Published in final edited form as:

ACS Infect Dis. 2016 August 12; 2(8): 579–591. doi:10.1021/acsinfecdis.6b00082.

Structure of the Essential *Mtb* FadD32 Enzyme: A Promising Drug Target for Treating Tuberculosis

Misty L. Kuhn^{†,‡}, Evan Alexander[‡], George Minasov[†], Holland J. Page[‡], Zdzislaw Warwzrak[§], Ludmilla Shuvalova[†], Kristin J. Flores[†], Daniel J. Wilson[§], Ce Shi[‡], Courtney C. Aldrich^{*,‡,§}, and Wayne F. Anderson^{*,†}

[†]Center for Structural Genomics of Infectious Diseases, Department of Biochemistry and Molecular Genetics, Northwestern University Feinberg School of Medicine, Chicago, Illinois 60611, United States

[‡]Department of Medicinal Chemistry, University of Minnesota, Minneapolis, Minnesota 55455, United States

[‡]Center for Drug Design, University of Minnesota, Minneapolis, Minnesota 55455, United States

[§]Department of Chemistry and Biochemistry, San Francisco State University, San Francisco, California 94132, United States

^{||}LS-CAT, Synchrotron Research Center, Northwestern University, Argonne, Illinois 60439, United States

Abstract

Mycolic acids are indispensable lipids of *Mycobacterium tuberculosis* (*Mtb*), the causative agent of tuberculosis (TB), and contribute to the distinctive architecture and impermeability of the mycobacterial cell envelope. FadD32 plays a pivotal role in mycolic acid biosynthesis by functionally linking fatty acid synthase (FAS) and polyketide synthase (PKS) biosynthetic pathways. FadD32, a fatty acyl-AMP ligase (FAAL), represents one of the best genetically and chemically validated new TB drug targets. We have determined the three-dimensional crystal structure of *Mtb* FadD32 in complex with a ligand specifically designed to stabilize the catalytically active adenylate-conformation, which provides a foundation for structure-based drug design efforts against this essential protein. The structure also captures the unique interactions of a

*Corresponding Authors aldri015@umn.edu., wf-anderson@northwestern.edu.

Author Contributions

M.L.K., H.J.P., C.C.A., and W.F.A. wrote the main article text. C.C.A. and E.A. prepared Figure 1, M.L.K. prepared Figures 2–7 and Table 1, H.J.P. prepared Figure 7, and G.M. prepared Figure 2 and Table 1. D.J.W. cloned the genes. E.A. and C.S. synthesized the inhibitor, purified PKS13 and FadD32 (for the biochemical studies), and performed the functional gel-based assay to measure inhibition of FadD32. K.J.F. purified the FadD32 protein for crystallization. M.L.K. crystallized the protein. L.S. and G.M. harvested the crystal and collected data at the synchrotron. G.M., Z.W., and W.F.A. determined the structure of the protein. All authors reviewed the article.

Notes

The authors declare no competing financial interest.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsinfecdis.6b00082.

¹H and ¹³C NMR spectra of Phu-AMS (PDF)

FAAL-specific insertion sequence and provides insight into the specificity and mechanism of fatty acid transfer.

Graphical abstract

Keywords

FadD32; *Mycobacterium tuberculosis*; fatty acyl-AMP ligase; mycolic acid biosynthesis

Mycobacterium tuberculosis (*Mtb*), the etiological agent of tuberculosis (TB), infects over one-third of humanity and is now the leading cause of infectious disease mortality by a single pathogen.¹ TB is a devastating disease characterized by a persistent cough, general malaise and fatigue, hemoptysis, and severe weight loss as the disease progresses. In medieval times, TB was referred to as consumption because of its ability to totally consume an individual's vitality. Most TB infections originate in the lungs via inhalation of *Mtb*-containing aerosolized droplets from an actively infected individual, but extrapulmonary involvement is also common. *Mtb* has extraordinary metabolic plasticity that allows it to endure under a wide variety of conditions and external stresses, it can evade immune surveillance, and it exists largely in granulomatous lesions, whose heterogeneous nature is believed to contribute to differential drug susceptibility.^{2–4} Consequently, TB is extremely difficult to treat relative to other bacterial infections, and chemotherapy for the simplest drug-sensitive TB requires a four-drug regimen composed of isoniazid, rifampicin, ethambutol, and pyrazinamide for 2 months, followed by a continuation phase of isoniazid and rifampicin for another 4 months. Multi-drug-resistant TB (MDR-TB) is defined as being resistant to both isoniazid and rifampin and requires the use of second-line TB drugs, which are generally less efficacious and have greater adverse effects. MDR-TB has been observed in all six WHO regions (Africa, the Americas, Southeast Asia, Europe, the Eastern Mediterranean, and the Western Pacific) and accounts for nearly a half million cases per year. The emergence and dissemination of extensively drug-resistant TB (XDR-TB) that possesses the MDR phenotype but is additionally resistant to two of the most important second-line drug classes represents a major global public health threat.⁵ To bring TB back under control, a coordinated effort will be required to develop improved diagnostics, effective vaccines, and new chemotherapeutic agents, ideally with new mechanisms of action that allow shortening the treatment duration and exhibit synergy with newer antitubercular drugs.

A hallmark of *Mtb* is its extraordinarily fortified waxy cell wall, whose major lipid components are the mycolic acids. These lipids are the longest in nature with chain lengths of up to 90 carbons and have a multitude of important structural and immunological functions.^{6–8} Because mycolic acids are critical to the survival of the bacterium and contribute to intrinsic resistance, their biosynthetic enzymes are attractive targets for drug development.^{7,9} Current drugs that target mycolic acid biosynthesis include isoniazid, ethionamide, isoxyl, and thioacetazone.⁹ Biosynthesis of the mycolic acids is accomplished through a combination of fatty acid synthase (FAS) and polyketide synthase (PKS) pathways. The involvement of a PKS in primary metabolism is unprecedented as all previously characterized PKS gene products are part of nonessential secondary metabolic pathways.¹⁰ A fatty acyl-AMP ligase encoded by the gene *fadD32* functionally links the FAS and PKS mycolate pathways in *Mtb* and is required for mycolic acid production.^{8,11–14}

FadD32 is one of 35 fatty acid adenylating enzymes (FadDs) in *Mtb* and a member of the fatty acyl-AMP ligase (FAAL) class of FadDs involved in lipid biosynthesis, which activate and transfer fatty acids onto PKS proteins for further construction to yield the highly diverse and complex lipids of *Mtb*.¹⁰ The second class of FadDs is the fatty acyl-CoA ligases (FACLs) involved in lipid degradation and remodeling. Whereas FACLs use coenzyme A (CoA) as a common substrate and therefore exhibit functional redundancy, each FAAL has a unique protein acceptor substrate, which for FadD32 is the PKS protein known as PKS13 that catalyzes the final biosynthetic step of the mycolic acids. FadD32 has been shown to prefer long-chain fatty acids as substrates and transfers the activated acyl-adenylate to the pantetheine moiety of the N-terminal acyl-carrier protein (ACP) domain of PKS13.^{8,13} The functional difference between the FAALs and FACLs is correlated with an ~20 amino acid insertion sequence within the N domain that is found only in FAALs and is thought to stabilize the presumptive closed acyl-adenylate enzyme conformation.^{15,16} The fidelity and efficiency of acyl chain transfer between an FAAL and its cognate PKS are likely governed by specific protein–protein interactions.¹⁷

Mtb FadD32 was first chemically validated as a drug target through the observation that adenosine 5'-dodecyl-phosphate (C12-AMP) specifically inhibited the enzyme activity and growth of mycobacteria.⁸ Additionally, *fadD32* and *pkc13* genes were found to be essential and highly conserved in mycobacteria.^{12,18} Further studies showed that the partial depletion of FadD32 in *Mtb* was bactericidal and also sensitized *Mtb* (2- to greater than 32-fold) to several antibiotics.¹⁹ Whole-cell phenotypic screening against *Mtb* identified diarylcoumarin compounds as additional FadD32 inhibitors whose on-target mechanism of action was confirmed by genetic, biochemical, and metabolic studies.^{20,21} Collectively, these studies establish FadD32 as a promising target for therapeutic development.

To provide a structural framework for structure-based drug discovery, we have determined the crystal structure of *Mtb* FadD32 in the adenylation conformation bound to a bisubstrate ligand. The 5'-O-[N-(11-phenoxyundecanoyl)-sulfamoyl]adenosine (PhU-AMS) ligand used for the structural studies was also shown to inhibit the FadD32-catalyzed loading of PKS13. Our structure reveals two new features that are particular to FadD32, including an insertion located in the C-terminal domain of the protein and the positioning of the C-terminal tail, which we hypothesize blocks the entrance for the pantetheine moiety of the

acyl carrier protein domain (ACP) of PKS13 during the adenylation reaction. Our motivation for determining the structure of the *Mtb* FadD32 protein was largely due to its importance in mycolic acid biosynthesis, its confirmed essentiality in mycobacteria, and the lack of structural information for this high-priority TB drug target. Despite significant efforts over the past several years, no group has been able to determine the structure of *Mtb* FadD32. Here we present the first structure of the *Mtb* FadD32 protein, which is critical for future structure-based therapeutic development. During the preparation and review of our article, two groups published full-length FadD32 ortholog structures from *M. smegmatis* (*Msm*) (PDB IDs 5D6J, 5EY8) and *M. marinum* (*Mma*) (PDB ID: 5EY9),^{22,23} and a third *Msm* structure was deposited into the PDB without an associated publication (PDB ID: 5ICR). After the presentation of our *Mtb* FadD32 structure, we will discuss how it compares to these other structures that have been very recently reported.

RESULTS AND DISCUSSION

Crystallization of FadD32 with PhU-AMS Inhibitor

Initial attempts to crystallize FadD32 were hampered by contamination with copurified GroEL protein and difficulties with growing crystals of suitable size and quality for data collection. We therefore synthesized a bisubstrate ligand to stabilize the protein and facilitate crystallization. FadD32 catalyzes the ATP-dependent ligation of its substrate fatty acid onto PKS13 via an acyl-adenylate intermediate (Figure 1). The key design element of the ligand is based on the observation that such acyl-adenylate species are typically tightly bound owing to the large number of interactions between the acyl-adenylate and both substrate-binding pockets (fatty acid and ATP). Replacement of the phosphate of the acyl-adenylate with the isosteric sulfamate moiety generates a chemically stable ligand, which can also be considered to be a dead-end inhibitor because it cannot undergo subsequent transfer onto PKS13. Guided solely by biochemical inhibition studies, we identified 5'-O-[N-(11-phenoxyundecanoyl)sulfamoyl]-adenosine (PhU-AMS) as the optimal ligand for FadD32. PhU-AMS has an excellent balance of potency, FadD selectivity, and solubility, with the latter property proving critical for crystallization trials. The ability of PhU-AMS to inhibit FadD32 is shown in Figure 1 using a functional assay that directly measures the loading of a fluorescently labeled fatty acid onto PKS13. The PhU-AMS inhibitor possesses optimal potency and selectivity toward FAAL enzymes, whereas simple inhibitors similar to C12-AMP²² are nonselective and inhibit all FadDs (FAALs and FACLS). Moreover, the C12-AMP and C20-AMP inhibitors described by Guillet and co-workers²² are not whole-cell active because they contain a highly polar charged phosphate moiety that precludes cellular penetration, whereas PhU-AMS employs a sulfamate isostere that leads to improved permeability and corresponding whole-cell activity. On the basis of these considerations, we selected PhU-AMS for crystallization trials.

Structure of the Full-Length *Mtb* FadD32

We determined the three-dimensional structure of *Mtb* FadD32 in complex with PhU-AMS (Figure 2). The protein crystallized in the orthorhombic P2₁2₁2₁ space group with one molecule in the asymmetric unit and allowed the collection of X-ray diffraction data to 2.25 Å resolution (Figure 2, Table 1, Methods). The FadD32 structure has a larger N-terminal

domain (residues 1–492) and a smaller, more flexible, C-terminal domain (residues 493–636) (Figure 2a) characteristic of class 1 members of the ANL (Acyl-CoA synthetases, NRPS adenylation domains, and Luciferase) superfamily of adenylation enzymes.^{24,25} Proteins from this family typically crystallize with the C-terminal domain in one of two conformations that correspond to either the formation of the acyl adenylate or acyl transfer to CoA or ACP and are designated either as the adenylation or thiolation conformation, respectively. The *Mtb* FadD32 FAAL protein crystallized in the adenylation conformation, which is common for most FAAL proteins. Exceptions include the *Mtb* FadD10²⁶ and *Legionella pneumophila*²⁷ FAAL proteins that show open or intermediate conformations of the C-terminal domain. Many FAAL structures have been determined in both adenylation and thiolation conformations, but no FAAL protein has been crystallized in the thiolation conformation.

The *Mtb* FadD32 structure is the first *Mtb* FAAL protein to show a complete picture of how an insertion motif and both N- and C-terminal domains interact. The core components of the *Mtb* FadD32 protein are structurally similar to all crystallized *Mtb* FadD proteins: FadD10,²⁶ FadD13,^{15,28} and the N-terminal domain of FadD28.^{15,16} Among the two *Mtb* FAAL FadDs structurally characterized, *Mtb* FadD10 does not contain the insertion motif characteristic of other *Mtb* FAAL proteins,²⁶ whereas the *Mtb* FadD28 structure lacks the entire C-terminal domain.^{15,16} Notable characteristics of the *Mtb* FadD32 structure include the following: (1) a P-loop motif conserved in a variety of ATP-binding proteins,²⁹ (2) a hinge region that connects and allows movement between the N- and C-terminal domains,²⁷ (3) an insertion motif that is a distinctive feature of FAAL proteins and has been shown to be important for dictating substrate specificity,¹⁵ (4) the conformation of the C-terminal tail of the FadD32 that we hypothesize is important for blocking the entrance to the pantetheine of ACP during the adenylation reaction, and (5) an insertion that is found only in FadD32 (Figure 3).

One of the most novel features of the *Mtb* FadD32 structure is the FAAL-specific insertion motif (Figure 3). *Mtb* FadD32 is the first structure to show direct interactions between the C-terminal domain of an *Mtb* FAAL protein and its FAAL-specific insertion motif. R375 from the FAAL-specific insertion motif forms direct H-bonds with the backbone oxygen of G585 and the side-chain oxygens of D506 from the C-terminal domain (Figure 3c). In FAAL proteins, this region typically has bulky side chains that would directly interfere with the insertion motif seen in FAAL proteins. The FAAL-specific insertion motif in FadD32 also forms interactions between itself and the N-terminal domain of the protein. These include T369, I371, E385, Q396, and R382 from the insertion motif and F481, Y486, and Q423 from the N-terminal domain (Figure 3d). The FAAL-specific insertion motif does not interact directly with the hinge region of the protein, but it fits snugly into a cleft created by the hinge region of the protein and may restrict the conformational flexibility of the C-terminal domain compared to other ANL superfamily enzymes (Figure 3e).

The C-terminal tail of the protein in the FadD32 structure is uniquely positioned relative to other FadD proteins (Figure 4). Direct H-bonding interactions occur between T635 of the C-terminal tail with R519 of the C-terminal domain and F633 of the C-terminal tail with R603 of the C-terminal domain. Alignment of the *Mtb* FadD32 structure with other ANL

superfamily enzyme structures containing a bound ligand (CoA or a bisubstrate CoA-AMP) in the active site shows that the C-terminal tail of the *Mtb* FadD32 protein obstructs access to the active site from this surface of the protein (Figure 4). Thus, we hypothesize that the conformation of the C-terminal tail must undergo rearrangement in order for the thiolation reaction to proceed. Other ANL enzyme family structures do not show the C-terminal tail of the protein in this orientation. This may be due to the tail region of the protein being disordered in the structure or the structure being altered because of the presence of a C-terminal polyhistidine tag in these ANL family structures.

Another distinctive feature of the *Mtb* FadD32 structure is the presence of an FadD32-specific insertion sequence. The function of this insertion is unclear; however, it resides in a visibly distinct region of the C-terminal domain of the protein and is not found in other *Mtb* FAAL proteins (Figure 3a). Because this insertion is unique to FadD32 and each FAAL has a specific partner, we hypothesize that it may be important for interactions with or the recognition of PKS13.

Ligand Interactions with FadD32

The presence of the PhU-AMS inhibitor in the structure of FadD32 locks the protein into the adenylation conformation. Well-defined density for the entire PhU-AMS inhibitor was seen within a long channel in the N-terminal domain of the protein (Figure 2c). The location of the inhibitor defines the binding location of the fatty acyl-adenylate. Hydrogen bonds were observed between PhU-AMS and FadD32 residues H238, D239, S321, S349, A353, D476, and K608, and hydrophobic interactions were observed between PhU-AMS and FadD32 residues V218, L222, G230, T244, L247, A248, L251, F285, A287, L318, E322, P323, Y350, G351, L352, L357, and I487 (Figure 2d). The placement of the ligand in FadD32 is similar to that seen for the structure of the *L. pneumophila* FAAL protein in complex with a C12 acyl adenylate.²⁷

The structure of *Mtb* FadD32 also showed two distinct regions of the protein that interact specifically with ligands other than PhU-AMS. Although these are not substrates or products of the catalyzed reaction, their locations and interactions may mimic biologically relevant ligands and are thus noted here. The first region is the channel where the pyrophosphate product of the adenylation reaction probably exits the active site. Here, two residues (R490 and K492) of the hinge region between the N- and C-domains and one residue (S195) in the P-loop coordinate a phosphate molecule and distort the conformation of the P-loop. Within the same channel, one poly(ethylene glycol) (PEG) molecule interacts with R126 from the N-terminal domain and T604 and S605 from the C-terminal domain (Figure 2e). The second region is in the vicinity of the CoA binding site in homologous FAAL structures where CoA enters the active site for the thiolation reaction. In the *Mtb* FadD32 structure, the protein binds two additional ligands: phosphate and glycerol (GOL). This phosphate molecule interacts with residues R265 and R266 in the N-terminal domain, and GOL interacts with a water molecule and residues R519 and V520 in the C-terminal domain near the C-terminal tail (Figure 2f). The GOL molecule is closer to the likely entrance site for the pantetheine moiety of ACP than the phosphate.

Comparison of Ortholog FadD32 Structures with *Mtb* FadD32

Crystallizing and determining the structure of the *Mtb* FadD32 protein has been extremely challenging for many groups around the world, including ours. As a result, mycobacterial FadD32 orthologs have been used as substitutes to gain a greater understanding of the structure/function relationship of *Mtb* FadD32. Most recently, a complete biochemical and biophysical characterization of four FadD32 proteins from *Mtb*, *Mma*, *Msm*, and *Corynebacterium glutamicum* (*Cgl*) has been performed, and structures of the *Msm* and *Mma* FadD32 proteins were determined.^{8,22,30} The *Cgl*, *Msm*, and *Mma* proteins have 39, 74, and 92% sequence identities to *Mtb*, respectively.²² Several structures of the *Msm* FadD32 protein have been determined within a few months of each other: (1) the N-domain alone (PDB ID: 5D6N),²³ (2) the full-length protein in complex with ATP and Mg²⁺ and in the presence of a *Saccharomyces cerevisiae* SUMO domain (PDB ID: 5D6J),²³ and (3) two structures of the full-length protein in complex with bisubstrate inhibitors (PDB ID: 5EY8, 5ICR).²² The structure of the full-length *Mma* protein was also determined in complex with a bisubstrate inhibitor (PDB ID: 5EY9).²² We performed a structural alignment between the *Mtb* FadD32 protein and other recently determined mycobacterial FadD32 ortholog structures and observed a relatively good alignment of the core protein structures. Variability between structures was primarily seen in the loops (Figure 5). The rmsd values from these alignments were 1.298 Å (5D6J; *Msm* full length with SUMO), 0.973 Å (5ICR chain A; *Msm* full length with inhibitor), 0.898 Å (5EY8 chain F; *Msm* full length with inhibitor), 0.697 Å (5D6N; *Msm* N-domain), and 0.560 Å (5EY9 chain A; *Mma* full length with inhibitor).

Overall, relatively minor differences were seen between *Mtb* and the other FadD32 mycobacterial structures (Figure 5). There were, however, four regions that varied among the structures, which included the N-terminus, the P-loop, a helix that lines the hydrophobic tunnel where the fatty acid of the substrate lies, and the C-terminal tail. The main difference seen in the N-terminus was better ordered in the other mycobacterial structures than in the *Mtb* protein. The opposite was largely true for the P-loop where all residues were present for the *Mtb* and *Msm* 5D6J structures but were not defined in the *Msm* 5ICR and 5EY3 and *Mma* 5EY9 structures. Slight conformational differences were observed in this loop for the *Mtb* compared to the *Msm* 5D6J structure, which is likely due to the presence of a bisubstrate inhibitor compared to ATP and Mg²⁺ alone. Significant movement of a histidine residue (H252 *Mtb*, H246 *Msm*, or H245 *Mma*) on a helix that is present at the end of the hydrophobic substrate tunnel was observed depending upon the type of substrate or bisubstrate inhibitor that was bound (Figure 5). For instance, the conformation of this histidine residue when a bisubstrate inhibitor that contained a phenyl ring at the end of the alkyl chain (*Mtb* 5HM3 and *Msm* 5ICR) was bound displayed a more open conformation compared to when the inhibitor did not have a phenyl ring at its end (*Msm* 5EY8 and *Mma* 5EY9). The slight movement of a phenylalanine residue (F255 *Mtb*, F248 *Msm*, or F247 *Mma*) in the *Mma* structure was also observed in the absence of a phenyl substituent at the end of the bisubstrate inhibitor. The most significant difference between the homologous structures was the placement of the C-terminal tail (Figure 5). The majority of the C-terminal tail was present in the *Mtb* and *Msm* (5ICR, 5D6J) structures but was not completely resolved in the *Mma* and *Msm* 5EY9 structures. The *Msm* 5D6J structure shows

a significant shift in the location of the C-terminal tail compared to other orthologs (*Mtb*, *Msm*, and *Mma*), which is likely due to the presence of the SUMO protein.

Rationale for FadD32 Membrane Localization

Although the length of the acyl chain of the PhU-AMS inhibitor is somewhat longer than a 14-carbon fatty acid, the native fatty acid is significantly longer than the protein can possibly accommodate and must protrude past the protein surface and into the membrane. This observation is similar to that for *Mtb* FadD13, where its fatty acid substrate is longer than the protein.²⁸ For the fatty acid to gain access to the active site of *Mtb* FadD32, F285 that stabilizes the phenyl ring on the bisubstrate inhibitor must rotate to provide this access. Specific residues in FadD13 were shown to be important to membrane binding,²⁸ and because similar residues are present in *Mtb* FadD32 (R40, K43, R47, Q225, R228, H253, R276, K277, and F285), we hypothesize that FadD32 also interacts with the membrane at this location to accommodate long-chain fatty acids (Figure 3a). *Mtb* FadD32 has been identified in the membrane fraction in several studies^{31–33} and in both the membrane fraction and cytoplasm in one study.³¹

Domain Interactions between Adenylation and Thiolation Conformations

For the activated fatty acid to be transferred to PKS13, a conformational change in the FadD32 C-domain to the thiolation conformation must occur. This conformational change has to allow the pantetheine of ACP to access the active site. In general, it has been more challenging to obtain structures of the thiolation conformation of FAALs because the thiolation conformation is likely stabilized by the ACP domain.²⁴ The ACP–ANL enzyme interface is very large, and the pantetheine cofactor, which inserts into the ANL enzyme active site, also contributes to binding. In the case of FAALs, the cognate ACPs are separate proteins that are embedded in large multidomain proteins such as PKS13 (186 kDa). Moreover, the unique insertion sequence present in FAALs also serves to stabilize the adenylation conformation as demonstrated functionally by Gokhale and co-workers¹⁶ and corroborated by the structural studies described herein. Both of these factors likely work together to favor crystallization in the adenylation conformation.

Because we were unable to obtain crystals of the *Mtb* FadD32 protein in the thiolation conformation, we modeled the FadD32 thiolation conformation on the basis of the structure of a homologous 4-chlorobenzoyl-CoA ligase enzyme (PDB ID: 3CW9)³⁴ (Figure 6). We selected the 3CW9 structure to model the FadD32 thiolation conformation for several reasons. First, the 4-chlorobenzoyl-CoA ligase protein has been crystallized in both adenylation (PDB ID: 3CW8) and thiolation conformations (PDB ID: 3CW9), so we have evidence for the relative orientation of the C-domain in the same protein for both conformations. Second, the interactions formed between the N- and C-domains of the 3CW8 and FadD32 structures in the adenylation conformation were similar, and the thiolation conformation of the C-domain of the homologous 4-chlorobenzoyl-CoA ligase would provide access to the pantetheine of ACP of PKS13 to approach the FadD32 active site. Therefore, it is reasonable to presume that the thiolation conformation of FadD32 will likely participate in interactions common to the 3CW9 structure. The adenylation structure and the modeled thiolation conformation of *Mtb* FadD32 show that during the adenylation reaction

the entrance to the active site is restricted by the conformation of the C-terminal domain (Figure 6b) and during the thiolation reaction the C-terminal domain rotates to allow access to the active site for the PKS13 ACP domain to approach (Figure 6d).

We analyzed the specific interactions formed and disrupted between adenylation and thiolation conformations of the homologous 4-chlorobenzoyl-CoA ligase and compared them to the Mtb FadD32 crystallized and modeled proteins (Figure 7). In the 4-chlorobenzoyl-CoA ligase, residues in the hinge (N411 and H413) H-bond to E350 in the adenylation conformation (3CW8), whereas in the thiolation conformation (3CW9) this H-bond interaction was disrupted by the large rotation of the C-domain away from the location of E350 on the N-domain. The corresponding residues in FadD32 are N501, Y503, and I424, respectively. However, no H-bonds were seen between these residues in the crystal structure. The presence of the FAAL-specific insertion motif in FadD32 shifts the position of these secondary structural elements compared to the 3CW8 structure. Another key residue for the transition between adenylation and thiolation conformations is K89 in the 4-chlorobenzoyl-CoA ligase enzyme. In the 3CW8 structure, this residue does not form specific interactions, but when the protein undergoes the transition to the thiolation conformation (3CW9), it forms a set of H-bonds with the C-domain that includes residues R466, D472, S468, and L470. The corresponding residue in the FadD32 protein appears to be H123, which in the thiolation conformation may form interactions with residues in the region between R579 and V588. The interactions that would need to form between the N- and C-domains in the thiolation conformation for the FadD32 protein are not completely clear because the helix may be flexible and could unwind to create the specific H-bonds. In the adenylation conformation of the protein, the region of FadD32 where H123 resides is on a longer flexible loop than K89 in the 3CW8 structure and forms direct and more extensive interactions with the C-terminal tail of the FadD32 protein. The lack of interactions between K89 and the C-terminal tail of the 4-chlorobenzoyl-CoA ligase protein may be due to its expression with a C-terminal polyhistidine tag, which was not removed prior to crystallization.^{34,35}

Although specific interactions between the N- and C-domains of the FadD32 protein are unknown for the thiolation conformation, we propose that the following interactions can be predicted on the basis of the 3CW9 structure and modeled FadD32 protein in the thiolation conformation. First, the H-bonding interactions between K608 and the bisubstrate inhibitor in the adenylation conformation will be disrupted during the transition to the thiolation conformation. As a result, a helix and loop region composed of residues G607–L619 (red and orange in the figure) may reorient to form interactions with the FAAL-specific insertion motif to anchor the C-terminal domain in a distinct conformation for acyl transfer to PKS13 (Figure 7).

Implications for Drug Discovery Efforts

Mtb encodes for 35 FadD fatty acid adenyating enzymes, and prior to this study, only three *Mtb* structures had been determined: FadD10, FadD13, and the N-terminal domain of FadD28.^{15,16,26,28} Because FadD32 has been identified as a viable candidate drug target to treat tuberculosis, elucidating its structure was important in improving our functional

understanding of this enzyme. In addition to bisubstrate analog inhibitors of FadD32, coumarin compounds have been identified as antimycobacterial compounds.²¹ Residues that were found to confer resistance to the coumarin compounds in FadD32 are E120 and F291. These residues are located at the interface between the N- and C-domains of the protein at the entrance to the pantetheine tunnel adjacent to the C-terminal tail (Figure 4d). Because the coumarin compounds were found to inhibit transfer of the acyl group to the PKS13 ACP domain but did not inhibit the FadD32 adenylation reaction,²¹ it is plausible that they block the entrance to the phosphopantetheine tunnel used by PKS13 ACP to approach the active site or prevent the transition to the thiolation conformation.

The structure of the *Mtb* FadD32 enzyme in complex with PhU-AMS, an analog of the acyl adenylate intermediate, provides important new information that will be useful for structure-guided drug discovery efforts. The adenine nucleobase is recognized by hydrogen bonds with the side chain of S349 and backbone of S321. There is some space adjacent to the C-2 position of the nucleobase indicating that small nonpolar groups may be tolerated at this position. (See Figure 1C for atom numbering of the ligand.) D476 interacts with the 2'- and 3'-hydroxyls of the ribose, and each hydroxyl also forms a hydrogen bond with a water molecule, suggesting that both alcohols are important for binding, but replacement with isosteric fluorines, which are known to impart improved physicochemical properties, may be tolerated. Although the binding interactions of PhU-AMS and the related C12-AMP ligand reported by Guillet and co-workers²² are highly conserved in the nucleoside moiety of both ligands, the interactions with the alkanoyl side chain of PhU-AMS diverge from the linear alkyl side chain of C12-AMP. K608 forms a tripartite interaction with the ribofuranose 4' and 5'-oxygen atoms as well as the C-1'' carbonyl moiety of PhU-AMS stabilizing the bent compact conformation of the ligand. C12-AMP, which does not contain a carbonyl group, lacks this crucial interaction. The presence of the carbonyl lowers the pK_a of the nitrogen atom of the acyl-sulfamide linker to ~3 so that it is likely in the deprotonated amide form and thus interacts with the corresponding protonated form of the side chain of D239. These interactions imply that PhU-AMS more closely mimics the transition state of FadD32 when forming the acyl-adenylate compared to the C12-AMP ligand. The terminal phenyl ring in PhU-AMS additionally participates in more hydrophobic interactions than C12-AMP because it remains buried in the hydrophobic binding pocket. The lack of polar interactions within the fatty acid binding pocket suggests that the introduction of heteroatoms within the alkyl chain will be poorly tolerated. Finally, the structure reveals positions within the alkyl chain where unsaturation or cyclopropanation may constrain the conformation of the molecule to further improve the specificity and membrane permeability by reducing the number of rotatable bonds. Taken together, the cocrystal structure of PhU-AMS and C12-AMP with Fad32 reveals a constellation of hydrogen bonds and extensive hydrophobic interactions that provide numerous opportunities to enhance inhibitor selectivity and potency.

Our results provide structural evidence as to how the FAAL-specific insertion motif, which dictates substrate specificity, in *Mtb* FAAL proteins interacts with and restricts the movement of the C-terminal domain and stabilizes the adenylation conformation of the protein compared to FACL proteins. Because the transfer of activated acyl chains is specific between FAALs and their cognate PKS proteins, the FadD32-specific insertion motif may be

important in the recognition of PKS13. A defined protein–protein interaction between FadD32 and PKS13 could trigger the transition of FadD32 from the adenylation to thiolation conformation. During this process, the FadD32-specific insertion may interact with PKS13 and the FAAL-specific insertion motif may anchor FadD32 in a strict position for acyl transfer to ACP. Collectively, the FadD32 structures provided insight into the molecular details of the FadD32-specific insertion motif, which had not been possible previously. This insertion, in addition to the placement of the C-terminal tail, will likely be important for protein–protein interactions with PKS13 and elucidating the transition from the adenylation to thiolation conformation of FadD32 upon PKS13 binding. We speculate that the entire assembly of the activated fatty acid and FadD32 and PKS13 proteins is localized at the membrane for the efficient threading and transfer of the acyl chain through the FadD32 and PKS13 proteins and back into the membrane to form the mature mycolic acid.

METHODS

Cloning

The *fadD32* gene (locus tag Rv3801c) and *pkc13* gene from *M. tuberculosis* H37Rv (locus tag Rv3800c) were cloned into the pET28b and pET24b vectors (Novagen), which produce proteins that have either an N-terminal polyhistidine tag followed by a thrombin cleavage site or a C-terminal polyhistidine tag, respectively. The *fadD32* gene was amplified from genomic DNA using standard PCR methods with the following two primers CACCCATATGTTTGTGACAGGAGAG and GGAAGCTTTCAGTCCGAAGTGGCGAAGAC. The product was cloned into pENTR/TEV/D-TOPO (Life Technologies) using the manufacturer's instructions and sequenced to confirm the absence of PCR errors. pENTR/TEV-fadD32 was digested using *NdeI* and *HindIII* (sites underlined in the primer sequence above) and ligated into pET28b cut with the same enzymes to produce pET28b-fadD32.

The plasmid Litmus 28 (NEB) carrying the *pkc13* gene was kindly provided by Prof. Rajesh Gokhale, CSIR-Institute of Genomics and Integrative Biology. The gene was recloned by digestion with *NdeI* and *EcoRI* into pET24b cut with the same enzymes.

Protein Expression and Purification

FadD32—The kanamycin-resistant pET28b-fadD32 plasmid was cotransformed with the chloramphenicol-resistant pGRO7 plasmid (Takara) into BL21(DE3) cells for protein expression. The cotransformed cells were used to inoculate 15 mL cultures of LB containing 35 µg/mL kanamycin and 25 µg/mL chloramphenicol. The cultures were grown overnight at 37°C with shaking at 200 rpm. The next day, the starter culture was used to inoculate 1.5 L of selenomethionine (SelMet) medium (6 g/L Na₂HPO₄, 3 g/L KH₂PO₄, 1 g/L NH₄Cl, 0.5 g/L NaCl, 1 mM MgSO₄, 0.1 mM CaCl₂, 12.5 g/L Fe₂SO₄, 0.4% glucose, 0.00005% thiamine, 0.00005% Biotin) in plastic 2 L bottles (0.5 L per bottle) as described previously.³⁶ The medium was supplemented with 25 µg/mL chloramphenicol, 35 µg/mL kanamycin, and 0.5 mg/mL L-arabinose. Cells were grown at 37 °C to an OD_{600 nm} of 1.2 with shaking at 200 rpm. Once cells reached the appropriate OD, amino acid supplements (100 mg/L L-lysine, L-phenylalanine, and L-threonine, and 50 mg/L L-isoleucine, L-leucine,

L-valine, and L-SelMet) were added to the cultures and incubated at RT for 15 min. Protein expression was induced with 1 mM IPTG at 25 °C overnight with shaking at 200 rpm.

Cells were harvested by centrifugation as described previously,³⁶ resuspended in lysis buffer (1.5 mM Mg-(C₂H₃O₂)₂, 1 mM CaCl₂, 100 mM (NH₄)₂SO₄, 40 mM Na₂HPO₄, 3.25 mM citric acid, 5 mM imidazole, 0.08% (w/v) *n*-dodecyl β-D-maltoside (DDM), 0.5 mM TCEP, 100 mM KCl, and 10% glycerol), and sonicated. The cleared lysate was loaded onto a 5 mL HisTrapFF column (GE Healthcare) equilibrated with loading buffer (10 mM Tris-HCl pH 8.3, 100 mM KCl, 0.5 mM TCEP, and 5% glycerol) at 4 °C using an AKTExpress system (GE Healthcare). Unbound sample was washed out of the column using loading buffer plus 25 mM imidazole. Protein was eluted with loading buffer plus 500 mM imidazole and subsequently loaded onto a HiPrep200 26/60 gel-filtration column (GE Healthcare) equilibrated with loading buffer. The affinity tag was not removed.

After gel-filtration chromatography, the protein eluted as two separate peaks. The first peak of *Mtb* FadD32 protein was contaminated with GroEL protein, which was necessary for soluble protein expression, and the second peak contained only the FadD32 protein (data not shown). We determined that the first peak of protein contained GroEL because during our crystallization trials we crystallized GroEL instead of the FadD32 protein. On the basis of these results, subsequent crystallization trials included only protein from the second peak after gel-filtration chromatography.

PKS13—The kanamycin-resistant pET24b-pks13 plasmid was cotransformed with the chloramphenicol-resistant pDHS4032 plasmid (pLysS lac UV5 promoter for constitutive expression) encoding the phosphopantetheinyltransferase Sfp³⁷ (kindly provided by Prof. David Sherman, Life Sciences Institute, University of Michigan) into BL21(DE3) cells for protein expression. The cotransformed cells were used to inoculate 5 mL of LB containing 50 μg/mL kanamycin and 30 μg/mL chloramphenicol. The cultures were grown overnight at 37 °C with shaking at 200 rpm. The next day, the starter culture was used to inoculate 1 L of ZYM-5052 medium³⁸ containing 50 μg/mL kanamycin and 30 μg/mL chlorophenicol at 20 °C for 72 h at 300 rpm.

Cells were harvested by centrifugation as described previously,³⁶ resuspended in lysis buffer (50 mM HEPES, 300 mM NaCl, 10% v/v glycerol, pH 7.9), and sonicated. The cleared lysate was loaded onto a 5 mL HisTrapHP column (GE Healthcare) equilibrated with loading buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.5 mM TCEP) at 4 °C using an AKTExpress system (GE Healthcare). Unbound sample was washed out of the column using loading buffer plus 10 mM imidazole. Pks13 was eluted using a linear gradient from 10 to 500 mM imidazole in loading buffer over 20 min. The fractions were pooled and concentrated using a 30 kDa centrifugal filter. The concentrated protein was loaded onto a HiPrep200 26/60 gel-filtration column (GE Healthcare) equilibrated with running buffer (50 mM sodium phosphate buffer, pH 7.1, 150 mM NaCl). The affinity tag was not removed. The peak corresponding to full-length holo-Pks13 was pooled, exchanged in storage buffer (20 mM Tris, 150 mM NaCl, 10% v/v glycerol, pH 8.0), and concentrated using a 30 kDa centrifugal filter.

Crystallization of FadD32

The NeXtal CubicPhase Kit (Qiagen) was used to obtain crystals of the apo-FadD32 SelMet protein for further seeding trials. We followed the manufacturer's protocol for the manual setup of a lipidic cubic-phase experiment using a multichannel pipet with the Cubic Phase I and Cubic Phase II Suites. Briefly, a NeXtal CubicPhase microplate containing monoolein was warmed to RT, and 1 μL of 6 mg/mL FadD32-SelMet protein was pipetted directly onto the dried monoolein. The plates were sealed and allowed to incubate at RT for 3 h.

Afterward, 70 μL of the screening solutions was added to the reservoir wells of the plates and 0.8 μL of a 1:7 dilution of each screening solution was pipetted into the protein well of the plate. The plates were sealed and incubated at 22 °C. After 2–4 days, clusters of needles grew under Cubic Phase I Suite conditions A10, B8, B10, B11, C5, C10, C11, D5, D8, D11, D12, E5, E8, E10, E11, E12, F8, F10, F11, G8, G11, and H11. These conditions contained primarily 2.4 M Na/K phosphate at varying pH or 0.1 M Hepes, Tris, MES, or sodium acetate at varying pH and a high concentration of $(\text{NH}_4)_2\text{SO}_4$ (1.4–2.6 M). No crystals were suitable for data collection, so we used them as seeds for optimization trials.

On the same day as the monoolein plates were set, we also screened the Cubic Phase I and II suites in the absence of monoolein. We set screens with the FadD32 SelMet protein (6 mg/mL) in the presence of 0.1 mM PhU-AMS inhibitor. The reservoir contained 80 μL of screening solution and the protein well contained 1 μL of protein and 1 μL of a 1:2 dilution of the reservoir. The plates were allowed to incubate at 22 °C for 5 days and then were cross-seeded with a cat whisker using seeds from the G11 Cubic Phase I suite apo-fadD32 monoolein plates (0.1 M Tris pH 8.2, 2.6 M $(\text{NH}_4)_2\text{SO}_4$). The newly seeded plates were incubated at 22 °C. A suitable crystal for data collection grew approximately 2 weeks after cross-seeding in 2.4 M Na/K phosphate at pH 7.6. The crystal was harvested, transferred to a solution containing 2.4 M Na/K phosphate at pH 7.6 and 25% sucrose, and flash frozen in liquid nitrogen.

Data Collection and Refinement

Data were collected at Life Science Collaborative Access Team (LS-CAT) beamline 21ID-D at the Argonne National Laboratory Advanced Photon Source (APS). Six hundred images were collected from a rodlike single crystal at the selenium peak ($\lambda = 0.97875$), which represented 360° of the spindle rotation. After careful analysis of the images, two data sets were generated: data set 1 with highly redundant data for phasing (430 of the 600 images were scaled and merged) and data set 2 for model refinement (only 200 images were scaled and merged) (Table 1). Data was indexed, integrated, and scaled with HKL-3000.³⁹ PHENIX⁴⁰ was used for phasing, and the initial model was checked and manually corrected using Coot.⁴¹ The corrected model was refined against data set 2 in REFMAC.⁴² After several rounds of refinement and manual correction, solvent molecules were added to the model using ARP/wARP,⁴³ and the inhibitor was fitted into the density followed by rounds of refinement in Refmac and manual corrections in Coot. The final model was checked and validated using SFCHECK⁴⁴ and MolProbity.⁴⁵ Crystals belonged to the orthorhombic space group $P2_12_12_1$ with unit cell dimensions *a*, *b*, and *c* as 62.78, 74.56, and 127.76, respectively, with all angles equal to 90°. The asymmetric unit accommodates one polypeptide chain. The data were scaled to 2.25 Å resolution, and statistics for the final

model are listed in Table 1. The first 43 residues of the N-terminus are quite mobile, and the following residues of the protein were disordered: MGSSHHHHHHSSGLVPRGSH (residues from the tag and cleavage site), ESGMAYH (residues 6–12), and PAN (residues 24–26). The structure factors and coordinates of the final model were deposited into the Protein Data Bank (PDB) with accession code 5HM3. All figures were produced using the PyMOL Molecular Graphics System, version 1.5.0.4 software (Schrodinger LLC).

Synthesis of 5'-O-[N-(11-Phenoxyundecanoyl)-sulfamoyl]adenosine (PhU-AMS)

To a solution of 11-phenoxyundecanoic acid (181 mg, 0.65 mmol, 1.0 equiv) in CH₂Cl₂ (10 mL) at 23 °C was added *N*-hydroxysuccinimide (75 mg, 0.65 mmol, 1.0 equiv) and *N,N'*-dicyclohexylcarbodiimide (134 mg, 0.65 mmol, 1.0 equiv). The reaction mixture was stirred for 16 h, filtered, and concentrated under reduced pressure to provide the *N*-hydroxysuccinimidyl ester of 11-phenoxyundecanoate that was used without further purification.

To a solution of the resulting *N*-hydroxysuccinimidyl ester prepared above (240 mg, 0.65 mmol, 2.0 equiv) in DMF (4 mL) were added *N*⁶,*N*⁶-bis(*tert*-butoxycarbonyl)-5'-O-sulfamoyl-2',3'-O-isopropylideneadenosine⁴⁶ (189 mg, 0.32 mmol, 1.0 equiv) and cesium carbonate (313 mg, 0.96 mmol, 3.0 equiv) at 0 °C. The mixture was warmed to room temperature and stirred for 16 h. The resulting mixture was concentrated in vacuo, resuspended in 10:1 MeOH-CH₂Cl₂ (20 mL), and filtered over Celite. The filtrate was concentrated under reduced pressure to afford crude 5'-O-[*N*-(11-phenoxyundecanoyl)sulfamoyl]-2',3'-O-isopropylideneadenosine that was used without further purification.

The crude product prepared above was stirred in 4:1 trifluoroacetic acid-H₂O (5 mL) at 0 °C for 3 h. The mixture was concentrated in vacuo and purified by flash chromatography on silica gel, eluting with 79:20:1 EtOAc-MeOH-Et₃N to afford the title compound as the montriethylammonium salt (145 mg, 64%): HPLC purity, 99.86%; *t*_R = 10.44 min, *k'* = 2.2 (Phenomenex Gemini C18, 250 × 4.6 mm², 5 μm, 110 Å, flow rate = 1 mL/min, detection at 254 nm, 50 mM aqueous triethylammonium bicarbonate at pH 7.5 (solvent A) and 100% acetonitrile (solvent B), with a linear gradient from 5% B to 70% B over 10 min and then 70% B to 90% B from 10 to 12 min, followed by an isocratic elution with 90% B from 12 to 16 min). *R*_f = 0.29 (4:1 EtOAc-MeOH). ¹H NMR (600 MHz, CD₃OD) δ 1.27–1.30 (m, 21H), 1.43 (p, *J* = 7.2 Hz, 2H), 1.59 (p, *J* = 7.2 Hz, 2H), 1.73 (p, *J* = 7.2 Hz, 2H), 2.19 (t, *J* = 7.2 Hz, 2H), 3.18 (q, *J* = 7.2 Hz, 6H), 3.93 (t, *J* = 7.2 Hz, 2H), 4.26–4.28 (m, 2H), 4.31–4.34 (m, 1H), 4.38–4.40 (m, 1H), 4.69 (t, *J* = 6.0 Hz, 1H), 6.08 (d, *J* = 6.0 Hz, 1H), 6.85–6.88 (m, 3H), 7.23 (t, *J* = 7.2 Hz, 2H), 8.20 (s, 1H), 8.50 (s, 1H). ¹³C NMR (150 MHz, CD₃OD) δ 9.4, 27.3, 27.6, 30.55, 30.62, 30.67, 30.72, 30.77, 30.79, 40.4, 48.0, 69.0, 69.3, 72.5, 76.2, 84.7, 89.3, 115.6 (2C), 120.3, 121.6, 130.5 (2C), 141.3, 151.0, 154.0, 157.4, 160.7, 183.3. HRMS (ESI) calculated for C₂₇H₃₇N₆O₈S [M – H][–] 605.2399, found 605.2391 (error 1.3 ppm).

Biochemical Activity Assay for FadD32

Reactions were performed in triplicate under initial velocity conditions in a total volume of 15 μL with 2% DMSO. A seven-point 2-fold dilution series of PhU-AMS was prepared to provide final assay concentrations of 50 μM down to 0.78 μM . A positive control (DMSO only) and three negative controls (50 mM EDTA, (–) FadD32, (–) PKS13) were also included in the experiment. The final reaction mixtures consisted of 32 nM FadD32, 50 μM 4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene-3-hexadecanoic acid (BODIPY FL C₁₆, Invitrogen, catalog no. D-3821), 2 mM ATP, and 2 μM Pks13 in assay buffer (50 mM HEPES, 5 mM MgCl₂, 1 mM DTT, 0.01% Tween-20, pH 7.2), and varying amounts of inhibitors or controls. The reactions were preincubated for 5 min and then initiated with ATP and incubated for 2 h at 30 °C. The reaction were quenched with 2 $\times$ SDS (15 μL), and 15 μL of the reaction mixtures was loaded onto a mini-protean TGX precast gel and a 4–20% precast polyacrylamide gel (Biorad) using 2.5 mM Tris, 19.2 mM glycine, 0.01% SDS, and pH 8.3 running buffer. The labeled FL-Pks13 was detected and quantified using a Typhoon FLA 9000 gel imager. The fluorescence was quantified (relative fluorescence units, RFU) using the ImageQuant software provided with the instrument. The IC₅₀ value was calculated by plotting the normalized band intensity versus PhU-AMS concentration and fitting to the four-parameter Hill equation using GraphPad Prism (version 6.0).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This project has been funded in whole or in part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under contract nos. HHSN272200700058C and HHSN272201200026C (to W.F.A.). C.C.A. acknowledges support from the NIH (1-U54-AI-057153, AI070291). Some support was also provided as a subgrantee of the SGC under a Bill and Melinda Gates Foundation grant for the Structure-Guided Drug Discovery Coalition (SDDC). Additional funding for this project includes San Francisco State University Startup funds (to M.L.K.). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. The use of LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (grant 085P1000817).

ABBREVIATIONS

Mtb	Mycobacterium tuberculosis
TB	tuberculosis
FAS	fatty acid synthase
PKS	polyketide synthase
FAAL	fatty acyl-AMP ligase
MDR-TB	multi-drug-resistant TB
XDR-TB	extensively drug-resistant TB

FadDs	fatty acid adenylating enzymes
FACLs	fatty acyl-CoA ligases
CoA	coenzyme A
ACP	acyl-carrier protein
Msm	<i>Mycobacterium smegmatis</i>
Mma	<i>Mycobacterium marinum</i>
PDB	Protein Data Bank
PhU-AMS	5'-O-[N-(11-phenoxyundecanoyl)sulfamoyl]adenosine

REFERENCES

1. WHO. Tuberculosis mortality nearly halved since 1990. [http:// www.who.int/mediacentre/news/releases/2015/tuberculosis-mortality/en/](http://www.who.int/mediacentre/news/releases/2015/tuberculosis-mortality/en/) (accessed May 11, 2016)
2. Lenaerts A, Barry CE, Dartois V. Heterogeneity in tuberculosis pathology, microenvironments and therapeutic responses. *Immunol. Rev.* 2015; 264:288–307. [PubMed: 25703567]
3. Dartois V. The path of anti-tuberculosis drugs: from blood to lesions to mycobacterial cells. *Nat. Rev. Microbiol.* 2014; 12:159–67. [PubMed: 24487820]
4. Russell DG, Lee W, Tan S, Sukumar N, Podinovskaia M, Fahey RJ, VanderVen BC. The Sculpting of the *Mycobacterium tuberculosis* Genome by Host Cell-Derived Pressure. *Microbiol. Spectrum.* 2014; 2 OI:10.1128/microbiolspec.MGM2-0016-2013.
5. WHO. Antimicrobial resistance: global report on surveillance 2014. <http://www.who.int/drug-resistance/documents/surveillance-report/en/> (accessed May 11, 2016).
6. Takayama K, Wang C, Besra GS. Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. *Clin. Microbiol. Rev.* 2005; 18:81–101. [PubMed: 15653820]
7. Marrakchi H, Lanèlle M-A, Daffè M. Mycolic acids: structures, biosynthesis, and beyond. *Chem. Biol.* 2014; 21:67–85. [PubMed: 24374164]
8. Lèger M, Gavalda S, Guillet V, van der Rest B, Slama N, Montrozier H, Mourey L, Quémard A, Daffè M, Marrakchi H. The dual function of the *Mycobacterium tuberculosis* FadD32 required for mycolic acid biosynthesis. *Chem. Biol.* 2009; 16:510–9. [PubMed: 19477415]
9. North EJ, Jackson M, Lee RE. New approaches to target the mycolic acid biosynthesis pathway for the development of tuberculosis therapeutics. *Curr. Pharm. Des.* 2014; 20:4357–78. [PubMed: 24245756]
10. Mohanty D, Sankaranarayanan R, Gokhale RS. Fatty acyl-AMP ligases and polyketide synthases are unique enzymes of lipid biosynthetic machinery in *Mycobacterium tuberculosis*. *Tuberculosis.* 2011; 91:448–455. [PubMed: 21601529]
11. Trivedi OA, Arora P, Sridharan V, Tickoo R, Mohanty D, Gokhale RS. Enzymic activation and transfer of fatty acids as acyl-adenylates in mycobacteria. *Nature.* 2004; 428:441–5. [PubMed: 15042094]
12. Portevin D, de Sousa-D'Auria C, Montrozier H, Houssin C, Stella A, Lanèlle M-A, Bardou F, Guilhot C, Daffè M. The acyl-AMP ligase FadD32 and AccD4-containing acyl-CoA carboxylase are required for the synthesis of mycolic acids and essential for mycobacterial growth: identification of the carboxylation product and determination of the acyl-CoA carboxylase component. *J. Biol. Chem.* 2005; 280:8862–74. [PubMed: 15632194]
13. Gavalda S, Lèger M, van der Rest B, Stella A, Bardou F, Montrozier H, Chalut C, Burlet-Schiltz O, Marrakchi H, Daffè M, Quémard A. The Pks13/FadD32 crosstalk for the biosynthesis of mycolic acids in *Mycobacterium tuberculosis*. *J. Biol. Chem.* 2009; 284:19255–64. [PubMed: 19436070]

14. Gavalda S, Bardou F, Laval F, Bon C, Malaga W, Chalut C, Guilhot C, Mourey L, Daffè M, Quémard A. The polyketide synthase Pks13 catalyzes a novel mechanism of lipid transfer in mycobacteria. *Chem. Biol.* 2014; 21:1660–9. [PubMed: 25467124]
15. Goyal A, Verma P, Anandhakrishnan M, Gokhale RS, Sankaranarayanan R. Molecular basis of the functional divergence of fatty acyl-AMP ligase biosynthetic enzymes of *Mycobacterium tuberculosis*. *J. Mol. Biol.* 2012; 416:221–38. [PubMed: 22206988]
16. Arora P, Goyal A, Natarajan VT, Rajakumara E, Verma P, Gupta R, Yousuf M, Trivedi OA, Mohanty D, Tyagi A, Sankaranarayanan R, Gokhale RS. Mechanistic and functional insights into fatty acid activation in *Mycobacterium tuberculosis*. *Nat. Chem. Biol.* 2009; 5:166–73. [PubMed: 19182784]
17. Buchholz TJ, Geders TW, Bartley FE, Reynolds KA, Smith JL, Sherman DH. Structural basis for binding specificity between subclasses of modular polyketide synthase docking domains. *ACS Chem. Biol.* 2009; 4:41–52. [PubMed: 19146481]
18. Portevin D, De Sousa-D'Auria C, Houssin C, Grimaldi C, Chami M, Daffè M, Guilhot C. A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. *Proc. Natl. Acad. Sci. U. S. A.* 2004; 101:314–9. [PubMed: 14695899]
19. Carroll P, Faray-Kele M-C, Parish T. Identifying vulnerable pathways in *Mycobacterium tuberculosis* by using a knockdown approach. *Appl. Environ. Microbiol.* 2011; 77:5040–3. [PubMed: 21642404]
20. Kawate T, Iwase N, Shimizu M, Stanley SA, Wellington S, Kazyanskaya E, Hung DT. Synthesis and structure-activity relationships of phenyl-substituted coumarins with anti-tubercular activity that target FadD32. *Bioorg. Med. Chem. Lett.* 2013; 23:6052–9. [PubMed: 24103299]
21. Stanley SA, Kawate T, Iwase N, Shimizu M, Clatworthy AE, Kazyanskaya E, Sacchettini JC, Ioerger TR, Siddiqi NA, Minami S, Aquadro JA, Grant SS, Rubin EJ, Hung DT. Diarylcoumarins inhibit mycolic acid biosynthesis and kill *Mycobacterium tuberculosis* by targeting FadD32. *Proc. Natl. Acad. Sci. U. S. A.* 2013; 110:11565–70. [PubMed: 23798446]
22. Guillet V, Galandrin S, Maveyraud L, Ladevèze S, Mariaule V, Bon C, Eynard N, Daffè M, Marrakchi H, Mourey L. Insight into Structure-Function Relationships and Inhibition of the Fatty Acyl-AMP Ligase (FadD32) orthologs from *Mycobacteria*. *J. Biol. Chem.* 2016; 291:7973–89. [PubMed: 26900152]
23. Li W, Gu S, Fleming J, Bi L. Crystal structure of FadD32, an enzyme essential for mycolic acid biosynthesis in mycobacteria. *Sci. Rep.* 2015; 5:15493. [PubMed: 26628098]
24. Gulick AM. Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem. Biol.* 2009; 4:811–27. [PubMed: 19610673]
25. Schmelz S, Naismith JH. Adenylate-forming enzymes. *Curr. Opin. Struct. Biol.* 2009; 19:666–71. [PubMed: 19836944]
26. Liu Z, Ioerger TR, Wang F, Sacchettini JC. Structures of *Mycobacterium tuberculosis* FadD10 protein reveal a new type of adenylate-forming enzyme. *J. Biol. Chem.* 2013; 288:18473–83. [PubMed: 23625916]
27. Zhang Z, Zhou R, Sauder JM, Tonge PJ, Burley SK, Swaminathan S. Structural and functional studies of fatty acyl adenylate ligases from *E. coli* and *L. pneumophila*. *J. Mol. Biol.* 2011; 406:313–24. [PubMed: 21185305]
28. Andersson CS, Lundgren CAK, Magnúsdóttir A, Ge C, Wieslander A, Martinez Molina D, Högbom M. The *Mycobacterium tuberculosis* very-long-chain fatty acyl-CoA synthetase: structural basis for housing lipid substrates longer than the enzyme. *Structure.* 2012; 20:1062–7. [PubMed: 22560731]
29. May JJ, Kessler N, Marahiel MA, Stubbs MT. Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. *Proc. Natl. Acad. Sci. U. S. A.* 2002; 99:12120–5. [PubMed: 12221282]
30. Galandrin S, Guillet V, Rane RS, Lèger MNR, Eynard N, Das K, Balganesch TS, Mourey L, Daffè M, Marrakchi H. Assay development for identifying inhibitors of the mycobacterial FadD32 activity. *J. Biomol. Screening.* 2013; 18:576.

31. Mawuenyega KG, Forst CV, Dobos KM, Belisle JT, Chen J, Bradbury EM, Bradbury ARM, Chen X. Mycobacterium tuberculosis functional network analysis by global subcellular protein profiling. *Mol. Biol. Cell.* 2005; 16:396–404. [PubMed: 15525680]
32. Målen H, Pathak S, Sjøfteland T, de Souza GA, Wiker HG. Definition of novel cell envelope associated proteins in Triton X-114 extracts of Mycobacterium tuberculosis H37Rv. *BMC Microbiol.* 2010; 10:132. [PubMed: 20429878]
33. Gu S, Chen J, Dobos KM, Bradbury EM, Belisle JT, Chen X. Comprehensive proteomic profiling of the membrane constituents of a Mycobacterium tuberculosis strain. *Mol. Cell. Proteomics.* 2003; 2:1284. [PubMed: 14532352]
34. Reger AS, Wu R, Dunaway-Mariano D, Gulick AM. Structural characterization of a 140 degrees domain movement in the two-step reaction catalyzed by 4-chlorobenzoate:CoA ligase. *Biochemistry.* 2008; 47:8016–25. [PubMed: 18620418]
35. Wu R, Reger AS, Cao J, Gulick AM, Dunaway-Mariano D. Rational redesign of the 4-chlorobenzoate binding site of 4-chlorobenzoate: coenzyme A ligase for expanded substrate range. *Biochemistry.* 2007; 46:14487–99. [PubMed: 18027984]
36. Millard CS, Stols L, Quartey P, Kim Y, Dementieva I, Donnelly MI. A less laborious approach to the high-throughput production of recombinant proteins in Escherichia coli using 2-liter plastic bottles. *Protein Expression Purif.* 2003; 29:311–20.
37. Quadri LE, Weinreb PH, Lei M, Nakano MM, Zuber P, Walsh CT. Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry.* 1998; 37:1585–95. [PubMed: 9484229]
38. Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein Expression Purif.* 2005; 41:207–34.
39. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2006; 62:859–66. [PubMed: 16855301]
40. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2010; 66:213–21. [PubMed: 20124702]
41. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2010; 66:486–501. [PubMed: 20383002]
42. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 1997; 53:240–55. [PubMed: 15299926]
43. Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* 1999; 6:458–63. [PubMed: 10331874]
44. Vaguine AA, Richelle J, Wodak SJ. SFCHECK: a unified set of procedures for evaluating the quality of macro-molecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 1999; 55:191–205. [PubMed: 10089410]
45. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2010; 66:12–21. [PubMed: 20057044]
46. Ikeuchi H, Meyer ME, Ding Y, Hiratake J, Richards NGJ. A critical electrostatic interaction mediates inhibitor recognition by human asparagine synthetase. *Bioorg. Med. Chem.* 2009; 17:6641–50. [PubMed: 19683931]

Figure 1.

α-Mycolic acid structure, biosynthesis mediated by FadD32, and FadD32 inhibition by Phu-AMS. (a) Structure of α-mycolic acid (1), the most abundant lipid produced by *Mycobacterium tuberculosis* H37Rv, which is biosynthetically derived from meromycolic acid (2). (b) FadD32 catalyzes the penultimate step in mycolic acid biosynthesis through loading meromycolic acid (2) onto the N-terminal acyl carrier protein (ACP1) domain of PKS13, which is accomplished in two partial reactions. FadD32 first condenses meromycolic acid (2) and ATP to afford acyl-adenylate (3). In the second partial reaction, FadD32 binds acceptor protein PKS13 and transfers the activated meromycolic acid onto the pantetheine arm of ACP1 of PKS13 to provide meromycolic-bound PKS13 (4).¹⁴ (c) Chemical structures of FadD32 ligands used to co-crystallize FadD32 homologues. Phu-AMS is described in these studies, and C12-AMP was reported by Guillet and co-workers²² (d) Inhibition of FadD32-catalyzed loading of BODIPY FL C₁₆ onto full-length PKS13 by Phu-AMS. This panel displays a representative gel showing the loading of BODIPY FL C₁₆ onto full-length PKS13 catalyzed by FadD32 in the presence of varying concentrations of Phu-AMS from 0 to 50 μM. A positive control (DMSO only) and three negative controls [50 mM EDTA, (-) FadD32, and (-) PKS13] were also included in the experiment. Reaction conditions are described in Methods. The final reaction mixtures consisted of 32 nM FadD32, 50 μM BODIPY FL C₁₆, 2 mM ATP, and 2 μM PKS13 in assay buffer (50 mM HEPES, 5 mM MgCl₂, 1 mM DTT, 0.01% Tween-20, pH 7.2) and varying amounts of inhibitors or controls. The reaction mixture was separated on SDS-PAGE, and the fluorescent intensity of labeled PKS13 was measured by gel fluorescence. The fluorescence was quantified (relative fluorescence units, RFU) using the ImageQuant software provided with the instrument. (e) Concentration-response plot of fluorescent band intensity as a

function of PhU-AMS concentration. The curve represents the best nonlinear fit to the four-parameter Hill equation. The data points represent the mean plus standard deviation ($IC_{50} = 6.8 \pm 1.1 \mu M$) of triplicate experiments. The IC_{50} value was calculated by plotting the normalized band intensity versus PhU-AMS concentration and fitting to the four-parameter Hill equation using GraphPad Prism (version 6.0).

Figure 2.

FadD32 structure with PhU-AMS bound to active site. (a) Overall structure of FadD32. N-terminal domain in green, C-terminal domain in cyan, PhU-AMS in yellow, PEG and glycerol in gray, and phosphate in orange. (b) Stereo representation of the FAAL-specific insertion motif (residues 369–397) in the FadD32 structure. The stereo view of the autobuilt model using Phenix contoured at 1.0σ with initial experimental electron density maps. (c) Active site with bound inhibitor. The Fo-Fc electron density omit map is shown in gray contoured at 3.0σ and was calculated using the final model where the ligand was removed. (d) Ligplot representation of PhU-AMS bound to the protein. Residues forming hydrophobic interactions with the ligand are labeled in black near red eyelashes, and residues that form H-bonds with the ligand are labeled in green. (e, f) Regions of the protein where additional ligands are bound. Residues that participate in H-bonds with PEG, glycerol (GOL), and phosphate are shown.

Figure 3.

Important regions of the FadD32 protein structure, including the interactions between the N- and C-terminal domains and FAAL-specific insertion motif. (a) FadD32 structure with the FAAL-specific insertion motif in red, hinge in dark blue, FadD32-specific insertion motif in teal, P-loop in orange, C-terminal tail in purple, PhU-AMS ligand in yellow, and residues on the N-terminal domain that we propose interact with the membrane in magenta. The N-domain is dark gray, and the C-domain is light gray. (b) Protein sequence of FadD32 colored by the same regions of the protein in panel a: P-loop in orange (194–204), FAAL-specific insertion motif in red (residues 369–397), hinge in dark blue (490–508), FadD32-specific insertion motif in teal (529–554), and C-terminal tail in purple (625–637). (c) Interaction between the FAAL-specific insertion motif and C-terminal domain. The insertion motif is red, the C-terminal domain is light gray, the hinge is dark blue, and the N-terminal domain is dark gray. A water molecule is shown as a gray sphere. (d) Hydrophobic interactions between the FAAL-specific insertion motif and N-terminal domain. (e) Surface representation of the C-terminal domain and how it interacts with the FAAL-specific insertion motif (cartoon representation). Colors are the same as indicated in panel a.

Figure 4.

C-terminal tail of FadD32 blocks the proposed entrance site for the pantetheine of ACP. (a) FadD32 structure in the adenylation conformation. The C-terminal tail is in purple, the hinge is in dark blue, the FAAL-specific insertion motif is in red, the P-loop is in orange, and the PhU-AMS inhibitor is in yellow. (b) FadD32 structure with CoA ligand in pink overlaid from the 1PG4 structure after aligning both structures on the basis of their N-terminal domains. (c) FadD32 structure with DG9 ligand in light blue overlaid with the 4DG9 structure after aligning both structures on the basis of their N-terminal domains. (d) FadD32 structure with the N-terminal domain in green and C-terminal domain in cyan. The DG9 ligand from the 4DG9 structure is shown in light blue, and the PhU-AMS inhibitor is in yellow. Residues in dark purple were identified as ones that conferred resistance to coumarin compounds¹⁸ and are at the interface of the N- and C-terminal domains at the proposed entrance site for the pantetheine of ACP on PKS13.

Figure 5.

Structural alignment and comparison of FadD32 homologous crystal structures. All recently determined crystal structures of mycobacterial FadD32 proteins were aligned using the align command in pymol. Structures include *Mtb* in gray, *Mma* in light purple, *Msm* 5ICR in dark purple, *Msm* 5EY8 in cyan, *Msm* 5D6J in light pink, and *Msm* 5D6N (N-domain only) in magenta. Each of the four regions of the protein that were variable between structures is indicated on the panel at the right of the aligned structures. The carbon atoms of the inhibitors are colored the same as the structures in which they were observed.

Figure 6.

Adenylation and modeled thiolation conformations of FadD32. The N-terminal domain of FadD32 is shown as a surface representation in dark gray. Residues proposed to interact with the membrane are in magenta. The C-terminal domain is rainbow-colored, with the N-terminus in blue and the C-terminus in red. (a, b) Crystallized adenylation conformation of FadD32 and (c, d) the modeled thiolation conformation of FadD32, which was determined by aligning the C-terminal domain of the FadD32 structure with the C-terminal domain of the 3CW9 structure. CoA is shown in pink sticks from the 1PG4 structure.

Figure 7.

Interactions between N- and C-terminal domains in the crystallized adenylation conformation and modeled thiolation conformation of FadD32. (a) FadD32 crystallized in the adenylation conformation. The N-terminal domain is in gray and the C-terminal domain is rainbow-colored from blue to red for the N- to C-terminus of the C-terminal domain. The PhU-AMS inhibitor is shown in yellow. The FAAL-specific insertion motif is indicated with an arrow, and the hinge is also indicated. K608 is shown with a stick representation in orange. (b) FadD32 modeled thiolation conformation based on the 3CW9 structure. All colors are the same as in panel a. (c) Hinge movement of structures from the ANL superfamily of enzymes compared to the FadD32 structure. The proposed progression of movement of the hinge from adenylation to thiolation conformations is shown from purple to red and numbered accordingly. All structures were aligned by their N-terminal domains in pymol. Only the hinge regions are shown: (1) 3CW8 in dark purple, (2) FadD32 in dark blue, (3) 3R44 in cyan, (4) 3IPL in green, (5) 3KXW in yellow, (6) 3CW9 in orange, and (7) 1PG4 in red.

Table 1Data Collection, Refinement, and Validation Statistics for the FadD32 Structure^a

	data set 1, used for structure solution	data set 2, used for structure refinement
Data Collection Statistics		
space group		$P2_12_12_1$
unit cell: a , b , c (Å); α , β , γ (deg)	62.78, 74.56, 127.76, 90, 90, 90	
wavelength (Å)		0.97875
resolution range (Å)	30.00–2.50 (2.54– 2.50)	30.00–2.25 (2.29–2.25)
number of observed reflections	21 379 (1027)	29 261 (1419)
R_{merge} (%)	7.3 (54.1)	7.4 (58.3)
completeness (%)	100.0 (100.0)	99.9 (100.0)
$I/\sigma(I)$	35.3 (6.1)	18.4 (2.8)
redundancy	10.4 (10.6)	4.9 (4.9)
Wilson B factor	40.6	49.5
phasing method	SAD (Se-methionine)	
Refinement and Validation: Final Model		
$R_{\text{work}}/R_{\text{free}}$ (%)		18.5/23.9 (25.5/31.1)
number of protein residues		625
ligands		
heteroatoms		inhibitor, 1 μB_E , 2 Cl, 3 PO_4 , 1 PEG, 2 GOL
solvent molecules		131 H_2O
average B factor for protein		63.7 (5029 atoms)
average B factor for inhibitor		41.7 (42 atoms)
rmsd bond length (Å)		0.009
rmsd bond angle (deg)		1.52
Ramachandran (%)		
most favored regions		90.0
additionally allowed regions		9.0
generously allowed regions		0.8
disallowed regions		0.2
PDB ID		5HM3

^aHigh-resolution shell statistics are shown in parentheses, and there is one molecule in the asymmetric unit of the structure. See Methods for more information.