



<http://www.diva-portal.org>

## Postprint

This is the accepted version of a paper published in *Structure*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Andersson, C., Lundgren, C., Magnusdottir, A., Ge, C., Wieslander, Å. et al. (2012)  
The Mycobacterium tuberculosis Very-Long-Chain Fatty Acyl-CoA Synthetase: Structural Basis  
for Housing Lipid Substrates Longer than the Enzyme.  
*Structure*, 20(6): 1062-1070  
<http://dx.doi.org/10.1016/j.str.2012.03.012>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-79995>

Structure. 2012 Jun 6;20(6):1062-70. doi: 10.1016/j.str.2012.03.012.  
Epub 2012 May 3.

**The *Mycobacterium tuberculosis* very-long-chain fatty acyl-CoA synthetase: structural basis for housing lipid substrates longer than the enzyme.**

Andersson CS, Lundgren CA, Magnúsdóttir A, Ge C, Wieslander A, Martinez Molina D, Högbom M.

Stockholm Center for Biomembrane Research, Department of Biochemistry and Biophysics, Stockholm University, Arrhenius Laboratories for Natural Sciences C4, SE-10691 Stockholm, Sweden.

# **The *Mycobacterium tuberculosis* very-long-chain fatty acyl-CoA synthetase - structural basis for housing lipid substrates longer than the enzyme**

Charlotta S. Andersson, Camilla A. K. Lundgren, Auður Magnúsdóttir, Changrong Ge, Åke

Wieslander, Daniel Martinez Molina and Martin Högbom\*

Stockholm Center for Biomembrane Research, Department of Biochemistry and Biophysics, Stockholm University, Arrhenius Laboratories for Natural Sciences C4, SE-10691 Stockholm, Sweden.

\* To whom correspondence should be addressed. E-mail: [hogbom@dbb.su.se](mailto:hogbom@dbb.su.se)

Running title: Very-long-chain fatty acyl-CoA synthetase

## **Summary**

The *Mycobacterium tuberculosis* acid-induced operon *MymA* encodes the fatty acyl-CoA synthetase FadD13 and is essential for virulence and intracellular growth of the pathogen. Fatty acyl-CoA synthetases activate lipids before entering into the metabolic pathways and are also involved in transmembrane lipid transport. Unlike soluble fatty acyl-CoA synthetases, but like the mammalian integral-membrane very-long-chain acyl-CoA synthetases, FadD13 accepts lipid substrates up to the maximum length tested (C<sub>26</sub>). Here we show that FadD13 is a peripheral membrane protein. The structure and mutational studies reveal an arginine and aromatic-rich surface patch as the site for membrane interaction. The protein accommodates a hydrophobic tunnel that extends from the active site toward the positive patch and is sealed by an arginine-rich lid-loop at the protein surface. Based on this and previous data we propose a structural basis for accommodation of lipid substrates longer than the enzyme and transmembrane lipid transport by vectorial CoA-esterification.

## **Highlights:**

- *M. tuberculosis* FadD13 is a peripheral membrane protein
- Mutagenesis reveals an arginine-rich patch as the site for membrane interaction
- A hydrophobic tunnel extends from the active site to the positively charged patch
- The architecture lets substrates reside partly in the membrane during catalysis

## ***Introduction***

Fatty acids are activated before introduced into different metabolic pathways. This is universally achieved by fatty acyl-CoA synthetases (ACSs), in bacteria usually denoted FadD or FACS proteins. ACSs activate fatty acids, and other hydrophobic substrates, by a two-step mechanism, schematically illustrated in Figure 1. First, the substrate is reacted with ATP, forming an acyl-adenylate intermediate (acyl-AMP) with the concomitant release of pyrophosphate. The acyl-AMP intermediate is subsequently reacted with CoA to produce a fatty acyl-CoA thioester, the substrate for numerous anabolic and catabolic pathways, including membrane biogenesis. The central position in lipid uptake and metabolism also associate the ACS proteins with lipotoxicity and conditions such as insulin resistance, although the direct biochemical connection remains to be established (Li et al., 2010).

Three main groups of soluble ACS enzymes have been categorized based on their ability to activate different length substrates; short- (C2-C4), medium- (C4-C12) and long- (C12-C22) chain ACSs. Enzymes from the different groups share a two-domain structure with a common fold and show conformational changes and domain movements between the steps of the mechanism (Figure 1). The proteins form substrate-binding pockets of suitable size that explains the substrate-size specificity and, importantly, the inability to accommodate longer substrates (Hisanaga et al., 2004; Kochan et al., 2009; Reger et al., 2008). The ACS protein groups can be partly assigned based on sequence similarity but substrate specificity cannot be conclusively defined and must be experimentally determined (Black and DiRusso, 2003; Khurana et al., 2010; Watkins, 2008; Watkins et al., 2007).

There is also a group of very-long-chain ACSs, mainly characterized in eukaryotes (denoted ACSVL) (Mashek et al., 2004; Watkins, 2008). ACSVL enzymes have the capacity to accept

lipids longer than C<sub>22</sub>, usually tested by the ability to activate C<sub>24</sub> substrates. An interesting conundrum that arises when approaching lipid substrates of this size is that the length of the acyl chain is well beyond the expected dimensions of the protein, likely precluding the formation of a sufficiently large internal hydrophobic pocket to accommodate the substrate. Unlike other ACS proteins, the ACSVLs generally appear to be integral membrane proteins (Watkins, 2008; Gimeno, 2007; Uchiyama et al., 1996). A second subfamily “bubblegum” has also been identified and different members have been assigned as integral (Fraisl et al., 2004; Fraisl et al., 2006), or peripheral (Steinberg et al., 2000), membrane proteins. Interestingly, ACSVL enzymes were first identified as fatty acid transport proteins (denoted FATP family), moving exogenous lipids over the membrane (Hirsch et al., 1998; Schaffer and Lodish, 1994), thus linking assimilation and activation of fatty acids. It is still not completely resolved if the transport function is distinct from the activation or if it is a consequence of metabolic capture of the lipid by CoA-esterification on the cytosolic side, leading to net vectorial transport (DiRusso et al., 2008; Gimeno, 2007; Watkins, 2008). No structure is presently available for a full length ACSVL, likely because of the difficulties associated with membrane protein production and crystallization.

*Mycobacterium tuberculosis* is the causative agent of tuberculosis (TB). Despite the existence of antibiotic treatments and a vaccine it remains one of the worst global killers with an estimated 1.8 million yearly deaths and a third of the world’s population infected. There is a rapid development of multi drug-resistant (MDR) and extensively drug-resistant (XDR) strains. It is estimated that 1.3 million MDR-TB cases will have to be treated between 2010 and 2015 at a projected cost of over 16 billion US\$ (WHO, 2010). Identification of new anti-TB drugs and drug targets is imperative (WHO, 2009).

The bacterium has a unique lipid-rich cell envelope that is key to its virulence, inherent drug resistance and ability to multiply within the macrophage (Cole et al., 1998). The envelope consists of several complex lipids, in particular mycolic acids (MAs), long chain (C<sub>60</sub>-C<sub>90</sub>)  $\beta$ -hydroxy fatty acids. For this reason, lipid synthesis is a key drug target for present drugs and future drug design efforts (Barry et al., 2007; Hoffmann et al., 2008; Kaur et al., 2009). A number of studies show that the *MymA* operon, comprising seven proteins, is important for virulence and production of envelope lipids. Removal of the transcriptional regulator (*VirS*) or disruption the first gene in the operon leads to: altered morphology of the cell wall; less MAs and an altered MA composition and increased sensitivity to antibiotics, detergents, and acidic pH (Singh et al., 2005a; Singh et al., 2003; Singh et al., 2005b). In addition, the *MymA* encoded mRNAs are the most highly induced of all when bacteria are grown under acidic conditions in order to mimic the environment of the phagosome (Fisher et al., 2002). It has also been shown that the operon is required for *in vivo* growth of *M. tuberculosis* in macrophages (Cheruvu et al., 2007) and that the encoded proteins are highly upregulated in phagocytosed bacteria (Singh et al., 2003).

*M. tuberculosis* encodes no less than 36 ACS homologues, denoted FadD1-FadD36. A subset of these has also been shown to catalyze only the first step of the reaction (Arora et al., 2009; Trivedi et al., 2004). The FadD proteins thus play a central role in supporting the diverse and extensive lipid metabolism of the organism. FadD13 has attracted particular attention as a putative drug target because of its location in the *MymA* operon and the biochemical function as a node point in lipid metabolism (Jatana et al., 2010; Khare et al., 2009; Singh et al., 2005a; Singh et al., 2003; Singh et al., 2005b; Cheruvu et al., 2007; Fisher et al., 2002). *M. tuberculosis* FadD13 has been thoroughly biochemically characterized both in terms of substrate specificity and effects of protein variants on catalytic parameters (Khare et al., 2009). FadD13 can accept

substrates of different lengths and the activity increases with lipid length up to the limit tested (C<sub>26</sub>) (Khare et al., 2009), functionally placing it firmly among the ACSVL enzymes. Interestingly, unlike other ACSVL proteins, FadD13 is predicted to be a soluble protein. Homology models have been produced (Jatana et al., 2010; Khare et al., 2009) but a rationale for how the enzyme is able to accommodate lipid substrates significantly longer than the expected distance from the active site to the protein surface is still lacking.

To gain insight into substrate accommodation as well as provide a foundation for drug design against *M. tuberculosis*, we here present the 1.8 Å crystal structure of *M. tuberculosis* FadD13, the first structure of a full length ACSVL enzyme. We also show that FadD13 associates to the membrane when expressed in *Escherichia coli* and that the interaction can be partially reversed under certain buffer conditions, such as high concentrations of salt, phosphate or a high pH. The structure reveals a single exposed positively charged surface patch, rich in arginines and aromatic residues. This patch is shown to be the location for membrane interaction based on mutational studies. The protein contains a hydrophobic tunnel protruding from the active site toward the positive patch, capped at the surface by an arginine-rich lid-loop. Based on this and previous data we propose a structural basis for the accommodation of lipid substrates longer than the enzyme and vectorial lipid CoA-esterification by ACSVL enzymes.

## **Results**

### **FadD13 is a peripheral membrane protein**

Unlike canonical ACSVL proteins, the FadD13 protein sequence does not possess properties that would suggest a hydrophobic anchoring to the membrane. We hypothesized that FadD13 could

still associate to the membrane as a peripheral membrane protein, reversibly interacting through charge and/or moderately hydrophobic interactions.

To test this hypothesis we overexpressed the protein in *E. coli*. After mechanical disruption of the cells we prepared soluble and membrane fractions (see Experimental Procedures). The membrane fraction was resuspended and washed with buffer (80x the pellet volume) with different salt concentrations and pH before finally being repelleted by ultracentrifugation. The repelleted membrane fractions were dissolved in detergent (FC-12 or DDM) and applied to an IMAC column to capture any membrane-associated FadD13 protein and finally analyzed by SDS-PAGE. As shown in Figure 2, when membranes were washed with buffer at physiological pH and 150 mM NaCl, a large amount of FadD13 remained attached to the membrane and could be isolated upon membrane solubilization (per liter of culture, about 2 mg of pure FadD13 was recovered from isolated membranes). The amount of FadD13 that was isolated from the membrane fraction was reduced to about half when membranes were pre-washed with a high-pH (pH 11) buffer or with buffer containing 200 mM phosphate. Similarly, a high salt (NaCl) concentration during the membrane wash step also reduced the amount of protein isolated by approximately 40%. The protein is highly overexpressed and a significant amount of protein could also be isolated from the soluble fraction. Together, these results show that FadD13 indeed associates to the membrane, but not as an integral membrane protein. The relative ratio between the soluble and membrane fractions differs depending on the buffer conditions indicating an equilibrium between a membrane-bound and soluble state. Given the high expression obtained, the relative ratios may also be influenced by a possible saturation effect of the membrane.

### **Overall structure of FadD13**

Purification under high-salt conditions allowed us to produce a protein sample that was homogenous and could be crystallized. The 1.8 Å resolution structure of FadD13 includes residues -2 to 502. No electron density was observed for parts of a loop structure (residues 166-168) and the last residue of the protein (Lys 503). Similar to previous structures of ACS enzymes, the protein consists of two domains. The larger N-terminal domain (residues 1-395) is connected via a six-amino acid linker to the smaller C-terminal domain (residues 402-503) (Figure 3A and Figure S1) The relative orientation of the domains is consistent with the apo-form of shorter-chain ACS proteins (Hisanaga et al., 2004; Kochan et al., 2009; Reger et al., 2008). As in these proteins, the N-terminal domain folds into an  $\alpha + \beta$  topology and can be further divided into 2 subdomains. The smaller C-terminal domain has one two-stranded and one three-stranded antiparallel  $\beta$ -sheet, with three helices at its faces.

### **Active site structure**

The active site is located close to the hinge region between the N- and C-terminal domains and consists of three subsites. The ATP binding site is open and basic in nature (Figure 3B). FadD13 contains both ATP/AMP binding signature motifs normally found in ACS proteins, the first sequence (<sup>163</sup>YTSGTTGHPKGV<sup>174</sup>) is partially disordered in the structure, likely because of the absence of bound ATP. The sequence corresponding to the second motif (<sup>300</sup>GYALTE<sup>305</sup>), involved in binding of the adenine group, is structurally virtually identical to those of other ACS proteins. The CoA binding site in ACS proteins is formed in the junction between the domains upon a 140° rotation of the C-terminal domain after the initial adenylation step (Hisanaga et al., 2004; Kochan et al., 2009; Reger et al., 2008) (Figure 1, last step). Modeling of the CoA-bound

conformation based on homologous structures shows a conserved CoA binding site. The sequence and structural features thus suggest that FadD13 will bind the ATP and CoA substrates in the same manner as previously described ACS proteins. Despite extensive trials we have been unsuccessful in obtaining fatty acid or ATP analogue complex structures of FadD13. The lack of success may be due to the high-salt conditions needed to produce a crystallizable sample or because membrane interaction is needed to promote substrate binding (see Discussion).

### **Arginine- and aromatic-rich surface patch and lipid binding site**

The lipid binding site and surrounding structure display a number of interesting features and differences compared to shorter-chain ACS proteins. For example, in the long-chain ACS from *T. thermophilus* (ttLC-FACS), binding of AMP-PNP induces a rotamer change of Trp234, to interact directly with the beta phosphate of AMP-PNP. This shift opens a “dead-end branch” which houses the omega end of its substrate myristic acid (C<sub>14</sub>), and positions the lipid head group for activation (Hisanaga et al., 2004). A corresponding tryptophan residue is not present in FadD13 and the opening of a similar lipid-accommodating cavity is not expected to occur. In apo FadD13 there is a cavity in a similar position, but this is only large enough to house substrates of maximally around 14 carbon atoms in length, and thus not nearly large enough to accommodate a C<sub>26</sub> substrate. However, in apo FadD13, a hydrophobic cavity also extends from the ATP binding site towards the protein/solvent interface where it narrows and is sealed at the surface (Figure 3C). The cavity protrudes between the parallel beta sheet  $\beta 9$ - $\beta 14$  and helices  $\alpha 8$ - $\alpha 9$  (grey in Figure 4) and is lined by hydrophobic residues, leaving only a small number of possible H-bond interactions (mainly from the protein backbone). Each of the secondary structure elements, lining the fatty acid cavity, end in arginine-rich loops at the protein surface (Figure 4) and make up the

center of a larger patch, rich in arginines and aromatic residues that provides the only positively charged surface on this negatively charged protein (theoretical pI 5.1), aside from the ATP-binding cavity (Figure 5A). The patch comprises 14 surface-exposed arginines, and 10 hydrophobic aromatics (five tyrosines, three phenylalanines and two tryptophan residues) (Figure 5B). Of particular interest is a loop structure (residues 193-199) that links  $\alpha 8$  and the central strand of the  $\beta$ -sheet (red in Figure 4 and 5B) and seals the hydrophobic cavity at the protein surface. The loop is located centrally in the positively charged patch and itself contains one tyrosine and three arginine residues.

### **The arginine- and aromatic-rich surface patch is involved in membrane binding**

Given the particular chemical nature of the surface patch we suspected that it was involved in membrane binding. Aromatic residues are the ones most prone to bind to lipid bilayer (hydrophobic) interfaces (Wimley and White, 1996) while arginines yields affinity to both zwitterionic and anionic lipid polar headgroups. Both types are present in lipid-binding sites in membrane proteins (Palsdottir and Hunte, 2004) and peptides (Yang et al., 2003). To investigate this, we used a screening protocol for binding of the wild-type protein to liposomes (of various compositions) and to commercial Membrane Lipid Strips™, as described in (Ge et al., 2011). The protein revealed a weak binding to (zwitterionic) phosphatidylcholine which was enhanced somewhat by mixing in the anionic phosphatidylglycerol and especially by cardiolipin, both present in *M.tuberculosis* (Sartain et al., 2011). However, several strip lipid spots containing various, single (100%) anionic lipids, did not confer binding. These findings were quantified by surface plasmon resonance (SPR) analyses, monitoring the binding of the wild type protein and a quintuple residue variant of the protein where the charge of the patch had been reduced while its

hydrophobicity was increased (R9A, R17A, R195A, R197A, R244A). The interaction to a phosphatidylcholine:cardiolipin 70:30 model membrane bilayer-surface coated sensor chip was investigated using a BSA-coated surface as the reference (Figure 6 and Figure S2A and B). The data show that both wild-type and mutant proteins bind to the model membrane surface in a concentration dependent manner, consistent with the binding observed to the *E. coli* membrane, and ruling out the requirement for additional proteins for membrane interaction. Reducing plus-charge and increasing hydrophobicity of the surface patch led to an increase in total binding to the membrane surface. The fact that the strength of the membrane binding is altered when altering the patch shows that it is involved in the binding. Even by this rather extensive mutagenesis, removing positive charges that are likely important to initially recruit the protein to the membrane, it is clear that the membrane association of the enzyme was not prevented. This may not be very surprising, considering the large size of the patch and the great number of involved amino acids and potential binding motifs (e.g. only 5 out of 14 arginines were removed). The fact that the association in total (thus including the contribution from both on- and off-rates) became stronger, also points to an important role for hydrophobic interactions to keep the protein bound at the membrane.

### **Structural basis of catalytic properties in protein variants.**

In a recent study Khare and coworkers (Khare et al., 2009) investigated the catalytic properties of numerous FadD13 protein variants, of which the most important are discussed here. The K172A protein variant shows ~ 3 fold lower  $k_{cat}$  and ~ 10-fold increase in  $K_m$  for ATP while the K487A variant shows ~ 30 fold lower  $k_{cat}$ . K172 is located in the first ATP-binding signature motif while K487 is located in a loop of the C-terminal domain. When ATP is modeled into FadD13 these residues form charge-charge interactions with the phosphates and are thus likely directly involved

in ATP binding (Supplemental Figure 1A). W377 is partially surface exposed and located distant from the substrate-binding sites, still the W377A protein variant show ~ 20-fold reduction in both  $k_{\text{cat}}$  and  $V_{\text{max}}$ . The structure reveals that W377 forms a small hydrophobic core, linking three secondary structure elements (Supplemental Figure 1B) and its importance is likely due to structural integrity of the enzyme. The V209D, A211G, A302G and D382A protein variants all primarily influence  $K_m$  for the lipid substrate. V209D, A211G and D382A decrease the binding affinity by ~ 2, 2 and 4 fold, respectively, while A302G increases it ~ 2.5 fold. All these residues influence the secondary structure elements and the internal surface of the lipid-binding tunnel entrance at the ATP-binding site, D382 via a H-bond to Y301 (Supplemental Figure 1C).

## ***Discussion***

Unlike most other biological molecules, lipids are apolar and show very low water solubility. This presents particular challenges for the largely water solution-based chemistry of biological systems. Lipids partition into membranes and must be activated and carried by specific lipid-binding proteins to be entered into cellular metabolism. This problem is accentuated for very long lipids that are not readily accommodated within the hydrophobic core of proteins. One possible solution to this problem is to allow part of the lipid to reside in the membrane while the other part is enzymatically modified. This mechanism appears to be in effect in the integral-membrane ACSVL enzymes. These enzymes also function in the transmembrane transport of lipids (Hirsch et al., 1998; Schaffer and Lodish, 1994), linking assimilation and activation of fatty acids. At least one mechanism for net transfer appears to be metabolic capture of the lipid by CoA-esterification on the cytosolic side (Watkins, 2008; DiRusso et al., 2008; Gimeno, 2007).

FadD13 activates lipids, with increasing activity, up to the limit tested (C<sub>26</sub>) (Khare et al., 2009), firmly placing it functionally among the ACSVL proteins. The structure confirms that FadD13 is unable to accommodate substrates of this length within the protein in any way similar to the binding modes previously observed (Hisanaga et al., 2004; Kochan et al., 2009; Reger et al., 2008). However, here we also show that FadD13 is a peripheral membrane protein, likely reversibly associating with the membrane. The protein surface contains distinctive patch, rich in arginines and aromatics, that we show is involved in the membrane association. Arginines and aromatic residues are commonly found in protein/membrane interfaces and also make up distinct lipid-binding motifs i.e. for cardiolipin aromatic residues are quite common (Palsdottir and Hunte, 2004). However, this plus-charged patch is located surrounded by many negatively charged residues (see Figure 5A), so both attractive and repellent forces by the negatively charged membrane bilayer surface, act on the protein (-14e net charge), cf. (Mulgrew-Nesbitt et al., 2006). It is likely that a critical balance between quenching ions (salts), local pH, and local lipid minus charge concentration, will affect the amounts of protein associated (“attracted”) to the bilayer or staying in the in the cytoplasm (“repelled”). It is well established that e.g. the anionic cardiolipin can have a heterogenous lateral distribution in bacterial membranes.

The structure also reveals a hydrophobic tunnel, extending from the active site towards the positively charged surface patch, capped by an arginine-rich lid-loop. Based on these structural and biochemical features we hypothesize that the lid loop opens upon interaction with the membrane, allowing binding of substrate with the lipid tail remaining partly in the membrane during catalysis (Figure 7). The positive nature of this region of the protein surface may also contribute to recruit fatty acid substrates to the enzyme. A partial accommodation of the substrate within the membrane is also consistent with the enzyme’s observed non-strict specificity for fatty

acid chain length (Khare et al., 2009). Despite significant effort we have not been able to obtain lipid complex structures by co-crystallization or soaking, possibly because of the need for a membrane interaction to induce a substrate-binding conformation of the hydrophobic channel

*M. tuberculosis* FadD13 perform the same function as the mammalian integral-membrane ACSVL enzymes but an important remaining question is if FadD13 can be used also as a structural model for the catalytic domains of these. Even though phylogenetic analysis indicate a common ancestry (Watkins et al., 2007), a conclusive assignment of the catalytic domains of ACSVL enzymes to the same fold as other ACS proteins has been hampered by the borderline sequence similarity (Uchiyama et al., 1996; Hirsch et al., 1998; Lewis et al., 2001; Obermeyer et al., 2007; Richards et al., 2003). However, benefiting from the more sophisticated bioinformatic tools and larger sequence and structure base presently available, this assignment now appears unambiguous. In addition to convincing sequence alignment reliability scores for homology, structure prediction/fold recognition using all tested methods, including for example Swissmodel (Arnold et al., 2006) (<http://swissmodel.expasy.org/>), Psipred (Bryson et al., 2005) (<http://bioinf.cs.ucl.ac.uk/psipred/>) and threading using the Wurst server (Torda et al., 2004) (<http://www.zbh.uni-hamburg.de/wurst>), suggest this fold with very high confidence scores. The membrane-bound ACSVLs commonly contain N-terminal extensions compared to FadD13. Experiments and topology/hydrophobicity predictions suggest that the most common architecture of ACSVL proteins include one N-terminal transmembrane helix and other hydrophobic parts of the protein that provide additional interfacial anchoring sites to the membrane (Lewis et al., 2001; Obermeyer et al., 2007; Richards et al., 2003). The hydrophobic segments of these proteins are located mainly in the N-terminal half of the sequence, consistent with the membrane interaction geometry as proposed here for FadD13 via mainly hydrophobic interactions. This is also

consistent with the topological models previously proposed for ACSVL/FATP proteins (Lewis et al., 2001; Obermeyer et al., 2007; Richards et al., 2003). For the reasons described above, we propose that the structure and membrane association topology of FadD13 is a good general model for the catalytic domains of the integral membrane ACSVL enzymes.

From a lipid metabolism perspective, the linking of transmembrane transport and activation of lipids is particularly interesting. In the original publications defining the group called fatty acid transport proteins (Hirsch et al., 1998; Schaffer and Lodish, 1994), now known to be ACSVL enzymes, one *M. tuberculosis* homologue (FadD6) was identified. However, recently it was shown that disruption of another homologue, FadD5, gives an attenuated strain that is impaired in utilizing mycolic acids provided in the growth medium as a carbon source. No difference in growth rate was observed when shorter-chain fatty acids such as palmitate (C<sub>16</sub>) or oleate (C<sub>18:1 cis-9</sub>) were provided (Dunphy et al., 2010). It thus appears that more ACS variants than initially thought may be capable of specific lipid transport. Out of the 36 FadDs in *M. tuberculosis* FadD5 is the closest sequence homologue (~ 37% identity) to FadD13, described here, providing additional clues to the exact metabolic context of the *MymA* operon, known to influence virulence, lipid metabolism and survival. The proposed membrane-protein co-accommodation of substrate is likely a feature that should be considered when designing inhibitors for these proteins and can possibly be exploited to improve specificity for ACSVL enzymes.

## **Experimental Procedures**

### **Cloning, Variant construction, Protein Expression, and Purification**

The gene encoding FadD13 was cloned from *M. tuberculosis* strain H37Rv into the pET-46 Ek/LIC vector using ligation free cloning, introducing a N-terminal His<sub>6</sub>-tag. The resulting plasmid pET-46/FadD13 was transformed into *E. coli* BL21(DE3) after sequence verification. The cells were grown in TB medium supplemented with carbenicillin at 37° C to A<sub>600</sub> = 0.8-1.0, and overexpression was induced by 0.5 mM isopropyl thio-β-d-1-galactoside (IPTG) at 20°C over night. Cells were collected, and periplasmic material was removed before lysis (Magnusdottir et al., 2009). The protein was purified by IMAC (immobilized metal affinity chromatography) chromatography using Ni-loaded profinity resin (*Biorad*) in 20 mM HEPES pH 7.5, 500 mM NaCl, 10% glycerol, 10 mM imidazole, and 0.5 mM TCEP and washed with 25 mM imidazole before elution with 400 mM imidazole. The second purification step, size exclusion chromatography was done on a HiLoad 16/60 Superdex 200 (*GE healthcare*) at an increased salt concentration, 600 mM NaCl. The protein was concentrated to 11 mg/ml concentrated in a vivaspin 30K (*Sartorius biotech*) and was finally stored in 20 mM HEPES pH 7.5, 600 mM NaCl, 10% glycerol, and 0.5 mM TCEP. A 5-point protein variant was constructed by substitution of five arginines into alanine residues (R9A, R17A 195A, R197A, R244A), to create a less charged and more hydrophobic surface patch. The sequence verified synthetic gene, provided by GeneArt, was and moved into the expression vector and used for protein production as described above.

### **Membrane association studies**

Transformed cells were grown in Terrific-Broth media at 37 °C. Overexpression was initiated

with 0.1 mM isopropyl-beta-D-thiogalactopyranoside (Anatrace) at OD<sub>600</sub> 0.9 - 1.0 and carried out at 20 °C for 18 h. The cells were harvested by 15 min. centrifugation at 7500xg. Cells were resuspended in 20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP, Benzonase and EDTA-free protease inhibitor cocktail (Buffer A) and lysed with three rounds of freeze thawing followed by emulsiflex (high-pressure homogenization apparatus, Avestin Inc.) treatment. The lysate was clarified by centrifugation at 11000xg for 30 minutes and the supernatant divided into three equal fractions (A, B and C). The membranes were pelleted by ultra centrifugation at 140000xg for one hour. The resulting pellets were resuspended in buffers with varying composition; A in Buffer A; B in 200 mM Na-phosphate buffer at pH 7,5 and; C in 100 mM Na<sub>2</sub>CO<sub>3</sub> at pH 11. All samples contained 150 mM NaCl and were supplemented with 0.5 mM TCEP and EDTA-free protease inhibitor cocktail. After 1h of gentle shaking at 4°C, the membranes were re-pelleted by ultracentrifugation and again washed by resuspension in the respective buffer and left on ice for one hour. The membranes were re-pelleted and solubilized with 1% detergent FC-12 (foscholine-12) in Buffer A supplemented with 20 mM imidazole for three hours at 4° C with gentle stirring. Unsolubilized material was pelleted by ultracentrifugation and solubilized material was loaded on an Ni-loaded NTA-resin (*Qiagen*). Bound material was washed with Buffer A supplemented with 20 mM imidazole before elution with Buffer A supplemented with 500 mM imidazole. All fractions of the IMAC purification were followed by SDS-PAGE (4-12% Bis-Tris, Invitrogen) and stained with Coomassie blue.

To see the effect of increasing salt concentrations in the wash steps, FadD13 was expressed as described above, the cells were divided into four equal parts, lysed and washed as described above, but varying the amounts of salt of Buffer A (0 mM, 150 mM, 300 mM and 600 mM NaCl). Prior to solubilization, all membrane pellets were resuspended in Buffer A supplemented

with 20 mM imidazole and with a final NaCl concentration of 600 mM in all samples. Detergent (DDM) (n-Dodecyl- $\beta$ -maltoside) was added to each sample to a concentration of 1% and left to solubilize on gentle stirring over night at 4° C. Unsolubilized material was pelleted and the detergent solubilized membrane fraction was purified as described above. Also here, all fractions of the Ni-IMAC were followed by SDS-PAGE.

### **Liposome Preparation**

Lipids were dissolved in chloroform and mixed to yield a 10 mM mixture of dioleoyl-phosphatidylcholine (DOPC) and cardiolipin (CL) at a ratio of 70:30 (mole:mole). The solvent was evaporated under N<sub>2</sub> gas, and the resulting lipid film was further dried under vacuum overnight. The dried lipid film was dispersed in running buffer (10 mM Hepes pH 7.5, 150 mM NaCl, 5 % glycerol, 0.1 % DMSO) by extensive vortexing. The hydrated lipids were incubated for 1 h at room temperature. Liposomes were prepared using a 1 ml extruder (Avestin LiposoFast) and the lipid suspension was pushed through 0.1  $\mu$ m membrane 21 times to ensure unilamellar vesicles.

### **Surface Plasmon Resonance (SPR) Binding Analysis**

The binding analysis was performed using a Biacore3000 instrument with a L1 Sensor Chip (GE Healthcare) and 10 mM Hepes at pH 7.5, 150 mM NaCl, 5 % glycerol, 0.1 % DMSO as running buffer. The sensor chip surface was washed with 40 mM n-octyl- $\beta$ -glycoside (20  $\mu$ l, 20  $\mu$ l/min) prior to liposome injection. Liposomes were injected (30  $\mu$ l, 3  $\mu$ l/min) and immobilized to the surface yielding a response level of 1800-2700 RU. Unbound vesicles were washed away with running buffer (125  $\mu$ l, 50  $\mu$ l/min). The surface was allowed to stabilize (40  $\mu$ l, 5  $\mu$ l/min) prior to sample injection. Bovine serum albumin, BSA (40  $\mu$ g/ml), was covalently immobilized via

amine coupling and used as a reference surface. Sample was injected over the surfaces (25  $\mu$ l, 5  $\mu$ l/min) and association and dissociation was followed for 5 min respectively. Thereafter the surface was regenerated with 50 % 2-propanol in 50 mM NaOH.

### **Crystallization, Data Collection, Structure Determination**

FadD13 was crystallized in the presence of 10 mM CoA and 1 mM ATP, using the sitting drop vapor diffusion method. The protein was first mixed with ATP and reached a concentration of 11 mg/ml. In the final crystallization conditions 0.5  $\mu$ l protein solution was mixed with 1.5  $\mu$ l reservoir solution containing 100 mM Hepes pH 7.3, 20 mM MgCl<sub>2</sub>, 23% polyacrylic acid, and 10 mM CoA, crystals grew in 1-2 days. Attempts to setup crystallization in the absence of either ATP or CoA generated none or were poorly diffracting crystals. Crystals were cryoprotected by soaking for 10 minutes in mother liquor supplemented with 60% w/v Na<sub>2</sub> malonate. Data were collected to 1.8 Å at 100K at beamline ID14-2 at the ESRF synchrotron in Grenoble, France. Diffraction data were processed and scaled using MOSFLM (Leslie, 1992) and SCALA from the CCP4 suite (CCP4, 1994). Data statistics are presented in table 1. Initial phases were obtained using molecular replacement with Phaser (McCoy et al., 2007) using the *Thermus thermophilus* tt0168 structure (PDB id 1ULT) as the template for model generation. The initial molecular replacement solution was improved with Phenix Autobuild (Adams et al., 2010) and Arp/Warp (Perrakis et al., 1999) followed by cycles of manual building in COOT (Emsley and Cowtan, 2004). Refinement was carried out in Refmac5 (Vagin et al., 2004). Model statistics are given in Table 2. Molecular surfaces and cavities were calculated using the program MOLE (Petrek et al., 2007). All figures were made in PyMol (<http://www.pymol.org>). The model and structure factors have been deposited in the PDB with id 3R44.

## References

- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* *66*, 213-221.
- Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* *22*, 195-201.
- Arora, P., Goyal, A., Natarajan, V.T., Rajakumara, E., Verma, P., Gupta, R., Yousuf, M., Trivedi, O.A., Mohanty, D., Tyagi, A., *et al.* (2009). Mechanistic and functional insights into fatty acid activation in *Mycobacterium tuberculosis*. *Nat Chem Biol* *5*, 166-173.
- Barry, C.E., Crick, D.C., and McNeil, M.R. (2007). Targeting the formation of the cell wall core of *M. tuberculosis*. *Infect Disord Drug Targets* *7*, 182-202.
- Black, P.N., and DiRusso, C.C. (2003). Transmembrane movement of exogenous long-chain fatty acids: proteins, enzymes, and vectorial esterification. *Microbiol Mol Biol Rev* *67*, 454-472, table of contents.
- Bryson, K., McGuffin, L.J., Marsden, R.L., Ward, J.J., Sodhi, J.S., and Jones, D.T. (2005). Protein structure prediction servers at University College London. *Nucleic Acids Res* *33*, W36-38.
- CCP4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* *50*, 760-763.
- Cheruvu, M., Plikaytis, B.B., and Shinnick, T.M. (2007). The acid-induced operon Rv3083-Rv3089 is required for growth of *Mycobacterium tuberculosis* in macrophages. *Tuberculosis (Edinb)* *87*, 12-20.
- Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., *et al.* (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* *393*, 537-544.
- DiRusso, C.C., Darwis, D., Obermeyer, T., and Black, P.N. (2008). Functional domains of the fatty acid transport proteins: studies using protein chimeras. *Biochim Biophys Acta* *1781*, 135-143.
- Dunphy, K.Y., Senaratne, R.H., Masuzawa, M., Kendall, L.V., and Riley, L.W. (2010). Attenuation of *Mycobacterium tuberculosis* functionally disrupted in a fatty acyl-coenzyme A synthetase gene *fadD5*. *J Infect Dis* *201*, 1232-1239.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* *60*, 2126-2132.
- Fisher, M.A., Plikaytis, B.B., and Shinnick, T.M. (2002). Microarray analysis of the *Mycobacterium tuberculosis* transcriptional response to the acidic conditions found in phagosomes. *J. Bacteriol.* *184*, 4025-4032.

- Fraisl, P., Forss-Petter, S., Zigman, M., and Berger, J. (2004). Murine bubblegum orthologue is a microsomal very long-chain acyl-CoA synthetase. *Biochem J* 377, 85-93.
- Fraisl, P., Tanaka, H., Forss-Petter, S., Lassmann, H., Nishimune, Y., and Berger, J. (2006). A novel mammalian bubblegum-related acyl-CoA synthetase restricted to testes and possibly involved in spermatogenesis. *Arch Biochem Biophys* 451, 23-33.
- Ge, C., Georgiev, A., Ohman, A., Wieslander, A., and Kelly, A.A. (2011). Tryptophan residues promote membrane association for a plant lipid glycosyltransferase involved in phosphate stress. *J Biol Chem* 286, 6669-6684.
- Gimeno, R.E. (2007). Fatty acid transport proteins. *Curr Opin Lipidol* 18, 271-276.
- Hirsch, D., Stahl, A., and Lodish, H.F. (1998). A family of fatty acid transporters conserved from mycobacterium to man. *Proc Natl Acad Sci U S A* 95, 8625-8629.
- Hisanaga, Y., Ago, H., Nakagawa, N., Hamada, K., Ida, K., Yamamoto, M., Hori, T., Arai, Y., Sugahara, M., Kuramitsu, S., *et al.* (2004). Structural basis of the substrate-specific two-step catalysis of long chain fatty acyl-CoA synthetase dimer. *J Biol Chem* 279, 31717-31726.
- Hoffmann, C., Leis, A., Niederweis, M., Plitzko, J.M., and Engelhardt, H. (2008). Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proc Natl Acad Sci U S A* 105, 3963-3967.
- Jatana, N., Jangid, S., Khare, G., Tyagi, A.K., and Latha, N. (2010). Molecular modeling studies of Fatty acyl-CoA synthetase (FadD13) from Mycobacterium tuberculosis-a potential target for the development of antitubercular drugs. *J Mol Model*.
- Kaur, D., Guerin, M.E., Skovierova, H., Brennan, P.J., and Jackson, M. (2009). Chapter 2: Biogenesis of the cell wall and other glycoconjugates of Mycobacterium tuberculosis. *Adv Appl Microbiol* 69, 23-78.
- Khare, G., Gupta, V., Gupta, R.K., Gupta, R., Bhat, R., and Tyagi, A.K. (2009). Dissecting the role of critical residues and substrate preference of a Fatty Acyl-CoA Synthetase (FadD13) of Mycobacterium tuberculosis. *PLoS One* 4, e8387.
- Khurana, P., Gokhale, R.S., and Mohanty, D. (2010). Genome scale prediction of substrate specificity for acyl adenylate superfamily of enzymes based on active site residue profiles. *BMC Bioinformatics* 11, 57.
- Kochan, G., Pilka, E.S., von Delft, F., Oppermann, U., and Yue, W.W. (2009). Structural snapshots for the conformation-dependent catalysis by human medium-chain acyl-coenzyme A synthetase ACSM2A. *J Mol Biol* 388, 997-1008.
- Leslie, A.G.W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* 26.
- Lewis, S.E., Listenberger, L.L., Ory, D.S., and Schaffer, J.E. (2001). Membrane topology of the murine fatty acid transport protein 1. *J Biol Chem* 276, 37042-37050.
- Li, L.O., Klett, E.L., and Coleman, R.A. (2010). Acyl-CoA synthesis, lipid metabolism and lipotoxicity. *Biochim Biophys Acta* 1801, 246-251.

- Magnusdottir, A., Johansson, I., Dahlgren, L.G., Nordlund, P., and Berglund, H. (2009). Enabling IMAC purification of low abundance recombinant proteins from *E. coli* lysates. *Nat Methods* 6, 477-478.
- Mashek, D.G., Bornfeldt, K.E., Coleman, R.A., Berger, J., Bernlohr, D.A., Black, P., DiRusso, C.C., Farber, S.A., Guo, W., Hashimoto, N., *et al.* (2004). Revised nomenclature for the mammalian long-chain acyl-CoA synthetase gene family. *J Lipid Res* 45, 1958-1961.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J Appl Crystallogr* 40, 658-674.
- Mulgrew-Nesbitt, A., Diraviyam, K., Wang, J., Singh, S., Murray, P., Li, Z., Rogers, L., Mirkovic, N., and Murray, D. (2006). The role of electrostatics in protein-membrane interactions. *Biochim Biophys Acta* 1761, 812-826.
- Obermeyer, T., Fraisl, P., DiRusso, C.C., and Black, P.N. (2007). Topology of the yeast fatty acid transport protein Fat1p: mechanistic implications for functional domains on the cytosolic surface of the plasma membrane. *J Lipid Res* 48, 2354-2364.
- Palsdottir, H., and Hunte, C. (2004). Lipids in membrane protein structures. *Biochim Biophys Acta* 1666, 2-18.
- Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* 6, 458-463.
- Petrek, M., Kosinova, P., Koca, J., and Otyepka, M. (2007). MOLE: a Voronoi diagram-based explorer of molecular channels, pores, and tunnels. *Structure* 15, 1357-1363.
- Reger, A.S., Wu, R., Dunaway-Mariano, D., and Gulick, A.M. (2008). Structural characterization of a 140 degrees domain movement in the two-step reaction catalyzed by 4-chlorobenzoate:CoA ligase. *Biochemistry* 47, 8016-8025.
- Richards, M.R., Listenberger, L.L., Kelly, A.A., Lewis, S.E., Ory, D.S., and Schaffer, J.E. (2003). Oligomerization of the murine fatty acid transport protein 1. *J Biol Chem* 278, 10477-10483.
- Sartain, M.J., Dick, D.L., Rithner, C.D., Crick, D.C., and Belisle, J.T. (2011). Lipidomic analyses of *Mycobacterium tuberculosis* based on accurate mass measurements and the novel "Mtb LipidDB". *J Lipid Res* 52, 861-872.
- Schaffer, J.E., and Lodish, H.F. (1994). Expression cloning and characterization of a novel adipocyte long chain fatty acid transport protein. *Cell* 79, 427-436.
- Singh, A., Gupta, R., Vishwakarma, R.A., Narayanan, P.R., Paramasivan, C.N., Ramanathan, V.D., and Tyagi, A.K. (2005a). Requirement of the *mymA* operon for appropriate cell wall ultrastructure and persistence of *Mycobacterium tuberculosis* in the spleens of guinea pigs. *J. Bacteriol.* 187, 4173-4186.
- Singh, A., Jain, S., Gupta, S., Das, T., and Tyagi, A.K. (2003). *mymA* operon of *Mycobacterium tuberculosis*: its regulation and importance in the cell envelope. *FEMS Microbiol. Lett.* 227, 53-63.
- Singh, R., Singh, A., and Tyagi, A.K. (2005b). Deciphering the genes involved in pathogenesis of *Mycobacterium tuberculosis*. *Tuberculosis* 85, 325-335.

- Steinberg, S.J., Morgenthaler, J., Heinzer, A.K., Smith, K.D., and Watkins, P.A. (2000). Very long-chain acyl-CoA synthetases. Human "bubblegum" represents a new family of proteins capable of activating very long-chain fatty acids. *J Biol Chem* 275, 35162-35169.
- Torda, A.E., Procter, J.B., and Huber, T. (2004). Wurst: a protein threading server with a structural scoring function, sequence profiles and optimized substitution matrices. *Nucleic Acids Res* 32, W532-535.
- Trivedi, O.A., Arora, P., Sridharan, V., Tickoo, R., Mohanty, D., and Gokhale, R.S. (2004). Enzymic activation and transfer of fatty acids as acyl-adenylates in mycobacteria. *Nature* 428, 441-445.
- Uchiyama, A., Aoyama, T., Kamijo, K., Uchida, Y., Kondo, N., Orii, T., and Hashimoto, T. (1996). Molecular cloning of cDNA encoding rat very long-chain acyl-CoA synthetase. *J Biol Chem* 271, 30360-30365.
- Vagin, A.A., Steiner, R.A., Lebedev, A.A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G.N. (2004). REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D Biol Crystallogr* 60, 2184-2195.
- Watkins, P.A. (2008). Very-long-chain acyl-CoA synthetases. *J Biol Chem* 283, 1773-1777.
- Watkins, P.A., Maignel, D., Jia, Z., and Pevsner, J. (2007). Evidence for 26 distinct acyl-coenzyme A synthetase genes in the human genome. *J Lipid Res* 48, 2736-2750.
- WHO (2009). WHO global report 2009.
- WHO (2010). Multidrug and extensively drug-resistant TB (M/XDR-TB): 2010 Global report on surveillance and response.
- Wimley, W.C., and White, S.H. (1996). Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct Biol* 3, 842-848.
- Yang, S.T., Shin, S.Y., Lee, C.W., Kim, Y.C., Hahm, K.S., and Kim, J.I. (2003). Selective cytotoxicity following Arg-to-Lys substitution in tritrypticin adopting a unique amphipathic turn structure. *FEBS Lett* 540, 229-233.

## **Acknowledgments**

We would like to acknowledge assistance from the Staff at Beamline ID14-2, ESRF, Grenoble. Genomic DNA was provided by Colorado State University as part of NIH, NIAID Contract No. HHSN266200400091C, entitled "Tuberculosis Vaccine Testing and Research Materials,". This work was supported by grants from the Swedish Research Council (2010-5061), the Carl-Trygger foundation, and the Swedish Foundation for Strategic Research.

## Figure legends

Figure 1. Schematic general mechanism in ACS proteins. Binding of ATP induces structural changes promoting binding of the hydrophobic substrate. Formation of an acyl-adenylate intermediate induces a 140° rotation of the small domain and binding of CoA for production of the final product, a fatty acyl-CoA thioester.

Figure 2. FadD13 associates to the membrane. Coomassie blue stained gel showing the protein that remains bound to the membrane fraction after it has been washed with the buffers indicated above the lane. After membrane isolation, the membrane pellet was resuspended and washed twice with buffer with varying composition (text) before being repelleted and solubilized in detergent to isolate the protein associated to the membrane using IMAC affinity chromatography. Washes with phosphate, alkaline pH or high salt reduced the amount of FadD13 associated to the membrane.

Figure 3. Overall structure of *M. tuberculosis* FadD13. A, rainbow coloring from blue (N-terminal) to red (C-terminal), showing the 2-domain structure. B, surface contact potential showing the basic ATP-binding cavity (center). C, Hydrophobic tunnel protruding from the ATP binding site toward the protein surface. The channel is capped at the surface by a loop structure (red), see also Figure S1.

Figure 4. Stereo figure showing the arginine-rich lid loop capping the hydrophobic channel at the protein surface (in red). The channel is lined by sidechains provided by a parallel beta sheet and two helices, in grey.

Figure 5. Details of the arginine-rich surface patch. A, surface contact potential. B, Positioning of arginines and aromatic sidechains in relation to the cavity-lining secondary structure elements (grey) and the lid loop (red).

Figure 6. Binding analysis by SPR. A 75  $\mu$ M sample of FadD13 (solid lines) and the R9A, R17A, R195A, R197A, R244A variant (dotted lines) was injected over a DOPC/CL (70:30, mole:mole) surface, association and dissociation was followed for 5 min respectively. Duplicate experiments are shown (gray and black). Sensorgrams show values subtracted against a BSA-coated reference surface, see also Figure S2.

Figure 7. Proposed model for catalysis and lipid accommodation in FadD13 and ACSVL enzymes. The anchoring to the membrane may be polar and reversible (FadD13) or hydrophobic and permanent (mammalian integral-membrane ACSVL enzymes, N-terminal transmembrane segments, commonly found in ACSVL enzymes, are indicated with a dashed line). Large lipid substrates are expected to partition to the membrane and may be membrane lipids or originate intracellularly or extracellularly. Net vectorial transport is achieved by metabolic capture as the acyl-CoA thioester is formed intracellularly.

## Tables

**Table 1**

<b>Data Collection</b>	
Beamline	ID14-2 ESRF
Wavelength (Å)	0.933
Space group	P3 <sub>1</sub> 21
Cell parameters: (Å)	57.00 ; 57.00 ; 249.18
(°)	90 ; 90 ; 120
Resolution (Å)	32-1.80 (1.90-1.80)
Completeness (%)	96.7 (87.7)
Redundancy	3.06 (2.15)
R <sub>meas</sub> (%)	6.4 (60.4)
I/σI	12.71 (2.01)

**Table 2**

<b>Refinement statistics</b>	
Resolution (Å)	19.86-1.80 (1.846-1.800)
No. of unique reflections	41486
No. of reflections in test set	2205
R <sub>work</sub> (%)	16.2 (23.7)
R <sub>free</sub> (%)	19.9 (29.1)
No. of atoms	
Protein	7729
Solvent	292
RMSD from ideal values	
Bond length (Å)	0.022
Bond angle (°)	1.967
Ramachandran outliers, %	0.4

Figure 1  
[Click here to download high resolution image](#)

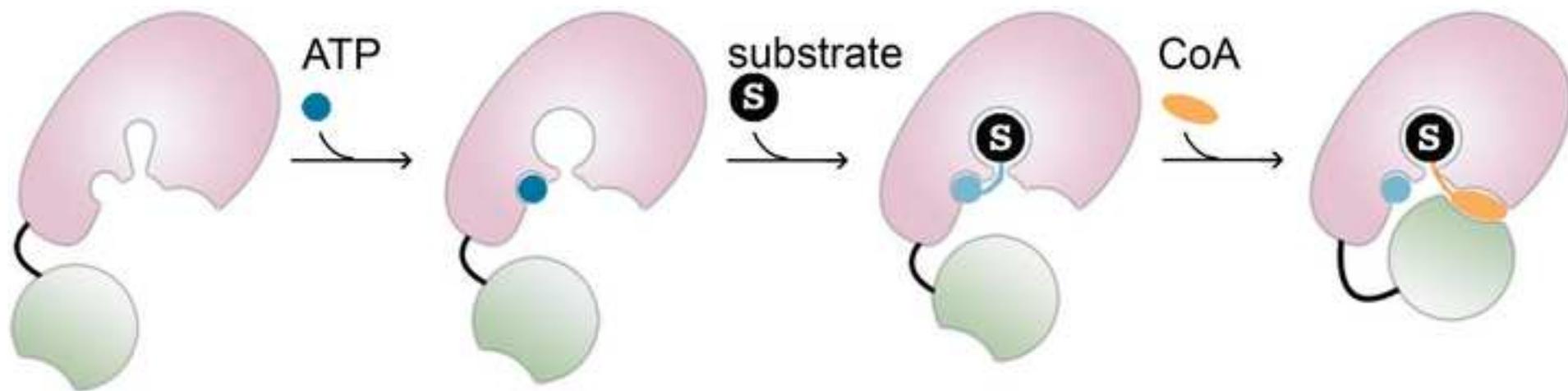
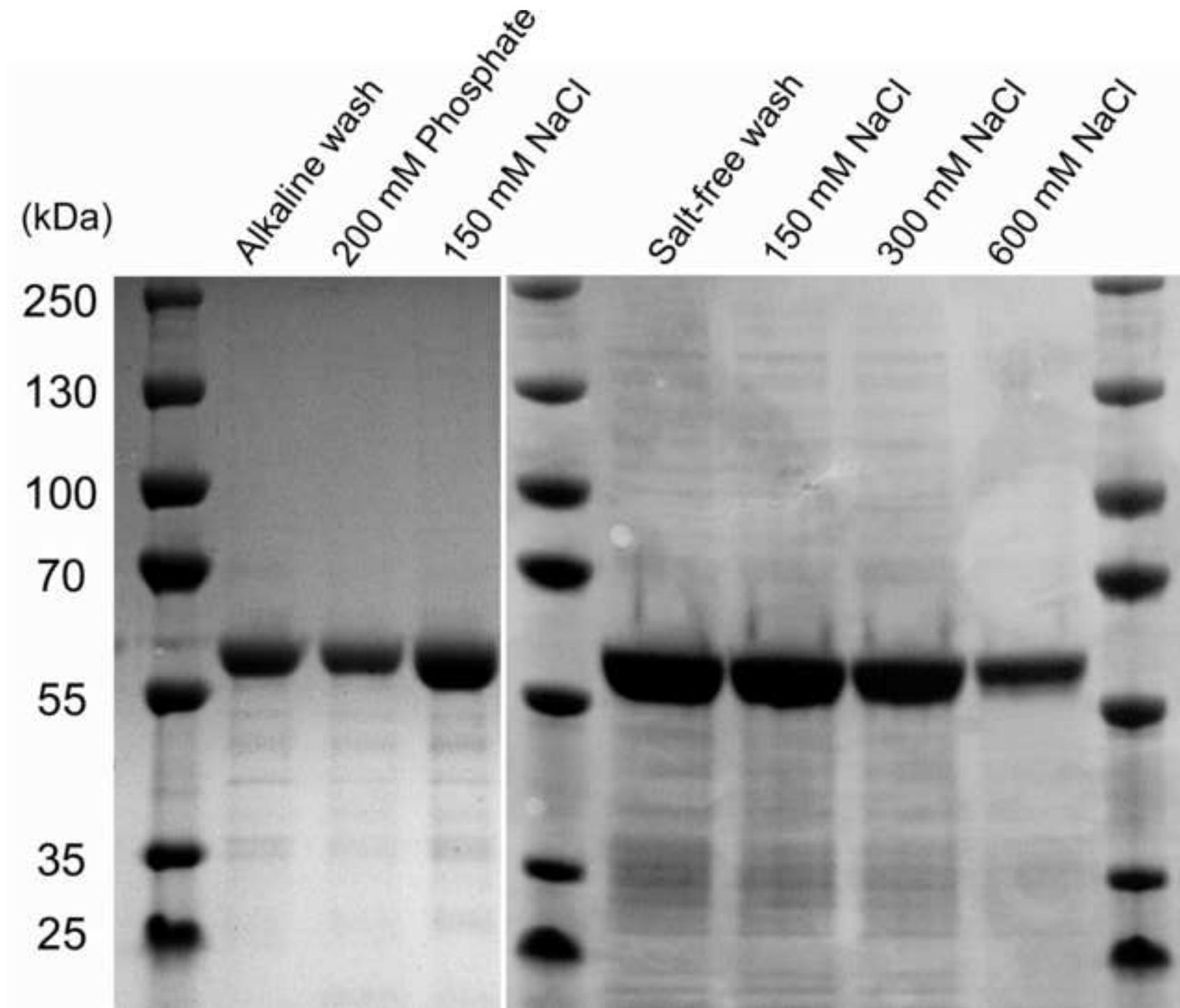


Figure 2  
[Click here to download high resolution image](#)



**Figure 3**  
[Click here to download high resolution image](#)

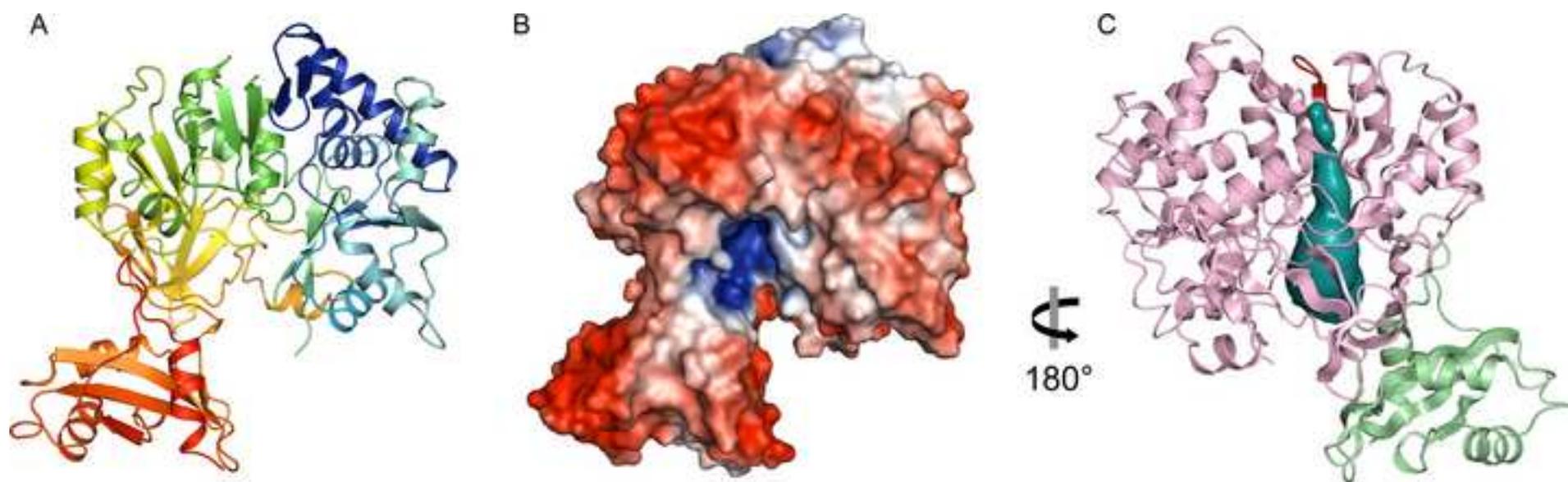


Figure 4  
[Click here to download high resolution image](#)

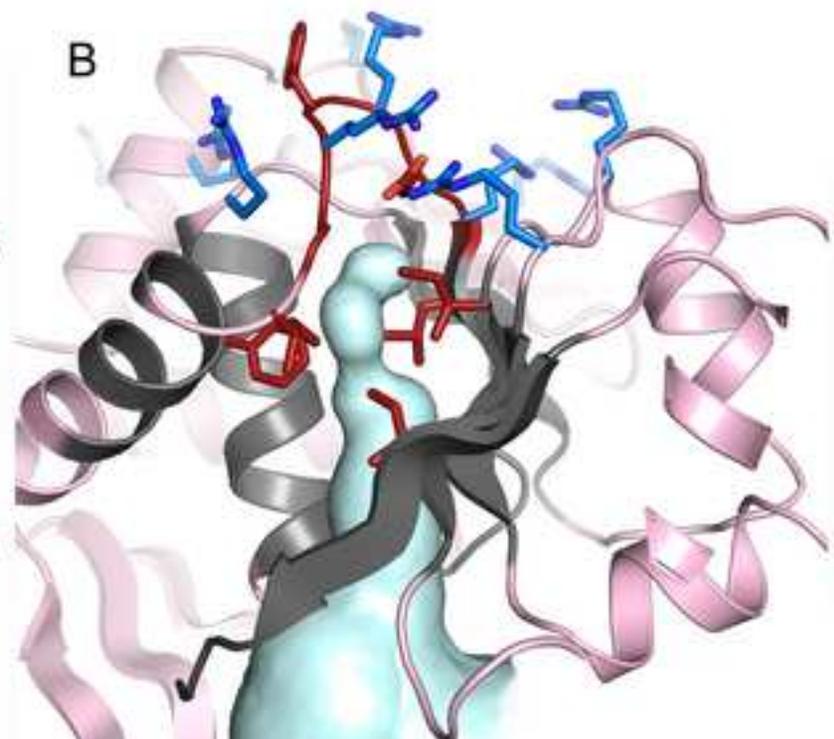
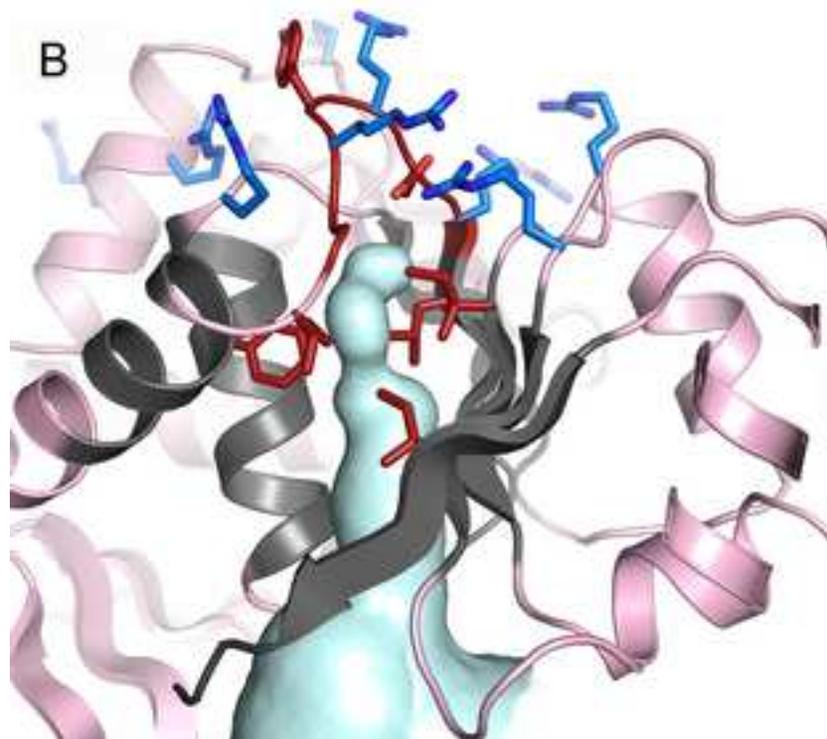
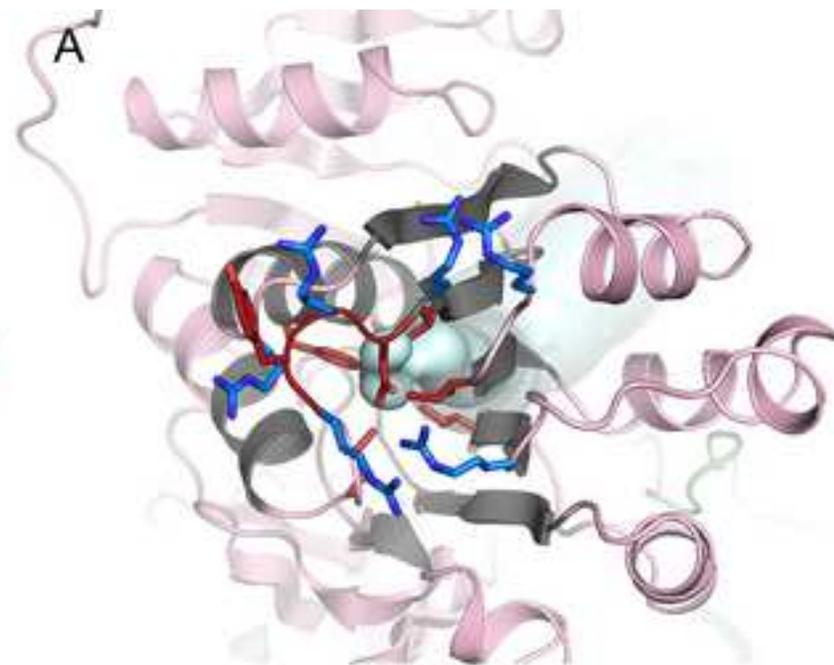
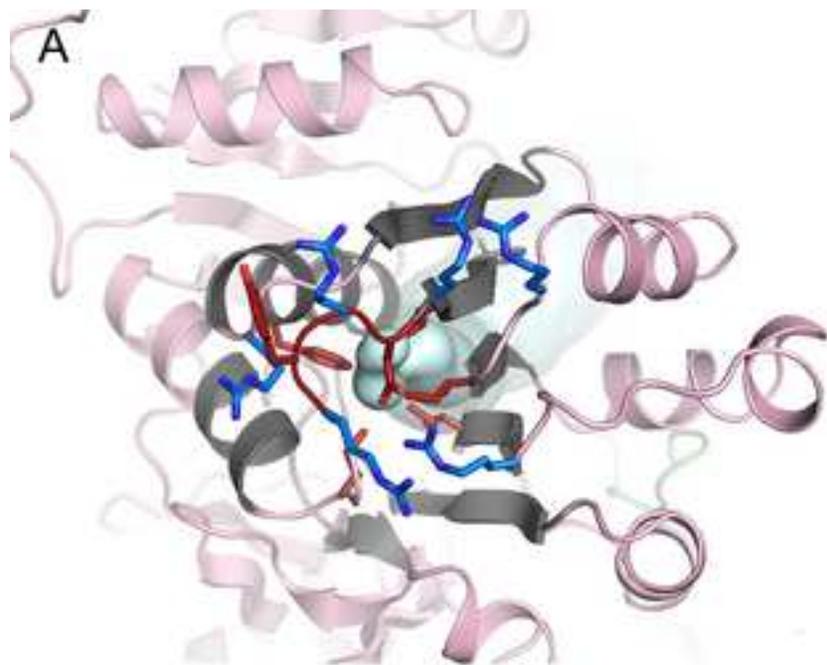
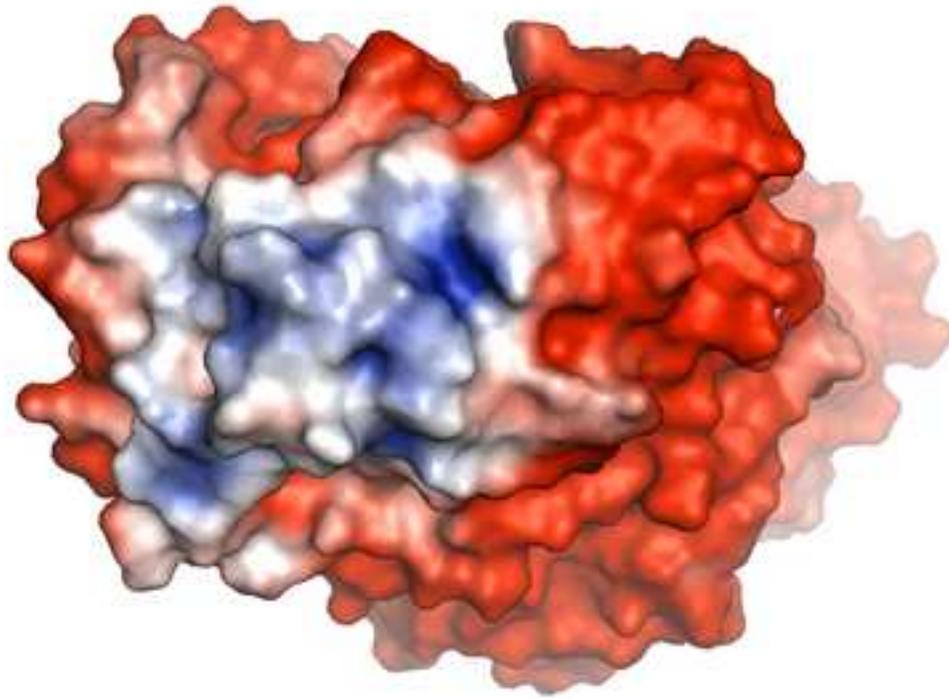


Figure 5  
[Click here to download high resolution image](#)

A



 90°

B

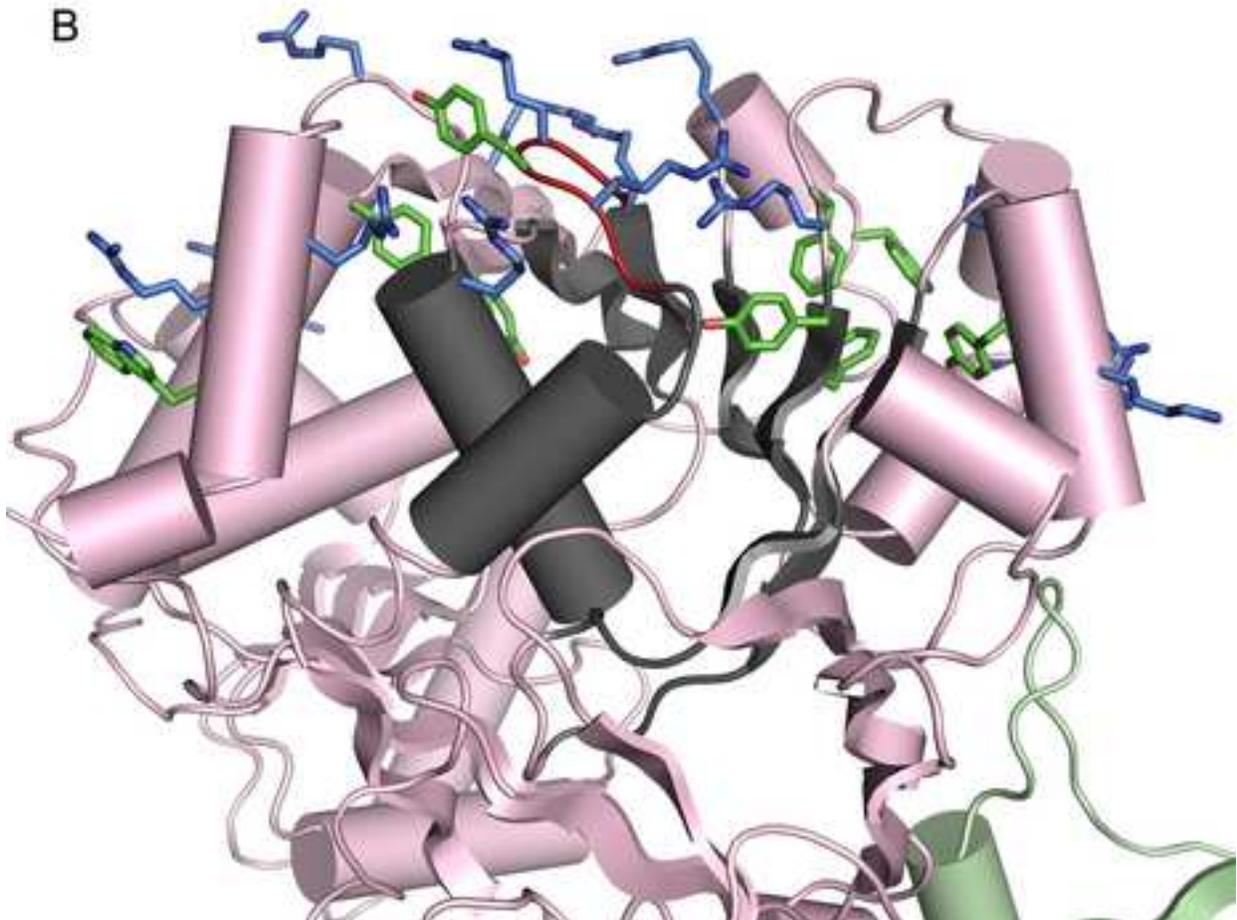


Figure 6  
[Click here to download high resolution image](#)

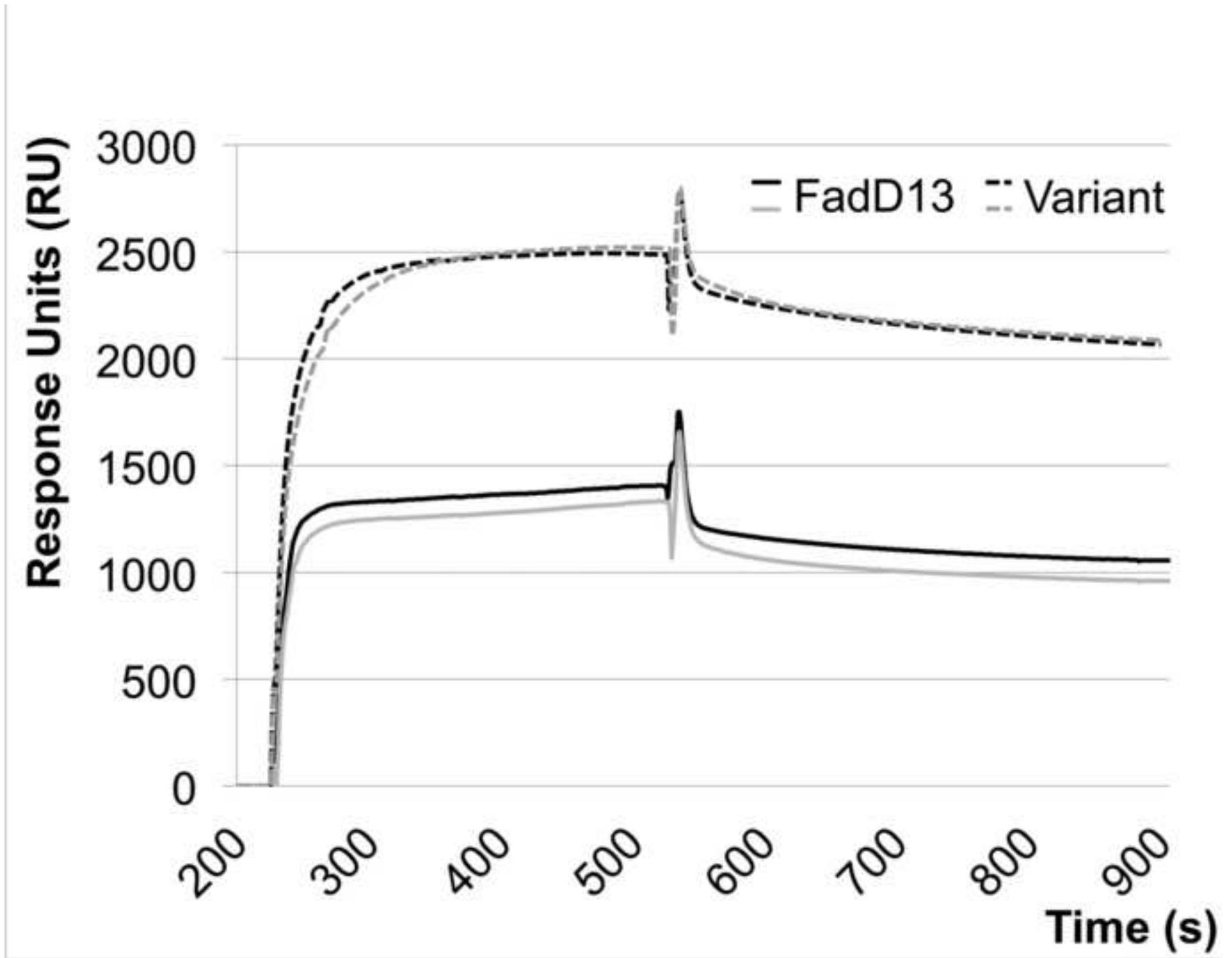
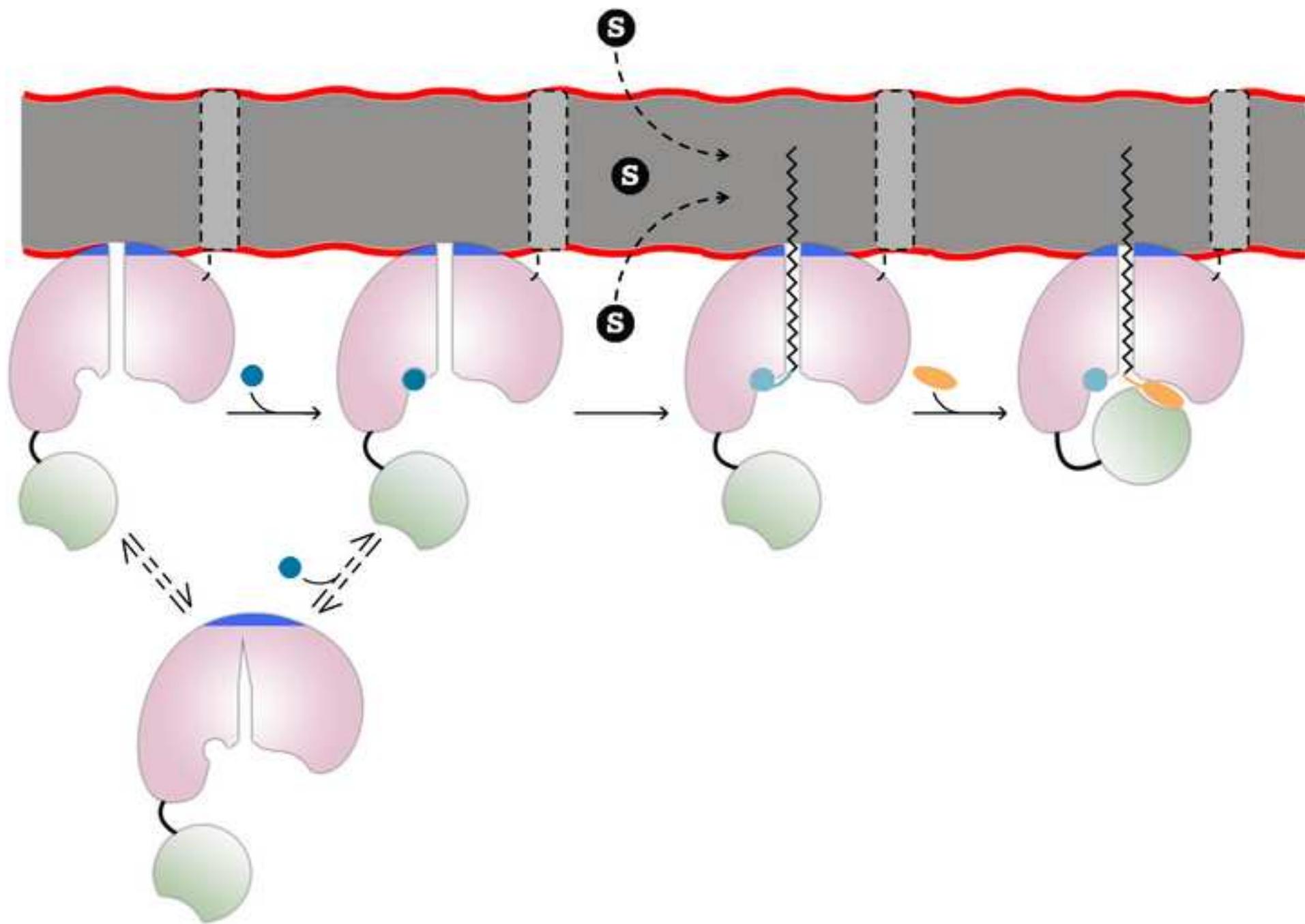


Figure 7  
[Click here to download high resolution image](#)



Supplemental information inventory

**The *Mycobacterium tuberculosis* very-long-chain fatty acyl-CoA synthetase - structural basis for housing lipid substrates longer than the enzyme**

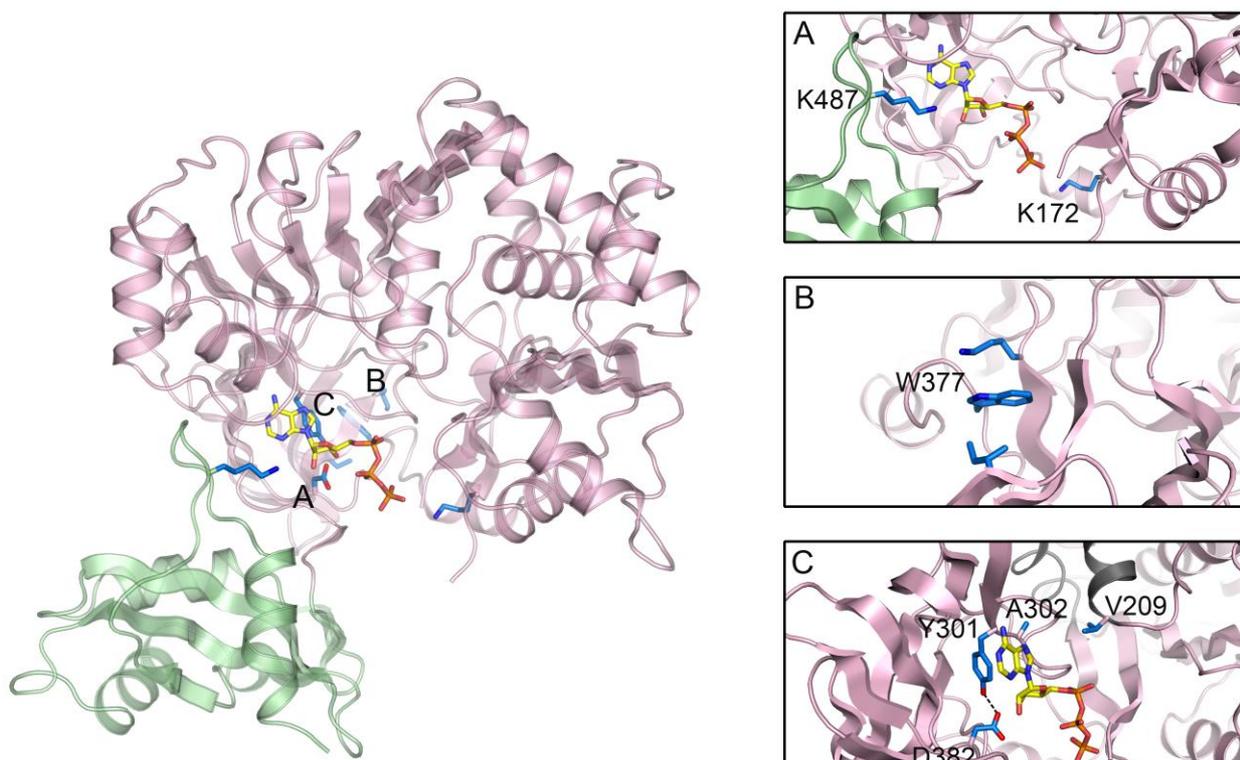
Charlotta S. Andersson, Camilla A. K. Lundgren, Auður Magnúsdóttir, Changrong Ge, Åke Wieslander, Daniel Martinez Molina and Martin Högbom

The following supplemental information is provided:

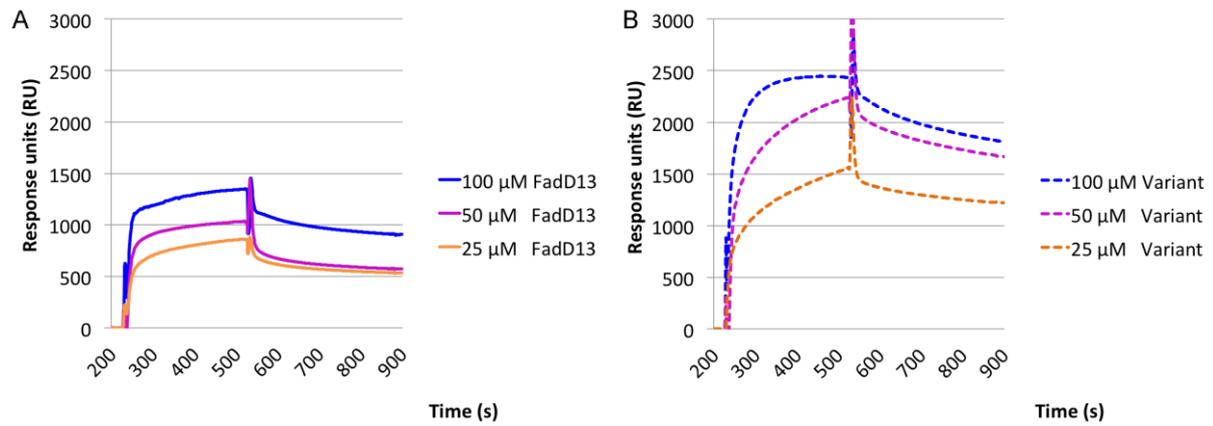
- Supplemental Figure 1, related to figure 3.
- Supplemental Figure 2, related to figure 6.

**The *Mycobacterium tuberculosis* very-long-chain fatty acyl-CoA synthetase - structural basis for housing lipid substrates longer than the enzyme**

Charlotta S. Andersson, Camilla A. K. Lundgren, Auður Magnúsdóttir, Changrong Ge, Åke Wieslander, Daniel Martinez Molina and Martin Högbom



Supplemental Figure 1, related to figure 3. Structural features of residues affecting activity.



Supplemental Figure 2, related to figure 6. Sensorgrams for: (A) wild-type FadD13 and (B) the R9A, R17A, R195A, R197A, R244A variant. Each protein is injected at 25, 50 and 100  $\mu\text{M}$  over the immobilized model-membrane surface. Sensorgrams show subtracted values using a BSA-coated reference surface.