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).

† Authors have contributed equally
•Corresponding Authors.

SUPPLEMENTARY INFORMATION

Contents:
Supplementary Figures (5 figures) 2

Supplementary Tables (4 tables) 8

Supplementary Method 12

a

```
FAAL23   QTTAYL QYTSGSTRTPAGVMITYKNILANFQQ 194
FAAL28   PSTAYL QYTSGSTRTPAGVVMSHQNVRVNFEQ 194
FAAL24   PETAYL QYTSGSTRTPAGVMVSNKNVFANFEQ 199
FAAL25   PDILYL QYTSGSTRTPAGVMVSNKNLFANFEQ 198
FAAL21   PSAAYL QYTSGSTRAPAGVMISHRNLQANFQQ 193
FAAL26   TGAAYL QYTSGSTRTPAGVIVSHTNVIANVTQ 239
FAAL29   PSTAYL QYTSGSTRAPAGVVLSHKNVITNCVQ 234
FAAL30   DWSSYL QYTSGSTANPRGVVLSMRNVTENVDQ 194
FAAL31   DAVSHL QYTSGSTRPPVGVEITHRAVGTNLVQ 224
FAAL32   ETVAYL QYTSGSTRIPSGVQITHLNLPTNVVQ 217
FACL3    QDVSDI LFTSGTTGRSKGVLCAHRQSLSASAS 200
FACL35   SDPVNI QYTSGTTAYPKGVTLSHRNILNNGYL 219
FACL2    RRASII ILTSGTTGTPKGANRNTPPTLAPIGG 250
FACL12   SHGKVI LLTSGTTGTPKGA-RHSGGGIGTLKA 222
FACL19   ADAIYL LYTGGTTGFPKGVMWRHEDIYRVLFG 197
FACL5    DSPALI MYTSGTTGRPKGAVLTHANLTGQAMT 214
FACL13   DDNLFI MYTSGTTGHPKGVVHTHESVHSAASS 187
FACL1    MDPFMM IFTSGTSGNPKAVPVSHLMATFAGRS 181
FACL17   ADLFML IFTSGTSGDPKAVKCSHRKVAIAGVT 180
FACL8    DQVIGL TYTGGTTGKPKGVIGTAQSIATMTSI 239
FACL33   EGPAVL QGTAGSTGAPRTAILSPGAVLSNLRG 178
FACL7    PDDAMI MFTGGTTGLPKMVPWTHANIASSVRA 194
FACL4    REGDLL QYSSGTTGRPKGIKRELPHVSPDAAP 215
FACL6    KDTAFY IFTSGTTGFPKASVMTHHRWLRALAV 228
FACL22   DALAYA TYTSGTTGPPKAAIHRHADPLTFVDA 176
FACL34   DDDAYV QYTSGSTAAPRGVVITYRNLLSNMRA 200
FACL10   EDPLAM IFTSGTTGEPKAVLLANRTFFAVPDI 200
FACL15   TDPATL IYTSGTTGRPKGCQLTQSNLVHEIKG 210
```

FAAL (FAAL23–FAAL32)

FACL (FACL3–FACL15)

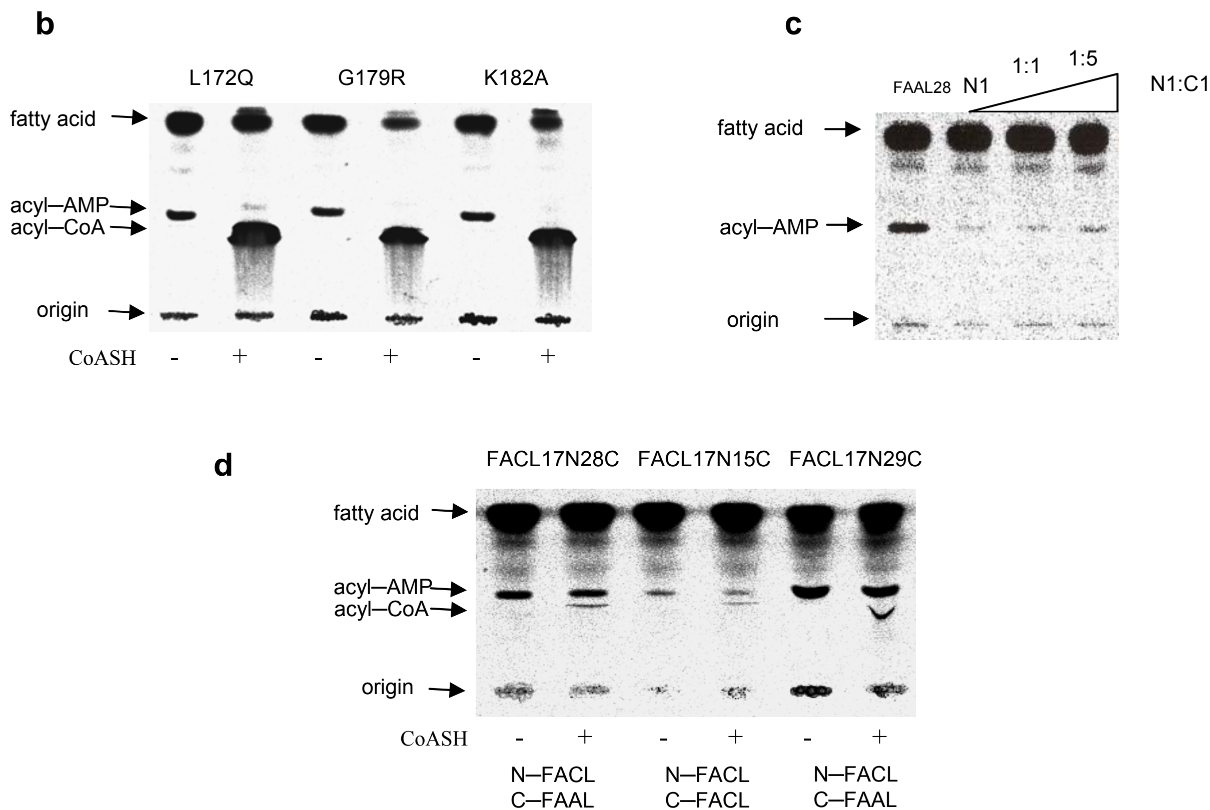


Fig S1. Sequence based analyses of FAAL and FACL proteins

(**S1a**)**.** Sequence alignment of representative FAAL and FACL homologues were carried out using ClustalW and a part of the alignment is depicted for clarity. The class-specific conservation of residues in the nucleotide binding motif is indicted by arrows.

(**S1b**)**.** Enzyme assay carried out with FACL19 mutant proteins, FACL19L172Q (lane1 and 2), FACL9G179R (lane 3 and 4) and FACL19K192A (lane 5 and 6) in the absence or presence of 1mM CoASH respectively and the products were separated on TLC. Acyl-AMP and acyl-CoA products were identified based on Rf.

(**S1c**). Enzyme assay performed with FAAL28 full length protein (lane 1), N1 protein (FAAL28; 1–460 amino acids) (lane 2), N1 protein with equimolar amount of C1 (FAAL28 460–580 amino acids) (lane 3), N1 and C1 in 1:5 molar ratio (lane 4) and the products were separated on TLC. Acyl-AMP and acyl-CoA products were identified based on Rf.

(**S1d**)**.** Enzyme assay carried out with FACL17 hybrid proteins, FACL17N28C (lane1 and 2), FACL17N15C (lane 3 and 4) and FACL17N29C (lane 5 and 6) in the absence or presence of 1mM CoASH respectively and the products were separated on TLC. Acyl–AMP and acyl–CoA products were identified based on Rf.

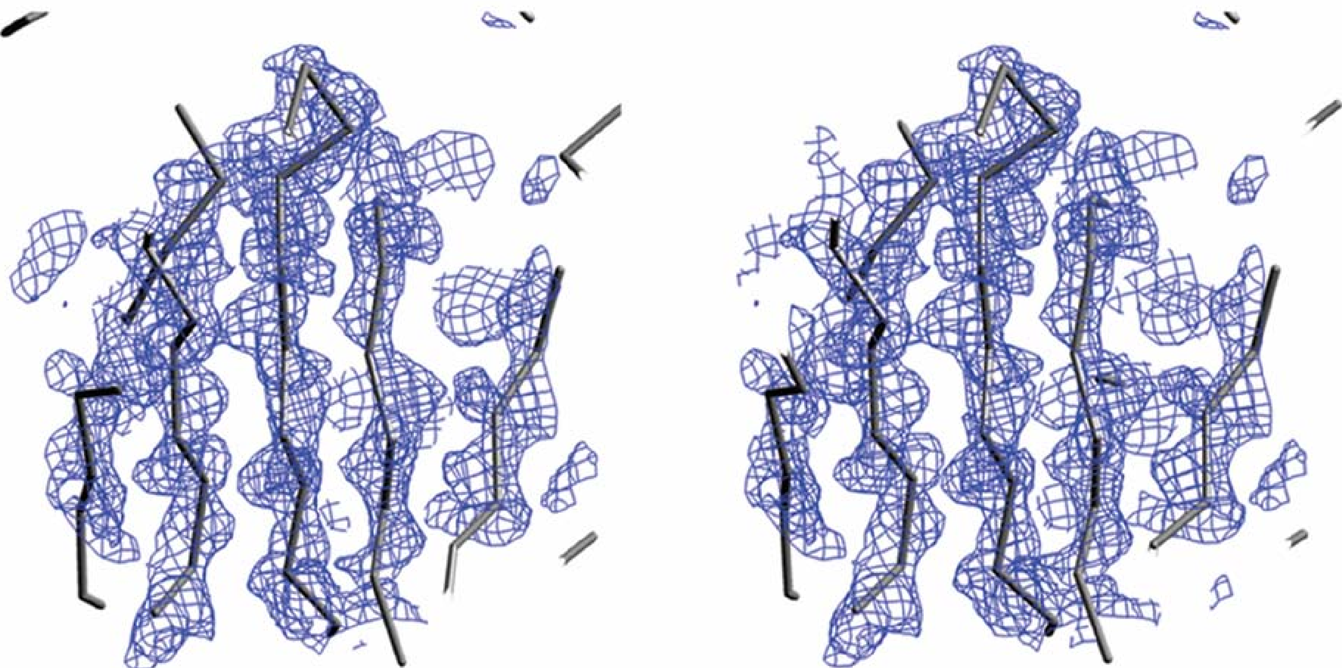


Fig S2. Structure of FAAL28 N terminal domain

A stereo representation of the experimental electron density map of a beta sheet of FAAL28 N terminal domain contoured at 0.8 σ

a

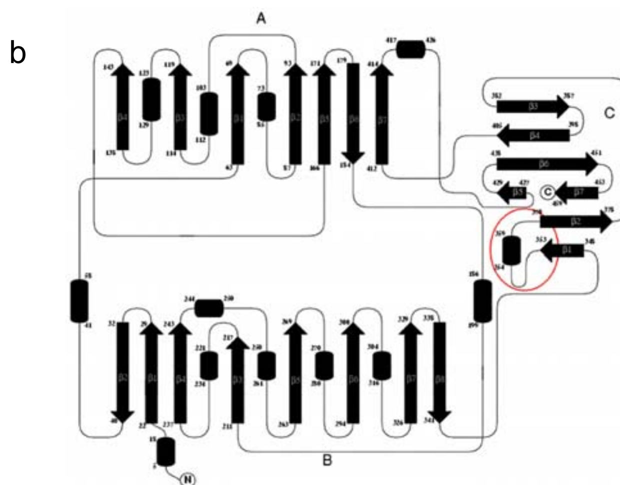
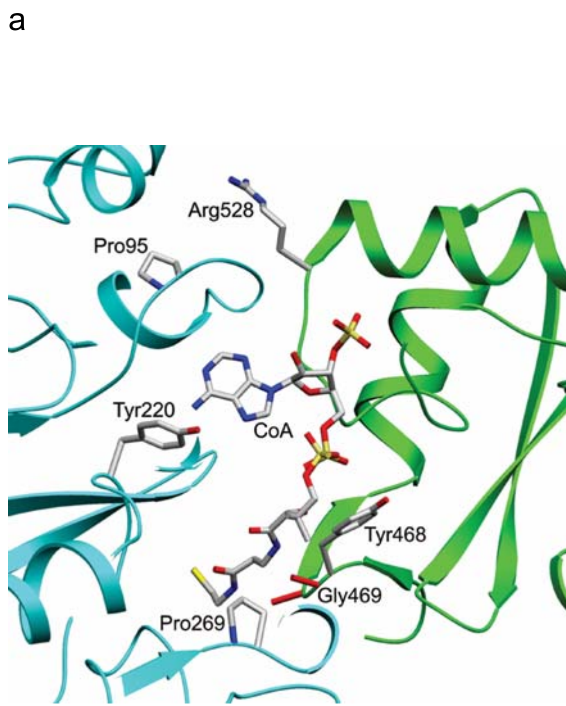


Fig S3. Structure based sequence comparison and topology of FAAL28

(**S3a**). Various FAAL proteins like FAAL25 and FAAL32, a FACL member (FACL19) and other structurally characterized homologous protein sequences were compared. The FAAL specific insertion is indicated by a line on top of the FAAL28 sequence spanning residues 351−372. Residues (354−365) deleted in FAAL28Δ mutant, used for biochemical studies, are indicated in magenta.

(**S3b**). Topology diagram of N1 depicts the secondary structural elements and their connectivity in the N terminal domain of FAAL28, the FAAL specific insertion is encircled.

a



b

	N-terminal domain residues			C-terminal domain residues		
ACS	F163	G165	A357	S523	G524	R584
FAAL28	Y220	P95	P269	Y468	G469	R528
FAAL21	Y219	P95	P268	Y467	G468	N526
FAAL23	Y220	P95	P269	Y468	G469	R524
FAAL24	Y225	P95	P274	Y473	G474	R532
FAAL25	Y224	P97	P273	Y472	G473	R531
FAAL26	Y220	P97	P269	D470	G471	R529
FAAL29	Y257	P132	P306	D506	G507	R565
FAAL30	Y220	P95	P269	D472	G473	R531
FAAL31	Y245	P125	P296	D508	G509	R573
FAAL32	F237	P118	P289	D498	G499	R579
FAAL33	Y203	P87	P251	A421	G422	L477
FAAL34	H220	R90	P269	H437	G438	R489

Fig S4 : Conservation of CoASH binding residues in mycobacterial FAAL proteins

(**S4a**). CoASH molecule docked into the active site of FAAL28 with C - terminal domain modeled in CoA binding conformation.

(**S4b**). Conserved residues involved in CoA binding highlighted across different FAALs.

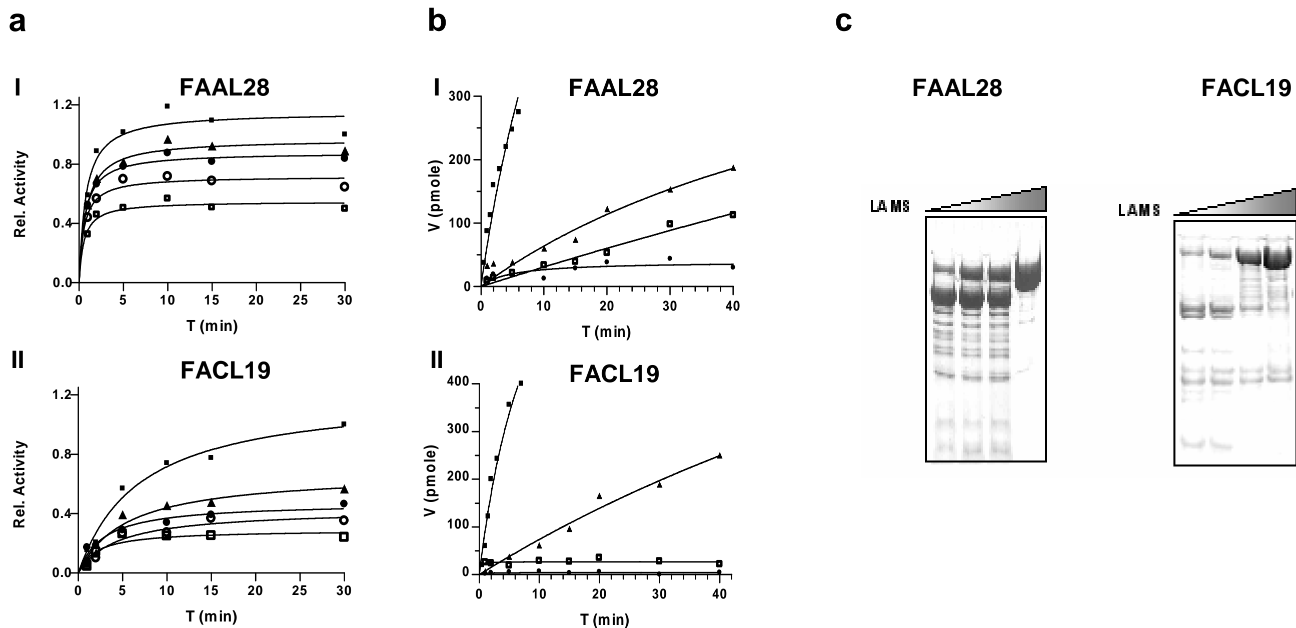


Fig S5. Kinetic analysis of LAMS inhibition

(**S5a**). The Time dependent inhibition of FAAL (I) and FACL (II) proteins as a function of LAMS concentration. The concentrations of LAMS tested were: (■) 0nM, (▲) 25nM, (●) 50nM, (○) 75nM and (□) 200nM.

(**S5b**). Enzymatic turnover recovery assays. FAAL28 (I) and FACL19 (II) were incubated with LAMS for (▲) 15, (□) 30 and (●) 60 min. The control assays were performed with proteins incubated in the absence of LAMS for (■) 60min.

(**S5c**). Limited proteolysis of FAAL28 and FACL19 proteins in the presence of increasing concentrations of LAMS

Supplementary Table 1 : FAAL and FACL proteins generated for the inter-conversion studies and their properties

Protein	Mutation (Insertion/ deletion)	Protein expression and solubility status	Expected activity	Observed activity	Comments
FACL19	-	soluble	Acyl-CoA forming	Acyl-CoA forming	Km 24µM
$FACL19_I$	FAAL28 residues 342 to 372 inserted in place of residues 348 to 357 in FACL19	Soluble and comparable expression to WT	Acyl-AMP-forming	Acyl-AMP forming (at low CoASH concentrations)	Km for CoASH could not be determined as saturation of CoA formation was not observed
FACL6	-	Soluble	Acyl-CoA forming	Acyl-CoA forming	
$FACL6_i$	FAAL28 residues 342 to 372 inserted in place of residues 362 to 376 in FACL6	Poorly expressed and present in inclusion body	Acyl-AMP-forming	N.D	
FAAL28	-	Soluble	Acyl-AMP-forming	Acyl-AMP-forming	
$FAAL28_{\Delta 0}$	Deletion of residues 342 to 372	Expressed but is present in inclusion body	Acyl-CoA forming	N.D	
$FAAL28_{\Delta 1}$	Deletion of residues 341 to 372	Expressed but is present In inclusion body	Acyl-CoA forming	N.D	
$FAAL28_{\Delta 2}$	Deletion of residues 350 to 372	Expressed in small amounts and does not purify to homogeneity	Acyl-CoA forming	Acyl-CoA forming	No adenylate detected, and in the presence of CoASH synthesizes CoA
$FAAL28_{\Delta 3}$	Deletion of residues 357 to 372	Expressed in small amounts and does not purify to homogeneity	Acyl-CoA forming	Acyl-CoA forming	No adenylate detected and in the presence of CoASH synthesizes CoA
$FAAL28_{\Delta}$ (used for kinetics)	Deletion of residues 354 to 365	Expresses and purified to homogeneity	Acyl-CoA forming	Acyl-CoA forming	Km for CoASH >500µM
FAAL30	-	Soluble	Acyl-AMP-forming	Acyl-AMP-forming	
$FAAL30_{\Delta}$	Deletion of residues 363 to 372	Expressed in inclusion body	Acyl-CoA forming	N.D	
$FACL19_{AS}$	A A S A A S S A S S A A S sequence of amino acids inserted in place of residues 348 to 357 in FACL19	Soluble and comparable expression to WT	Acyl-AMP-forming	FAAL makes very small amounts of CoA	Km for CoASH could not be determined as saturation of CoA formation was not observed

Supplementary Table 2 : Fatty acid substrate specificity determined biochemically for different FAAL and FACL proteins

FadD	C_6	C_{12}	C_{16}
FACL6	+	+	+
FACL8	+	+	+
FAAL9	N.D.	+	+
FACL10	-	+++	++
FACL13	N.D.	+++	++
FACL15	-	+	+
FACL17	+	+	+
FACL19	+	+	+
FAAL23	-	+	+
FAAL26	-	+	+
FAAL28	-	+	+
FAAL29	N.D.	+	N.D.
FAAL30	N.D.	+	-
FAAL32	N.D.	-	+

Supplementary Table 3 : Primers used in the study

Mutant protein	Primers sequence
FACL19L172Q	5’GAC GCC ATC TAT CTG CAG TAC ACC GGC GGC ACC 3’ 5’GGT GCC GCC GGT GTA CTG CAG ATA GAT GGC GTC 3’
FACL19G179R	5’CC GGC GGC ACC ACC CGT TTC CCC AAG GGT GTG 3’ 5’CAC ACC CTT GGG GAA ACG GGT GGT GCC GCC GG 3’
FACL19K182A	5’CC ACC GGT TTC CCC GCG GGT GTG ATG TGG CG 3’ 5’CG CCA CAT CAC ACC CGC GGG GAA ACC GGT GG 3’
Hybrid protein	
FACL17N28C	5'GGC GGC GTC TAC CAC ACC GGT GAC CTC GCC TAT CGC 3' 5'GCG ATA GGC GAG GTC ACC GGT GTG GTA GAC GCC GCC 3‘ and 5'GGT CCT TGG CTA AGA ACC GGT GAC TCA GGT TTC GTC3' 5'GAC GAA ACC TGA GTC ACC GGT TCT TAG CCA AGG ACC 3'
FACL17N15C	5'GGC GTC TAC CAC AGT GGA GAT CTC GCC TAT CGC GAC 3' 5'GTC GCG ATA GGC GAG ATC TCC ACT GTG GTA GAC GCC 3‘ and 5'GAC GGC TGG TTC AAG ACC GGA GAT CTC GGT GCG GTG 3' 5'CAC CGC ACC GAG ATC TCC GGT CTT GAA CCA GCC GTC 3'
FACL17N29C	5'GGC GTC TAC CAC AGT GGA GAT CTC GCC TAT CGC GAC 3' 5'GTC GCG ATA GGC GAG ATC TCC ACT GTG GTA GAC GCC 3‘ and 5'TGG CTG CGG ACC GGA GAT CTC GGC GTC ATT TTC GAG G 3' 5'C CTC GAA AAT GAC GCC GAG ATC TCC GGT CCG CAG CCA 3'
FadD homolog	
Rhodococcus RHA1 ro_4064	5’ TT GAA TTC CTA TTC GGA AAC GCT ATC GGG GAA 3’ 5’ TTG AAT TCG TTC GGA AAC GCT ATC GGG GAA GGC

Insertion/ deletion	Primer sequence	Engineered restriction site
FACL19$_i$	5’ CGG TAC CAG CGT CGC TAG CGC CGG GCA GGC GC 3’ 5’ GCG CCT GCC CGC CGC TAG CGA CGC TGG TAC CG	Nhe I
	5’ GGC CCC GGG CTA GCA TCG ACC ATC GCA CCG 3’ 5’ CGG TGC GAT GGT CGA TGC TAG CCC GGG GCC 3’	Nhe I
FAAL28$_{\Delta 0}$	5’ GCT AGC GCA CGG CTT CGC ATG GCC GGC GGA 3’	Nhe I
FAAL28$_{\Delta 1}$	5’ GCT AGC TAA ACT TTC AGT ATC GAA GTC GAC 3’	Nhe I
FAAL28$_{\Delta 2}$	5’ GCT AGC GAC GGT CTC CGG TGG TTG ACC GCC AAA 3’	Nhe I
FAAL28$_{\Delta 3}$	5’ GGC GGC GCT ACT AGT TTG ATC AGC TAC ATG 3’ 5’ CAT GTA GCT GAT CAA ACT AGT AGC GCC GCC 3’	Nhe I
	5’ GGC GGC GCT ACT AGT TTG ATC AGC TAC ATG 3’ 5’ CAT GTA GCT GAT CAA ACT AGT AGC GCC GCC 3’	Spe I
FAAL28$_{\Delta}$	5’ CCG TCG ACT TCG ATG CTA GCA GTT TAT CCG CCG GCC 3’ 5’ GGC CGG CGG ATA AAC TGC TAG CAT CGA AGT CGA CGG 3’	Nhe I
	5’CCG GCC ATG CGA AGA CTA GTG CAG GCG GCG GCG C 3’ 5’ GCG CCG CCG CCT GCA CTA GTC TTC GCA TGG CCG G 3’	Spe I
FACL19$_{AS}$	5’ CTA GTG CTG CCA GTG CAG CGA GCT CTG CTT CCT CAG CCT CGG CAG CGG 3’ 5’ CTAG CCG CTG CCG AGG CTG AGG AAG CAG AGC TCG CTG CAC TGG CAG C 3’	Nhe I

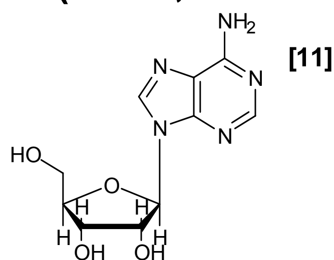
Supplementary Table 4 : Data collection and refinement statistics for FAAL28 N-terminal domain MAD (SeMet) structure.

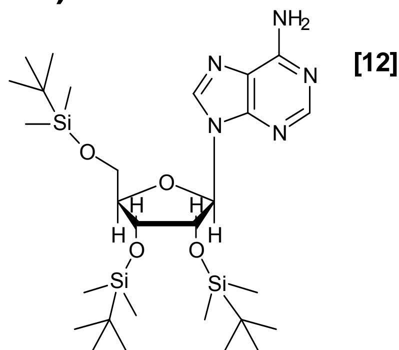
		FAAL28 N-terminal domain SeMet*	
Data collection			
Space group		P2(1)2(1)2(1)	
Cell dimensions			
a, *b*, *c* (Å)		50.97, 60.74, 136.54	
α, β, γ (°)		90, 90, 90	
	Peak	*Inflection*	*Remote*
Wavelength	0.9792	0.9800	0.9776
Resolution (Å)	20.0-2.35 (2.39-2.35)	20.0-2.35 (2.39-2.35)	20.0-2.35 (2.39-2.35)
R_{sym} or R_{merge}	9.3 (43.4)	7.7 (41.4)	8.4 (42.7)
$I / \sigma I$	19.72 (3.44)	23.13 (3.54)	21.87 (3.36)
Completeness (%)	99.5 (93.8)	99.3 (93.2)	99.7 (95.9)
Redundancy	7.0 (5.5)	6.9 (5.5)	6.3 (4.9)
Refinement			
Resolution (Å)		20-2.35 (2.43-2.35)	
No. reflections		17986 (1436)	
R_{work} / R_{free}		0.207/0.267 (0.286/0.356)	
No. atoms			
Protein		3,231	
Ligand/ion		-	
Water		240	
B-factors			
Protein		31.5	
Ligand/ion		-	
Water		38.3	
R.m.s deviations			
Bond lengths (Å)		0.006	
Bond angles (°)		1.3	

*Data collected from one crystal. **Highest-resolution shell is shown in parentheses.

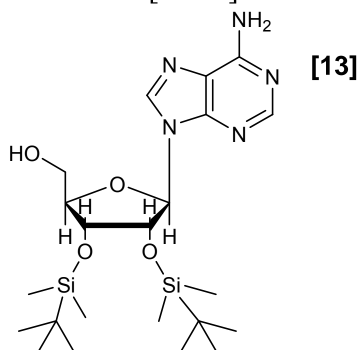
The coordinates of FAAL28 N-terminal domain have been deposited in the Protein Data Bank with access code 3E53.

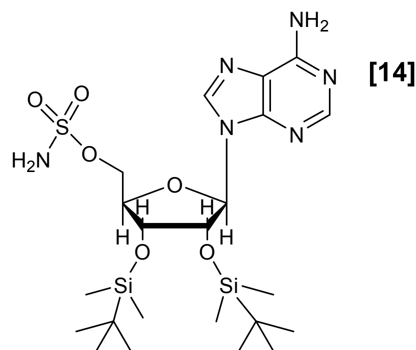
Supplementary method : Synthesis of acyl-AMS analogues (LAMS, HAMS and AAMS)





2', 3', 4'-O,O,O-tris(t-butyldimethylsilyl)adenosine [12].
tert-butyldimethylsilyl chloride (3.5 equiv) in anhydrous DMF was added to a solution of (-) adenosine **[10]** imidazole (9.0 equivalents) in anhydrous DMF and stirred overnight. The reaction mixture was then diluted with CH_2Cl_2, washed thrice with saturated $NaHCO_3$ and dried with $MgSO_4$, filtered and concentrated. The white residue was characterized as 2', 3', 4'-O,O,O-tris(t-butyldimethylsilyl)adenosine **[12].**
Compound **[12]** is identical to the one reported in[1].
ESI-MS m/z of 610.2 $[M+H]^+$ was obtained for a calculated mass of 610.36

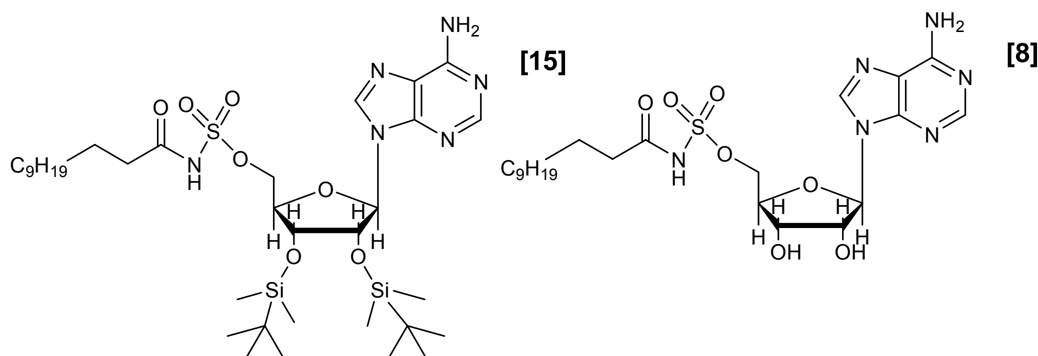




2,'3'-bis(t-butyldimethylsilyl)adenosine [13]
A mixture of TFA and H_2O (1:1) was added to a cooled solution of **[12]**. The reaction mixture was stirred at 0 °C for 5 h. The aqueous layer was extracted with ethyl acetate after adding $NaHCO_3$ solution. This extract was washed with H_2O and dried with $MgSO_4$, filtered and concentrated. The white solid obtained was characterized as 2', 3'-O,O-bis(t-butyldimethylsilyl)-O-sulfamoyladenosine **[13]**.
Compound **[13]** is identical to the one reported in[1].
ESI-MS m/z of 496.3 $[M+H]^+$ was obtained for a calculated mass of 496.27

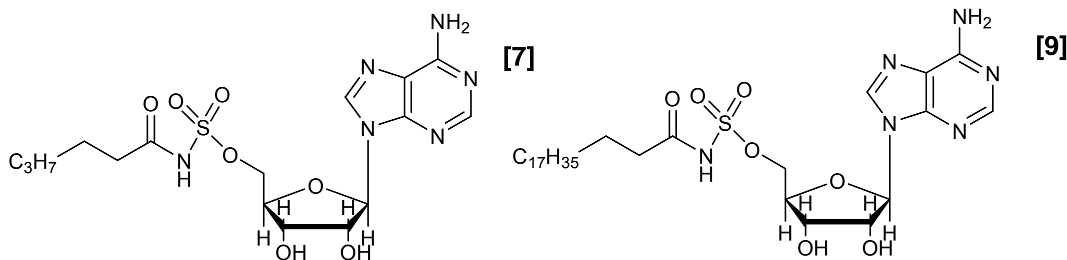
2', 3'-O,O-bis(t-butyldimethylsilyl)-O-sulfamoyladenosine [14]
Bis(tri-butyltin)oxide (3.5 euivalents) was added dropwise to a solution of [13] in anhydrous benzene. The resulting white suspension was refluxed with stirring for 2 h and sulfamoyl chloride (4.5 equivalents) in dioxane was added dropwise to the reaction mixture, stirred for an additional hour at 5 °C. The residue obtained by vacuum concentration was rinsed with hot (40 °C) hexane. The solid was washed with 1 N NH_3 solution in MeOH and dried.

The residue was characterized as 2’, 3’- O,O-bis(t-butyldimethylsilyl)-O-(N-dodecanoyl)sulfamoyladenosine **[14]**.
Compound **[14]** is identical to the one reported in[1].
ESI-MS m/z (pos) 575.3 $[M+H]^+$ for a calculated mass of 575.25



2’, 3’-O,O-bis(t-butyldimethylsilyl)-O-(N-dodecanoyl) sulfamoyladenosine [15]
A solution of lauric acid (3.0 equivalents) and 1,1- carbonyldiimidazole (3.6 equivalents) in anhydrous acetonitrile was stirred at 60 °C for 2 h under Argon atmosphere. The reaction mixture was cooled to room temperature. A mixture of **[14]** and DBU (1.5 equivalents) was then added dropwise to the reaction mixture. The resulting yellow solution was again stirred at 60 °C. After an additional 30 min, the reaction mixture was diluted with H_2O. The aqueous layer was extracted thrice with ethyl acetate, washed with 1N HCl, saturated aqueous solution of $NaHCO_3$, dried with $MgSO_4$, filtered and concentrated. The residue was purified by flash chromatography and was characterized as 2’,3’-O,O-bis(t-butyldimethylsilyl)-O-(N-dodecanoyl)sulfamoyladenosine **[15]**.
ESI-MS m/z (pos) 757.3 $[M+H]^+$ for a calculated mass of 757.41

5’-O-(N-dodecanoylsulfamoyl) adenosine (LAMS) [8]
TBAF (1.0 equivalents in THF, 2.5 equivalents) was added dropwise to a solution of **[15]** in anhydrous THF. The solvent was evaporated after stirring for 30 min. The residue was purified by flash chromatography to yield 5’-O-(N-dodecanoylsulfamoyl) adenosine **[8]** as a white solid.
ESI-MS m/z (pos) 529.7 $[M+H]^+$ for a calculated mass of 529.24
^{1}H-NMR (300 MHz, DMSO-d6) δ 8.4 (s, 1H), δ 8.13 (s, 1H), δ 7.25 (s, 2H), 5.90 (d, 1H, J=5.7), 5.45 (d, 1H, J= 6.0), 5.30 (d, 1H, J=4.2), 4.59 (m, 1H), δ 4.15 – 3.90 (m, 4H, J= 29.7), δ 1.96 (t, 2H), δ 1.22 (m, 20H), δ 0.83 (t, 3H)
^{13}C-NMR (75 MHz, DMSO-d6) 178.15, 155.99, 152.57, 149.64, 139.38,118.86, 115.35, 86.79, 82.76, 73.61, 70.90, 67.05, 31.28, 29.07, 29.04, 25.89, 22.07, 13.94



5'-O-(N-hexanoylsulfamoyl) adenosine (HAMS) **[7]** and 5'-O-(N-icosanoylsulfamoyl) adenosine (AAMS) **[9]** were synthesized in a manner identical to **[8]**, from **[14]** using hexanoic acid and icosanoic acid respectively and characterized using mass spectrometry and NMR as follows:

5'-O-(N-hexanoylsulfamoyl) adenosine (HAMS) [7]
ESI-MS m/z (pos) 445.5 $[M+H]^+$ for a calculated mass of 445.15
^{1}H-NMR (300 MHz, CD_3OD) δ 8.68 (s, 1H) , δ 8.12 (s, 1H) , δ 6.03 (d, 1H) , δ 3.9-3.6 (m, 1H) , δ 2.18 (t, 2H) , δ 1.57 (m, 2 H) , δ 1.29 (m, 4H), δ 0.96 (t, 3H),.
^{13}C-NMR (300 MHz, CD_3OD) δ 156.1, 152.4, 149.9, 140.4, 119.5, 90.8, 80.1, 75.4,73.8, 62.4, 31.9, 29.7, 25.1, 22.8

5'-O-(N-icosanoylsulfamoyl) adenosine (AAMS) [9]
ESI-MS m/z (pos) 641.8 $[M+H]^+$ for a calculated mass of 641.37
^{1}H-NMR (300 MHz, DMSO-d6) δ 8.30 (s, 1H), 8.13 (s, 1H), δ 7.25 (s, 2H), δ 5.90 (d, 1H J=6.3), δ 5.54 (d, 1H, J= 6.9), δ 5.31(d, 1H, J=6.0), δ 4.59 (m, 1H), δ 4.14-3.97 (m, 10H), δ 3.16 – 3.17 (m, 14H), δ 1.97 (t, 2H), δ 1.43 (s, 3H), δ 1.21 (m, 42 H), δ 0.82 (t, 4H)
^{13}C-NMR (75 MHz, DMSO-d6) 179.04, 158.14, 157.73, 156.46, 153.06, 150.08, 139.88, 119.66, 119.35, 115.69, 87.38, 74.03, 71.30, 67.72, 31.74, 29.51, 29.47, 26.33, 22.53, 14.37

Reference

1. Ferreras, J. A., Ryu, J. S., Di Lello, F., Tan, D. S. & Quadri, L. E. Small-molecule inhibition of siderophore biosynthesis in Mycobacterium tuberculosis and Yersinia pestis. *Nat Chem Biol* **1**, 29-32 (2005).