The mycolic acid reductase Rv2509 has distinct structural motifs and is essential for growth in slow-growing mycobacteria

Javid, Asma; Cooper, Charlotte; Singh, Albel; Schindler, Steffen; Hänisch, Milena; Marshall, Robert L; Kalscheuer, Rainer; Bavro, Vassiliy N; Bhatt, Apoorva

DOI: 10.1111/mmi.14437

License: Creative Commons: Attribution (CC BY)

Citation for published version (Harvard):
Javid, A, Cooper, C, Singh, A, Schindler, S, Hänisch, M, Marshall, RL, Kalscheuer, R, Bavro, VN & Bhatt, A 2019, 'The mycolic acid reductase Rv2509 has distinct structural motifs and is essential for growth in slow-growing mycobacteria', Molecular Microbiology. https://doi.org/10.1111/mmi.14437

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

• Users may freely distribute the URL that is used to identify this publication.
• Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
• Users may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
• Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 05. May. 2020
INTRODUCTION

The cell walls of mycobacteria including the tuberculosis-causing Mycobacterium tuberculosis contain distinct long-chain fatty acids termed mycolic acids. These \( \alpha \)-alkyl, \( \beta \)-hydroxy fatty acids form an integral part of the cell wall of mycobacteria, and of other related bacteria of the suborder Corynebacterineae, including the genera Nocardia, Corynebacterium and Rhodococcus (Marrakchi, Laneelle, & Daffé, 2014). Mycolic acids can be found covalently linked to cell wall arabinogalactan and as part of the glycolipids' trehalose monomycolate (TMM), trehalose dimycolate and glucose monomycolate. They are essential for mycobacterial viability (Marrakchi et al., 2014; Nataraj et al., 2015) and virulence and are synthesised by a complex
array of enzymes that includes a mammalian-like Fatty Acid Synthase I (FAS-I), and a multi-enzyme complex of Fatty Acid Synthase-II (FAS-II) (Marrakchi et al., 2014). One of the late stages of mycolic acid biosynthesis involves the Claisen condensation of a FAS-II-derived long meromycolate chain (C_{62}-C_{62}) with a short FAS-I-derived fatty acid (C_{24}-C_{24}) to yield α-alkyl, β-keto fatty acid intermediate, catalysed by the polyketide synthase Pks13 (Gande et al., 2004; Portevin et al., 2004). The final step in the formation of a mature mycolic acid involves the enzymatic reduction of the β-keto group of the product of Pks13 to produce α-alkyl, β-hydroxy fatty acids (mature mycolic acid). In M. tuberculosis, the reductase required for this final step is encoded by Rv2509 (Bhatt, Brown, Singh, Minnikin, & Besra, 2008; Lea-Smith et al., 2007). Initial studies in Corynebacterium glutamicum, a species that can survive the loss of mycolic acids, indicated that deletion of cmr (NCgl2385), the Rv2509 orthologue, resulted in a strain that accumulated the α-alkyl, β-keto fatty acid precursor instead of mature mycolic acids (Lea-Smith et al., 2007). Mycolic acid biosynthesis genes are essential in mycobacteria; however, surprisingly, we were able to generate a viable mutant of MSMEG4722, the homologue of Rv2509 in the fast-growing Mycobacterium smegmatis (Bhatt et al., 2008). Similar to the Cl. glutamicum NCgl2385 mutant, the M. smegmatis MSMEG4722 mutant produced α-alkyl, β-keto fatty acyl precursors of mycolic acids (3-oxo-mycolic acid precursors) which were transported and esterified to arabinogalactan in the cell wall (Bhatt et al., 2008). While the precise knowledge of the substrate for Rv2509 is lacking, functional analysis of the preceding enzyme in the biosynthetic pathway, Pks13, indicated that while this polyketide synthase primarily catalyses the formation of the α-alkyl, β-keto fatty acyl precursor, it also contains enzymatic motifs that facilitate the release of the nascent fatty acyl chain and its subsequent transfer to a trehalose residue to produce a trehalose residue esterified with the α-alkyl, β-keto fatty acyl precursor of mycolic acid (Gavalda et al., 2014). These findings suggest that the trehalose-bound mono 3-oxo-mycolic acid precursor may likely be the substrate for Rv2509, which catalyses its conversion to TMM (Figure 1). In this study, we set out to do a detailed study of the amino acid sequence of Rv2509 with the aim of deducing its structure and functional connections, which led to the identification of series of unique sequences, and by proxy, novel structural features. We then probed the role of one distinct structural feature of the mycobacterial mycolyl reductase in vitro by deletion analysis. We also probed the essentiality of Rv2509 for the growth and viability of the slow-growing M. tuberculosis complex using Mycobacterium bovis BCG and Corynebacterium. While the two genera share a common set of enzymes required for mycolic acid biosynthesis, select differences also exist. For example, while the meromycolate chains are synthesised by a Type-II fatty acid synthase complex in species of Mycobacterium, the same function is carried out by a mammalian-like Type-I fatty acid synthase in Corynebacterium species (Radmacher et al., 2005). Similarly, differences exist in the transport of mycolic acids between the two genera (Varela et al., 2012; Yang et al., 2014). To investigate whether the reduction of premycolates by an exclusive mycolyl reductase, was an enzymatic process conserved across other mycolic acid-producing species, we used the amino acid sequence of Rv2509 as a source template for a BLASTp search of genome sequences of other known mycolic acid-producing bacteria. Homologues of Rv2509 were found across the mycolate-producing genera (Figure 2). A phylogenetic analysis of Rv2509, its homologues in other mycobacteria and in select mycolic acid-producing species revealed that mycolyl reductases from Corynebacterial species branched early (Supporting information Figure S1). Corynebacteria produce relatively short chain mycolic acids (Radmacher et al., 2005). To test whether mycolyl reductases across the mycolata had evolved specificities to accommodate differing chain lengths, we transformed the M. smegmatis ΔMSMEG4722 strain with a plasmid-containing C. glutamicum mycolyl reductase gene NCgl2385 to generate the strain ΔMSMEG4722-CNCgl2385. Mature mycolic acid production was restored in the transformed strain, indicating that NCgl2385 could reduce M. smegmatis 3-oxo-mycolic acid precursors and consequently fully complement the mutant M. smegmatis strain. This outcome suggested that mycolyl reductases across the mycolata likely did not evolve specificities for differing chain lengths (Figure 3).

2.2 | Predicted structure of M. tuberculosis Rv2509

We attempted to deduce possible functional properties of Rv2509 using a comparative sequence and structural analysis with proteins of known function. Sequence analysis of Rv2509 using the NCBI conserved domain database (Marchler-Bauer et al., 2015) categorises the protein as a part of the DlTE family of short-chain dehydrogenases (cluster COG0300), which are in turn members of the cl27753 short-chain dehydrogenase/reductase (SDR) superfamily of NAD(P)-dependent oxidoreductases. This is a diverse group, containing over 80,000 unique sequences (Fujisawa, Nagata, & Misono, 2003). However, despite this diversity, all SDR-superfamilies display the classical Rossmann fold typical of NAD+/NADP+-binding proteins and hence can be classified as belonging to the cl21454 family of topologically and structurally related proteins.

Next, we screened the peptide sequence of Rv2509 against the PDB database to find structural templates for homology modelling. The serine dehydrogenase YdfG (UniProt ID P39831; PDB ID 3ASU/3AV) (Yamazawa, Nakajima, Mushiaké, Yoshimoto, & Ito, 2011) had the highest identity score; however, structural analysis revealed several gaps in the pairwise alignment and the C-terminal
region was shorter than that observed in Rv2509, which necessitated additional template input. Therefore, the structures of the four highest-scoring experimental SDRs (3ASU/3ASV, 1XG5, 4X54 and 4BMV) were used in parallel as a basis for comparative homology modelling of Rv2509 (Figure 4) using the I-TASSER package (Yang et al., 2015) with specific template selection enabled. Notably, while, three of the templates yielded models with roughly equivalent (confidence) C-scores, the one based on the SDR from Sphingobium yanoikuyae (4BMV.pdb; GenBank ID ACB78183.1) stood out with the highest C-score of 0.6, indicating a high quality of prediction.

As visualised by the structural alignment (Figure 2) which shows the predicted secondary structure elements of Rv2509 aligned to its Corynebacterial homologues, the core structure of the Rv2509 model presents a typical NADP-binding protein Rossmann fold (residues 1–235) (Persson, Kallberg, Oppermann, & Jornvall, 2003) with seven parallel β-strands forming a central β-sheet, which is
sandwiched between two layers of α-helices—three on each side (Figure 4). Although an earlier study has correctly classified Rv2509 within the short-chain reductase (SDR) superfamily and identified crucial sequence signatures such as the potential NAD(H)/NADP(H)-binding motif and active site in residues 157–161 (Lea-Smith et al., 2007), it did not include further structural prediction and analysis, and was limited in terms of sequences used. We, therefore, sought to expand this initial analysis by performing a systematic analysis of the Rv2509 orthologues from a wide range of Corynebacterineae to represent all known mycolate-producing genera.

We initiated these studies by analysing the cofactor binding site. SDRs can be divided into two large families, “Classical” SDRs which are around 250 amino acid residues long and “Extended” SDRs which contain around 350 residues (Jornvall et al., 1995), exhibiting differences in the arrangement of glycine residues within the coenzyme binding sites. Coenzyme preference-predictions based on the strong conservation of the signature residues (Persson et al., 2003) classified the Rv2509 into the group of the NADP(H)-binding SDRs, consistent with its mycolyl reductase function. However, it was difficult to categorically allocate Rv2509 into a specific group of SDRs as the predicted NADP-binding site of Rv2509 (G16XXQ19N20I21G22) displays a unique and novel variation of the (T)-G-X-X-X-G-I-G motif described by Brakoulias and Jackson (2004) (Figure 2). The variation seems to be restricted to the long-chain mycolates-producing species of Mycobacteriaceae, as short-chain mycolates-producing Corynebacteria retain the more established G-X-X-X-G-I-G motif. (Figure 2). Key residue positions in the NADP(H) site suggest that Rv2509 is a cP2 member of the cP family of SDRs. (Figure 2).
the “Classic” SDR family, according to the classification system of Persson et al. (Brakoulias & Jackson, 2004), which is further reinforced by the presence of two specific discriminator residues, R-R, found in positions 41–42 in both \textit{Rv2509} and \textit{YdfG} from \textit{Escherichia coli} (3ASV.pdb), as these residues are associated with coordination of the phosphate group of the NADP(H) cofactor. Further analysis of \textit{Rv2509} suggests that it has a Class IV NADP(H)-binding site (with consensus [AVIC]-[LIV]-[VIL]-T-G-[ASGC]-X$_3$-[GR]-[ILF]-G-X$_5$-[LFY] (Hua, Wu, Sargsyan, & Lim, 2014). Thus, our analysis firmly places the \textit{Rv2509} family within the NADP(H)-binding subfamily of SDRs.

The active site in SDRs is formed by a tight cleft, lined by \(\alpha\)-helix 4, \(\beta\)-strand 4 and \(\alpha\)-helix 5 (Filling et al., 2002; Oppermann et al., 1997). \textit{Rv2509} exhibits the “classical” SDR motif Yx[AS][ST]K in \(\alpha\)-helix 5 (Y157 and K161 being the central catalytic residues), while the GxgxgS$^/$T motif in \(\beta\)-strand 4 (G137, S144) and the central residue of \(\alpha\)-helix 4, N116 (Figure 2) are also instantly recognisable. In addition, it is possible that some contact with the substrate may be provided by the \(\beta\)-strand 4.

Rather intriguingly, the consensus residues “NNAG” involved in the stabilisation of the central \(\beta\)-sheet of the Rossmann fold in the “Classical” SDRs and located at the center of the protein are not fully conserved in \textit{Rv2509} but take the form of a novel “ANAG” sequence (see Figure 2). Furthermore, this novel sequence is paired with a preceding Cys residue, which is also correlated with the presence of a C-terminal Thr (in position 91 in \textit{M. tuberculosis}). This unique “CANAGT” signature makes \textit{Rv2509} orthologues in \textit{Mycobacteriaceae} instantly recognisable (Figure 2).

In SDR proteins, the substrate-binding loop links the sixth \(\beta\)-sheet with the seventh \(\alpha\)-helix and is one of the larger loops. In the predicted structure of \textit{Rv2509} this loop spans roughly from P$_{187}$ to S$_{214}$ and appears largely conserved within the mycolic acid-producing species (Figure 2). While the loop’s length and position corresponds closely to the length of the substrate-binding loop in both 4BMV (P183-E206) and 3ASV (P178-T208), there was little conservation of the loop’s sequence relative to \textit{Rv2509} in these and other divergent SDRs (see ConSurf analysis below (Ashkenazy et al., 2016); Supporting information Figure S2), suggesting, as expected, major differences in the structure of the substrate for \textit{Rv2509}.

The most intriguing finding upon the analysis of the \textit{Rv2509} orthologues was that all of them seem to present a C-terminal domain which seems to be highly specific to the \textit{Corynebacterineae}. It proved also the most difficult to model reliably, as in the majority of the available SDR structural templates, the C-terminal region, appears to be either truncated or disorganized. ConSurf mapping using the multiple sequence alignment shown in Figure 2 and multiple sequence alignment based on 150 nonredundant sequences of remote SDR homologues revealed that the C-terminus of \textit{Rv2509} is uniquely a feature of the subfamily of Corynebacteriales (Supporting information Figure S3). Both pairwise and multiple sequence analyses indicated that the C-terminal region of the 4BMV.pdb is closely related to \textit{Rv2509} and thus can be tentatively used as a template for that region. This was further corroborated that by comparative modelling (Song et al., 2013) and pure \textit{ab-initio} modelling (Raman et al., 2009), as implemented in the NewRobetta server (http://new.robetta.org). To validate this hypothesis, we conducted additional \textit{ab initio} secondary structure predictions of \textit{Rv2509} (236–268) using JPred4 (Drozdzetskiy, Cole, Procter, & Barton, 2015) and PSI-PRED (Buchan & Jones, 2019; Jones, 1999). Both of these bioinformatics tools supported a high alpha-helical propensity of the C-terminal region of \textit{Rv2509}, suggesting that
4BMV as a reliable template for the region (Supporting information Figure S3). Indeed, the C-terminus of the 4BMV.pdb structure, similarly to Rv2509, appears to be extended and is well-structured, presenting an α-helical organisation.

2.3 | The extended C-terminal of Rv2509 is essential for the function

Mature mycolic acid production could be restored in the M. smegmatis ΔMSMEG4722 mutant on complementation with the M. tuberculosis orthologue Rv2509 (Bhatt et al., 2008). The mutant strain thus provided us with a means of identifying residues and domains required for Rv2509 function, by assessing the ability of mutated constructs of plasmid-borne Rv2509 to complement the ΔMSMEG4722 strain by restoring mature mycolic acid biosynthesis. To test if the extended C-terminal of Rv2509 was required for function, we generated truncated versions of Rv2509 cloned in the replicative shuttle plasmid pMV261 (Stover et al., 1991). Individual plasmids were then introduced by electroporation into the M. smegmatis ΔMSMEG4722 mutant and the transformed strain analysed for the restoration of mature mycolic acid biosynthesis. Deletion of 14 amino acid residues (Ala_{255}-Ser_{268}) from
the C-terminal of Rv2509 resulted in a complete loss of function as the strains transformed with the truncated construct failed to restore mature mycolic acid biosynthesis in the ΔMSMEG4722 strain (Figure 5; ΔMSMEG4722-CRVdI). A construct with a shorter truncation (Tyr262-Ser268 deleted) also failed to complement the ΔMSMEG4722 strain (Figure 5; ΔMSMEG4722-CRVdI). Expression of Rv2509 and truncated constructs could be detected by RT-PCR confirming that the loss of Rv2509 function was linked to the deletion of the C-terminal residues (Supporting information Figure S4). These results demonstrated that the C-terminus of Rv2509 spanning terminal residues Ala255-Ser268 was critical for function. The findings suggested an essential role for the unique, C-terminal extension in the activity of Rv2509.

To further query whether this C-terminal extension played a role in either protein stability, protecting Rv2509 from degradation, or in oligomerisation, we expressed N-terminal His-tagged versions of Rv2509 and Rv2509-D1 (Rv2509 with Ala255-Ser268 deleted) in E. coli. Western blot analysis of native polyacrylamide gels showed that there were no patterns of degradation detected in Rv2509-D1, indicating that the loss of function due to the deletion of this region was not related to protein stability (Supporting information Figure S5).

Furthermore, we did not observe additional slower migrating bands, indicating dimerisation (or oligomerisation) of full-length or truncated Rv2509. Furthermore, cross-linking Rv2509 with glutaraldehyde did not reveal additional slower migrating bands, suggesting that Rv2509 did not oligomerise in solution (Supporting information Figure S5).

SECTION 2.4 | Rv2509 is an essential mycobacterial gene

Given the nonessentiality of the mycolyl reductase gene in M. smegmatis and prediction of transposon-site hybridisation screens of a slow-growing M. tuberculosis Rv2509 mutant, we aimed to generate a null mutant of Rv2509 in M. tuberculosis using Specialised Transduction (Bardarov et al., 2002). Such a mutant would also allow us to test the role of the mycolyl reductase in virulence utilising macrophage and mouse models of infection. However, repeated attempts failed to generate transductants, suggesting that Rv2509 was essential for growth in M. tuberculosis. Indeed, subsequent transposon mutagenesis screens utilising deep sequencing predicted Rv2509 to be essential for in vitro growth of M. tuberculosis (Griffin et al., 2011). We then utilised a promoter replacement strategy to demonstrate the essentiality of Rv2509 in slow-growing mycobacterial vaccine strain Mycobacterium bovis BCG. The Pmyc1 promoter from M. smegmatis engineered to contain four tetO operator sites (Korte et al., 2016) was inserted immediately upstream of the start codon of BCG2529, the Rv2509 homologue in M. bovis BCG-Pasteur, using Specialised Transduction generating the strain BCG::PTet<sup>-</sup>BCG2529. Controlled gene expression of the BCG2529 gene was achieved using a plasmid-borne synthetic gene (rev-tetR) derived from Tn10 tetR encoding a mutated TetR protein with reversed binding affinity to tetO sites upon the binding of tetracycline. The addition of anhydrotetracycline (ATc) results in the loss of expression.

The addition of ATc resulted in the loss of growth of BCG::PTet<sup>-</sup>BCG2529 in broth, indicating that the expression of the gene encoding the mycolyl reductase was essential for the in vitro growth of slow-growing mycobacteria (Figure 6a). Furthermore, while concentrated cultures of BCG::PTet<sup>-</sup>BCG2529 showed confluent growth on 7H10 agar plates, no confluent growth was observed on plates containing ATc (Figure 6b). We did observe a few scattered colonies on the ATc-containing plates and these are likely to be suppressors allowing leaky expression of PTet<sup>-</sup>driven BCG2529. These results demonstrated that the mycolyl reductase was essential for the growth and viability of slow-growing mycobacteria.

SECTION 2.5 | Conditional depletion of the mycolyl reductase results in the loss of mature mycolic acids

The M. smegmatis mycolyl reductase mutant was unable to synthesise fully mature mycolic acids and instead produced 3-oxo-mycolate precursors that were found esterified to the cell wall arabinogalactan and to trehalose (Bhatt et al., 2008). These labile precursors gave rise to palmitone derivatives following alkali treatment, which were visualised on Thin Layer Chromatography (TLC) plates migrating close to the
solvent front in a solvent system used for separating methyl esters of mycolic acids and fatty acids (FAMEs and MAMEs) (Bhatt et al., 2008).

To test the effects of conditional depletion of the mycolyl reductase we labelled the cultures of BCG::PTet-BCG2529 with 14C-acetic acid following grown in the presence and absence of ATc, and extracted MAMEs from the cultures. TLC analysis of the extracted MAMES showed a decrease in the levels of α and keto MAMES and an accumulation of the degradation products derived from 3-oxo-mycolate precursors close to the solvent front (Figure 7). These results showed that, as in M. smegmatis, loss of the mycolyl reductase function led to the loss of mature mycolic acid production and accumulation of the premature 3-oxo-mycolate precursors in slow grow mycobacteria.

3 | DISCUSSION

As summarised above, the predicted structure of the Rv2509 displays a Rossmann fold with seven parallel β-strands forming a central beta-sheet, which is sandwiched between two layers of alpha-helices—three on each side (Figure 4). While largely consistent with the "classical" family of SDRs (Persson et al., 2003), in Mycobacteriaceae, the NADP(H)-binding motif (G-x-x-Q-N-I-G), alongside the core stabilisation consensus (C-A-N-A-G-T) is significantly divergent from the consensus, to instantly identify this group of SDRs, and in our view is distinct enough to be separated into a distinct family.

While we currently lack knowledge about the exact mode of co-factor and ligand binding within the Rv2509 family, some information could be inferred from the structurally related FabG family of β-ketoacyl reductases which are classified as "Complex family SDRs" (Persson et al., 2003). Given that the structural overlay of Rv2509 models with existing ligand-bound FabG members yields an RMSD of around 1 Å over the C-alpha backbone atoms, and that FabG share common active site residues with Rv2509 it is tempting to suggest that these structures share common ligand-binding and NADP(H) orientations towards the active site residues with that observed in the aceto-acyl-CoAs. Furthermore, the predicted 3D structure of Rv2509 also revealed a unique α-helical C-terminal extension, which we subsequently showed to be essential for function—constructs with deletions in the C-terminal failed to complement (restore mycolate biosynthesis) in the M. smegmatis mutant.

The role of this C-terminal domain is difficult to assign with certainty; however, in related SDRs similar C-terminal extensions play a number of functions. In the NADP(H)-dependent tetrameric serine-dehydrogenase YdfG from E. coli (3ASV.pdb) (Yamazawa et al., 2011), the C-terminal region protrudes from the main core of the protein and is required for tetramer formation, while it also plays a role in the substrate binding via stabilisation of the substrate loop. It is also notable that in FabG family of related SDRs discussed above also organise as tetrameric assemblies with the participation of their C-terminal domains (Javidpour et al., 2014). Furthermore, in the mycobacterial FabG4 dimer, Arg146 and Arg445 of one protomer interact with the C-terminus of the second protomer and play an essential role in the substrate association and catalysis (Dutta, Bhattacharyya, Roychowdhury, Biswas, & Das, 2013). However, our experiments with purified full-length and truncated Rv2509 do not seem to suggest that the mycolyl reductase forms oligomers.

While it isn’t in itself surprising that Rv2509 presents a conserved SDR fold, our analysis has identified it for the first time decisively as a member of the NADPH-binding family and indicated that it possesses a number of unique features. Taken together with the unique signatures of the NADPH-binding site, core stabilisation domains, as well as C-terminal domain extension, the unique features of Rv2509 possibly merit its inclusion into a distinct family.

Our previous studies with M. smegmatis suggested that the mycolyl reductase would likely not be essential for the growth in M. tuberculosis and related slow-growing mycobacteria. However, this work demonstrated that gene encoding the reductase, Rv2509, was an essential gene required for the growth of slow-growing mycobacteria in laboratory media. This, along with the identification of unique domains in the predicted structure of Rv2509 highlights its potential as a novel anti-TB drug target.
4 | EXPERIMENTAL PROCEDURES

4.1 | Growth conditions, strains, phages and plasmids

*M. smegmatis* mc²155 and derived strains was cultured in Tryptic Soy Broth (TSB) or on TSB-agar, *M. bovis* BCG strains were grown in 7H9 broth or on 7H10 agar plates. *E. coli* strains were grown in LB broth or LB agar. Hygromycin was used at a concentration of 100 mg/ml for *E. coli* and mycobacterial strains. Kanamycin was used at a concentration of 25 mg/ml and 50 mg/ml for mycobacteria and *E. coli* respectively. Recombinant mycobacterial strains, phages and plasmid constructs used in this study are listed and described in Table 1.

4.2 | Sequence analysis and homology modelling

The multiple sequence alignments (MSA) were prepared using MAFFT and NJ/UPGMA phylogeny algorithms as implemented in MAFFT v.7 server (https://mafft.cbrc.jp/ (Katoh, Misawa, Kuma, & Miyata, 2002)) and the resulting phylogenetic trees were visualised using Archaeopteryx (Han & Zmasek, 2009). Multiple sequence alignment visualisation and structural annotations were performed using ESPript 3 (Gouet, Robert, & Courcelle, 2003). For the purposes of homology modelling, we employed I-TASSER (Yang et al., 2015) with the assignment of templates. Outputs were independently corroborated by ab-initio modelling using the New Robetta server (http://new.robetta.org). Sequence conservation analysis was performed using ConSurf (Ashkenazy et al., 2016) and visualised using PyMOL (The PyMOL Molecular Graphics System, Version 1.71 Schrödinger, LLC) with additional local manual refinement and structural superposition performed in Coot (Emsley, Lohkamp, Scott, & Cowtan, 2010). Secondary structure predictions for the C-terminal domain were performed with JPred4 (Cole et al., 2008) and PSIPRED (Buchan et al., 2013).

4.3 | Expression, purification and analysis of Rv2509 and its C-terminal truncated derivative

Rv2509 was cloned in the expression vector pET28a by PCR from *M. tuberculosis* H37Rv genomic DNA using the primer pair Pet28a_Rv2509_F2 (5'-TTTACATATGCCGAGCGCGCCGCCC-3') and Pet28a_Rv2509_R1 (5'-ATTAAAGCTTCTAGCTGCCCCCAA...
TABLE 1 Plasmids and bacterial strains used in this study

| Description | Reference/Source |
|-------------|------------------|
| **Plasmids** |                  |
| pMV261      | E. coli-Mycobacterium shuttle plasmid vector with hsp60 promoter and Kan<sup>6</sup> cassette (aph) | Stover et al. (1991) |
| pMV261-Rv2509 | Rv2509 cloned in pMV261 | Bhatt et al. (2008) |
| pMV261-NCgl2385 | NCgl2385 cloned in pMV261 | This work |
| pMV261-Rv2509D1 | Deletion construct of Rv2509 cloned in pMV261 encodes a truncated protein missing the C-terminal residues Ala<sub>255</sub>-Ser<sub>268</sub> | This work |
| pMV261-Rv2509D2 | Deletion construct of Rv2509 cloned in pMV261 encodes a truncated protein missing the C-terminal residues Tyr<sub>262</sub>-Ser<sub>268</sub> | This work |
| pMV261::rev-tetR-RBS-D | Episomal E. coli-mycobacterium shuttle plasmid pMV261::rev-tetR-RBS-D providing constitutive rev-tetR gene expression from the HSP60 promoter in mycobacteria | Korte et al. (2016) |
| pET28a-Rv2509 | Plasmid for expressing Rv2509 in E. coli | This work |
| pET28a-Rv2509D1 | Plasmid for expressing C-terminal truncated Rv2509 in E. coli | This work |
| **Bacterial strains** |                  |
| mc<sup>155</sup> | Electroperoration-proficient pmt mutant of M. smegmatis strain mc<sup>2</sup>6 | Snapper, Melton, Mustafa, Kieser, and Jacobs (1990) |
| ΔMSMEG4722 | Deletion mutant of mc<sup>155</sup> in which MSMEG4722 is replaced by <sup>hyg</sup> | Bhatt et al. (2008) |
| ΔMSMEG4722-CrV | ΔMSMEG4722-containing pMV261-Rv2509 | Bhatt et al. (2008) |
| ΔMSMEG4722-NCgl2385 | ΔMSMEG4722-containing pMV261-NCgl2385 | This work |
| ΔMSMEG4722-CrVd1 | ΔMSMEG4722-containing pMV261-Rv2509D1 | This work |
| ΔMSMEG4722-CrVd2 | ΔMSMEG4722-containing pMV261-Rv2509D2 | This work |
| BCG::P<sub>tet</sub>-BCG2529 | BCG Pasteur strain containing the Pmyc1 promoter from M. smegmatis engineered to contain four tetO operator sites, inserted immediately upstream of the start codon of BCG2529, the BCG homologue of Rv2509 | This work |

GCCTCTTG-3′). Similarly, the C-terminal truncated version of Rv2509 (missing the last 14 amino acids) was PCR amplified using the primer pair Pet28a_Rv2509_F2 (5′-TTTACATA GGCTCTTG-3′) and Pet28a_Rv2509_delR1 (5′-ATTAAAAGCTTCTACA CGATGGCGCGGGAGC-3′). The constructs were cotransformed into BL21 cells with M. tuberculosis chaperone GroES 60.2. Cultures were seeded 1:100 with an overnight starter culture and grown at 37°C in terrific broth supplemented with ampicillin (100 µg/ml) and kanamycin (50 µg/ml) to OD<sub>600</sub> = 0.6 and induced with 0.5 mM IPTG overnight at 37°C or 16°C respectively. Cells were harvested at 4,000 rpm at 4°C, washed with PBS and frozen at −80°C until further use. Full-length protein was purified on cobalt-IMAC in buffer 50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 7.5, 500 mM NaCl, 10% glycerol and truncated in the same buffer composition at pH 7.0. The eluted fraction of 200 mM imidazole was taken for native-PAGE analysis.

A native polyacrylamide gel with the separated samples was transferred to a Hybond<sup>TM</sup> nitrocellulose membrane (GE Healthcare Life Sciences) which was then blocked in 5% milk 20 mM Tris pH 7.5, 150 mM NaCl TBS 0.05% Tween for 1h at room temperature followed by incubation with 0.1% milk TBS-tween with penta-his antibody BSA free (Qiagen). Following washing with TBS-tween the membrane was incubated with 0.1% milk TBS-tween with anti-mouse IgG-Alkaline Phosphatase (Sigma Aldrich) for 30 min at room temperature, washed with TBS-tween and a final wash with TBS. The membrane was stained for 5 min with BCI/P/NBT (SIGMAFAST™) in water for visualisation.

Cross-linking of full-length Rv2509 was carried out in 50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 7.5, 10% glycerol at a final concentration of 0.1% glutaraldehyde and allowed to progress at room temperature for 2 min. The reaction was halted with 1M Tris pH 8 and subsequently run on SDS-PAGE for visualisation using SYPRO-ruby (Molecular Probes).

4.4 | Conditional depletion of the Rv2509 homologue BCG2529 in M. bovis BCG-Pasteur

For establishing the regulated expression of the BCG2529 gene, a synthetic gene cassette (hyg-Pmyc1-4XtetO) comprising a hygromycin-resistance gene and the Pmyc1 promoter from M. smegmatis engineered to contain four tetO operator sites (Korte et al., 2016) was inserted immediately upstream of the BCG2529 start codon in M. bovis BCG-Pasteur. Targeted gene knock-in was achieved by specialised transduction employing temperature-sensitive mycobacteriophages essentially as described previously (Korte et al., 2016). Briefly, for the generation of an allelic exchange construct for site-specific insertion
of the hyg-Pmyc1-4XtetO cassette in *M. bovis* BCG-Pasteur, upstream- and downstream DNA regions flanking the BCG2529 start codon were amplified by PCR employing the oligonucleotide pairs Mb2537_LL 5′-TTTTCCATAAATTTGGAAGGCTTACGATGAAACCC-3′ and Mb2537_LR 5′-TTTTCCATTTCCTTGGCGGATTTCTGCGAGCCCGCCG-3′ as well as Mb2537_RL 5′-TTTTCCATAAGTTGATGCCGATACCCGCGCCCAGCCCGC-3′ and Mb2537_RR 5′-TTTTTCCATTTTGGGCCGCTGTTGGGAGGATACCTCGAG-3′ respectively. The gene BCG2528c is localised in antilinear orientation upstream of BCG2529 with only 84 bp between the start codons of both genes. Since it could not be excluded that the Mb2536c promoter region might overlap with the BCG2529 coding region, the upstream flanking region contained a duplication of 66 bp of the 5′-end of BCG2529 to conserve 150 bp in front of BCG2529 that likely comprised its promoter. Subsequently, the upstream and downstream flanks were digested with VanII (restriction sites underlined) and ligated with VanII-digested pcRv1327c-4XtetO vector arms (Korte et al., 2016). The resulting knock-in plasmid was then linearised of the episomal -mycobacterium shuttle plasmid pMV261::E. coli tetO sites upon the TetR protein with reversed binding affinity to the HSP60 promoter in mycobacteria using solid medium containing 50 mg/L of hygromycin (50 mg/L) for selection, resulting in the site-specific insertion of the hyg-Pmyc1-4XtetO cassette. The obtained *M. bovis* c-BCG_2529-4XtetO knock-in mutant was verified by the diagnostic PCR of genomic DNA using the oligonucleotide pair BCG_2529_L_fw 5′-GTCAGGTAGACGGAGAACAC-3′ and BCG_2529_L_rev 5′-AGCTCACCAGCGCAGATTCC-3′ binding outside the allelic exchange substrates used to generate this mutant and subsequent sequencing of PCR products (Supporting information Figure S6). For achieving the controlled gene expression of the Mb2537 gene, a synthetic gene (rev-tetR) derived from Tn10 tetR encoding a mutated TetR protein with reversed binding affinity to tetO sites upon the binding of tetracycline (Klotzsche, Ehrth, & Schnappinger, 2009) was heterologously expressed in the knock-in mutant by electroporation of the episomal E. coli-mycobacterium shuttle plasmid pMV261::rev-tetR-RBS-D providing constitutive rev-tetR gene expression from the HSP60 promoter in mycobacteria using solid medium containing 50 mg/L of hygromycin and 20 mg/L of kanamycin for selection (Famulla et al., 2016). This yielded the conditional mutant BCG-Pasteur c-BCG2529-4XtetO pMV261::rev-tetR-RBS-D (referred to as BCG::P<sub>Tet</sub>BCG2529) allowing silencing of the BCG2529 gene in the presence of anhydrotetracycline (ATc).

For silencing experiments, precultures of the BCG::P<sub>Tet</sub> BCG2529 strain were first grown in liquid medium containing 0.3 µg/ml of ATc, before these precultures were used to inoculate the test cultures containing 0.5 µg/ml of ATc. Growth in 96-well microtitre plates was quantified after incubation at 37°C for five days by adding 10% of 100 µg/ml of resazurin solution. After further incubation at ambient temperature for 16 hr, cells were fixed for 30 min by formalin addition (5%, v/v, final concentration) and fluorescence was quantified using a microplate reader (excitation 540 nm, emission 590 nm). For the growth on 7H10 agar, nonpermissive conditions were obtained by adding ATc to a final concentration of 5 µg/ml.

### 4.5 | Functional complementation studies

Deletion constructs for *Rv2509* were generated by PCR amplification of selected regions of *Rv2509* using pMV261-Rv2509 as a template, incorporating a premature stop codon in the reverse primer. The plasmid constructs generated in this manner were called pMV261-Rv2509D1 and pMV261-Rv2509D2, encoding a truncated protein missing the C-terminal residues Ala<sub>255</sub>-Ser<sub>268</sub> and Tyr<sub>262</sub>-Ser<sub>268</sub> respectively. The *C. glutamicum* Rv2509 orthologue, NCgl2385 was PCR amplified using *C. glutamicum* genomic DNA as a template and the PCR product was cloned into pMV261 using primer incorporated BamHI and EcoRI sites generating the plasmid pMV261- NCgl2385. Plasmid clones were electroporated into the *M. smegmatis* ΔMSMEG4722 mutant strain using previously described protocols (Snapper et al., 1990). Kanamycin-resistant transformants were cultured in TSB and FAMEs and MAMEs were extracted from the cell pellets and separated by TLC as previously described (Vilcheze & Jacobs, 2007).

### 4.6 | RT-PCR of *Rv2509* transcripts

*M. smegmatis* strains were grown to mid-log phase and RNA was extracted by acid phenol and chloroform extraction. The purified RNA was treated with DNase and converted to cDNA, which was used as a template for quantitative PCR. PCR products were run on a 1% agarose gel to detect the expression of *Rv2509*. The gene sigA was monitored as a constitutively expressed gene. The primers used for PCR amplification were Rv2509RT-f (5′-ACAAGTACCGCGTCACGGTC-3′) and Rv2509RT-r (5′-AAGTCCGGCACCAGCTTCTC-3′) for the amplification of nondeleted sections of *Rv2509*.

### ACKNOWLEDGEMENTS

CC is funded by a BBSRC PhD studentship provided by the BBSRC Midlands Integrative Biosciences Training Partnership (MIBTP). AB acknowledges previous grant support from the Medical Research Council (UK). RK acknowledges financial support from the Jürgen Manchot Foundation.

### AUTHOR CONTRIBUTIONS

AB, RK and VB designed the study. AJ, CC, AS, SS, MH and RM conducted experiments and acquired data. AJ, CC, AS, SS, MH, RM, AB, RK and VB interpreted the data. AB, RK and VB wrote the manuscript.

### ORCID

Vassily N. Bavro https://orcid.org/0000-0003-2330-8924
Apoorva Bhatt https://orcid.org/0000-0002-6655-1636
Persson, B., Kallberg, Y., Oppermann, U., & Jornvall, H. (2003). Coenzyme-based functional assignments of short-chain dehydrogenases/reductases (SDRs). *Chemico-Biological Interactions*, 143-144, 271-278. https://doi.org/10.1016/S0009-2797(02)00223-5

Portevin, D., De Souza-D’Auria, C., Houssin, C., Grimaldi, C., Chami, M., Daffe, M., & Guilhot, C. (2004). A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 314–319. https://doi.org/10.1073/pnas.030549101

Radmacher, E., Alderwick, L. J., Besra, G. S., Brown, A. K., Gibson, K. J., Sahm, H., & Eggeling, L. (2005). Two functional FAS-I type fatty acid synthases in *Corynebacterium glutamicum*. *Microbiology*, 151, 2421–2427. https://doi.org/10.1099/mic.0.28012-0

Raman, S., Vernon, R., Thompson, J., Tyka, M., Sadreyev, R., Pei, J., ... Baker, D. (2009). Structure prediction for CASP8 with all-atom refinement using Rosetta. *Proteins*, 9, 89–99.

Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T., & Jacobs, W. R., Jr. (1990). Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Molecular Microbiology*, 4, 1911–1919.

Song, Y., DiMaio, F., Wang, R. Y., Kim, D., Miles, C., Brunette, T., ... Baker, D. (2013). High-resolution comparative modeling with RosettaCM. *Structure*, 21, 1735–1742. https://doi.org/10.1016/j.str.2013.08.005

Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., ... Bloom, B. R. (1991). New use of BCG for recombinant vaccines. *Nature*, 351, 456–460. https://doi.org/10.1038/351456a0

Varela, C., Rittmann, D., Singh, A., Krumbach, K., Bhatt, K., Eggeling, L., ... Bhatt, A. (2012). MmpL genes are associated with mycolic acid metabolism in mycobacteria and corynebacteria. *Chemistry & Biology*, 19, 498–506. https://doi.org/10.1016/j.chembiol.2012.03.006

Vilcheze, C., & Jacobs, W. R. (2007) Isolation and analysis of *Mycobacterium tuberculosis* mycolic acids. *Current Protocols in Microbiology, Chapter 10, Unit 10A.3.

Yamazawa, R., Nakajima, Y., Mushiake, K., Yoshimoto, T., & Ito, K. (2011). Crystal structure of serine dehydrogenase from *Escherichia coli*: Important role of the C-terminal region for closed-complex formation. *Journal of Biochemistry*, 149, 701–712. https://doi.org/10.1093/jb/mvr024

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, 12, 7–8. https://doi.org/10.1038/nmeth.3213

Yang, L., Lu, S., Belardinelli, J., Huc-Claustre, E., Jones, V., Jackson, M., & Zgurskaya, H. I. (2014). RND transporters protect *Corynebacterium glutamicum* from antibiotics by assembling the outer membrane. *Microbiologyopen*, 3, 484–496.

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section. FigS1-S6

---

**How to cite this article:** Javid A, Cooper C, Singh A, et al. The mycolic acid reductase Rv2509 has distinct structural motifs and is essential for growth in slow-growing mycobacteria. *Mol Microbiol*. 2019;00:1–13. https://doi.org/10.1111/mmi.14437