

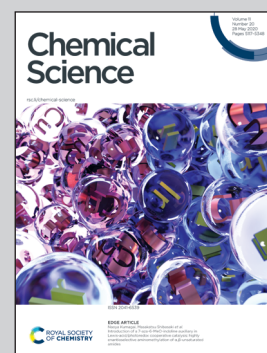


Showcasing research from Dr. Gerald Lackner's laboratory, Leibniz Institute for Natural Product Research and Infection Biology and Friedrich Schiller University, Germany.

Structure elucidation of the redox cofactor mycofactocin reveals oligo-glycosylation by MftF

A team of researchers around Dr. Gerald Lackner discovered mycofactocin, a redox cofactor from pathogenic mycobacteria involved in alcohol metabolism. It is the first cofactor whose existence had been postulated on the basis of mere bioinformatics and which had remained cryptic for almost a decade. Intensive metabolomics studies eventually led to the identification of the molecule. The elucidation of its chemical structure revealed an unusual sugar decoration. Last but not least, mycofactocin could be a promising target for fighting mycobacterial infections. The authors cordially thank Luo Yu for the artwork.

As featured in:



See Gerald Lackner *et al.*,  
*Chem. Sci.*, 2020, 11, 5182.

Cite this: *Chem. Sci.*, 2020, **11**, 5182

All publication charges for this article have been paid for by the Royal Society of Chemistry

# Structure elucidation of the redox cofactor mycofactocin reveals oligo-glycosylation by MftF†

Luis Peña-Ortiz,<sup>ab</sup> Ana Patricia Graça,<sup>ab</sup> Huijuan Guo,<sup>c</sup> Daniel Braga,<sup>ab</sup> Tobias G. Köllner,<sup>d</sup> Lars Regestein,<sup>e</sup> Christine Beemelmans<sup>c</sup> and Gerald Lackner<sup>ab\*</sup>

Mycofactocin (MFT) is a redox cofactor belonging to the family of ribosomally synthesized and post-translationally modified peptides (RiPPs) and is involved in alcohol metabolism of mycobacteria including *Mycobacterium tuberculosis*. A preliminary biosynthetic model had been established by bioinformatics and *in vitro* studies, while the structure of natural MFT and key biosynthetic steps remained elusive. Here, we report the discovery of glycosylated MFT by <sup>13</sup>C-labeling metabolomics and establish a model of its biosynthesis in *Mycolicibacterium smegmatis*. Extensive structure elucidation including NMR revealed that MFT is decorated with up to nine β-1,4-linked glucose residues including 2-O-methylglucose. Dissection of biosynthetic genes demonstrated that the oligoglycosylation is catalyzed by the glycosyltransferase MftF. Furthermore, we confirm the redox cofactor function of glycosylated MFTs by activity-based metabolic profiling using the carveol dehydrogenase LimC and show that the MFT pool expands during cultivation on ethanol. Our results will guide future studies into the biochemical functions and physiological roles of MFT in bacteria.

Received 26th February 2020

Accepted 18th April 2020

DOI: 10.1039/d0sc01172j

rsc.li/chemical-science

## Introduction

Coenzymes are small molecules indispensable for the catalytic activity of many enzymes. While coenzymes like NAD<sup>+</sup> or FAD are ubiquitous in nature and are essential for the core metabolism of all forms of life, specialized cofactors like pyrroloquinoline quinone (PQQ)<sup>1</sup> and coenzyme F<sub>420</sub><sup>2</sup> are restricted to certain microbial phyla, but typically involved in extraordinary metabolic processes like methylotrophy, methanogenesis, or detoxification processes. Moreover, specialized cofactors serve as model systems for the evolution of cofactors

and their co-evolution with their associated enzyme families. They can be regarded as examples of low-molecular weight natural products that modify, extend or enhance microbial metabolism. Mycobacteria are particularly rich in unusual redox cofactors and antioxidants that contribute to redox balance and metabolic plasticity. For instance, mycothiol<sup>3,4</sup> or ergothioneine<sup>5</sup> protect *Mycobacterium tuberculosis* from oxidative stress and support detoxification pathways. Coenzyme F<sub>420</sub> is involved, *e.g.*, in cell wall biosynthesis<sup>6</sup> or defense against nitrosative stress in mycobacteria.<sup>7</sup> Moreover, some antimycobacterial drugs like pretomanid<sup>8</sup> are administered as prodrugs and will only develop bioactivity after biotransformation by a coenzyme F<sub>420</sub>-dependent reductase.<sup>9</sup> Mycofactocin (MFT) is a putative redox-cofactor whose existence has been postulated on the basis of comparative genomics and bioinformatics.<sup>10,11</sup> Its molecular identity and structure, however, have remained elusive to date. The MFT biosynthetic gene cluster is highly conserved and widespread among mycobacteria. The inactivation of the MFT gene locus in the model species *Mycolicibacterium smegmatis* (synonym: *Mycobacterium smegmatis*) as well as *M. tuberculosis* resulted in the inability of the mutants to utilize ethanol as a sole source of carbon and further disturbances of mycobacterial redox homeostasis were revealed.<sup>12</sup> Involvement of MFT in methanol metabolism was reported as well.<sup>13</sup> These recent results strongly support the hypothesis that MFT is a redox cofactor and might represent a fitness factor of mycobacteria during some stages of infection.

\*Junior Research Group Synthetic Microbiology, Leibniz Institute for Natural Product Research and Infection Biology (HKI), Beutenbergstr. 11a, 07745 Jena, Germany. E-mail: gerald.lackner@leibniz-hki.de

<sup>b</sup>Friedrich Schiller University, Beutenbergstr. 11a, 07745 Jena, Germany

<sup>c</sup>Junior Research Group Chemical Biology of Microbe-Host Interactions, Leibniz Institute for Natural Product Research and Infection Biology (HKI), Beutenbergstr. 11a, 07745 Jena, Germany

<sup>d</sup>Department of Biochemistry, Max Planck Institute for Chemical Ecology, Hans-Knöll-Str. 8, 07745 Jena, Germany

<sup>e</sup>Bio Pilot Plant, Leibniz Institute for Natural Product Research and Infection Biology (HKI), Beutenbergstr. 11a, 07745 Jena, Germany

† Electronic supplementary information (ESI) available: ESI Appendix: materials and methods, supplementary figures, supplementary results and discussion (.PDF). Data Set 1: <sup>13</sup>C-labeled compounds (.xlsx). Data Set 2: comparative metabolomics comparing *M. smegmatis* WT, mutant and complement strains (.xlsx). Data Set 3: comparative metabolomics of *M. smegmatis* WT treated with glucose or ethanol (.xlsx). Data Set 4: activity-based metabolic profiling (.xlsx). See DOI: 10.1039/d0sc01172j



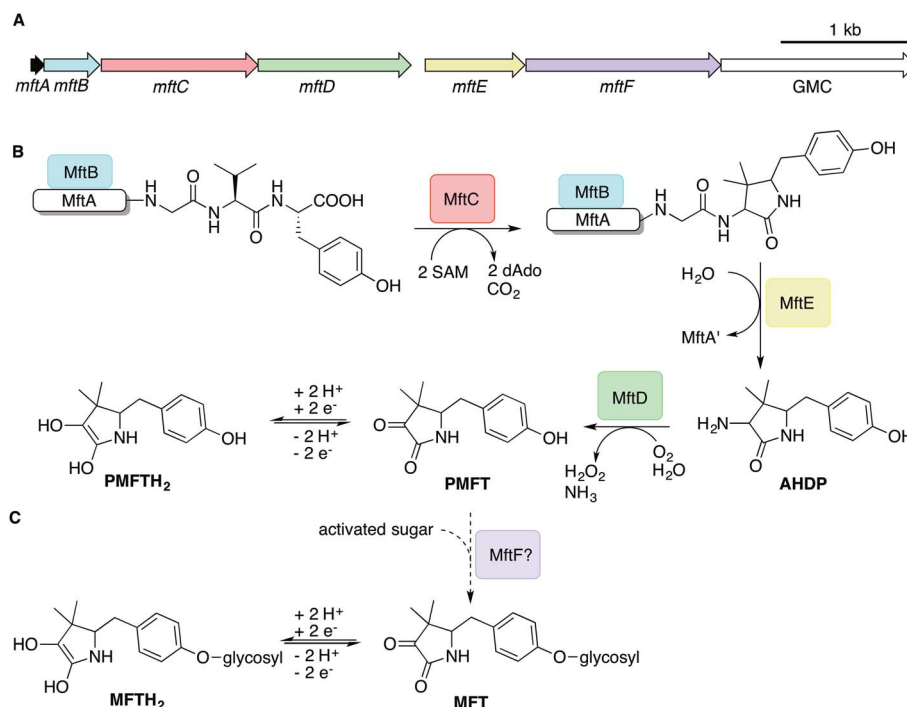


Fig. 1 Biosynthesis of mycofactocin. (A) Schematic representation of the MFT biosynthetic gene cluster of *M. smegmatis*. Arrows present genes *mftA*–*F*. The scale bar indicates 1000 base pairs. (B) Current biosynthesis model of MFT revealed by *in vitro* studies. The precursor peptide MftA (WP\_029104568.1) is bound by its chaperone MftB. The rSAM enzyme MftC catalyzes oxidative decarboxylation and cyclization of the core peptide consisting of a C-terminal Val–Tyr dipeptide. The peptidase MftE releases the cyclized core to form AHDP. MftD performs oxidative deamination of AHDP yielding pre-mycofactocin (PMFT), the presumed redox-active core. (C) The putative glycosyltransferase MftF (WP\_011727662.1) was hypothesized to glycosylate premycofactocins. (P)MFT is reduced to (P)MFTH<sub>2</sub> (mycofactocinol) by oxidoreductases. dAdo: 5'-deoxyadenosine, GMC: glucose-methanol-choline oxidoreductase, SAM: S-adenosyl methionine.

The architecture of the MFT gene cluster (Fig. 1A) suggested that the resulting natural product is a ribosomally synthesized and post-translationally modified peptide (RiPP).<sup>14</sup> Several *in vitro* studies have contributed to a preliminary biosynthetic model of MFT (Fig. 1B): the precursor peptide MftA of *M. smegmatis* consisting of 31 amino acids is produced by the ribosome and bound by its chaperone MftB. Subsequently, the terminal core peptide consisting of Val and Tyr is oxidatively decarboxylated and cyclized by the radical SAM enzyme MftC.<sup>15–17</sup> The resulting cyclic core structure is released by the peptidase MftE<sup>18</sup> forming 3-amino-5-[(*p*-hydroxyphenyl)methyl]-4,4-dimethyl-2-pyrrolidinone (AHDP).<sup>19</sup> Just recently, it was shown that MftD, an enzyme homologous to the *L*-lactate dehydrogenase LldD2<sup>20</sup> catalyzes the oxidative deamination of AHDP to yield pre-mycofactocin (PMFT).<sup>21</sup> The same study demonstrated by voltammetry that the  $\alpha$ -keto amide moiety of PMFT is redox-active and can be reduced to PMFTH<sub>2</sub> (midpoint potential: –255 mV). Efficient reduction was also achieved by the action of carveol dehydrogenase using carveol as an electron donor *in vitro*.<sup>21</sup> Therefore, PMFT likely represents the redox-center of MFT, as riboflavin is the redox-active core of FMN and FAD.

Although these current hypotheses are plausible, all these known metabolic intermediates have only been observed *in vitro* and could therefore represent artifacts, making the verification of their relevance *in vivo* urgently desired. Furthermore,

additional steps of MFT biosynthesis, the function of the *mftF* gene as well as the chemical structure of natural MFT awaited experimental clarification. In this study, we confirm the current biosynthetic model of MFT *in vivo*, detected several novel oligoglycosylated MFT congeners and elucidated their structure. We show that MFTs are decorated with a  $\beta$ -1,4-glucan chain and provide genetic evidence that glycosylation is performed by the glycosyltransferase MftF. Finally, we show dependence of MFT formation on ethanol and corroborate its cofactor function by activity-based metabolic profiling.

## Results and discussion

### Discovery of mycofactocins by metabolomics

In order to identify potential mycofactocin congeners in mycobacteria, we used the fast-growing and weakly pathogenic species *M. smegmatis* MC<sup>2</sup> 155 as a model organism and developed a metabolomics approach combining metabolic induction and labeling to specifically trace MFT congeners. Assuming that MFT production would be stimulated by alcohols, we cultivated bacteria in media containing 10 g L<sup>-1</sup> ethanol. Furthermore, we used stable isotope labeling to obtain candidate molecules compatible with the proposed biosynthetic pathway: since the C-terminal core peptide of MftA is composed of Val and Tyr, we reasoned that MFT congeners could be specifically labeled by feeding *L*-Val-<sup>13</sup>C<sub>5</sub> and *L*-Tyr-<sup>13</sup>C<sub>9</sub>,



Intracellular contents were extracted and analyzed by liquid chromatography coupled with high-resolution mass spectrometry (LC-MS). Compounds were detected by *in silico* grouping of co-eluting isotopic peaks and adducts (feature finding). Afterwards,  $^{13}\text{C}$ -labeled compounds were deduced computationally (Data Set 1†). According to the established biosynthetic pathway we expected 13 carbons to remain  $^{13}\text{C}$ -labeled after oxidative decarboxylation of the Val-Tyr core peptide. We therefore searched for compounds that displayed an exchange of 13 carbons, resulting in a mass shift of +13.04362 Da (ESI Fig. S1†). This approach revealed a list of only twelve candidate compounds. Strikingly, the exact mass and proposed sum formula of three of these compounds corresponded to known intermediates of MFT, recently described *in vitro*, namely AHDP, PMFT as well as PMFTH<sub>2</sub>. In addition to these compounds, several labeled molecules with increasing molecular weight were detected. Some co-eluting candidates with a mass difference of +17.02654 could be explained as NH<sub>4</sub><sup>+</sup> adducts of each other. The remaining nine candidate compounds (Table 1) were grouped based on their chromatographic retention times, eluting closely to either PMFT or PMFTH<sub>2</sub> (approx. 7.2 min and 6.9 min, respectively). Intriguingly, members of the two groups could be arranged in pairs with a mass difference of two hydrogen atoms leading us to assume that each group represented derivatives of either PMFT or PMFTH<sub>2</sub>. Thus, we termed these two clusters of molecules mycofactocinones (MFT) and mycofactocinols (MFTH<sub>2</sub>), respectively. Notably, some mycofactocinols eluted as two chromatographically separated isomers. For instance, the dominant PMFTH<sub>2</sub> eluted at 6.8 min, while the minor isomer eluted at 6.5 min. These two compounds displayed highly similar MS/MS spectra (ESI Fig. S2†) and most likely represent tautomeric forms. For reasons of

simplicity, only the more prevalent isomer was considered during metabolomics studies. We then performed MS/MS networking (Fig. 2), an approach that clusters compounds based on similarity of their MS/MS fragmentation pattern and therefore potentially related chemical scaffolds.<sup>22</sup>

Interestingly, candidates retrieved from  $^{13}\text{C}$ -labeling experiments clustered with further putative MFT congeners. The mass difference between the first candidate mycofactocinol (MFT-1H<sub>2</sub>) with an exact mass of 397.17395 Da and PMFTH<sub>2</sub> was +162.05303 Da, which corresponded to a hexose sugar.

Furthermore, the MS/MS spectrum of MFT-1H<sub>2</sub> (Fig. 2) showed a fragment ion that corresponded to the mass of the putative aglycon ( $m/z$  236.13 [M + H]<sup>+</sup>), thus supporting the assumption that MFT-1H<sub>2</sub> was a glycosylated derivative of PMFTH<sub>2</sub>. MS/MS networking also revealed a recurrent mass difference of 14.01565 between compounds, indicating that methylation might occur as well. We thus assumed that the MFT candidate molecules could be explained as glycosylated or glycosylated and monomethylated species of PMFT(H<sub>2</sub>). In analogy to coenzyme F<sub>420-n</sub>, where *n* indicates the number of glutamyl residues in the side chain,<sup>2</sup> we named the glycosylated molecules MFT-*n*(H<sub>2</sub>) with *n* representing the number of sugar moieties. Monomethylated species were termed methylmycofactocinones (MMFT-*n*) and methylmycofactocinols (MMFT-*n*H<sub>2</sub>), respectively. A targeted search for theoretical mass traces revealed additional members of the MFT-*n*(H<sub>2</sub>) and MMFT-*n*(H<sub>2</sub>) series (Data Set 2†). As expected, the mycofactocinones exhibited MS/MS fragments with a systematic shift by -2.0016 (e.g.,  $m/z$  234.11, 396.16, 572.24) (Fig. 2A) demonstrating that the reduction/oxidation indeed takes place in the PMFT moiety. Oligoglycosylation with up to *n* = 9 saccharide units was detected, while seven and eight units appeared to be the most dominant forms. We observed both methylated (MMFT) and unmethylated (MFT) sugar chains, with the methylated series being more prominent. Only monomethylated species were found. Mass fragmentation of MMFT-*n*(H<sub>2</sub>) species was well in agreement with the assumption that the second sugar was the hotspot for methylation. For instance, MS/MS fragmentation of MMFT-8H<sub>2</sub> yielded peaks corresponding to ions of MFT-1H<sub>2</sub> (398.18011) and MMFT-2H<sub>2</sub> (574.24640) suggesting that the methyl group is present in the second sugar moiety (Fig. 2A).

Table 1 MFT candidate molecules obtained by stable isotope labeling of *M. smegmatis* with L-Val- $^{13}\text{C}_9$  and L-Tyr- $^{13}\text{C}_9$ <sup>a</sup>

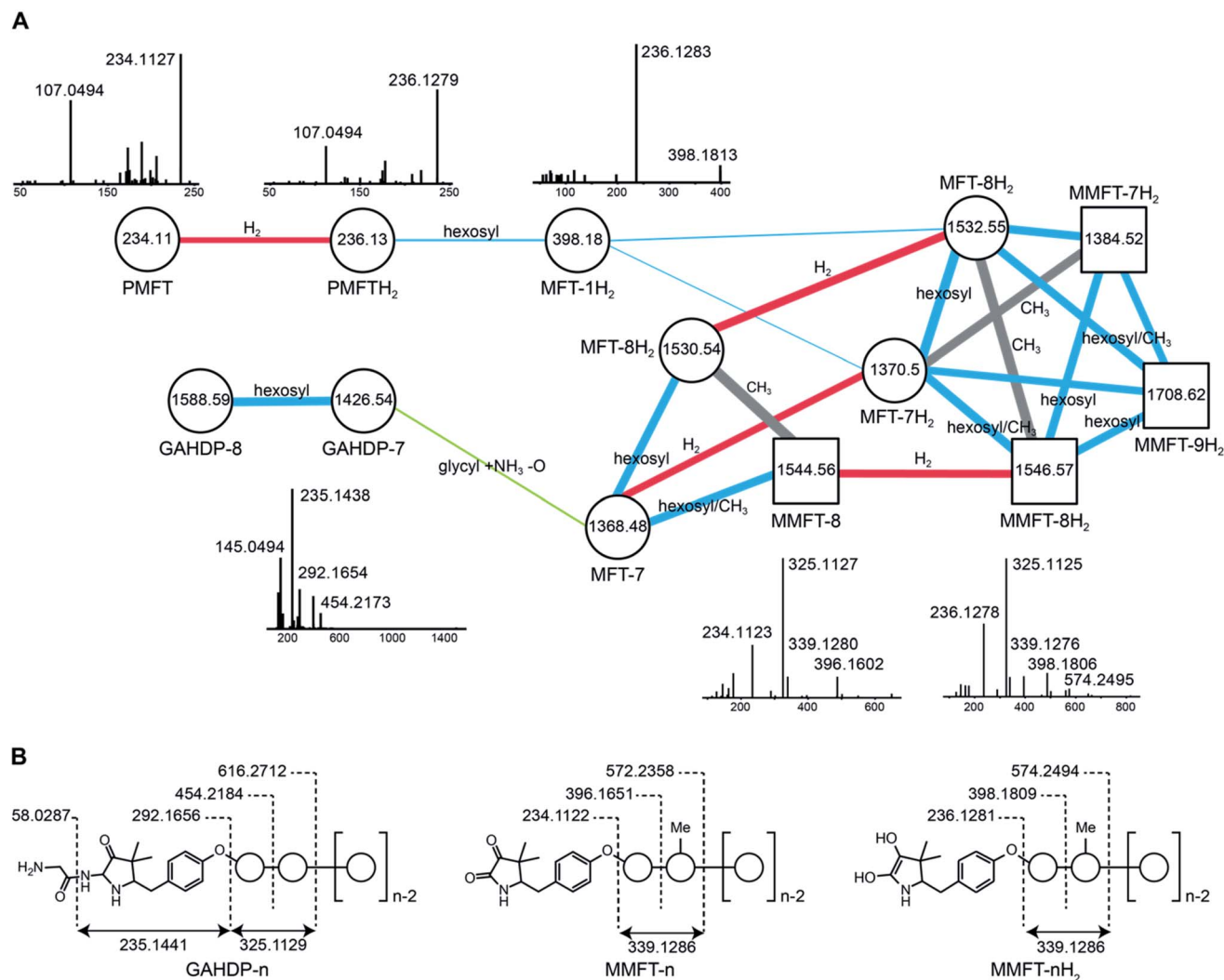
Name	Sum formula	Exact mass (measured)	RT [min]	Area (mean)
<b>Aglycons</b>				
AHDP	C <sub>13</sub> H <sub>18</sub> N <sub>2</sub> O <sub>2</sub>	234.13683	6.53	44 862
PMFTH <sub>2</sub>	C <sub>13</sub> H <sub>17</sub> NO <sub>3</sub>	235.12084	6.88	49 905
PMFT	C <sub>13</sub> H <sub>15</sub> NO <sub>3</sub>	233.10519	7.22	100 832
<b>Mycofactocinols (MFT-<i>n</i>H<sub>2</sub>)</b>				
MFT-1H <sub>2</sub>	C <sub>19</sub> H <sub>27</sub> NO <sub>8</sub>	397.17367	6.88	53 413
<b>Methylmycofactocinols (MMFT-<i>n</i>H<sub>2</sub>)</b>				
MMFT-2H <sub>2</sub>	C <sub>26</sub> H <sub>39</sub> NO <sub>13</sub>	573.24214	6.89	5722
MMFT-7H <sub>2</sub>	C <sub>56</sub> H <sub>89</sub> NO <sub>38</sub>	1383.50626	6.89	340 064
MMFT-8H <sub>2</sub>	C <sub>62</sub> H <sub>99</sub> NO <sub>43</sub>	1545.55908	6.86	550 390
<b>Methylmycofactocinones (MMFT-<i>n</i>)</b>				
MMFT-7	C <sub>56</sub> H <sub>87</sub> NO <sub>38</sub>	1381.49061	7.25	185 407
MMFT-8	C <sub>62</sub> H <sub>97</sub> NO <sub>43</sub>	1543.54343	7.21	335 701

<sup>a</sup> Mycofactocinols (MFT-*n*H<sub>2</sub>) and mycofactocinones (MFT-*n*) represent glycosylated forms of PMFTH<sub>2</sub> and PMFT, respectively. MMFT-*n*: methylmycofactocinones, MMFT-*n*H<sub>2</sub>: methylmycofactocinols (*n*: number of saccharide moieties). Area values represent the mean of 4 biological replicates. All labeled compounds are shown in Data Set 1, all MFT congeners revealed in this study are shown in Data Set 2.

### Structure elucidation of the oligosaccharide moiety

To determine the structure of elongated mycofactocins, we conducted large-scale cultivations. The dominant mycofactocin exhibited the same mass and fragmentation pattern as MMFT-2H<sub>2</sub>, but eluted at a slightly shifted retention time (Fig. S3†). We therefore named this species MMFT-2bH<sub>2</sub>. Due to the low yields and co-elution of contaminants, structural analysis by nuclear magnetic resonance (NMR) was not possible at this stage. However, cellulase (β-1,4-glucanase) treatment degraded the sugar chain of mycofactocin species (*n* > 2), while amylase (α-1,4-glucanase) did not exhibit any effect (Fig. 3A). This finding strongly suggested that the oligosaccharide chain represents a β-1,4-glucan. Intriguingly, isomer MMFT-2bH<sub>2</sub>, but not



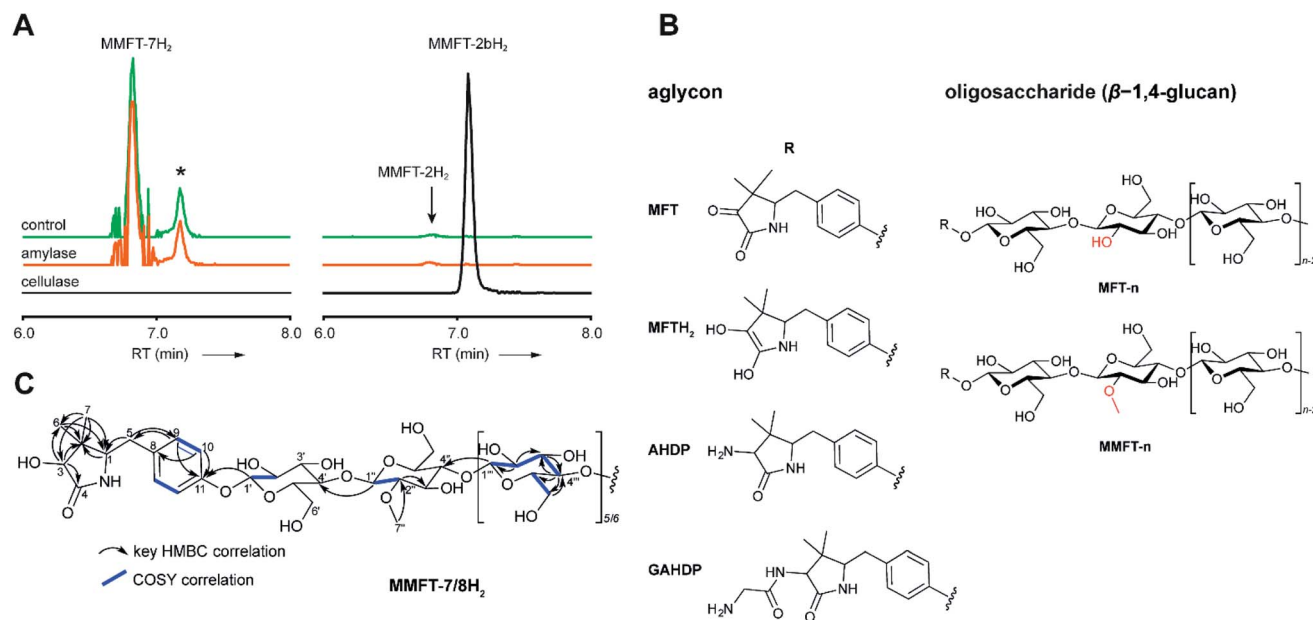


**Fig. 2** Discovery and tandem mass spectrometry of MFT congeners. (A) Molecular network of MFT congeners. Nodes (circles) represent chemical compounds. Internal node labels display the precursor mass of compounds ( $m/z$  [ $M + H^+$ ]). External node labels show proposed compound annotations. Edges represent relationships in terms of shared MS/MS fragments. Edge labels show proposed modifications based on precursor mass shifts (blue: hexosylation, red: oxidation/reduction, grey: methylation). Line widths of edges mirror cosine distances. Representative MS/MS spectra of corresponding precursor ions are shown above or below nodes. (B) Schematic representation of mass fragmentation patterns of GAHDP- $n$ , MMFT- $nH_2$  and MMFT- $n$ . Numbers indicate mass-to-charge ratios ( $m/z$ ) of fragments observed. Circles represent hexose moieties. Me: methyl group.

MMFT-2H<sub>2</sub>, accumulated after the enzymatic digest, suggesting that MMFT-2b(H<sub>2</sub>) represents the product of cellulase digestion of MMFT-7/8(H<sub>2</sub>) and shares the identical disaccharide anchor. To further elucidate the structure of MMFT- $nH_2$ , we analyzed enriched fractions of MMFT-2bH<sub>2</sub> and MMFT- $nH_2$  by chemical derivatization and gas chromatography coupled with mass spectrometry (GC-MS). Monosaccharides were released by acid hydrolysis and derivatized by trimethylsilylation (TMS). Comparative analysis of peaks arising from the MMFT- $nH_2$  and MMFT-2bH<sub>2</sub> fractions and carbohydrate standards confirmed the presence of D-glucose (Fig. S4†) and revealed that the methylated sugar present in MMFT- $n(H_2)$  is 2-O-methyl-D-glucose (Fig. S5†). To confirm the glycosidic linkage positions, the oligosaccharide was permethylated before hydrolysis so that only hydroxyl groups involved in glycosidic bond formation

would be free for silylation.<sup>23</sup> This experiment (Fig. S6†) led to the formation of glucose with 2,3,6-O-methyl-1,4-O-TMS modification confirming the 1,4-glycosidic linkage. Additional modification experiments (methanolysis and permethylation) supported the assignments (Fig. S7–S19†). After repeated cultivation we finally obtained MMFT-7/8H<sub>2</sub> in sufficient amounts to record 1D and 2D-NMR spectra (ESI results and discussion, Fig. S20–S27, Tables S1 and S2†). The <sup>1</sup>H NMR spectrum of MMFT-7/8H<sub>2</sub> exhibited a similar five-membered lactam moiety as present in AHDP, but an isolated methine group was shifted to low-field ( $\delta_{H-3}$  4.28 ppm/ $\delta_{C-3}$  76.16 ppm) compared to AHDP ( $\delta_{H-3}$  3.30 ppm/ $\delta_{C-3}$  61.64 ppm). This indicated the amine group connected to C-3 was replaced by a hydroxyl group. The HMBC correlation between H-1' to C-11 suggested the sugar chain to be attached to the hydroxyl group of the tyrosine moiety (Fig. 3B). The  $\beta$ -1,4-glycosidic linkage was





**Fig. 3** Structure of mycofactocins. (A) Enzymatic degradation of MMFT-*n* by cellulase. Extracted ion chromatograms (XIC,  $[M + H]^+$ ) of extract of *M. smegmatis* corresponding to MMFT-7H<sub>2</sub> ( $m/z$  1383.50626, left stack) and MMFT-2H<sub>2</sub> or MMFT-2bH<sub>2</sub> ( $m/z$  574.24640, right stack) after treatment with cellulase, amylase, or buffer (control) are shown. Asterisk designates a peak corresponding to the  $M + 2$  isotope of MMFT-7. Digestion by cellulase ( $\beta$ -1,4-glucanase) consumes MMFT-*n*H<sub>2</sub> and produces MMFT-2bH<sub>2</sub> suggesting that the oligosaccharide consists of  $\beta$ -1,4-linked glucose. (B) Key COSY and HMBC correlations of MMFT-7/8H<sub>2</sub>. (C) Proposed chemical structures of key mycofactocins and biosynthetic congeners. Mycofactocins are glycosylated by sugar chains consisting of up to nine  $\beta$ -1,4-linked glucose units ( $n \leq 9$ ). In methylated mycofactocins (MMFT) the second hexose is methylated (2-*O*-methyl- $\beta$ -D-glucose). The aglycon is PMFT or PMFTH<sub>2</sub> in mycofactocinones or mycofactocinols, respectively. The aglycon is AHDP or GAHDP in biosynthetic precursors AHDP-*n* and GAHDP-*n*, respectively.

confirmed by the HMBC correlations of H-1'' to C-4' and H-1''' to C-4'' and the configuration of the glucose moiety was assigned as  $\beta$ -form by the large coupling constant of anomeric protons ( $J_{H-1''-H-2'} = 8.0$  Hz;  $J_{H-1'''-H-2''} = 8.0$  Hz;  $J_{H-1''''-H-2'''} = 8.0$  Hz). The position of the methylated glucose was determined by the observation of a methoxy moiety ( $\delta_{H-7''} = 3.64$  ppm/ $\delta_{C-7''} = 60.55$  ppm) and HMBC correlation of H-7'' to C-2''. The planar structure of MMFT-7/8H<sub>2</sub> is presented in Fig. 3C.

In summary, we propose that the oligosaccharide moiety of MFT is a  $\beta$ -1,4-glucane (cellulose). The methylated hexose present in MMFT-*n*(H<sub>2</sub>) and MMFT-2b(H<sub>2</sub>) was shown to be 2-*O*-methylglucose. The fact that MMFT-2 and MMFT-2b (digested MMFT-*n*) are distinct in retention times points to some degree of structural diversity within MMFTs. Notably, cellulose was shown to be produced by *M. tuberculosis* as a constituent of biofilms after exposure to reductive stress.<sup>24</sup> The production of methylated glucans, like 6-*O*-methylglucose lipopolysaccharides (MGPL), albeit with  $\alpha$ -1,4 linkage, is well described in *Mycobacteria*, 2-*O*-methylglucose appears to be less common.<sup>25</sup> Glycosylation is a relatively uncommon modification of cofactors. The most important examples are mycothiol<sup>4</sup> and bacillithiol.<sup>26</sup>

### Glycine-derived intermediates of MFT biosynthesis

Surprisingly, the MS/MS network (Fig. 2A) revealed two additional compounds ( $m/z$  1426.54 and 1588.59) with an unusual mass shift compared to the MFT-*n*(H<sub>2</sub>) candidates. Their mass differences and MS/MS spectra indicated that they represented hepta- and octaglycosylated species sharing a head moiety closely related to

PMFT and PMFT(H<sub>2</sub>). The molecular masses and MS/MS spectra of the compounds could be explained by the assumption that the aglycon corresponded to glycy-AHDP (GAHDP) and these compounds represented the oligoglycosylated forms GAHDP-7 and GAHDP-8. Since the VY core peptide of MftA is preceded by a glycine residue at its N-terminal side it appeared highly likely that the GAHDP-*n* species corresponded to premature cleavage products of the MftC-processed precursor peptide. To corroborate this hypothesis, we fed *M. smegmatis* cultures with a combination of fully <sup>13</sup>C-labeled Gly-<sup>13</sup>C<sub>2</sub>, Val-<sup>13</sup>C<sub>5</sub>, and Tyr-<sup>13</sup>C<sub>9</sub>. Indeed, GAHDP-derived molecules underwent a mass shift of +15.05033 Da, indicating the incorporation of Gly-<sup>13</sup>C<sub>2</sub> (+2.00671 Da) in addition to the decarboxylated Val-Tyr moiety (+13.04362 Da) (Fig. S28†). Targeted searches for GAHDP-*n* as well as AHDP-*n* (lacking the glycy residue) and MAHDP-*n* candidates (AHDP decorated with monomethylated oligosaccharide) revealed three series of oligoglycosylated compounds with similar retention times within each series (Fig. 4, Data Set 2†).

### Dissection of MFT biosynthesis

In order to test if all MFT candidate compounds were related to MFT biosynthesis, we investigated mutants ( $\Delta mftC$ ,  $\Delta mftD$ ,  $\Delta mftE$ ,  $\Delta mftF$ ) created previously<sup>12</sup> of the MFT biosynthesis pathway for the production of candidate molecules (Fig. 4A, Data Set 2†). Indeed, none of the aglycons, nor any of the glycosylated candidates were detected in the  $\Delta mftC$  strain (Fig. 4B). This finding, together with the fact that the genetically complemented strain  $\Delta mftC$ -Comp restored production of MFT



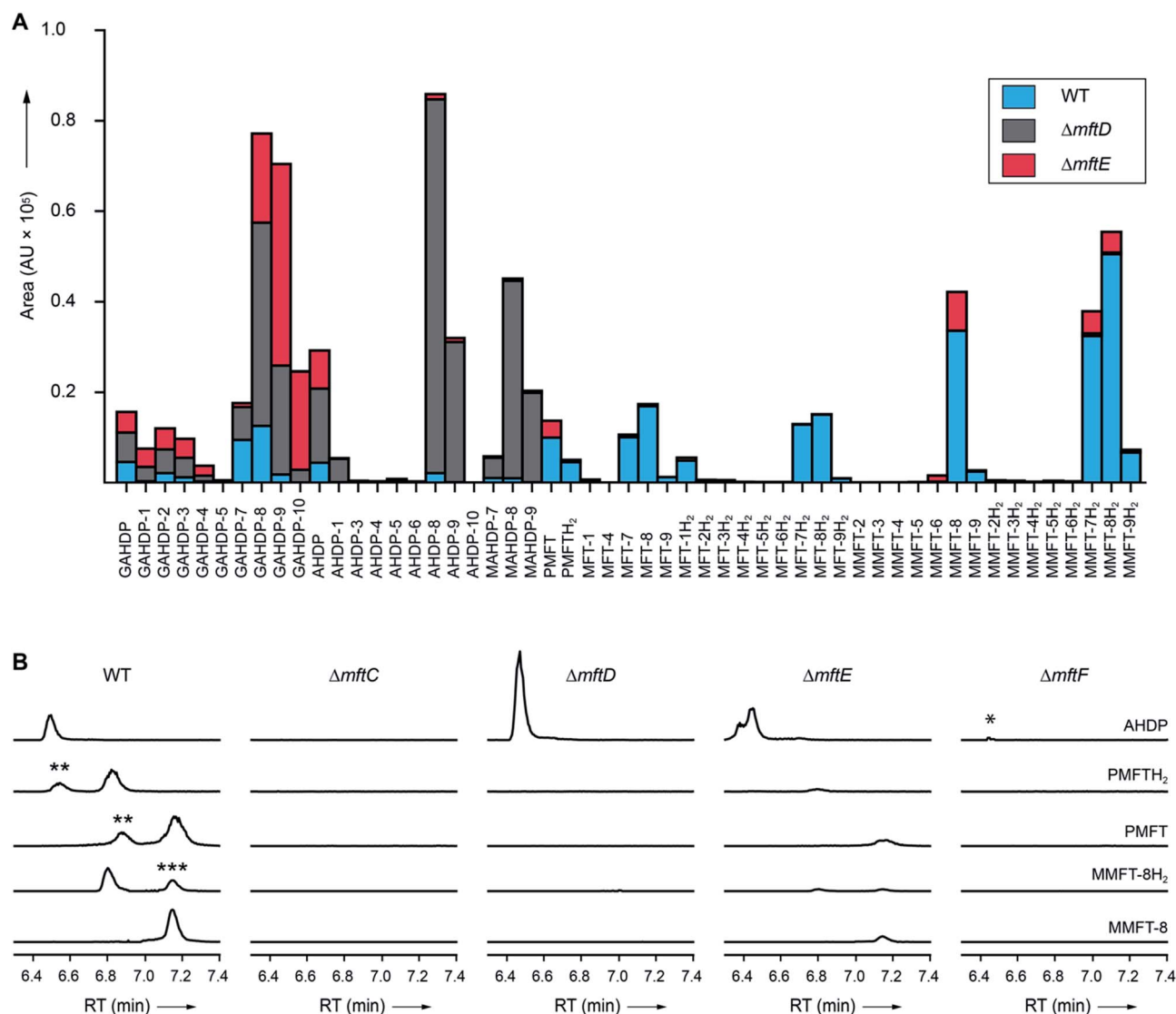


Fig. 4 Metabolic profile of MFT congeners present in *M. smegmatis*. (A) Distribution of proposed MFT congeners as determined by LC-MS (Data Set 2†). Bars indicate area under the curve of designated species (average of three biological replicates,  $n = 3$ ). Blue: WT, red:  $\Delta mftE$ , gray:  $\Delta mftD$ . The  $\Delta mftE$  mutant produces significantly reduced amounts of MFT congeners compared to WT, but accumulates incorrectly cleaved products (GAHDP- $n$  series).  $\Delta mftD$  is unable to produce PMFT(H<sub>2</sub>) and glycosylated (M)MFT- $n$ (H<sub>2</sub>), thus accumulating AHDP- $n$  congeners. (B) Extracted ion chromatograms (XIC, [M + H]<sup>+</sup>) of WT and mutants ( $\Delta mftC$ ,  $\Delta mftD$ ,  $\Delta mftE$ ,  $\Delta mftF$ ) corresponding to AHDP ( $m/z$  235.14411), PMFT ( $m/z$  234.11247), PMFTH<sub>2</sub> ( $m/z$  236.12812), MMFT-8 ( $m/z$  1544.55072) and MMFT-8H<sub>2</sub> ( $m/z$  1546.56637). \*\*marks minor isomeric forms (Fig. S2†). \*\*\*marks a peak corresponding to the M + 2 isotope of MMFT-8.  $\Delta mftC$  is blocked in biosynthesis of all MFT intermediates,  $\Delta mftF$  abolishes most of the MFT products, but forms trace amounts of AHDP (\*).  $\Delta mftE$  produces most MFT species in lower amounts, while intermediates like AHDP are increasing.  $\Delta mftD$  strongly accumulates AHDP, while MFT congeners are abolished.

congeners (Data Set 2†) represented strong evidence that we indeed identified *bona-fide* MFT-derivatives.

The  $\Delta mftE$  mutant was able to produce mycofactocins, albeit in significantly lower amounts, explaining the previously unexpected phenotypic observation that the  $\Delta mftE$  mutant was able to grow on ethanol, but slower than WT.<sup>12</sup> Intriguingly, the pool of GAHDP- $n$  was strongly increased in the  $\Delta mftE$  strain (Fig. 4A). We thus conclude that MftE can be complemented by an unknown peptidase present in the metabolic background of mycobacteria. Theoretically, an aminopeptidase would be sufficient to degrade the N-terminus of MftA, releasing the AHDP-like core. Peptidases

encoded outside the biosynthetic gene cluster have been observed in other RiPP biosyntheses as well.<sup>27</sup>

However, the removal of the glycine residue might be an apparent bottleneck of the alternative maturation pathway in *M. smegmatis*. Alternatively, GAHDPs could represent shunt products that cannot be further processed. In full agreement with the *in vitro* finding that MftD consumes AHDP to form PMFTH<sub>2</sub>,<sup>21</sup> all metabolites downstream of (M)AHDP- $n$  were abrogated in the  $\Delta mftD$  strain, whereas AHDP- $n$  and GAHDP- $n$  accumulated (Fig. 4). The fact that GAHDP- $n$  increased might suggest that the MftE step is impeded in the absence of MftD as



well. Genetic dysregulation or cooperative effects between the two enzymes, like complex formation and substrate channeling, might account for this result.

### Glycosylation of MFT is mediated by MftF

It has been speculated that the putative glycosyltransferase MftF catalyzes a final glycosylation of PMFT (Fig. 1C) to yield the mature cofactor.<sup>21</sup> Our results at this point showed that multiple glucose residues are indeed attached to the aglycon *in vivo*. However, glycosylation appeared already at an early stage as mirrored by the presence of the glycosylated (G)AHDP-*n* series. In order to link oligoglycosylation to a given gene product, we analyzed the  $\Delta mftF$  mutant for the production of glycosylated MFT congeners. Indeed, all glycosylated MFT congeners were abolished in the  $\Delta mftF$  metabolome. Unexpectedly,  $\Delta mftF$  mutants additionally ceased to produce the aglycons PMFT and PMFTH<sub>2</sub>. MftF did, however, produce trace amounts of AHDP, thus showing that at least residual MftC activity was present in the mutant (Fig. 4B). To exclude polar effects, we complemented  $\Delta mftF$  by re-introduction of the *mftF* gene under control of the *mftA* promoter. The restoration of the full MFT metabolite spectrum (Data Set 2†) excluded polar effects and thus verified that MftF was the glycosyltransferase responsible for oligoglycosylation of MFT congeners. The appearance of glycosylated (G)AHDP species in WT together with the drastic decrease of aglycons in  $\Delta mftF$  can be interpreted in a scenario where either glycosylation or the MftF protein itself are essential for the MftD step to efficiently take place *in vivo*. If missing, the biosynthetic machinery may fail to assemble a functional complex or may be unable to recruit the unglycosylated metabolic precursors. The finding that the *mftF* gene is a conserved constituent of MFT biosynthetic loci among different phyla supports the importance of this modification.<sup>11</sup>

The deduced MftF protein of *M. smegmatis* (MSMEG\_1426) consists of 470 amino acids (aa) and belongs to the glycosyltransferase 2 family (GT2) according to PFAM (PF00535) and CAZy searches. These enzymes are known for an inverting mechanism of oligoglycoside formation. This is well in agreement with the proposed  $\beta$ -configuration of the MFT oligosaccharide chain. Sequence alignment (Fig. S29A†) showed a high degree of sequence conservation among mycobacterial species and other actinomycetes (*e.g.*, 92% similarity to MftF of *M. tuberculosis* H37Rv). Prediction of transmembrane domains revealed a single helix spanning residues 324–346 with the N-terminus being located outside of the membrane (Fig. S29B†). The MMFT biosynthetic machinery, however, appears not to be fully encompassed within the MFT cluster since no methyltransferase was found. Future studies are warranted to identify the enzymes involved in MFT oligosaccharide methylation.

### Cofactor role of mycofactocin

After discovery of the glycosylated mycofactocins, we examined to which extent their production was actually dependent on the presence of ethanol in culture media. We therefore systematically compared the metabolome of *M. smegmatis* WT after ethanol treatment with glucose controls (Data Set 3†).

The results demonstrated that all MFT congeners or intermediates were strongly upregulated upon cultivation on ethanol (median: 34-fold upregulation) (Fig. 5A). These data perfectly support a recent report that MFT is involved in alcohol metabolism.<sup>12</sup>

Finally, we sought to confirm that the MFT congeners identified in this study are indeed coenzymes of MFT-dependent enzymes. To assess this question, we turned to activity-based metabolic profiling.<sup>28</sup> We incubated the extracted metabolome of *M. smegmatis* with the recombinant *l*-carveol dehydrogenase LimC (CAB54559.1) from *Rhodococcus erythropolis* (Fig. S30†), a nicotinoprotein with a non-exchangeable NADH cofactor.<sup>29</sup> This enzyme was proposed to require MFT as an external electron acceptor.<sup>11</sup> A recent study showed that carveol dehydrogenase from *M. smegmatis* was able to reduce PMFT to PMFTH<sub>2</sub> using carveol and internally bound NADH as an electron donor.<sup>21</sup> Likewise, we observed full reduction of all mycofactocinones to mycofactocinols (Fig. 5B, Data Set 4†) by LimC when combined with carveol as a substrate. Controls lacking enzyme or substrate showed weak and no turnover, respectively. The low turnover by LimC alone can be explained by internally bound NADH as reported

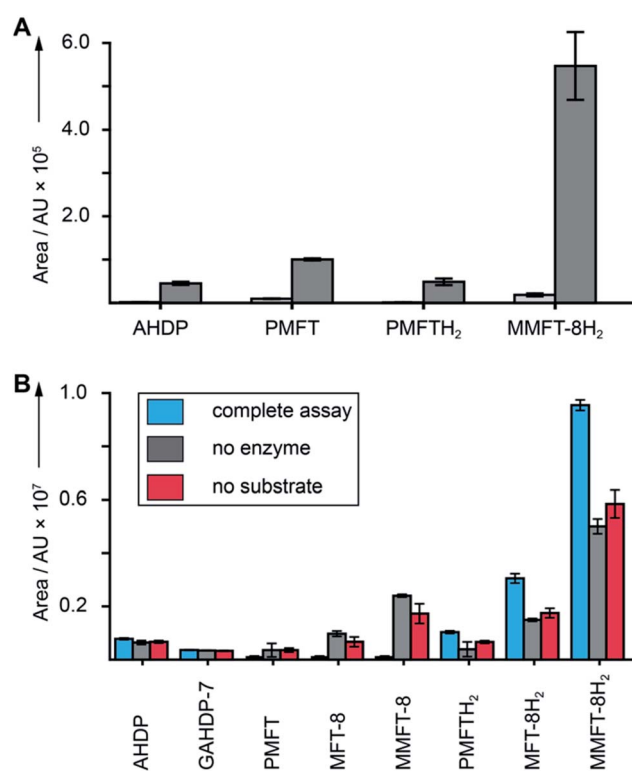


Fig. 5 Cofactor role of mycofactocin. (A) MFT congeners are strongly upregulated (MMFT-8H<sub>2</sub>: 26-fold) on ethanol-containing media. Area under the curve of MFT species produced by *M. smegmatis* treated with ethanol (dark gray) versus glucose controls (light gray) are shown (Data Set 3†). (B) Reduction of mycofactocinones to mycofactocinols by treatment of extracts with LimC and carveol. Blue: complete assay with enzyme and *l*-carveol as substrate. Red: control without substrate, gray: control without enzyme. Bars represent average area under the curve, error bars standard deviation of 3 biological replicates (*n* = 3) for in both charts.





before.<sup>21</sup> Both the aglycon PMFT as well as the oligoglycosylated MFT-*n* and MMFT-*n* species were completely turned over, while redox-inactive AHDP congeners remained unaffected. These data further validate the notion that all MFT candidates presented here are mycofactocins with full cofactor function. It remains to be clarified if there is a preference for the glycosylated coenzymes or their aglycons in the bacterial cell.

## Conclusion

The redox cofactor mycofactocin has attracted considerable interest since it was postulated by bioinformatics. Despite recent progress made by *in vitro* studies, evidence for mycofactocin congeners in living microorganisms has been missing so far. Our integrated metabolomics approach combining stable isotope labeling, metabolite induction, MS/MS networking as well as genetic dissection of the biosynthetic pathway turned out to be a powerful approach to identify RiPP congeners in bacteria and could inspire similar projects in the future. Using this technique, we discovered natural MFT and found that it is decorated with oligosaccharides consisting of up to nine  $\beta$ -1,4-linked glucose units. Analyses of  $\Delta mftF$  mutants and complement strains revealed that MftF is the glycosyltransferase responsible for the oligoglycosylation observed. Mycofactocins can be isolated in oxidized (mycofactocinones) and reduced forms (mycofactocinols) and are co-substrates of enzymatic reduction by carveol dehydrogenase. These data provide strong evidence that mycofactocins are indeed redox cofactors as proposed earlier.<sup>11,12,21</sup> We, therefore, conclude that we have finally discovered the family of compounds that was tentatively called “mycofactocin” and thus close an important gap of knowledge in the field. Our results will guide further studies into the occurrence, physiological role, and biochemistry of mycofactocins in microorganisms. Finally, these and other studies will inspire future efforts to exploit mycofactocin, *e.g.*, as a disease marker or as a potential drug target for the treatment mycobacterial infections.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We would like to thank the Carl-Zeiss Foundation and the Leibniz Association as well as the European Regional Development Fund for financial support. CB was kindly supported by the CRC ChemBioSys 1127 (Deutsche Forschungsgemeinschaft). We thank Stefan Kaufmann and Gopinath Krishnamoorthy for kindly providing *M. smegmatis* mutants. We also thank Heike Heinecke for NMR measurements.

## Notes and references

1 J. P. Klinman and F. Bonnot, *Chem. Rev.*, 2014, **114**, 4343–4365.

- 2 C. Greening, F. H. Ahmed, A. E. Mohamed, B. M. Lee, G. Pandey, A. C. Warden, C. Scott, J. G. Oakeshott, M. C. Taylor and C. J. Jackson, *Microbiol. Mol. Biol. Rev.*, 2016, **80**, 451–493.
- 3 N. A. Buchmeier, G. L. Newton, T. Koledin and R. C. Fahey, *Mol. Microbiol.*, 2003, **47**, 1723–1732.
- 4 G. L. Newton, N. Buchmeier and R. C. Fahey, *Microbiol. Mol. Biol. Rev.*, 2008, **72**, 471–494.
- 5 V. Saini, B. M. Cumming, L. Guidry, D. A. Lamprecht, J. H. Adamson, V. P. Reddy, K. C. Chinta, J. H. Mazorodze, J. N. Glasgow, M. Richard-Greenblatt, A. Gomez-Velasco, H. Bach, Y. Av-Gay, H. Eoh, K. Rhee and A. J. C. Steyn, *Cell Rep.*, 2016, **14**, 572–585.
- 6 E. Purwantini and B. Mukhopadhyay, *PLoS One*, 2013, **8**, e81985.
- 7 E. Purwantini and B. Mukhopadhyay, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 6333–6338.
- 8 C. K. Stover, P. Warren, D. R. VanDevanter, D. R. Sherman, T. M. Arain, M. H. Langhorne, S. W. Anderson, J. A. Towell, Y. Yuan, D. N. McMurray, B. N. Kreiswirth, C. E. Barry and W. R. Baker, *Nature*, 2000, **405**, 962–966.
- 9 R. Singh, U. Manjunatha, H. I. Boshoff, Y. H. Ha, P. Niyomrattanakit, R. Ledwidge, C. S. Dowd, I. Y. Lee, P. Kim, L. Zhang, S. Kang, T. H. Keller, J. Jiricek and C. E. Barry III, *Science*, 2008, **322**, 1392–1395.
- 10 R. Ayikpoe, V. Govindarajan and J. A. Latham, *Appl. Microbiol. Biotechnol.*, 2019, **103**, 2903–2912.
- 11 D. H. Haft, *BMC Genomics*, 2011, **12**, 21.
- 12 G. Krishnamoorthy, P. Kaiser, L. Lozza, K. Hahnke, H. J. Mollenkopf and S. H. E. Kaufmann, *mBio*, 2019, **10**, e00190-19.
- 13 A. A. Dubey and V. Jain, *Biochem. Biophys. Res. Commun.*, 2019, **516**, 1073–1077.
- 14 P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, P. D. Cotter, D. J. Craik, M. Dawson, E. Dittmann, S. Donadio, P. C. Dorrestein, K. D. Entian, M. A. Fischbach, J. S. Garavelli, U. Goransson, C. W. Gruber, D. H. Haft, T. K. Hemscheidt, C. Hertweck, C. Hill, A. R. Horswill, M. Jaspars, W. L. Kelly, J. P. Klinman, O. P. Kuipers, A. J. Link, W. Liu, M. A. Marahiel, D. A. Mitchell, G. N. Moll, B. S. Moore, R. Muller, S. K. Nair, I. F. Nes, G. E. Norris, B. M. Olivera, H. Onaka, M. L. Patchett, J. Piel, M. J. Reaney, S. Rebuffat, R. P. Ross, H. G. Sahl, E. W. Schmidt, M. E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Stein, R. D. Sussmuth, J. R. Tagg, G. L. Tang, A. W. Truman, J. C. Vederas, C. T. Walsh, J. D. Walton, S. C. Wenzel, J. M. Willey and W. A. van der Donk, *Nat. Prod. Rep.*, 2013, **30**, 108–160.
- 15 N. A. Bruender and V. Bandarian, *Biochemistry*, 2016, **55**, 2813–2816.
- 16 B. Khaliullin, P. Aggarwal, M. Bubas, G. R. Eaton, S. S. Eaton and J. A. Latham, *FEBS Lett.*, 2016, **590**, 2538–2548.
- 17 B. Khaliullin, R. Ayikpoe, M. Tuttle and J. A. Latham, *J. Biol. Chem.*, 2017, **292**, 13022–13033.
- 18 N. A. Bruender and V. Bandarian, *J. Biol. Chem.*, 2017, **292**, 4371–4381.



- 19 R. Ayikpoe, J. Salazar, B. Majestic and J. A. Latham, *Biochemistry*, 2018, **57**, 5379–5383.
- 20 S. Billig, M. Schneefeld, C. Huber, G. A. Grassl, W. Eisenreich and F. C. Bange, *Sci. Rep.*, 2017, **7**, 6484.
- 21 R. S. Ayikpoe and J. A. Latham, *J. Am. Chem. Soc.*, 2019, **141**, 13582–13591.
- 22 M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kapon, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A. V. Melnik, M. J. Meehan, W. T. Liu, M. Crusemann, P. D. Boudreau, E. Esquenazi, M. Sandoval-Calderon, R. D. Kersten, L. A. Pace, R. A. Quinn, K. R. Duncan, C. C. Hsu, D. J. Floros, R. G. Gavilan, K. Kleigrew, T. Northen, R. J. Dutton, D. Parrot, E. E. Carlson, B. Aigle, C. F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B. T. Murphy, L. Gerwick, C. C. Liaw, Y. L. Yang, H. U. Humpf, M. Maansson, R. A. Keyzers, A. C. Sims, A. R. Johnson, A. M. Sidebottom, B. E. Sedio, A. Klitgaard, C. B. Larson, C. A. Boya P., D. Torres-Mendoza, D. J. Gonzalez, D. B. Silva, L. M. Marques, D. P. Demarque, E. Pociute, E. C. O'Neill, E. Briand, E. J. N. Helfrich, E. A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J. J. Kharbush, Y. Zeng, J. A. Vorholt, K. L. Kurita, P. Charusanti, K. L. McPhail, K. F. Nielsen, L. Vuong, M. Elfeki, M. F. Traxler, N. Engene, N. Koyama, O. B. Vining, R. Baric, R. R. Silva, S. J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P. G. Williams, J. Dai, R. Neupane, J. Gurr, A. M. C. Rodriguez, A. Lamsa, C. Zhang, K. Dorrestein, B. M. Duggan, J. Almaliti, P. M. Allard, P. Phapale, L. F. Nothias, T. Alexandrov, M. Litaudon, J. L. Wolfender, J. E. Kyle, T. O. Metz, T. Peryea, D. T. Nguyen, D. VanLeer, P. Shinn, A. Jadhav, R. Muller, K. M. Waters, W. Shi, X. Liu, L. Zhang, R. Knight, P. R. Jensen, B. O. Palsson, K. Pogliano, R. G. Linington, M. Gutierrez, N. P. Lopes, W. H. Gerwick, B. S. Moore, P. C. Dorrestein and N. Bandeira, *Nat. Biotechnol.*, 2016, **34**, 828–837.
- 23 A. Debettigniesdutz, G. Reznicek, B. Kopp and J. Jurenitsch, *J. Chromatogr.*, 1991, **547**, 299–306.
- 24 A. Trivedi, P. S. Mavi, D. Bhatt and A. Kumar, *Nat. Commun.*, 2016, **7**, 11392.
- 25 G. Stadthagen, T. Sambou, M. Guerin, N. Barilone, F. Boudou, J. Kordulakova, P. Charles, P. M. Alzari, A. Lemassu, M. Daffe, G. Puzo, B. Gicquel, M. Riviere and M. Jackson, *J. Biol. Chem.*, 2007, **282**, 27270–27276.
- 26 G. L. Newton, M. Rawat, J. J. La Clair, V. K. Jothivasan, T. Budiarto, C. J. Hamilton, A. Claiborne, J. D. Helmann and R. C. Fahey, *Nat. Chem. Biol.*, 2009, **5**, 625–627.
- 27 S. Chen, B. Xu, E. Chen, J. Wang, J. Lu, S. Donadio, H. Ge and H. Wang, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 2533–2538.
- 28 L. P. de Carvalho, H. Zhao, C. E. Dickinson, N. M. Arango, C. D. Lima, S. M. Fischer, O. Ouerfelli, C. Nathan and K. Y. Rhee, *Chem. Biol.*, 2010, **17**, 323–332.
- 29 M. J. van der Werf, C. van der Ven, F. Barbirato, M. H. Eppink, J. A. de Bont and W. J. van Berkel, *J. Biol. Chem.*, 1999, **274**, 26296–26304.

