


# SCIENTIFIC REPORTS



OPEN

## Novel Type III Polyketide Synthases Biosynthesize Methylated Polyketides in *Mycobacterium marinum*

Amreesh Parvez<sup>1</sup>, Samir Giri<sup>1,2</sup>, Gorkha Raj Giri<sup>1</sup>, Monika Kumari<sup>1,3</sup>, Renu Bisht<sup>1</sup> & Priti Saxena<sup>1</sup> 

*Mycobacterial* pathogenesis is hallmarked by lipidic polyketides that decorate the cell envelope and mediate infection. However, factors mediating persistence remain largely unknown. Dynamic cell wall remodeling could facilitate the different pathogenic phases. Recent studies have implicated type III polyketide synthases (PKSs) in cell wall alterations in several bacteria. Comparative genome analysis revealed several type III *pks* gene clusters in mycobacteria. In this study, we report the functional characterization of two novel type III PKSs, MMAR\_2470 and MMAR\_2474, in *Mycobacterium marinum*. These type III *pks* belong to a unique *pks* genomic cluster conserved exclusively in pathogenic mycobacteria. Cell-free reconstitution assays and high-resolution mass spectrometric analyses revealed methylated polyketide products in independent reactions of both proteins. MMAR\_2474 protein exceptionally biosynthesized methylated alkyl-resorcinol and methylated acyl-phloroglucinol products from the same catalytic core. Structure-based homology modeling, product docking, and mutational studies identified residues that could facilitate the distinctive catalysis of these proteins. Functional investigations in heterologous mycobacterial strain implicated MMAR\_2474 protein to be vital for mycobacterial survival in stationary biofilms. Our investigations provide new insights into type III PKSs conserved in pathogenic mycobacterial species.

*Mycobacteria* genus comprises of some of the deadliest pathogenic species<sup>1</sup>. While *Mycobacterium tuberculosis* (Mtb) and *Mycobacterium leprae* have been banes of humankind, members of related species are known to infect other organisms, including cattle, birds, fish and other marine hosts<sup>2</sup>. *Mycobacterium marinum* (Mmar), the water-dwelling mycobacteria, causes systemic granulomas and lesions in fish and frog and is the causative agent of aquarium granulomas in humans. Mmar shares 85% genome-wide amino acid sequence identity with Mtb and is a model organism for dissecting mycobacterial pathogenesis<sup>3</sup>. Last two decades have seen tremendous progress in our understanding of mycobacterial virulence<sup>4–7</sup>. Mega-synthetic polyketide synthases (PKSs), in conjunction with fatty acid synthases, generate a repertoire of polyketide lipids that surface Mtb cell envelope and aid virulent phases of tuberculosis<sup>6</sup>. Efforts, however, are on to dissect molecular mechanisms that mediate mycobacterial persistence<sup>4,8–10</sup>.

Comparative genome analyses has revealed conservation of several biosynthetic *pks* genes across various mycobacterial species<sup>11,12</sup>. Many of these *pks* are clustered in biosynthetic operons and are uniquely conserved in genomes of pathogenic mycobacteria. Mmar genome harbors 34 ORFs homologous to *pks* genes, with four being type III *pks*<sup>3</sup>. Type III polyketide synthases have shown immense potential to biosynthesize architecturally diverse and bio-functionally distinctive metabolites<sup>13–19</sup>. These small, homodimeric proteins utilize a single catalytic cysteine residue in each monomer and iteratively condense activated coenzyme-A (CoA) analogs of mono-carboxylic aliphatic/aromatic starter substrates with repeating di-carboxylic malonyl-/methylmalonyl-CoA extender units<sup>20–22</sup>. It is interesting to note that these proteins can cyclize common reaction intermediates through at least three chemically distinct mechanisms, leading to an array of products with

<sup>1</sup>Chemical Biology Group, Faculty of Life Sciences and Biotechnology, South Asian University, New Delhi, 110021, India. <sup>2</sup>Present address: Department of Ecology, School of Biology, University of Osnabrück, Osnabrück, 49076, Germany. <sup>3</sup>Present address: Department of Biochemistry, University College of Medical Sciences, Delhi, 110095, India. Amreesh Parvez and Samir Giri contributed equally to this work. Correspondence and requests for materials should be addressed to P.S. (email: [psaxena@sau.ac.in](mailto:psaxena@sau.ac.in))

varied chemical scaffolds<sup>23,24</sup>. Substrate promiscuity, variable iterations, and multiple cyclization potential confer these proteins with the ability to generate a palette of diverse products utilizing limited substrate pool<sup>17,22,23,25–31</sup>. Variability is further generated by post-synthesis modifications of the polyketide products by other modifying enzymes that impart biological activity to the nascent polyketide cores<sup>32</sup>. Notably, genes for polyketide modifying enzymes are known to be clustered with type III *pks* genes in microbial genomes<sup>15,33,34</sup>. Several of these type III *pks* clusters have been recently implicated in generating functionally important metabolites. *srs* cluster in *Streptomyces griseus* produces methylated polyketides that confer antibiotic resistance to the species<sup>34</sup>. Phenolic lipids in *Azotobacter vinelandii* are biosynthesized by a cluster comprising of type I *pks*s and two type III *pks*s<sup>35</sup>. These molecules replace membrane phospholipids during encystment of dormant cells and are important components of cyst exine<sup>35</sup>. Recently, methylated polyketide quinones from *Mycobacterium smegmatis* (Msmeg) have been shown to facilitate anaerobic respiration of Msmeg cells in stationary biofilms<sup>36</sup>. These molecules are produced by a genomic cluster of Msmeg homologous to the *srs* operon. While an orthologous genomic cluster is not present in Mtb, polyketide quinones were isolated from Mtb biofilms and a homologous type III PKS was implicated in their biosynthesis. Growth in multicellular aggregates, like biofilms, provides a secure niche for developing mycobacterial persistence and advantages of multi-drug resistance<sup>37,38</sup>.

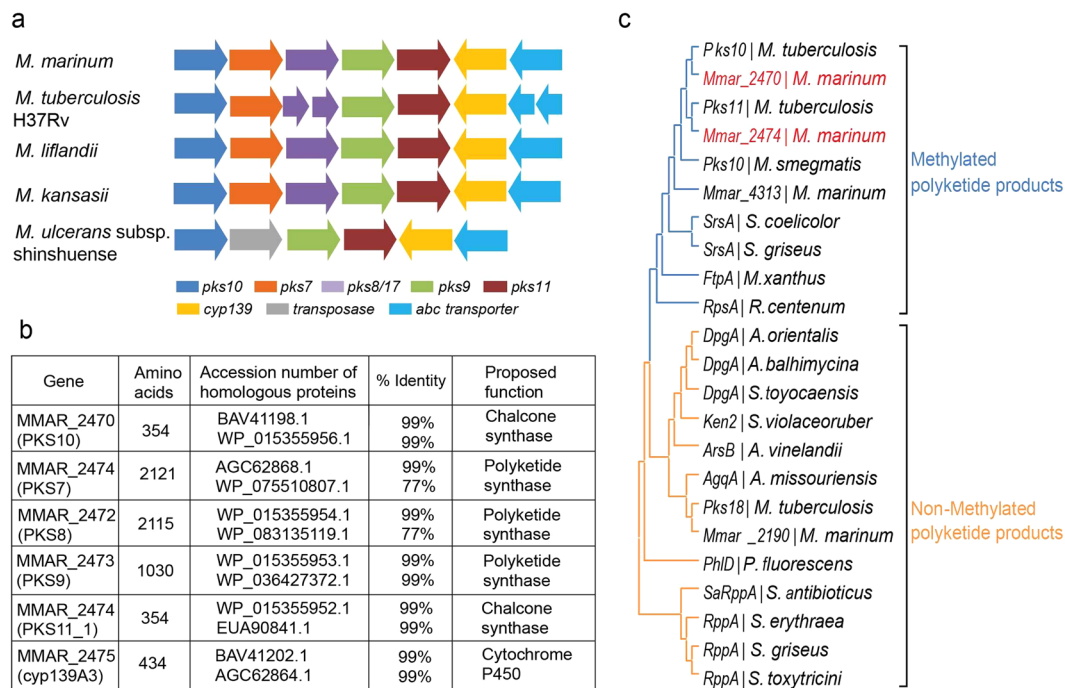
Here, we report the functional characterization of two novel type III PKSs from Mmar. These type III *pks*s belong to *Mycobacterium* specific 2x type III *pks* genomic cluster conserved exclusively in pathogenic species. MMAR\_2470 and MMAR\_2474 proteins biosynthesized methylated polyketide metabolites from acyl-CoA starter units selectively utilizing two different extender substrates. Distinctively, MMAR\_2474 protein catalyzed the biosynthesis of methylated alkyl-resorcinols and methylated acyl-phloroglucinols from the same protein core. Homology-based structural modeling, molecular docking and site directed mutagenesis identified catalytically crucial active site residue positions and revealed clues to the structure-function relationship in Mmar type III PKSs. High-resolution metabolomic investigations of heterologous complemented type III *pks* knockout mycobacterial strain revealed the physiological importance of methylated polyketides and implicated the biosynthetic proteins in aiding mycobacterial survival in biofilms.

## Results

**Genomic conservation of 2x type III *pks* cluster in pathogenic mycobacteria.** Mycobacterial genome sequencing has revealed a plethora of genes homologous to polyketide synthases (*pks*s). Many of these *pks*s are clustered with other polyketide related genes in the genomes. Functional characterization of several *pks* clusters in Mtb in the last two decades have implicated several multi-domain, mega-synthetic type I PKSs in the biosynthesis of unique virulent mycobacterial polyketide lipids<sup>6</sup>. Lately, type III polyketides from diverse microorganisms have exhibited remarkable functional capabilities. Comparative genome analyses revealed novel type III *pks* clusters in several mycobacterial species. We investigated the occurrence of type III *pks* genes in Mmar. This organism shares substantial number of conserved genes with human pathogen Mtb<sup>3,39</sup>. Genome sequencing identified four type III *pks* genes in Mmar. Interestingly, three of these genes, *mmar\_2190*, *mmar\_2470* and *mmar\_2474* were exclusively conserved in pathogenic mycobacteria whereas; the fourth homolog, *mmar\_4313*, could be identified in several environmental mycobacteria. The *mmar\_4313* genomic cluster of Mmar is orthologous to the polyketide quinone biosynthetic gene cluster recently characterized in Msmeg<sup>36</sup>. *mmar\_2190* is an ortholog of alkyl pyrone biosynthesizing *pks18* gene of Mtb. The *mmar\_2470* and *mmar\_2474* genes constitute a unique genomic cluster in Mmar as shown in Fig. 1a. This gene cluster contains three multi-functional type I *pks* genes flanked on either side by the type III *pks*s. The cluster shows conservation of cytochrome P450 and ABC transporter genes downstream of *mmar\_2474*. It is interesting to note that an orthologous 2x type III *pks* genomic organization, though identified to be non-essential in Mtb genome<sup>40,41</sup>, is exclusively conserved in almost all pathogenic mycobacteria with 77–99% amino acid sequence identity (Fig. 1b), suggesting the possibility of its involvement in mycobacterial pathogenesis.

Type III PKSs demonstrate promiscuity in substrate selection, leading to variability in chemical profiles of the biosynthesized polyketide products. A number of iterative condensations of growing polyketide chain and cyclization chemistry facilitated by the protein to generate biosynthesized products further augment product diversity. We attempted to predict the biosynthetic potential of MMAR\_2470 and MMAR\_2474 proteins by comparing these with experimentally characterized bacterial type III PKSs. Figure 1c shows a phylogenetic tree constructed using amino acid sequences of 23 bacterial type III PKS proteins exhibiting varied enzymatic catalysis. It is interesting to note that sequences cluster in three distinct groups, based on substrate preference and chemical architecture of the biosynthesized products. Proteins with the ability to condense several molecules of di-carboxylic malonyl-CoA units, without requirement of a mono-carboxylic starter substrate, are closely placed in phylogenetic tree, next to the group that represents sequences of proteins with a specific preference for malonyl-CoA as extender substrate. Proteins in both these groups catalyze formation of non-methylated polyketide scaffolds. It can be noted that of the four type III PKSs of Mmar, MMAR\_2190 protein groups with malonyl-CoA accepting proteins, while the other three Mmar type III PKSs distinctly cluster with proteins that biosynthesize methylated polyketide products. These C-methylated polyketides are biosynthesized by incorporation of methylmalonyl-CoA along with malonyl-CoA extender units during polyketide chain elongation. Our phylogenetic analyses could reliably group Mmar proteins along with their orthologous sequences from other species and provides clues to the substrate preference and product profiles of the Mmar type III PKS proteins. Based on these analyses we predicted that MMAR\_2470 and MMAR\_2474 proteins would exhibit the potential to utilize both malonyl- and methylmalonyl-CoA extender units and biosynthesize methylated polyketide metabolites.

**Homology modeling and structure/function analyses of Mmar type III PKSs.** We further investigated the possibility of biosynthesis of methylated metabolites by mycobacterial proteins using structure-based homology modeling. Three-dimensional homology models were generated for MMAR\_2470 and MMAR\_2474

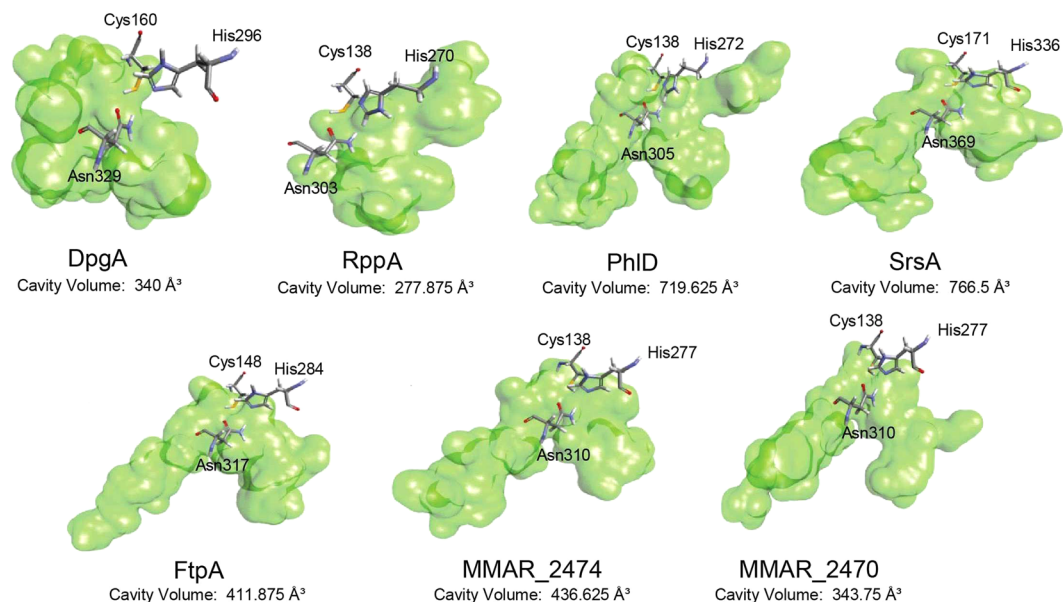


**Figure 1.** Conservation of 2x type III *pks* genomic cluster and phylogenetic analysis of *M. marinum* (Mmar) type III PKSs. **(a)** Conservation of the 2x type III *pks* genomic cluster in pathogenic mycobacterial species. **(b)** Sequence analysis of Mmar 2x type III *pks* genomic cluster. **(c)** Phylogenetic analyses of Mmar type III PKS proteins with functionally characterized microbial type III PKSs. Type III PKSs biosynthesizing non-methylated polyketides are represented in orange color, while type III PKSs biosynthesizing methylated polyketide products are represented in blue color. The Mmar type III PKSs discussed in this study are highlighted in red color. The source organism is mentioned along with the protein names.

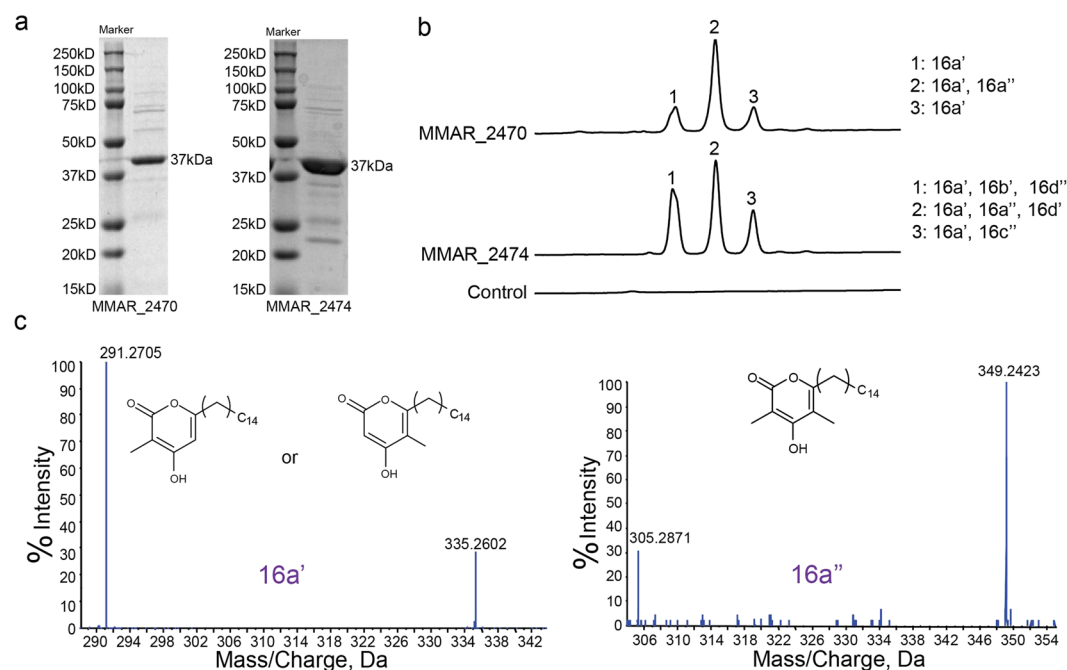
proteins using the x-ray crystallographic structure of Mtb PKS11 protein (4JAJ\_D)<sup>42</sup> as template. Functionally characterized FtpA, DpgA, PhlD and SrsA type III PKSs from *M. xanthus*<sup>43</sup>, *S. lividans*<sup>44</sup>, *P. fluorescens*<sup>45</sup> and *S. griseus*<sup>18</sup>, respectively, were modeled using closest structural homolog for each. Details of templates used for generating each model and the calculated model energies are provided in Supplementary Table S1. Computationally generated models along with the crystal structure of THNS protein (1U0M\_B) from *S. erythraea*<sup>20</sup> were probed for differences in volume capacity of the product binding cavities in core of the proteins. As can be observed from Fig. 2, the product binding cavities of MMAR\_2470 and MMAR\_2474 are strikingly similar to the cavity of FtpA protein. FtpA utilizes both malonyl- and methylmalonyl-CoA extender units and synthesizes methylated phenolic lipids with long-chain acyl-CoA starter molecules. Based on our *in silico* analyses, we speculated that MMAR\_2470 and MMAR\_2474 proteins would be able to utilize methylmalonyl-CoA as an extender unit along with malonyl-CoA and biosynthesize methylated polyketide products.

**Biochemical characterization of Mycobacterial type III PKSs.** The two type III PKSs, MMAR\_2470 and MMAR\_2474 proteins in Mmar, share 74% sequence identity with each other. We cloned the *mmar\_2470* and *mmar\_2474* genes from genomic DNA of Mmar and over-expressed these proteins in *Escherichia coli* under the control of T7 expression system. The recombinant hexa-histidine-tagged proteins were purified by Ni<sup>2+</sup>-nitrilotriacetic acid affinity chromatography as single protein bands of ~37 kDa each, as determined by SDS-PAGE (Fig. 3a).

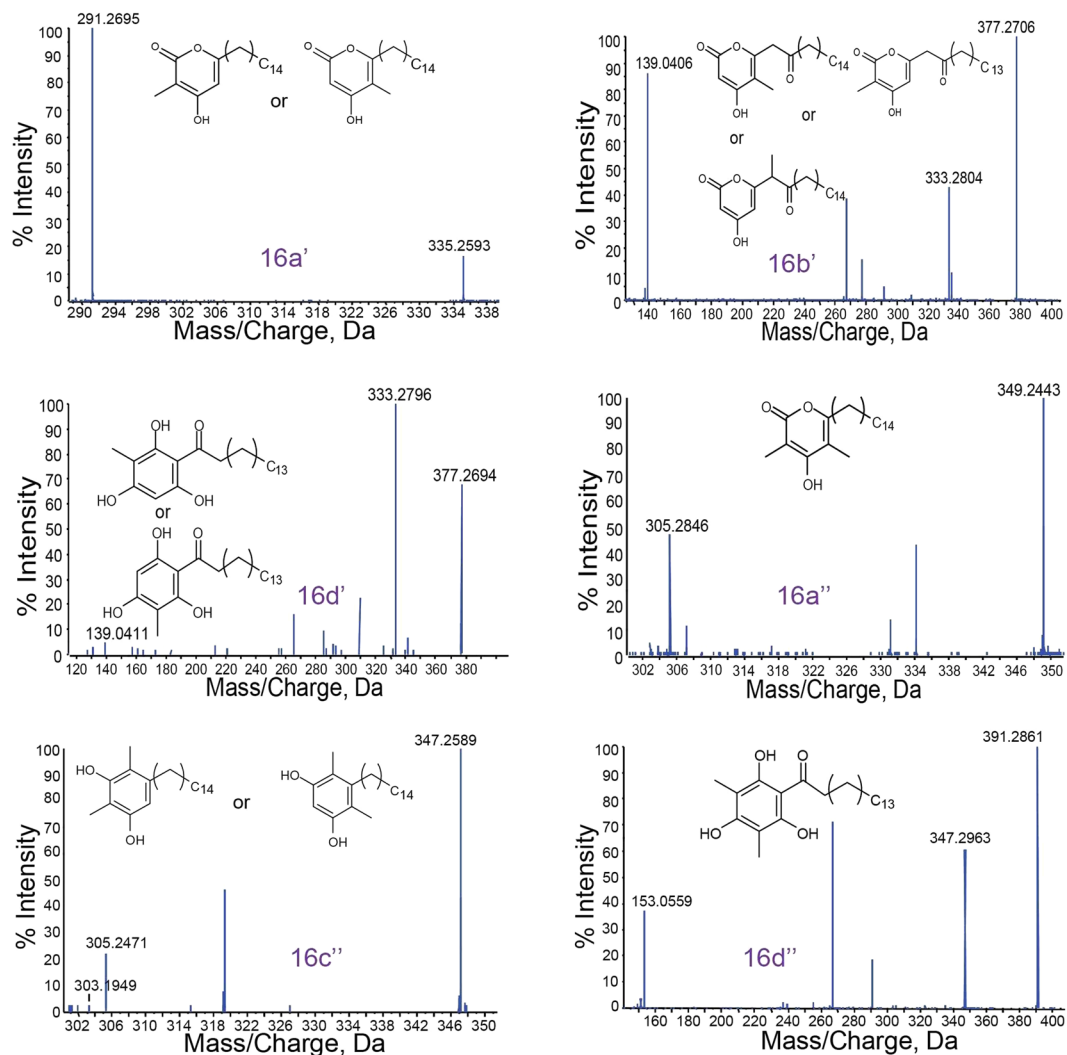
Our *in silico* phylogenetic analyses of MMAR\_2470 and MMAR\_2474 proteins showed a preference of these mycobacterial proteins to group with FtpA, SrsA, and Msmeg PKS10 proteins. Notably, these latter iterative type III PKSs display potentials to selectively condense malonyl- and methylmalonyl-CoA units as extender substrates with variable acyl-CoA starter units to generate a series of methylated polyketide products. We investigated biochemical functions of the two mycobacterial proteins by carrying out cell-free enzymatic assays. Purified recombinant MMAR\_2470 and MMAR\_2474 proteins were incubated independently with malonyl-CoA and/or methylmalonyl-CoA as extender substrates in the presence of long-chain acyl-CoA starter units. Reaction products were extracted and separated using ultra-fast liquid chromatography (UFLC) and chemically characterized by high-resolution mass spectrometry (HRMS). Mass spectrometric analyses corroborated our computational studies and revealed the identity of methylated products for both mycobacterial proteins. Mycobacterial MMAR\_2470 and MMAR\_2474 proteins exhibited the ability to utilize malonyl- or methylmalonyl-CoA alone as extender units to mainly biosynthesize non-methylated and di-methylated triketide  $\alpha$ -pyrone products utilizing each extender unit (Supplementary Figure S1, Supplementary Tables S2, S3). It was interesting to note that reactions with both malonyl- and methylmalonyl-CoA as extender units resulted in more than one reaction products. The UFLC chromatogram of enzymatic assays primed with palmitoyl-CoA (C<sub>16</sub>-CoA) and a combination of malonyl- and methylmalonyl-CoA



**Figure 2.** Homology modeling and comparison of ligand/product binding cavities of MMAR\_2470 and MMAR\_2474 proteins with experimentally characterized bacterial type III PKSs. Homology models of the type III PKSs were generated using Biovia Discovery Studio 4.5. The homology models along with x-ray crystallographic structure of RppA (PDB id 1U0M\_B) were probed for differences in architecture and volumes of the product binding cavities in each protein. The *Cys*, *His* and *Asn* catalytic triad is shown along with the identified cavity. Cavity volumes (in Å<sup>3</sup>) are mentioned below the protein names.

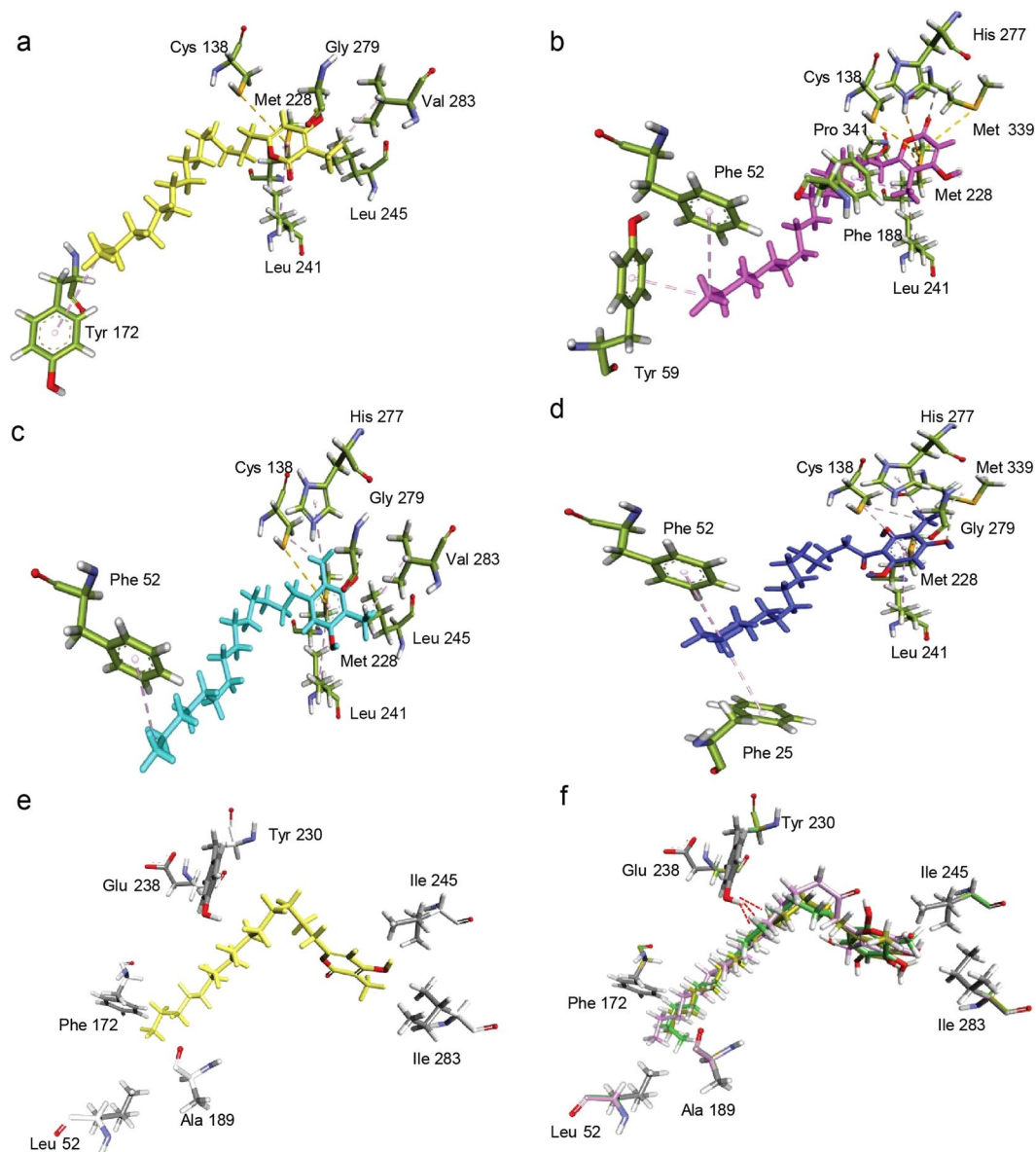


**Figure 3.** Biochemical analyses of MMAR\_2470 and MMAR\_2474 proteins and mass spectrometric characterization of reaction products. **(a)** SDS-PAGE profiles of purified MMAR\_2470 and MMAR\_2474 proteins. **(b)** UFLC chromatograms at 280 nm for reaction products of enzymatic assays carried out independently for both the Mmar type III PKS proteins. Reactions were primed with palmitoyl-CoA (C<sub>16</sub>-CoA) as starter substrate and both malonyl- and methylmalonyl-CoA as extender substrates. Control chromatogram represents reaction carried out in the absence of Mmar proteins. **(c)** Tandem MS/MS analysis of reaction products of MMAR\_2470 protein.



**Figure 4.** Triple-TOF HRMS identification of novel methylated tetraketide reaction products of MMAR\_2474 protein. Tandem MS/MS spectra for each product of MMAR\_2474 protein from assays carried out with palmitoyl-CoA (C<sub>16</sub>-CoA) as starter substrate and both malonyl- and methylmalonyl-CoA as extenders.

can be visualized in Fig. 3b. The UFLC profiles for reactions with MMAR\_2470 and MMAR\_2474 proteins showed three product peaks **1**, **2** and **3** at 280 nm in each case. The product from each peak was subjected to HRMS analyses. In the MMAR\_2470 reaction, a molecular ion peak of [M-H]<sup>-</sup> at *m/z* 335.2602 could be observed in all three chromatographic peaks, while peak **2** in the chromatogram revealed an additional molecular ion of [M-H]<sup>-</sup> at *m/z* 349.2423. A tandem MS/MS analyses confirmed identities of the two products as mono- and di-methylated palmitoyl-triketide  $\alpha$ -pyrones, 16a' and 16a'', respectively (Fig. 3c, Supplementary Table S2). Mass spectrometric identification of similar molecules has been previously reported for reaction products of PKS10 and RpsA proteins from *Msmeg* and *R. centenum*, respectively<sup>36,46</sup>. Both these products could be identified in the three chromatographic peaks observed for MMAR\_2474 reactions. In addition to the ability of MMAR\_2470 protein to extend starter units with two rounds of condensations to form triketide products, MMAR\_2474 protein performed three rounds of condensations with the extender units to generate tetraketide products. Molecular ion peaks of [M-H]<sup>-</sup> at *m/z* 377.2706 in peak 1 and [M-H]<sup>-</sup> at *m/z* 347.2589 in peak 3 could be expected from a mono-methylated palmitoyl-tetraketide  $\alpha$ -pyrone (16b') and a di-methylated palmitoyl-resorcinol (16c'') molecule, respectively. The identity of the products was confirmed using tandem MS/MS analyses as shown in Fig. 4 and Supplementary Table S3. To our surprise, we observed two additional molecular ions of [M-H]<sup>-</sup> at *m/z* 377.2694 and *m/z* 391.2861, respectively, in peaks 2 and 1 of the chromatogram. Tandem MS/MS analysis confirmed the identity of these molecular masses as mono- and di-methylated palmitoyl-phloroglucinols, 16d' and 16d'' (Fig. 4, Supplementary Table S3). Mass spectrometric characterization of non-methylated acyl-phloroglucinols has previously been reported for products of RppA protein from *S. griseus*<sup>29</sup>. The concomitant biosynthesis of methylated alkyl-resorcinol and methylated acyl-phloroglucinol from the same catalytic pocket is a unique observation in the microbial type III PKS protein family. Our biochemical analyses of MMAR\_2470 and MMAR\_2474 proteins from *Mmar* establish the biosynthetic functions of these mycobacterial proteins.



**Figure 5.** Protein-ligand/product docking studies to delineate structural basis for cyclization specificity in mycobacterial type III PKSs. The products of MMAR\_2474 protein were docked in MMAR\_2474 homology model. Docking was performed with (a) mono-methylated triketide  $\alpha$ -pyrone (16a') (b) mono-methylated tetraketide  $\alpha$ -pyrone (16b'), (c) di-methylated alkyl-resorcinol (16c'') and (d) mono-methylated acyl-phloroglucinol (16d'). Interactions with amino acid side chains of MMAR\_2474 protein are shown in each case. To identify the structural basis for lack of tetraketide products from MMAR\_2470 protein, the cavity lining residues in the MMAR\_2474 model were systematically replaced with corresponding amino acid residues in MMAR\_2470 protein. (e) Shows a mono-methylated triketide  $\alpha$ -pyrone (16a') docked without clashes in the reduced cavity of the mutated MMAR\_2474 model. (f) Replacement of *Trp* at position 230 in MMAR\_2474 with corresponding *Tyr* in MMAR\_2470 results in a clash with all three tetraketide products (16b', 16c'' and 16d') docked in the cavity.

### Computational docking studies provide clues to unique bio-functionality of Mmar Type III PKSs.

Biochemical characterization of Mmar type III PKSs revealed that while both proteins actively biosynthesized mono- and di-methylated products, MMAR\_2474 protein exhibited remarkable ability to synthesize methylated alkyl-resorcinols and methylated acyl-phloroglucinols in the same reaction. We probed product formation by docking biochemically identified polyketide products of MMAR\_2474 on to the three-dimensional homology model generated for this protein. Figure 5a–d, respectively show docked 16a', 16b', 16c'' and 16d' reaction products and interacting side chains of cavity lining residues with each of the products. As can be observed from the figure, docked products displayed a variation in the orientation of the alkyl-chain placed in substrate-binding tunnel, leading to differences in interacting residues of the MMAR\_2474 protein. List of interacting residues with each docked substrate is given in Supplementary Table S4. The alkyl-chains showed interactions with hydrophobic *Tyr*

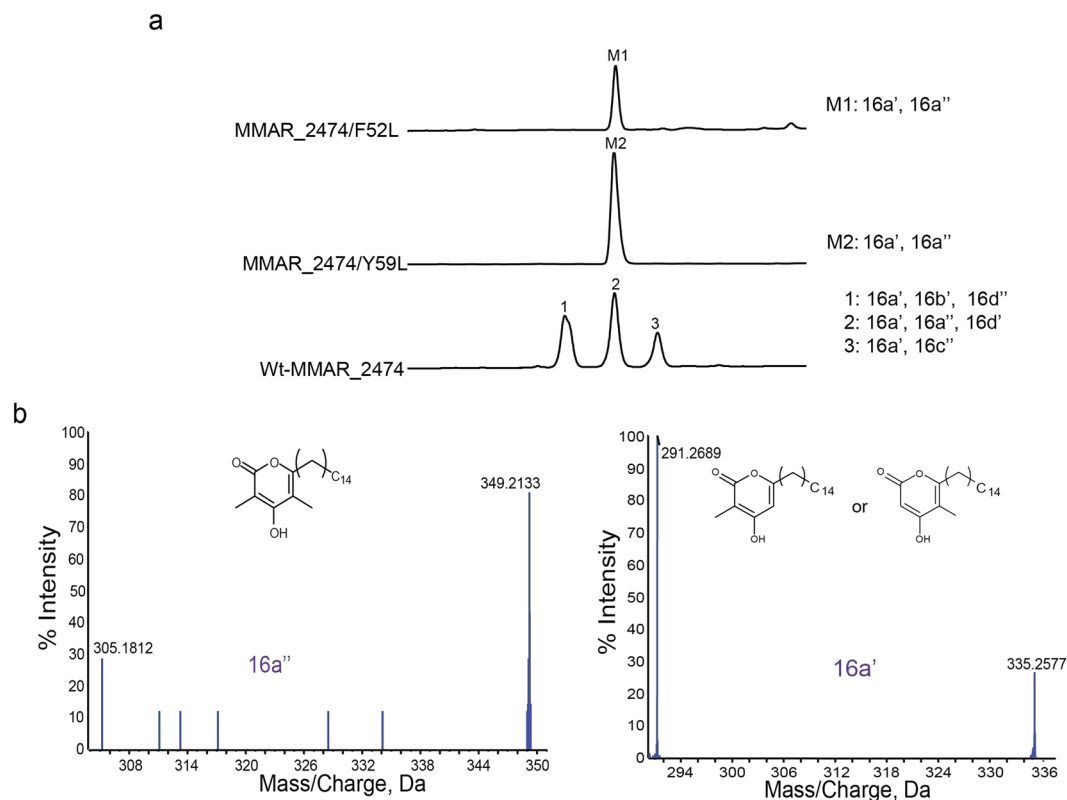
and *Phe* residues of the substrate-binding tunnel. These interactions could stabilize the tetraketide intermediate in an orientation necessary for each type of cyclization to occur. We noticed that a specific *Phe52* residue in the substrate-binding tunnel showed interaction with all three tetraketide products, 16b, 16c<sup>”</sup> and 16d<sup>’</sup> docked in MMAR\_2474 protein. Interestingly, this *Phe52* residue is uniquely present in the substrate-binding tunnel in MMAR\_2474 protein. The corresponding position in functionally related microbial type III PKSs is occupied by a conserved *Leu* residue (Supplementary Figure S2). It is tempting to speculate that the interactions of growing polyketide intermediate with *Phe52* residue in substrate-binding pocket in MMAR\_2474 protein stabilize an orientation of the tetraketide intermediate that facilitates cyclization by three cyclization chemistries, leading to methylated alkyl-tetraketide  $\alpha$ -pyrone, methylated alkyl-resorcinol and methylated acyl-phloroglucinol formation. During dynamic functioning of the protein, additional hydrophobic interactions with *Tyr59* or *Phe25* might aid the tetraketide intermediate in undergoing, respectively, lactonization or claisen condensation in the cyclization pocket. As can be observed from the multiple sequence alignment in Supplementary Figure S2, *Tyr* at position 59 in MMAR\_2474 is conserved in the methylated tetraketide  $\alpha$ -pyrone forming proteins. Our docking studies revealed differences in placement and interactions of cyclized ring system of all docked reaction products in the cyclization pocket. A close inter-link between the substrate-binding pocket and the cyclization cavity has been previously known in PKSIII<sub>Nc</sub> protein from *Neurospora crassa* to influence orientation of the growing polyketide intermediates and dictate cyclization chemistry of reaction products<sup>47</sup>. A close examination of active site residues revealed a *Trp* residue at position 230 in MMAR\_2474 protein. *Trp* at the same position in ArsB protein from *Azotobacter* has been remarkably shown to be essential for biosynthesis of resorcinol products<sup>35</sup>. Conservation of *Trp* at this position has been reported for functionally characterized alkyl-resorcinol forming, bacterial FtpA, SrsA, PKS10 and PKS11 proteins from *M. xanthus*<sup>43</sup>, *S. griseus*<sup>18</sup>, Msmeg and Mtb<sup>36</sup>, respectively.

MMAR\_2470 protein displayed a restricted product profile and biosynthesized methylated alkyl-triketide  $\alpha$ -pyrones as the sole catalytic product, a property not yet reported for microbial type III PKSs. Gene inactivation studies recently implicated CPZ6, a type III PKS from *S. coelicolor*, biosynthesizing triketide  $\alpha$ -pyrones from shuffling sulfate group during biosynthesis of caprazamycin antibiotics<sup>32</sup>. However, the ability of CPZ6 to form additional products was not investigated in this study. In an attempt to delineate the structural basis for MMAR\_2470 enzymatic activity, we carefully compared the active sites of MMAR\_2470 and MMAR\_2474 proteins. While most of the cavity-lining residues were conserved in the proteins, variations were observed at certain positions. Using MMAR\_2474 homology model as backbone, we systematically replaced side chains at *Phe52*, *Tyr172*, *Gly189*, *Trp230*, *Gln238*, *Leu245* and *Val283* of MMAR\_2474 with *Leu*, *Phe*, *Ala*, *Tyr*, *Glu*, *Ile* and *Ile*, respectively, as present at equivalent positions in the MMAR\_2470 protein. As can be observed from Fig. 5e, while these side chain changes drastically reduce cavity volume, docked methylated triketide  $\alpha$ -pyrone fits well in the cavity. It is interesting to note that change from *Trp* to *Tyr* at position 230 exhibited hindrance and clashed with the three docked tetraketide reaction products (Fig. 5f). Presence of a *Tyr* at 230 position in MMAR\_2470 protein likely impedes and limits elongation of the growing poly- $\beta$ -keto chain to a triketide intermediate which is then cyclized and released as an alkyl-triketide  $\alpha$ -pyrone product. Our product docking studies identify several residue positions that could play decisive roles in determining the catalytic products of the Mmar type III PKS proteins.

**Site-directed mutagenesis to investigate active site in MMAR\_2474 protein.** *In silico* product docking studies with the Mmar type III PKSs implicated several residue positions in playing important roles in determining and influencing the product profile. Product docking studies with MMAR\_2474 protein revealed interactions of all three methylated tetraketide products with a specific *Phe52* residue. The methylated tetraketide  $\alpha$ -pyrone product showed additional interactions with *Tyr59* residue (Fig. 5b).

We attempted to investigate the role of these residues in stabilizing the tetraketide intermediate to form various methylated tetraketide products by site-directed mutagenesis. Both *Phe52* and *Tyr59* were mutated to *Leu* side chains as present in other homologous type III PKSs. Mutant proteins were expressed in *E. coli* using wild-type MMAR\_2474 (wt-MMAR\_2474) protein expression conditions. Purified F52L and Y59L mutant proteins were independently assayed for product formation using palmitoyl-CoA (C<sub>16</sub>-CoA) as starter substrate and a combination of malonyl-CoA and methylmalonyl-CoA as extender substrates. Fig. 6a illustrates the UFLC chromatogram of separated reaction products. As can be visualized, reactions with the mutant proteins displayed an altered product profile resulting in single product peaks (M1 or M2) in both the chromatograms. Peaks M1 and M2 corresponded to peak 2 of the chromatogram for wt-MMAR\_2474 protein suggesting possible presence of the products 16a<sup>’</sup>, 16a<sup>”</sup> and 16d<sup>’</sup> as identified above in Fig. 3. HRMS analyses of both M1 and M2 peaks, revealed molecular ions of [M-H]<sup>-</sup> at *m/z* 335.2577 and *m/z* 349.2133 corresponding to mono- and di-methylated palmitoyl-triketide  $\alpha$ -pyrones, 16a<sup>’</sup> and 16a<sup>”</sup>, respectively. Identity of the reaction products was further confirmed by tandem MS/MS analyses (Fig. 6b). Notably, both the mutant proteins were impaired in catalyzing biosynthesis of tetraketide products. Biochemical characterization of mutant proteins corroborated our *in silico* findings and established a catalytically important role of *Phe52* and *Tyr59* side chains for tetraketide product formation in MMAR\_2474 protein.

**Probing physiological importance of MMAR\_2470 and MMAR\_2474 proteins.** Mycobacterial MMAR\_2470 and MMAR\_2474 proteins belong exclusively to the pathogenic mycobacterial species. Interestingly, while MMAR\_2470 shows a restricted product profile, the MMAR\_2474 protein could catalyze the formation of methylated alkyl-resorcinol and methylated acyl-phloroglucinol products from the same catalytic pocket. We investigated the physiological relevance of the two mycobacterial type III PKS proteins by independently over-expressing these in a genetically modified, type III *pks* knockout strain ( $\Delta pks10$ ) of Msmeg. Msmeg genome harbors a single type III *pks* gene, *pks10*, in a three-gene operon. In a recent study, PKS10 protein from Msmeg was characterized to biosynthesize long-chain alkyl-resorcinols, which upon *o*-methylation and oxidation, led to the formation of novel polyketide quinones. These molecules facilitate mycobacterial survival



**Figure 6.** Biochemical characterization of F52L and Y59L mutants of MMAR\_2474 protein. **(a)** UFLC chromatograms at 280 nm for reaction products of enzymatic assays carried out with palmitoyl-CoA (C<sub>16</sub>-CoA) as starter substrate and malonyl- and methylmalonyl-CoA as extenders. Reactions carried out with wild-type MMAR\_2474 protein were used as control. **(b)** Tandem MS/MS analysis of reaction products of F52L and Y59L mutant proteins.

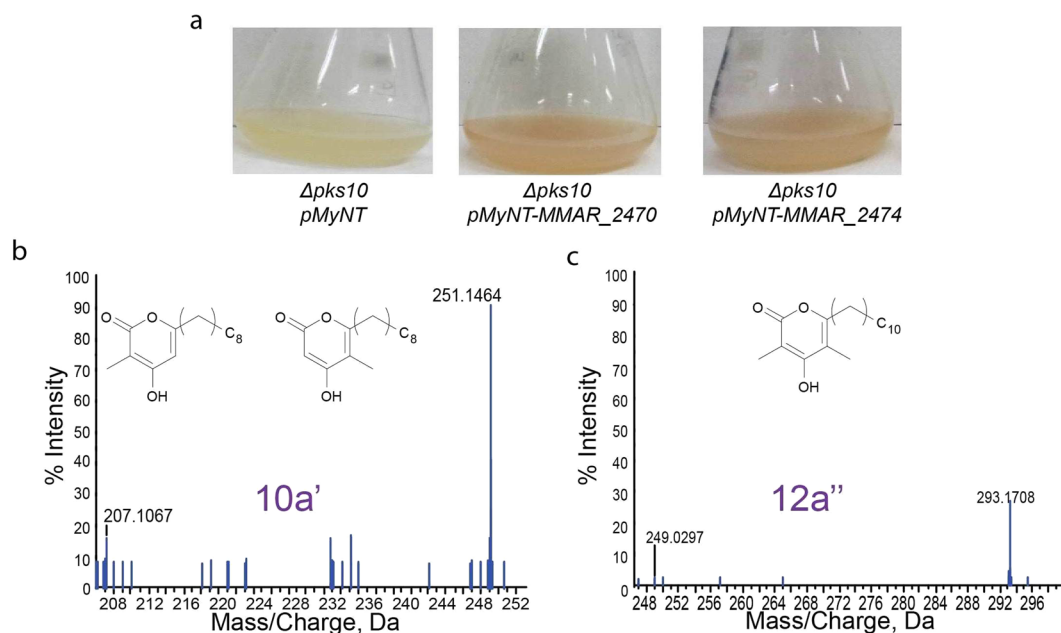
in oxygen-deficient niches by participating in electron transport during anaerobic respiration. *mmar\_2470* and *mmar\_2474* genes were cloned in inducible mycobacterial expression vector *pMyNT* and expressed in  $\Delta pks10$  Msmeg mutant strain<sup>36</sup>. Planktonic cultures over-expressing Mmar type III PKSs displayed brownish coloration upon induction with acetamide (Fig. 7a). Filtrates of these cultures were extracted and subjected to HRMS metabolomic analyses. As can be seen in Fig. 7b,c, tandem MS/MS analyses identified mono- and di-methylated alkyl-triketide  $\alpha$ -pyrone products, respectively, in extracts of  $\Delta pks10$  Msmeg cultures over-expressing the MMAR\_2470 and MMAR\_2474 proteins (Supplementary Tables S2, S3).

We further probed a possible role of the Mmar type III PKS proteins in facilitating mycobacterial growth in biofilms. Mycobacteria are naturally known to exist as stationary aggregates in biofilms. Development of biofilm pellicle was carried out for  $\Delta pks10$  Msmeg strains over-expressing *pMyNT*-MMAR\_2470 or *pMyNT*-MMAR\_2474 proteins and analyzed by scanning electron microscopy (SEM). As reported earlier, the mutant  $\Delta pks10$  Msmeg strain produced a fragile pellicle as compared to the reticulated architecture of the wild-type Msmeg biofilm<sup>36</sup>. As a compelling observation, over-expression of MMAR\_2474 protein in  $\Delta pks10$  Msmeg mutant strain restored the porous community structure of the biofilm to moderately resemble the wild-type Msmeg pellicle (Fig. 8). Similar restoration of the biofilm phenotype has been reported by Anand *et al.*, upon complementation of the  $\Delta pks10$  Msmeg biofilms with alkyl-benzoquinones generated from modification of alkyl-resorcinol product of PKS10 protein in Msmeg<sup>36</sup>. As MMAR\_2474 protein biosynthesizes methylated alkyl-resorcinol products, we investigated the presence of polyketide quinones in the biofilm of  $\Delta pks10$ -*pMyNT*-MMAR\_2474 strain by analyzing the biofilm pellicle extracts by high resolution metabolomics analyses. However, despite best efforts, we could not observe the quinone molecules. Over-expression of MMAR\_2474 protein in the  $\Delta pks10$  Msmeg strain could not regenerate the biofilm pellicle, owing probably to the absence of methylated alkyl-resorcinol products in this strain. Our results suggest physiological importance of type III polyketide metabolites and implicate Mmar type III PKSs in assisting mycobacterial survival in stationary aggregates.

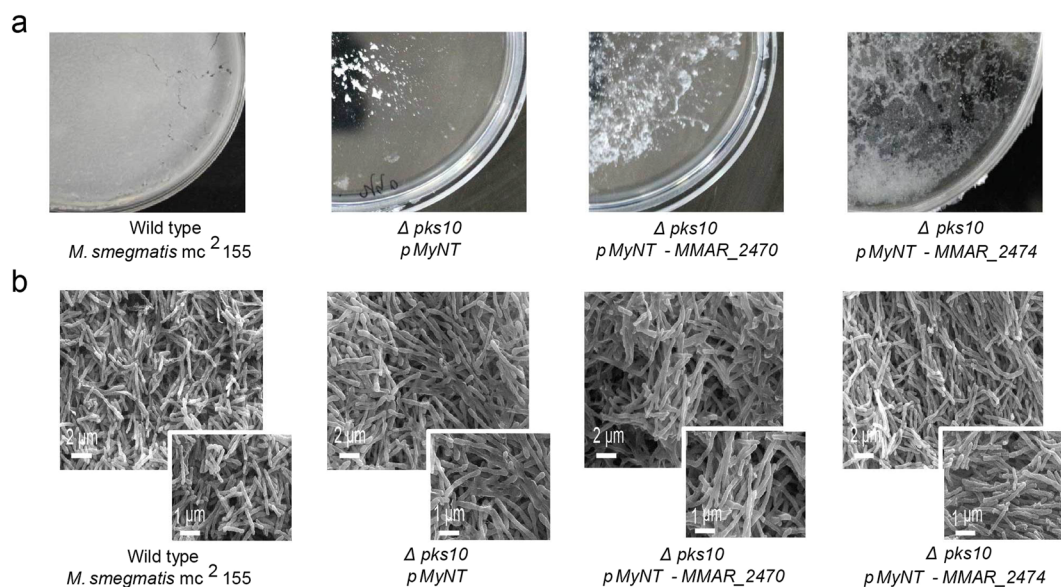
## Discussion

Mycobacterial polyketide gene clusters are involved in the biosynthesis of architecturally and functionally unique lipids that play major roles as effectors of virulence<sup>6,48</sup>. Several studies have provided understanding on PKS mediated mechanisms employed by pathogenic mycobacterial species to armor against the host defenses and establish virulent pathology. Efforts, however, are on to identify factors that determine persistence and dormancy in mycobacteria<sup>8–10</sup>. The existence of mycobacteria in self-assembled multi-cellular and structurally robust biofilms in host provides unique advantages of antimicrobial resistance<sup>49</sup>, host immune modulation<sup>50</sup>, and persistence<sup>8,51</sup>.





**Figure 7.** Over-expression of MMAR\_2470 and MMAR\_2474 proteins in  $\Delta pks10$  *Mycobacterium smegmatis* strain. **(a)** Planktonic cultures over-expressing MMAR\_2470 and MMAR\_2474 proteins show brownish coloration. Triple-TOF HRMS metabolomic analysis of extracted culture filtrates of planktonic cultures over-expressing the Mmar type III PKSs identified **(b)** mono- and **(c)** di-methylated alkyl-triketide  $\alpha$ -pyrones.



**Figure 8.** Biofilm development of  $\Delta pks10$  *Mycobacterium smegmatis* strains complemented with Mmar type III PKSs. **(a)** Pellicle development and, **(b)** scanning electron microscopy (SEM) of biofilms of wild-type *M. smegmatis* (Msmeg),  $\Delta pks10$ -pMyNT Msmeg,  $\Delta pks10$ -pMyNT-MMAR\_2470 Msmeg and  $\Delta pks10$ -pMyNT-MMAR\_2474 Msmeg strains.

Type III polyketide synthase genomic clusters from several bacteria have been lately associated with diverse bio-functionalities including, conferring antibiotic resistance in *Streptomyces*<sup>18</sup>, development of exine in dormant *Azotobacter* cells<sup>35</sup>, and survival capability of mycobacteria in anaerobic biofilms<sup>36</sup>.

In this study, we have functionally characterized two novel type III PKSs, MMAR\_2470 and MMAR\_2474 proteins from *Mycobacterium marinum* (Mmar) by a combination of computational, biochemical, mutational and physiological studies. Genes for both these proteins are present in a unique 2x type III *pks* genomic cluster, that includes genes for other type I modular/iterative PKSs. Presence of two type III *pks*s in a single genomic cluster has been previously reported only in the case of *ars* cluster in *Azotobacter*<sup>35</sup>. Our comparative genome analyses

revealed a conservation of Mmar 2x type III *pks* cluster exclusively in the genomes of pathogenic mycobacterial species (Supplementary Table S5). It is interesting to note that this genomic organization is not observed in any other genera outside *Mycobacterium*, suggesting an extremely specialized function of the cluster in mycobacterial pathogenesis. Functional characterization of MMAR\_2470 and MMAR\_2474 proteins corroborated our phylogenetic findings and both the proteins could accept malonyl-CoA as well as methylmalonyl-CoA to biosynthesize mono- and di-methylated polyketide products (Supplementary Figure S3). MMAR\_2474 exceptionally displayed the capacity to catalyze concomitant biosynthesis of methylated acyl-phloroglucinols and methylated alkyl-resorcinols. Our three-dimensional homology modeling and product docking studies identified interactions of tetraketide products with specific *Phe52* and *Tyr59* residues uniquely positioned in the substrate-binding tunnel in MMAR\_2474 protein. Systematic site-directed mutagenesis studies revealed the inability of F52L and Y59L mutant proteins to synthesize tetraketide products, establishing their involvement in dictating the product profile of MMAR\_2474 protein.

Mycobacterial pathogenicity is associated with the ability of the pathogen to grow and differentiate in multi-cellular assemblages, enveloped in extracellular matrix that provides a secure antimicrobial-resistant niche for propagation and development of persistence<sup>37,49,52,53</sup>. Alkyl-resorcinols from PKS10 in *Msmeg*, upon *o*-methylation and oxidation, form polyketide quinones that have been recently shown to facilitate growth and survival in anaerobic biofilms. Notably, over-expression of MMAR\_2474 protein in a biofilm-growth defective  $\Delta pks10$  *Msmeg* mutant strain revived growth in stationary biofilms. While the polyketide quinones could not be detected in the extracts of revived biofilms, the biosynthetic activity of the over-expressing MMAR\_2474 protein likely provided the alkyl-resorcinol precursor needed for quinone formation in the revived biofilms. Orthologous PKS11 protein in *Mtb* biosynthesizes similar precursors and has been implicated in biofilm generation<sup>36</sup>. Investigations in this study unravel a unique 2x type III *pks* genomic organization exclusively conserved in pathogenic mycobacterial species and emphasize the importance of type III polyketide synthases in aiding mycobacterial survival in complex communities. Our findings reveal unique insights on biosynthetic potentials of type III polyketide synthases from *M. marinum* and identify a new class of type III PKSs conserved in pathogenic mycobacteria.

## Methods

**Bacterial strains and materials.** *Escherichia coli* XL-1 blue and BL21 (DE3) strains were used as cloning and expression strains, respectively. The sequenced *Mycobacterium marinum* (strain ATCC BAA-535/M) was kindly provided by Y. Singh (IGIB, India) and was used for genomic DNA isolation. The type III *pks* knockout, a  $\Delta pks10$  strain of *Mycobacterium smegmatis* mc<sup>2</sup>155 was a kind gift from R. S. Gokhale (NII, India) and was used for over-expressing *M. marinum* type III PKSs. The mycobacterial expression vector, *pMyNT* was kindly donated by Annabel Parret and Matthias Wilmanns (Addgene plasmid # 42191). Acyl-CoA starters and extender substrates were procured from Sigma. Middlebrook 7H9, Middlebrook 7H11 media and Sauton's Fluid Media Base were purchased from Himedia. Restriction endonucleases and PCR master mix were purchased from New England Biolabs while PCR cleanup kit and Ni<sup>2+</sup>-NTA was purchased from Qiagen. Quik Change site-directed mutagenesis-XL kit was procured from Stratagene. All analytical, HPLC and MS grade solvents were purchased from Merck and Sigma.

**Cloning, expression, and purification of MMAR\_2470 and MMAR\_2474 proteins.** Type III PKS genes, *mmar\_2470* and *mmar\_2474*, were amplified by PCR from genomic DNA of *M. marinum* using gene-specific forward primers (5'TTCATATGAGCGTCATCGCAGGTGTG3') and (5'TTCATATGAGCGTTATCGCAGGTGTG3') for *mmar\_2470* and *mmar\_2474*, respectively, containing NdeI restriction enzyme site and reverse primers (5'GGGAATTCCTGCGCAACGCAGCAGTACTAG3') and (5'AAGAATTCCTGCGCAGCGCAGTAACACCAG3') for *mmar\_2470* and *mmar\_2474*, respectively, containing EcoRI restriction enzyme site and cloned into pBluescript SK (+) (Stratagene). The authenticity of clones was confirmed by automated nucleotide sequencing. These *mmar\_2470* and *mmar\_2474* genes were sub-cloned into a *pET28c* (Novagen) and *pET21c* (Novagen) expression vectors, respectively for protein purification. The MMAR\_2470 protein was expressed as both N- and C-terminal hexa-histidine-tagged protein whereas MMAR\_2474 was expressed as C-terminal hexa-histidine-tagged protein in the BL21/ (DE-3) strain of *E. coli*. The recombinant *E. coli* cells harboring the expression plasmids were grown in Luria Bertini broth at 30 °C until an OD 600 of 0.5 units was reached. The culture was induced with 0.5 mM isopropyl-1-thio- $\beta$ -D galactopyranoside (IPTG) and further incubated in a shaker at 22 °C for 16–20 h. After harvesting, the cells were re-suspended in lysis buffer (50 mM Tris pH 8.0, 10% glycerol, 0.15 M NaCl) and disrupted using 8 cycles of sonication at 30% amplitude. Cell debris was removed by centrifugation at 13,000 rpm for 30 min at 4 °C. 1 ml of Ni<sup>2+</sup>-NTA slurry per liter of culture was added to the supernatant and was incubated at 4 °C for 1 hour. 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 an increasing concentration of imidazole (5 mM to 250 mM). Expression and purification of both MMAR\_2470 and MMAR\_2474 proteins were performed using modified protocol mentioned in earlier literature<sup>27,36</sup>.

**Enzymatic assay and product characterization.** The standard reaction was carried out using 100  $\mu$ M starter CoA and 50  $\mu$ M malonyl-CoA and/or 50  $\mu$ M methylmalonyl-CoA extender with 50  $\mu$ g of protein at 30 °C for 120 min. The reaction was quenched with 5% acetic acid. Products were extracted with 2x 300  $\mu$ l ethyl acetate and dried under vacuum. Products from at least three reactions were pooled after vacuum drying by dissolving in 20  $\mu$ l of methanol. Products were resolved on an analytical C5-reverse phase-UFLC column using a 5–100% gradient of acetonitrile containing 1% acetic acid for 60 min. Resolved products were further characterized using SCIEX Triple-Time Of Flight (Triple-TOF) 6600 HRMS.

**Site-Directed Mutagenesis.** Site-directed mutagenesis was carried out using Quik Change site-directed mutagenesis-XL kit from Stratagene. The recombinant plasmid containing the wild-type *mmar\_2474* gene was used as the DNA template. The plasmids containing the mutant genes were used to transform *E. coli* BL21 (DE3), and recombinant colonies were used for protein expression. Mutant proteins were expressed and purified using experimental conditions same as used for wild-type MMAR\_2474 protein.

***M. marinum* type III *pks* complementation of  $\Delta pks10$  *Mycobacterium smegmatis* mutant strain and metabolomic analyses of planktonic culture and biofilm extracts.** The *mmar\_2470* and *mmar\_2474* genes were cloned into a mycobacterial expression vector, *pMyNT*, using NdeI site. The  $\Delta pks10$  mutant strain of *M. smegmatis* mc<sup>2</sup>155<sup>36</sup> was electroporated with *pMyNT-mmar\_2470* and *pMyNT-mmar\_2474* plasmids and used for over-expression of the *M. marinum* type III PKSs. Transformed colonies were screened in 7H11 media with hygromycin and kanamycin as selectable markers. Planktonic cultures were grown in 7H9 media supplemented with ADC, 0.18% glycerol and 0.05% tween 80 with 100  $\mu$ g/ml hygromycin and 50  $\mu$ g/ml kanamycin and were induced with 1 mM acetamide when the OD at 600 nm reached 0.8. Biofilm pellicles were grown in Sauton's media supplemented with 2% glucose and 2% glycerol and with 100  $\mu$ g/ml hygromycin and 50  $\mu$ g/ml kanamycin and induced with 1 mM acetamide at the start of the culture and incubated at 37 °C for 14 days. Cells from grown planktonic culture and biofilm pellicle were harvested and resuspended in 100 mM Tris-Cl, pH 8.0. After resuspension the culture supernatants and pellicle were acidified with acetic acid to pH 4 and polyketide derived lipid molecules were extracted by adding a double volume of ethyl acetate. The organic layer containing the extracted products was separated and concentrated using multi-vapor. The extracts were dissolved in methanol and products were characterized using SCIEX Triple-TOF 6600 HRMS using multiple-reaction-monitoring (MRM) and information dependent acquisition (IDA) methods.

**Scanning electron microscopy.** Scanning electron microscopy was performed for the biofilm pellicle samples using standard protocols. Briefly, the pellicle was fixed over-night with a fixative containing 2.5% (vol/vol) glutaraldehyde and 4% paraformaldehyde and washed with sodium cacodylate buffer at room temperature. The pellicle was then osmicated with 1% Osmium tetroxide (wt/vol) in 0.1 M sodium cacodylate buffer for 40 min at room temperature in dark and washed with milli-Q water at room temperature. Pellicle was then gradually dehydrated with increasing concentrations of ethanol followed for 15 min and incubated in dark with hexamethyldisilazane (HMDS). The samples were mounted on stubs, sputter coated with gold particles for 300 s, and imaged on a Zeiss-EVO LS15 Scanning Electron Microscope (Carl Zeiss).

**In silico studies.** The sequences of mycobacterial type III PKS clusters were retrieved in FASTA format from UniProt. Subsequently, a homology search was done using NCBI BLASTp program<sup>54</sup> to identify potentially related proteins. Cluster conservation was identified using UniProt and microbrowser search. Experimentally characterized type III PKS amino acid sequences were downloaded from UniProt. The phylogenetic tree was constructed using maximum likelihood method in MEGA7<sup>55</sup>.

Homology modeling of type III PKSs was carried out using Biovia Discovery Studio version 4.5. FtpA, MMAR\_2470, MMAR\_2474 and SrsA proteins were modeled using a 4JAO\_D<sup>42</sup> template whereas DpgA model was generated using the 1TED\_B<sup>21</sup> template (Supplementary Table S1). 1U0M\_B<sup>20</sup> was used for PhlD model preparation (Supplementary Table S1). Ligand/product binding pockets and respective cavity volumes of homology models of type III PKSs were identified using Biovia Discovery Studio 4.5.

Predicted-ligand/product library was generated using ChemDraw software and ligands/products were prepared in Biovia Discovery Studio 4.5 along with protein preparation for protein-ligand/product docking studies. Protein ligand/product docking was done using CDOCKER of Biovia Discovery Studio 4.5. CDOCKER is the CHARMM-based docking method that generates highly accurate docked poses<sup>56</sup>.

**Data availability.** All data generated or analysed during this study are available from the corresponding author on reasonable request.

## References

- Brosch, R., Gordon, S. V., Eiglmeier, K., Garnier, T. & Cole, S. T. Comparative genomics of the leprosy and tubercle bacilli. *Research in microbiology* **151**, 135–142 (2000).
- Palmer, M. V., Welsh, M. D. & Hostetter, J. M. Mycobacterial diseases of animals. *Veterinary medicine international* **2011** (2011).
- Stinear, T. P. *et al.* Insights from the complete genome sequence of *Mycobacterium marinum* on the evolution of *Mycobacterium tuberculosis*. *Genome research* **18**, 729–741 (2008).
- Forrellad, M. A. *et al.* Virulence factors of the *Mycobacterium tuberculosis* complex. *Virulence* **4**, 3–66 (2013).
- Gordon, S. & Andrew, P. Mycobacterial virulence factors. *Journal of Applied Microbiology* **81** (1996).
- Trivedi, O. A. *et al.* Dissecting the mechanism and assembly of a complex virulence mycobacterial lipid. *Molecular cell* **17**, 631–643 (2005).
- Agarwal, N., Lamichhane, G., Gupta, R., Nolan, S. & Bishai, W. R. Cyclic AMP intoxication of macrophages by a *Mycobacterium tuberculosis* adenylate cyclase. *Nature* **460**, 98–102 (2009).
- Pandey, A. K. & Sasseti, C. M. Mycobacterial persistence requires the utilization of host cholesterol. *Proceedings of the National Academy of Sciences* **105**, 4376–4380 (2008).
- Meena, L. S. Survival mechanisms of pathogenic *Mycobacterium tuberculosis* H37Rv. *The FEBS journal* **1277**, 2416–2427 (2010).
- Boshoff, H. I., Singh, R. & Barry, C. E. Virulence and Persistence Mechanisms of *Mycobacterium tuberculosis*. *Handbook of Tuberculosis*, 151–191 (2008).
- Cole, S. *et al.* Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544 (1998).
- Kneitz, S. *et al.* Pathogenomics: Genome analysis of pathogenic microbes (2006).
- Yu, D., Xu, F., Zeng, J. & Zhan, J. Type III polyketide synthases in natural product biosynthesis. *IUBMB life* **64**, 285–295 (2012).
- Ghosh, R. *et al.* Dissecting the Functional Role of Polyketide Synthases in *Dictyostelium discoideum* Biosynthesis of The Differentiation Regulating Factor 4-Methyl-5-Pentylbenzene-1, 3-Diol. *Journal of Biological Chemistry* **283**, 11348–11354 (2008).

15. Bangera, M. G. & Thomashow, L. S. Identification and characterization of a gene cluster for synthesis of the polyketide antibiotic 2, 4-diacetylphloroglucinol from *Pseudomonas fluorescens* Q2-87. *Journal of Bacteriology* **181**, 3155–3163 (1999).
16. Li, T.-L., Choroba, O. W., Hong, H., Williams, D. H. & Spencer, J. B. Biosynthesis of the vancomycin group of antibiotics: characterisation of a type III polyketide synthase in the pathway to (S)-3, 5-dihydroxyphenylglycine. Electronic supplementary information (ESI) available: electrospray mass spectrum of DhpaS. See, <http://www.rsc.org/suppdata/cc/b1/b106638b>. *Chemical Communications*, 2156–2157 (2001).
17. Chemler, J. A. *et al.* Biochemical and structural characterization of germicidin synthase: analysis of a type III polyketide synthase that employs acyl-ACP as a starter unit donor. *Journal of the American Chemical Society* **134**, 7359–7366 (2012).
18. Funabashi, M., Funai, N. & Horinouchi, S. Phenolic lipids synthesized by type III polyketide synthase confer penicillin resistance on *Streptomyces griseus*. *Journal of Biological Chemistry* **283**, 13983–13991 (2008).
19. Sun, L. *et al.* Identification of a type III polyketide synthase involved in the biosynthesis of spiroloxine. *Applied microbiology and biotechnology* **100**, 7103–7113 (2016).
20. Austin, M. B. *et al.* Crystal structure of a bacterial type III polyketide synthase and enzymatic control of reactive polyketide intermediates. *Journal of Biological Chemistry* **279**, 45162–45174 (2004).
21. Sankaranarayanan, R. *et al.* A novel tunnel in mycobacterial type III polyketide synthase reveals the structural basis for generating diverse metabolites. *Nature structural & molecular biology* **11**, 894–900 (2004).
22. Lim, Y. P., Go, M. K. & Yew, W. S. Exploiting the Biosynthetic Potential of Type III Polyketide Synthases. *Molecules* **21**, 806 (2016).
23. Satou, R. *et al.* Structural Basis for Cyclization Specificity of Two Azotobacter Type III Polyketide Synthases A Single Amino Acid Substitution Reverses Their Cyclization Specificity. *Journal of Biological Chemistry* **288**, 34146–34157 (2013).
24. Katsuyama, Y. & Ohnishi, Y. Type III polyketide synthases in microorganisms. *Methods Enzymol* **515**, 359–377 (2012).
25. Nobutaka, F., Ohnishi, Y., Ebizuka, Y. & Horinouchi, S. Alteration of reaction and substrate specificity of a bacterial type III polyketide synthase by site-directed mutagenesis. *Biochemical Journal* **367**, 781–789 (2002).
26. Chan, Y. A., Podevels, A. M., Kevany, B. M. & Thomas, M. G. Biosynthesis of polyketide synthase extender units. *Natural product reports* **26**, 90–114 (2009).
27. Saxena, P., Yadav, G., Mohanty, D. & Gokhale, R. S. A new family of type III polyketide synthases in *Mycobacterium tuberculosis*. *Journal of Biological Chemistry* **278**, 44780–44790 (2003).
28. Tsai, S.-C. A fine balancing act of type III polyketide synthase. *Chemistry & biology* **11**, 1177–1178 (2004).
29. Funai, N., Ohnishi, Y., Ebizuka, Y. & Horinouchi, S. Properties and substrate specificity of RppA, a chalcone synthase-related polyketide synthase in *Streptomyces griseus*. *Journal of Biological Chemistry* **277**, 4628–4635 (2002).
30. Miyana, A., Funai, N., Awakawa, T. & Horinouchi, S. Direct transfer of starter substrates from type I fatty acid synthase to type III polyketide synthases in phenolic lipid synthesis. *Proceedings of the National Academy of Sciences* **105**, 871–876 (2008).
31. Arora, P., Vats, A., Saxena, P., Mohanty, D. & Gokhale, R. S. Promiscuous fatty acyl CoA ligases produce acyl-CoA and acyl-SNAC precursors for polyketide biosynthesis. *Journal of the American Chemical Society* **127**, 9388–9389 (2005).
32. Tang, X. *et al.* A two-step sulfation in antibiotic biosynthesis requires a type III polyketide synthase. *Nature chemical biology* **9**, 610–615 (2013).
33. Nakano, C., Ozawa, H., Akanuma, G., Funai, N. & Horinouchi, S. Biosynthesis of aliphatic polyketides by type III polyketide synthase and methyltransferase in *Bacillus subtilis*. *Journal of bacteriology* **191**, 4916–4923 (2009).
34. Nakano, C., Funai, N., Ohnishi, Y. & Horinouchi, S. The O-methyltransferase SrsB catalyzes the decarboxylative methylation of alkylresorcylic acid during phenolic lipid biosynthesis by *Streptomyces griseus*. *Journal of bacteriology* **194**, 1544–1551 (2012).
35. Funai, N., Ozawa, H., Hirata, A. & Horinouchi, S. Phenolic lipid synthesis by type III polyketide synthases is essential for cyst formation in *Azotobacter vinelandii*. *Proceedings of the National Academy of Sciences* **103**, 6356–6361 (2006).
36. Anand, A. *et al.* Polyketide quinones are alternate intermediate electron carriers during mycobacterial respiration in oxygen-deficient niches. *Molecular cell* **60**, 637–650 (2015).
37. Ojha, A. K. *et al.* Growth of *Mycobacterium tuberculosis* biofilms containing free mycolic acids and harbouring drug-tolerant bacteria. *Molecular microbiology* **69**, 164–174 (2008).
38. Chen, L. & Wen, Y.-m. The role of bacterial biofilm in persistent infections and control strategies. *International journal of oral science* **3**, 66–73 (2011).
39. Ramakrishnan, L. Using *Mycobacterium marinum* and its hosts to study tuberculosis. *Current Science* 82–92 (2004).
40. Sasseti, C. M., Boyd, D. H. & Rubin, E. J. Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular microbiology* **48**, 77–84 (2003).
41. Griffin, J. E. *et al.* High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS pathogens* **7**, e1002251 (2011).
42. Gokulan, K. *et al.* Crystal structure of *Mycobacterium tuberculosis* polyketide synthase 11 (PKS11) reveals intermediates in the synthesis of methyl-branched alkylpyrones. *Journal of Biological Chemistry* **288**, 16484–16494 (2013).
43. Hayashi, T., Kitamura, Y., Funai, N., Ohnishi, Y. & Horinouchi, S. Fatty Acyl-AMP Ligase Involvement in the Production of Alkylresorcylic Acid by a *Myxococcus xanthus* Type III Polyketide Synthase. *Chembiochem* **12**, 2166–2176 (2011).
44. Pfeifer, V. *et al.* A Polyketide Synthase in Glycopeptide Biosynthesis The Biosynthesis of The Non-Proteinogenic Amino Acid (S)-3, 5-Dihydroxyphenylglycine. *Journal of Biological Chemistry* **276**, 38370–38377 (2001).
45. Achkar, J., Xian, M., Zhao, H. & Frost, J. Biosynthesis of phloroglucinol. *Journal of the American Chemical Society* **127**, 5332–5333 (2005).
46. Awakawa, T. *et al.* 4-Hydroxy-3-methyl-6-(1-methyl-2-oxoalkyl) pyran-2-one Synthesis by a Type III Polyketide Synthase from *Rhodospirillum centenum*. *Chembiochem* **14**, 1006–1013 (2013).
47. Goyal, A. *et al.* Structural insights into biosynthesis of resorcinolic lipids by a type III polyketide synthase in *Neurospora crassa*. *Journal of structural biology* **162**, 411–421 (2008).
48. Gokhale, R. S., Saxena, P., Chopra, T. & Mohanty, D. Versatile polyketide enzymatic machinery for the biosynthesis of complex mycobacterial lipids. *Natural product reports* **24**, 267–277 (2007).
49. Falkinham, J. O. III Growth in catheter biofilms and antibiotic resistance of *Mycobacterium avium*. *Journal of medical microbiology* **56**, 250–254 (2007).
50. Queiroz, A. & Riley, L. W. Bacterial immunostat: *Mycobacterium tuberculosis* lipids and their role in the host immune response. *Revista da Sociedade Brasileira de Medicina Tropical* **50**, 9–18 (2017).
51. Pang, J. M. *et al.* The polyketide Pks1 contributes to biofilm formation in *Mycobacterium tuberculosis*. *Journal of bacteriology* **194**, 715–721 (2012).
52. Marsollier, L. *et al.* Impact of *Mycobacterium ulcerans* biofilm on transmissibility to ecological niches and Buruli ulcer pathogenesis. *PLoS pathogens* **3**, e62 (2007).
53. Islam, M. S., Richards, J. P. & Ojha, A. K. Targeting drug tolerance in mycobacteria: a perspective from mycobacterial biofilms. *Expert review of anti-infective therapy* **10**, 1055–1066 (2012).
54. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of molecular biology* **215**, 403–410 (1990).
55. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular biology and evolution* **33**, 1870–1874 (2016).
56. Wu, G., Robertson, D. H., Brooks, C. L. & Vieth, M. Detailed analysis of grid-based molecular docking: A case study of CDOCKER—A CHARMM-based MD docking algorithm. *Journal of computational chemistry* **24**, 1549–1562 (2003).

## Acknowledgements

P.S. gratefully acknowledges Innovative Young Biotechnologist Award (IYBA) of Department of Biotechnology (DBT), Ministry of Science and Technology, Government of India and South Asian University (SAU) Startup Research Grant, SAU, India for providing funds and support for this study. A.P. acknowledges CSIR for Senior Research Fellowship. We are thankful to Rajesh S. Gokhale (NII, India) for kindly providing the  $\Delta pks10$  *Mycobacterium smegmatis* mutant strain. We thank Sciex, Gurugram, India for Mass Spectrometry data collection and the Scanning Electron Microscopy (SEM) Facility at CSIR-Institute of Genomics and Integrative Biology for SEM related studies and Ms. Hemlata for the technical help in SEM analyses. We thank Kailash Prasad Prajapati for his assistance in bioinformatic analyses.

## Author Contributions

P.S. conceptualized and designed the experiments. P.S., A.P., S.G., G.R.G. and M.K. performed the experiments. R.B. provided technical support and prepared figures. P.S., A.P., S.G. and G.R.G. analyzed and interpreted the data. P.S., A.P. and S.G. wrote the manuscript. P.S. supervised the study. All the authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-24980-1>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018