Kupyaphores are counter regulatory zinc homeostatic metallophores required for Mycobacterium tuberculosis colonization

Kritee Mehdiratta
CSIR-Institute of Genomics and Integrative Biology
Shubham Singh
Indian Institute of Science Education and Research Pune
Sachin Sharma
National Institute of Immunology
Rashmi Bhosale
National Institute of Immunology
Rahul Choudhury
CSIR-Institute of Genomics and Integrative Biology
Dattatraya P. Masal
CSIR-Institute of Genomics and Integrative Biology
Alzu Manocha
National Institute of Immunology
Bhushan Dilip Dhamale
National Institute of Immunology
Naseem Khan
Translational Health Science and Technology Institute
Vivekanand A.
CSIR-Institute of Genomics and Integrative Biology
Pooja Sharma
CSIR-Institute of Genomics and Integrative Biology
Melanie Ikeh
School of Natural Sciences, University of California
Amanda C. Brown
Texas A&M University
Tanya Parish
Seattle Children's Research Institute
Anil Ojha
University at Albany
Joy Sarojini Michael

Christian Medical College, Vellore
Mohammed Faruq
CSIR-Institute of Genomics and Integrative Biology
Guruprasad R. Medigeshi
Translational Health Science and Technology Institute
Debasisa Mohanty
National Institute of Immunology
D Srinivasa Reddy
CSIR-National Chemical Laboratory https://orcid.org/0000-0003-3270-315X
Vivek Natarajan
CSIR-Institute of Genomics and Integrative Biology
Siddhesh Kamat
Indian Institute of Science Education and Research Pune https://orcid.org/0000-0001-6132-7574
Rajesh Gokhale (rsg@nii.ac.in)
National Institute of Immunology

Article

Keywords: Tuberculosis (TB), zinc, infectious disease

Posted Date: February 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-226693/v1

Kupyaphores are counter regulatory zinc homeostatic metallophores required for *Mycobacterium tuberculosis* colonization

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]

[1]CSIR-Institute of Genomics and Integrative Biology, Mathura Road, New Delhi; [2]Academy of Science and Innovative Research, Ghaziabad, India, [3]Department of Biology, Indian Institute of Science Education and Research, Pashan, Pune, Maharashtra, India; [4]Immunometabolism Laboratory, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi, India; [5]CSIR-National Chemical Laboratory, Pune, India; [6]Vaccine and Infectious Disease Research Center, Translational Health Science and Technology Institute, Faridabad, Haryana; [6]Queen Mary University of London, Barts & The London School of Medicine and Dentistry, London E1 2AT, United Kingdom; [8]Division of Genetics , Wadsworth Center, New York State Department of Health, Albany, New York 12208, USA, [9]Department of Biomedical Sciences, School of Public Health, University at Albany, Albany, New York 12208, USA, [10]Department of Microbiology, Christian Medical College, Vellore, India.

[+]Current address: School of Natural Sciences, University of California Merced, CA 95343.

[$] Current address: Texas A&M Veterinary Medical Diagnostic Laboratory, Dept. of Animal Science, Texas A&M University, College Station, TX 77841, USA

[^]Current address: Center for Global Infectious Disease Research, Seattle Children's Research Institute, Seattle, USA

***** Corresponding authors

ABSTRACT

Tuberculosis (TB) patients suffer from progressive and debilitating loss of muscle mass and function, referred to as cachexia. Though a multifactorial condition, cachexia in cancer is promoted by systemic zinc redistribution and accumulation in muscles. Clinical studies with TB patients indeed show zinc dyshomeostasis. We therefore set out to understand mechanisms by which *Mycobacterium tuberculosis* (*Mtb*) govern zinc metallostasis at the host-pathogen interface. Here, we report a novel zinc metallophore from *Mtb* that restores zinc metabolic imbalance. These diisonitrile lipopeptides, named kupyaphores are transiently induced early-on during macrophage infection and also in infected mice lungs. Kupyaphores protects bacteria from host-mediated nutritional deprivation and intoxication. Kupyaphore *Mtb* mutant strain cannot mobilize zinc and shows reduced fitness in mice. Further, we characterize *Mtb* encoded isonitrile hydratase that could mediate intracellular zinc release through covalent modification of kupyaphores. Our studies could provide a molecular link between TB-induced altered zinc homeostasis and associated cachexia.

MAIN

Cachexia or wasting syndrome, marked by progressive and debilitating loss of muscle mass and function, is a characteristic complication observed in almost all active tuberculosis (TB) patients[1, 2]. Classically, muscle mass loss in TB patients was associated with poor nutrition since TB was considered to be a poor man's disease[3]. However, it is becoming evident that TB-associated cachexia cannot be reversed by conventional nutritional support[4, 5]. The progressive functional impairment culminates into deterioration of lung diaphragm and cardiac muscles leading to premature death. Despite being a key indicator of poor prognosis in TB patients, the aetiology of tuberculosis-associated cachexia has not been addressed. Understanding the causal mechanism of this syndrome in TB would provide opportunities for the development of novel interventions and also improve treatment response, quality of life, and survival.

Cancer-associated cachexia (CAC) studies suggest that multiple organs contribute to the wasting of skeletal muscle[6]. Tumour triggers a complex set of interactions including release of inflammatory cytokines and host metabolic changes that result in a major increase in resting energy expenditure, the presence of inflammation, and increased protein catabolism[7]. It is worthwhile to note that the known mediators of CAC such as TNF-α, IL-1β, IFN-γ, and IL-6 cytokines, are all often increased in the serum of TB patients[8]. More recently, the dysregulated organismal zinc distribution with reduced serum zinc levels and concomitant accumulation of zinc in tissues has emerged as a driving factor for muscular atrophy[9]. Interestingly, several clinical studies report zinc serum levels to be substantially low in active pulmonary TB subjects and that the administration of oral zinc did not provide any clinical benefits to the patients[10, 11, 12]. Whereas in cancer, cell intrinsic mechanisms would regulate zinc redistribution, the TB-perturbed zinc homeostasis can be directly attributed to *Mycobacterium tuberculosis* (*Mtb*) infection. We therefore set out to understand *Mtb*-specific strategies that may be able to regulate and impact zinc mobilization upon infection.

Mtb colonization of macrophages is known to initiate a competitive conflict for utilization of host nutrients, including elements such as iron, zinc, copper, magnesium and selenium[13]. By limiting transition metal availability to invading pathogens, the host is known to restrict pathogen proliferation, a physiological concept monikered 'nutritional immunity'[14]. Paradoxically, human macrophages have been shown to intoxicate intracellular *Mtb* with a burst of free zinc[15] and/or copper[16] through upregulation of zinc importers like Zip8[17]. Thus an

active and dynamic zinc flux changes are initiated in the host and bacteria upon infection which can have systemic ramifications.

The adaptive response strategy of bacterial pathogens to host-imposed zinc scarcity and poisoning is primarily governed by metal-sensing metalloregulatory proteins, metal transporters and pumps. As in most bacteria, *Mtb* contains two key zinc sensor proteins that co-ordinately function as uptake or efflux repressors. *Mtb* Zinc uptake repressor (Zur, Rv2359) regulate expression of several genes, including those involved in zinc uptake[18], while the efflux regulator ZntR (Rv3334) regulates transcription of genes encoding three types of exporters[19]. *Mtb* has been shown to utilize P1-type ATPases to neutralize the toxic effects of zinc in macrophages[14]. *Mtb* also employs Zn sparing to overcome zinc starvation, e.g. remodeling of the ribosome 70S subunit[20]. Similar mechanisms of metallostasis also govern the intracellular levels of iron during the host-pathogen interaction[21, 22]. However, the essential bacterial strategy of iron acquisition involving siderophore-mediated scavenging is yet to be recognized for other transition metals. We therefore attempted to identify mechanisms that will actively facilitate *Mtb* to acquire zinc during intracellular survival and in-turn modulate host zinc dynamics.

In this study, we identify a zinc-specific metallophore, which is produced on demand by *Mtb* to maintain bacterial fitness under varying zinc environments. These diacyl-diisonitrile lipopeptides, named kupyaphores are specifically induced during infection and can be move in and out of cells scavenging zinc. Further, we identify a novel isonitrile hydratase homologue in *Mtb* that is expressed in low zinc conditions and probably facilitates zinc release from kupyaphores. Detection of kupyaphores from infected mouse lungs and *Mtb* clinical strains propose pathophysiological relevance in zinc mobilization that could mediate cachectic metabolic end state.

RESULTS

Transcriptional responses maintains zinc homeostasis

To examine mycobacterial adaptive responses to conditions of varying zinc levels, we performed transcriptome analysis. The Middlebrook 7H9 medium used for culturing *Mtb* contains 6μM concentration of $ZnSO_4$ and we assessed *Mtb* growth kinetics at a log-order lower- and higher-zinc concentrations in the minimal Sauton's medium. *Mtb* growth curves at 0.1μM (low), 6μM (optimal) and 50μM (high) of zinc showed no significant differences [Fig 1a]. *Mtb* intracellular total zinc levels, as measured by inductively coupled plasma mass

spectrometry (ICP-MS), were also buffered at all three conditions [Fig 1b]. RNA was isolated from log phase cultures and transcriptome analysis was performed using Illumina MiSeq sequencing platform for three biological replicates and the quality of reads was assessed [Fig S1a]. Principal component analysis (PCA) with biological replicates showed good clustering, except for one from optimal conditions, which was not taken into account for further analysis. Interestingly, greater correlation was observed between transcriptional features of low and high zinc condition data sets, than with the optimal zinc condition [Fig 1c]. We compared differential expression analysis between 50μM and 0.1μM with cut off $\log_2$-fold change >0.5 and p-value <0.05 and visualized the data using MA plot [Fig 1d]. This analysis identified a small set of 39 genes to be differentially regulated. Interestingly, several of these genes have been previously known to be zinc (or metal) responsive genes. For example, upregulated genes included Zn transcription repressors *smtb* and *zur*[18] and efflux zinc pump *ctpC*[15], Rv2025c[23] and *cadI*[24]. The metal storage proteins *bfrB*[25], Rv0958 and *rpoB*[26] were downregulated. Circos plot analysis of 30 important operonic loci for all the three conditions of low-, high- and optimal-zinc levels were performed to understand co-regulated genes from *Mtb* [Fig 1e]. It is interesting to note dramatic increase in the levels of *ctpC* operon at high-zinc and downregulation at low-zinc. This is concordant with the established role of *ctpC* in zinc detoxification. Genes belonging to *esx-1* and *esx-3* export system show highest downregulation at both high- and low-zinc, when compared to optimal levels. Previous study had suggested role of these clusters in metal-dependent regulation[27]. We also observed regulation of biosynthetic clusters that are known to produce mycobactin and pthiocerol dimycocerosate. Careful examination of genes from the cryptic biosynthetic cluster Rv0097 to Rv0101 showed modest upregulation at both low- and high- zinc conditions. Since this cluster has been proposed to produce unknown metabolite, we assessed the induced expression of these genes by qRT-PCR analysis. This cluster indeed showed upregulation at both low and high conditions, when compared with optimum levels of zinc [Fig 1f]. We set out to understand the relevance of this cluster in zinc metallostasis.

Role of *Mtb* nrps in mycobacterial physiology

Rv0097-Rv0101 *Mtb* gene cluster spans 10.8 kbp, which over the years has been implicated in *Mtb* virulence, without exactly delineating their biochemical function[28, 29, 30]. The largest open reading frame (Rv0101) of Rv0097-Rv0101 gene cluster codes for a bimodular non-ribosomal peptide synthetase (NRPS). We generated a NRPS knockout strain in *Mtb* (Δ *nrps*) using phage-based transduction[31] to assess the role of this gene cluster in metal homeostasis [Fig

S2a]. The mutant strain showed no growth profile differences under planktonic conditions in the 7H9 medium [Fig S2b]. However, Δ*nrps Mtb* mutant exhibited differences under biofilm growth conditions as reported earlier[30] and formed a fragile film lacking the characteristic reticulation of WT *Mtb* biofilms. This mutant phenotype could be reversed by complementation with an integrative shuttle cosmid vector (containing an *Mtb* H37Rv cosmid fragment spanning Rv0096 to Rv0109) into the Δ*nrps* mutant [Fig S2c, d]. We then examined the growth profiles of WT *Mtb* and Δ*nrps* strains in the metal-limited chelated Sauton's medium. *Mtb* Δ*nrps* strain displayed defective growth in the Sauton's medium, which contains only glycerol and asparagine as sole carbon and nitrogen sources [Fig 2a]. The recovery of this growth defect was evaluated by addition of various divalent metals, one at a time, at 6μM concentration. Δ*nrps* growth could be restored only when zinc was supplemented in Sauton's medium and addition of copper, magnesium, manganese and iron could not reverse the phenotype [Fig 2b]. This mutant, however, exhibited compromised growth at low- [Fig 2c] and high- zinc [Fig 2d] levels, with almost no growth observed at 0.1μM. The total intracellular metal ion concentration measured by ICP-MS indeed exhibited major differences in the intracellular concentration for zinc ions for Δ*nrps*, as compared to WT *Mtb*. At low zinc, the Δ*nrps* mutant strain had lower intracellular zinc levels [Fig 2e], while at high zinc conditions mutant strain accumulated zinc [Fig 2f]. Δ*nrps Mtb* complemented strain reversed the phenotype observed to WT. No significant differences were observed for copper, magnesium, cobalt, nickel and manganese under these conditions for Δ*nrps* and WT *Mtb* [Fig S2e]. Together, the induction of Rv0097-Rv0101 biosynthetic pathway followed by altered intracellular zinc concentration provide evidence for the involvement of the unknown metabolite produced by this cryptic gene cluster in maintenance of bacterial zinc homeostasis.

Pathophysiological Relevance of *nrps* gene cluster in *Mtb* macrophage and murine infection

To establish the role of Rv0097-Rv0101 biosynthetic cluster in the context of *Mtb* pathogenesis, *ex-vivo* infection experiments were performed with both the primary murine bone marrow derived macrophages and immortalized murine bone marrow-derived RAW 246.7 macrophage cell line with WT and Δ*nrps Mtb* strains. Within 2 hours of infection remarkable up regulation of Rv0097-Rv0101 gene cluster could be noted in WT *Mtb* cells [Fig S3a]. This upregulation response dampens in 24 hours in the macrophage infection assays,

indicating a role for the metabolic end product of Rv0097-Rv0101 cluster in early reprogramming of host cells for sustained survival.

Further, the intracellular unbound zinc concentration for WT and Δ*nrps Mtb* were measured using FluoZin™-3AM (FZ3AM) indicator, a zinc-selective indicator. Infection of macrophages with WT *Mtb* results in the increase in punctate signals, as quantitated by confocal microscopy [Fig 3a]. Δ*nrps* mutant infected macrophage cells exhibit weak fluorescence signals for free zinc. We then compared the zinc pools of macrophages and bacteria by performing differential lysis of the two cells. The total zinc levels were measured by ICP-MS analysis. Uninfected macrophages at 2 and 24 hour time points maintain a steady levels of zinc. Macrophages infected with WT *Mtb* showed significant decrease in total zinc levels at 24 hours in comparison to 2 hour time point. Despite these changes in the host zinc levels, WT *Mtb* is able to maintain its intracellular levels. Mutant-infected macrophages exhibit no difference in zinc levels at either of these two time points as compared to uninfected macrophages [Fig 3b]. However, the intracellular zinc pools of Δ*nrps* mutant cells show reduced total zinc levels at both 2 and 24 hours [Fig 3c]. No significant difference could be observed for any other divalent metal cation in both host and bacterial lysates [Fig S3b, c]. Previous mice infection studies with this mutant (ΔNRPS) in immunocompetent and immunodeficient mice showed reduced lung pathology with significantly increased mice survival rates. This phenotype was attributed to early infection impairment (20 days) in the mutant *Mtb* strain. We carefully dissected the very early infection events (day 1 to day 21) and mutant strain showed severely compromised growth till week 3 post infection as compared to wild type [Fig 3d], highlighting the pivotal role of this gene cluster in successful establishment of TB infection.

Computational and Biochemical analysis of *Mtb nrps* biosynthetic gene cluster

To predict metabolite produced by the biosynthetic operon Rv0097-Rv0101, we carried out detailed *in silico* analysis and also confirmed biochemical functions for three enzymes by performing *in vitro* assays. Retro-biosynthetic analysis of the *Mtb* cluster postulates FAAL10 (Rv0099) to activate α/β unsaturated long chain fatty acids as acyl-adenylates and transfer them onto the thiol group of Ppant ACP (Rv0100) protein. This ACP-bound lipid chain is likely to be acted upon by Rv0097 and Rv0098. *Mtb* NRPS (Rv0101) contains two modules, each consisting of enzymatic domains-condensation (C), adenylation (A) and thiolation (T). The terminal domain of this NRPS protein contains a reductase (R) domain that releases thioester-

bound lipopeptides as corresponding alcohols[32] [Fig 4a]. It is interesting to note that while the A and T domains from the two modules of *Mtb* NRPS have high (~70%) sequence identity, the two C domains (C1 and C2) show weak sequence homology (similarity/identity- 35/36%) [Fig S4a]. All the species belonging to *Mycobacterium Tuberculosis Complex* (*MTB*C), *M. tuberculosis*, *M. bovis, M. canettii, M. pinnipedi, M. caprae, M. micorti and M. africanum* possess bimodular NRPS. While other pathogenic mycobacterial species like *M. marinum, M. leprae, M. abscessus,* and *M. immunogenum* have unimodular organization. This cluster is absent in non-pathogenic *Mycobacteria*. Dendrogram-based analysis of all these C domains of NRPS proteins (unimodular or bimodular) along with typical C domains from other classical NRPS proteins from *Actinomycetes* results in classification into two discrete clusters. All the C domains from NRPS corresponding to these biosynthetic loci forms a separate branch (we refer to these as C1*), while the C2 found in bimodular NRPS from *MTB*C group cluster with typical C domains [Fig 4b]. Though both the domains are of similar length and share the conserved HHXXXDG motif, structural modelling of C1* domains show unique conservation of amino acid residues around the active site pocket, suggestive of functional divergence from the cognate C domains [Fig S4b, c]. These structural models suggest substantial homology of C1* to CoA-dependent acyltransferases. Based on reported studies of metabolites isolated from the heterologous overexpression systems for *M. marinum*[33] and *S. thioluteus*[34], we reasoned that this unusual C1* domain may be involved in the transfer of acyl chains on both the α-amino and ε-amino groups of a lysine.

The four other genes from this cluster encode for an oxidoreductase (Rv0097), a thioesterase-like protein (Rv0098), a fatty acyl-AMP ligase (FAAL) (Rv0099) and an acyl carrier protein (ACP) (Rv0100). Our previous studies showed FAAL10 (Rv0099) to activate long-chain fatty acids as acyl-adenylates[35] and transfer them onto the thiol group of Ppant ACP (Rv0100) protein[32]. We confirmed similar activity for α/β unsaturated long chain fatty acids. This ACP-bound lipid chain is likely to be acted upon by Rv0097 and Rv0098. We cloned and expressed Rv0097 and Rv0098 in *E. coli* and purified the proteins using affinity chromatography. Enzymatic assays were then performed by using phosphopantetheine mimics, *N*-acetyl cysteamine (NAC) thioesters of α/β unsaturated fatty acids of different carbon lengths- C4, C8 and C12 [Fig S5a-k]. Enzymatic activities were confirmed by liquid chromatography coupled to mass spectrometry (LC-MS) assays. Addition of glycine to NAC-thioester of dodecenoic acid in the presence of enzyme Rv0098 identified a molecular ion peak at m/z of 374.2239 corresponding to Michael addition of glycine across unsaturated fatty acyl thioester. Tandem

mass spectrometric analysis of this ion resulted in fragment ions of m/z $[M-H]^-$= 254.176, 74.025 and 58.030 providing confidence to the structure [Fig 4c]. No activity was observed for NAC-thioester of crotonoic and minor products were obtained with octenoic acids. When the purified protein Rv0097 was included in above assays, a new peak corresponding to m/z of 326.2085 could be observed in high-resolution LC-MS. This molecular ion along with fragment ions corresponded to the isonitrile modification of the of C12 fatty acyl- SNAC moiety [Fig 4d]. This isonitrile modified C12 ACP-bound acyl chain could be condensed with first amino acid that is selected by the adenylation domain of the first module of the NRPS protein (C1*-A1-T1). Computational algorithms predict the adenylation A1 to activate lysine, while the second A2 domain shows specificity for phenylalanine[36, 37]. Finally, the acyl chain would be reductively released to the corresponding alcohol, thus producing a novel isonitrile lipopeptide.

Identification and Characterization of Diisonitrile Lipopetides from *Mtb*

In order to isolate these molecules from *Mtb*, we extracted metabolites from *Mtb* cells grown in Middlebrook 7H9 medium, Sauton's medium and also from the 35-day biofilms using ethyl acetate. These samples were analyzed by LC-MS using a recently developed method from our laboratory[38]. Analysis of *Mtb* biofilm extract detected a cluster of unique peaks with m/z 707.5350, 721.5506, 735.5663, 763.5967, 833.6758, 847.6915, 861.7071, which were absent in the *Mtb* planktonic cultures grown in Middlebrook 7H9 medium [Fig 5a]. The differences between parent ion masses (**I**-**VII**) were in the multiple of 14, which is typical of series with different $–CH_2$ group and can be attributed to different acyl chain lengths. The previously reported molecular formula $C_{40}H_{73}N_6O_5$ with m/z of 717.5612 was absent in our biofilm analyzed samples[30]. Tandem MS fragmentation patterns of parent ions revealed a common set of fragments with m/z values of $[M+H]^+$= 266.1863, 152.1080, 133.099, 89.1073 [Fig 5b]. The m/z of 266.1863 along with the fragments of m/z 152.1080 and 133.099 can be assigned to phenylalaninol ion and ornithine ion respectively. Peak of 89.1073 correspond to decarboxylation of the ornithine ion, which suggests backbone to be composed of a common core of ornithine and phenylalaninol dipeptide. Negative ion mode LC-MS analysis of the biofilm extract revealed two types of metabolites- symmetric (with identical acyl chains on the two $–NH_2$ groups of ornithine) and asymmetric (with different acyl chains on the two $–NH_2$ groups of ornithine) [Fig S6a, b]. To establish chemical structure, we treated these metabolites with formic acid. Previous studies had reported facile conversion of nitriles into amines on reaction with mild acid[39]. We therefore synthesized an analogue of C18 lipopeptide with amine

functionality at the β−position. Towards this end, protected β-amino fatty acid was coupled with *L*-ornithine and *L*-phenylalaninol sequentially, followed by deprotections resulting in desired lipopeptide [Fig S7a-k]. Addition of formic acid to the biofilm metabolite extract resulted in disappearance of m/z for $[M+H]^+$ ion 848.6987 (**VI**) having retention time of 36 minutes and a new peak could be observed at m/z for $[M+H]^+$ ion 828.73 with retention time of 34 minutes. This new peak corresponded to the chemical synthetic standard compound [Fig 5c]. Additionally, both the parent peak and synthetic standard had identical MS/MS fragments [Fig S5l]. We have thus identified and characterized series of diacyl diisontirile lipopeptides from *Mtb* that consists of a dipeptide core of ornithine and phenylalaninol, where the amino-groups are acylated with isonitrile-containing fatty acyl chains ranging from C_{13} to C_{19}. Based on their metal acquisition functionality, we propose to name these *Mtb* diisonitrile lipopeptides as -'kupyaphores' – ('Kupya' in Sanskrit refers to rare metals and 'phores' means carrier) and this gene cluster could be recognized as *kupya A-E*.

Characterization of Kupyaphores in Infected Mouse Lungs and Human Clinical Strains

Finally, we set out to understand the temporal regulation of this zinc acquisition machinery in the mouse infection model. After aerosol challenge in Balb/C mice with WT *Mtb*, lungs were harvested from these mice on day 1, and weeks 1, 2, 4 and 6. We performed CFU analysis with part of lung tissue to confirm infection [Fig S8a] and the other portion of the lung tissue were utilized for extraction of metabolites using ethyl acetate. LC-MS analysis performed with tissue organic extracts showed clear signals for kupyaphores during week 1. These metabolite peaks were not there on day 1, which diminished dramatically at week 2 and disappeared by week 4. Two series of kupyaphores could be identified based on their masses, one symmetrical corresponding to two C18 acyl chains of this lipopeptide and other one asymmetrical, both modified with isonitrile groups [Fig 5d]. Semi-quantitative measurement using multiple reaction monitoring (MRM) LC-MS, shows that kupyaphores are indeed an early response of *Mtb* to the host challenge and these metabolites disappear as the host-pathogen establish a more mutually co-existential harmony.

We analyzed early passage clinical strains obtained from TB subjects for the expression of kupyaphores. In all the four clinical *Mtb* strains, we could detect robust expression of several of the metallophores could be detected [Fig S8b]. Our studies demonstrate significance of kupyaphores in maintaining *Mtb* zinc homeostasis.

Putative mechanism of Zinc release from kupyaphore by *Mtb* isonitrile hydratase

Zinc acquisition will involve translocation of zinc-kupyaphore complex across cell envelope, analogous to iron siderophores. *Mtb* cells grown in Sauton's medium showed very low levels of **I** to **VII** metabolites, while substantial enrichment was observed in the media supernatant under both low and high zinc conditions [Fig 6a]. Dimeric kupyaphore zinc bound metabolites of m/z $[M+H]^+$=1507.48 could be detected only in supernatant of WT *Mtb* grown with zinc supplementation and were absent in cultures grown in chelated Sauton's media. This suggested that these metabolites are efficiently exported out of the cells and may be involved in zinc acquisition under low zinc and zinc neutralization under high zinc conditions to restore metal metabolic imbalance.

Since zinc is redox-inert metal, the release of zinc from metallophores cannot follow the classical reductive release mechanism known for iron-siderophore. In the case of enterobactin, siderophore hydrolase is known to modify the scaffold mediating iron release[40]. We therefore analyzed the *Mtb* genome to investigate the presence of putative isonitrile modifying enzyme. This enzyme activity (InhA) was first reported from *Pseudomonas putida*[41] and was recently characterized (SfaF) from SF2768 biosynthetic cluster of *S. thioluteus*[42]. Protein sequence analysis of *Mtb* H37Rv genome coding sequence with InhA identified Rv0052 with an E value score of $5e^{-31}$. Careful analysis of Rv0052 sequence with the three-dimensional structure (PDB 3NON) indicated the absence of an important catalytic site (D17) from the annotated Rv0052. The gene Rv0052 is located between 57410 bp to 57973 bp in the genome of H37Rv. Analysis of upstream sequences revealed another start site 57 bp upstream in-frame with annotated Rv0052. Protein coding sequence of this new open reading frame (ORF) shows conservation of the catalytic residue D17. This protein was conserved across all *MTB*C strains and was absent from non-pathogenic mycobacterial sequences, where kupyaphore cluster is absent [Fig S9a]. To examine biochemical function, we cloned and expressed Rv0052 in BL21-DE3. Protein was purified and enzymatic assays were performed along with Rv0098 & Rv0097 [Fig 6b]. A clear peak corresponding to m/z of 243.176 for N-formamide product could be observed only in the reactions supplemented with Rv0052. MS/MS fragmentation in the negative ion mode also further provided confidence to the hydration of isonitrile unit [Fig 6c]. This modification of isonitrile is likely to reduce the Lewis base character and thus will diminish formation of stable coordinate complex bond with zinc. Interestingly, Rv0052 expression was found to be significantly upregulated by qRT-PCR only in low zinc condition and showed no induction under high zinc condition [Fig 6d]. We thus propose that this could be a putative mechanism of zinc release from kupyaphores under low-zinc conditions.

DISCUSSION

Given that there is virtually no free zinc in the cell and that this micronutrient cannot be produced de-novo, it becomes mandatory for pathogens to acquire these from the host pools. Bacteria have thus evolved rather intricate molecular mechanisms of metal ion sensing, uptake, efflux and allocation to maintain homeostasis[43]. We here decipher a novel mechanism of zinc acquisition whereby *Mtb* secretes zinc-specific metallophores on demand to support its intracellular survival. While siderophore-mediated iron acquisition is well established in several bacterial and fungal systems[44, 45], metabolites that selectively chelate and transport zinc have not been characterized thus far. As pathogen proliferate and establish infection, the demand of zinc continue to accrue and micronutrients like zinc becomes a shared resource for both host and pathogen. This competitive conflict culminates into an altered host zinc homeostasis[46, 47, 48]. In addition, the inflammatory response of host during *Mtb* pathogenesis also mediate redistribution of metabolic nutrients. Together, these events could play a pivotal role in the cachexia development, as recognized in cancer.

Our studies show that diisonitrile lipopeptides, named kupyaphores produced by *Mtb* are critical to maintain zinc metallostasis. The chemical features of kupyaphores are typical of *Mtb* metabolic repository. Long-fatty acyl chains facilitates these molecules to transverse complex *Mtb* cell envelope. The electron rich isonitrile functionality forms coordinate bond and the affinity for chelation to zinc metal ions is probably governed by the structural architecture. In the low zinc conditions, zinc-kupyaphore complex is shuttled into mycobacterial cells by using transporters that are yet to be characterized. The intracellular release of zinc ions may be mediated by *Mtb*-coded isonitrile hydratase (Rv0052) that modifies isonitriles to corresponding formamides. It is possible that kupyaphores are then recycled back into the system by regenerating the active isonitrile groups. Recently, opine-type metallophores staphylopine from *Staphylococcus aureus*[49] and pseudopaline from *pseudomonas aeruginosa*[50] that chelate broad range of metal ions including zinc have been reported. This family of metallophores have been referred to as zincophore[51] despite the fact that these metabolites possess a wide range of metal ion specificity.

Kupyaphores show strict specificity for zinc and *kupya* biosynthetic cluster is transcriptionally activated during both the low- and high-levels of zinc. We therefore propose that kupyaphores may be potentially playing a role in both zinc acquisition as well as quenching of toxic zinc levels [Fig 6e]. Secretion of kupyaphores under limiting zinc conditions would allow *Mtb* to

scavenge zinc from environment with concomitant intracellular release through isonitrile hydratase. While detailed mechanism needs to be worked out, based on RNAseq data we propose that classical zinc suppressor Zur may be involved in this activation during the low zinc conditions. The expression of *zur* mRNA are substantially higher when *Mtb* is grown at high zinc levels as compared to low levels. Indeed, classical Zur-binding sites can be mapped (61bp upstream of Rv0096 and 354bp upstream of Rv0099) in the promoter region of *kupya* biosynthetic cluster. As previously recognized, binding of zinc to Zur stabilizes the formation of dimers that binds strongly to these binding sites, thus suppressing the genes in their respective regulons. During depleting zinc conditions, Zur falls off from these sites and promotes expression of kupya proteins. It is not clear how higher toxic levels of zinc also induce expression of this biosynthetic pathway. Expression of kupyaphores could chelate extracellular toxic zinc levels thus contributing to metal detoxification by reducing the concentration of free metal in the cytosol.

The *kupya* biosynthetic cluster is conserved only in pathogenic mycobacteria. This new family of diisonitrile-containing metabolites, SF2768 and isonitrile lipopeptides (INLP), have been characterized from *Streptomyces thioluteus*[34] and *M. marinum*[33] by heterologous over-expression of cognate gene clusters in *S. lividans* and *E. coli* respectively. SF2768 from S. thioluteus has been shown to chelate and transport copper in the *in vitro* studies. In both these instances diisonitrile metabolites have not been isolated from their natural hosts and thus their functional significance in the microbial physiology remains obscure. Kupyophore lipopeptide is biosynthesized by genomic cluster of 5 genes, Rv0097-Rv0101. The non-ribosomal peptide synthetases produces the dipeptide core consisting of ornithine and phenylalanine, which is N-acylated with long chain isonitrile-fatty acyl. It is rather peculiar that the identification of these zinc chelating metallophores took this long, the iron siderophores mycobactins of *Mtb* were discovered in 1965[52]. Our studies show strict controlled inducible expression of *kupya* biosynthetic cluster such that these diisonitrile lipopeptides are produced transiently during the early phase of infection.

Micronutrients play a central part in metabolism and in the maintenance of tissue functions[53]. Cachexia, which is a hallmark of pulmonary TB, unfortunately does not manifest in murine and guinea pig TB infection models. With the discovery of kupyaphores as an important *Mtb* zinc-specific metallophores, it may be worthwhile to evaluate the role of these metabolites in appropriate animal models to understand systemic zinc mobilization and redistribution that may be associated with cachexia during TB pathogenesis.

METHODS

Bacterial culture

Mycobacterium tuberculosis H37Rv were grown in Middlebrook 7H9 culture medium (BD Difco, Cat# 271310) supplemented with 10% oleic acid-albumin-dextrose-catalase (BD, Cat# 211886), 0.05% Tween-80 (Amresco, Cat# 0442) at 37°C under shaking conditions or on Middlebrook 7H11 agar (BD Difco, Cat# 283810) supplemented with 10% OADC unless stated.

For metal studies, *Mtb* strains were grown in Sauton's media (HiChem Life Sciences, Cat# M1276) prepared in double distilled water, previously treated for 12 hours at 4°C with 10g/L of chelex 100 resin (Sigma Aldrich, Cat# C7901), filtered and autoclaved. Log phase grown culture were pelleted, washed thrice with 1X PBS (HiChem Life Sciences, Cat# TL1101) and re-suspended in chelated Sauton's media. Growth was then monitored after supplementation with specific concentration of metal ions (Sigma Aldrich) and 0.025% tyloxapol (Sigma Aldrich, Cat# 25301). All glassware used was treated with 0.01% EDTA (Sigma Aldrich) overnight and rinsed with 0.2N HCL (Merck, Cat# 100317) next day and subsequently washed with double distilled water before using.

Mtb biofilms were grown in Sauton's medium (without tween 80 or tyloxapol) by incubation without shaking at 37°C for 5 weeks under humidified conditions. The dishes were wrapped with parafilm during incubation.

RNA-seq library preparation

Total DNA-free RNA sample was depleted of bacteria rRNA with Ambion's MICROBExpress kit (AM1905) per the manufacturer's instructions. Total bacterial RNA ~1ug 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, Santa Clara, CA, USA). 300x2bp 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-seq data analysis

Mycobacterium tuberculosis genome and its annotations (gff/fna/ptt/rnt) were downloaded using an R package, Progenome (v0.0.7) [54] and the gist used for extraction is in the following reference [55]. Further, raw reads were quality trimmed using Trim-Galore (v0.6.4)[56] and the resulting trimmed reads were

aligned using Salmon (v0.14.1) [57] and estimated differential gene expression using DESeq2 (v1.26.0) [58] 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) [59] and used ComplexHeatmap (v2.2.0)[60] for plotting the expression of genes in operons.

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 HMMs. 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.

Biochemical assays of Rv0097, Rv0098 and Rv0052

Recombinant Rv0097, Rv0098 and Rv0052 were expressed and purified as detailed in the SI Methods. Enzymatic reactions were set up at pH-7.5 (Tris-HCl) (Sigma Aldrich, Cat# 648317) with 200mM glycine (Sigma Aldrich, Cat# G7126), 5mM (C12-NAC) substrate and 50μM enzyme- Rv0098 or 400μM iron sulphate (Sigma Aldrich, Cat# V000119) and 500μM alpha-ketoglutarate (Sigma Aldrich, Cat# 75890) and 4mM ascorbate (Sigma Aldrich, Cat# 134032) with 50uM 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, 50μM 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

Middlebrook 7H9 grown planktonic and biofilm cultures of WT *Mtb* and Δ*nrps* mutant were harvested and weighed. Low-molecular-weight molecules were then extracted by adding five times the volume of MS grade ethyl acetate (Honeywell, Cat# 319902) and left for 24-48 hours on a stirrer. The organic layer was then separated by centrifugation at 3500rpm for 5 minutes and evaporated to dryness. The residual material was dissolved in a minimal volume of methanol and analyzed by LC-MS.

WT *Mtb* strain was also utilized for metabolite extraction from planktonic cultures of chelated Sauton's media grown bacteria. Briefly, 1L of chelated Sauton's media grown culture was pelleted down at 3500rpm for 15 mins. Supernatant and pellet were separately collected. To cell pellet, five times volume of MS grade ethyl acetate was added whereas equal volume of MS grade ethyl acetate was added to the collected supernatant. Samples were left for stirring for 24-48 hours. The organic layer was then separated and evaporated to dryness, and the residual material was dissolved in minimal volume of methanol(Honeywell, Cat# 14262) 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 using a previously reported method[38]. 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 **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.

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) (Sigma Aldrich, Cat# 64186) 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.

Determination of total metal ions in *invitro* bacterial cultures

Total metal concentrations were measured by ICP-MS. Briefly 0.8-1.0 OD WT *Mtb* and *nrps* KO cultures grown in 7H9-OADC were pelleted and washed with 1X PBS 3 times and grown in chelated Sauton's media with 0.025% tyloxapol supplemented with 0.1μM, 6μMand 50μM $ZnSO_4$ (Sigma Aldrich, Cat# 221376) for 1 generation time. Subsequently cultures were pelleted and subjected to lysis by boiling in 0.1% SDS (Sigma Aldrich, Cat# L3771) and 0.2% Triton-X 100 (Sigma Aldrich, Cat# T8787) for 15 minutes in trace-element free 1.5ml micro-centrifuge tubes. Volume were made up to 1ml using MS grade water (Honeywell, Cat# 39253) and samples were filtered through 0.2μ PTFE

syringe filters before running on ICP-MS (Thermo Xcaliber II). Absolute ppb counts were normalized to protein concentration estimated for each sample and are plotted.

Gene expression analysis by RT-PCR

Using NucleoSpin RNA kit (Macherey-Nagel, Cat# 740955), total RNA from desired cultures of the indicated strains of *Mtb* were isolated from 10 mL cultures obtained from mid-exponential growth phase in 7H9-OADC, Sauton's medium or biofilm. Using the PrimeScript First Strand cDNA Synthesis Kit (Takara Bio, Cat# 6110A), cDNA was generated from 1μg 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, Cat# A25742) 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.

Cell culture and infection

The murine macrophage cell line RAW264.7 or primary Bone-marrow derived macrophages were infected with WT and *nrps* bacteria at a multiplicity of infection (MOI) of 5 bacteria per cell (for CFU, imaging and ICP-MS studies) 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 media for indicated time points. Details for culture maintenance provided in Supplementary Methods Information.

Determination of free and total zinc in macrophage infection studies

Macrophage cultures were washed with PBS and incubated with 1μM Fluozin 3-AM (Life Technologies, Cat# F24195) for 1hour at 37°C. After two washes with PBS, cells were fixed with 4% formaldehyde in 1× PBS. Fixed cells were then mounted with DAPI (Life Technologies, Cat#S36938). DAPI-positive and FZ3 green-positive cells were counted manually after taking images of five representative fields of each of the triplicate samples for each time point by confocal microscopy using Leica SP8 system.

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 4000rpm 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% Triton-X 100 for 15 minutes in trace-element free 1.5ml micro-centrifuge tubes. Volume were made up to 1ml 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.

Mtb metabolite isolation upon murine TB infection

For *Mtb* metabolite extraction from mice, 0.1g 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 0.1μ zirconium bead in a bead beater followed by centrifugation at 13000g for 5 minutes. 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. Details of mice infection studies with WT *Mtb* are provided in Supplementary Information.

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.

References

1 Eddleston M, Davidson R, Brent A, Wilkinson R (2008) Oxford Handbook of Tropical Medicine, 3rd Edition. Oxford, UK: Oxford University Press. 856 p.

2 Fauci AS, Braunwald E, Kasper DL, Hauser SL (2009) Harrison's Principles of Internal Medicine, 17th Edition. New York: McGraw-Hill Professional. 2958 p.

3 Sacks, L. V. and Pendle, S. (1998) Arch Intern Med, 158, 1916-22.

4 Chang, S. W. et al., Gut hormones, appetite suppression and cachexia in patients with pulmonary TB. PLoS One 8 (1), e54564.

5 Gullett, N. P., Mazurak, V. C., Hebbar, G., and Ziegler, T. R., Nutritional interventions for cancer-induced cachexia. Curr Probl Cancer 35 (2), 58.

6 Dhanapal, R., Saraswathi, T., and Govind, R. N., Cancer cachexia. J Oral Maxillofac Pathol 15 (3), 257.

7 Bennani-Baiti, N. and Davis, M. P., Cytokines and cancer anorexia cachexia syndrome. Am J Hosp Palliat Care 25 (5), 407 (2008).

8 Stallings, C. L., Host response: Inflammation promotes TB growth. Nat Microbiol 2, 17102.

9 Wang, G. et al., Metastatic cancers promote cachexia through ZIP14 upregulation in skeletal muscle. Nat Med 24 (6), 770.

10 Ali, W. et al., Serum zinc levels and its association with vitamin A levels among tuberculosis patients. J Nat Sci Biol Med 5 (1), 130.

11 Ghulam, H. et al., Status of zinc in pulmonary tuberculosis. J Infect Dev Ctries 3 (5), 365 (2009).

12 Boloorsaz, M. R. et al., Impact of anti-tuberculosis therapy on plasma zinc status in childhood tuberculosis. East Mediterr Health J 13 (5), 1078 (2007).

13 Marcela Rodriguez, G. and Neyrolles, O., Metallobiology of Tuberculosis. Microbiol Spectr 2 (3).

14 Hood, M. I. and Skaar, E. P., Nutritional immunity: transition metals at the pathogen-host interface. Nat Rev Microbiol 10 (8), 525.

15 Botella, H. et al., Mycobacterial p(1)-type ATPases mediate resistance to zinc poisoning in human macrophages. Cell Host Microbe 10 (3), 248.

16 Wolschendorf, F. et al., Copper resistance is essential for virulence of Mycobacterium tuberculosis. Proc Natl Acad Sci U S A 108 (4), 1621.

17 Pyle, C. J. et al., Elemental Ingredients in the Macrophage Cocktail: Role of ZIP8 in Host Response to Mycobacterium tuberculosis. Int J Mol Sci 18 (11).

18 Goethe, E. et al., Critical Role of Zur and SmtB in Zinc Homeostasis of Mycobacterium smegmatis. mSystems 5 (2).

19 Riccardi, G., Milano, A., Pasca, M. R., and Nies, D. H., Genomic analysis of zinc homeostasis in Mycobacterium tuberculosis. FEMS Microbiol Lett 287 (1), 1 (2008).

20 Li, Y. et al., Zinc depletion induces ribosome hibernation in mycobacteria. Proc Natl Acad Sci U S A 115 (32), 8191.

21 Chao, A., Sieminski, P. J., Owens, C. P., and Goulding, C. W., Iron Acquisition in Mycobacterium tuberculosis. Chem Rev 119 (2), 1193.

22 Krithika, R. et al., A genetic locus required for iron acquisition in Mycobacterium tuberculosis. Proc Natl Acad Sci U S A 103 (7), 2069 (2006).

23 Campbell, D. R. et al., Mycobacterial cells have dual nickel-cobalt sensors: sequence relationships and metal sites of metal-responsive repressors are not congruent. J Biol Chem 282 (44), 32298 (2007).

24 Festa, R. A. et al., A novel copper-responsive regulon in Mycobacterium tuberculosis. Mol Microbiol 79 (1), 133.

25 Pandey, R. and Rodriguez, G. M., A ferritin mutant of Mycobacterium tuberculosis is highly susceptible to killing by antibiotics and is unable to establish a chronic infection in mice. Infect Immun 80 (10), 3650

26 Eckelt, E. et al., Identification of a lineage specific zinc responsive genomic island in Mycobacterium avium ssp. paratuberculosis. BMC Genomics 15, 1076.

27 Tinaztepe, E. et al., Role of Metal-Dependent Regulation of ESX-3 Secretion in Intracellular Survival of Mycobacterium tuberculosis. Infect Immun 84 (8), 2255.

28 Giovannini, D. et al., A new Mycobacterium tuberculosis smooth colony reduces growth inside human macrophages and represses PDIM Operon gene expression. Does an heterogeneous population exist in intracellular mycobacteria? Microb Pathog 53 (3-4), 135.

29 Bhatt, K. et al., A Nonribosomal Peptide Synthase Gene Driving Virulence in Mycobacterium tuberculosis. mSphere 3 (5).

30 Richards, J. P. et al., Adaptation of Mycobacterium tuberculosis to Biofilm Growth Is Genetically Linked to Drug Tolerance. Antimicrob Agents Chemother 63 (11).

31 Parish, T. and Stoker, N. G., Use of a flexible cassette method to generate a double unmarked Mycobacterium tuberculosis tlyA plcABC mutant by gene replacement. Microbiology 146 (Pt 8), 1969 (2000).

32 Chhabra, A. et al., Nonprocessive [2 + 2]e- off-loading reductase domains from mycobacterial nonribosomal peptide synthetases. Proc Natl Acad Sci U S A 109 (15), 5681.

33 Harris, N. C. et al., Biosynthesis of isonitrile lipopeptides by conserved nonribosomal peptide synthetase gene clusters in Actinobacteria. Proc Natl Acad Sci U S A 114 (27), 7025.

34 Wang, L. et al., Diisonitrile Natural Product SF2768 Functions As a Chalkophore That Mediates Copper Acquisition in Streptomyces thioluteus. ACS Chem Biol 12 (12), 3067.

35 Arora, P. et al., Mechanistic and functional insights into fatty acid activation in Mycobacterium tuberculosis. Nat Chem Biol 5 (3), 166 (2009).

36 Ansari, M. Z., Yadav, G., Gokhale, R. S., and Mohanty, D., NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases. Nucleic Acids Res 32 (Web Server issue), W405 (2004).

37 Bachmann, B. O. and Ravel, J., Chapter 8. Methods for in silico prediction of microbial polyketide and

38 Kelkar, D. S. et al., A chemical-genetic screen identifies ABHD12 as an oxidized-phosphatidylserine lipase. Nat Chem Biol 15 (2), 169.

39 Ames, B. D. et al., Crystal structure and biochemical studies of the trans-acting polyketide enoyl reductase LovC from lovastatin biosynthesis. Proc Natl Acad Sci U S A 109 (28), 11144.

40 Larsen, N. A. et al., Structural characterization of enterobactin hydrolase IroE. Biochemistry 45 (34), 10184 (2006).

41 Goda, M. et al., Isonitrile hydratase from Pseudomonas putida N19-2. Cloning, sequencing, gene expression, and identification of its active acid residue. J Biol Chem 277 (48), 45860 (2002).

42 Zhu, M. et al., Tandem Hydration of Diisonitriles Triggered by Isonitrile Hydratase in Streptomyces thioluteus. Org Lett 20 (12), 3562.

43 Lopez, C. A. and Skaar, E. P., The Impact of Dietary Transition Metals on Host-Bacterial Interactions. Cell Host Microbe 23 (6), 737.

44 Chaguza, C., Bacterial survival: evolve and adapt or perish. Nat Rev Microbiol 18 (1), 5.

45 Kramer, J., Ozkaya, O., and Kummerli, R., Bacterial siderophores in community and host interactions. Nat Rev Microbiol 18 (3), 152.

46 Sheldon, J. R. and Skaar, E. P., Metals as phagocyte antimicrobial effectors. Curr Opin Immunol 60, 1.

47 Subramanian Vignesh, K. and Deepe, G. S., Jr., Immunological orchestration of zinc homeostasis: The battle between host mechanisms and pathogen defenses. Arch Biochem Biophys 611, 66.

48 Gammoh, N. Z. and Rink, L., Zinc in Infection and Inflammation. Nutrients 9 (6).

49 Song, L. et al., Mechanistic insights into staphylopine-mediated metal acquisition. Proc Natl Acad Sci U S A 115 (15), 3942.

50 Lhospice, S. et al., Pseudomonas aeruginosa zinc uptake in chelating environment is primarily mediated by the metallophore pseudopaline. Sci Rep 7 (1), 17132.

51 Morey, J. R. and Kehl-Fie, T. E., Bioinformatic Mapping of Opine-Like Zincophore Biosynthesis in Bacteria. mSystems 5 (4).

52 Snow, G. A., Isolation and structure of mycobactin T, a growth factor from Mycobacterium tuberculosis. Biochem J 97 (1), 166 (1965).

53 Motamed, S. et al., Micronutrient intake and the presence of the metabolic syndrome. N Am J Med Sci 5 (6), 377.

54 Niu,Y. Github: ProGenome.

55 Extract gff/fna/ptt/rnt annotation of Mycobacterium tuberculosis - https://gist.github.com/viv3kanand/c26615f96975f7c71b8510fb4ac88826.

56 Krueger,F. (2016) Trim Galore. Babraham Bioinforma.

57 Patro,R., Duggal,G., Love,M.I., Irizarry,R.A. and Kingsford,C. (2017) Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods, 10.1038/nmeth.4197.

58 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.

59 Tjaden,B. (2020) A computational system for identifying operons based on RNA- seq data. Methods, 10.1016/j.ymeth.2019.03.026.

60 Gu,Z., Eils,R. and Schlesner,M. (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics, 10.1093/bioinformatics/btw313.

Acknowledgements

R.S.G acknowledges support from J.C. Bose fellowship and Department of Biotechnology (DBT) [grant number BT/PR20085/MED/291213/2017]. This work was supported by the DBT/Wellcome Trust India Alliance [grant number IA/I/15/2/502058] to S.S.K., and a Department of Science & Technology – Funds for Improvement of S&T Infrastructure Development (DST-FIST) grant to Department of Biology, IISER Pune. We thank Dr. Apoorva Bhatt for providing us with *Mtb* H37Rv cosmid spanning Rv0096-Rv0109 for our complementation studies. We thank the Tuberculosis Aerosol Challenge Facility (DBT-TACF) staff at International Centre for Genetic Engineering and Biotechnology (New Delhi, India) for their kind help. We also thank DBT for institutional support provided to NII.

Contributions

K.M, T.N.V, S.S.K, R.S.G conceptualized, designed and coordinated this study. S.S.K and S.S (IISER) carried out mass spectrometry studies and analysis. S.S (NII), R.S.B and A.K performed biochemical studies. R.C, D.P.M, D.S.R were involved with all aspects of chemical synthesis. B.D.D and D.M performed computational analysis of condensation domains. N.K and G.R.M supported ICP-MS studies. T.P generated the Mtb KO strain used in the study. P.A, V.A, M.F performed RNA seq and data analysis. J.S.M provided clinical strains from TB subjects. A.O provided early scientific inputs in this study. K.M and R.S.G co-wrote the paper and all authors read and approved the final manuscript.

Ethics Declaration

All mouse studies described in this paper received formal approval from the National Institute of Immunology – Institutional Animal Ethics Committee (NII- IAEC 440/17) following the guidelines outlined by the Committee for the Purpose of Control and Supervision of Experiments on Animals (CPCSEA), Government of India. Studies on clinical strains isolated from TB subjects described in this paper has approval from Christian Medical College, Vellore and CSIR-Institute of Genomics and Integrative Biology.

Code and Data availability

All codes used for data collection are submitted in github (https://github.com/viv3kanand/MTU-Manuscript/blob/main/Operon/ann/Extract_ann.R). Detailed codes used for the analysis are submitted in github repository (https://github.com/viv3kanand/MTU-Manuscript/). The datasets generated during and/or analysed during the current study can be accessed using the NCBI BioProject ID PRJNA701877 .

Competing Interests

The authors declare no competing interests.

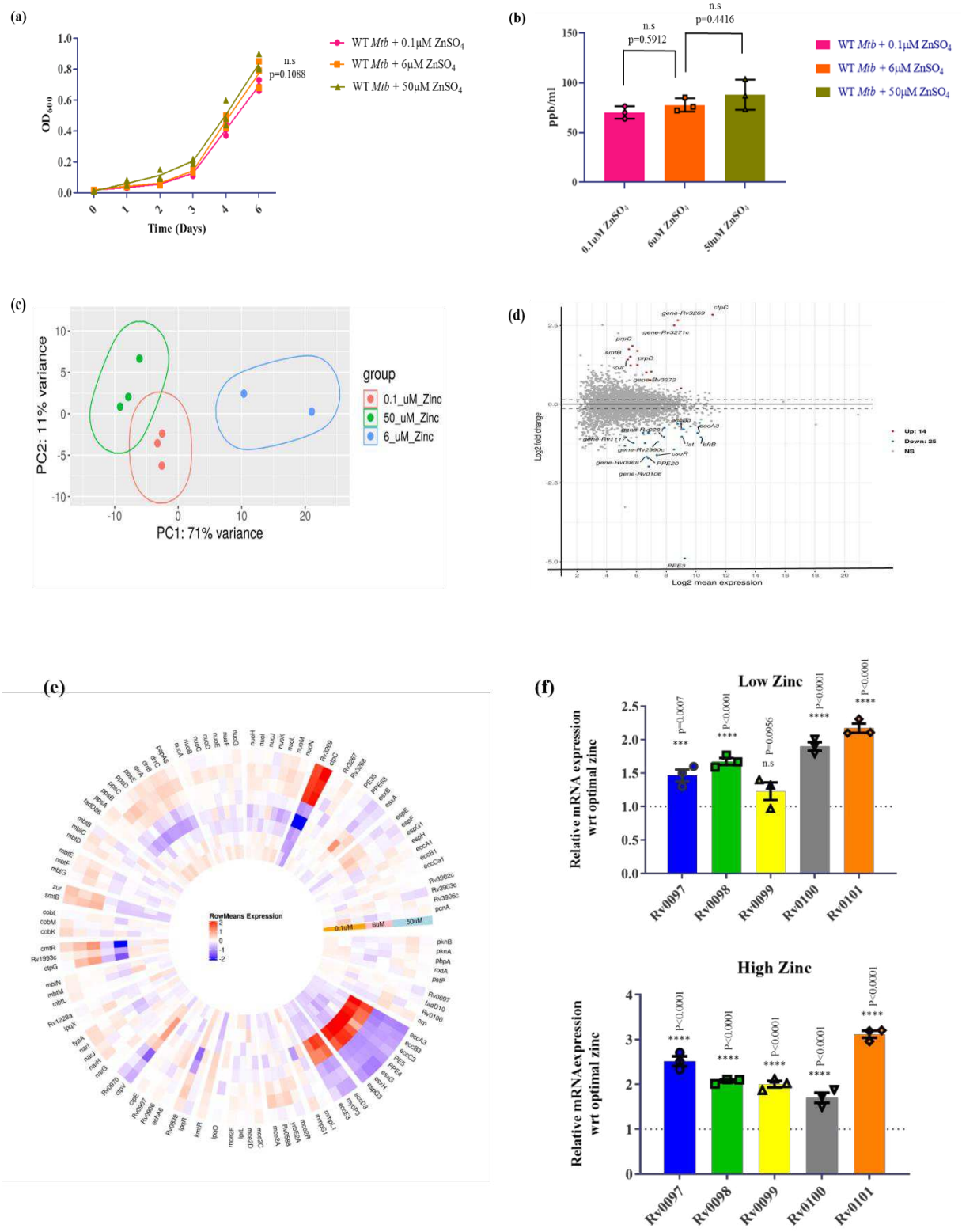


Figure 1:Remodeling of *Mtb* transcriptional responses maintains zinc homeostasis (a) Growth kinetics of WT *Mtb* grown in chelated Sauton's media supplemented with varying zinc concentrations – 0.1μM, 6μM and 50μM. p-values are determined by two-way ANOVA with Bonferroni post-tests relative to 6μM zinc supplemented condition. **(b)** ICP-MS analysis of total intracellular zinc content of *Mtb* grown in these three conditions. p-values are determined by one-way ANOVA with Dunnett post-

tests relative to 6μM zinc supplemented condition. **(c)** 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. **(d)** MA plot of differential gene expression analysis in 50μM zinc supplemented condition as compared to 0.1μM zinc supplementation with *p*-value <0.05 and log2 fold change>0.5 as cut-off for significance. Genes marked in red are significantly upregulated, while genes marked in blue are significantly downregulated. **(e)** Circos plot for differential expression of few metal responsive *Mtb* operons in three different zinc conditions. **(f)** Gene expression analysis of Rv0097-Rv0101 through qRT-PCR under low and high zinc conditions as compared to optimal zinc concentration. p-values are determined by one-way ANOVA with Dunnett post-tests relative to 6μM zinc supplemented condition. Data represents mean ± s.e.m (n=3 biological replicates, averaged for 2 technical replicates for each biological replicate).

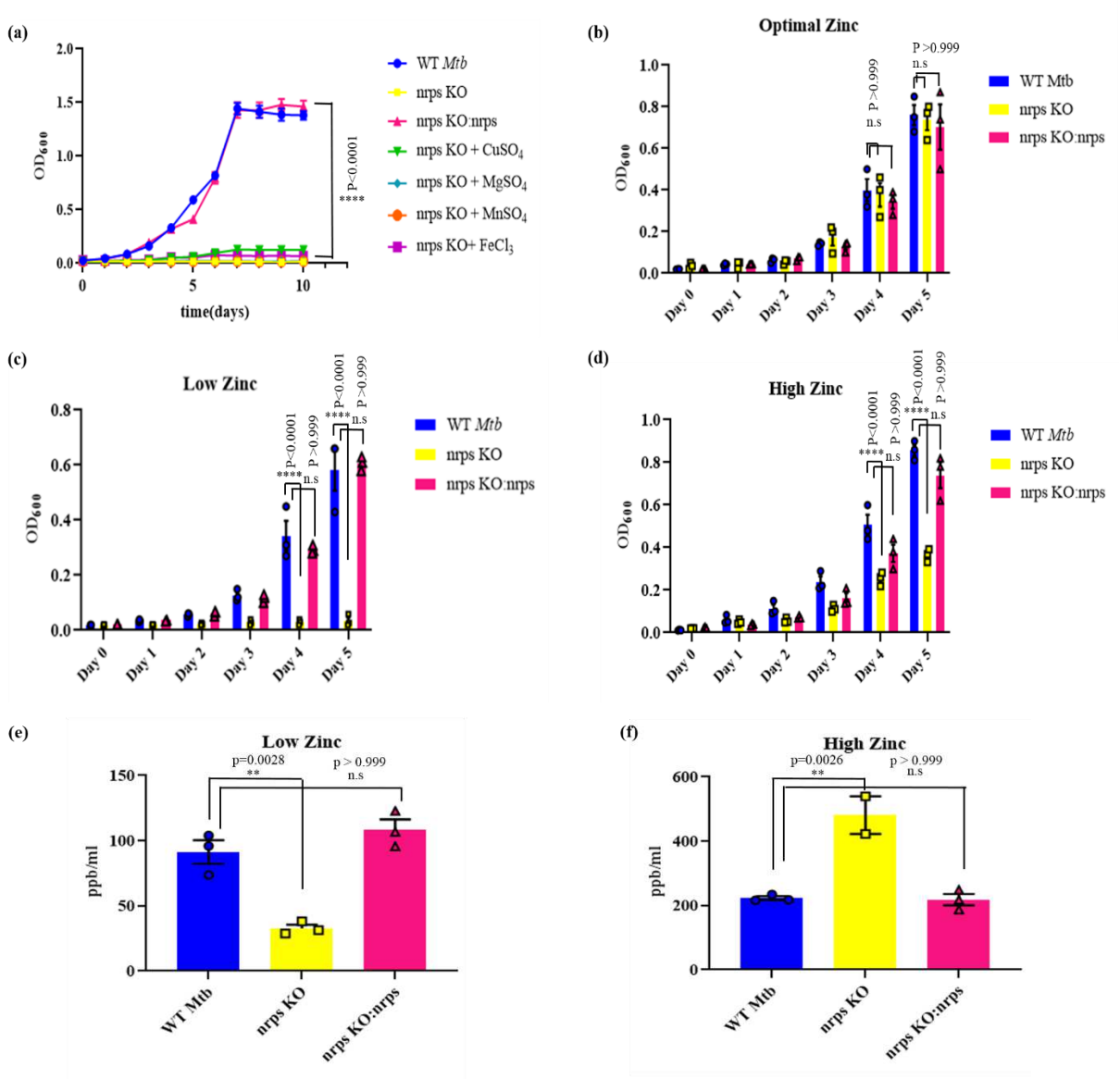


Figure 2: Role of *Mtb nrps* in mycobacterial physiology (**a**) Growth kinetics of wild type (WT) and *Mtb* strain knocked out for *nrps* show mutant to be severely compromised in nutrient deficient condition of chelated Sauton's Media. Addition of other key transition elements - copper, magnesium manganese or iron to Sauton's media failed to revive mutant significantly. p-values are determined by two-way ANOVA with Bonferroni post-tests relative to WT *Mtb* growth kinetics. (**b**) Only supplementation with 6µM zinc could restore growth profile of *Mtb nrps* knockout (KO) strain. (**c**), (**d**) Titration of Sauton's media with different concentrations of zinc show compromised growth pattern of mutant under both low (0.1µM) and high (50µM) zinc concentrations for *nrps* KO strain as compared to WT *Mtb*. Complementation of *nrps* KO strain with wild type copy (*nrps* KO:*nrps*) reversed the phenotype. p-values are determined by two-way ANOVA with Bonferroni post-tests relative to WT *Mtb* growth kinetics. (**e**), (**f**) ICP-MS analysis of WT, *nrps* KO and *nrps* KO:*nrps Mtb* strain to measure intracellular levels of trace elements under low, and high zinc conditions. p-values are determined by one-way ANOVA with Dunnett post-tests relative to WT *Mtb* zinc levels. Data represents mean ± s.e.m (n=3 biological replicates, averaged for 2 technical replicates for each biological replicate).

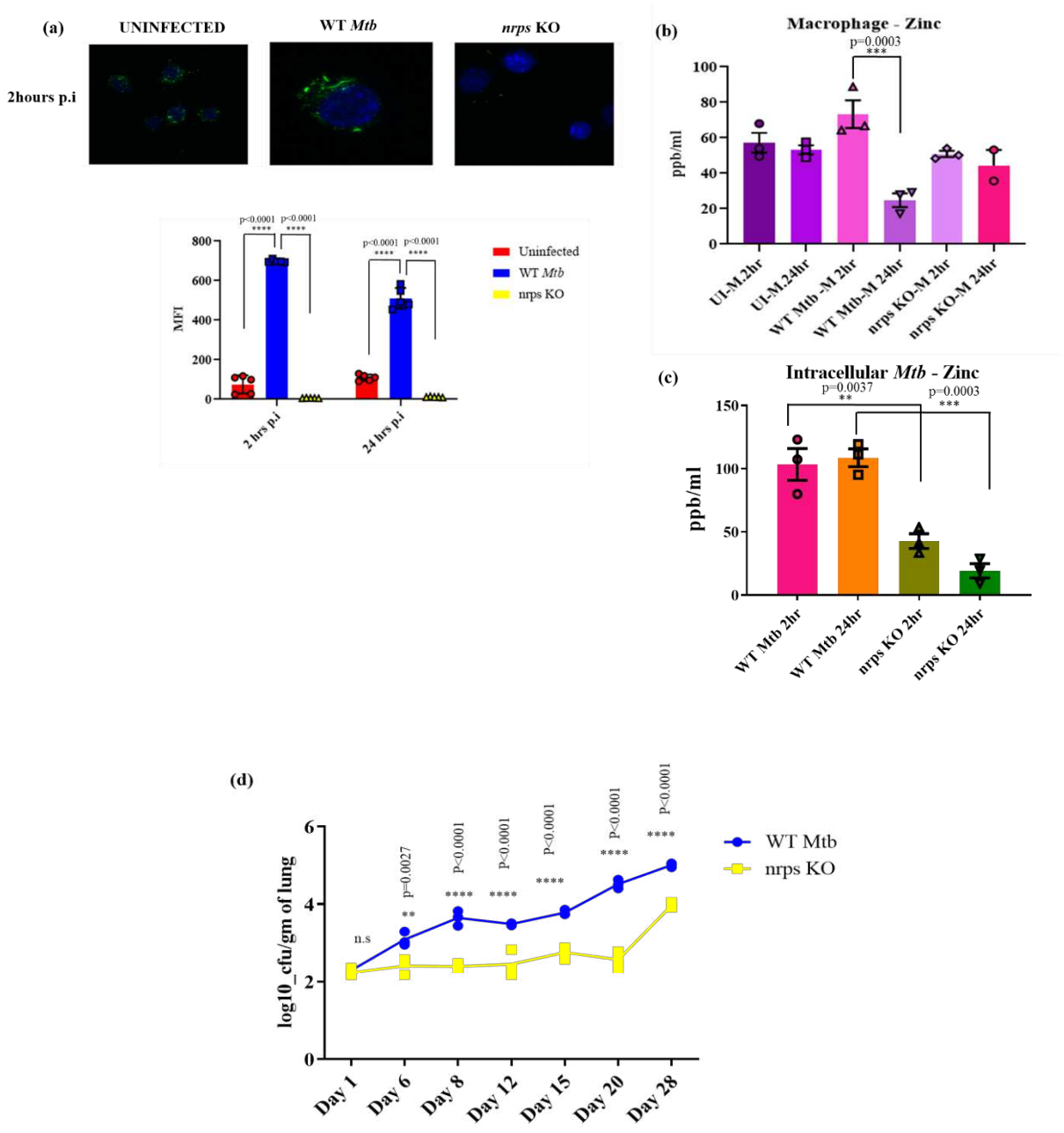


Figure 3: Pathophysiological Relevance of *nrps* gene cluster in *Mtb* macrophage and murine infection (**a**) Free zinc measured by using Fluozin-3AM dye based confocal microscopy for uninfected, WT *Mtb* and *nrps* KO *Mtb* strains infected in murine macrophages. p-values are determined by two-way ANOVA with Bonferroni post-tests relative to WT *Mtb* mean fluorescence intensity. (**b**) ICP-MS analysis showed no significant changes in total zinc pool of macrophages upon *nrps* KO *Mtb* infection as compared to uninfected macrophages. Total macrophage zinc pools significantly decreased 24 hours post WT *Mtb* infection. (**c**) ICP-MS analysis of *Mtb* harvested post macrophage infection show significantly decreased intracellular levels of total zinc in *nrps* KO *Mtb* strain as compared to WT. p-values are determined by one-way ANOVA with Dunnett post-tests relative to WT *Mtb* zinc content. Data represents mean ± s.e.m (n=3 biological replicates, averaged for 2 technical replicates for each biological replicate). (**d**) Balb/c mice (n=3 for each strain) were infected by aerosol exposure to 200 CFU WT or *nrps* KO *Mtb* strains. At the indicated time points post infection, the lungs of infected mice were collected and the bacterial burden was determined by CFU enumeration. p-values are determined by two-way ANOVA with Bonferroni post-tests relative to WT *Mtb*.

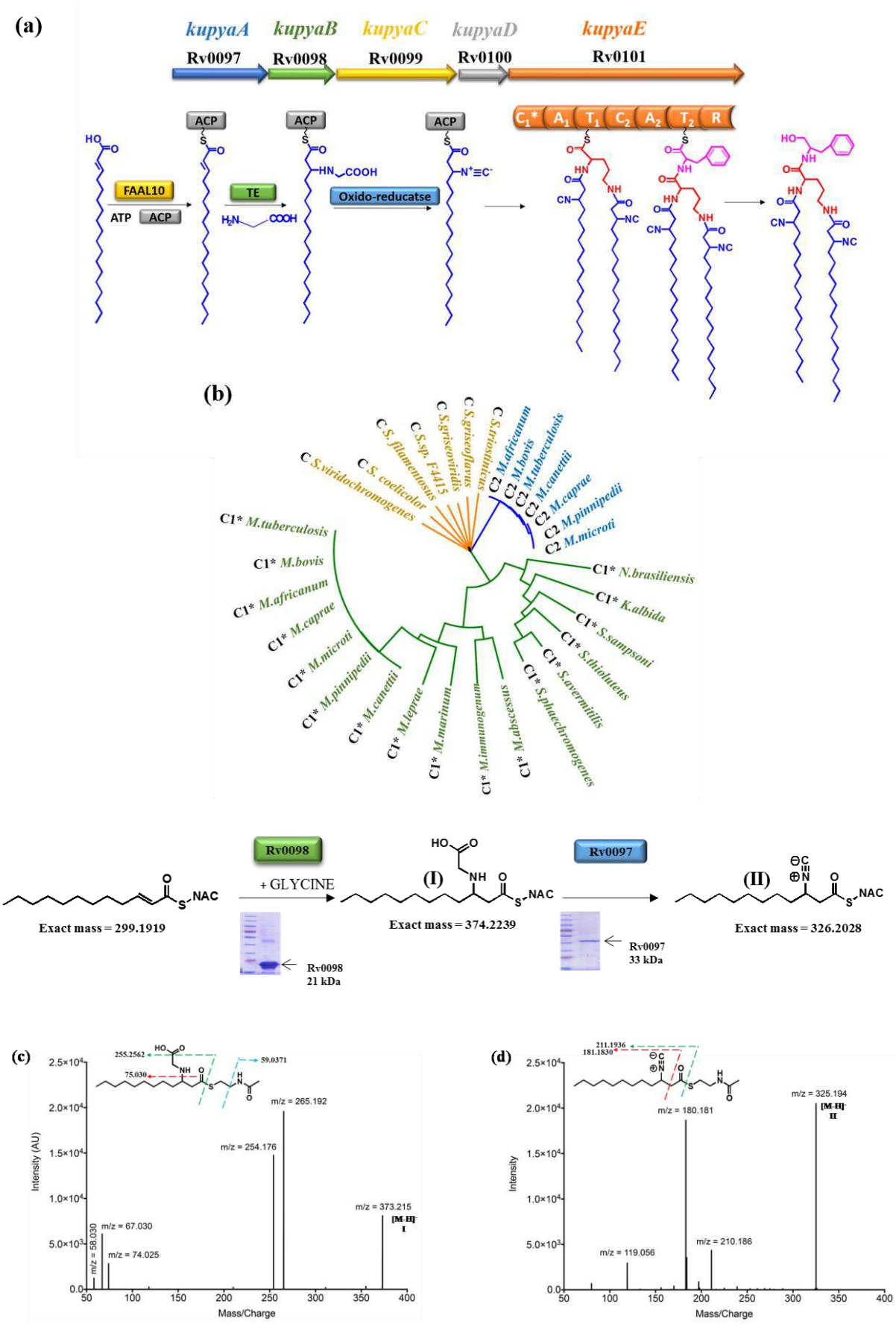
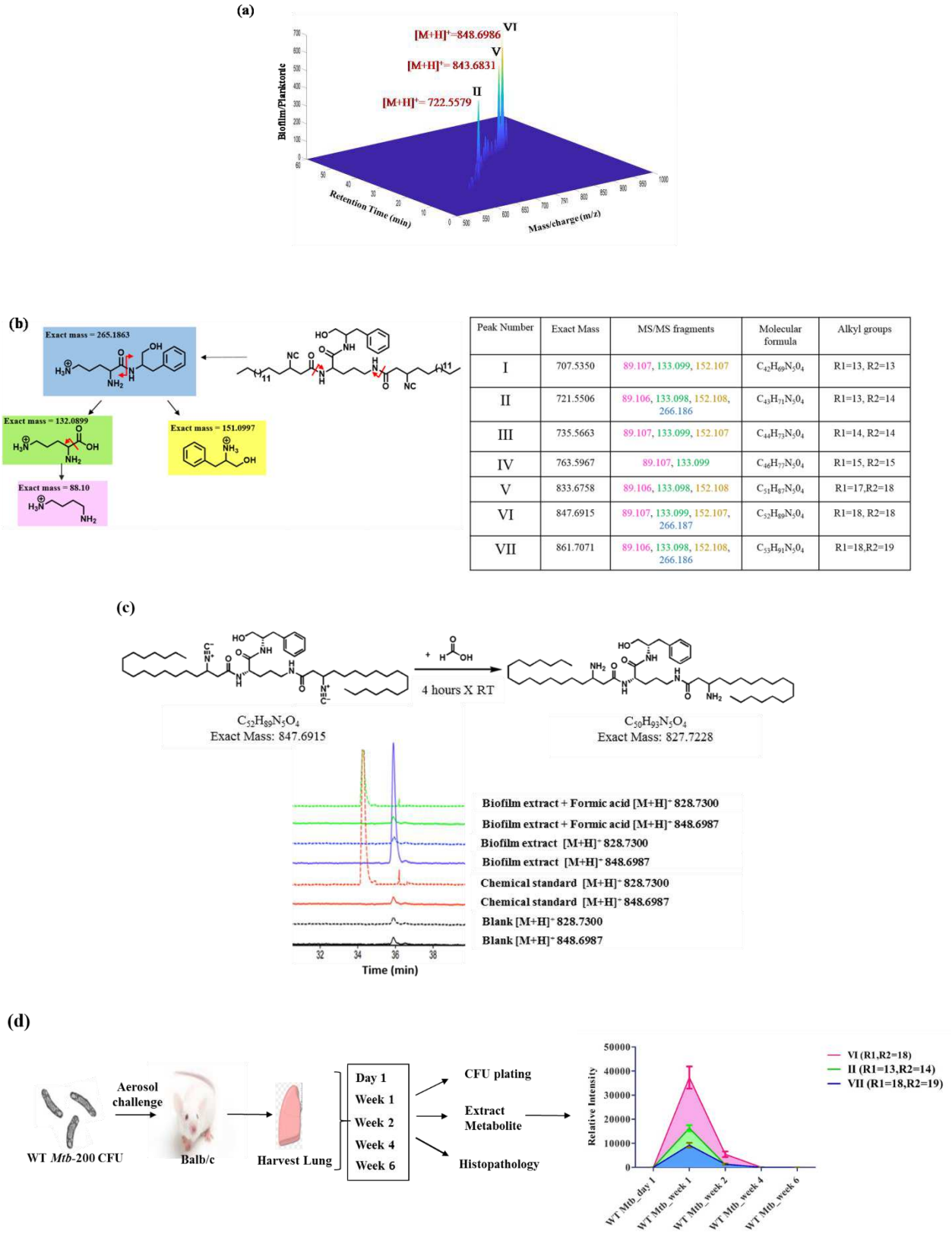

(a)
kupyaA
kupyaB
kupyaC
kupyaD
kupyaE
Rv0097
Rv0098
Rv0099
Rv0100
Rv0101
ACP
FAAL10
ATP
TE
Oxido-reducatse
C1* A1 T1 C2 A2 T2 R
(b)
C S.griseoflavus
C S.griseoviridis
C S.sp. F4415
C S. filamentosus
C S. coelicolor
C S.viridochromogenes
C2 M.africanum
C2 M.bovis
C2 M.tuberculosis
C2 M.canettii
C2 M.caprae
C2 M.pinnipedii
C2 M.microti
C1* M.tuberculosis
C1* M.bovis
C1* M.africanum
C1* M.caprae
C1* M.microti
C1* M.pinnipedii
C1* M.canettii
C1* M.leprae
C1* M.marinum
C1* M.immunogenum
C1* M.abscessus
C1* S.phaechromogenes
C1* S.avermitilis
C1* S.thioluteus
C1* S.sampsoni
C1* K.albida
C1* N.brasiliensis
Rv0098
+ GLYCINE
Rv0097
NAC
(I)
(II)
Exact mass = 299.1919
Exact mass = 374.2239
Exact mass = 326.2028
Rv0098
21 kDa
Rv0097
33 kDa
(c)
Intensity (AU)
Mass/Charge
m/z = 67.030
m/z = 58.030
m/z = 74.025
m/z = 254.176
m/z = 265.192
m/z = 373.215
[M-H]- I
255.2562
59.0371
75.030
(d)
Intensity
Mass/Charge
211.1936
181.1830
m/z = 180.181
m/z = 325.194
[M-H]- II
m/z = 119.056
m/z = 210.186

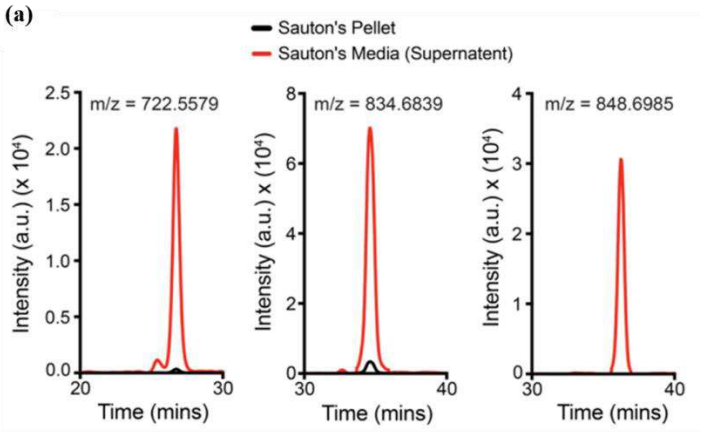
Figure 4: Phylogenetic and Biochemical analysis of *Mtb nrps* gene cluster (**a**) Schematic representation of biosynthesis of kupyaphores by *Mtb* Rv0097-Rv0101 gene cluster. *Mtb* harbors bimodular NRPS with 7 catalytic domains. Putative biosynthetic steps are illustrated here. (**b**) Dendrogram-based clustering analysis of condensation domains that are present in the NRPS protein of similar 5 gene *kupya* biosynthetic loci from other *Actinomycetes* along with classical condensation domain of NRPS genes. The starter condensation domain from these *kupya* NRPS are distinctively different (C1*) from other classical condensation domains (C). The second condensation domain (C2) present in bimodular *kupya* NRPS protein clusters with classical C domains. (**c**) MS/MS spectra showing fragmentation products of glycine adduct of 2-dodecenoic-SNAC detected in the enzymatic assay with Rv0098 purified protein. No product could be detected in control reactions with no protein or no glycine. (**d**) MS/MS spectra of additional peak of m/z 326.2028 corresponding to isonitrile adduct of 2-dodecanoic-SNAC observed upon addition of Rv0097 purified protein to the Rv0098 assay conditions. The calculated masses for intermediate metabolites **I** and **II** are within 5 ppm mass error tolerance at MS1. All assays in (**c**) and (**d**) were performed in triplicates with reproducible results each time.

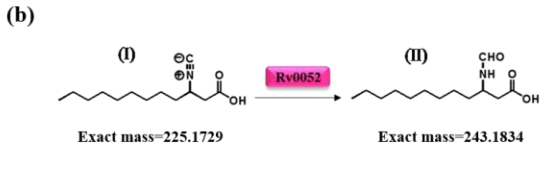


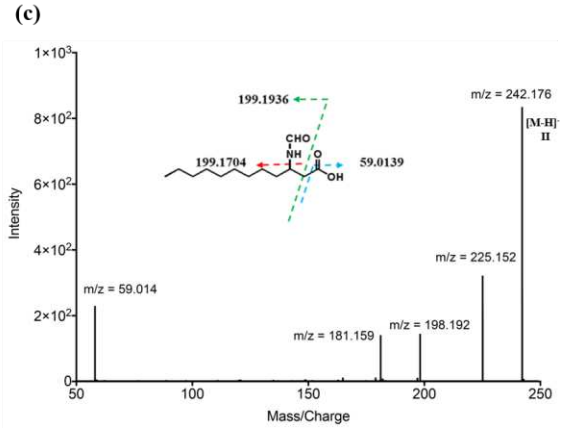
Peak Number	Exact Mass	MS/MS fragments	Molecular formula	Alkyl groups
I	707.5350	89.107, 133.099, 152.107	$C_{42}H_{69}N_5O_4$	R1=13, R2=13
II	721.5506	89.106, 133.098, 152.108, 266.186	$C_{43}H_{71}N_5O_4$	R1=13, R2=14
III	735.5663	89.107, 133.099, 152.107	$C_{44}H_{73}N_5O_4$	R1=14, R2=14
IV	763.5967	89.107, 133.099	$C_{46}H_{77}N_5O_4$	R1=15, R2=15
V	833.6758	89.106, 133.098, 152.108	$C_{51}H_{87}N_5O_4$	R1=17,R2=18
VI	847.6915	89.107, 133.099, 152.107, 266.187	$C_{52}H_{89}N_5O_4$	R1=18, R2=18
VII	861.7071	89.106, 133.098, 152.108, 266.186	$C_{53}H_{91}N_5O_4$	R1=18,R2=19

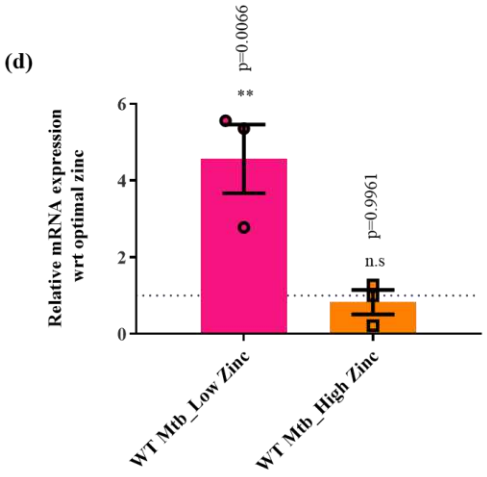
Figure 5: Identification and characterization of Kupyaphores from *Mtb* (**a**) 3D heat map plot shows masses uniquely identified from biofilm extracts of *Mtb* cultures but not planktonic cultures at various retention times. 3 peaks with the highest fold change between biofilm and planktonic cultures corresponded to m/z $[M+H]^+$= 722.55, 834.68, 848.69 are marked here. (**b**) MS/MS fragmentation

pattern revealed common dipeptide backbone of ornithine and phenylalaninol in all the 7 ions observed in the complete spectra. Common fragments identified in the positive ion mode for all these masses show identical pattern. The predicted fragments are color coded along with their masses in the table. (**c**) Mass spectrometry based comparative analysis of EIC for biofilm organic extract and synthetic chemical standard at $[M+H]^+$ = 828.7300 and 848.6987. Addition of formic acid to biofilm organic extract shift the mass of parent ion (m/z-848.6987) to that of synthetic chemical standard of amino substituted lipopeptide (m/z-828.73) concomitant with conversion of isonitrile to amine upon acid hydrolysis. (**d**) Plot shows MRM-based semi-quantitative LC-MS measurements of Kupyaphore metabolites for the three masses detected in lung homogenate, produced by *Mtb* only in the lungs of infected mice (n=3). The observed masses for **I–VII** are within 10 ppm mass error tolerance at MS1 or MS/MS for each mass.


(a)
Sauton's Pellet
Sauton's Media (Supernatent)
m/z = 722.5579
m/z = 834.6839
m/z = 848.6985
Intensity (a.u.) (x 10⁴)
Time (mins)


(b)
(I)
(II)
Rv0052
Exact mass=225.1729
Exact mass=243.1834


(c)
199.1936
199.1704
59.0139
m/z = 242.176
[M-H]⁻ II
m/z = 225.152
m/z = 59.014
m/z = 181.159
m/z = 198.192
Intensity
Mass/Charge


(d)
p=0.0066
**
p=0.9961
n.s
Relative mRNA expression wrt optimal zinc
WT Mtb_Low Zinc
WT Mtb_High Zinc

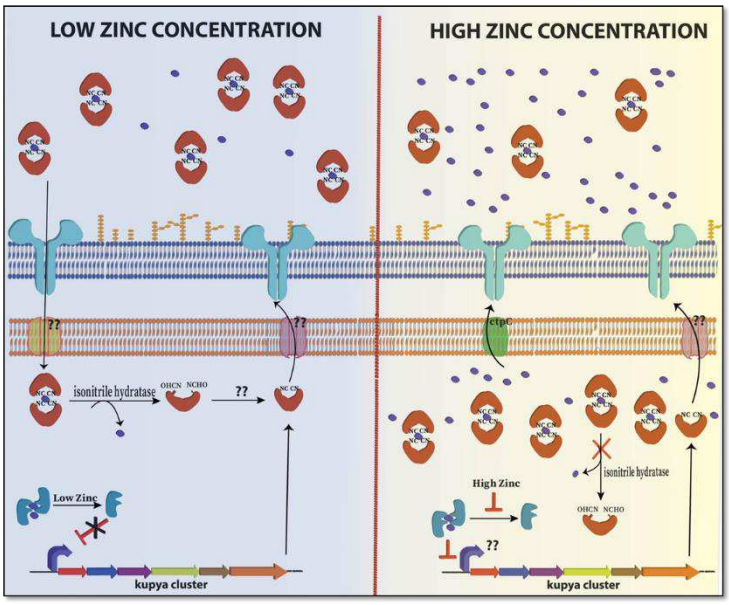

(e)
LOW ZINC CONCENTRATION
HIGH ZINC CONCENTRATION
isonitrile hydratase
Low Zinc
High Zinc
kupya cluster

Figure 6: Release of zinc from Kupyaphores is mediated by *Mtb* isonitrile hydratase (a) Extracted Ion Chromatogram (EIC) of Kupyaphore of m/z $[M+H]^+$= **II**, **V** and **VI** detected from organic extracts from the supernatant of adapted *Mtb* WT cultures grown in Sauton's media under planktonic conditions relative to the cell pellet organic extracts. **(b)** Reaction scheme for *invitro* assay of isonitrile hydratase. **(c)** MS/MS spectra with fragmentation pattern and products marked, confirming the presence of formamide product in assay. The observed masses are within 10 ppm mass error tolerance at MS1 or MS/MS for each mass. **(d)** qRT based mRNA expression analysis of RV0052 in WT *Mtb* under low and high zinc condition as compared to optimal zinc growth condition. Data represents mean ± s.e.m (n=3 biological replicates, averaged for 2 technical replicates for each biological replicate). p-values are determined by two-way ANOVA with Bonferroni post-tests relative to WT *Mtb* supplemented with 6μM zinc. **(e)** 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.

Figures

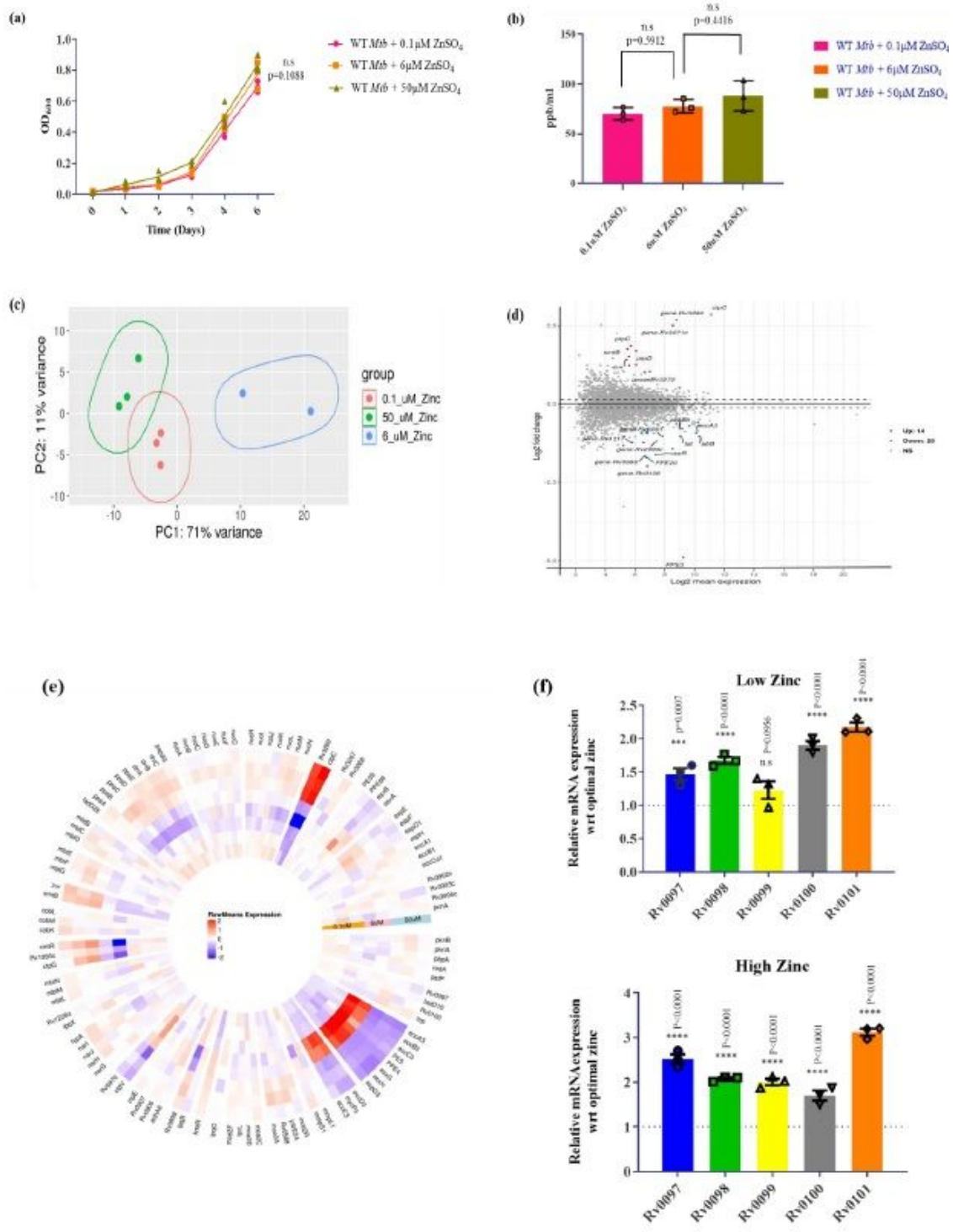


Figure 1

Remodeling of Mtb transcriptional responses maintains zinc homeostasis (a) Growth kinetics of WT Mtb grown in chelated Sauton's media supplemented with varying zinc concentrations – 0.1uM, 6uM and 50uM. p-values are determined by two-way ANOVA with Bonferroni post-tests relative to 6uM zinc

supplemented condition. (b) ICP-MS analysis of total intracellular zinc content of Mtb grown in these three conditions. p-values are determined by one-way ANOVA with Dunnett post-tests relative to 6uM zinc supplemented condition. (c) 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. (d) MA plot of differential gene expression analysis in 50uM zinc supplemented condition as compared to 0.1uM zinc supplementation with p-value <0.05 and log2 fold change>0.5 as cut-off for significance. Genes marked in red are significantly upregulated, while genes marked in blue are significantly downregulated. (e) Circos plot for differential expression of few metal responsive Mtb operons in three different zinc conditions. (f) Gene expression analysis of Rv0097-Rv0101 through qRT-PCR under low and high zinc conditions as compared to optimal zinc concentration. p-values are determined by one-way ANOVA with Dunnett post-tests relative to 6uM zinc supplemented condition. Data represents mean ± s.e.m (n=3 biological replicates, averaged for 2 technical replicates for each biological replicate).

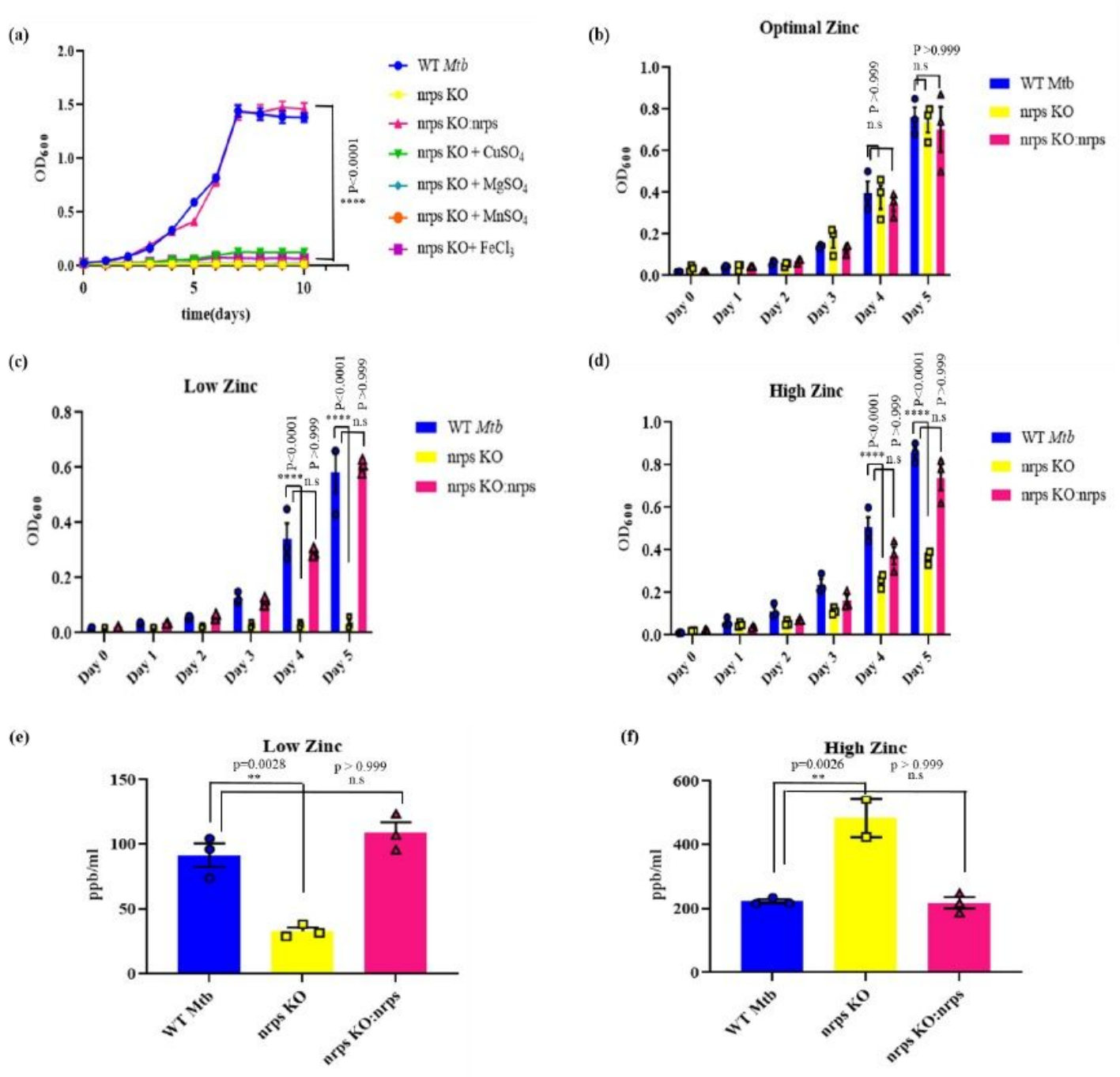


Figure 2

Role of Mtb nrps in mycobacterial physiology (a) Growth kinetics of wild type (WT) and Mtb strain knocked out for nrps show mutant to be severely compromised in nutrient deficient condition of chelated Sauton's Media. Addition of other key transition elements - copper, magnesium manganese or iron to Sauton's media failed to revive mutant significantly. p-values are determined by two-way ANOVA with Bonferroni post-tests relative to WT Mtb growth kinetics. (b) Only supplementation with 6μM zinc could restore growth profile of Mtb nrps knockout (KO) strain. (c), (d) Titration of Sauton's media with different concentrations of zinc show compromised growth pattern of mutant under both low (0.1μM) and high (50μM) zinc concentrations for nrps KO strain as compared to WT Mtb. Complementation of nrps KO

strain with wild type copy (nrps KO:nrps) reversed the phenotype. p-values are determined by two-way ANOVA with Bonferroni post-tests relative to WT Mtb growth kinetics. (e), (f) ICP-MS analysis of WT, nrps KO and nrps KO:nrps Mtb strain to measure intracellular levels of trace elements under low, and high zinc conditions. p-values are determined by one-way ANOVA with Dunnett post-tests relative to WT Mtb zinc levels. Data represents mean ± s.e.m (n=3 biological replicates, averaged for 2 technical replicates for each biological replicate).

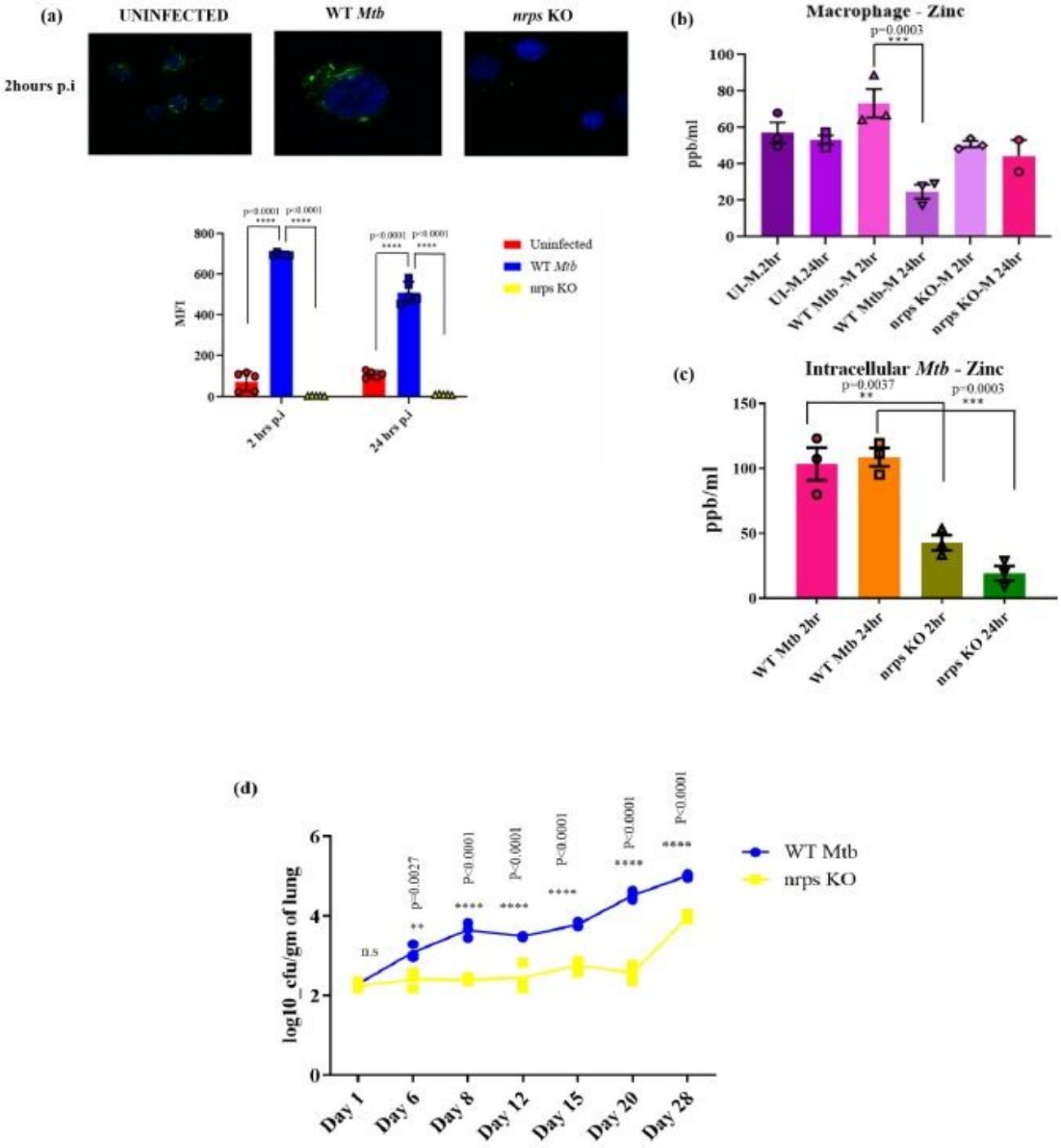


Figure 3

Pathophysiological Relevance of nrps gene cluster in Mtb macrophage and murine infection (a) Free zinc measured by using Fluozin-3AM dye based confocal microscopy for uninfected, WT Mtb and nrps KO Mtb strains infected in murine macrophages. p-values are determined by two-way ANOVA with Bonferroni post-tests relative to WT Mtb mean fluorescence intensity. (b) ICP-MS analysis showed no significant changes in total zinc pool of macrophages upon nrps KO Mtb infection as compared to uninfected macrophages. Total macrophage zinc pools significantly decreased 24 hours post WT Mtb infection. (c) ICP-MS analysis of Mtb harvested post macrophage infection show significantly decreased intracellular levels of total zinc in nrps KO Mtb strain as compared to WT. p-values are determined by one-way ANOVA with Dunnett post-tests relative to WT Mtb zinc content. Data represents mean ± s.e.m (n=3 biological replicates, averaged for 2 technical replicates for each biological replicate). (d) Balb/c mice (n=3 for each strain) were infected by aerosol exposure to 200 CFU WT or nrps KO Mtb strains. At the indicated time points post infection, the lungs of infected mice were collected and the bacterial burden was determined by CFU enumeration. p-values are determined by two-way ANOVA with Bonferroni post-tests relative to WT Mtb.

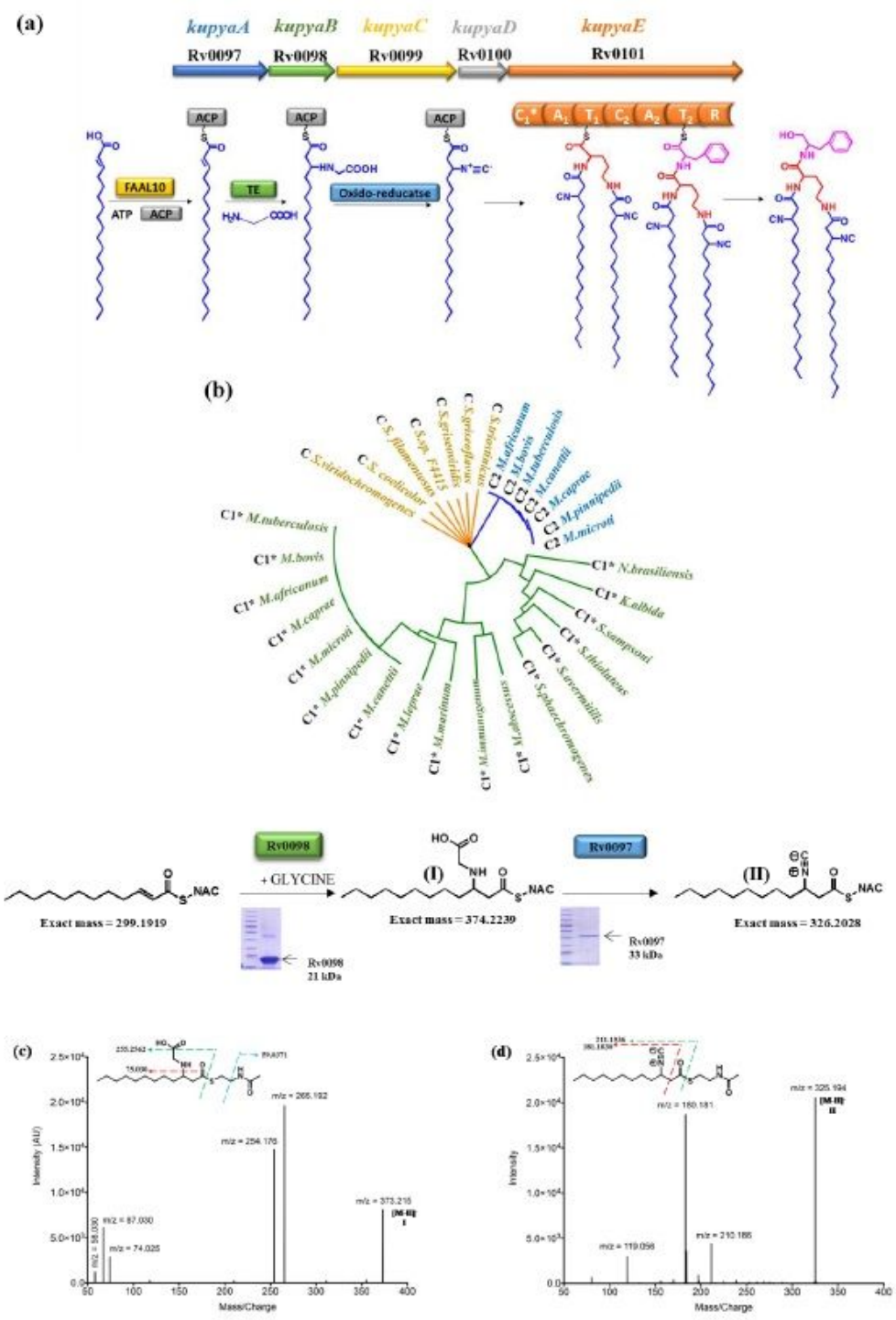
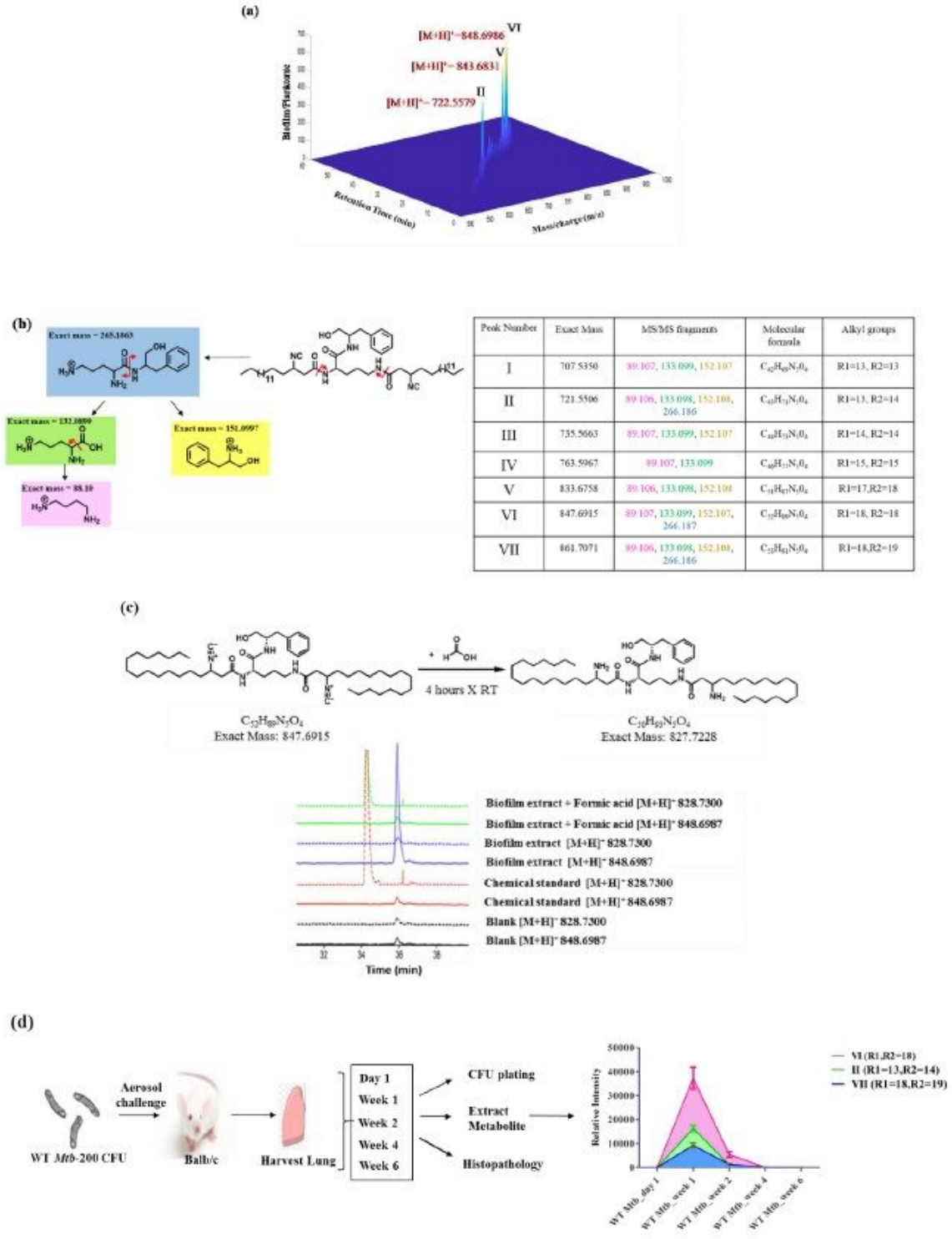


Figure 4

Phylogenetic and Biochemical analysis of Mtb nrps gene cluster (a) Schematic representation of biosynthesis of kupyaphores by Mtb Rv0097-Rv0101 gene cluster. Mtb harbors bimodular NRPS with 7 catalytic domains. Putative biosynthetic steps are illustrated here. (b) Dendrogram-based clustering analysis of condensation domains that are present in the NRPS protein of similar 5 gene kupya biosynthetic loci from other Actinomycetes along with classical condensation domain of NRPS genes.

The starter condensation domain from these kupya NRPS are distinctively different (C1*) from other classical condensation domains (C). The second condensation domain (C2) present in bimodular kupya NRPS protein clusters with classical C domains. (c) MS/MS spectra showing fragmentation products of glycine adduct of 2-dodecenoic-SNAC detected in the enzymatic assay with Rv0098 purified protein. No product could be detected in control reactions with no protein or no glycine. (d) MS/MS spectra of additional peak of m/z 326.2028 corresponding to isonitrile adduct of 2-dodecanoic-SNAC observed upon addition of Rv0097 purified protein to the Rv0098 assay conditions. The calculated masses for intermediate metabolites I and II are within 5 ppm mass error tolerance at MS1. All assays in (c) and (d) were performed in triplicates with reproducible results each time.



Peak Number	Exact Mass	MS/MS fragments	Molecular formula	Alkyl groups
I	707.5350	89.107, 133.099, 152.107	$C_{42}H_{69}N_7O_4$	R1=13, R2=13
II	721.5506	89.106, 133.098, 152.108, 266.186	$C_{43}H_{71}N_7O_4$	R1=13, R2=14
III	735.5663	89.107, 133.099, 152.107	$C_{44}H_{73}N_7O_4$	R1=14, R2=14
IV	763.5967	89.107, 133.099	$C_{46}H_{77}N_7O_4$	R1=15, R2=15
V	833.6758	89.106, 133.098, 152.108	$C_{51}H_{85}N_7O_4$	R1=17,R2=18
VI	847.6915	89.107, 133.099, 152.107, 266.187	$C_{52}H_{87}N_7O_4$	R1=18, R2=18
VII	861.7071	89.106, 133.098, 152.108, 266.186	$C_{53}H_{89}N_7O_4$	R1=18,R2=19

Figure 5

Identification and characterization of Kupyaphores from Mtb (a) 3D heat map plot shows masses uniquely identified from biofilm extracts of Mtb cultures but not planktonic cultures at various retention times. 3 peaks with the highest fold change between biofilm and planktonic cultures corresponded to m/z [M+H]+= 722.55, 834.68, 848.69 are marked here. (b) MS/MS fragmentation pattern revealed common dipeptide backbone of ornithine and phenylalaninol in all the 7 ions observed in the complete spectra.

Common fragments identified in the positive ion mode for all these masses show identical pattern. The predicted fragments are color coded along with their masses in the table. (c) Mass spectrometry based comparative analysis of EIC for biofilm organic extract and synthetic chemical standard at [M+H]+ = 828.7300 and 848.6987. Addition of formic acid to biofilm organic extract shift the mass of parent ion (m/z-848.6987) to that of synthetic chemical standard of amino substituted lipopeptide (m/z-828.73) concomitant with conversion of isonitrile to amine upon acid hydrolysis. (d) Plot shows MRM-based semi-quantitative LC-MS measurements of Kupyaphore metabolites for the three masses detected in lung homogenate, produced by Mtb only in the lungs of infected mice (n=3). The observed masses for I–VII are within 10 ppm mass error tolerance at MS1 or MS/MS for each mass.

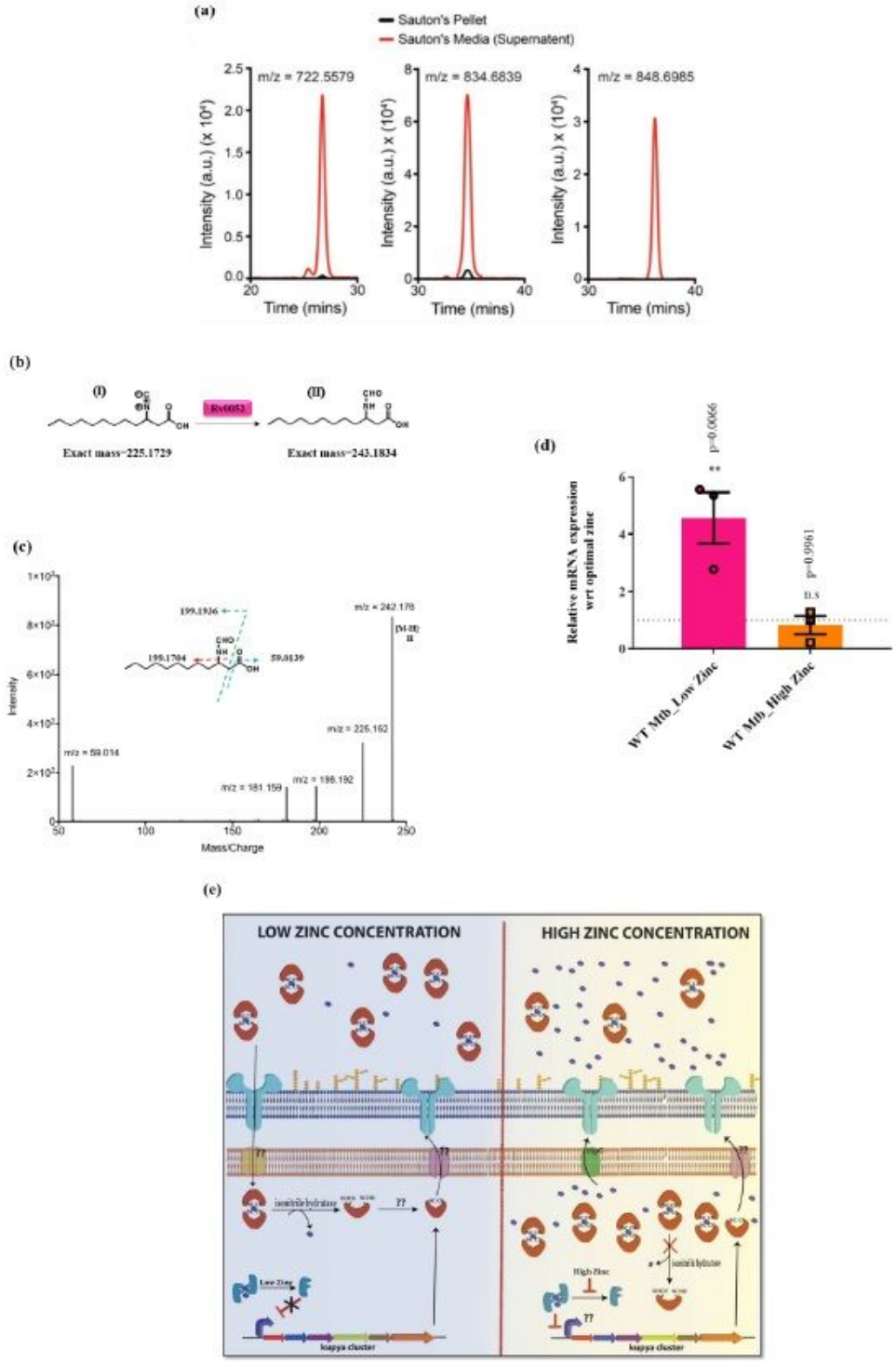


Figure 6

Release of zinc from Kupyaphores is mediated by Mtb isonitrile hydratase (a) Extracted Ion Chromatogram (EIC) of Kupyaphore of m/z [M+H]+= II, V and VI detected from organic extracts from the supernatant of adapted Mtb WT cultures grown in Sauton's media under planktonic conditions relative to the cell pellet organic extracts. (b) Reaction scheme for invitro assay of isonitrile hydratase. (c) MS/MS spectra with fragmentation pattern and products marked, confirming the presence of formamide product

in assay. The observed masses are within 10 ppm mass error tolerance at MS1 or MS/MS for each mass. (d) qRT based mRNA expression analysis of RV0052 in WT Mtb under low and high zinc condition as compared to optimal zinc growth condition. Data represents mean ± s.e.m (n=3 biological replicates, averaged for 2 technical replicates for each biological replicate). p-values are determined by two-way ANOVA with Bonferroni post-tests relative to WT Mtb supplemented with 6uM zinc. (e) 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 Files

This is a list of supplementary files associated with this preprint. Click to download.

- SImodified.pdf