PNAS
www.pnas.org

Supplementary Information for

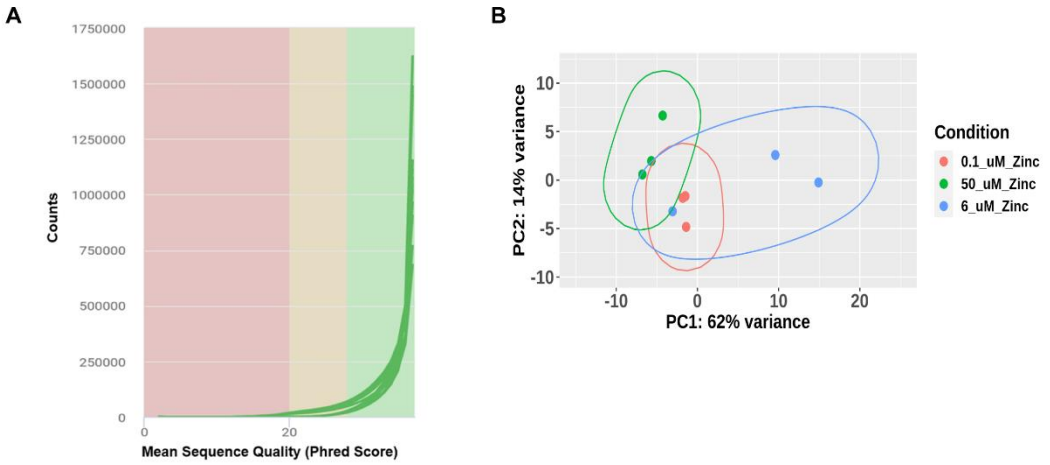
Kupyaphores are zinc homeostatic metallophores required for colonization of *Mycobacterium tuberculosis*

Kritee Mehdiratta[1, 2], Shubham Singh[3], Sachin Sharma[4], Rashmi S. Bhosale[4], Rahul Choudhury[5], Dattatraya P. Masal[5], Alzu Manocha[4], Bhushan Dilip Dhamale[4], Naseem Khan[6], Vivekanand A. [1, 2], Pooja Sharma[1, 2], Melanie Ikeh[7%], Amanda C. Brown[7$], Tanya Parish[7^], Anil Ojha[8, 9], Joy Sarojini Michael[10], Mohammed Faruq[1, 2], Guruprasad R. Medigeshi[6], Debasisa Mohanty[4], D. Srinivasa Reddy[5], Vivek T. Natarajan[1, 2], Siddhesh S. Kamat[3]*, Rajesh S. Gokhale[4]*

Corresponding authors*: Rajesh S. Gokhale and Siddhesh S. Kamat
Email: rsg@nii.ac.in; siddhesh@iiserpune.ac.in

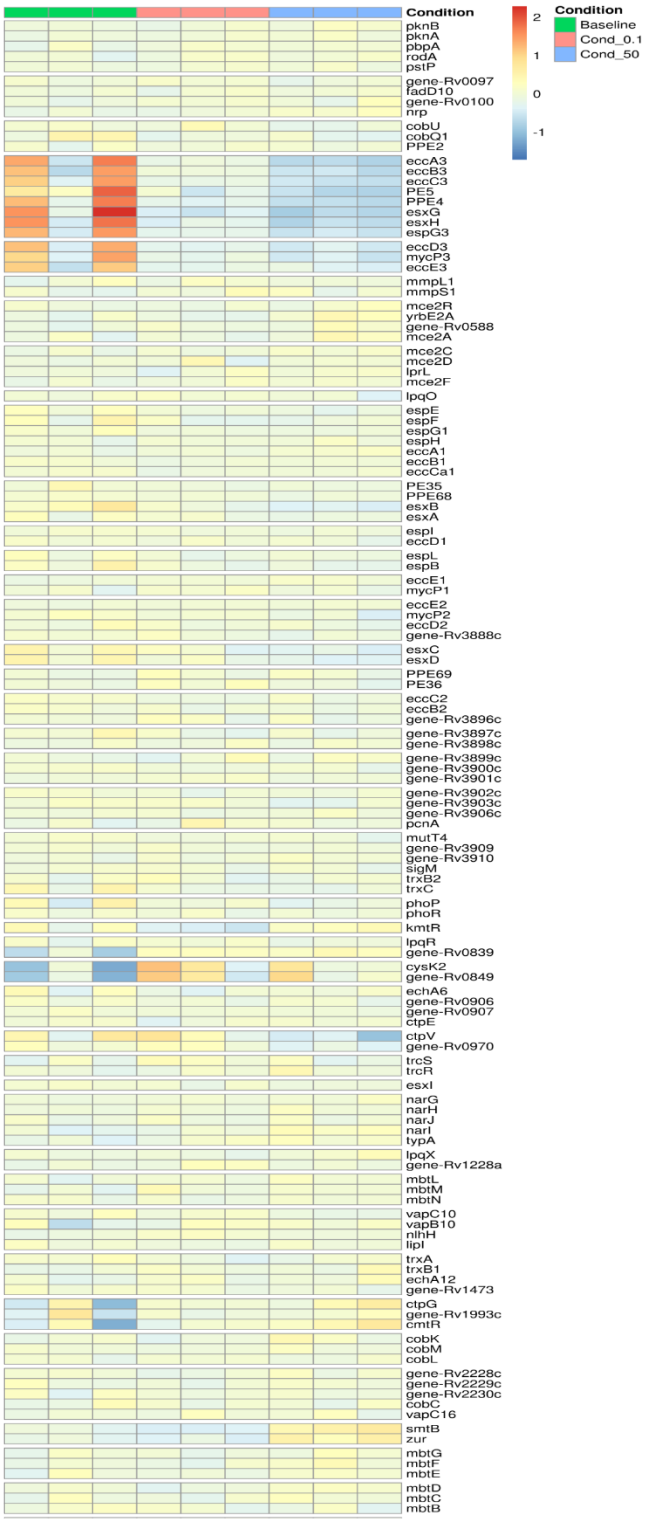
This PDF file includes:

25


A
1750000
1500000
1250000
1000000
750000
500000
250000
0
Counts
0
20
Mean Sequence Quality (Phred Score)
B
10
5
0
-5
-10
PC2: 14% variance
-10
0
10
20
PC1: 62% variance
Condition
0.1_uM_Zinc
50_uM_Zinc
6_uM_Zinc

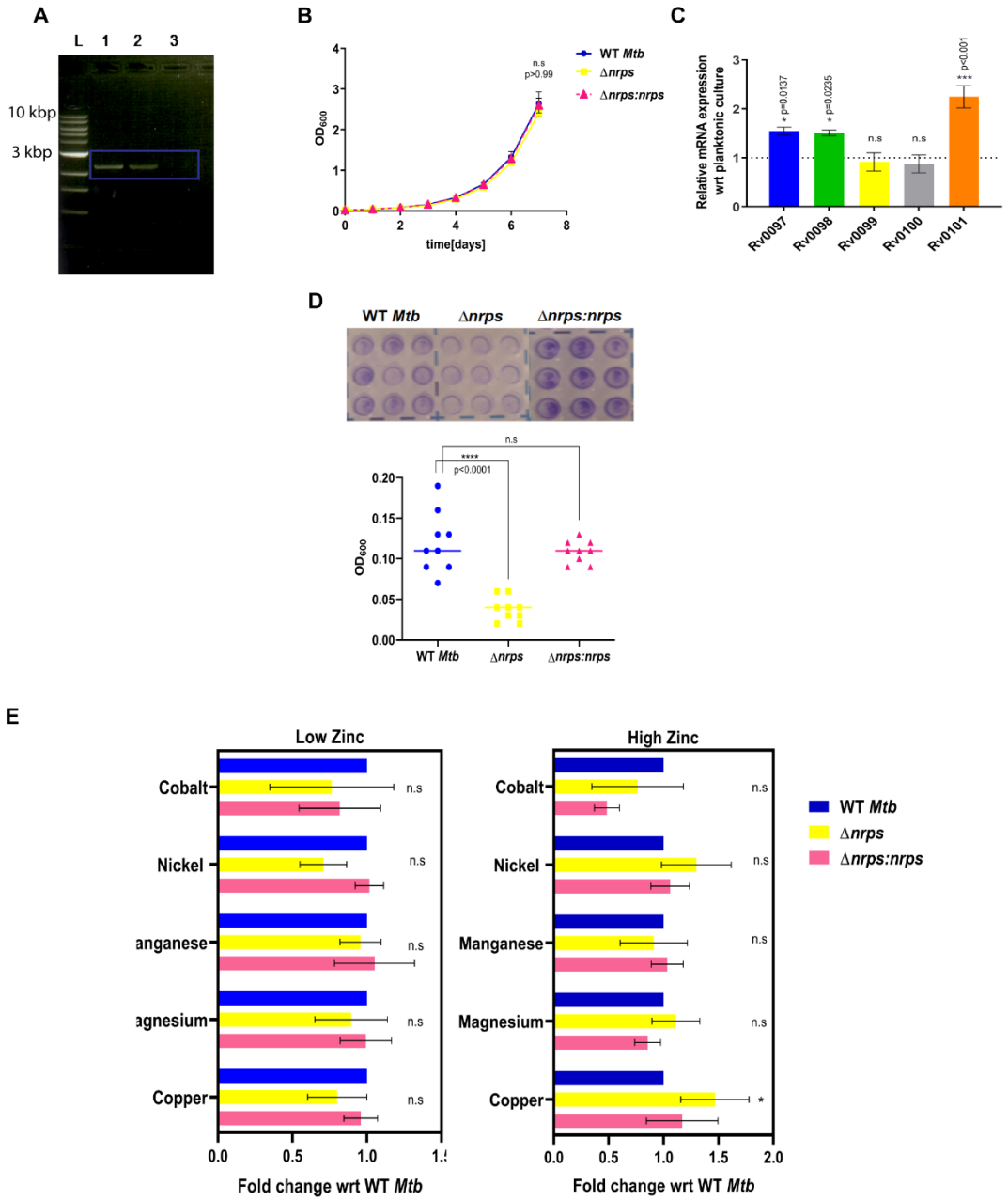
26

C



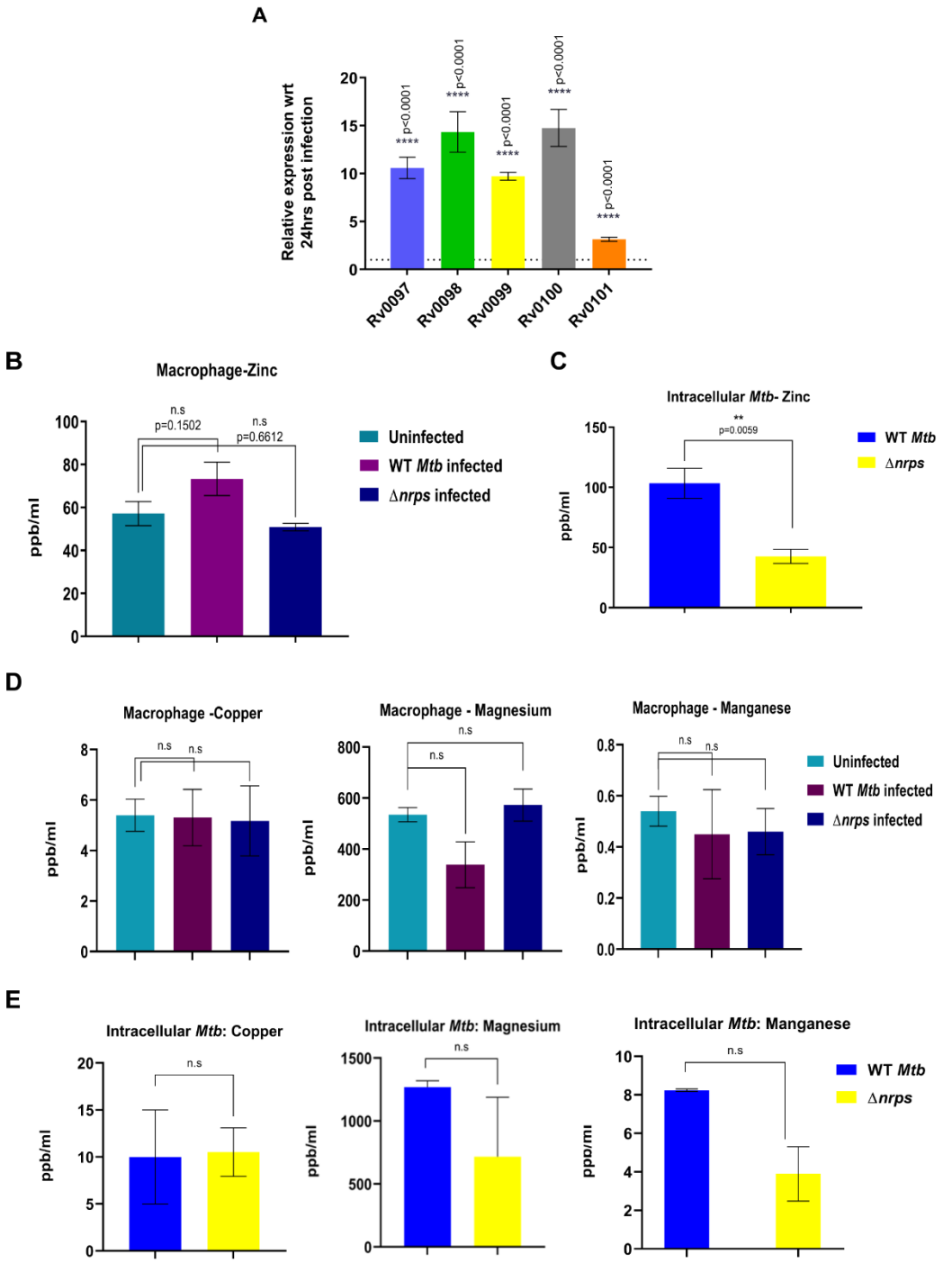
Supplementary Figure 1:
(A) Plot represents per sequence quality scores to depict if a subset of reads has poor quality. Y axis represents count and mean sequence quality, calculated as Phred score is plotted on x-axis. **(B)** Principal Component Analysis (PCA) plot of gene expression data where samples with similar gene expression profiles cluster together. Sample groups are indicated by using different colors as

detailed in the legend provided, here each point represents a biological replicate. (**C**) Heat map for differential expression of few metal responsive and FAAL-PKS/NRPS *Mtb* operons in three different zinc conditions.



Supplementary Figure 2: (**A**) PCR confirmation of *nrps* deletion and complementation strain generation using primers specific to region of Rv0101 gene. L: 1kb ladder, Lane 1- WT *Mtb* gDNA, Lane 2- Δ*nrps*:*nrps* gDNA and Lane 3- Δ*nrps* gDNA. (**B**) Growth kinetics of WT *Mtb*, Δ*nrps* and Δ*nrps*:*nrps* strains in complete 7H9 media show no growth profile differences for either strain. (**C**) Gene expression analysis of Rv0097-Rv0101 operon in WT *Mtb* planktonic and biofilm cultures (**D**) Crystal violet staining for biomass estimation of biofilm cultures of WT *Mtb*, Δ*nrps* and Δ*nrps*:*nrps* strains with quantitation of same by spectrophotometric measurement of absorbance. (**E**) ICP-MS

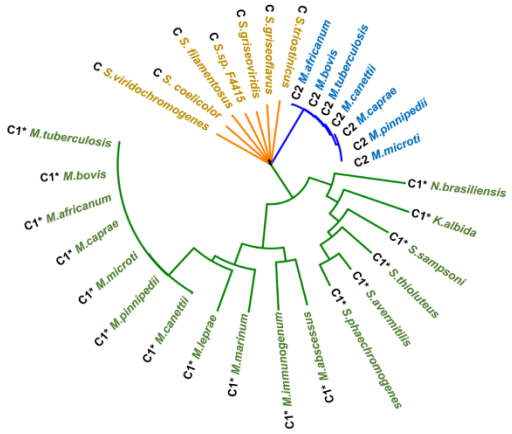
analysis of WT, Δ*nrps* and Δ*nrps:nrps Mtb* strains show no significant difference in intracellular levels of copper, magnesium manganese, cobalt and nickel in cultures grown under low zinc-Sauton's medium condition. Similarly, no significant difference could be observed in intracellular levels of magnesium, manganese, cobalt and nickel in cultures grown under high zinc-Sauton's medium condition. Some differences though could be observed in total levels of copper as compared to WT *Mtb*. Each experiment was done with 3 biological replicates. Data represents mean ± s.e.m with p values indicated at each data point.



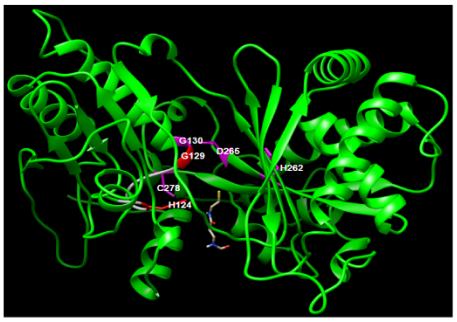
Supplementary Figure 3: (**A**) Relative gene expression of Rv0097-Rv0101 cluster in WT *Mtb* isolated from infected RAW 264.7 murine macrophages at 2 hours with respect to 24 hours post infection. (**B**) ICP-MS analysis showed no significant changes in total zinc pool of macrophages upon WT or Δ*nrps Mtb* infection as compared to uninfected macrophages. (**C**) ICP-MS analysis of *Mtb* harvested from macrophages post 2 hours of infection show significantly decreased intracellular levels of zinc in Δ*nrps Mtb* strain as compared to WT *Mtb*. (**D**) ICP-MS analysis of macrophage lysate post infection with WT *Mtb* and Δ*nrps* show no differences in intracellular levels

of copper, magnesium and manganese post 2 hours infection as compared to WT *Mtb*. (**E**) ICP-MS analysis for measurement of levels of copper, magnesium and manganese in intracellular *Mtb* harvested post infection. No differences could be observed in the total levels of copper, magnesium and manganese. Each experiment was done with 3 biological replicates. Data represents mean ± s.e.m, with p value indicated at each data point.

A



B

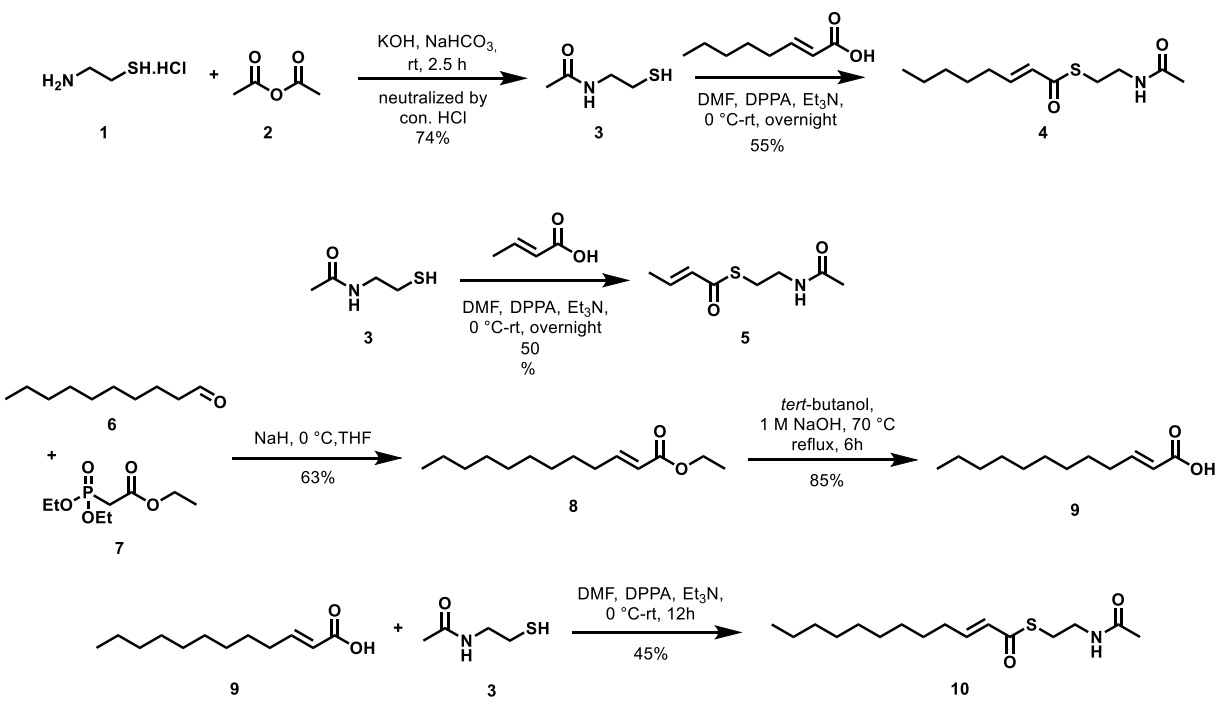


C

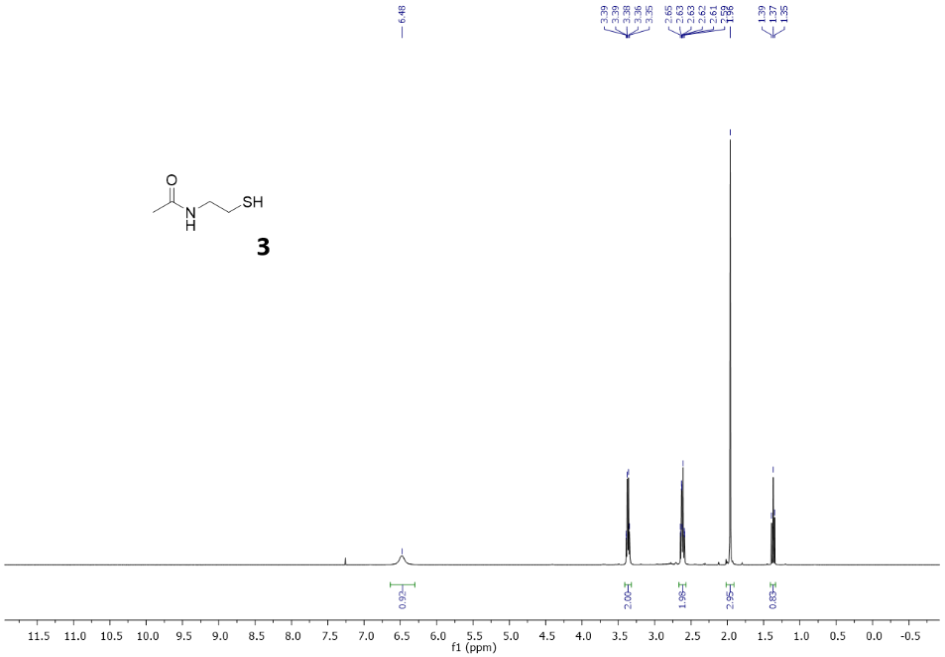
	9	13	25	26	28	124	127	128	129	130	261	262	263	264	265	266	278	279	280	281	282	283	313	319	320	321	333	337	338	339	365
C1*_Mtb_F28	S	L	L	Y	I	H	L	D	G	G	V	H	T	V	D	N	C	L	V	N	S	-	R	H	Y	R	V	T	L	N	V
C1*_Mtb_H37Rv	S	L	L	Y	I	H	L	D	G	G	V	H	T	V	D	N	C	L	V	N	S	-	R	H	Y	R	V	T	L	N	V
C1*_Mbovi_BCG	S	L	L	Y	I	H	L	D	G	G	V	H	T	V	D	N	C	L	V	N	S	-	R	H	Y	R	V	T	L	N	V
C1*_Mafri	S	L	L	Y	I	H	L	D	G	G	V	H	T	V	D	N	C	L	V	N	S	-	R	H	Y	R	V	T	L	N	V
C1*_Mcane	S	L	L	Y	I	H	L	D	G	G	V	H	T	V	D	N	C	L	V	N	S	-	R	H	Y	R	V	T	L	N	V
C18_Mabsc	S	M	L	Y	I	H	L	D	G	G	A	R	A	A	P	L	C	D	G	H	R	-	R	-	-	-	-	-	-	-	-
C1*_Mimmu	S	M	L	Y	I	H	F	D	G	G	I	H	A	V	D	N	C	L	V	N	S	-	-	-	-	-	-	-	-	-	-
C1*_Nbras	A	V	A	F	I	H	L	D	G	Y	A	L	P	V	M	L	M	T	L	N	G	V	-	-	-	-	-	-	-	-	-
C1*_Kalbi	-	-	-	-	-	H	L	D	G	Y	G	M	P	V	M	G	T	V	L	N	A	V	-	-	-	-	-	-	-	-	-
C1*_Ssamp	A	I	A	Y	T	H	L	D	G	F	S	V	P	A	M	G	M	V	R	N	V	L	-	-	-	-	-	-	-	-	-
C1*_Saver	A	I	A	Y	T	H	L	D	G	F	S	L	P	A	M	G	M	V	R	N	I	L	-	-	-	-	-	-	-	-	-
C2_Mycotuber	S	S	V	N	I	H	G	D	H	W	G	A	P	V	A	G	F	F	I	N	I	V	S	L	F	D	-	V	V	H	G
C2_Mtb_H37Rv	A	L	V	Y	M	H	A	D	G	W	G	F	P	I	A	G	F	F	V	N	T	L	-	-	-	-	-	-	-	-	-
C2_Mcane	A	L	V	Y	M	H	A	D	G	W	G	F	P	I	A	G	F	F	V	N	T	L	-	-	-	-	-	-	-	-	-
C2_Mbovi	A	L	V	Y	M	H	A	D	G	W	G	F	P	I	A	G	F	F	V	N	T	L	-	-	-	-	-	-	-	-	-
C2_Mafri	A	L	V	Y	M	H	A	D	G	W	G	F	P	I	A	G	F	F	V	N	T	L	-	-	-	-	-	-	-	-	-
C_Bsubt	A	L	T	Y	I	H	G	D	G	W	G	S	P	I	A	G	L	F	I	N	T	L	S	L	F	Q	-	M	L	A	L
C_Bbrev	A	L	T	Y	I	H	G	D	G	W	G	S	P	I	A	G	L	F	I	N	T	L	S	L	F	Q	-	M	L	A	L
C_Paeru_m_1	A	L	A	Y	I	H	A	D	G	W	G	V	P	S	A	N	F	F	I	N	T	L	-	L	F	Q	-	L	F	N	L
C_Paeru	A	M	A	Y	V	H	T	E	G	W	G	A	P	V	A	N	A	F	L	N	T	Q	S	L	F	Q	-	M	C	N	L
C_Paeru_ser	A	M	A	Y	L	H	S	D	G	W	G	V	P	I	A	N	L	F	V	N	T	Q	S	L	F	Q	-	M	Y	N	V
C_Paeru_dhb	A	Q	A	Y	I	H	S	D	G	W	G	V	P	V	A	N	F	F	V	N	T	Q	S	L	F	Q	-	L	F	N	V
C_Paeru_pvdD	A	Q	A	Y	I	H	S	D	G	W	G	V	P	I	A	N	F	F	V	N	T	Q	N	L	F	Q	-	L	F	N	L
C_Paeru_pvdD	A	Q	A	Y	I	H	S	D	G	W	G	V	P	I	A	N	F	F	V	N	T	Q	S	L	F	Q	-	L	F	N	F

Supplementary Figure 4: (**A**) Dendrogram-based clustering analysis of condensation domains that are present in the NRPS protein of similar 5 gene biosynthetic loci from other *Actinomycetes* along with classical condensation domain of *NRPS* genes. The starter condensation domain from these 5-genes NRPS are distinctively different (C1*) from other classical condensation domains (C). The second condensation domain (C2) present in bimodular 5-genes NRPS protein clusters with classical C domains. (**B)** Modelling based structure prediction of unique starter condensation domain (C1*) with residues possibly contributing to isonitrile stability colour coded. Magenta–Residues within 10 Å of –SH of 4'-phosphopantethiene, Red – H124, G129 (catalytic residues), White – HHxxDG (conserved motif). Distance between them is 7.6 Å. (**C**) Sequence analysis of C domains in *kupya* operon and classical NRPS of *Actinomycetes* reveal 31 residues having at least one atom within 10 Å radius of last atom of 4'-phosphopantethiene sidechains to be conserved uniquely in C_1*.

A

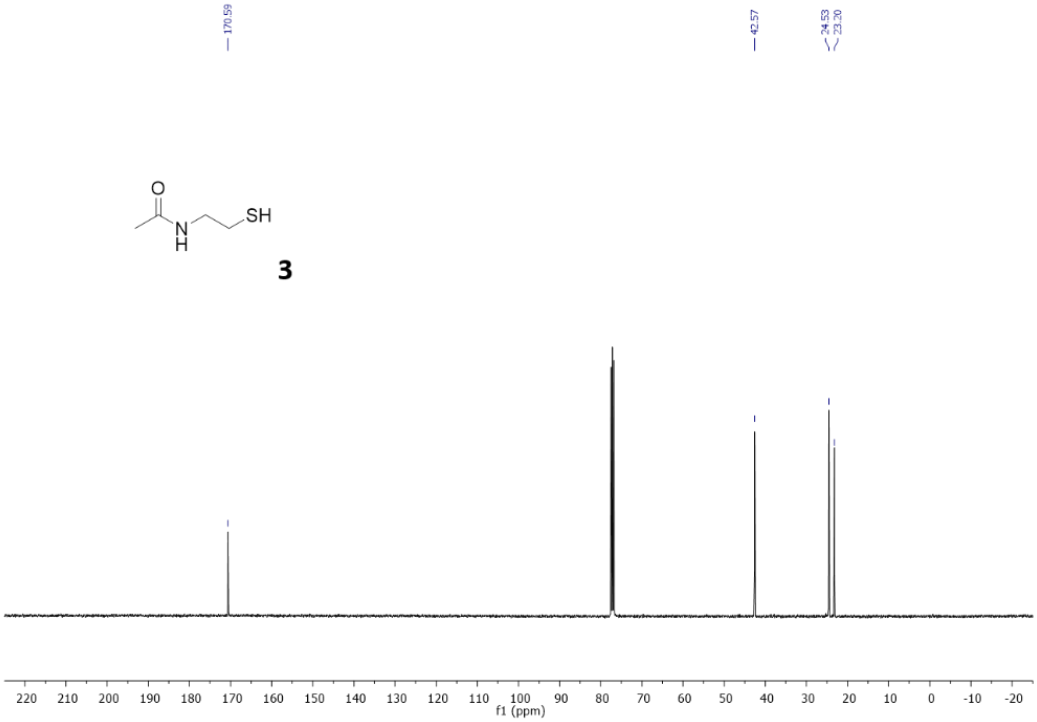

KOH, NaHCO3,
rt, 2.5 h
neutralized by
con. HCl
74%
DMF, DPPA, Et3N,
0 °C-rt, overnight
55%
DMF, DPPA, Et3N,
0 °C-rt, overnight
50
%
NaH, 0 °C,THF
63%
tert-butanol,
1 M NaOH, 70 °C
reflux, 6h
85%
DMF, DPPA, Et3N,
0 °C-rt, 12h
45%

B



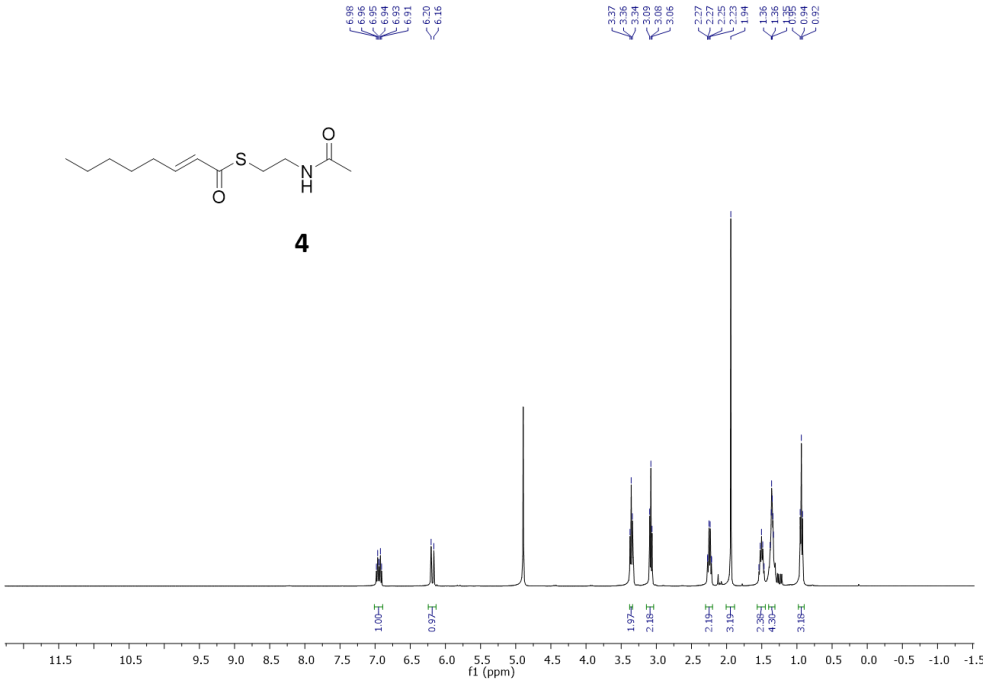
^{1}H NMR analysis of **03** in $CDCl_3$ at 400MHz

C



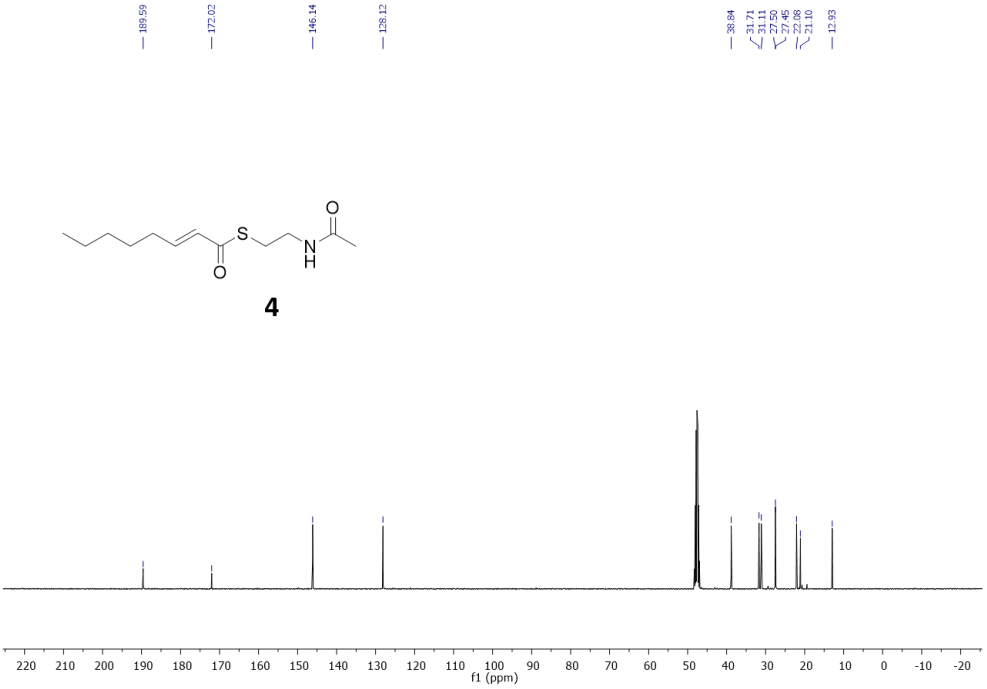
^{13}C NMR analysis of **03** in $CDCl_3$ at 100MHz

D



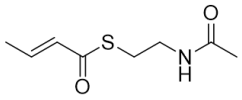
^{1}H NMR analysis of **04** in CD_3OD at 400MHz

E

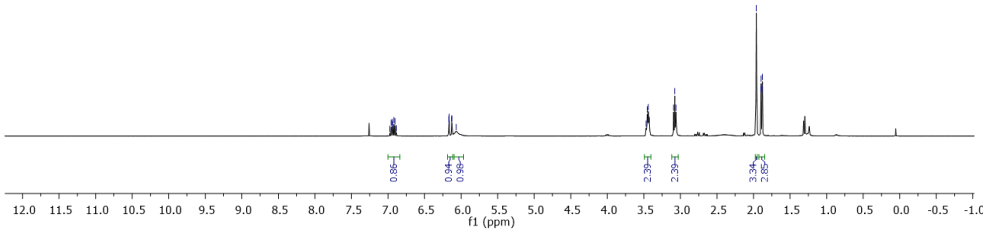


^{13}C NMR analysis of **04** in CD_3OD at 100MHz

F

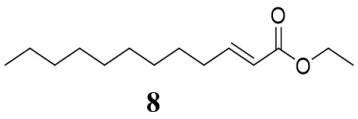


5

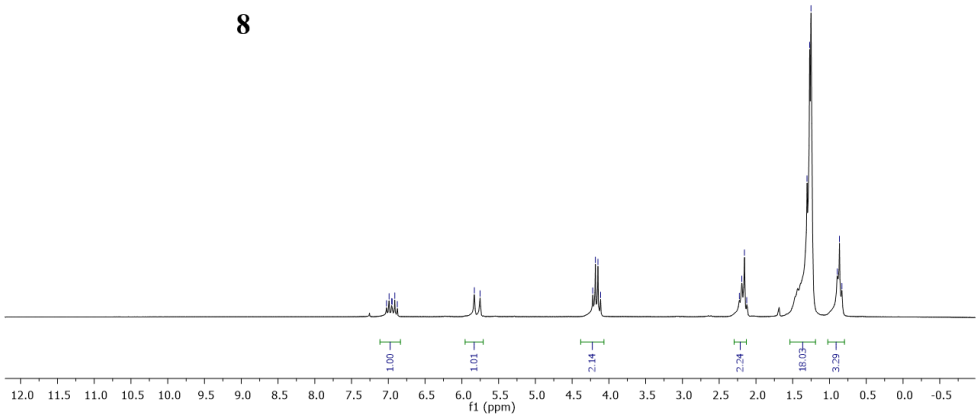


^{1}H NMR analysis of **05** in $CDCl_3$ at 400MHz

G

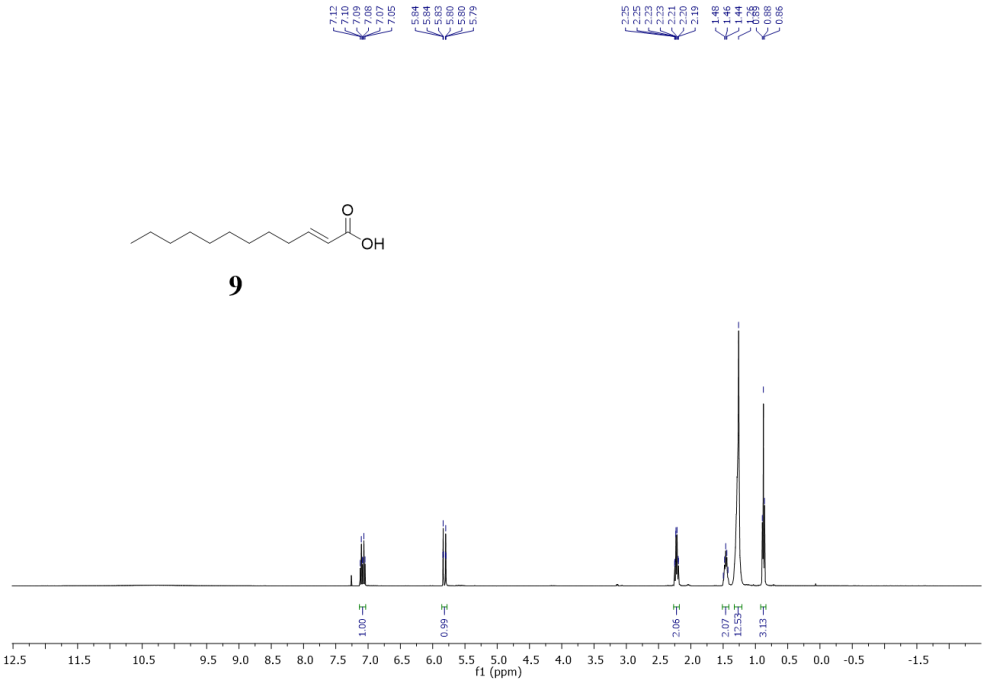


8



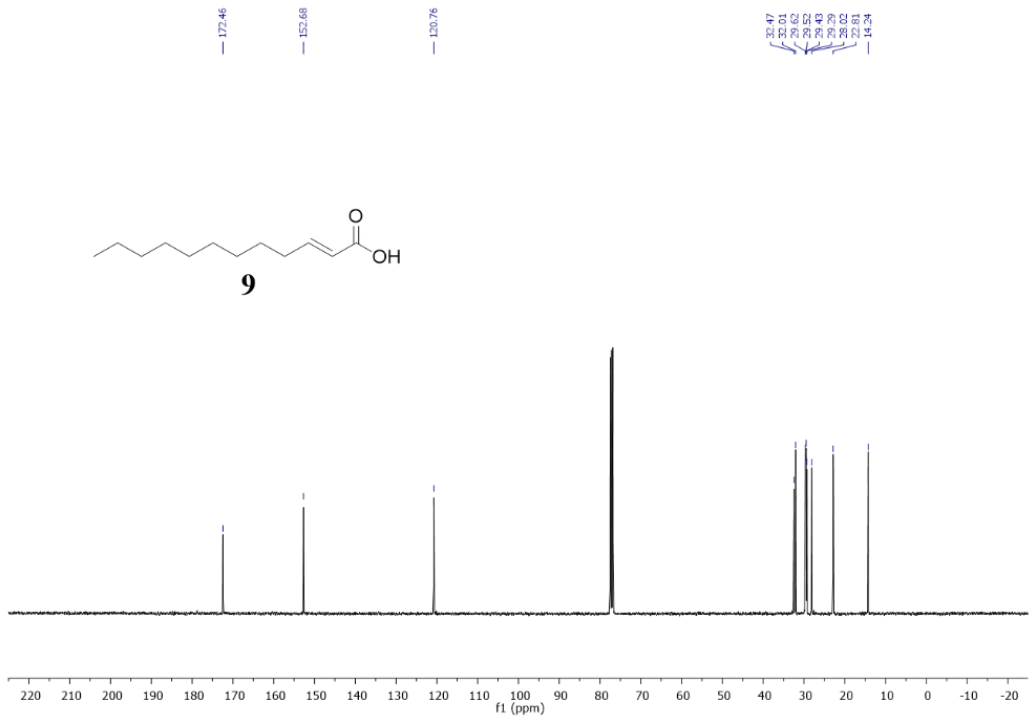
^{1}H NMR analysis of **08** in $CDCl_3$ at 200MHz

H



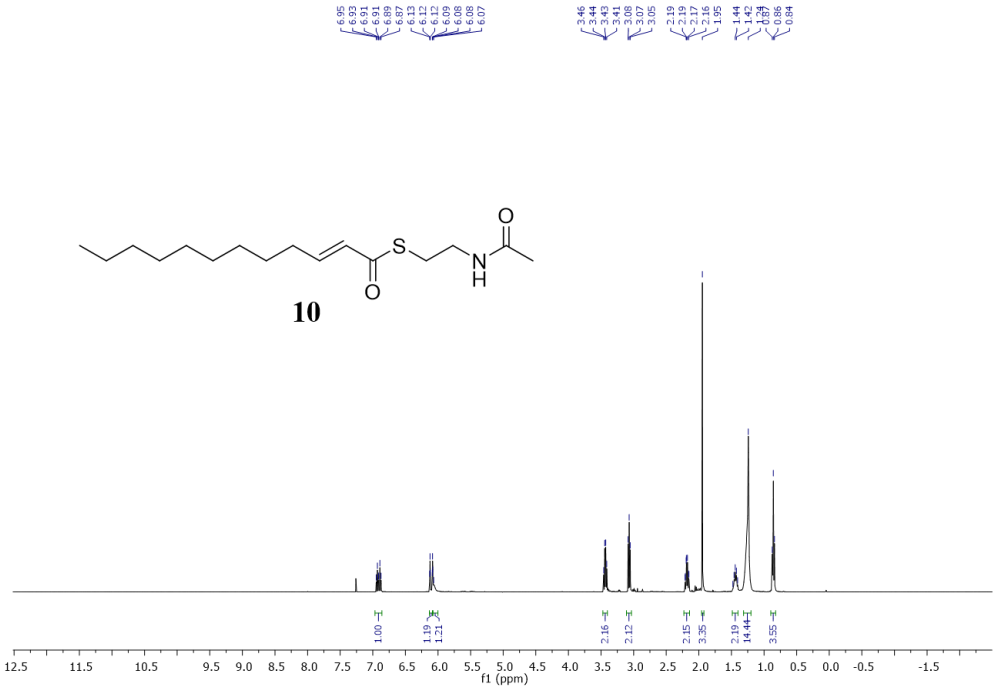
^{1}H NMR analysis of **09** in $CDCl_3$ at 400MHz

I



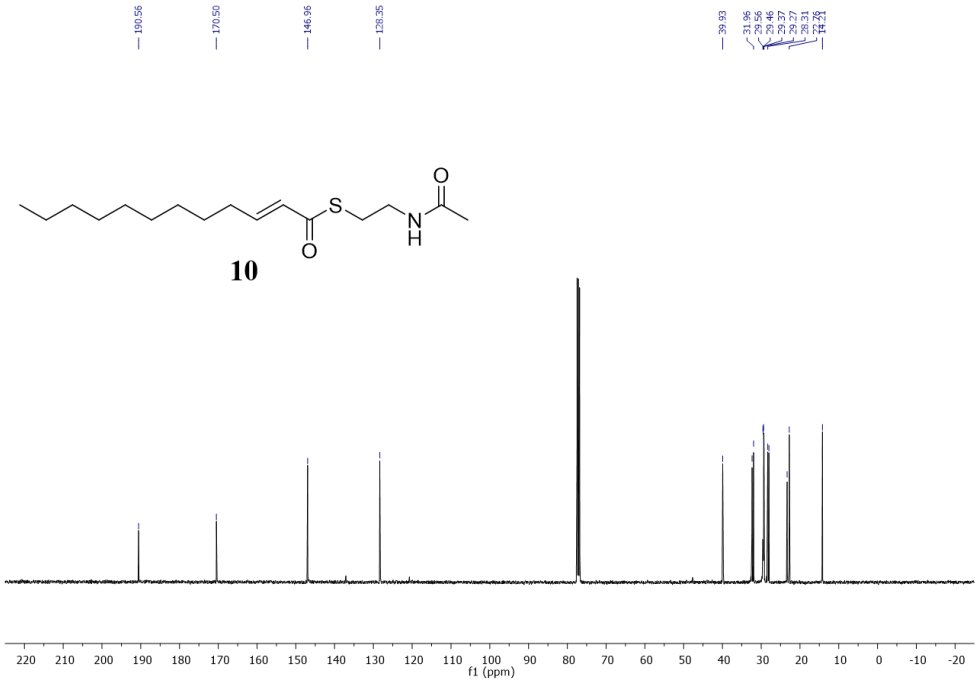
^{13}C NMR analysis of **09** in $CDCl_3$ at 100MHz

J



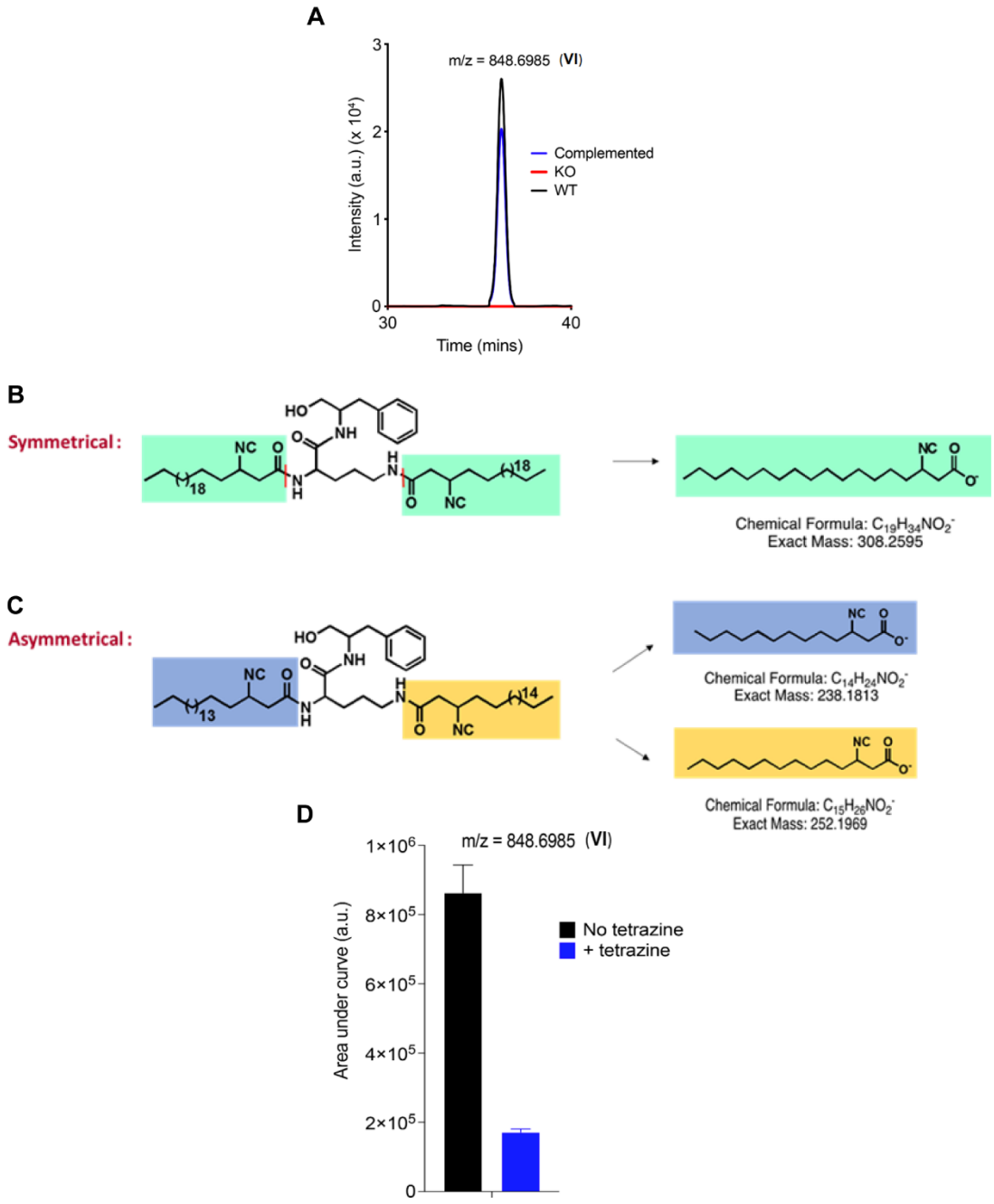
^{1}H NMR analysis of **10** in $CDCl_3$ at 400MHz

K



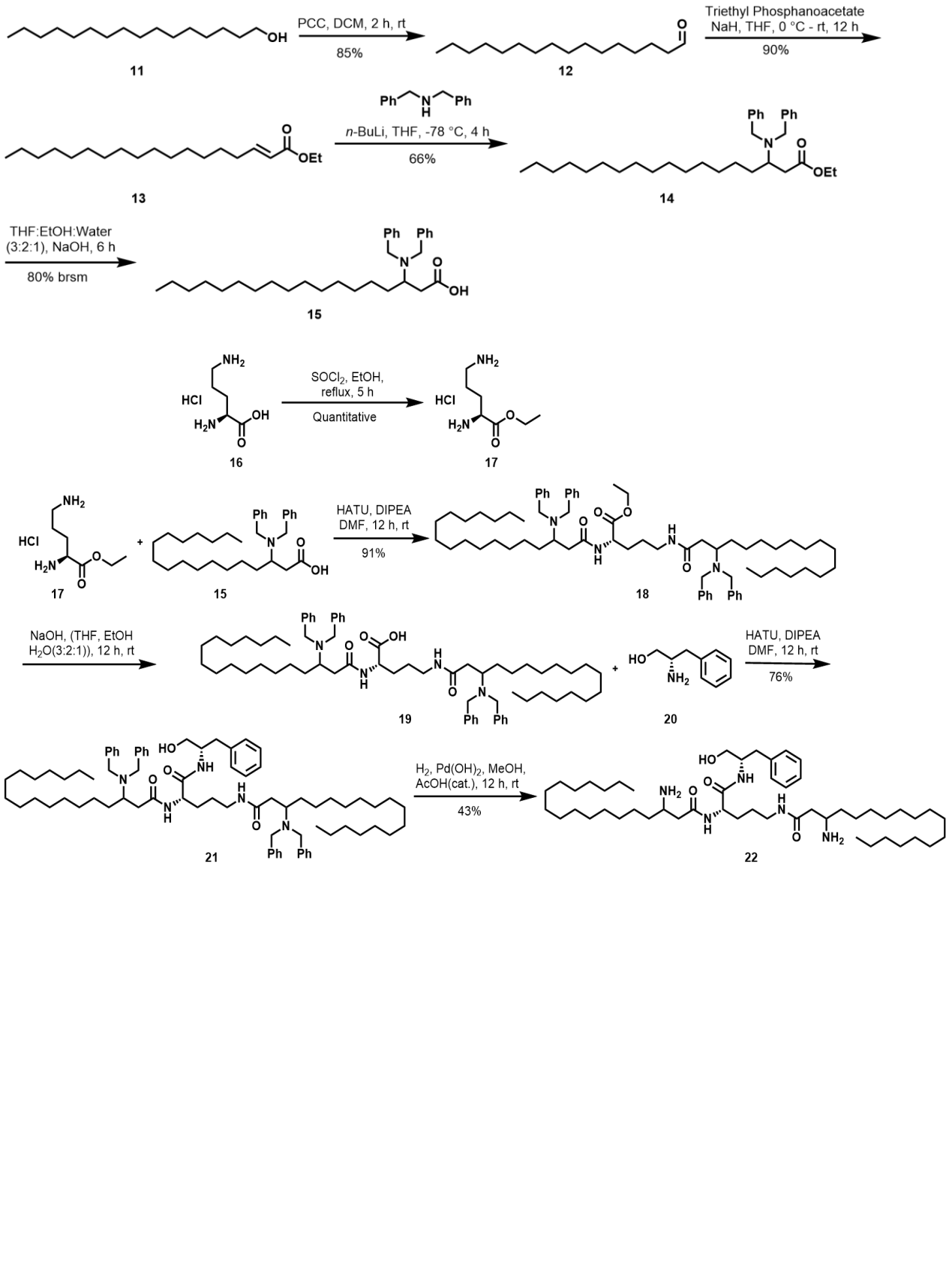
^{13}C NMR analysis of **10** in $CDCl_3$ at 100MHz

Supplementary Figure 5: (**A**) Reaction Scheme for synthesis of 2-dodecanoyl-S-N-acetylcysteamine **10** with intermediate metabolites labelled **1** to **10**. ^{1}H (**B**), (**D**), (**F**), (**G**), (**H**), (**J**) and ^{13}C NMR analysis (**C**), (**E**), (**I**), (**K**) of intermediates **3**, **4**, **5**, **8**, **9**, **10** and **3**, **8**, **9**, **10** respectively**.**

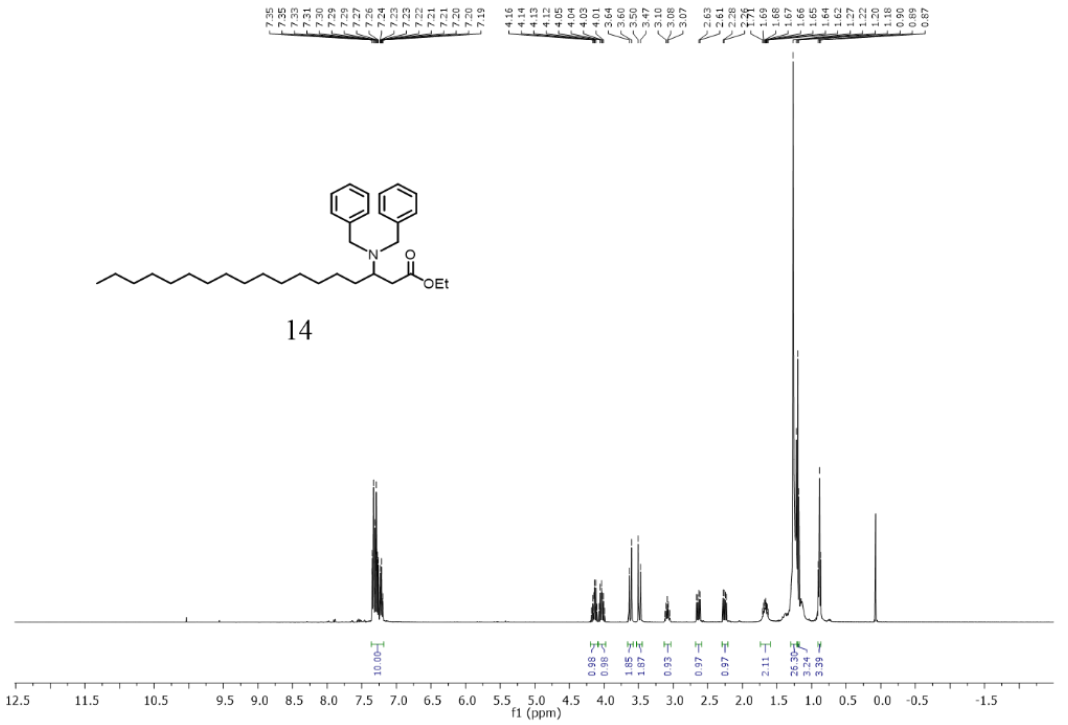


Supplementary Figure 6: (**A**) Comparative EIC of kupyaphore **VI** of m/z for $(M+H)^+$ = 848.6985 from biofilm extracts of WT, Δ*nrps*, Δ*nrps:nrps Mtb* strains. (**B**), (**C**) Negative ion mode HRMS analysis of parent lipopetide revealed 2 series of acyl chain substitutions – symmetrical and asymmetrical. Symmetrical kupyaphores possibly have identical acyl chains substituted at both the amine groups of ornithine. Asymmetrical kupyaphores though differ in their acyl substitutions either in terms of chain length or branched substitutions resulting in mass differences. (**D**) Relative abundance of kupyaphore **VI** with or without tetrazine addition to WT *Mtb* culture extract. Each experiment was performed in triplicates with mass error tolerance of 10 ppm.

A

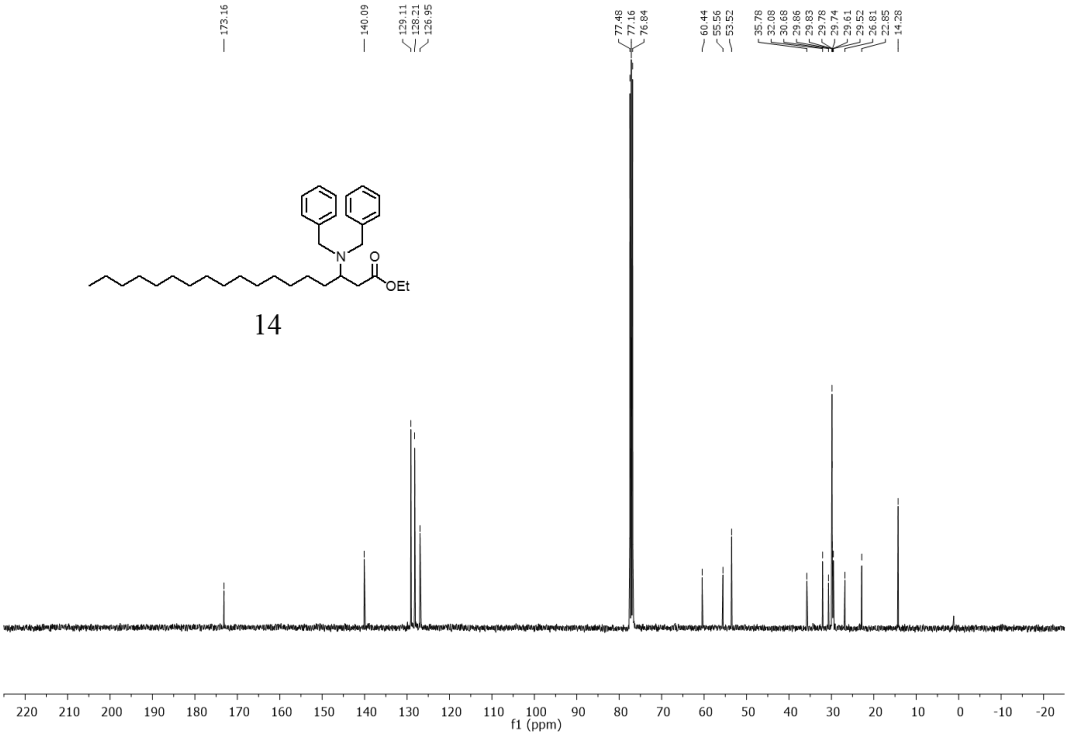

PCC, DCM, 2 h, rt
85%
Triethyl Phosphanoacetate
NaH, THF, 0 °C - rt, 12 h
90%
n-BuLi, THF, -78 °C, 4 h
66%
THF:EtOH:Water
(3:2:1), NaOH, 6 h
80% brsm
SOCl2, EtOH,
reflux, 5 h
Quantitative
HATU, DIPEA
DMF, 12 h, rt
91%
NaOH, (THF, EtOH
H2O(3:2:1)), 12 h, rt
HATU, DIPEA
DMF, 12 h, rt
76%
H2, Pd(OH)2, MeOH,
AcOH(cat.), 12 h, rt
43%
11
12
13
14
15
16
17
18
19
20
21
22

B



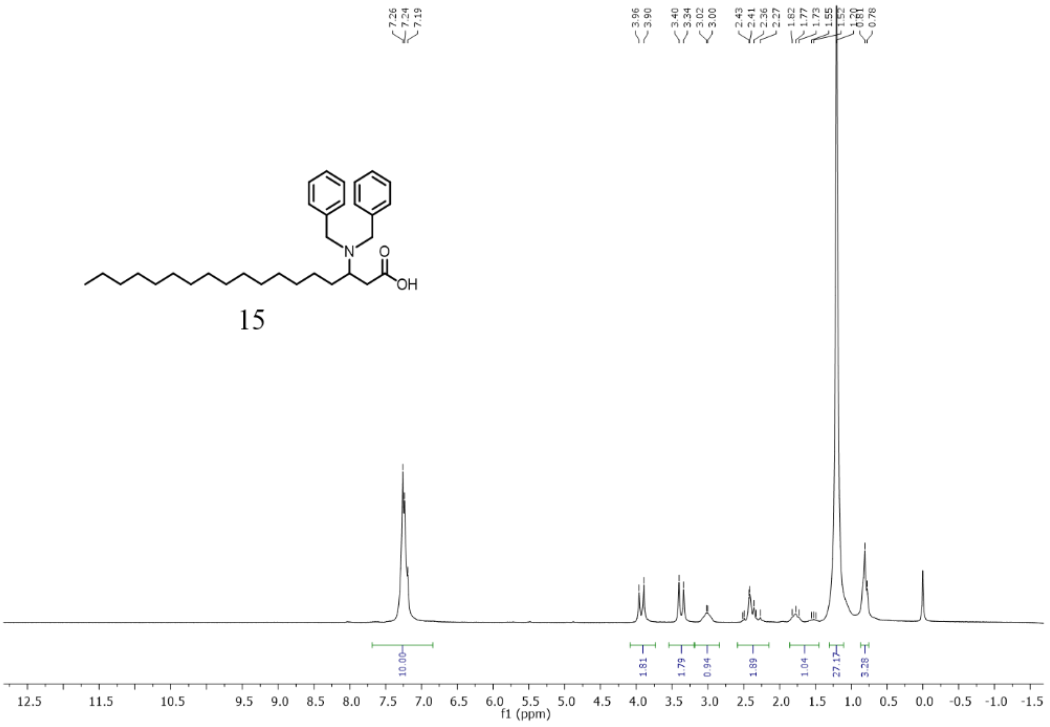
^{1}H NMR analysis of **14** in $CDCl_3$ at 400MHz

C



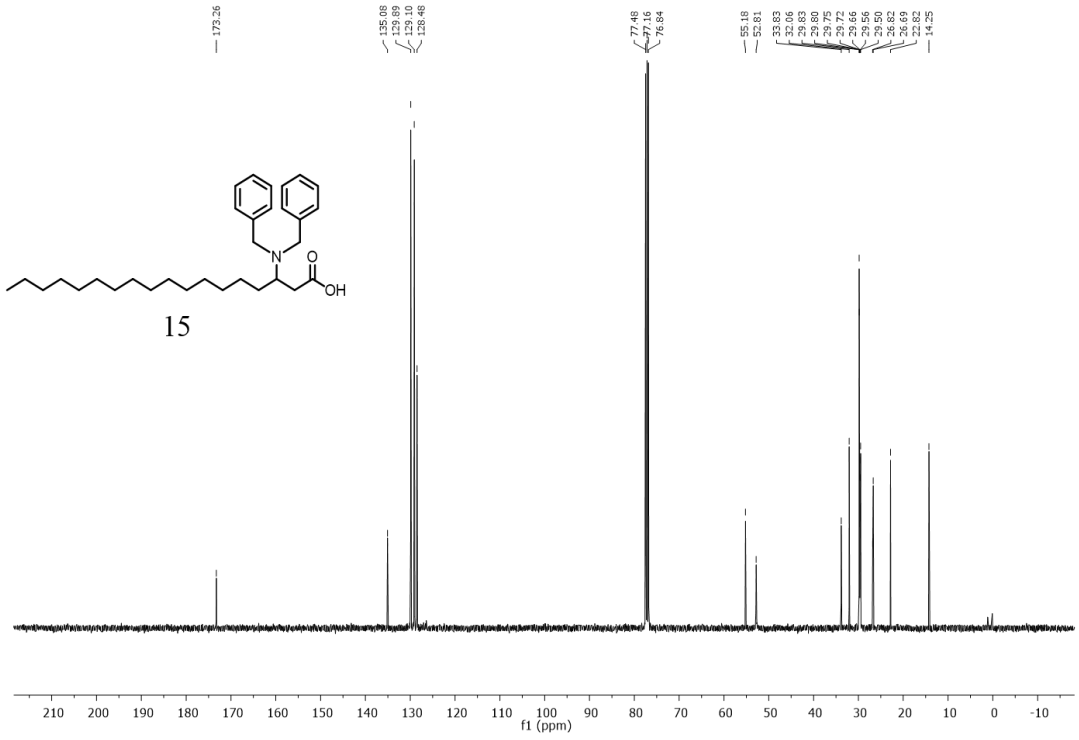
^{13}C NMR analysis of **14** in $CDCl_3$ at 100MHz

D



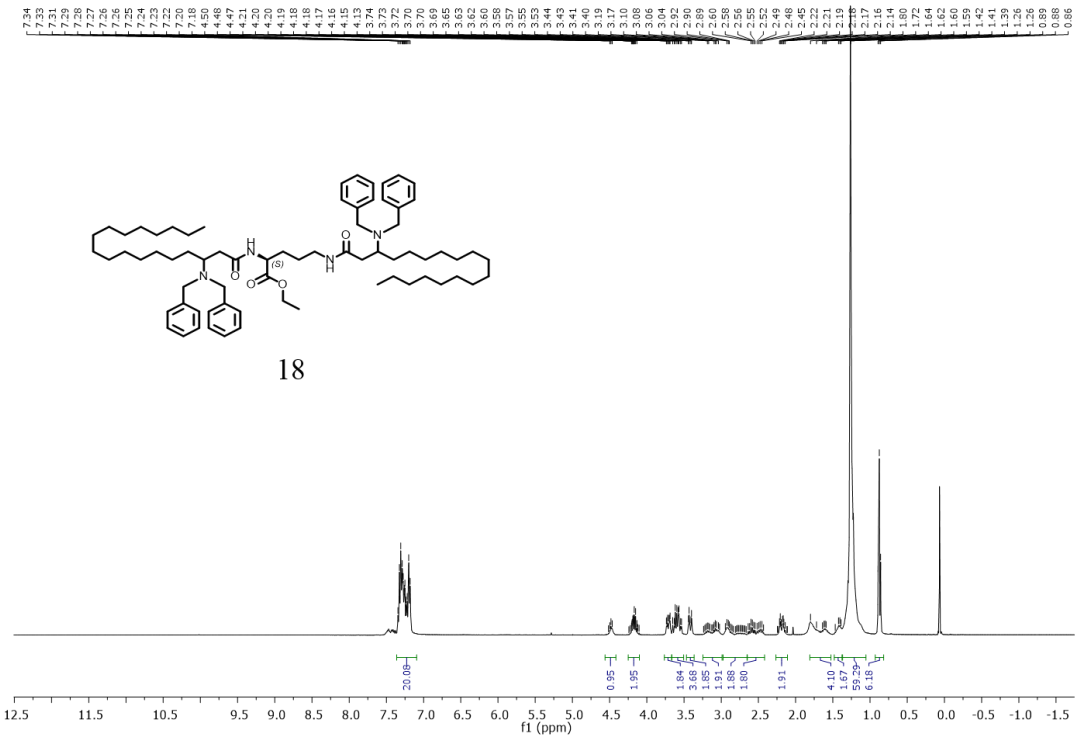
^{1}H NMR analysis of **15** in $CDCl_3$ at 200MHz

E



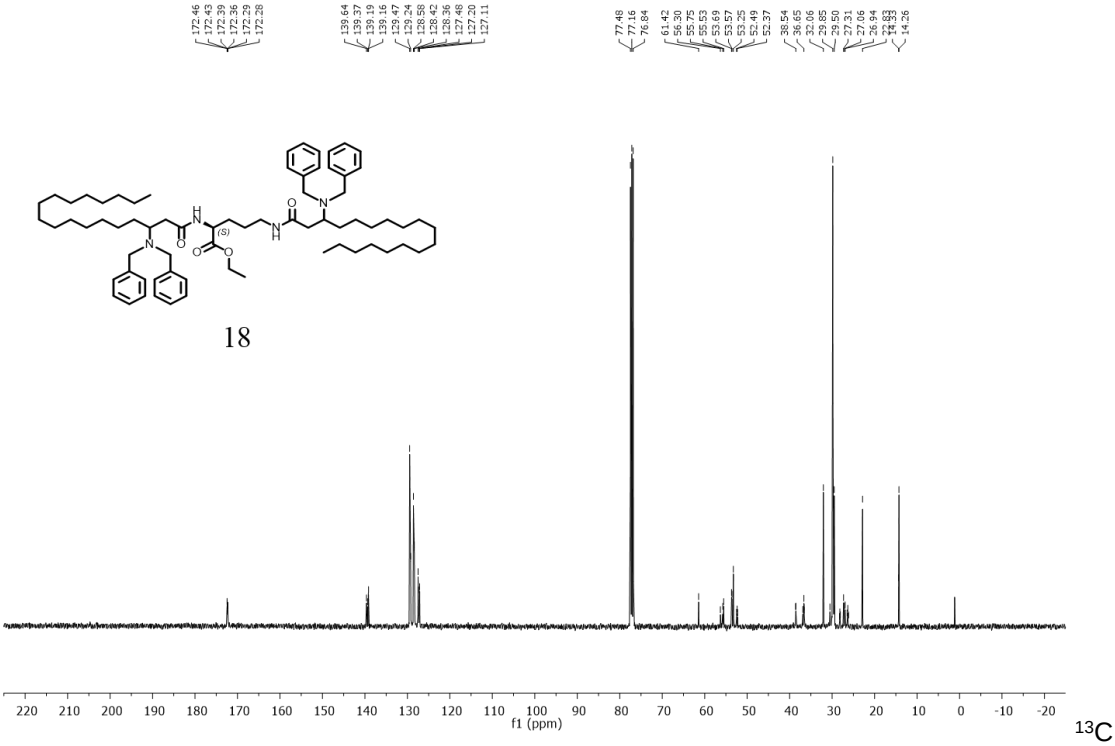
^{13}C NMR analysis of **15** in $CDCl_3$ at 100MHz

F



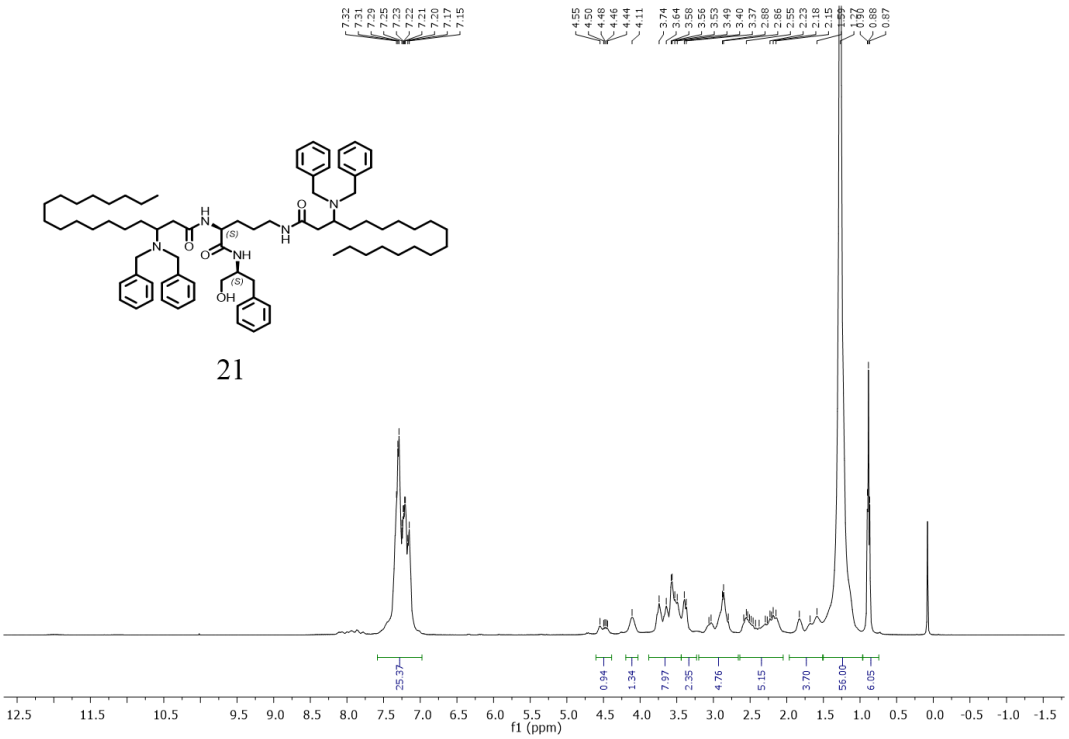
^{1}H NMR analysis of **18** in $CDCl_3$ at 400MHz

G



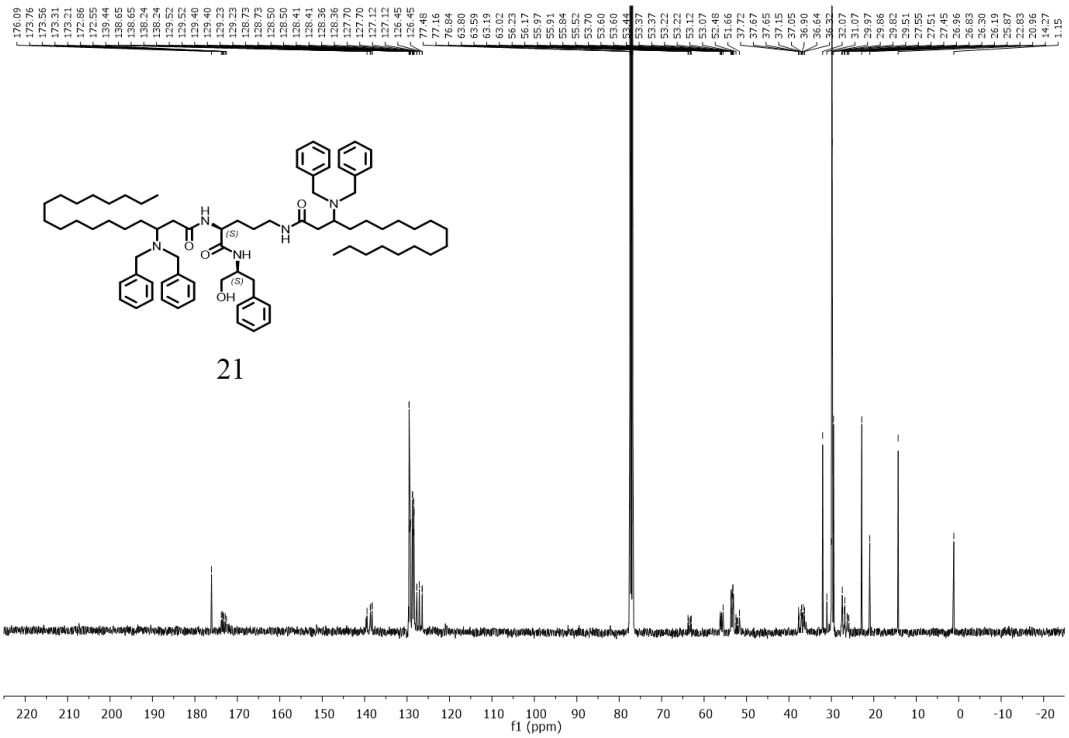
^{13}C NMR analysis of **18** in $CDCl_3$ at 100MHz

H



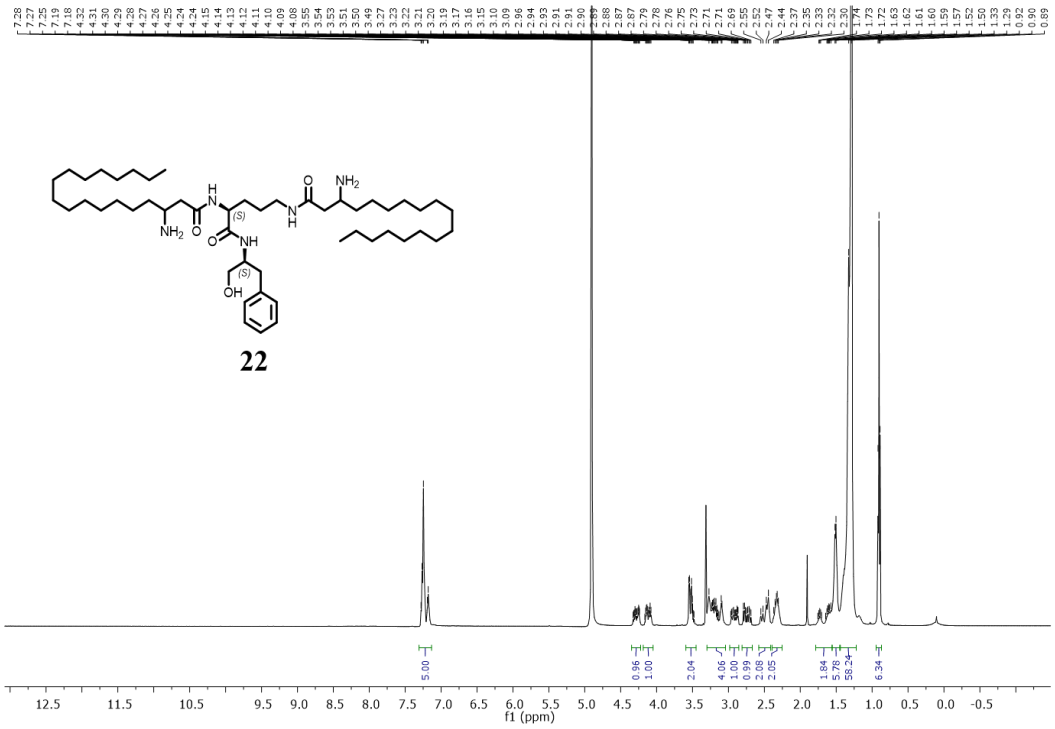
^{1}H NMR analysis of **21** in $CDCl_3$ at 400MHz

I



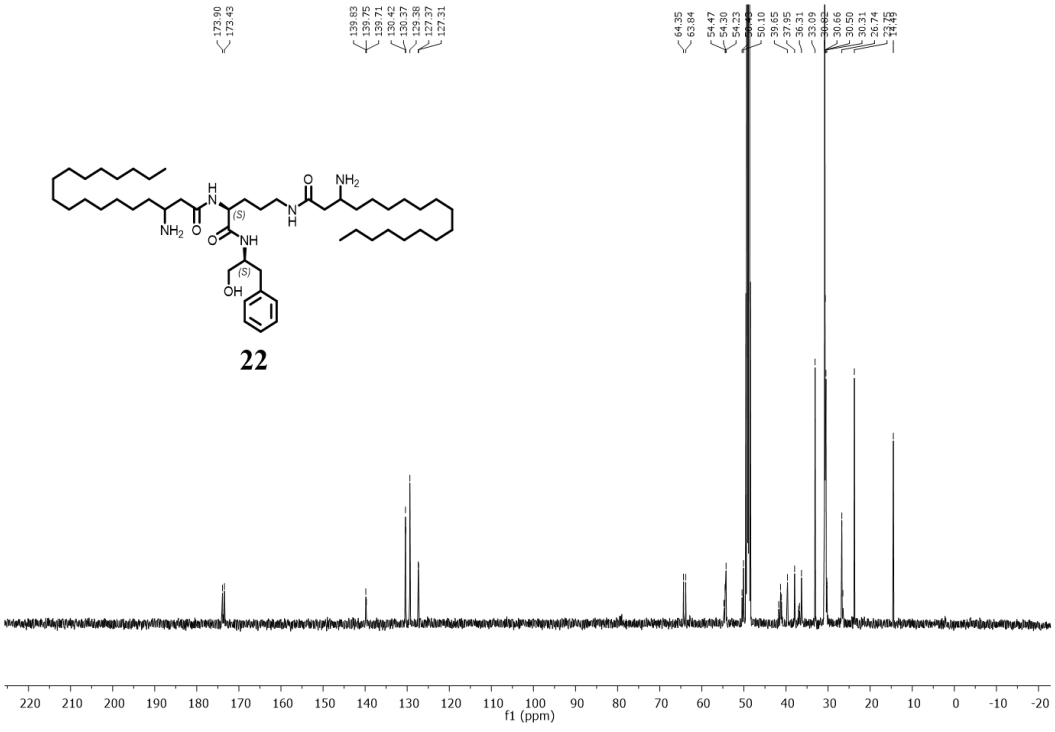
^{13}C NMR analysis of **21** in $CDCl_3$ at 100MHz

J



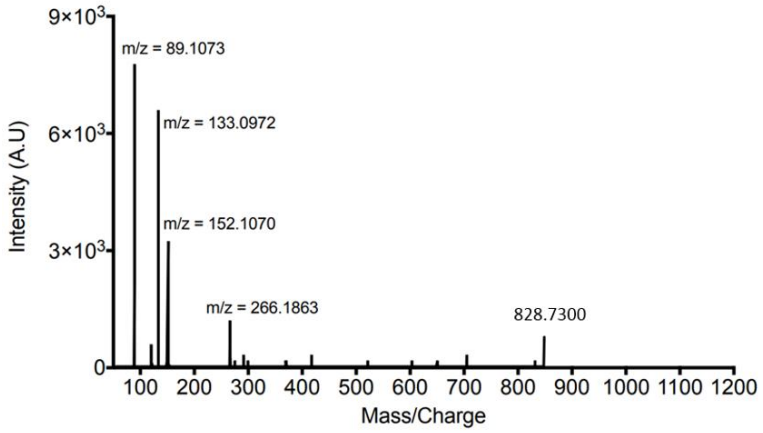
^{1}H NMR analysis of **22** in CD_3OD at 500MHz

K



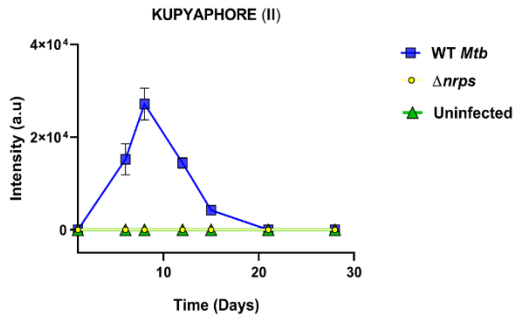
^{13}C NMR analysis of **22** in CD_3OD at 125MHz

L

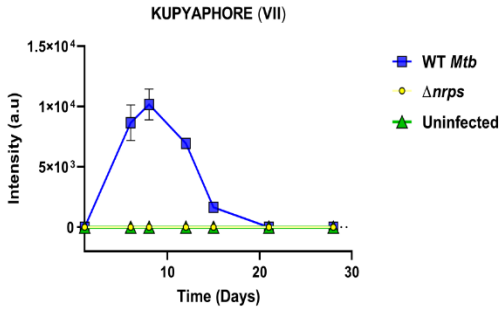


Supplementary Figure 7: (**A**) Reaction scheme and confirmation of synthesis of β-amine substituted isonitrile lipopeptide **22** with intermediates labelled from **11** to **22.** ^{1}H (**B**), (**D**), (**F**), (**H**), (**J**) and ^{13}C NMR analysis (**C**), (**E**), (**G**), (**I**), (**K**) of key intermediates **14**, **15**, **18**, **21**, **22**. (**L**) Positive ion mode MS/MS analysis of chemically synthesized standard **22** show identical spectra as *Mtb* endogenous kupyaphore lipopetides.

A



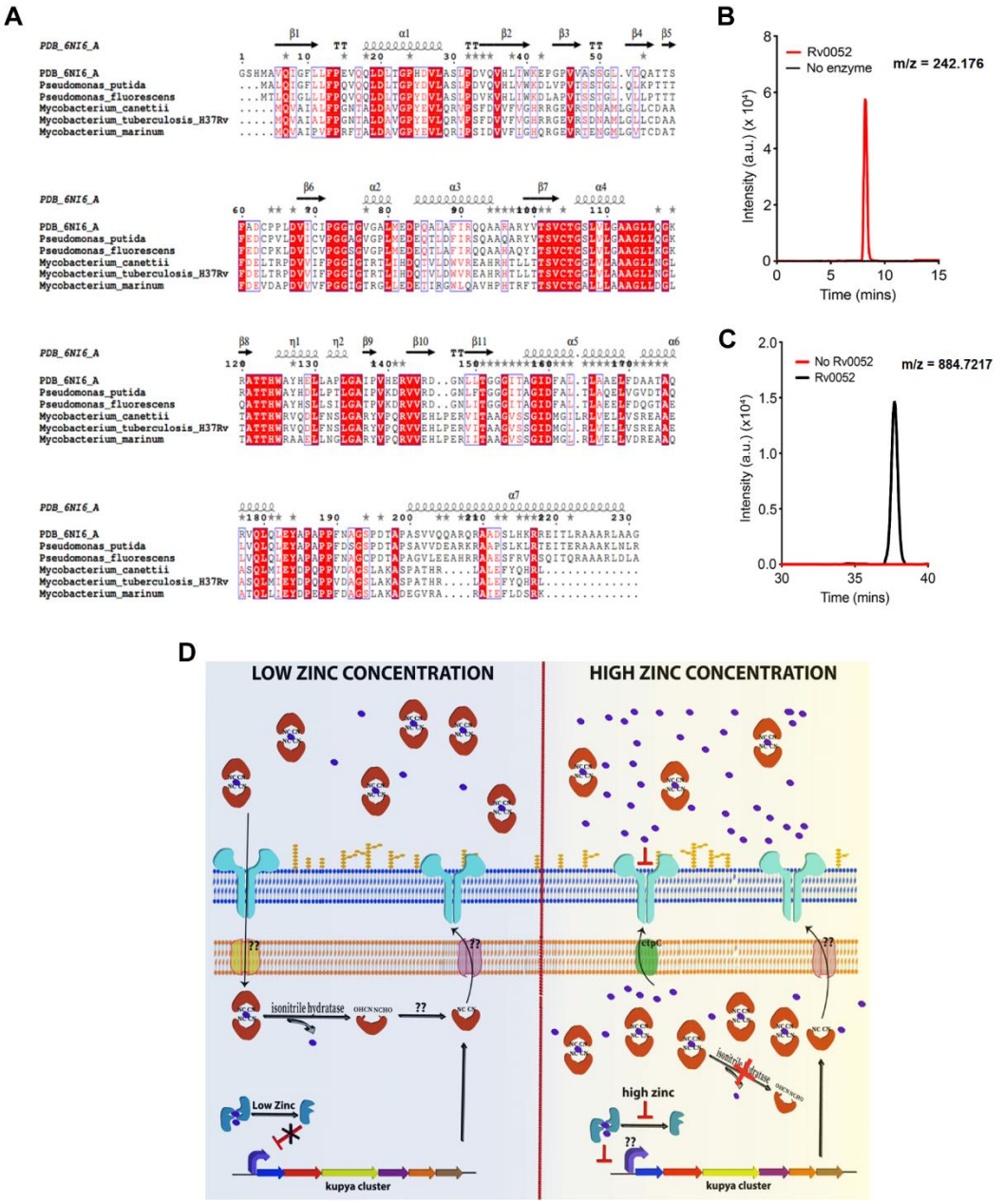
B



C

	TB_1	TB_2	TB_3	TB_4
I	1488288	1762445	10981	9912
II	1973505	1761535	101298	77563
III	1990992	238063	23145	17817
IV	ND	0	144159	167163
V	2527675	150775	88903	99752
VI	775995	504890	109814	144445
VII	ND	37925	126534	156123

Supplementary Figure 8: (**A**), (**B**) Plots show MRM-based semi-quantitative LC-MS measurements of kupyaphore species – **II** and **VII** of m/z for $[M+H]^+$ = 722.55 and 848.69 respectively, from lungs of uninfected, WT and Δ*nrps Mtb* infected mice. Data represents mean ± s.e.m (n=3 biological replicates per strain at each time point). (**C**) Relative abundance of kupyaphores species **I-VII** isolated from early passage clinical strains of *Mtb* isolated from TB patients.



Supplementary Figure 9: (A) Multiple sequence alignment of isonitrile hydratase across members of pathogenic mycobacteria family with characterized isonitrile hydratase from *Pseudomonas*. **(B)** EIC for formamide product of m/z for $(M\text{-}H)^-$ = 242.176 in *in vitro* assay of C12-isonitrile substrate with or without purified Rv0052 protein. **(C)** EIC of diformamide kupyphore analogue of m/z for $(M+H)^+$ = 884.7217 in *in vitro* reaction of kupyaphores with or without purified Rv0052 protein. **(D)** Proposed model for kupyaphore mediated nutritional passivation of zinc in *Mtb*. Secretion of kupyaphores under limiting zinc conditions would allow *Mtb* to scavenge zinc from environment with concomitant intracellular release through isonitrile hydratase. Conversely, under toxic zinc conditions, kupyaphores would chelate intracellular accumulated zinc levels thus contributing to metal detoxification by reducing the concentration of free zinc in the cytosol. Detailed mechanisms for kupyaphore regulation and export under these two strikingly opposite conditions would be interesting to explore further.

Supplementary Table 1: List of clones and primers

Clones generated in study –

ID	Clone details
pAM1	pet28c+Rv0097
pAM2	pet28c+Rv0098
pRSB1	pet28c+Rv0052

Primers used for cloning-

Primer ID	Sequence
Rv0097 FP	GGAGTACTATGACGCTTAAGGTCAAAGGCG
Rv0097 RP	GGAAGCTTTCATGCCGCGTATCCCGGCGT
Rv0098 FP	GGAGTACTATGAGCCACACCGACTTGACG
Rv0098 RP	GGAAGCTTTTACGGAATGTTGAGGGCCGC
Rv0052 FP	TTCATATGCAGGTCGCGATCGCA
Rv0052 RP	TTGAATTCGGCAAACGATGCTGATAGAA
EUF20	CTGCAGAGATCGAAGCAGGTGCTGTT
EUR20	AAGCTTTAACGCGGGATTGTTATCCT
EUF21	AAGCTTCAACTGCTCGGACTGCTGTAG
EUR21	CGCGGCCGCACACCGACGTCAGGATTACC

Primers used for gene expression analysis-

Primer ID	Sequence
Rv0097 FP	CCAAAGACCGGCCAAGAGAT
Rv0097 RP	ATGTCGCCAACCTGGTAGTG
Rv0098 FP	AACGCGGTCGAACTGATTCT
Rv0098 RP	TGGTGTTGGCAGTAGTCGTC
Rv0099 FP	ATCTGTGGTGGCGTGAACAT

Rv0099 RP	TCAGGAATCTCGTAGCACGC
Rv0100 FP	TTGATGGCGACGAAACGGAT
Rv0100 RP	GTTCGGATTGTCGGTTCACG
Rv0101 FP	TCGGAGTGTGGTTCAAGAGC
Rv0101 RP	AAACGACGACACTCCCAGAC

Supplementary Information Methods:

Bacterial culture

For metal studies, *Mtb* strains were grown in chelated Sauton's media. All glassware used for chelated media preparation were washed with teepol soap solution and dried in oven at 80°C for 3-4 hours. The glass bottles were then treated with 0.01% EDTA overnight and rinsed with 0.1N HCL the next day. This was followed by rigorous washing with double distilled water 6-8 times and drying in oven at 80°C for another hour. Sauton's media was then prepared in double distilled water and autoclaved. Upon cooling, media was treated for 12 hours with 10 g/L of chelex 100 resin at 4°C. The media was subsequently filtered and 0.025% tyloxapol and metal supplements were added as stated.

RNA-sequencing library preparation

Total DNA-free RNA sample was depleted of bacteria rRNA with Ambion's MICROBExpress kit as per the manufacturer's instructions. Total bacterial RNA ~1 μg was processed using Truseq Stranded Total RNA seq protocol (Illumina Inc). Depletion of Bacterial rRNA was done using biotinylated, target specific oligos combined with Ribo-Zero Gold rRNA removal beads. The rRNA depleted total RNA was purified, fragmented and primed for cDNA 1st- and 2nd-strand synthesis followed by A-tailing and ligation of adapters with multiplexing indexes. The products were amplified with 15 PCR cycles and purified using Agencourt AMPure XP magnetic beads (Beckman Coulter, Brea, CA, USA) according to the manufacturer's instructions. The quality of cDNA libraries was checked with Agilent DNA1000 chips (2100 Bioanalyzer; Agilent Technologies). 300x2 bp paired end sequencing was performed using V3 flow cell on Illumina MiSeq. A total of 23 million reads were generated for the three biological samples (each in triplicates)

RNA- sequencing data analysis

Mycobacterium tuberculosis genome and its annotations (gff/fna/ptt/rnt) were downloaded using an R package, Progenome (v0.0.7) (1) and the gist used for extraction is in the following reference (2). Further, raw reads were quality trimmed using Trim-Galore (v0.6.4) (3) and the resulting trimmed reads were aligned using Salmon (v0.14.1) (4) and estimated differential gene expression using DESeq2 (v1.26.0) (5) after filtering genes with count of 5 or more for at least 4 or more samples. For predicting the operons from the RNA-Seq data we used Rockhopper (v2.03) (6) and used ComplexHeatmap (v2.2.0) (7) for plotting the expression of genes in operons.

Gene expression analysis by qRT-PCR

Using NucleoSpin RNA kit, total RNA from desired cultures of the indicated strains of *Mtb* were isolated from 5 mL cultures as per manufacturer's instruction. Using the PrimeScript First Strand cDNA Synthesis Kit (Takara Bio), cDNA was generated from 1 μg of total RNA from each specified sample. RT-qPCR reactions were prepared using 1 μL of cDNA reaction mixture for each gene-specific primer per reaction with SYBR Green Master Mix (Applied Biosystems) as per the manufacturer's instruction in a Roche LightCycler 480 instrument II. Template normalization was performed by dividing the absolute gene expression of specific genes by the absolute gene expression of 16S rRNA. The sequences of primers used for qRT-PCR are provided in SI Table 1. Reactions without the cDNA were used as no-template negative control.

Generation of Δ*nrps* and Δ*nrps:nrps Mtb* strains

The *nrps* deletion strain was constructed by two step homologous recombination as previously described (8). A suicide vector was constructed by amplifying the regions upstream and downstream of the gene to create an unmarked deletion using primer pairs EUF20/EUR20 and

EUF21/EUR21, cloning into p2NIL and adding the selection cassette from pGOAL19 (*sacB*, *lacZ*, *hyg*, *kan*). Single crossovers were isolated on kanamycin and hygromycin after electroporation with 5 μg of plasmid DNA. Double cross overs were isolated by negative selection on sucrose and screening for lack of LacZ activity. The deletion strain was identified using PCR amplification and confirmed by Southern blotting. For generation of Δ*nrps:nrps Mtb* strain, a 28-kb *Mtb* H37Rv cosmid fragment (spanning genomic position 104933 bp to 133536 bp) packaged in the integrative shuttle cosmid vector pYUB412-Kan (shared by Dr. Apoorva Bhatt) was introduced into the Δ*nrps* strain by electroporation, and kanamycin-resistant transformants were confirmed by PCR amplification. This recombinant cosmid spans a region that extends from Rv0096 to Rv0109.

Determination of total metal ions in *in vitro* bacterial cultures

Total metal concentrations were measured by ICP-MS. Briefly 0.8-1.0 OD WT *Mtb* cultures grown in chelated Sauton's medium with 0.025% tyloxapol supplemented with $ZnSO_4$ were spun down and washed with 1X PBS 3 times. Subsequently cells were subjected to lysis by boiling in 0.1% SDS and 0.2% HNO_3 for 15 minutes in trace-element free 1.5 ml micro-centrifuge tubes. Volume were made up to 1 ml using MS grade water and samples were filtered through 0.2 μ PTFE syringe filters before running on ICP-MS (ThermoXcaliber II). For comparative metal content analysis of Δ*nrps* with WT and Δ*nrps:nrps Mtb* strains, log phase cultures from 7H9 were spun down and washed with 1X PBS thrice. *Mtb* strains were then grown in chelated Sauton's medium with 0.025% tyloxapol supplemented with either low- or high-zinc for one generation time. Cells were subsequently lysed and ICP-MS was performed as described above. Absolute part per billions (ppb) counts were normalized to protein concentration estimated for each sample and are plotted.

Cell culture and infection

The murine immortalized bone marrow derived macrophage cell line RAW 264.7 (source, European Collection of Authenticated Cell Cultures) was cultured in Dulbecco's modified Eagle's medium (DMEM) with 2 mM L-glutamine and 10% fetal calf serum (FCS) (Gibco) at 37°C under 5% CO_2. Cells were routinely collected by scraping. Cells were seeded at required density a day before infection. Cells were infected with either mCherry labelled WT, Δ*nrps* and Δ*nrps:nrps Mtb* at a multiplicity of infection (MOI) of 5 bacteria per cell (for imaging and ICP-MS studies) and MOI = 50 for mRNA expression studies of intracellular bacteria for 4 hours. After bacterial challenge, medium was changed to antibiotic containing one for 20 mins to get rid of extracellular bacteria. Cultures were then subsequently maintained in antibiotic free medium for indicated time points.

Determination of total zinc in macrophage infection studies

Total metal concentrations were measured by ICP-MS. At indicated time points, macrophage were lysed with 0.05% SDS (in M.S grade water) and spun down at 4000 rpm for 8 minutes. Supernatant was collected separately to measure macrophage total metal content while the bacterial pellet was subjected to lysis by boiling in 0.1% SDS and 0.2% nitric acid for 15 minutes in trace-element free 1.5ml micro-centrifuge tubes. Volume were made up to 1 ml using MS grade water and samples were filtered before running on ICP-MS (Thermo Xcaliber II). Absolute ppb counts were normalized to protein concentration estimated for each sample and plotted.

Determination of free zinc in macrophage infection studies

Macrophage cultures uninfected or infected with mCherry labelled *Mtb* strains were washed with PBS and incubated with 0.5 μM Fluozin 3-AM for 30 minutes at room temperature. After 30 minutes cells were washed with PBS to remove dye remnants. Cells loaded with FZ3-AM dye were then incubated in 1X PBS for another 30 mins at 37 °C. PBS was subsequently removed after incubation

and cells were fixed with 4% paraformaldehyde in 1X PBS. Fixed cells were then mounted with DAPI. DAPI-positive, mCherry-positive and FZ3 green-positive cells were selected for each of the triplicate samples by confocal microscopy using Zeiss LSM980 and mean pixel intensity for FZ3 signals was analyzed using ImageJ.

Computational analysis

10,932 whole bacterial genomes were downloaded in FASTA (.fna) format from RefSeq. The genomes spanned 1021 unique genera, and 2854 unique species. The dataset was non-redundant at strain level. The genomes were run through antiSMASH 4. The NRPS biosynthetic gene cluster containing GenBank (.gbk) files were extracted and converted to FASTA (.fa) format using an in-house script. These NRPS biosynthetic gene clusters were searched for all HotDog-fold (thioesterase) clan (Pfam: CL0050) family pHMMs alongside presence of TauD (Pfam: PF02668). Except FcoT (Pfam: PF10862) family containing biosynthetic gene clusters none of the other thioesterase families had an organization similar to the NRPS biosynthetic gene cluster. Hence, only FcoT containing biosynthetic gene clusters were used for further analysis. A representative dataset of 73 genomes was picked up from the putative NRPS biosynthetic gene cluster containing genomes. Among these genomes, we searched for *nrps* genes in which the first condensation domain ($C1^{*}$) was not being picked up by Pfam pHMMs. C-domains being detected by Pfam pHMMs were labelled C2. The sequences for all these condensation domains were subjected to Multiple Sequence Alignment (MSA) using Clustal Omega. The cladogram for these MSA sequences was constructed using iTOL.

Cloning, expression and purification

Rv0097, Rv0098 and Rv0052 were PCR amplified with template from BAC clone library using KOD polymerase from TOYOBO on Veriti thermal cycler. Amplification conditions for genes were, initial denaturation: 94°C for 2 minutes, denaturation: 94°C for 15 seconds, annealing: 62°C for 45 seconds and elongation: 68°C for 2 minutes. Amplified genes were further cloned in pET- 28c expression vector using Nde1 and HindIII (NEB) restriction enzymes and T4 DNA ligase (NEB). Rv0097, Rv0098 and Rv0052 were expressed and purified as His tagged protein from E. coli BL21 (DE3) competent cells. In brief, 2 L LB media for each protein was inoculated with respective overnight grown starter culture. Once OD reached 0.6 at 37°C, cultures were induced with 0.5 mM and kept at 30°C for 6 to 8 hours. Cells were spun down, washed, and resuspended in lysis buffer (pH-7.5, 50 mM Tris-HCl, 150 mM NaCl, 10% glycerol). Cells were the centrifuged to remove cell debris and supernatants were incubated for 1 hour with Ni-NTA agarose resin beads at 4°C. The mixture was loaded onto a column using gravity flow. The resin was washed with wash buffer (50 mM Tris pH 8.0, 10% glycerol) till the unbound proteins were removed. The protein was eluted using elution buffers containing increasing concentration of imidazole. Fractions containing the proteins of interest were pooled and 1 mM TCEP was added and stored at -80°C.

Biochemical assays of Rv0097, Rv0098 and Rv0052

Enzymatic reactions were set up at pH-7.5 (Tris-HCl) with 200 mM glycine, 5 mM (C4, C8 or C12-NAC) substrate and 50 μM enzyme- Rv0098 or 400 μM iron sulphate and 500 μM alpha-ketoglutarate and 4 mM ascorbate with 50 μM Rv0098 and Rv0097 each. Reactions were mixed gently and kept at 30°C for 6 hours. For biochemical characterization of Rv0052, to the above reaction mix of Rv0098 and Rv0097, 50 μM Rv0052 was additionally added. For each assay set, no-protein control reactions were also set. With equal volume of ethyl acetate, reaction products were extracted and dried. High resolution LC-MS analysis was performed using a Gemini 5U C-18 column (Phenomenex, 5 μm, 50 x 4.6 mm) using solvents, flow rates and MS parameters described

earlier on a Sciex X500R Quadrupole Time Of Flight (QTOF) mass spectrometer fitted with an ExionLC UPHLC system.

Extraction and analysis of kupyaphores from *Mtb* cultures

The organic layer collected from ethyl acetate based small molecule extraction of *Mtb* cultures was separated and evaporated to dryness. The residual material was dissolved in minimal volume of methanol and analyzed by an information dependent acquisition (IDA) scanning on a Sciex X500R QTOF mass spectrometer fitted with an ExionLC UPHLC system using the SciexOS software as per previously described method (9-10). The LC separation was achieved on a Gemini 5U C18 column (Phenomenex, 5 μm, 50 x 4.6 mm) coupled to a Gemini guard column (Phenomenex, 4 x 3 mm, Phenomenex security cartridge). All metabolites corresponding to kupyaphore species **I** to **VII** were analyzed by IDA scanning in both positive and negative ionization mode using an electrospray ionization (ESI) source with solvent systems, flow rates and a solvent gradient described earlier. The total scan time for both the MS1 and MS2 spectra was 2.5 s and the collision energy (volts) of 50 was used. The declustering potential and ion source voltage were set at 100 and 5500 volts respectively.

Click chemistry of *Mtb* culture extract with tetrazine

Dry extracts of WT *Mtb* biofilm culture were redissolved in 400 µL methanol, to which 100 mM 3,6-Di-2-pyridyl-1,2,4,5-tetrazine in dichloromethane was added to a final concentration of 10 mM. The mixture was stirred at room temperature for 3 hours. The universal product was separated and analyzed using previously described method (11).

Acid catalyzed hydrolysis for isonitrile functional group confirmation

Ethyl acetate metabolite extracts of WT *Mtb* biofilm culture were treated with formic acid (5% v/v) for 4 hours at room temperature. Extracted ion analysis was performed for the parent lipopeptide peak (m/z for $[M+H]^+$ ion 848.6987) and synthesized β-amine lipopetide (m/z for $[M+H]^+$ ion 828.7300) using the LC-MS protocol described in the earlier section.

Mouse Infection studies

Six to eight-week-old female Balb/c mice were housed in individually ventilated cages at the animal house facility (TACF) of ICGEB, New Delhi, India. Mice were infected with indicated *Mtb* strain by aerosol route to implant nearly 200 CFU per lung. Deposition of bacterium inside lung was assessed by sacrificing three mice 24 hours post infection and plating the lungs homogenates in triplicates on 7H11-OADC agar plates. Bacillary loads in lung were evaluated at different time points after aerosol infection to follow the course of infection. A part of the lung was used for metabolite analysis. At each time point, corresponding uninfected mice were taken as controls.

Extraction and analysis of *Mtb* metabolites from mouse infection studies

For *Mtb* metabolite extraction from mice, 0.1 g of lung tissue were taken from left apical lobe of uninfected and infected mice at indicated time points. 5 times the volume of ethyl acetate was added. The sample were then homogenized using 0.5 μ zirconium bead in a bead beater. The organic layer was collected, transferred to fresh tube, dried and subjected to targeted LC-HRMS studies. All the species analyzed were quantified using the multiple reaction monitoring high resolution (MRM-HR) LC-MS method on a Sciex X500R QTOF mass spectrometer fitted with an ExionLC UHPLC system. All data was collected and analyzed using the SciexOS software. All metabolite estimations were performed using an ESI source, with the following MS parameters: curtain gas = 20 L/min, ion spray voltage = 5500 V, temperature = 500 °C.

Supernatant exchange for Δ*nrps* growth rescue

Δ*nrps Mtb* strain was inoculated in Sauton's medium at starting OD of 0.02. After one generation time, culture were spun down and medium was removed. The cells were then resuspended in either fresh Sauton's medium or in supernatant collected after spinning down WT *Mtb* cells grown in Sauton's medium for 7 days. Growth was then monitored in both the conditions by measuring optical density at 600 nm for 6 days.

Radioactive zinc 65 uptake assay

Logarithmic cultures of WT, Δ*nrps* and Δ*nrps:nrps Mtb* strains were spun down and washed thrice with 1X PBS. Equal number of cells were then incubated with 0.5 μM of radioactive zinc-65 for 4 hours. After 4 hours, cells were spun down and washed once with 1X PBS to get rid of exogenous zinc 65. Subsequently cells were lysed in 1X PBS supplemented with 5% glycerol by bead beating for 3 cycles. The supernatant was then collected and spotted on blotting sheet. Intracellular radioactive count was then measured by autoradiography and quantitated using ImageJ analysis. Mean pixel intensity so obtained was then normalized to protein content estimated by BCA for each sample.

Selection criteria for TB patients and isolation of clinical *Mtb* strains from TB patients

The initial assessment of bone/pulmonary TB patients was done by doctors on the basis of clinical suspicion. The symptoms included weight loss, persistent pain, cough and bone damage near pathogenic foci. TB subjects were then tested by chest X-ray, sputum culture test, radiological examination of infected tissue and by microscopic examination for bacterial presence by acid-fast bacilli (AFB) staining. HIV infected subjects and patients undergoing anti tubercular therapy were excluded from the study. The pulmonary mycobacterial strains were isolated from sputum of infected subjects. Mycobacterial strains from bone were isolated from infected skeletal tissue. The infected tissue was then lysed with 4% NaOH and inoculated onto LJ slants. The slants were then incubated at 37°C until visible colonies were observed.

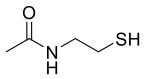
***Mtb* metabolite extraction and analysis from clinical *Mtb* strains isolated from TB patients**

Colonies grown on LJ slant were scraped and collected. 5 times the volume of ethyl acetate was added. The organic layer was collected, transferred to fresh tube, dried and subjected to targeted LC-HRMS studies as described previously. All the species analyzed were quantified using the multiple reaction monitoring high resolution (MRM-HR) LC-MS method on a Sciex X500R QTOF mass spectrometer fitted with an ExionLC UHPLC system. All data was collected and analyzed using the SciexOS software. All metabolite estimations were performed using an ESI source, with the following MS parameters: curtain gas = 20 L/min, ion spray voltage = 5500 V, temperature = 500 °C.

Statistical Analysis

GraphPad Prism 8 software was used for statistical analysis. Statistical significance was analyzed by Student's t-test or one-way or two-way ANOVA with $p > 0.05$ (not significant), $*p < 0.05$, and $**p < 0.01$ ***p<0.005 when applicable. Data were plotted as the mean, with error bars representing SEM of three biological replicates. Schematic diagrams were adapted from biorender.com.

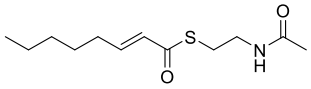
Chemical Synthesis of *N*-acetyl cysteamine thioester of 2-dodecenoic acid



***N*-(2-mercaptoethyl)acetamide (03):** Compound (03) was synthesized by following reported procedure.[1]

IR υ_{max} (film): 3285, 1645, 1550, 1371, 1292 cm^{-1}

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.48 (bs, 1H), 3.39 – 3.35 (m, 2H), 2.65 – 2.59 (m, 2H), 1.96 (s, 3H), 1.37 (t, *J* = 8.5 Hz, 1H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ 170.59, 42.57, 24.53, 23.20.

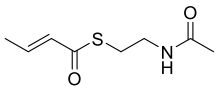


(*E*)-2-octenoyl-acyl-*N*-acetylcysteamine (04): Compound (04) was synthesized by following reported procedure.[2, 3]

IR υ_{max} (film): 3285, 2925, 1644, 1546, 1368, 1286, 1019, 807 cm^{-1}

^{1}H NMR (400 MHz, CD_3OD) δ 6.94 (dt, *J* = 15.2, 7.0 Hz, 1H), 6.18 (d, *J* = 15.5 Hz, 1H), 3.36 (t, *J* = 6.7 Hz, 2H), 3.08 (t, *J* = 6.7 Hz, 2H), 2.24 (td, *J* = 8.0, 1.1 Hz, 2H), 1.94 (s, 3H), 1.50 (dt, *J* = 14.3, 7.3 Hz, 2H), 1.35 (m, 4H), 0.94 (t, *J* = 6.8 Hz, 3H). **^{13}C NMR (100 MHz, CD_3OD)** δ 189.59, 172.02, 146.14, 128.12, 38.84, 31.71, 31.11, 27.50, 27.45, 22.08, 21.10, 12.93.

HRMS (ESI): m/z calculated for $C_{12}H_{22}O_2NS$ $[M+H]^+$ 244.1366 observed 244.1365.

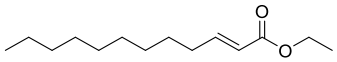


***S*-(2-acetamidoethyl) (*E*)-but-2-enethioate (05):** Compound (05) was synthesized by following reported procedure.[2]

IR υ_{max} (film): 3282, 1646, 1547, 1437, 1286, 1046, 812 cm^{-1}

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.93 (dq, *J* = 20.9, 6.9 Hz, 1H), 6.19 – 6.12 (m, 1H), 6.07 (bs, 1H), 3.45 (dd, *J* = 11.9, 5.9 Hz, 2H), 3.08 (t, *J* = 6.4 Hz, 2H), 1.96 (s, 3H), 1.89 (dd, *J* = 6.8, 1.2 Hz, 3H).

HRMS (ESI): m/z calculated for $C_8H_{14}O_2NS$ $[M+H]^+$ 188.0740 observed 188.0739.

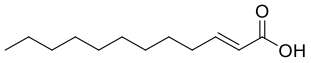


Ethyl (*E*)-dodec-2-enoate (08): Compound (08) was synthesized by following reported procedure.[4]

IR υ_{max} (film): 2925, 2855, 1721, 1655, 1461, 1265, 1180, 981, 750 cm^{-1}

^{1}H NMR (200 MHz, $CDCl_3$) δ 6.95 (dt, *J* = 15.4, 6.9 Hz, 1H), 5.79 (d, *J* = 15.6 Hz, 1H), 4.17 (q, *J* = 7.1 Hz, 2H), 2.17 (q, *J* = 6.3 Hz, 2H), 1.56 – 1.19 (m, 17H), 0.86 (t, *J* = 6.2 Hz, 3H).

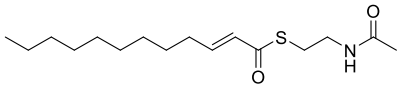
HRMS (ESI): m/z calculated for $C_{14}H_{27}O_2$ $[M+H]^+$ 227.2006 observed 227.2004.



(*E*)-dodec-2-enoic acid (09): Compound (09) was synthesized by following reported procedure.[5]

IR υ_{max} (film): 2924, 1694, 1286, 753 cm^{-1}

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.08 (dt, *J* = 15.6, 7.0 Hz, 1H), 5.82 (dt, *J* = 15.6, 1.5 Hz, 1H), 2.25 – 2.19 (m, 2H), 1.50 – 1.42 (m, 2H), 1.26 (s, 12H), 0.88 (t, *J* = 6.9 Hz, 3H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ 172.46, 152.68, 120.76, 32.47, 32.01, 29.62, 29.52, 29.43, 29.29, 28.02, 22.81, 14.24.



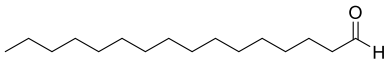
***S*-(2-acetamidoethyl) (*E*)-dodec-2-enethioate (10):** (*E*)-dodec-2-enoic acid (09) (0.858 g; 4.33 mmol) was added to 15 mL of DMF at 0 °C and treated with diphenylphosphoryl azide (DPPA) (0.930 mL; 4.3 mmol) and triethylamine (1.11 mL; 7.9 mmol) for 2 h. *N*-(2-mercaptoethyl)acetamide (03) (0.430 g; 3.61 mmol) in 5 ml DMF was added to the solution and stirred at room temperature for overnight. The reaction was quenched with 50 mL of ice-cold water and exacted twice with ethyl acetate. The organic phase was combined, dried, and evaporated, and *S*-(2-acetamidoethyl) (*E*)-dodec-2-enethioate was purified with silica gel (100-200 mesh) chromatography to give 485 mg (45% yield) of pale yellow solid.[6]

IR υ_{max} (film): 3282, 2923, 2856, 1660, 1547, 1451, 1287, 977, 808 cm^{-1}

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.91 (dt, *J* = 15.5, 6.9 Hz, 1H), 6.10 (dt, *J* = 15.5, 1.4 Hz, 1H), 6.07 (bs, 1H), 3.44 (dd, *J* = 12.4, 6.2 Hz, 2H), 3.07 (t, *J* = 6.4 Hz, 2H), 2.21 – 2.15 (m, 2H), 1.95 (s, 3H), 1.48 – 1.41 (m, 2H), 1.24 (s, 14H), 0.86 (t, *J* = 6.8 Hz, 3H). **^{13}C NMR (100 MHz, $CDCl_3$)** δ 190.56, 170.50, 146.96, 128.35, 39.93, 32.36, 31.96, 29.56, 29.46, 29.37, 29.27, 28.31, 28.01, 23.31, 22.76, 14.21.

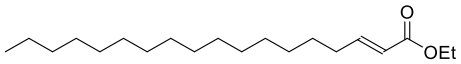
HRMS (ESI): m/z calculated for $C_{16}H_{30}O_2NS$ $[M+H]^+$ 300.1992 observed 300.1989.

Chemical Synthesis of β-amine substituted lipopeptide



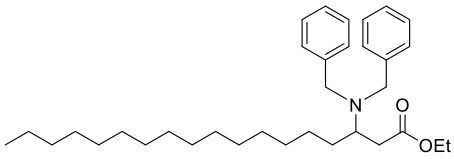
Hexadecanal (12): Compound **12** was prepared by following the literature procedure.[7]

Yield- 85 % as a white solid.



Ethyl (*E*)-octadec-2-enoate (13): Compound **13** was prepared by following the literature procedure.[8]

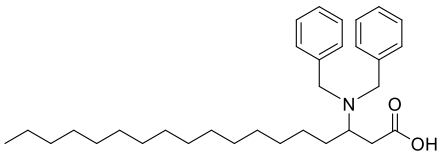
Yield- 90 % as a colorless oil**.**



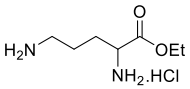
Ethyl 3-(dibenzylamino) octadecanoate (14): Dibenzylamine 18.56 mL (96.61 mmol) was added in THF (150 mL) at -78 °C followed by drop wise addition of *n*-butyl lithium 66.41 mL (106.26 mmol, 1.6 M in Hexane) over a period of 15 min under inert condition. Ethyl (*E*)-octadec-2-enoate **13** (15 g, 48.30 mmol) in THF (50 mL) was added after 30 min. The resulting solution was stirred at -78 °C for 4 h. The reaction mixture was quenched with saturated ammonium chloride and extracted with ethyl acetate (3 x 70 mL). The combined organic extracts were dried over anhydrous sodium sulphate, filtered, and concentrated in *vacuo*. The organic residue was purified by column chromatography to afford Ethyl 3-(dibenzylamino) octadecanoate **14** (16.2 g, yield-66 %) as a pale yellow oil.[9]

IR υ_{max} (thin film): 2923, 2852, 1734, 1454, 1367, 1178 cm^{-1}.

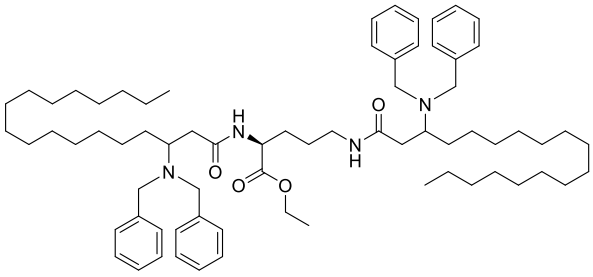
^{1}H NMR (400 MHz, $CDCl_3$): δ 7.35 – 7.19 (m, 10H), 4.14 (dq, *J* = 10.7, 7.1 Hz, 1H), 4.03 (dq, *J* = 10.7, 7.1 Hz, 1H), 3.62 (d, *J* = 13.6 Hz, 2H), 3.49 (d, *J* = 13.6 Hz, 2H), 3.08 (p, *J* = 7.0 Hz, 1H), 2.64 (dd, *J* = 13.7, 6.5 Hz, 1H), 2.25 (dd, *J* = 13.7, 7.4 Hz, 1H), 1.67 (dq, *J* = 12.0, 7.2 Hz, 2H), 1.27 (s, 26H), 1.20 (t, *J* = 7.1 Hz, 3H), 0.89 (t, *J* = 6.8 Hz, 3H); **^{13}C NMR (100 MHz, $CDCl_3$):** δ 173.16, 140.09, 129.11, 128.21, 126.95, 60.44, 55.56, 53.52, 35.78, 32.08, 30.68, 29.86, 29.83, 29.78, 29.74, 29.61, 29.52, 26.81, 22.85, 14.28.
HRMS (ESI): m/z calculated for $C_{34}H_5O_2N$ $[M+H]^+$ 508.4149 observed 508.4152.



3-(dibenzylamino) octadecanoic acid (15): Ethyl 3-(dibenzylamino) octadecanoate **14** (16 g, 31.50 mmol) was dissolved in a mixture of solvents THF, ethanol and water in a ratio of (3:2:1) followed by addition of NaOH (7.56 g, 189.00 mmol) at 0 °C. The resulting solution was stirred at room temperature for 6 h and completion of reaction was monitored by TLC. The reaction mass was concentrated and adjusted the pH of reaction mass at ~ 3-4 and extracted with ethyl acetate (2 x 100 mL). The combined organic layer was dried over anhydrous sodium sulphate, filtered and concentrated in *vacuo*. The organic residue was purified by column chromatography to afford 3-(dibenzylamino) octadecanoic acid **15** (8.50 g, yield-80 % brsm) as yellowish oil.
IR υ_{max} (thin film): 3028, 2922, 2852, 1709, 1456, 1377 cm^{-1}.
^{1}H NMR (200 MHz, $CDCl_3$): δ 7.26 – 7.19 (m, 10H), 3.93 (d, *J* = 13.0 Hz, 2H), 3.37 (d, *J* = 13.0 Hz, 2H), 3.01 (d, *J* = 4.0 Hz, 1H), 2.51 – 2.27 (m, 2H), 1.65 (dt, *J* = 12.5, 8.0 Hz, 1H), 1.20 (s, 27H), 0.79 (t, *J* = 6.5 Hz, 3H); **^{13}C NMR (100 MHz, $CDCl_3$):** δ 173.26, 135.08, 129.89, 129.10, 128.48, 55.18, 52.81, 33.83, 32.06, 29.83, 29.80, 29.75, 29.72, 29.66, 29.56, 29.50, 26.82, 26.69, 22.82, 14.25.
HRMS (ESI): m/z calculated for $C_{32}H_{50}O_2N$ $[M+H]^+$ 480.3836 observed 480.3839.



Ethyl 2, 5-diaminopentanoate (17): Compound **17** was prepared by following the literature procedure.[12]



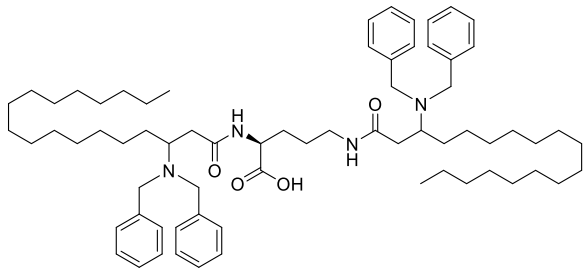
Ethyl 2, 5-bis(3-(dibenzylamino)octadecanamido)pentanoate (18): 3-(dibenzylamino) octadecanoic acid **15** (6.098 g, 12.711 mmol) was dissolved in dry DMF (30 mL) at 0 ºC. HATU (4.833 g, 12.711 mmol) and *N*, *N*-diisopropylethylamine (4.428 mL, 25.422 mmol) were added accordingly. After 10 min ethyl 2, 5-diaminopentanoate **17** (1 g, 5.084 mmol) was added. The temperature of reaction mixture was maintained at 0 ºC for 2 h. After completion, the reaction mixture was diluted with ethyl acetate and washed with saturated brine solution (1 x 50 mL) and water (2 x 25 mL). The resulting organic layer was dried over anhydrous sodium sulphate, filtered,

and concentrated in vacuo. The organic residue was purified by column chromatography to afford Ethyl 2, 5-bis (3-(dibenzylamino)octadecanamido)pentanoate **18** (5 g, yield- 91%) as a colorless oil.[10]

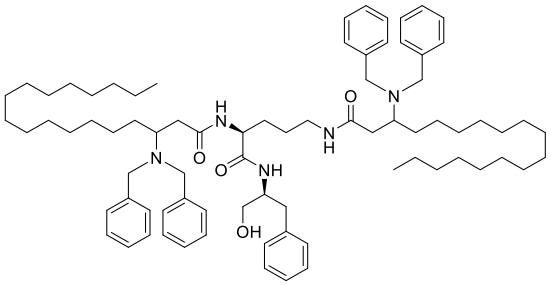
IR υ_{max} (thin film): 3298, 2923, 1738, 1646, 1539, 1454 cm $^{-1}$.

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.34 – 7.18 (m, 20H), 4.49 (dd, *J* = 13.1, 7.0 Hz, 1H), 4.24 – 4.10 (m, 2H), 3.74 – 3.69 (m, 2H), 3.59 (ddd, *J* = 18.2, 13.6, 6.7 Hz, 4H), 3.42 (dd, *J* = 13.2, 3.3 Hz, 2H), 3.24 – 3.02 (m, 2H), 2.94 – 2.66 (m, 2H), 2.64 – 2.43 (m, 2H), 2.24 – 2.11 (m, 1H), 1.68 (m, 4H), 1.47 – 1.39 (m, 1H), 1.26 (d, *J* = 2.0 Hz, 59H), 0.88 (t, *J* = 6.8 Hz, 6H); **^{13}C NMR (100 MHz, $CDCl_3$): (characteristic peaks)** δ 172.46, 172.36, 172.28, 139.64, 139.37, 139.19, 129.47, 129.24, 128.58, 127.48, 127.11, 61.42, 56.30, 55.75, 53.69, 52.49, 38.59, 36.89, 36.61, 32.06, 30.51, 29.85, 28.25, 27.31, 26.94, 26.31, 22.83, 14.33.

HRMS (ESI): m/z calculated for $C_{71}H_{111}O_4N_4$ $[M+H]^+$ 1083.8600 observed 1083.8602.



2, 5-bis(3-(dibenzylamino)octadecanamido)pentanoic acid (19): Ethyl 2,5-bis(3-(dibenzylamino)octadecanamido)pentanoate **18** (5.0 g, 4.613 mmol) was dissolved in a mixture of solvents THF, ethanol and water in a ratio of (3:2:1) at 0 °C. followed by addition of NaOH (1.107 g, 27.683 mmol) at same temperature. The resulting solution was stirred at room temperature for 6 h, completion of reaction was monitored by TLC. The reaction mass was concentrated and the pH was adjusted at ~ 3-4 by 1N HCl. Extracted with ethyl acetate (2 x 50 mL). The combined organic layer was dried over anhydrous sodium sulphate, filtered and concentrated in *vacuo*. Crude (organic residue) was forwarded into next step without purification.

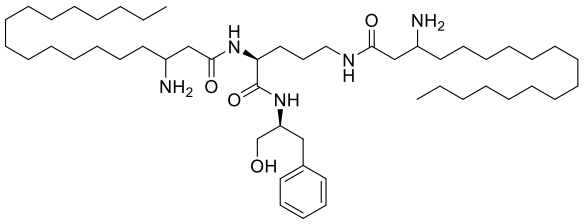


***N*,*N*'-(5-((1-hydroxy-3-phenylpropan-2-yl)amino)-5-oxopentane-1,4-diyl)bis(3-(dibenzylamino)octadecanamide) (21):** 2,5-bis(3-(dibenzylamino)octadecanamido)pentanoic acid **19** (4.0 g, 3.789 mmol) was dissolved in DMF (5 mL) at 0 ºC. HATU (2.881 g, 7.578 mmol) and *N*, *N*-diisopropylethylamine (1.98 mL, 11.367 mmol) were added to it. After 10 min, (*S*)-2-amino-3-phenylpropan-1-ol **20** (0.716 g, 4.736 mmol) was added. The temperature of reaction mixture was maintained at 0 ºC for 2 h. After completion, the reaction mixture was diluted with ethyl acetate and washed with saturated brine (50 mL) and water (2 x 50 mL). The resulting organic layer was dried over anhydrous sodium sulphate, filtered, and concentrated in vacuo. The organic residue was purified by column chromatography to afford *N*, *N*'-(5-((1-hydroxy-3-phenylpropan-2-yl) amino)-5-oxopentane-1,4-diyl)bis(3-(dibenzylamino)octadecanamide) **21** (3.4 g, yield- 76 %) as a yellowish sticky oil.[10]

IR υ_{max} (thin film): 3277, 2954, 2361, 1648, 1540, 1457 cm $^{-1}$.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.32 – 7.15 (m, 25H), 4.55 – 4.44 (m, 1H), 4.11 (s, 1H), 3.74 – 3.49 (m, 8H), 3.38 (d, J = 10.7 Hz, 2H), 3.06 – 2.80 (m, 5H), 2.59 – 2.15 (m, 5H), 1.83 – 1.59 (m, 4H),

1.27 (s, 56H), 0.88 (t, *J* = 6.4 Hz, 6H); **^{13}C NMR (100 MHz, $CDCl_3$) (characteristic peaks)** δ 176.09, 173.76, 172.55, 138.65, 138.24, 129.40, 128.73, 128.41, 127.70, 127.12, 126.45, 63.80, 63.23, 56.23, 55.97, 55.52, 53.70, 53.60, 53.44, 53.22, 52.02, 51.66, 37.72, 37.67, 37.65, 37.05, 36.64, 32.07, 31.07, 29.51, 27.55, 26.19, 25.87, 22.83, 20.96, 14.27.
HRMS (ESI): m/z calculated for $C_{78}H_{118}O_4N_5$ $[M+H]^+$ 1188.9178 observed 1188.9178.



***N*,*N*'-(5-(((S)-1-hydroxy-3-phenylpropan-2-yl)amino)-5-oxopentane-1,4-diyl)bis(3-aminooctadecanamide) (22):** Compound **21** (1 g, 0.841 mmol) was dissolved in dry MeOH (15 mL) and transferred into Parr reactor by cannula. $Pd(OH)_2$ (10 mol%) and acetic acid(cat.) were added. The reactor was closed properly and filled with hydrogen gas (pressure 250 psi). The reaction mass was stirred at room temperature for 24 h. After completion of reaction, the reaction mixture was filtered through celite pad and concentrated to afford *N*, *N*'-(5-(((S)-1-hydroxy-3-phenylpropan-2-yl)amino)-5-oxopentane-1,4-diyl)bis(3-aminooctadecanamide) **22** (0.3 g, yield-43%) as a colorless sticky oil.[11]
IR υ_{max} (thin film): 3406, 3360, 3321, 2924, 1708, 1460, 1082 cm^{-1}.
^{1}H NMR (500 MHz, CD_3OD) δ 7.28 – 7.18 (m, 5H), 4.33 – 4.24 (m, 1H), 4.16 – 4.08 (m, 1H), 3.52 (ddd, *J* = 17.5, 12.4, 5.4 Hz, 2H), 3.27 – 3.09 (m, 4H), 2.97 – 2.87 (m, 1H), 2.74 (ddd, *J* = 22.6, 13.5, 8.5 Hz, 1H), 2.50 (dd, *J* = 37.9, 14.6 Hz, 2H), 2.37 – 2.30 (m, 2H), 1.76 – 1.57 (m, 2H), 1.51 (d, *J* = 6.4 Hz, 6H), 1.31 (d, *J* = 18.7 Hz, 56H), 0.90 (t, *J* = 6.9 Hz, 6H); **^{13}C NMR (125 MHz, CD_3OD) (characteristic peaks)** δ 173.90, 173.43, 139.83, 130.42, 129.38, 127.37, 64.35, 63.84, 54.76, 41.73, 41.08, 39.65, 37.95, 36.31, 33.09, 30.66, 26.47, 23.75, 14.49.
HRMS (ESI): m/z calculated for $C_{50}H_{94}O_4N_5$ $[M+H]^+$ 828.7300 observed 828.7305.

Supplementary Information References:

1. Niu,Y. Github: ProGenome.
2. Extract gff/fna/ptt/rnt annotation of Mycobacterium tuberculosis - https://gist.github.com/viv3kanand/c26615f96975f7c71b8510fb4ac88826.
3. Krueger,F. Trim Galore. Babraham Bioinforma (2016).
4. Patro,R., Duggal,G., Love,M.I., Irizarry,R.A. and Kingsford,C. Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods, 10.1038/nmeth.4197 (2017).
5. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol., 10.1186/s13059-014-0550-8 (2014).
6. Tjaden,B. A computational system for identifying operons based on RNA- seq data. Methods, 10.1016/j.ymeth.2019.03.026 (2020).
7. Gu,Z., Eils,R. and Schlesner,M. Complex heat maps reveal patterns and correlations in multidimensional genomic data. Bioinformatics, 10.1093/bioinformatics/btw313 (2016).
8. Parish T & Stoker NG, Use of a flexible cassette method to generate a double unmarked Mycobacterium tuberculosis tlyA plcABC mutant by gene replacement. *Microbiology* 146 (Pt 8), 1969-1975 (2000).
9. Kelkar DS, Ravikumar G, Mehendale N, Singh S, Joshi A, Sharma AK, Mhetre A, Rajendran A, Chakrapani H, & Kamat SS, A chemical-genetic screen identifies ABHD12 as an oxidized-phosphatidylserine lipase. *Nat Chem Biol* 15, 169-178 (2019).
10. Khandelwal N, Shaikh M, Mhetre A, Singh S, Sajeevan T, Joshi A, Balaji KN, Chakrapani H, & Kamat SS, Fatty acid chain length drives lysophosphatidylserine-dependent immunological outputs. Cell Chem Biol 28, 1169-1179 e1166 (2021).
11. Huang YB, Cai W, Del Rio Flores A, Twigg FF, & Zhang W, Facile Discovery and Quantification of Isonitrile Natural Products via Tetrazine-Based Click Reactions. Anal Chem 92, 599-602 (2020).