Published in final edited form as:
Nat Chem Biol. 2009 March ; 5(3): 166–173. doi:10.1038/nchembio.143.

Mechanistic and functional insights into fatty acid activation in *Mycobacterium tuberculosis*

Pooja Arora[1,†], **Aneesh Goyal**[2,†], **Vivek T Natarajan**[1,†], **Eerappa Rajakumara**[2], **Priyanka Verma**[1], **Radhika Gupta**[3], **Malikmohamed Yousuf**[2], **Omita A. Trivedi**[1], **Debasisa Mohanty**[1], **Anil Tyagi**[3], **Rajan Sankaranarayanan**[2,*], and **Rajesh S. Gokhale**[1,4,*]

[1]National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067, India

[2]Centre for Cellular and Molecular Biology, Council of Scientific and Industrial Research, Uppal Road, Hyderabad-500 007, India

[3]Department of Biochemistry, University of Delhi South Campus, Benito Juarez Road, New Delhi 110021, India

[4]Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore, India (by courtesy).

Abstract

The recent discovery of fatty acyl-AMP ligases (FAALs) in *Mycobacterium tuberculosis* (Mtb) provided a new perspective to fatty acid activation dogma. These proteins convert fatty acids to corresponding adenylates, which is an intermediate of acyl-CoA-synthesizing fatty acyl-CoA ligases (FACLs). Presently, it is not evident how obligate pathogens like Mtb have evolved such new themes of functional versatility and whether the activation of fatty acids to acyl-adenylates could indeed be a general mechanism. Here, based on elucidation of the first structure of a FAAL protein and by generating loss- as well as gain-of-function mutants that interconvert FAAL and FACL activities, we demonstrate that an insertion motif dictates formation of acyl-adenylate. Since FAALs in Mtb are crucial nodes in biosynthetic network of virulent lipids, inhibitors directed against these proteins provide a unique multi-pronged approach of simultaneously disrupting several pathways.

Introduction

In nature, fatty acids must be activated before they can be assimilated into various metabolic pathways. The universal mechanism of *n*-fatty acid activation involves conversion of fatty acids to their corresponding coenzyme A (CoASH [**1**])-derivatives by a family of omnipresent fatty acyl-CoA ligases (FACLs). This catalytic process occurs in two steps and involves requisite formation of an acyl-adenylate (Acyl-AMP [**2**]) intermediate. Recently, an

*Corresponding Authors. Correspondence should be addressed to: Dr. Rajesh S. Gokhale, Chemical Biology Group, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi-110067, India. Email: rsg@nii.res.in , Tel: +91-11-26703647, Fax: +91-11-26703647 Dr. Rajan Sankaranarayanan, Structural Biology Laboratory, Centre for Cellular and Molecular Biology (CCMB), Uppal Road, Hyderabad - 500 007, India. Email: sankar@ccmb.res.in Tel: +91-40-27192832, Fax: +91-40-27160591.
†Authors have contributed equally

Author Contributions

P.A., A.G., V.T.N., R.S. and R.S.G. designed experiments, analyzed data and wrote the manuscript. E.R. and M.Y. participated in structural studies. P.V., R.G. and O.A.T. participated in biochemical and mechanistic studies. D.M. and A.T. provided valuable advice.

Crystallographic data

The data collection statistics are reported in Goyal *et, al* (2006).

The coordinates of FAAL28 N-terminal domain have been deposited in the Protein Data Bank with access code 3E53.

Europe PMC Funders Author Manuscripts

alternate mechanism of fatty acid activation catalyzed by fatty acyl-AMP ligases (FAALs) was established in *Mycobacterium tuberculosis* (Mtb)1. FAAL proteins convert fatty acids to acyl-adenylates and do not perform the final transfer to CoASH. Instead the activated fatty acids are acylated onto the acyl carrier proteins (ACP) of polyketide synthases (PKSs) to biosynthesize lipidic metabolites1. This mode of activation is reminiscent of the adenylation domains of non-ribosomal peptide synthetases (NRPSs)2-4, which along with FAAL and FACL proteins constitute a large superfamily of acyl-activating enzymes (AAEs). Interestingly, the genome sequencing projects of several bacteria, fungi and plants have revealed a large number of fatty acid-activating enzymes. However, presently it is not feasible to discriminate between FAAL and FACL proteins based on their protein sequences.

Mtb is the causative agent of tuberculosis (TB) in humans. Although this pathogen has been known for centuries, TB still accounts for more than two million deaths every year5,6. Mtb possesses complex arsenal of virulence factors and has evolved elaborate strategies to escape host surveillance. The cell envelope of tubercle bacilli is endowed with complex lipids, many of which play an important role in its pathogenesis7,8. Recent investigations of lipid biosynthesis have demonstrated that polyketide synthases (PKSs) in combination with fatty acid synthases (FASs) in Mtb synthesize unusual acyl chains9-15. The coordination between FASs and PKSs is achieved by FAALs1. These proteins along with FACLs constitute 34 homologues annotated as FadD in the Mtb genome16. Emerging evidences suggest that Mtb might be utilizing FACL proteins to degrade alternate carbon sources during the latent phase of infection17,18. FAALs in contrast are likely to be essential during the growth phase of Mtb. It is thus interesting to note that both FAAL and FACL enzymes utilize fatty acid pools but channel them towards different metabolic fates in distinct stages of Mtb life cycle (Fig. 1).

Structural studies of AAEs have revealed a conserved fold despite the limited sequence homology among different members19. This conserved fold contain a large N-terminal and a small C-terminal domain, which undergo domain movements during various steps of catalytic cycle20. Mtb genome contains 6 adenylation domains as part of the multifunctional NRPS proteins, whereas the 34 FadD homologues (except for FAAL22 and FAAL9) are all present as independent proteins. Previous sequence-based dendrogram analysis of Mtb FadD proteins revealed two clades of FAAL and FACL enzymes. However, the molecular features which determine their biochemical functions have not been elucidated.

Our study here reveals a mechanism by which Mtb may have evolved FAAL proteins from the omnipresent FACLs. It is remarkable that FAAL proteins have retained CoA-binding pocket and this new catalytic function is generated by modifying substrate-induced conformational rearrangements. The existence of FAALs in other genomes is also demonstrated in this study. We have also developed small molecule inhibitors of these enzymes that simultaneously disrupt multiple pathways in Mtb. Our studies illustrate a novel multi-pronged approach that provides credence to the emerging ‘systems pharmacology’ approach for drug discovery.

Results

Analysis of FAAL and FACL specific determinants

The identification of large number of fatty acid-activating enzymes (annotated as FadD initially) in the Mtb genome was rather surprising16; however subsequent genome sequencing of other *Actinomycetes* have revealed even greater number of FadD homologues. The FadD nomenclature for these genes emerge from *Escherichia coli* literature and stands for fatty acid degradation (gene *D* from this operon). Interestingly,

several of the 34 FadD proteins in Mtb were recently demonstrated to be involved in biosynthesis of lipidic metabolites1. The common annotation of these Mtb homologues as FadD creates confusion and does not reflect its true function. We here propose to modify these annotations and directly refer to them as FAAL and FACL, for example FadD28 can be substituted by FAAL28 and FadD19 as FACL19.

FAAL proteins of Mtb are a group of closely related enzymes and show high sequence similarity in the range of 70 to 80 percent. Interestingly, the homology of these enzymes with other AAE proteins (FACL and adenylation domains) in Mtb is of the order of 20-30 percent. Since FAAL and FACL proteins also possess similar substrate specificity, we decided to focus our attention on these stand-alone proteins to delineate the mechanistic differences. Comparative analysis of FAAL and FACL proteins showed a striking class-specific variation in the conserved 12 residue nucleotide-binding motif (see supplementary **Fig. S1a** online). We decided to probe whether these amino acid changes could influence ATP [**3**] binding. However, mutagenic studies of this region did not change the catalytic function (see supplementary **Fig. S1b** online), suggesting that these residues are not determinants of FAAL function.

Since rotation about N- and C-terminal domains in AAEs are known to be correlated with catalysis19-22, we therefore decided to investigate into the functional contributions of N- and C-domains. Towards this, three variants of N-domains (N1: 1 to 460, N2: 1 to 465 and N3: 1 to 469) and two C-domains (C1: 460 to 580 and C2: 465 to 580) of FAAL28 protein were cloned and expressed in *E. coli*. Biochemical assays performed with engineered N-domain proteins showed poor catalytic activity towards formation of acyl-AMP and the addition of C-domain protein in different molar ratios also did not significantly enhance the activity (see supplementary **Fig. S1c** online). Another strategy involved generation of domain hybrids between N- and C-domains of FAAL and FACL protein. Hybrid proteins between FACL15 and FACL17 and FAAL28 and FAAL29 were constructed by splicing at a conserved A7 ST-GD motif. Enzymatic assays for the fusion proteins containing C-domain from FAAL and N-domain from FACL did not show any difference in their product profiles (see supplementary **Fig. S1d** online). This was surprising since the CoA-binding pocket is known to be present at the junction of N- and C-domains20. This data suggested a surprising tentative inference that C-domains of FAAL proteins retain the CoASH binding residues and in conjunction with FACL N-terminal domains perform the complete catalytic cycle to synthesize acyl-CoA [**4**].

Three-dimensional structure determination of FAAL28

Since sequence-based investigation did not provide a definitive basis for discriminating between FAAL and FACL activity, we therefore decided to solve the three-dimensional structure of one of the FAAL proteins. We obtained well diffracting crystals for FAAL N-terminal domain (1 to 460; N1), which includes most of the residues involved in the enzymatic activity and also the determinants of fatty acyl substrate specificity23. By using multi-wavelength anomalous dispersion method, we have determined crystal structure of N1 of FAAL28 protein (see supplementary **Fig. S2** online). The structure reveals 3 subdomains: subdomains A and B, fold into a α+β topology and the third subdomain C into a distorted β-barrel topology (see supplementary **Fig. S3b** online). The overall fold of FAAL28 is similar to acetyl-CoA synthetase (ACS)20, adenylation domain of gramicidin NRPS (PheA)24 and LC-FACS from *Thermus thermophilus*22 and can be superimposed with a root mean square deviation of 1.9Å for 311 Cα positions (22.3% sequence identity), 1.8 Å for 311 Cα positions (20.4% sequence identity) and 1.85 Å for 288 Cα positions (23.5% sequence identity) respectively.

Structural analysis revealed a 22 amino acid insertion (351-372) between β1 and β2 of C subdomain (Fig. 2a and see supplementary **Fig. S3** online). This insertion was conserved in other FAAL homologues and was absent in FACL proteins (see supplementary **Fig. S3a** online). In order to investigate into the functional role of this unique insertion motif, we then modeled the C-domain of FAAL28 by using coordinates from PheA24 and ACS structures20 (Fig. 2b). C-terminal domain modeled based on PheA (green in Fig. 2b) represents the acyl-AMP forming active conformation of FAALs. In this orientation, the C-domain is positioned closer to the insertion motif and could directly interact with this insertion. The C-domain modeled on the basis of ACS for the CoA-bound conformation (magenta in Fig. 2b) is substantially away from the insertion motif towards the active site to form the acyl-CoA product. The other face of insertion motif is stabilized by hydrophobic interactions with the rest of the protein (Fig. 2b) and is unlikely to be mobile. Based on these observations we argued that the insertion motif could modulate the C-domain movements, which could in turn disrupt the second half of the catalysis to synthesize acyl-CoA.

Insertion sequence converts FACL into an FAAL

In order to verify our structure-based hypothesis, we incorporated the 22 amino acid insertion sequence from FAAL28 at the corresponding structural position in the FACL19 protein. In order to convert a FACL activity to a FAAL, the protein must actively release the acyl-adenylate. Enzymatic assays performed with this mutant protein ($FACL19_{\mathbf{i}}$) showed dramatic decrease in its ability to synthesize acyl-CoA (Fig. 3a). Comparative kinetic analysis for the formation of acyl-CoA showed more than 25-fold increase in the K_M of CoASH for the mutant protein and no significant change in K_M was observed for ATP. We predicted that the conversion of FACL into a FAAL protein could be attributed to the inability of the mutant protein to generate a conformation that could bind CoASH effectively.

The comparative analysis of insertion sequences from other Mtb FAAL members did not reveal any conserved motif. Incorporation of a stretch of 15 amino acids containing Ala and Ser residues in FACL19 protein ($FACL19_{\mathbf{AS}}$) (see supplementary **Table 1** online) also predominantly synthesized acyl-AMP (Fig. 3b). Thus the physical presence of a polypeptide chain at this strategic position is sufficient to disable the formation of acyl-CoA. It is worthwhile to note that the FACL mutant proteins are capable of actively releasing acyl-AMP, which is an enzyme-sequestered intermediate during acyl-CoA formation. However, loss-of-function experiments could be difficult to rationalize, since these could result from secondary effects. An unambiguous proof for the importance of this insertion sequence in determining catalytic function would be to convert FAAL into FACL.

Gain-of-function in FAAL28 to synthesizes acyl-CoA

In order generate gain-of-function in FAAL scaffold, we perceived that this may require engineering a CoA-binding pocket in these proteins. Computational investigation of the CoA-binding pockets in FAALs surprisingly revealed significant conservation of CoA-binding residues (see supplementary **Fig. S4** online). This analysis along with the results from the domain fusion experiments suggested that FAAL proteins could indeed possess a CoA-binding pocket and that FAALs could be converted to FACLs by deleting the insertion region. Since such a deletion from FAAL28 would expose hydrophobic patches to solvent, we therefore generated a series of deletion mutants in FAAL28 by keeping in mind various secondary structural elements present in this region (Fig. 3d). The FAAL28 mutant with a deletion from 354 to 365 ($FAAL28_{\mathbf{\Delta}}$) remarkably showed capability to convert lauroyl-AMP [**5**] intermediate into lauroyl-CoA [**6**], even at 50 μM concentrations of CoASH (Fig. 3c). This engineered protein follows Michaelis-Menten saturation kinetics and the k_{cat} was estimated to be 19523 S^{-1}. Remarkably, the catalytic rate constant for $FAAL28_{\mathbf{\Delta}}$ is 85% of

the wild type FACL19 protein. However, the K_M for CoASH was estimated to be 10-fold higher than FACL19. This is not surprising because during the construction of FAAL28$_{\Delta}$, we have not optimized residues binding to CoASH. These studies unequivocally demonstrate that FAAL28 protein indeed possess a CoASH binding pocket. Based on these observations, we propose a model wherein the insertion motif alters the C-terminal domain movements and abrogates the second step of acyl-CoA synthesis (Fig. 3e). Recently, weak activity to synthesize aminoacyl-CoA was also reported for adenylation domains of NRPSs25.

Class-specific multi-target inhibitors

In combination with the structural and mechanistic understanding of FAAL proteins and the recognized importance of FAALs in biosynthesis of mycobacterial virulent lipids, we decided to develop inhibitors against these enzymes. FAALs are the key nodal points connecting FAS with different PKS systems and the inhibitors targeting multiple FAAL proteins could simultaneously disrupt many lipid biosynthetic pathways. Biochemical investigations of various FAALs and FACLs showed overlapping substrate specificity profile for 15 FAAL and FACL proteins (see supplementary **Table 2** online). This parallel substrate specificity provided a tempting opportunity to design small molecule inhibitors that could simultaneously target not only FAALs but also the FACL proteins. Since acyl adenylate is a reaction intermediate of FACL proteins and is a product of FAAL proteins, we hypothesized that non-hydrolyzable substrate analogues could act as a class-specific inhibitor of FAAL and FACL proteins (Fig. 4a). Based on these considerations three acyl-sulfamoyl (acyl-AMS) analogues of hexanoic (HAMS [**7**]), lauric (LAMS [**8**]) and arachidic (AAMS [**9**]) acids were synthesized (see supplementary Method online). The purity as well as identity of acyl-AMS compounds was confirmed by HPLC followed by mass spectrometry analysis. Acyl-sulfamoyl analogues have been successfully used to inhibit mechanistically related enzymes like the aminoacyl-tRNA synthetases26. Recently, adenylation domains from siderophore biosynthesis have also been demonstrated to be inhibited by sulfamoyl analogues27,28.

Tight binding inhibitors of FAAL and FACL enzymes

The inhibitory profile of these compounds was investigated with purified FAAL and FACL proteins by examining the enzymatic activity in presence of increasing concentrations of inhibitor. Acyl-AMS inhibition was tested with a panel of FAAL and FACL proteins and all of these proteins displayed inhibition in the nanomolar range of concentrations. The IC_{50} (concentration of compounds required to inhibit enzymatic activity by 50%) values for FAAL28 and FACL19 estimated from dose-response curve were less than 10 times the enzyme concentration used in the assay, a characteristic feature of tight binding inhibitors 29. Since tight binding inhibitors (TBIs) cannot be analyzed by the traditional steady-state kinetic methods, the data was analyzed by using Morrison equation (Eq 1) (Fig 4b). LAMS inhibition was also investigated by following progress curves. As expected for TBIs, increase in inhibitor concentration led to decrease in linear phase and resulted in saturation at lower steady state velocity (see supplementary **Fig. S5a** online). The stability of the enzyme-inhibitor complex was further examined by performing the enzyme turnover recovery assays (Fig. S5b). Based on dilution experiments, poor enzymatic activity could be recovered suggesting that formation of EI* complex was not reversible.

To rule out the possibility that covalent acylation has resulted in the formation of inactive enzyme, ethyl acetate extractions were carried out with purified EI* complex. The extracted LAMS could be detected by mass spectrometry, clearly indicating that the inhibitor interacts through non-covalent interactions. Inability of dilution assays to displace enzyme-bound acyl-AMS analogues suggested binding of inhibitor in a closed binding site, probably

generated through conformational changes. These induced conformational changes in the enzyme-inhibitor complex were probed by limited proteolysis experiments (see supplementary **Fig. S5c** online). With increasing LAMS concentration, protection against proteolysis could be observed. Similar protection against tryptic fragmentation was also observed for FACL19 protein. These experiments confirm that tight binding property of acyl-AMS arises from the binding induced conformational changes that result in the formation of dead end complex.

Effect of acyl-AMS inhibitors on Mtb growth in cultures

The potent inhibitory activity of acyl-AMS analogues with purified proteins prompted us to investigate whether these analogues will inhibit FAAL and FACL proteins under culture conditions as well. The susceptibility of mycobacteria to LAMS was assayed by following colony forming units (CFU) on plates containing various concentrations of this inhibitor. The concentration of acyl-AMS resulting in 10^2 reductions in colony forming unit was used as a parameter to calculate minimum inhibitory concentration (MIC_{99}). The growth inhibition of both Mtb H37Rv and *M. smegmatis* mc^2 was estimated to be 100 μM and 215 μM respectively (Fig. 4b).

Since FAAL proteins load starter fatty acid substrates on polyketide synthases and synthesize lipids, we analyzed lipid profiles by following the incorporation of radiolabelled acetate and propionate units at various concentrations of acyl-AMS analogues. Radio-TLC analysis showed clear decrease in biosynthesis of mycolic acids with increasing LAMS concentrations (Fig. 4c). The higher running band in the radio-TLC correspond to the fatty acid methyl esters and display a reverse trend indicating that these inhibitors are not targeting synthesis of fatty acids (Fig 4d). ^{14}C-propionate incorporation studies facilitate investigation of lipids that are biosynthesized by utilizing methyl malonyl-CoA [**10**]. Here again, incubation of mycobacterial culture with LAMS decreased the amount of extractable methylated lipids (Fig. 4e).

Transmission Electron Microscopy (TEM) studies also showed gross modification of the outer coat of mycobacteria (Fig. 4f) compares the cell wall architecture of treated and untreated mycobacteria. In TEM images, the mycobacterial cells have a dense granular cytoplasm (rich in ribosomes), a thin cytoplasmic membrane, an electron dense peptidoglycan layer and an equally thick electron-transparent layer that separates the PG from the thin outer wall layer. All these components could be observed in the untreated mycobacterial cells (Panel i & ii). This characteristic mycobacterial lipid core is not visible in the LAMS-treated bacterial cells and cells displayed irregular shape (Panel iii & iv)

Characterization of FAALs from other *Actinomycetes*

Genome sequencing projects have revealed a large number of FadD homologues in several other organisms; we wanted to probe if FAAL proteins could also be identified across genera. Many of the *Actinomycetes* produce mycolic acids and are grouped in taxon mycolata. A FAAL homologue is required for biosynthesis of mycolic acids in Mtb and analysis indeed revealed FAAL32 homologues in the genomes of *Nocardia farcinica* and *Rhodococcus* sp RHA1 (Fig. 5a). Sequence analysis also indicated the presence of an insertion sequence (Fig. 5b). The enzyme from *Corynebacterium glutamicum* involved in mycolic acid biosynthesis was recently reported15. Towards this Rhodococcus FAAL homolog ro_4064 was cloned and expressed in *E. coli*. Cell-free enzymatic assays demonstrated that this enzyme could only synthesize palmitoyl-AMP and formation of palmitoyl-CoA could not be detected even at 1 mM concentration of CoASH (Fig. 5c). It is interesting to note that all these organisms apart from utilizing mycolic acids as part of their cell envelope have also retained the novel mechanistic basis of fatty acid activation. A

FAAL homologue was recently characterized from *Bacillus subtilis* 30 and is proposed to be involved in the production of a lipopeptide mycosubtilin. Other lipopeptide-producing organisms also reveal FAAL-like domains in genomes of *Nostoc sp, Nodularia spumigena* and *Streptomyces coelicolor*. These observations emphasize that acyl-AMPs could routinely be employed during biosynthesis of lipidic metabolites in diverse organisms.

Discussion

The recent discovery of FAALs in Mtb has provided a new perspective to fatty acid activation dogma. These proteins by activating fatty acids as acyl-adenylates redirect the metabolic flux towards biosynthesis of complex lipidic metabolites10. The free fatty acids could either come from direct uptake mechanisms31 or by hydrolysis of acyl-CoA or acyl-ACP thioesters32. In this study, we delineate evolutionary relationship between FACL and FAAL proteins. Our studies show that FAAL proteins indeed possess CoASH-binding pocket and their inability to synthesize acyl-CoA manifests from the changes in domain movements. Such alteration is brought about by a novel insertion motif identified using the three-dimensional structure of an FAAL protein FAAL28. Remarkably, the inter-conversion between the catalytic functions can be achieved by mere insertion and deletion of this region. By using this criterion of an insertion motif, we show that FAAL homologues could be identified in genomes of other bacteria.

In most instances FAAL homologues are positioned next to their cognate PKS/NRPS proteins in the genomes. This conserved organization provides a dual advantage, both in terms of co-expression of the two proteins as well as in substrate channelling of the labile acyl-AMP intermediate. FAAL proteins also circumvent the need to sustain metabolic pools of fatty acyl-CoAs to produce diverse lipidic metabolites9,33. The promiscuous substrate specificity of Mtb FAALs can produce spectrum of metabolites, which could be a mechanism to generate phenotypic heterogeneity in intracellular pathogens. Recently, another important fatty acid activation mechanism was characterized in phospholipid biosynthesis34. The acyl-phosphates produced in this case were generated from acyl-ACP, which by itself also functions as acyl donor35.

Systems biology-based approaches in recent years are beginning to reveal new insights into the understanding of metabolic networks, which enables pathogens to overcome unfavourable conditions, including insults imparted by target-specific drugs36,37. It is because of this reason that pharmaceutical industry is thinking beyond the traditional ‘one-disease-one-drug-one-target’ paradigm and is interested to develop strategies that can simultaneously inhibit more than one target38. The pivotal role of several FAAL homologues in biosynthesis of virulent lipids in Mtb prompted us to develop inhibitors that would simultaneously target multiple pathways. Our studies with the bi-substrate acyl-sulfamoyl analogues, which resemble common reaction intermediate of FAAL and FACL proteins, showed dramatic effect on the cell surface architecture of Mtb. *In vitro* analysis indeed showed inhibitions of both FAAL and FACL proteins. Since FACL enzymes in Mtb are being implicated in providing alternate energy sources during dormancy stage of infection18,39,40, this strategy suggests a possible means to target different stages of Mtb infection.

Methods

Materials

BAC genomic library of *M. tuberculosis* H37Rv was obtained from Prof. Stewart Cole of Pasteur Research Institute, France, *Rhodococcus RHA1* strain was obtained from Prof. Lindsay Eltis at the University of British Columbia. ^{14}C-acetic acid, ^{14}C-propionic acid ^{14}C-

hexanoic acid, ^{14}C-lauric acid and ^{14}C-palmitic acid were procured from American Radiolabelled Chemicals, GE healthcare, and Perkin Elmer Life sciences. ATP and Coenzyme A were procured from Sigma; all other chemicals were of analytical grade.

Cloning and expression of proteins in *E. coli*

Cloning of FAAL28 and FACL19 genes is reported earlier1. The primers used for the cloning are tabulated in supplementary information (see supplementary **Table 3** online). In brief, the proteins were expressed as *His*-tagged fusion proteins in *E. coli*, purified by Ni^{2+}-NTA purification, followed by anion exchange chromatography using Resource Q column and used for further studies.

Site-directed mutagenesis

For generating deletion and insertion constructs, site directed mutagenesis was carried out using QuickChange-XL kit from Stratagene. Details of all the mutation, insertion and deletion constructs generated in this study are given in supplementary **Table 3** online. All mutants were confirmed by DNA sequencing.

Crystallization and structure determination

Crystallization of the N-terminal domain, comprising of the first 460 residues of FAAL28, was performed by hanging drop vapor diffusion method. The crystals belong to space group P2(1)2(1)2(1), with unit-cell parameters a = 50.97, b = 60.74, c = 136.54 angstroms. The crystal structure of the N-terminal domain of FAAL28 at 2.35 angstroms resolution has been solved using the MAD method23 (see supplementary **Table 4** online). Model building was performed using O version 10.0.0.41. After several rounds of iterative manual model building and refinement using CNS version 1.2,42 the structure has been refined to a final R_{cryst} and R_{free} of 20.7% and 26.7% respectively. The following short segments (1-3, 97-102, 131-137, 153-160 and 172-177 of subdomain A, 203-204 connecting subdomains A and B, 367-368 which are part of the FAAL specific insertion and 460 last residue at C-terminus) are missing in the electron density presumably due to disorder and hence are not modeled. The quality of the model was checked using PROCHECK. The ribbon figures were made by SETOR and multiple sequence alignment was prepared using ALSCRIPT 43

Enzyme activity assays and kinetics

Activation of fatty acids as acyl-AMP or acyl-CoA was analyzed by radio-TLC assay, as described earlier1. The identities of the bands in radio TLC assays were verified by their R_f as well as with standard reactions of FACL19, whose products were characterized by HPLC and mass spectrometry. Lauroyl-CoA (Sigma) was used as a standard for HPLC and mass spectrometric analysis.

Inhibitor synthesis and evaluation of kinetic parameters

Acyl-AMS analogues were synthesized according to the procedure described previously27,44 (see supplementary **Method** online). Assays were carried by pre-incubating LAMS and the dose response curves were fitted to Morrisons equation for tight binding inhibitors.

$$v_i/v_o = 1 - \frac{([E] + [I] + Ki^{app}) - \sqrt{\{([E] + [I] + Ki^{app})^2 - 4[E][I]\}}}{2[E]} \quad \text{Eq.1}$$

where [E] is the enzyme concentration, [I] is LAMS concentration, v_i is velocity in the presence of LAMS, v_o is the velocity in the absence of LAMS respectively Ki^{app} values were calculated with the equation for tight binding inhibitors using the above fit29.

For progressive kinetics, the resulting data was fitted to equation 2 for slow binding inhibition 45.

$$P_t = P_0 - V_s t - (V_i - V_s)(1 - \exp(-k_{obs} t)/k_{obs}) \quad \text{Eq.2}$$

Where P_t and P_0 are products formed at time t and time zero, k_{obs} is the pseudo first order constant for approach of steady state and v_i and v_s correspond to the initial and final slopes of the progress curves.

For enzymatic turnover recovery assay, 0.2 μM protein was incubated with 1 μM LAMS for specified times. Enzyme activities were then assayed after 10^4-fold dilution under saturating substrate conditions.

Effect of Fatty acyl-AMS on mycobacterial growth and lipid biosynthesis

M. tuberculosis H37Rv culture was serially diluted and plated on Middlebrook 7H10 agar plates supplemented with oleic acid-albumin-dextrose-citric acid (OADC) containing various concentrations of fatty acyl-AMS analogues. Colonies were counted after two weeks of incubation at 37°C. Minimum inhibitory concentration (MIC) was defined as the concentration of LAMS required for 10^2-fold reduction in colony forming units (CFU). For lipid analysis, *M. tuberculosis* H37Rv cultures were grown till mid-log phase. 5 μCi of ^{14}C-acetate or ^{14}C-propionate and different concentrations of LAMS were added to 10 ml cultures and incubation continued for 16 h at 37°C. Mycolic acids and total extractable lipids were analyzed by using the previously described methodologies 46-48.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Lindsay Eltis, from the Department of Microbiology and Immunology University, British Columbia, for *Rhodococcus* RHA1 strain and Prof. Stewart Cole of Pasteur Research Institute, France for *M. tuberculosis* BAC genomic DNA library. P.A, A.G and E.R are Senior Research Fellows of the Council of Scientific and Industrial Research, India. R.S is a Wellcome Trust International Senior Research Fellow in India. R.S.G is an HHMI International Fellow. This work is also partially supported by Swarnajayanti Fellowship from DST, India and from a Centre of Excellence Grant by Department of Biotechnology, India. None of the authors have a financial interest related to this work.

References

1. Trivedi OA, et al. Enzymic activation and transfer of fatty acids as acyl-adenylates in mycobacteria. Nature. 2004; 428:441–5. [PubMed: 15042094]
2. Weber T, Marahiel MA. Exploring the domain structure of modular nonribosomal peptide synthetases. Structure. 2001; 9:R3–9. [PubMed: 11342140]
3. Fischbach MA, Walsh CT. Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. Chem Rev. 2006; 106:3468–96. [PubMed: 16895337]
4. Stachelhaus T, Mootz HD, Marahiel MA. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem Biol. 1999; 6:493–505. [PubMed: 10421756]
5. Gomez JE, McKinney JD. M. tuberculosis persistence, latency, and drug tolerance. Tuberculosis (Edinb). 2004; 84:29–44. [PubMed: 14670344]
6. WHO Report 2008. 2008. Global tuberculosis control - surveillance, planning, financing.
7. Barry CE, Crick DC, McNeil MR. Targeting the formation of the cell wall core of M. tuberculosis. Infect Disord Drug Targets. 2007; 7:182–202. [PubMed: 17970228]

8. Brennan PJ, Crick DC. The cell-wall core of Mycobacterium tuberculosis in the context of drug discovery. Curr Top Med Chem. 2007; 7:475–88. [PubMed: 17346193]
9. Gokhale RS, Sankaranarayanan R, Mohanty D. Versatility of polyketide synthases in generating metabolic diversity. Curr Opin Struct Biol. 2007; 17:736–43. [PubMed: 17935970]
10. Gokhale RS, Saxena P, Chopra T, Mohanty D. Versatile polyketide enzymatic machinery for the biosynthesis of complex mycobacterial lipids. Nat Prod Rep. 2007; 24:267–77. [PubMed: 17389997]
11. Jackson M, Stadthagen G, Gicquel B. Long-chain multiple methyl-branched fatty acid-containing lipids of Mycobacterium tuberculosis: biosynthesis, transport, regulation and biological activities. Tuberculosis (Edinb). 2007; 87:78–86. [PubMed: 17030019]
12. Kolattukudy PE, Fernandes ND, Azad AK, Fitzmaurice AM, Sirakova TD. Biochemistry and molecular genetics of cell-wall lipid biosynthesis in mycobacteria. Mol Microbiol. 1997; 24:263–70. [PubMed: 9159514]
13. Krithika R, et al. A genetic locus required for iron acquisition in Mycobacterium tuberculosis. Proc Natl Acad Sci U S A. 2006; 103:2069–74. [PubMed: 16461464]
14. Trivedi OA, et al. Dissecting the mechanism and assembly of a complex virulence mycobacterial lipid. Mol Cell. 2005; 17:631–43. [PubMed: 15749014]
15. Portevin D, et al. 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 components. J Biol Chem. 2005; 280:8862–74. [PubMed: 15632194]
16. Cole ST, et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature. 1998; 393:537–44. [PubMed: 9634230]
17. Fontan P, Aris V, Ghanny S, Soteropoulos P, Smith I. Global transcriptional profile of Mycobacterium tuberculosis during THP-1 human macrophage infection. Infect Immun. 2008; 76:717–25. [PubMed: 18070897]
18. Van der Geize R, et al. A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into Mycobacterium tuberculosis survival in macrophages. Proc Natl Acad Sci U S A. 2007; 104:1947–52. [PubMed: 17264217]
19. Conti E, Franks NP, Brick P. Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. Structure. 1996; 4:287–98. [PubMed: 8805533]
20. Gulick AM, Starai VJ, Horswill AR, Homick KM, Escalante-Semerena JC. The 1.75 A crystal structure of acetyl-CoA synthetase bound to adenosine-5′-propylphosphate and coenzyme A. Biochemistry. 2003; 42:2866–73. [PubMed: 12627952]
21. 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]
22. Hisanaga Y, et al. Structural basis of the substrate-specific two-step catalysis of long chain fatty acyl-CoA synthetase dimer. J Biol Chem. 2004; 279:31717–26. [PubMed: 15145952]
23. Goyal A, et al. Crystallization and preliminary X-ray crystallographic studies of the N-terminal domain of FadD28, a fatty-acyl AMP ligase from Mycobacterium tuberculosis. Acta Crystallograph Sect F Struct Biol Cryst Commun. 2006; 62:350–2.
24. Conti E, Stachelhaus T, Marahiel MA, Brick P. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. Embo J. 1997; 16:4174–83. [PubMed: 9250661]
25. Linne U, Schafer A, Stubbs MT, Marahiel MA. Aminoacyl-coenzyme A synthesis catalyzed by adenylation domains. FEBS Lett. 2007; 581:905–10. [PubMed: 17303131]
26. Nakama T, Nureki O, Yokoyama S. Structural basis for the recognition of isoleucyl-adenylate and an antibiotic, mupirocin, by isoleucyl-tRNA synthetase. J Biol Chem. 2001; 276:47387–93. [PubMed: 11584022]
27. Ferreras JA, Ryu JS, Di Lello F, Tan DS, Quadri LE. Small-molecule inhibition of siderophore biosynthesis in Mycobacterium tuberculosis and Yersinia pestis. Nat Chem Biol. 2005; 1:29–32. [PubMed: 16407990]

28. Somu RV, et al. Rationally designed nucleoside antibiotics that inhibit siderophore biosynthesis of Mycobacterium tuberculosis. J Med Chem. 2006; 49:31–4. [PubMed: 16392788]
29. Copeland, RA. Enzymes: A practical introduction to Structure, Mechanism, and Data Analysis. Wiley-VCH, Inc Publications; 2000. Tight binding inhibitors; p. 305-317.
30. Hansen DB, Bumpus SB, Aron ZD, Kelleher NL, Walsh CT. The loading module of mycosubtilin: an adenylation domain with fatty acid selectivity. J Am Chem Soc. 2007; 129:6366–7. [PubMed: 17472382]
31. Black PN, Dirusso CC. Yeast acyl-CoA synthetases at the crossroads of fatty acid metabolism and regulation. Biochim Biophys Acta. 2006
32. Wang F, Langley R, Gulten G, Wang L, Sacchettini JC. Identification of a type III thioesterase reveals the function of an operon crucial for Mtb virulence. Chem Biol. 2007; 14:543–51. [PubMed: 17524985]
33. Arora P, Vats A, Saxena P, Mohanty D, Gokhale RS. Promiscuous fatty acyl CoA ligases produce acyl-CoA and acyl-SNAC precursors for polyketide biosynthesis. J Am Chem Soc. 2005; 127:9388–9. [PubMed: 15984864]
34. Lu YJ, et al. Acyl-phosphates initiate membrane phospholipid synthesis in Gram-positive pathogens. Mol Cell. 2006; 23:765–72. [PubMed: 16949372]
35. Jiang Y, Chan CH, Cronan JE. The soluble acyl-acyl carrier protein synthetase of Vibrio harveyi B392 is a member of the medium chain acyl-CoA synthetase family. Biochemistry. 2006; 45:10008–19. [PubMed: 16906759]
36. Kholodenko BN. Cell-signalling dynamics in time and space. Nat Rev Mol Cell Biol. 2006; 7:165–76. [PubMed: 16482094]
37. Strogatz SH. Exploring complex networks. Nature. 2001; 410:268–76. [PubMed: 11258382]
38. Morphy R, Rankovic Z. Designed multiple ligands. An emerging drug discovery paradigm. J Med Chem. 2005; 48:6523–43. [PubMed: 16220969]
39. Bloch H, Segal W. Biochemical differentiation of Mycobacterium tuberculosis grown in vivo and in vitro. J Bacteriol. 1956; 72:132–41. [PubMed: 13366889]
40. Jain M, et al. Lipidomics reveals control of Mycobacterium tuberculosis virulence lipids via metabolic coupling. Proc Natl Acad Sci U S A. 2007; 104:5133–8. [PubMed: 17360366]
41. Jones TA, Zou JY, Cowan SW, Kjeldgaard. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr A. 1991; 47(Pt 2):110–9. [PubMed: 2025413]
42. Brunger AT, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr. 1998; 54(Pt 5):905–21. [PubMed: 9757107]
43. Barton GJ. ALSCRIPT: a tool to format multiple sequence alignments. Protein Eng. 1993; 6:37–40. [PubMed: 8433969]
44. Castro-Pichel J, et al. Synthesis and antiviral activity of 5′-O-(substituted) sulfamoyl pyrimidine nucleosides. Arch Pharm (Weinheim). 1989; 322:11–5. [PubMed: 2730285]
45. Morrison JF, Walsh CT. The behavior and significance of slow-binding enzyme inhibitors. Adv Enzymol Relat Areas Mol Biol. 1988; 61:201–301. [PubMed: 3281418]
46. Phetsuksiri B, et al. Antimycobacterial activities of isoxyl and new derivatives through the inhibition of mycolic acid synthesis. Antimicrob Agents Chemother. 1999; 43:1042–51. [PubMed: 10223912]
47. Camacho LR, et al. Analysis of the phthiocerol dimycocerosate locus of Mycobacterium tuberculosis. Evidence that this lipid is involved in the cell wall permeability barrier. J Biol Chem. 2001; 276:19845–54. [PubMed: 11279114]
48. Cox JS, Chen B, McNeil M, Jacobs WR Jr. Complex lipid determines tissue-specific replication of Mycobacterium tuberculosis in mice. Nature. 1999; 402:79–83. [PubMed: 10573420]

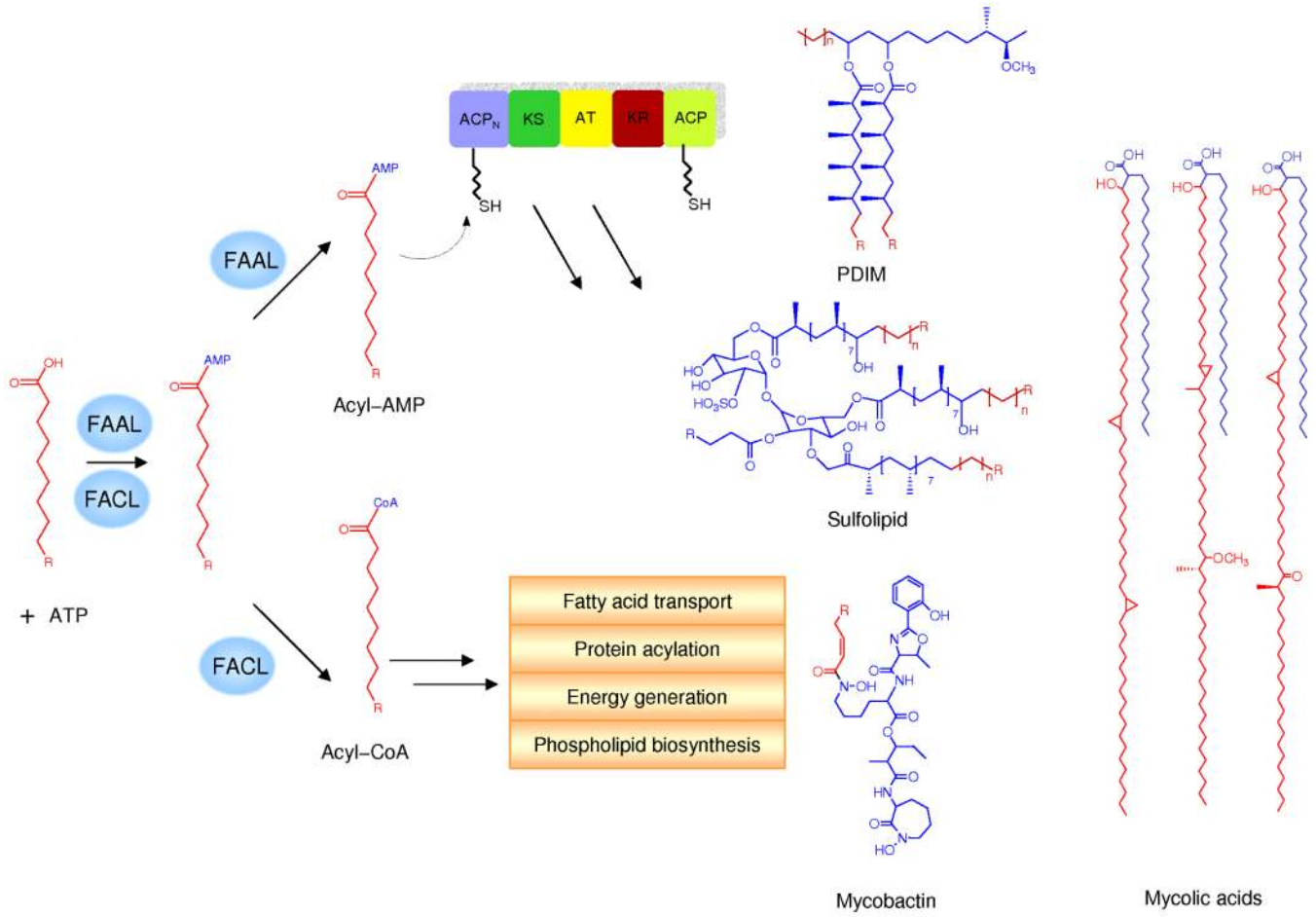


Figure 1. Dichotomy in the metabolic functions of FAALs and FACLs

FAALs and FACLs utilize fatty acid pools and activate them to a common acyl- adenylate intermediate. FACLs convert fatty acids to acyl-CoA and utilize them for fatty acid transport, protein acylation, energy generation, and phospholipid biosynthesis. FAAL produced acyl- adenylate is utilized by polyketide synthase enzymes for the synthesis of complex lipids like PDIM, sulpholipids, mycolic acids, and mycobactin.

Europe PMC Funders Author Manuscripts

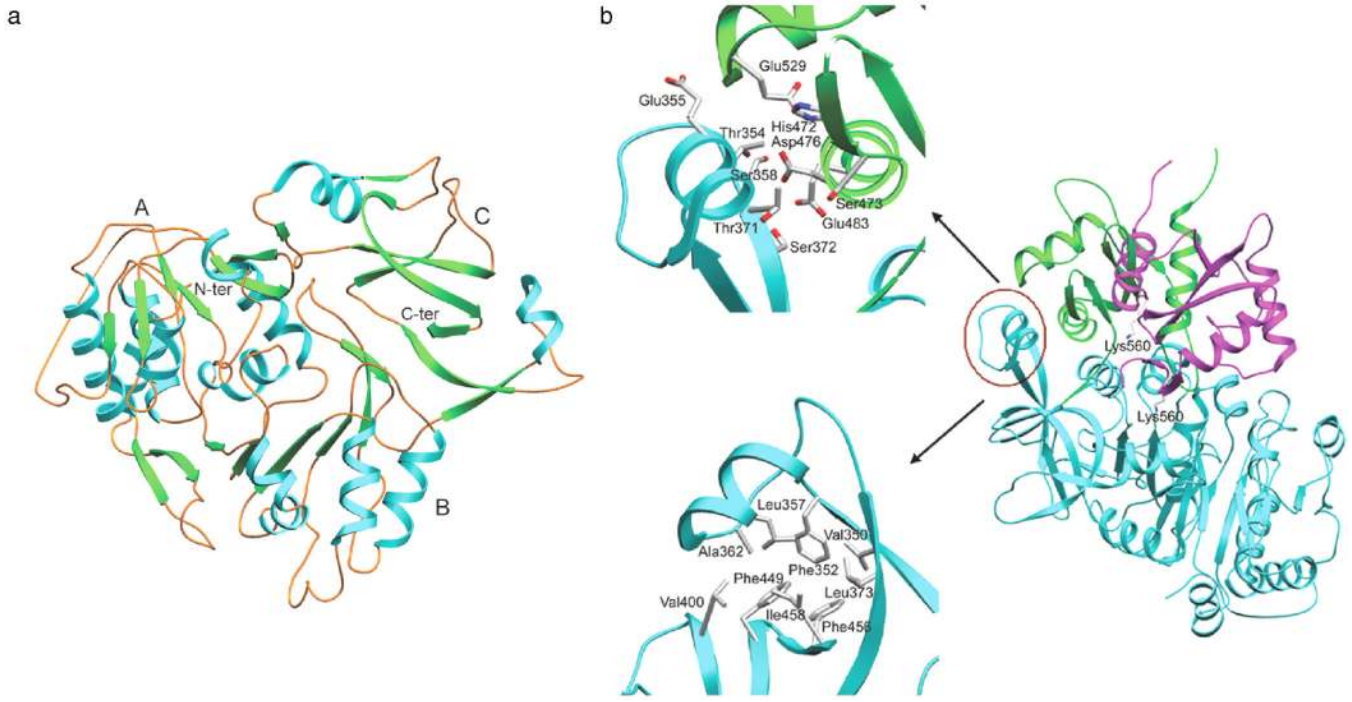


Figure 2. Crystal structure of FAAL28 N-terminal domain and structural analysis of the FAAL insertion

(**a**) Structure of FAAL28 N-terminal domain protein N1, demonstrates α+β topology of the A and B subdomain and a distorted β barrel for C subdomain, N and the C-termini are marked.

(**b**) FAAL28 C-terminal domain of FAAL28 modeled on the basis of PheA (1amu) structure in putative AMP bound conformation (green), acetyl CoA synthetase structure (1pg4) in putative CoA bound conformation (magenta), and FAAL28 N-terminal domain structure (cyan) with highlighted FAAL specific insertion. Inset highlights residues in the FAAL specific insertion facing towards the putative C-terminal domain (top) and strong hydrophobic interaction of residues in the FAAL specific insertion with the N-terminal domain (bottom).

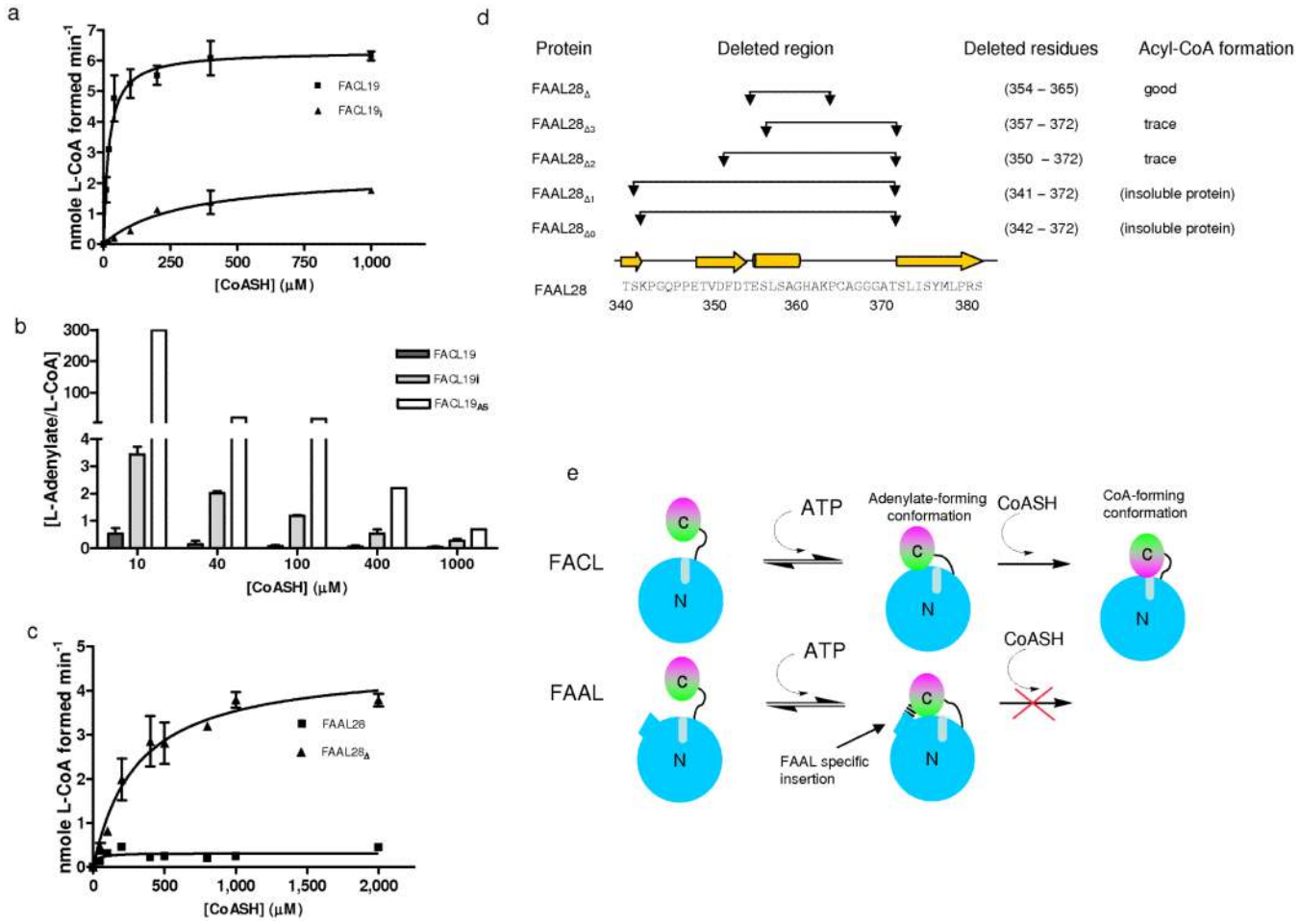


Figure 3. Interconversion of FACL and FAAL activities

(**a**) Kinetic analysis of CoA formation by wild type FACL19 and FACL19$_{\mathbf{i}}$ proteins by Michaelis-Menten plot. The data point is represents mean ± SEM of three independent experiments.

(**b**) The ratio of acyl-adenylate to acyl-CoA formed by FACL19, FACL19$_{\mathbf{i}}$ and FACL19$_{\mathbf{AS}}$ at different concentrations of CoASH substrate.

(**c**) Kinetic analysis of CoA formation by wild type FAAL28 and FAAL28$_{\mathbf{\Delta}}$ proteins by Michaelis-Menten plot. The data point is represents mean ±SEM of three independent experiments.

(**d**) Various deletion proteins of FAAL28 generated for the study mapped to the sequence and their corresponding biochemical characteristics.

(**e**) Schematic representation of proposed structural role of the FAAL specific insertion in acyl-AMP formation.

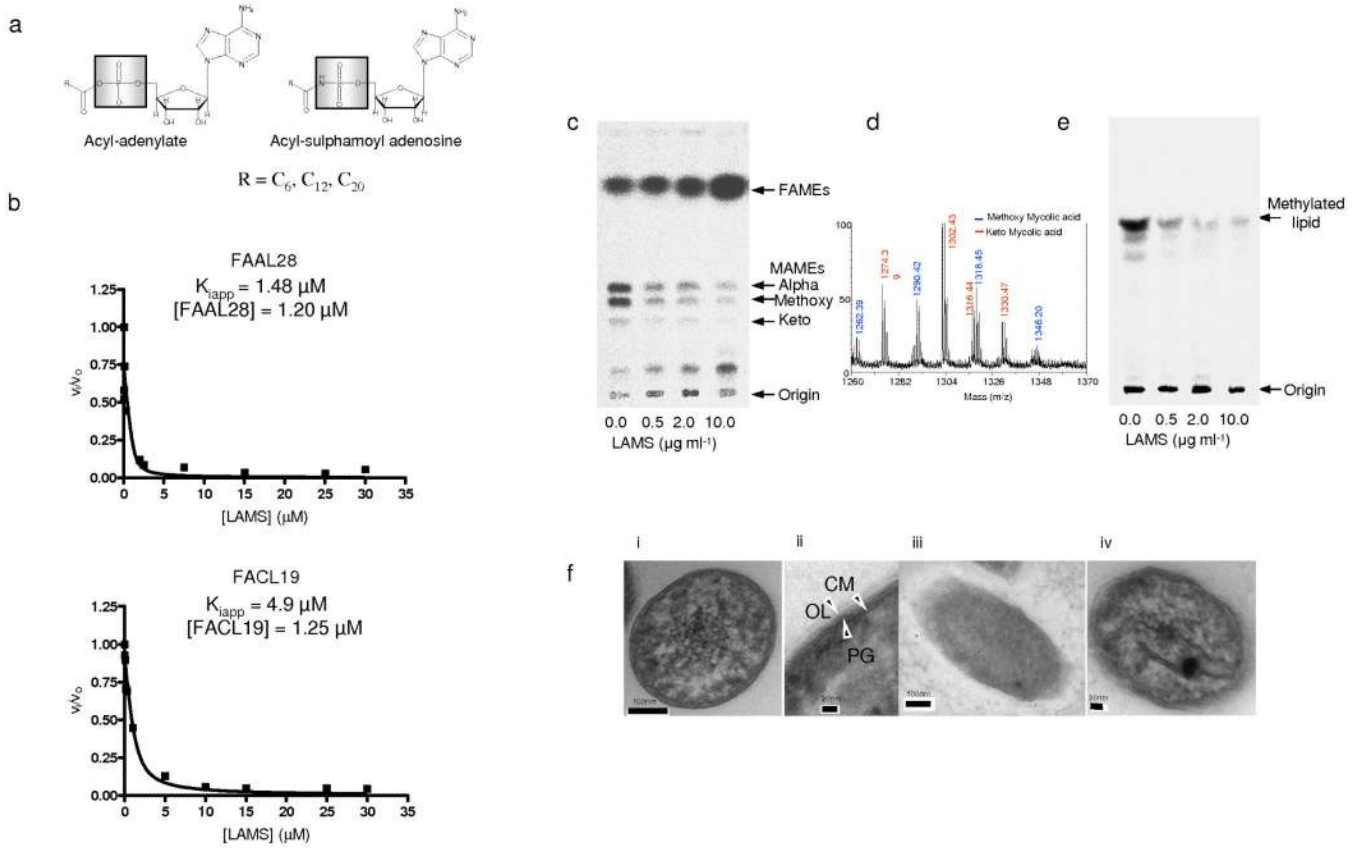


Figure 4. Inhibition of FAAL and FACL enzymes by acyl-sulfamoyl analogues

(a) Chemical structures of acyl-adenylate and the non-hydrolyzable acyl-sulfamoyl adenosine analogues.

(b) IC_{50} values of FAAL28 and FACL19 proteins with Lauroyl-sulfamate (LAMS).

(c) ^{14}C-acetate incorporation assay to analyze fatty acid and mycolic acid biosynthesis.

(d) Confirmation of mycolic acids by mass spectrometry

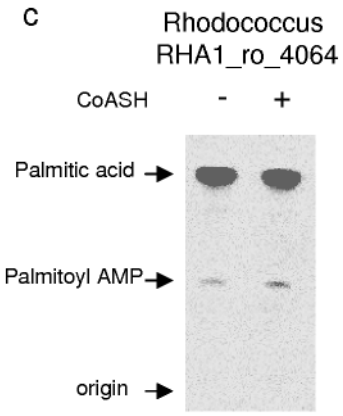
(e) ^{14}C-propionate labeling to study biosynthesis of extractable lipids like PDIM and SL-1.

(f) Ultrastructural studies of *M. tuberculosis* to study the effect of acyl-sulfamoyl analogues on cell morphology. (i) Transmission electron micrographs of thin section of untreated *M. tuberculosis* as control. (ii) Untreated cells show a thin cytoplasmic membrane (CM), and a thick electron-dense layer corresponding to the cell wall peptidoglycan (PG) and a thin and poorly stained outer layer (OL). (iii) and (iv) Cells treated with hexanoyl- and lauroyl-AMS respectively. These cells lack the typical lipid coat and have irregular shape.

a

Organism	FAAL	PKS
Rhodococcus sp RHA1	RHA1_ro_4064	RHA1_ro_4065
Mycobacterium tuberculosis	FAAL32	PKS13
Corynebacterium glutamicum	CG3179	CG3178
Nocardia farcinica	Nfa1880	Nfa1890

c



b

FAAL insertion

```
Nfa1880 VTTSSMKKFNEAFAPYGLPKTAIKPCYGMAEATLFVSATKAEDEAKVTYVDRNELNAGRMVKVDPSAENAIAQVSCGYVALSQWAV
ro_4064 VTTSSMRKFNEAFAPYGLPKTAIKPSYGMAEATLFVSSTRHSDEARVIYVDRDELNAGRVVKVDQDAENAIPQVSCGYVARSQWAA
FadD32  VSPASMRKFFEAFAPYGLKQTAVKPSYGLAEATLFVSTTPMDEVPTVIHVDRDELNNQRFVEVAADAPNAVAQVSAGKVGVSEWAV
cg3179  VTENALTTFREAFEPYGLPVQTLRPSYGLAEASLLVTTPQTENRPLISYFDREALAENRVELVEKGNNKAVAFVSNGQVAAPQQLV
FadD28  VQAATIKRFADRFARFNLQERVIRPSYGLAEATVYVATSKPGQPPETVDFDTESLSAG---HAKPCAGGGATSLISYMLPRSPIVR
FadD19  FSPSIKEKLLELLPNRVITDSIGSSETGFGGTSVVAAGQAHGGGP--------------------------------RVRIDHRTV
        .       : : :    :     .   *:. ::: .:       .                                  :
```

Figure 5. Characterization of FAAL in other *Actinomycetes*

(a) Conserved FAAL, PKS homologous pair across mycolates

(b) Multiple sequence alignment of Nfa1880 (*Nocardia farcinica*), ro_4064 (*Rhodococcus sp RHA1*), FAAL32 (*Mycobacterium tuberculosis*), cg3179 (*Corynebacterium glutamicum*) with the mycobacterial FAAL28 (acyl-AMP forming) and FACL19 (acyl-CoA forming). The bold line on top highlights the FAAL specific insertion.

(c) Enzyme assay of ro_4064 protein carried out using palmitic acid in the presence or absence of 1mM CoASH. The products were separated on TLC and the autoradiogram is depicted.

Europe PMC Funders Author Manuscripts

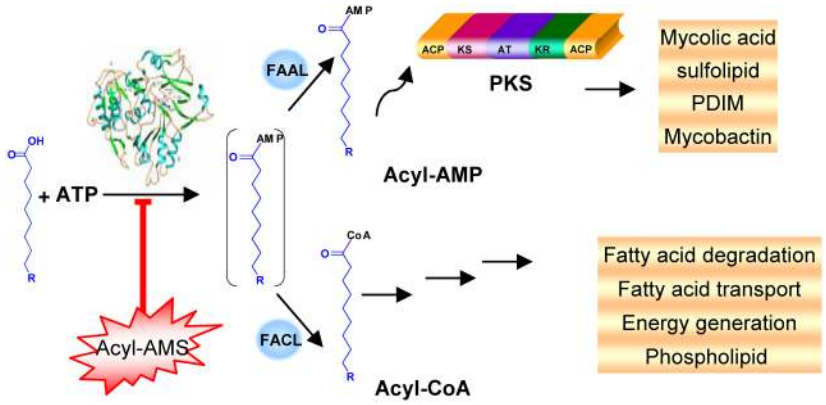


Figure 6.

Europe PMC Funders Author Manuscripts