part of CoA that is bound at this site is not built (Table S5). The distance of the adenine ring to the amide moiety of αGln629 and the indole ring of βTrp244 is 3.4 Å and 3.6 Å, respectively. This site is not close to any active site (Fig. 3) and it is not clear if this is a functional binding site. In any case the αGln629 and the βTrp244 residues are not conserved in TFE sequences of closely related species (Venkatesan and Wierenga, 2013).

Subsequently, a soaking experiment with a higher CoA concentration (10 mM CoA) was carried out and this structure is referred to as the MtTFE-CoA-B structure. In this structure full occupancy of bound CoA (ECH) and CoA(KAT) is observed. Apart from the CoA(HAD/KAT) binding sites, another CoA binding site, referred to as the CoA(ECH2) site (Fig. S7), was also identified in this structure (Table S5). The CoA (ECH2) binding site is not near crystal contacts and is present in both copies of the asymmetric unit. It is shaped by the helices H5 and H6 of the A-domain of the α-chain, forming a hydrophobic pocket for interactions with the sulfur and adenine moieties of the bound CoA. The interactions between CoA and protein are mostly van der Waals interactions between atoms of CoA and protein side chains, but include also hydrogen bonds and a salt bridge (between the side chain of αLys180, and the 3′-phosphate moiety). The pyrophosphate moiety points into bulk solvent. The distances from the sulfur of the CoA(ECH2) to the sulfurs of CoA(HAD) (as obtained from the superimposed complexed RnMFE1 structure) and CoA(KAT) are about 35 Å. The location of the CoA(ECH2) site with respect to the HAD and KAT active sites is visualized in Fig. 3.

In a soaking experiment with 2.7 mM *2E*-butenoyl-CoA the electron density map of the MtTFE-CoA-C structure shows the mode of binding of CoA in both ECH active sites as well as in both KAT active sites (Table S5). In addition, another ligand binding site of the α-chain is observed (Fig. S8). This site is not near crystal contacts and is present in both copies of the asymmetric unit. The extended shape of the electron density map suggests that the ligand is a bound CoA molecule. Since the 3′-phosphate-adenosine part is in weak density, this part has been included in the modelled CoA with lower occupancy. This site is referred to as the CoA(ECH/HAD) binding site. This binding site is an extended hydrophobic groove, shaped by the α-chain only, in particular by helices H5, H9B and H9A of domain A, and by the linker helix (domain B), together with the EH2 helix and the subsequent support loop of the HAD part (domain E). The interactions between the CoA and the protein consist of van der Waals interactions and hydrogen bonds. There are no salt bridge interactions. This site is located between the CoA(ECH2) site and the HAD active site (Fig. 3), and it is near the ECH pantetheine binding pocket (the pantetheine moieties of CoA(ECH/HAD) and CoA (ECH) are approximately 15 Å apart). The distances from the CoA sulfur of CoA(ECH/HAD) to the sulfurs of CoA(HAD) and CoA(KAT) are approximately 30 Å.

As shown in Fig. 3, the CoA(ECH2) and CoA(ECH/HAD) sites are on the same side of the complex as the HAD and KAT active sites. The nearby ECH active site is shaped by the same α-chain that forms the HAD active site and the surface path between the ECH and HAD active sites is a tunnel shaped connection (well visible in the stereo figure, Fig. S9), similarly as seen in RnMFE1 (Kasaragod et al., 2010). For RnMFE1 it has been speculated that this tunnel could be of functional relevance for its substrate channeling mechanism. The surface paths from the CoA (ECH2) and CoA(ECH/HAD) binding sites to the ECH, HAD and KAT active sites do not have negatively charged side chains and they are lined by positively charged side chains (Figs. 3, S9). Further MtTFE reaction mechanism studies are required to establish if the CoA(ECH2) and CoA (ECH/HAD) binding sites are functionally relevant as transient binding sites in substrate channeling.

3.4. The ECH active site switches into an active conformation on CoA binding

The ECH active site catalyses the stereo specific hydration of *2E*-

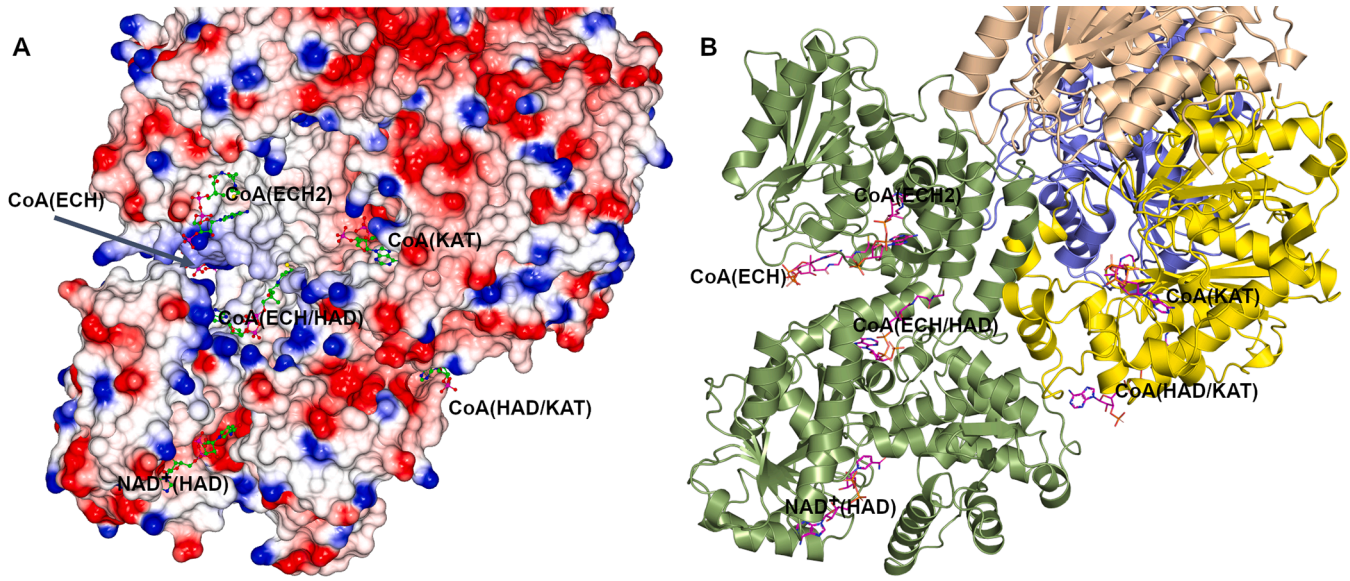

Fig. 3. The relative positions of the ECH, HAD and KAT active sites as well as the CoA (HAD/KAT), CoA(ECH2) and CoA(ECH/HAD) binding sites. (A) Electrostatic molecular surface of MtTFE. The color codes red, white and blue correspond to negative, neutral and positive electrostatic potential, respectively. The stereo representation of this figure is shown in Fig. S9. (B) MtTFE in ribbon presentation. Same view as in Fig. 3A. The colors green and wheat correspond to the A and B α-chains and the colors yellow and sky-blue correspond to the C and D β-chains. The green colored α-chain A is shown in similar view as in Fig. S2A. The ligands are shown in sticks. The ligand labels of CoA(ECH) (mostly invisible in Fig. 3A), NAD^+(HAD) and CoA(KAT) identify the locations of the ECH, HAD and KAT active sites, whereas CoA(HAD/KAT) (mostly invisible in Fig. 3A, B), CoA(ECH2) and CoA(ECH/HAD) identify the CoA(HAD/KAT), CoA(ECH2) and CoA(ECH/HAD) sites, respectively.

enoyl-CoA into 3S-hydroxyacyl-CoA. The substrate is activated for hydration because the thioester oxygen is bound in an oxyanion hole formed by loop-2 and loop-3. Comparison of the liganded and unliganded ECH active sites of MtTFE shows an interesting difference in the loop-2 conformation, residues αAla66-αGly67-αGly68 (Fig. 4). In the liganded structure the αAla66-αGly67 peptide plane is rotated such that O(αAla66) interacts with the N4P atom of the CoA moiety, as well as with the N6 atom of the adenine moiety of CoA (Fig. 4). In this conformational switch also the conformation of the αGly67-αGly68 peptide plane changes, such that N(αGly68) adopts its competent conformation in the liganded complex and can function as the hydrogen bond donor of the oxyanion hole. The other hydrogen bond donor of this oxyanion hole comes from the N-terminal peptide of helix H3, N (αGly116), as shown in the schematic Fig. 4C. The distance between N (αGly116) and N(αGly68) increases in the incompetent conformation (Table 4). A comparison with active sites of other hydratases (having also the crotonase fold) shows that in the CoA liganded structures the conformation of loop-2 is conserved, but in the unliganded structures the competent loop-2 conformation is not always preserved. In the incompetent conformation the distance between the two hydrogen bond donors of the oxyanion hole is increased or loop-2 is disordered (Table 4). The binding of the adenine moiety induces the competent conformation of loop-2 by the formation of the hydrogen bond between its N6 atom and O(αAla66), as well as with O(αAla68). This stabilization of loop-2 in its competent conformation on binding of the adenine moiety of CoA, rationalizes how the presence of CoA can activate the hydratase activity of *2E*-enoyl-CoA hydratase (crotonase) for the truncated *2E*-butenoyl-pantetheine substrate (Bahnson and Anderson, 1989; Waterson et al., 1972). This truncated substrate (not having the 3′-phosphate-ADP moiety) is a poor substrate, but in the presence of CoA the *2E*-butenoyl-pantetheine is a much better substrate of this enzyme than in the absence of CoA (Waterson et al., 1972).

3.5. *The conformational flexibility of the C-domain of the HAD active site*

The C-domains of the two α-chains are located at the far ends of the elongated tetramer (Fig. 1) and they have only a few interactions with the other parts of the MtTFE tetramer. The mode of binding of NAD^+ to the C-domain is captured in the MtTFE (NAD^+) structure (Fig. S10). The mode of binding is the same as reported for HsHAD and RnMFE1. In the MtTFE(NAD^+) structure NAD^+ is only interacting with the C-domain (Fig. 5A, B). There is considerable variation in the positioning of the C-domain with respect to the D/E-domains of the HAD-part in the various MtTFE structures, as shown in Fig. 5A. The comparison of the structures of the α subunits of MtTFE(NAD^+) and MtTFE(αβ-unliganded) also shows the rigid positioning of the A-domain with respect to the HAD part, being fixed in its position by the interactions with the thiolase β-chain (Fig. 1). This is in contrast to the conformational flexibility properties of the monomeric RnMFE1, in which case the A-domain positioning with respect to the D/E-domains of the HAD-part is variable (Table S6) (Sridhar et al., 2020). The functional relevance of the conformational flexibility of the C-domain of the monofunctional HsHAD has been described from structural studies where it has been proposed that only the fully closed HsHAD structure is competent for catalyzing the oxidation of the substrate. In this fully closed HsHAD structure the reference distance (Fig. 5B, Table S6) between domain C and domain D is 8.3 Å. This distance varies between 15.1 Å and 12.1 Å for the available wild type MtTFE structures (Table S6). Therefore, despite the variability of the positioning of the C-domain in MtTFE, each of these structures can be classified as having the "open" HAD conformation.

3.6. *TFE thiolase dimers are stable and catalytically competent only in the presence of the MtTFE α-chain*

The MtTFE thiolase and the *Mtb* FadA5 thiolase belong to the same thiolase sequence subfamily (Anbazhagan et al., 2014). A key difference with the FadA5 thiolase active site is the presence of helix H9A of the nearby α-chain in MtTFE (Fig. 6). This helix shields the MtTFE KAT active site from bulk solvent, making the active site more narrow. Therefore, the mode of binding of the cholesterol derivative as present in the FadA5 structure is unfavorable in MtTFE, as its keto-oxygen atom is

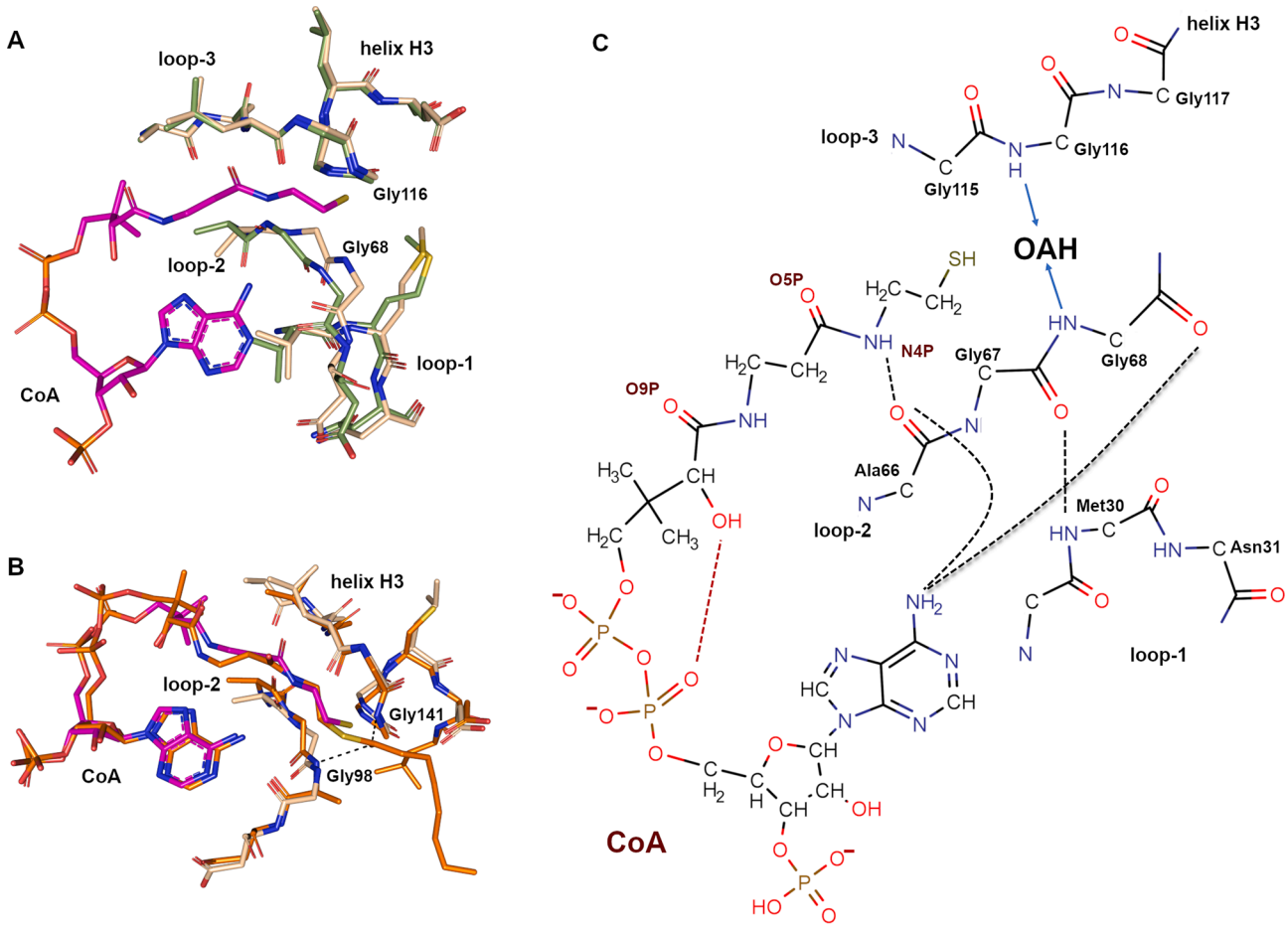

Fig. 4. The ECH active site. (A) Superposition of the unliganded active site of MtTFE(αβ-unliganded) (green) and the active site, complexed with CoA, of the MtTFE-CoA-B structure (wheat), in which loop-2 has adopted its competent conformation. The CoA(ECH) is shown in magenta sticks. (B) Superposition of the liganded ECH active sites of the MtTFE-CoA-B structure (wheat, with CoA, magenta sticks) and of rat, mitochondrial *2E*-enoyl-CoA hydratase (orange, PDB ID 2DUB), liganded with octanoyl-CoA (orange sticks). The hydrogen bond interactions of the thioester oxygen of octanoyl-CoA and its oxyanion hole hydrogen bond donors (NH(Ala98), NH (Gly141), Table 4 are highlighted by dotted lines. (C) Schematic figure of the mode of interaction of CoA(ECH) with loop-1, loop-2 and loop-3 of the MtTFE ECH active site, complexed with CoA. The oxyanion hole is formed by NH(αGly68) and NH(αGly116).

Table 4
Loop-2 sequences of hydratases that have the crotonase fold.

PDB ID (chain)	Ligand	loop-2 conformation	N-N distance# (Å)	loop-2 sequence[&,*]	loop-3 sequence (just before helix H3) [&]	Comment
2DUB (A)	octanoyl-CoA	competent	3.9	A-A-G-A98-D-VK*	L-G-G141-G-C-E	rat mitochondrial hydratase (Engel et al., 1998)
7O4S (A)	CoA	competent	4.4	F-A-G-G68-D-VK*	L-G-G116-G-L-E	MtTFE-CoA-B
7O4Q (A)	no ligand	incompetent	5.0	F-A-G-G68-D-VK	L-G-G116-G-L-E	MtTFE (αβ-unliganded)
1WDK (A)	detergent	incompetent	6.6	I-V-G-A69-D-IT	L-G-G117-G-L-E	PfTFE (Ishikawa et al., 2004)
6WYI (A)	no ligand	disordered	NA	C-A-G-M66-D-LK	I-A-G114-G-T-E	*Mtb* ECH-A19 (Bonds et al., 2020)
4F47 (A)	no ligand	disordered	NA	C-A-G-M77-D-LK	I-A-G125-G-T-E	*Mtb* ECH-A19 (Baugh et al., 2015)
3PZK (B)	no ligand	incompetent	5.0	A-A-G-A65-D-IK*	L-G-G108-G-C-E	*Mtb* ECH-A8 (unpublished)
3Q0G (F)	*2E*-butenoyl-CoA	competent	4.0	A-A-G-A65-D-IK*	L-G-G108-G-C-E	*Mtb* ECH-A8 (unpublished)

[#]The N-N distance is the distance between the N atoms of the peptide NH groups, that are the two hydrogen bond donors of the crotonase fold oxyanion hole. NA: not applicable.
[&] The peptide NH group of the numbered residue is the hydrogen bond donor of the oxyanion hole.
[*] The side chain of the conserved aspartate of loop-2 has a hydrogen bond interaction with the main chain NH group of a residue of the subsequent helical turn. The latter residue is highlighted with*.

pointing into a hydrophobic pocket formed by helix H9A residues (Fig. 6). This suggests that the substrate specificities of these two thiolases are different, in agreement with the kinetic data.

Another key difference between the MtTFE KAT active site and FadA5 and other thiolase active sites concerns the dynamic properties of the CB4-CB5 catalytic loop, as well as the nearby CB1-CA1 loop. These loops, which are close to the thiolase dimer interface (Fig. 6), are well defined in the previously reported thiolase structures, but have high B-

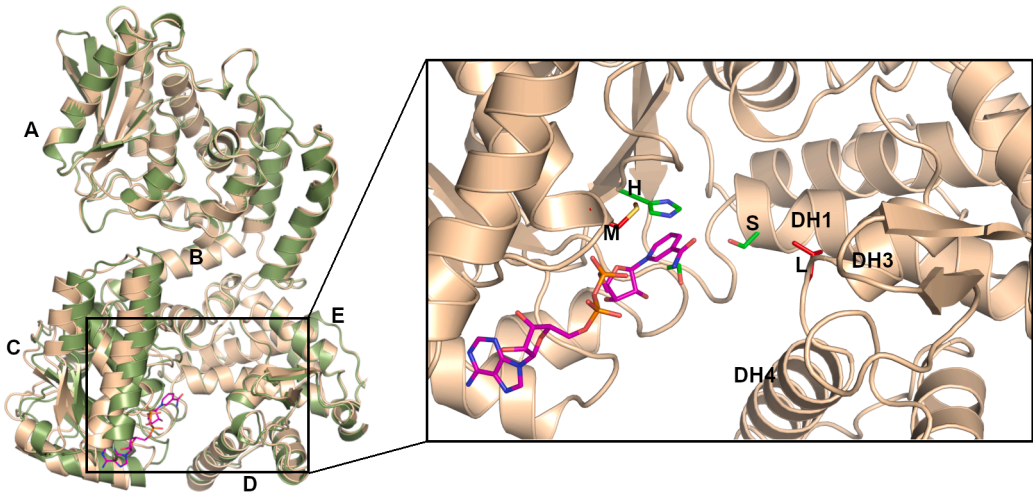

Fig. 5. The conformational flexibility of the C-domain of the HAD active site. The structures of the α-chains of MtTFE (NAD^+) (green) and MtTFE (αβ-unliganded) (wheat) are compared. The domains of the α-chain are identified by the labels A, B, C, D and E. The zoomed in view of the HAD part of MtTFE(NAD^+) shows the mode of binding of NAD^+ (magenta sticks) to domain C. The side chains αHis462 (H) and αSer512 (S) identify the catalytic site. The side chains of αMet335 (M) and αLeu555 (L) are also shown. The distance between the Cα-atoms of the latter two residues is used to define the open/closed state of domain C with respect to domain D, as discussed in the text and in Table S6.

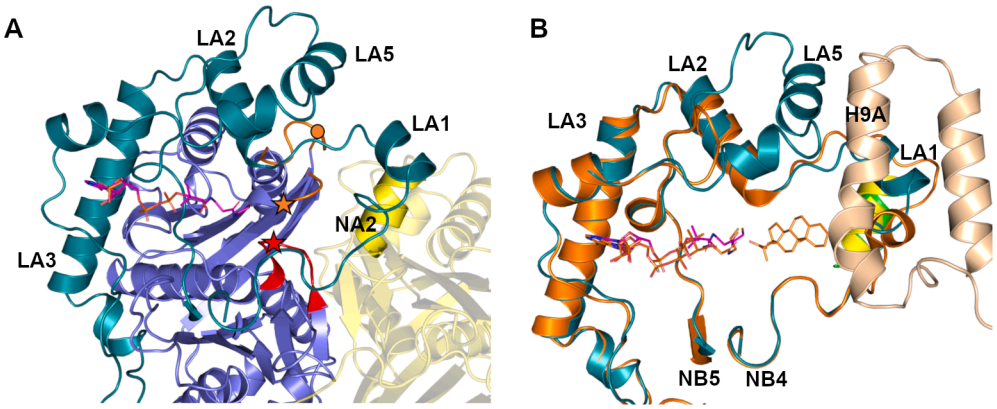

Fig. 6. The MtTFE KAT active site. (A) The KAT active site is near the β_2/thiolase dimer interface. The NA2 helix of the neighboring thiolase subunit is in yellow. The loop domain (turquoise blue) is running over the two core domains, encircling the bound CoA, shown in sticks. βCys92 (NB3-NA3 loop, red) is highlighted with a red star. βCys389 (CB4-CB5 loop, orange) is identified with an orange star. The CB1-CA1 loop is also shown in orange, highlighted by an orange dot. (B) The comparison of the active site structures of the MtTFE and FadA5 (PDB ID 4UBT) thiolases. The H9A helix of the neighboring MtTFE α-chain (wheat) and the NA2 helix of the neighboring β/thiolase subunit (yellow) are also shown. CoA (bound to MtTFE thiolase) is shown in magenta sticks. The FadA5 loop domain is shown in orange and the CoA and steroid molecule bound in the FadA5 active site are shown in orange sticks.

factors in the MtTFE structure (Fig. S11). The crystal structure of only the thiolase dimer (not assembled with the two α-chains) of MtTFE is not available, however the structure of an orthologue, the thiolase dimer of *Mycobacterium avium*, has been reported (PDB ID 3SVK). On close inspection of this structure, it can be noted that the 3SVK thiolase dimer is not a competent dimer, as the catalytic NB3-NA3 loop, the N-terminal region (LA1) and the C-terminal region (LA5) of the loop domain, the CB4-CB5 loop as well as the CB1-CA1 loops are disordered. There is 90% sequence identity between the MtTFE thiolase and the *M. avium* thiolase and the genes encoding for these thiolases are in the same location in the genome. Apparently, the competent MtTFE thiolase dimer interface is formed only in the presence of the alpha subunit.

The stability of a protein–protein interface can be predicted from its P-value of the PISA analysis (Krissinel and Henrick, 2007). The P-value depends on the hydrophobicity of the interface surface area. A protein–protein interface with a lower P-value is predicted to have a higher stability. The PISA analysis of the MtTFE thiolase dimer interface shows that the dimer interface area is rather extensive (2400 $Å^2$). However, the P-value for the dimer interface is 0.39. Similarly, the P-value for the PfTFE thiolase dimer is 0.54 suggesting that these TFE thiolase dimers are weak dimers. A similar PISA analysis of the *M. avium* thiolase dimer, as captured in the PDB ID 3SVK structure, calculates a P-value of 0.68, suggesting that this dimer has even lower stability than the competent dimer, as present in the MtTFE thiolase dimer. In addition, studies on EcTFE thiolase, which is a close homologue of the PfTFE thiolase, have shown that this thiolase is inactive and exists in a monomer–dimer mixture in the absence of the EcTFE-alpha subunit (Sah-Teli et al., 2020). Also, for some monofunctional thiolases, including FadA5, it has been reported that there exist two forms, a competent and an incompetent form. The equilibrium between the competent and the incompetent assemblies (in which catalytic cysteines are oxidized and occur as SS bridges) is influenced by the redox potential of the buffer solution (Kim et al., 2015; Lu et al., 2017; Pye et al., 2010). There are also other examples of thiolases, for example, zebrafish SCP2-thiolase, which have been shown to exist in solution in a dimer-monomer equilibrium (Kiema et al., 2019). In summary, these comparisons suggest that several thiolase dimers are weak dimers and in case of these TFE thiolases, a stable and competent TFE thiolase dimer is formed only in the presence of the alpha subunit.

3.7. The helix H9A of the α-chain is stabilized by the thiolase subunit in the MtTFE assembly

MtTFE-alpha (not complexed with the thiolase dimer) is less stable than the wild type MtTFE complex (Table S2), but it has catalytic ECH and HAD activity, although in particular the HAD specific activity is less than the wild type MtTFE (Table S2). It is a monomeric protein (Venkatesan and Wierenga, 2013). In the crystal structure of MtTFE-alpha (Table 1), obtained by cocrystallization in the presence of 2 mM NAD^+, there are two molecules per asymmetric unit. In both molecules

the structure of the C-domain is in an open conformation, like observed in the tetrameric $\alpha_2\beta_2$ MtTFE assembly, and the mode of binding of NAD^+ is identical as observed in the MtTFE (NAD^+) structure. However, the crystal structure of the monomeric MtTFE-alpha shows an important structural difference with the assembled MtTFE α-chain. The residues αPro236 till αLeu253, that adopt a kinked helical conformation (helix H9A, Fig. 1, Fig. S2A) in the assembled MtTFE α-chain, have adopted a partially disordered, random coil conformation in the monomeric MtTFE-alpha structure (Fig. 7). In the TFE assembly the H9A helix interacts with the thiolase subunit, near its active site, with two regions. It concerns the N-terminal region of the loop domain, residues αGly128 till αPhe146 (including helix LA1), and the C-terminal region of the loop domain, residues αGly227 till αLeu239 (helix LA5). Helix LA5 is a unique structural feature of the MtTFE thiolase subfamily (Venkatesan and Wierenga, 2013). These two regions (LA1 and LA5) are disordered in the structure of the non-competent dimer of *M. avium* thiolase, as described in the previous section. On assembly of the monomeric MtTFE α-chain with the thiolase subunit, the disordered H9A region of the α-chain and the LA1 and LA5 disordered regions of the thiolase loop domain adopt their ordered structures as seen in the MtTFE tetramer (Fig. 1). Such disorder-to-order transition of regions involved in protein–protein interfaces has been seen also for other assemblies (Perkins et al., 2010). In the tetrameric MtTFE complex, each α-chain interacts with both thiolase subunits with interface areas of about 1100 $Å^2$ (at the interface formed by LA1 and LA5) and 500 $Å^2$ (formed by other loops), respectively, thereby stabilizing the β_2-dimer.

4. Concluding remarks

The structural and enzymological data suggest that unbranched medium and long chain enoyl-CoA molecules are the preferred substrates of MtTFE and the analysis of the MtTFE structures shows that each of the ECH, HAD and KAT active sites has a binding pocket that can accommodate the extended carbon tail of medium and long chain acyl-CoA substrate molecules. Additional CoA binding sites have been identified. Two of these sites are located between the ECH, HAD and KAT active sites (Fig. 3). Further studies are required to understand the functional relevance of these sites. The comparison of the structures of the unliganded and liganded MtTFE ECH active sites shows that loop-2 adopts an incompetent conformation in the unliganded complex (Fig. 4). In each of the MtTFE structures the C-domain of the HAD part has adopted a slightly different but always open conformation with respect to the D/E-domains (Fig. 5), for which it is predicted that the CoA affinity of the HAD active site is low and the currently available MtTFE crystal structures have not captured the mode of binding of CoA in the HAD active site. However, in all MtTFE structures the A-domain is fixed with respect to the D/E-domains by the interactions of these domains with the β/thiolase subunit (Fig. 1). The structural analysis of the MtTFE thiolase subunit shows conformational flexibility of the catalytic CB4-CB5 loop as well as the adjacent CB1-CA1 loop. These loops are near the β_2-dimer interface (Fig. 6) and this conformational flexibility is correlated with the prediction that this β_2/thiolase dimer is a weak dimer, being stabilized by its interactions with the α-chains. In summary, the structural analysis presented here highlights how the binding of substrates and the assembly of the α-and β-subunits into the $\alpha_2\beta_2$ tetramer is important to form a fully functional MtTFE complex.

CRediT authorship contribution statement

Subhadra Dalwani: Conceptualization, Methodology, Investigation, Formal analysis, Data curation, Validation, Software, Writing - original draft. **Outi Lampela:** Methodology. **Pierre Leprovost:** Methodology. **Werner Schmitz:** Methodology. **André H. Juffer:** Methodology. **Rik K. Wierenga:** Conceptualization, Formal analysis, Supervision, Writing - original draft, Writing - review & editing. **Rajaram Venkatesan:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Project administration, Supervision, Writing - original draft, Writing - review & editing.

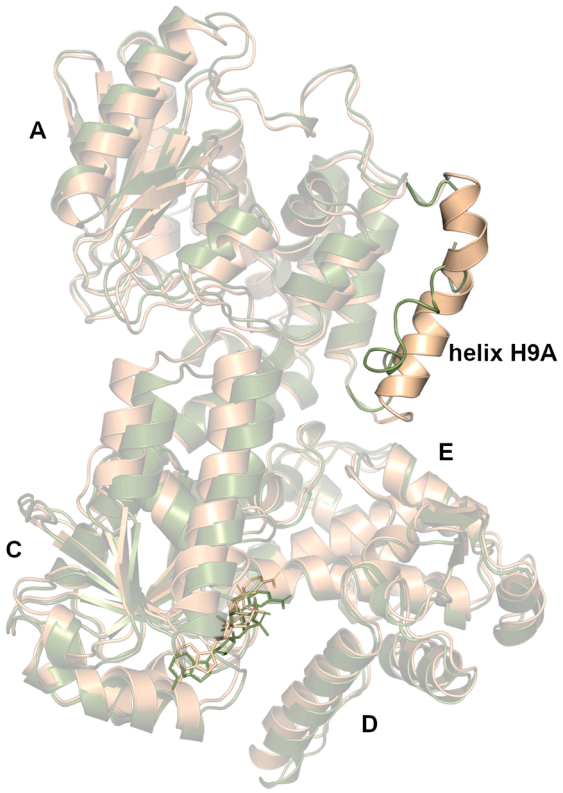

Fig. 7. The disordered H9A helix. Superposition of the MtTFE-alfa structure (molecule A, green) and the structure of the α-chain of MtTFE(NAD^+) (wheat). Helix H9A is disordered in the MtTFE-alfa structure. The mode of binding of NAD^+ to domain C of both structures is shown as sticks.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Nicole Sampson and Tianao Yuan of the Stonybrook University (USA) for providing 3,22-dioxo-chol-4-ene-24-oyl-CoA. We thank Gabriele Cordara for helping with data collection and data processing. The use of the facilities and expertise of the Biocenter Oulu Structural Biology core facility, a member of Biocenter Finland, Instruct-ERIC Centre Finland and FINStruct, is gratefully acknowledged. The support of the Biocomputing and the Protein Biophysical Biocenter Oulu core facilities, which both are part of Biocenter Finland, is also gratefully acknowledged. This research has been supported by grants of the Academy of Finland (293369, 289024 and 319194) to RV and SD. The synchrotron data sets have been collected at BESSY (Berlin, Germany), ESRF (Grenoble, France), DLS (Oxford, UK) and EMBL (using the PETRA III storage ring of DESY, Hamburg, Germany). We are most grateful for the skillful support provided by the beamline scientists for the data collection at the synchrotrons. The authors also wish to acknowledge the CSC – IT Center for Science, Espoo, Finland, for computational resources.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.

org/10.1016/j.jsb.2021.107776.

References

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., 2012. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. Sect. D Biol. Crystallogr. 68, 352–367. https://doi.org/10.1107/S0907444912001308.

Anbazhagan, P., Harijan, R.K., Kiema, T.R., Janardan, N., Murthy, M.R., Michels, P.A., Juffer, A.H., Wierenga, R.K., 2014. Phylogenetic relationships and classification of thiolases and thiolase-like proteins of Mycobacterium tuberculosis and Mycobacterium smegmatis. Tuberculosis (Edinburgh, Scotland) 94, 405–412. https://doi.org/10.1016/j.tube.2014.03.003.

Bahnson, B.J., Anderson, V.E., 1989. Isotope effects on the crotonase reaction. Biochemistry 28, 4173–4181. https://doi.org/10.1021/bi00436a008.

Barycki, J.J., O'Brien, L.K., Strauss, A.W., Banaszak, L.J., 2001. Glutamate 170 of human l-3-hydroxyacyl-CoA dehydrogenase is required for proper orientation of the catalytic histidine and structural integrity of the enzyme. J. Biol. Chem. 276, 36718–36726. https://doi.org/10.1074/jbc.M104839200.

Barycki, J.J., O'Brien, L.K., Strauss, A.W., Banaszak, L.J., 2000. Sequestration of the active site by interdomain shifting. Crystallographic and spectroscopic evidence for distinct conformations of L-3-hydroxyacyl-CoA dehydrogenase. J. Biol. Chem. 275, 27186–27196. https://doi.org/10.1074/jbc.M004669200.

Baugh, L., Phan, I., Begley, D.W., Clifton, M.C., Armour, B., Dranow, D.M., Taylor, B.M., Muruthi, M.M., Abendroth, J., Fairman, J.W., Fox, D., Dieterich, S.H., Staker, B.L., Gardberg, A.S., Choi, R., Hewitt, S.N., Napuli, A.J., Myers, J., Barrett, L.K., Zhang, Y., Ferrell, M., Mundt, E., Thompkins, K., Tran, N., Lyons-Abbott, S., Abramov, A., Sekar, A., Serbzhinskiy, D., Lorimer, D., Buchko, G.W., Stacy, R., Stewart, L.J., Edwards, T.E., Van Voorhis, W.C., Myler, P.J., 2015. Increasing the structural coverage of tuberculosis drug targets. Tuberculosis (Edinb) 95, 142–148. https://doi.org/10.1016/j.tube.2014.12.003.

Bayly, C.I., Cieplak, P., Cornell, W., Kollman, P.A., 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. J. Phys. Chem. 97, 10269–10280. https://doi.org/10.1021/j100142a004.

Bonds, A.C., Yuan, T., Werman, J.M., Jang, J., Lu, R., Nesbitt, N.M., Garcia-Diaz, M., Sampson, N.S., 2020. Post-translational succinylation of mycobacterium tuberculosis Enoyl-CoA Hydratase EchA19 slows catalytic hydration of cholesterol catabolite 3-Oxo-chol-4,22-diene-24-oyl-CoA. ACS Infect. Dis. 6, 2214–2224. https://doi.org/10.1021/acsinfecdis.0c00329.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 12–21. https://doi.org/10.1107/S0907444909042073.

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S., Barrell, B.G., 1998. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature 393, 537–544. https://doi.org/10.1038/31159.

Cox, J.A.G., Taylor, R.C., Brown, A.K., Attoe, S., Besra, G.S., Futterer, K., 2019. Crystal structure of Mycobacterium tuberculosis FadB2 implicated in mycobacterial beta-oxidation. Acta Crystallogr. Sect. D Struct. Boil. 75, 101–108. https://doi.org/10.1107/S2059798318017242.

Daniel, E., Maksimainen, M.M., Smith, N., Ratas, V., Biterova, E., Murthy, S.N., Rahman, M.T., Kiema, T.-R., Sridhar, S., Cordara, G., Dalwani, S., Venkatesan, R., Prilusky, J., Dym, O., Lehtiö, L., Koski, M.K., Ashton, A.W., Sussman, J.L., Wierenga, R.K., 2021. *IceBear* : an intuitive and versatile web application for research-data tracking from crystallization experiment to PDB deposition. Acta Crystallogr. D Struct. Biol. 77, 151–163. https://doi.org/10.1107/S2059798320015223.

Dupradeau, F.-Y., Cézard, C., Lelong, R., Stanislawiak, E., Pêcher, J., Delepine, J.C., Cieplak, P., 2008. R.E.DD.B.: a database for RESP and ESP atomic charges, and force field libraries. Nucleic Acids Res. 36, D360–367. https://doi.org/10.1093/nar/gkm887.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 486–501. https://doi.org/10.1107/S0907444910007493.

Engel, C.K., Kiema, T.R., Hiltunen, J.K., Wierenga, R.K., 1998. The crystal structure of enoyl-CoA hydratase complexed with octanoyl-CoA reveals the structural adaptations required for binding of a long chain fatty acid-CoA molecule. J. Mol. Biol. 275, 847–859. https://doi.org/10.1006/jmbi.1997.1491.

Evans, P.R., Murshudov, G.N., 2013. How good are my data and what is the resolution? Acta Crystallogr. Sect. D Biol. Crystallogr. 69, 1204–1214. https://doi.org/10.1107/S0907444913000061.

Ferdinandusse, S., Denis, S., Overmars, H., Van Eeckhoudt, L., Van Veldhoven, P.P., Duran, M., Wanders, R.J.A., Baes, M., 2005. Developmental changes of bile acid composition and conjugation in L- and D-bifunctional protein single and double knockout mice. J. Biol. Chem. 280, 18658–18666. https://doi.org/10.1074/jbc.M414311200.

Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Petersson, G.A., Nakatsuji, H., et, al, 2009. Gaussian 09, Revision B.1. Gaussian, Inc., Wallingford CT.

Haapalainen, M. Antti, Meriläinen, Gitte, Wierenga, K. Rik, 2006. The thiolase superfamily: condensing enzymes with diverse reaction specificities. Trends Biochem. Sci. 31 (1), 64–71. https://doi.org/10.1016/j.tibs.2005.11.011.

Imamura, S., Ueda, S., Mizugaki, M., Kawaguchi, A., 1990. Purification of the multienzyme complex for fatty acid oxidation from Pseudomonas fragi and reconstitution of the fatty acid oxidation system. J. Biochem. 107, 184–189. https://doi.org/10.1093/oxfordjournals.jbchem.a123023.

Incardona, M.-F., Bourenkov, G.P., Levik, K., Pieritz, R.A., Popov, A.N., Svensson, O., 2009. *EDNA* : a framework for plugin-based applications applied to X-ray experiment online data analysis. J. Synchrotron Rad. 16, 872–879. https://doi.org/10.1107/S0909049509036681.

Ishikawa, M., Tsuchiya, D., Oyama, T., Tsunaka, Y., Morikawa, K., 2004. Structural basis for channelling mechanism of a fatty acid beta-oxidation multienzyme complex. EMBO J. 23, 2745–2754. https://doi.org/10.1038/sj.emboj.7600298.

Kabsch, W., 2010. Xds. Acta crystallographica.Section D, Biological crystallography 66, 125–132. https://doi.org/10.1107/S0907444909047337.

Kabsch, W., Sander, C., 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22, 2577–2637. https://doi.org/10.1002/bip.360221211.

Kasaragod, P., Schmitz, W., Hiltunen, J.K., Wierenga, R.K., 2013. The isomerase and hydratase reaction mechanism of the crotonase active site of the multifunctional enzyme (type-1), as deduced from structures of complexes with 3S-hydroxy-acyl-CoA. FEBS J. 280, 3160–3175. https://doi.org/10.1111/febs.12150.

Kasaragod, P., Venkatesan, R., Kiema, T.R., Hiltunen, J.K., Wierenga, R.K., 2010. Crystal structure of liganded rat peroxisomal multifunctional enzyme type 1: a flexible molecule with two interconnected active sites. J. Biol. Chem. 285, 24089–24098. https://doi.org/10.1074/jbc.M110.117606.

Kiema, T.R., Thapa, C.J., Laitaoja, M., Schmitz, W., Maksimainen, M.M., Fukao, T., Rouvinen, J., Janis, J., Wierenga, R.K., 2019. The peroxisomal zebrafish SCP2-thiolase (type-1) is a weak transient dimer as revealed by crystal structures and native mass spectrometry. Biochem. J. 476, 307–332. https://doi.org/10.1042/BCJ20180788.

Kim, S., Jang, Y.-S., Ha, S.-C., Ahn, J.-W., Kim, E.-J., Lim, J.H., Cho, C., Ryu, Y.S., Lee, S. K., Lee, S.Y., Kim, K.-J., 2015. Redox-switch regulatory mechanism of thiolase from Clostridium acetobutylicum. Nat. Commun. 6, 8410. https://doi.org/10.1038/ncomms9410.

Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797. https://doi.org/10.1016/j.jmb.2007.05.022.

Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. 60, 2256–2268. https://doi.org/10.1107/S0907444904026460.

Krug, M., Weiss, M.S., Heinemann, U., Mueller, U., 2012. XDSAPP: a graphical user interface for the convenient processing of diffraction data using XDS. J. Appl. Crystallogr. 45, 568–572. https://doi.org/10.1107/S0021889812011715.

Kursula, P., Ojala, J., Lambeir, A.-M., Wierenga, R.K., 2002. The catalytic cycle of biosynthetic thiolase: a conformational journey of an acetyl group through four binding modes and two oxyanion holes. Biochemistry 41, 15543–15556. https://doi.org/10.1021/bi0266232.

Lee, W., VanderVen, B.C., Fahey, R.J., Russell, D.G., 2013. Intracellular mycobacterium tuberculosis exploits host-derived fatty acids to limit metabolic stress. J. Biol. Chem. 288, 6788–6800. https://doi.org/10.1074/jbc.M112.445056.

Liang, K., Li, N., Wang, X., Dai, J., Liu, P., Wang, C., Chen, X.-W., Gao, N., Xiao, J., 2018. Cryo-EM structure of human mitochondrial trifunctional protein. Proc. Natl. Acad. Sci. U.S.A. 115, 7039–7044. https://doi.org/10.1073/pnas.1801252115.

Liebschner, D., Afonine, P.V., Baker, M.L., Bunkoczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., Moriarty, N.W., Oeffner, R.D., Poon, B.K., Prisant, M.G., Read, R.J., Richardson, J.S., Richardson, D.C., Sammito, M.D., Sobolev, O.V., Stockwell, D.H., Terwilliger, T.C., Urzhumtsev, A.G., Videau, L.L., Williams, C.J., Adams, P.D., 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. Sect. D Struct. Biol. 75, 861–877. https://doi.org/10.1107/S2059798319011471.

Lu, R., Schaefer, C.M., Nesbitt, N.M., Kuper, J., Kisker, C., Sampson, N.S., 2017. Catabolism of the cholesterol side chain in mycobacterium tuberculosis is controlled by a redox-sensitive thiol switch. ACS Infect. Dis. 3, 666–675. https://doi.org/10.1021/acsinfecdis.7b00072.

Madeira, F., Park, Y. mi, Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., Lopez, R., 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Research 47, W636–W641. https://doi.org/10.1093/nar/gkz268.

Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., Simmerling, C., 2015. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J. Chem. Theory Comput. 11, 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674. https://doi.org/10.1107/S0021889807021206.

McNicholas, S., Potterton, E., Wilson, K.S., Noble, M.E.M., 2011. Presenting your structures: the CCP4mg molecular-graphics software. Acta Crystallogr. D Biol. Crystallogr. 67, 386–394. https://doi.org/10.1107/S0907444911007281.

Modis, Y., Wierenga, R.K., 2000. Crystallographic analysis of the reaction pathway of Zoogloea ramigera biosynthetic thiolase. J. Mol. Biol. 297, 1171–1182. https://doi.org/10.1006/jmbi.2000.3638.

Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A., 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 355–367. https://doi.org/10.1107/S0907444911001314.

Nada, M.A., Rhead, W.J., Sprecher, H., Schulz, H., Roe, C.R., 1995. Evidence for intermediate channeling in mitochondrial beta-oxidation. J. Biol. Chem. 270, 530–535. https://doi.org/10.1074/jbc.270.2.530.

Nazarova, E.V., Montague, C.R., Huang, L., La, T., Russell, D., VanderVen, B.C., 2019. The genetic requirements of fatty acid import by Mycobacterium tuberculosis within macrophages. Elife 8. https://doi.org/10.7554/eLife.43621.

Perkins, J.R., Diboun, I., Dessailly, B.H., Lees, J.G., Orengo, C., 2010. Transient protein-protein interactions: structural, functional, and network properties. Structure 18, 1233–1243. https://doi.org/10.1016/j.str.2010.08.007.

Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P.R., Jenkins, H.T., Keegan, R., Krissinel, E., Stevenson, K., Lebedev, A., McNicholas, S.J., Nicholls, R.A., Noble, M., Pannu, N.S., Roth, C., Sheldrick, G., Skubak, P., Turkenburg, J., Uski, V., von Delft, F., Waterman, D., Wilson, K., Winn, M., Wojdyr, M., 2018. CCP4i2: the new graphical user interface to the CCP4 program suite. Acta Crystallogr. Sect. D 74, 68–84. https://doi.org/10.1107/S2059798317016035.

Pye, V.E., Christensen, C.E., Dyer, J.H., Arent, S., Henriksen, A., 2010. Peroxisomal plant 3-ketoacyl-CoA thiolase structure and activity are regulated by a sensitive redox switch. J. Biol. Chem. 285, 24078–24088. https://doi.org/10.1074/jbc.M110.106013.

Robert, X., Gouet, P., 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42, W320–W324. https://doi.org/10.1093/nar/gku316.

Sah-Teli, S.K., Hynonen, M.J., Schmitz, W., Geraets, J.A., Seitsonen, J., Pedersen, J.S., Butcher, S.J., Wierenga, R.K., Venkatesan, R., 2019. Complementary substrate specificity and distinct quaternary assembly of the Escherichia coli aerobic and anaerobic beta-oxidation trifunctional enzyme complexes. Biochem. J. 476, 1975–1994. https://doi.org/10.1042/BCJ20190314.

Sah-Teli, S.K., Hynonen, M.J., Sulu, R., Dalwani, S., Schmitz, W., Wierenga, R.K., Venkatesan, R., 2020. Insights into the stability and substrate specificity of the E. coli aerobic beta-oxidation trifunctional enzyme complex. Journal of structural biology 210, 107494. S1047-8477(20)30060-5.

Salomon-Ferrer, R., Case, D.A., Walker, R.C., 2013. An overview of the Amber biomolecular simulation package: amber biomolecular simulation package. WIREs Comput. Mol. Sci. 3, 198–210. https://doi.org/10.1002/wcms.1121.

Sassetti, C.M., Rubin, E.J., 2003. Genetic requirements for mycobacterial survival during infection. Proc. Natl. Acad. Sci. U.S.A. 100, 12989–12994. https://doi.org/10.1073/pnas.2134250100.

Sato, S., Imamura, S., Ozeki, Y., Kawaguchi, A., 1992. Induction of enzymes involved in fatty acid beta-oxidation in Pseudomonas fragi B-0771 cells grown in media supplemented with fatty acid. J. Biochem. 111, 16–19. https://doi.org/10.1093/oxfordjournals.jbchem.a123711.

Schaefer, C.M., Lu, R., Nesbitt, N.M., Schiebel, J., Sampson, N.S., Kisker, C., 2015. FadA5 a thiolase from Mycobacterium tuberculosis: a steroid-binding pocket reveals the potential for drug development against tuberculosis. Structure (London, England : 1993) 23, 21–33. 10.1016/j.str.2014.10.010.

Schnappinger, D., Ehrt, S., Voskuil, M.I., Liu, Y., Mangan, J.A., Monahan, I.M., Dolganov, G., Efron, B., Butcher, P.D., Nathan, C., Schoolnik, G.K., 2003. Transcriptional adaptation of mycobacterium tuberculosis within macrophages: insights into the phagosomal environment. J. Exp. Med. 198, 693–704. https://doi.org/10.1084/jem.20030846.

Sparta, K.M., Krug, M., Heinemann, U., Mueller, U., Weiss, M.S., 2016. XDSAPP2.0. J. Appl. Crystallogr. 49, 1085–1092. https://doi.org/10.1107/S1600576716004416.

Sridhar, S., Schmitz, W., Hiltunen, J.K., Venkatesan, R., Bergmann, U., Kiema, T.R., Wierenga, R.K., 2020. Crystallographic binding studies of rat peroxisomal multifunctional enzyme type 1 with 3-ketodecanoyl-CoA: capturing active and inactive states of its hydratase and dehydrogenase catalytic sites. Acta Crystallogr. D Struct. Biol. 76, 1256–1269. https://doi.org/10.1107/S2059798320013819.

Srivastava, S., Chaudhary, S., Thukral, L., Shi, C., Gupta, R.D., Gupta, R., Priyadarshan, K., Vats, A., Haque, A.S., Sankaranarayanan, R., Natarajan, V.T., Sharma, R., Aldrich, C.C., Gokhale, R.S., 2015. Unsaturated lipid assimilation by mycobacteria requires auxiliary cis-trans enoyl coa isomerase. Chem. Biol. 22, 1577–1587. https://doi.org/10.1016/j.chembiol.2015.10.009.

Vagin, A., Teplyakov, A., 2010. Molecular replacement with MOLREP. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 22–25. https://doi.org/10.1107/S0907444909042589.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., Berendsen, H.J.C., 2005. GROMACS: fast, flexible, and free. J. Comput. Chem. 26, 1701–1718. https://doi.org/10.1002/jcc.20291.

Vanquelef, E., Simon, S., Marquant, G., Garcia, E., Klimerak, G., Delepine, J.C., Cieplak, P., Dupradeau, F.-Y., 2011. R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. Nucleic Acids Res. 39, W511–W517. https://doi.org/10.1093/nar/gkr288.

Venkatesan, R., Wierenga, R.K., 2013. Structure of mycobacterial beta-oxidation trifunctional enzyme reveals its altered assembly and putative substrate channeling pathway. ACS Chem. Biol. 8, 1063–1073. https://doi.org/10.1021/cb400007k.

Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., Case, D.A., 2004. Development and testing of a general amber force field. J. Comput. Chem. 25, 1157–1174. https://doi.org/10.1002/jcc.20035.

Waterson, R.M., Hass, G.M., Hill, R.L., 1972. Enoyl coenzyme A hydratase (crotonase). Enhancement of the rate of hydration of crotonylpantetheine by coenzyme A and related compounds. J. Biol. Chem. 247, 5252–5257.

Wilburn, K.M., Fieweger, R.A., VanderVen, B.C., 2018. Cholesterol and fatty acids grease the wheels of Mycobacterium tuberculosis pathogenesis. Pathogens Dis. 76 https://doi.org/10.1093/femspd/fty021.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 235–242. https://doi.org/10.1107/S0907444910045749.

Winter, G., 2010. xia2: an expert system for macromolecular crystallography data reduction. J. Appl. Crystallogr. 43, 186–190. https://doi.org/10.1107/S0021889809045701.

Winter, G., McAuley, K.E., 2011. Automated data collection for macromolecular crystallography. Methods 55, 81–93. https://doi.org/10.1016/j.ymeth.2011.06.010.

Winter, G., Waterman, D.G., Parkhurst, J.M., Brewster, A.S., Gildea, R.J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I.D., Sauter, N.K., Evans, G., 2018. DIALS: implementation and evaluation of a new integration package. Acta Crystallogr. Sect. D Struct. Biol. 74, 85–97. https://doi.org/10.1107/S2059798317017235.

Wipperman, M.F., Sampson, N.S., Thomas, S.T., 2014. Pathogen roid rage: cholesterol utilization by *Mycobacterium tuberculosis*. Crit. Rev. Biochem. Mol. Biol. 49, 269–293. https://doi.org/10.3109/10409238.2014.895700.

Xia, C., Fu, Z., Battaile, K.P., Kim, J.P., 2019. Crystal structure of human mitochondrial trifunctional protein, a fatty acid beta-oxidation metabolon. Proc. Natl. Acad. Sci. U.S.A. 116, 6069–6074. https://doi.org/10.1073/pnas.1816317116.

Yang, S.Y., Cuebas, D., Schulz, H., 1986. Channeling of 3-hydroxy-4-trans-decenoyl coenzyme A on the bifunctional beta-oxidation enzyme from rat liver peroxisomes and on the large subunit of the fatty acid oxidation complex from Escherichia coli. J. Biol. Chem. 261, 15390–15395.

Yaseen, A., Nijim, M., Williams, B., Qian, L., Li, M., Wang, J., Li, Y., 2016. FLEXc: protein flexibility prediction using context-based statistics, predicted structural features, and sequence information. BMC Bioinform. 17 (Suppl 8), 281. https://doi.org/10.1186/s12859-016-1117-3.

Zhang, H., Machutta, C.A., Tonge, P.J., 2010. Fatty Acid Biosynthesis and Oxidation. In: Comprehensive Natural Products Chemistry II: Chemistry and Biology. Elsevier, Oxford, pp. 231–275.