

ARTICLE

DOI: 10.1038/s41467-018-04772-x

OPEN

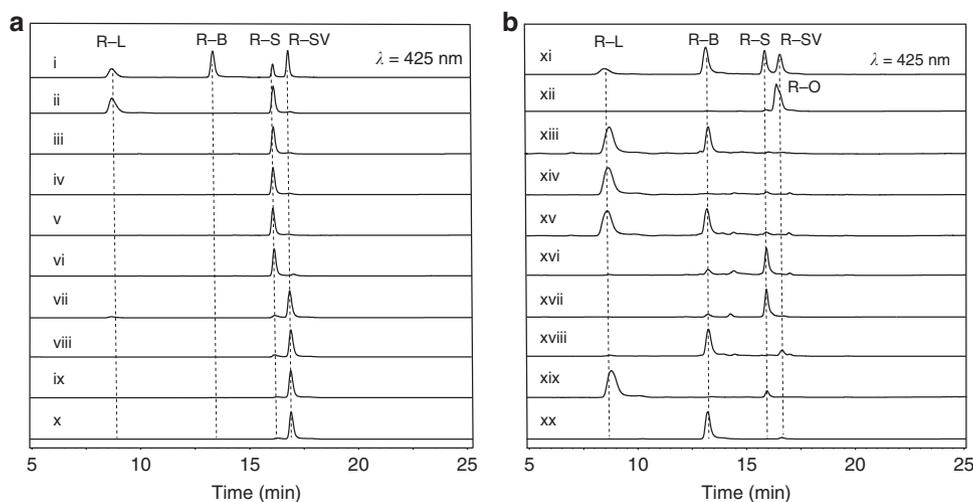
# Deciphering the late steps of rifamycin biosynthesis

Feifei Qi<sup>1</sup>, Chao Lei<sup>2</sup>, Fengwei Li<sup>1</sup>, Xingwang Zhang<sup>1</sup>, Jin Wang<sup>2</sup>, Wei Zhang<sup>1</sup>, Zhen Fan<sup>2</sup>, Weichao Li<sup>2</sup>, Gong-Li Tang<sup>3</sup>, Youli Xiao<sup>2,4</sup>, Guoping Zhao<sup>2,4</sup> & Shengying Li<sup>1,4</sup>

Rifamycin-derived drugs, including rifampin, rifabutin, rifapentine, and rifaximin, have long been used as first-line therapies for the treatment of tuberculosis and other deadly infections. However, the late steps leading to the biosynthesis of the industrially important rifamycin SV and B remain largely unknown. Here, we characterize a network of reactions underlying the biosynthesis of rifamycin SV, S, L, O, and B. The two-subunit transketolase Rif15 and the cytochrome P450 enzyme Rif16 are found to mediate, respectively, a unique C–O bond formation in rifamycin L and an atypical P450 ester-to-ether transformation from rifamycin L to B. Both reactions showcase interesting chemistries for these two widespread and well-studied enzyme families.

<sup>1</sup>Shandong Provincial Key Laboratory of Synthetic Biology, CAS Key Laboratory of Biofuels, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, Shandong 266101, China. <sup>2</sup>CAS Key Laboratory of Synthetic Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 200032 Shanghai, China. <sup>3</sup>State Key Laboratory of Bio-Organic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 200032 Shanghai, China. <sup>4</sup>University of Chinese Academy of Sciences, 100049 Beijing, China. These authors contributed equally: Feifei Qi, Chao Lei. Correspondence and requests for materials should be addressed to Y.X. (email: [ylxiao@sibs.ac.cn](mailto:ylxiao@sibs.ac.cn)) or to S.L. (email: [lishengying@qibebt.ac.cn](mailto:lishengying@qibebt.ac.cn))





**Fig. 2** HPLC analysis of the reactions catalyzed by Rif15 and Rif16. **a** Reactions catalyzed by the transketolase Rif15. (i) The mixed R-L, R-B, R-S, and R-SV standards; (ii) R-S with Rif15 in the presence of F-6-P, ThDP and  $Mg^{2+}$ ; (iii–vi) the negative controls of (ii) with the omission of Rif15a (iii), Rif15b (iv), ThDP (v), or  $Mg^{2+}$  (vi); (vii) R-SV with Rif15 in the presence of F-6-P, ThDP and  $Mg^{2+}$ ; (viii) R-SV in reaction buffer; (ix) R-SV in reaction buffer with the addition of 2 mM ascorbic acid; and (x) R-SV with Rif15 in the presence of F-6-P, ThDP,  $Mg^{2+}$ , and 2 mM ascorbic acid. **b** Reactions catalyzed by the cytochrome P450 enzyme Rif16. (xi) The mixed R-L, R-B, R-S, and R-SV standards; (xii) The freshly prepared R-O authentic standard; (xiii) R-L with Rif16 in the presence of *seFdx*, *seFdr*, and NADPH; (xiv) the negative control of (xiii) with the omission of NADPH; (xv) co-injection of (xiii) with 50  $\mu$ M R-B; (xvi) R-L with Rif16 in the presence of 20 mM  $H_2O_2$ ; (xvii) R-O in reaction buffer; (xviii) R-L with Rif16 in the presence of 20 mM  $H_2O_2$  and 1 mM NADPH; (xix) R-L with 20 mM  $H_2O_2$ ; (xx) R-O and 1 mM NADPH in reaction buffer

## Results and Discussion

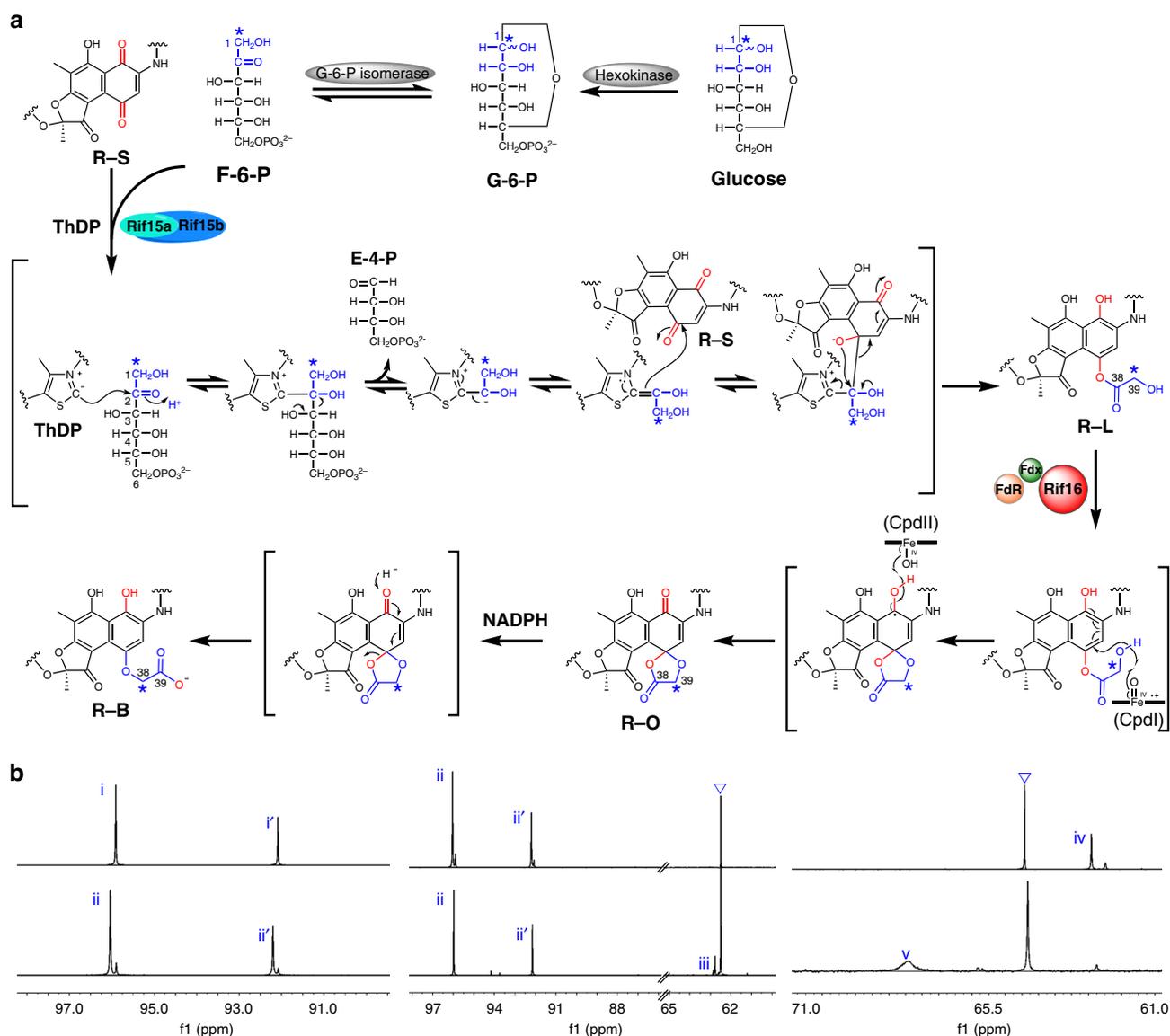
**Initial examination of the enzymatic activity of Rif16.** Generally, P450 enzymes catalyze oxidative reactions<sup>18,19</sup> and transketolases can, in the presence of the essential cofactor thiamine diphosphate (ThDP), transfer a  $C_2$  keto-containing fragment from a 2-ketose (e.g., fructose-6-phosphate (F-6-P), xylulose-5-phosphate (Xu-5-P), ribulose-5-phosphate (Ru-5-P), sedoheptulose-7-phosphate (S-7-P), dihydroxyacetone (DHA), etc.) to the first carbon atom of an aldose (e.g., ribose 5-phosphate, glyceraldehyde 3-phosphate, etc.)<sup>20</sup>. Based on the previous proposal that R-SV could be a biosynthetic precursor of R-B<sup>2</sup>, we first surmised that Rif16 (CYP105G1<sup>21</sup>) may oxidize R-SV to R-S and that R-S (containing a C-4 keto group) may be a substrate of Rif15. To test these enzymatic hypotheses, we sub-cloned *rif15a*, *rif15b*, and *rif16* (Supplementary Fig. 2) and heterologously expressed these genes in *Escherichia coli* Codon Plus (DE3)-RIPL. We then used Ni-NTA chromatography to purify the *N*-terminally His<sub>6</sub>-tagged Rif15a, Rif15b, and Rif16 to homogeneity (Supplementary Fig. 3). Notably, the *N*-His<sub>6</sub>-tagged Rif15a and the non-tagged Rif15b were able to be co-expressed and co-purified, suggesting a strong interaction between these two Rif15 subunits (Supplementary Fig. 3).

Purified Rif16 appeared to be a functional P450 enzyme, as it had the expected red color and showed a signature peak at 450 nm in its CO-reduced difference spectrum (Supplementary Fig. 4). To reconstitute the *in vitro* activity of Rif16, we used two surrogate redox partner proteins to shuttle electrons from NADPH to the heme-iron reactive center for P450 catalysis: the ferredoxin *seFdx* (SynPcc7942\_1499) and the ferredoxin reductase *seFdr* (SynPcc7942\_0978), both of which are from the cyanobacterium strain *Synechococcus elongatus* PCC 7942 and were here expressed heterologously in *E. coli* and purified<sup>22</sup>. Against our expectations, Rif16 was not able to catalyze the conversion from R-SV to R-S, while R-S was readily reduced to R-SV by addition of NADPH alone (Supplementary Fig. 5). Importantly, the hydroquinone R-SV was spontaneously oxidized to the quinone R-S by ambient  $O_2$ , and this transformation was dramatically accelerated by the presence of divalent metal ions

(e.g.,  $Cu^{2+}$ ,  $Mn^{2+}$ , etc.) (Supplementary Fig. 6), similar to previously reported findings<sup>23</sup>. However, we cannot exclude the possibility that an oxidase might be responsible for enzymatic oxidation of R-SV into R-S *in vivo*. Taken together, our results suggest that Rif16, rather than performing a normal bio-oxidation, may catalyze an atypical P450 reaction in rifamycin biosynthesis.

**Functional characterization of Rif15.** We next evaluated the *in vitro* activity of Rif15a/Rif15b at a 1:1 ratio (i.e., the reconstituted Rif15 transketolase) in the presence of R-S and F-6-P as the potential  $C_2$  keto acceptor and donor, respectively, with ThDP and  $MgCl_2$  as cofactors. As predicted, Rif15 converted R-S into a different product with higher polarity than R-B, while single subunits (that is, either Rif15a or Rif15b alone) were not able to catalyze the same transformation. Additionally, we found that both ThDP and  $Mg^{2+}$  were required for the catalytic activity of Rif15 (Fig. 2a, trace i–vi). This is unsurprising since the diphosphate moiety of ThDP is bound to the transketolase through a bivalent cation to form the catalytically active *holo*-enzyme from the *apo*-protein<sup>24,25</sup>. Protein sequence alignment of multiple transketolases shows that the residues involved in the interactions with ThDP and the metal ion are highly conserved regardless their origins and subunit organization modes (Supplementary Fig. 7).

High performance liquid chromatography-high resolution mass spectrometry (HPLC-HRMS) analysis revealed that the *m/z* value of the product was 754.3069 ( $[M-H]^-$ , deduced to be  $[C_{39}H_{48}NO_{14}]^-$ ) (Supplementary Fig. 8), which is consistent with that of R-L or R-B in negative ion mode (*calc.* 754.3069). Since the retention time of this product was distinct from that of R-B, we suspected that the product here is R-L. Both the 1D ( $^1H$ ,  $^{13}C$ , DEPT135) and 2D ( $^1H$ - $^1H$  COSY, HSQC, HMBC) NMR spectra (Supplementary Figs. 9–13, Supplementary Table 2) of the purified product were acquired, and spectral comparisons of the proton NMR data obtained from the product and the substrate R-S (Supplementary Fig. 9) revealed a new set of  $CH_2$  signals ( $\delta_{H-39}$



**Fig. 3** Putative mechanisms for Rif15 and Rif16 supported by  $^{13}\text{C}$ -tracer NMR results. **a** The proposed catalytic mechanisms for Rif15 and Rif16. The  $^{13}\text{C}$  labeled carbons marked by asterisks originate from  $[1-^{13}\text{C}]$ glucose. **b** The  $^{13}\text{C}$  NMR spectra of  $[1-^{13}\text{C}](\pm)$ -glucose (i, i'),  $[1-^{13}\text{C}](\pm)$ -G-6-P (ii, ii'),  $[1-^{13}\text{C}]$ F-6-P (iii) in  $\text{D}_2\text{O}$ , and  $[39-^{13}\text{C}]$ R-L (iv) and  $[38-^{13}\text{C}]$ R-B (v) in  $\text{CD}_3\text{OD}$ . The triangles indicate the carbon signals of residual glycerol derived from enzymatic reaction buffer

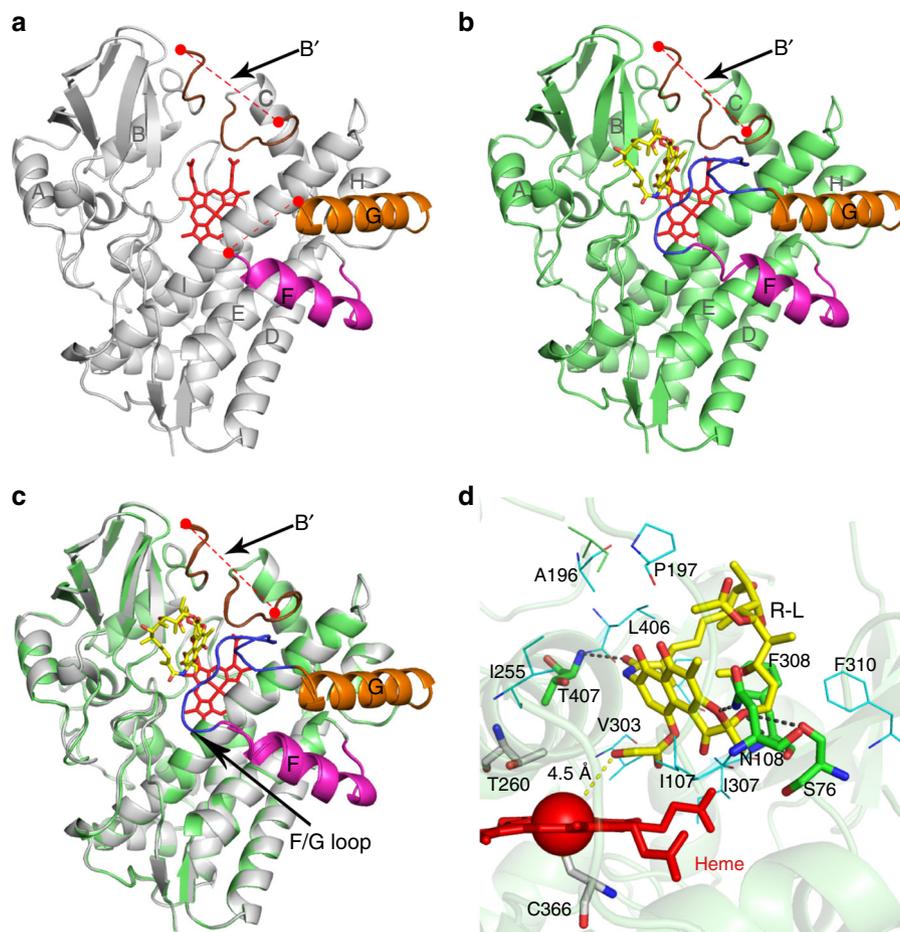
4.63, *d*,  $J = 17.1$  Hz, 1 H,  $\delta_{\text{H}-39}$  4.56, *d*,  $J = 17.1$  Hz, 1 H vs  $\delta_{\text{H}-38}$  4.72, *s*, 2 H in R-B;  $\delta_{\text{C}-39}$  62.3 vs  $\delta_{\text{C}-38}$  67.8 in R-B) from the product. Further extensive analyses of 2D NMR correlations of this methylene group confirmed it to be R-L. The steady-state kinetic constants of Rif15 were determined with  $K_m = 8.8 \pm 2.4$   $\mu\text{M}$  and  $k_{\text{cat}} = (2.2 \pm 0.1) \times 10^{-2} \text{ min}^{-1}$ . Moreover, Xu-5-P, Ru-5-P, S-7-P, and DHA were also able to serve as alternative  $\text{C}_2$  keto donors for Rif15, with Xu-5-P being optimal in terms of conversion ratios under the same conditions (Supplementary Fig. 14).

To examine whether R-SV is also a direct precursor of R-L as previously proposed<sup>2</sup>, similar Rif15 reactions were performed using R-SV as substrate, and, interestingly, we detected a small amount of R-L as a product (Fig. 2a, trace vii). In light of the spontaneous oxidation from R-SV to R-S by  $\text{O}_2$  that we had observed in earlier assays (in aqueous solution in the presence of  $\text{Mg}^{2+}$ , Supplementary Fig. 6), 2 mM ascorbic acid was added to the Rif15 reactions to protect R-SV from oxidation (Fig. 2a, trace viii, ix)<sup>26</sup>. Upon this addition, we no longer detected R-L as a

reaction product (Fig. 2a, trace x), establishing that the previously detected R-L was actually derived from the spontaneously formed R-S.

Collectively, our results from these *in vitro* assays demonstrate that Rif15 is a two-component transketolase that transforms a ketone (in R-S) into an ester (in R-L), a reaction that has not been found previously in natural product synthesis, to the best of our knowledge. Mechanistically, the deprotonation of ThDP at the thiazolium ring generates a carbanion, which is responsible for cleaving the C-2/C-3 bond in the 2-ketose. The resultant ThDP-bound dihydroxyethyl group then undertakes nucleophilic attack of the C-4 carbonyl carbon of R-S, which is followed by bond rearrangements and re-aromatization, ultimately yielding R-L (Fig. 3a).

**Biochemical, structural, and mechanistic characterization of Rif16.** Having characterized the R-SV→R-S→R-L transformation, we next sought to resolve the conversion of R-L into R-B. These



**Fig. 4** Structures of Rif16. **a** Substrate-free Rif16. The disordered B' loop (with 83-97 residues missing as shown with the dashed red line) that replaces the typical B' helix is shown in brown. The F and G helices are shown in magenta and orange, respectively. The disordered FG loop (192-205 residues) lacks electron density. The heme group is shown as a stick in red. **b** Substrate-bound Rif16. The B' loop in brown remains disordered. The FG loop (here ordered) is shown in blue. **c** Superimposition of the substrate-free (gray) and the substrate-bound (lime green) Rif16. The black arrows point out the regions that undergo conformational changes upon R-L binding. **d** The active site of Rif16. The substrate R-L and heme are shown as sticks in yellow and red, respectively, with the heme-iron depicted as a sphere. The key residues forming hydrogen bonds (black dashed lines) with R-L are shown as sticks in green. The residues within 5 Å around the substrate that constitute a large hydrophobic pocket are shown as sticks in silver. The conserved T260 and C366 are shown as sticks in silver. The distances (in angstroms) are indicated by the dashed yellow line

two rifamycin derivatives have the same oxidation state, but we still chose to test the activity of Rif16 against R-L, since this P450 enzyme was previously shown to be required for R-B biosynthesis<sup>17</sup>. Indeed, we found that R-L was significantly converted into R-B by Rif16 in the presence of *seFdx*, *seFdr*, and NADPH (Fig. 2b, trace xiii, xiv); the structure of the product was confirmed by the identical retention time of the product and the R-B authentic standard, co-elution with R-B in a co-injection experiment (Fig. 2b, trace xv), and the consistently observed *m/z* value of 754.3069 ( $[M-H]^+$ , *calc.*754.3069) (Supplementary Fig. 15). These results clearly establish that Rif16 is the long-sought R-B synthase of rifamycin biosynthesis.

To elucidate the catalytic mechanism for this atypical ester-to-ether transformation, the crystal structures of substrate-free Rif16 (PDB ID code: 5YSM, Fig. 4a) and R-L-bound Rif16 (PDB ID code: 5YSW, Fig. 4b) were solved at 1.90 Å and 2.60 Å resolution, respectively. In both of the structures, there was only one typical cytochrome P450 fold existing in an asymmetric unit. The BB' loop-B' helix-B'C loop region, which is known to be important for substrate specificity determination<sup>27</sup>, is significantly longer than those of many P450 enzymes that recognize smaller

substrates (Supplementary Fig. 16). The missing electron density of this region in both structures suggests the great structural flexibility. Both findings might help explain how Rif16 is able to accommodate its bulky substrate R-L, which represents one of the largest substrates for a P450 enzyme with the substrate-bound crystal structure available<sup>28</sup>. In the absence of substrate, Rif16 adopts an open conformation characterized by retraction of the F and G helices, loss of order in the B' helix, and missing electron density for the B'C and FG loops (Fig. 4a). A water molecule that is 2.5 Å away from the heme-iron forms the sixth axial ligand of  $Fe^{3+}$  (Supplementary Fig. 17a). Upon binding with R-L, Rif16's FG loop becomes ordered but the B' helix and the B'C loop remain disordered (Fig. 4b), thereby adopting a partially open conformation rather than the closed conformation observed in many substrate-bound P450 enzymes<sup>29,30</sup> (Fig. 4c).

In the substrate-bound structure (Fig. 4d and Supplementary Fig. 17b), R-L forms hydrogen bonds with residues S76, N108, F308, and T407, and additionally interacts with residues I107, A196, P197, I255, V303, P305, I307, F310, and L406 via hydrophobic interactions (all of these residues are within 5 Å of R-L). Critically, the axial water ligand is displaced and the

hydroxyl group at C-39 of R-L is closest (4.5 Å) to the heme-iron reactive center. These structural features suggest a possible mechanism for R-B production (Fig. 3a): the ferryl-oxo species (Compound I) likely abstracts the hydrogen atom of the C-39 hydroxyl group, leading to formation of a substrate radical and the ferryl-hydroxo Compound II. The resultant oxygen radical can then directly attack the neighboring arene to form a five-membered ring pendant, and the radical would be delocalized to the aromatic ring. Next, the relocation of the spirocyclic intermediate could induce the second hydrogen abstraction from the C-1 hydroxyl group by Compound II. This diradical mechanism might result in the formation of R-O. Notably, similar mechanisms—involving two alternative substrate binding poses being responsible for hydrogen abstractions from two distant sites—have been proposed for C–O coupling reactions catalyzed by a number of P450 enzymes<sup>31,32</sup>. Finally, the pentabasic cyclic compound R-O could be reduced to R-B (rather than R-L) by the NADPH-derived hydride, since the carboxylic acid is a better leaving group than the alcohol.

This R-L→R-B conversion reaction lacks net oxidation-reduction. To dissect this unusual P450 reaction experimentally, we elected to oxidize R-L by taking advantage of the peroxide shunt pathway of Rif16<sup>18</sup>, in which H<sub>2</sub>O<sub>2</sub> acts as the sole oxygen and electron donor of Rif16; this approach allowed us to bypass the dual roles of NADPH from our previous reaction system (its roles as an electron donor for the P450 enzyme and as a hydride provider for direct reduction of R-O). Interestingly, R-S was the dominant product from this reaction (Fig. 2b, trace xvi), which likely resulted from the spontaneous hydrolysis of the P450 product R-O<sup>15,33</sup> (Fig. 2b, trace xvii). The addition of NADPH into the Rif16/R-L/H<sub>2</sub>O<sub>2</sub> system led to predominant production of R-B (Fig. 2b, trace xviii, xix), as R-O can be reduced to R-B in the presence of NADPH (Fig. 2b, trace xx)<sup>34,35</sup>. Furthermore, the unstable compound R-O with the correct *m/z* value of 752.2920 ([M-H]<sup>-</sup>, *calc.* 752.2924) was observed in a time-course study (Supplementary Fig. 18). These results strongly suggest that R-O is the intermediate that enables the conversion of R-L to R-B.

To validate our proposed enzymatic reaction mechanisms, we performed a series of <sup>13</sup>C-tracer NMR experiments. First, [39-<sup>13</sup>C]R-L was prepared by mixing [1-<sup>13</sup>C]glucose, ATP, Mg<sup>2+</sup>, hexokinase, G-6-P isomerase, Rif15a/Rif15b, ThDP, and R-S in a one-pot reaction. We observed that [1-<sup>13</sup>C]glucose was phosphorylated to [1-<sup>13</sup>C]G-6-P by hexokinase, which was subsequently transformed into [1-<sup>13</sup>C]F-6-P by G-6-P isomerase (Fig. 3). The Rif15-mediated transfer of the <sup>13</sup>C-labeled glycolic acid C<sub>2</sub> moiety from [1-<sup>13</sup>C]F-6-P to R-S resulted in production of [39-<sup>13</sup>C]R-L, with an enriched C-39 signal of δ<sub>C</sub> 62.4 (Fig. 3b). The identity of this product was further confirmed by LC-HRMS analysis indicating an *m/z* value of 755.3106 ([M-H]<sup>-</sup>, *calc.* 755.3105, Supplementary Fig. 19), which is ~1 Da greater than that of unlabeled R-L [M-H]<sup>-</sup> = 754.3069). Next, Rif16, *seFdx/seFdr*, and NADPH were added into the above one-pot reaction. As expected, the P450 enzyme converted [39-<sup>13</sup>C]R-L into [38-<sup>13</sup>C]R-B, confirmed by LC-HRMS with an *m/z* value of 755.3100 ([M-H]<sup>-</sup>, *calc.* 755.3105, Supplementary Fig. 19) and by our observation that the <sup>13</sup>C-labeled carbon signal shifted downfield from δ<sub>C</sub> 62.4 to δ<sub>C</sub> 67.9 (Fig. 3b); both analytical results are consistent with the conversion of R-L to R-B via R-O (Fig. 3a).

It was previously reported that the R-SV high-producer *A. mediterranei* U32 has an R84W single mutation in Rif16<sup>16</sup>. The understanding of Rif16 mechanism allowed us to rationalize this industrially important phenotype. Specifically, the dissociation constant (*K<sub>d</sub>*) of R-L toward Rif16 was determined to be 1.3 ± 0.1 μM (Supplementary Fig. 20), while the purified Rif16<sub>R84W</sub> mutant (Supplementary Figs. 3 and 21) showed no detectable binding of R-L and lost the ability of catalyzing the

transformation from R-L to R-B (Supplementary Fig. 22). Since R84 is located at the B' loop of Rif16 (Supplementary Fig. 16), which is an important region for P450 substrate recognition<sup>27</sup>, its replacement by a tryptophan abolishes the productive substrate binding via a mechanism to be elucidated. Furthermore, according to the biosynthetic network shown in Fig. 1, the U32 mutant should accumulate R-L instead of the observed R-SV and R-S<sup>16</sup>. We reason that the ester R-L might be unstable, which could be hydrolyzed to R-SV either enzymatically or spontaneously (Supplementary Fig. 23).

Our elucidation of the network comprising the late steps of rifamycin biosynthesis revealed a unique C–O bond formation reaction mediated by a transketolase that involves both normal C–C bond formation and unusual bond rearrangements. Notably, transketolases primarily participate in central metabolic pathways such as pentose phosphate pathway and the Calvin cycle, and there have been few reports on transketolases that are involved in natural product biosynthesis<sup>36,37</sup>. The ether bond formation derived from the concomitant oxidation-reduction reactions and complex bond rearrangements also represents a highly atypical reaction for a P450 reaction system. The knowledge on the slow kinetics and the optimal C<sub>2</sub> keto donor of Rif15 could also help direct the future rational strain improvement. Finally, BLAST searches demonstrate that there exist other protein sequences with high similarity to Rif15 and Rif16 (Supplementary Table 3, Supplementary Fig. 24), suggesting that more Rif15-like and Rif16-like functionality could be further identified. Some of these enzymes come from rifamycin producing microorganisms<sup>38–41</sup>, which may suggest an effective method for discovery of more rifamycin producers by using Rif15 and Rif16 sequences as probes.

## Methods

**Chemicals.** Rifamycin SV and rifamycin O authentic standards were purchased from Sigma Aldrich (USA) and Toronto Research Chemicals (Canada), respectively. Rifamycin S and rifamycin B authentic standards were bought from National Institutes for Food and Drug Control (China).

**General DNA manipulation.** The *E. coli* DH5α strain was used for plasmid construction, storage, and isolation. Fast-digest restriction endonucleases (Thermo Fisher Scientific, USA) and T4 DNA ligase (Takara, Japan) were used for construction of vectors. PCR reactions were performed using I-5<sup>TM</sup> 2 × High-Fidelity Master Mix DNA polymerase (TsingKe Biotech, Beijing, China). Plasmid isolations from *E. coli* cells were performed using the Plasmid Miniprep Kit (TsingKe Biotech, Beijing, China). Purification of DNA fragments from agarose gels or PCR reactions was carried out using Gel Extraction Kit (Omega, USA) and Cycle Pure Kit (Omega, USA), respectively. Primers were synthesized by TsingKe (China).

**Molecular cloning.** The DNA sequences that encode the two-subunit transketolase Rif15, and the separated subunits Rif15a and Rif15b were amplified from the genomic DNA of *Amycolatopsis mediterranei* U32 (*A. mediterranei* U32 was deposited in Institute of Microbiology, Chinese Academy of Sciences designated as CGMCC4.5720) under standard PCR conditions using the primer pairs of *rif15-F/rif15-R*, *rif15a-pSj2F/rif15a-pSj2R*, and *rif15b-F/rif15b-R*, respectively (Supplementary Table 4). The P450 gene *rif16* and the mutant gene *rif16*<sub>R84W</sub> were amplified from *A. mediterranei* S699 and U32 (Professor Guoping Zhao's laboratory collection) gDNA, respectively, using a pair of primers *rif16-F* and *rif16-R* (Supplementary Table 4). The *rif15a* fragment was double digested by *Bam*HI/*Hind*III and cloned into the expression vector pSj2 (a derivative of pET21a, which is a gift from Professor Jiahai Zhou at Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences) using standard molecular cloning methods to generate pSj2-*rif15a* (Supplementary Fig. 2). The *rif15*, *rif15b*, *rif16*, and *rif16*<sub>R84W</sub> PCR products were double digested by *Nde*I/*Hind*III, *Nde*I/*Eco*RI, *Nde*I/*Xho*I, and *Nde*I/*Xho*I, respectively, and ligated into the corresponding pre-treated pET28b (Novagen, USA) to afford pET28b-*rif15*, pET28b-*rif15b*, pET28b-*rif16*, and pET28b-*rif16*<sub>R84W</sub> (Supplementary Fig. 2). All constructs were confirmed by DNA sequencing (Genewiz, China), and transformed into *E. coli* Codon Plus(DE3)-RIPL for expression of N-terminal His<sub>6</sub>-tagged recombinant proteins.

**Protein expression and purification.** The *E. coli* Codon Plus(DE3)-RIPL transformant carrying pSj2-*rif15a* was grown at 37 °C overnight in LB media supplemented with ampicillin (50 μg/mL), chloramphenicol (34 μg/mL), and

streptomycin (50 µg/mL). The transformant carrying pET28b-*rif15*, pET28b-*rif15b*, pET28b-*rif16* or pET28b-*rif16<sub>R84W</sub>* was cultured at 37 °C overnight in LB media supplemented with kanamycin (50 µg/mL), chloramphenicol (34 µg/mL), and streptomycin (50 µg/mL). The overnight seed culture was inoculated (1:100) into 1 L LB broth containing 10% glycerol, appropriate selective antibiotics and rare salt solution (6.75 mg/L FeCl<sub>3</sub>, 500 µg/L ZnCl<sub>2</sub>, CoCl<sub>2</sub>, Na<sub>2</sub>MoO<sub>4</sub>, 250 µg/L CaCl<sub>2</sub>, 465 µg/L CuSO<sub>4</sub>, and 125 µg/L H<sub>3</sub>BO<sub>3</sub>), and then cultured at 37 °C until OD<sub>600</sub> reached 0.6–1 (~3–4 h). Next, 0.1 mM isopropyl β-D-thiogalactoside (IPTG, Sigma, USA) was added to induce protein expression, and thiamin (1 mM, Sigma, USA) and δ-aminolevulinic acid (1 mM, Sigma, USA) were supplemented to support efficient expression of *holo*-form P450 enzymes. The cells were grown at 16 °C for 20 h. All recombinant proteins including Rif15, Rif15a, Rif15b, Rif16, and Rif16<sub>R84W</sub> were purified by Ni-NTA affinity chromatography<sup>42</sup>. Briefly, *E. coli* cells were collected by centrifugation at 5000 ×g for 5 min, the cell pellet was re-suspended with 20 mL lysis buffer (20 mM Tris-HCl, 300 mM NaCl, 10% (w/v) glycerol, and 10 mM imidazole, pH 8.0), and the cells were broken by sonication (5 s on/5 s off) for 30 min on ice. Next, the cellular debris was removed by centrifugation at 12,000 ×g for 30 min. To the supernatant 1 mL nickel-nitrilotriacetic acid resin (Qiagen, Germany) was added, and each mixture was incubated for 30 min with gentle shaking at 4 °C. The resin was loaded onto an empty column and washed with wash buffer (20 mM Tris-HCl, 300 mM NaCl, 10% (w/v) glycerol, and 20 mM imidazole, pH 8.0) until no protein could be detected in flow-through. The proteins bound to resin were eluted by 10 mL elution buffer (20 mM Tris-HCl, 300 mM NaCl, 10% (w/v) glycerol, and 250 mM imidazole, pH 8.0). Finally, the eluents were concentrated and buffer-exchanged to reaction buffer (20 mM Tris-HCl, 10% (w/v) glycerol, pH 7.4) via repeated ultrafiltration using Amicon Ultra-15 centrifugal filter units (Millipore, Ireland) with a 10-kDa cutoff. All protein purification steps were performed at 4 °C. The SDS-PAGE analysis (Supplementary Fig. 3) showed that Rif15, Rif15a, and Rif15b were >95% pure. However, the purity of Rif16 and Rif16<sub>R84W</sub> were unsatisfactory. Therefore, the collected Rif16 and Rif16<sub>R84W</sub> proteins were buffer-exchanged again to 20 mM Tris-HCl buffer (pH 8.0) by repeated ultrafiltration using Amicon Ultra-15 centrifugal filter units (Millipore, Ireland) with a 30-kDa cutoff. The buffer-exchanged proteins were further loaded onto a Mono Q 5/50 GL column (GE Healthcare, USA) pre-equilibrated with 20 mM Tris-HCl buffer (pH 8.0), and eluted with a gradient volume of 20 CV and an increasing ionic strength up to 1 M NaCl at a flow rate of 1 mL/min using ÄKTA purifier-100 (GE Healthcare, USA). Fractions containing proteins of interest were flash-frozen by liquid nitrogen and stored at –80 °C for later use.

**Protein concentration determination.** For Rif16 and Rif16<sub>R84W</sub>, the UV-visible spectra were recorded on a DU 800 spectrophotometer (Beckman Coulter, USA). The CO-bound reduced difference spectrum was employed to determine the functional concentration of P450 enzymes using the extinction coefficient ( $\epsilon_{450\text{nm}}^{\text{reduced}}$ ) of 91,000 M<sup>-1</sup>·cm<sup>-1</sup><sup>43</sup>. Briefly, CO was slowly bubbled through the Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> reduced P450 enzyme solution using a Pasteur pipette in a fume hood. The spectra of ferric, CO-bound, and CO-bound reduced forms of the P450 enzyme were recorded between 250 and 550 nm for generation of the CO-bound reduced difference spectrum. The protein concentrations of other proteins were determined using the Bradford assay with bovine serum albumin as standard<sup>44</sup>.

**Spectral substrate binding assays.** Spectral substrate binding assays were carried out on a UV-visible spectrophotometer 50 Bio (Cary, USA) at room temperature by titrating 100 µM rifamycin L DMSO solution (blank DMSO for the reference group) into 1 mL of 1 µM Rif16 or Rif16<sub>R84W</sub> solution in 1 µL aliquots, leading to the substrate concentrations ranging from 0.1 to 1.2 µM. The series of Type I difference spectra were used to deduce  $\Delta A$  ( $A_{\text{peak}(390\text{ nm})} - A_{\text{trough}(420\text{ nm})}$ ). Then, the  $\Delta A$  data versus substrate concentrations were fit to Michaelis–Menten equation to calculate the dissociation constant  $K_d$ <sup>45</sup>.

**In vitro enzymatic assays of Rif15 and Rif16.** The Rif15 reaction mixture contained 10 µM Rif15a, 10 µM Rif15b, 200 µM R-S (or other rifamycin derivatives, National Institutes for Food and Drug Control, China), 0.5 mM ThDP (Sigma, USA), 2.5 mM MgCl<sub>2</sub>, and 2 mM F-6-P (J&K Scientific, China) (or Xu-5-P, Ru-5-P, S-7-P, DHA from Sigma), in 100 µL of reaction buffer (20 mM Tris-HCl, 10% glycerol, pH 7.4). The Rif16 reaction assay was carried out in 100 µL of the same reaction buffer containing 2 µM Rif16, 200 µM R-L, 20 µM seFdx plus 10 µM seFdxR, and 1 mM NADPH (Solarbio, China) or 2 µM Rif16, 200 µM R-L, and 20 mM H<sub>2</sub>O<sub>2</sub>. The Rif15 and Rif16 reactions were incubated at 28 °C for 4 h and 1 h (unless otherwise specified), respectively, and quenched by mixing with the same volume of methanol. After high-speed centrifugation (20,000 ×g) for 15 min, the supernatants were analyzed on an Agilent 1260 infinity HPLC system (Agilent Technologies, USA) equipped with an ultraviolet detector. Compounds were separated by SB-C18 reverse-phase column (Thermo, 5 µm, 46 mm, USA) in a gradient system consisting of ddH<sub>2</sub>O + 0.1% trifluoroacetic acid as solvent A and acetonitrile as solvent B. The program of solvent gradient is as follows: 40% B for 5 min, 40–80% B over 15 min, and 80–100% B over 5 min at a flow rate of 1 mL/min. The wavelength of detection was 425 nm for all rifamycin derivatives except for R-O (370 nm). The peak identity in each HPLC trace was assigned by comparison with authentic standards regarding the retention time and the

UV-visible spectrum, and confirmed by co-injection experiments and LC-HRMS analysis.

**Kinetic analysis of Rif15.** The kinetic analysis of Rif15 was carried out using 2–10 µM Rif15a/Rif15b, 10–100 µM R-S as substrate, 2 mM F-6-P as the C<sub>2</sub> keto donor, 2.5 mM MgCl<sub>2</sub>, and 0.5 mM ThDP in 100 µL of reaction buffer. The reactions were performed 15–30 min at 28 °C and quenched by mixing with the same volume of methanol. After high-speed centrifugation (20,000 ×g) for 15 min, the supernatants were analyzed on an Agilent 1260 infinity HPLC system as described above. Each peak area of R-L was used to calculate the product concentration based on the standard curve of R-L. The triplicated data were fit to the Michaelis–Menten equation to determine the  $k_{\text{cat}}$  and  $K_m$  values using Origin 9.0.

**Preparation of <sup>13</sup>C labeled F-6-P, R-L, and R-B.** To prepare [1-<sup>13</sup>C]-F-6-P, the reaction mixture containing 5 mM [1-<sup>13</sup>C] glucose (Sigma, USA), 6.25 mM MgCl<sub>2</sub>, 12.5 mM ATP, 1 U hexokinase (Sigma, USA), 1 U phosphoglucose isomerase (Sigma, USA) was incubated at 28 °C for 16 h<sup>46,47</sup>. To prepare the <sup>13</sup>C-labeled R-L, 5 mM [1-<sup>13</sup>C]-glucose, 8.75 mM MgCl<sub>2</sub>, 12.5 mM ATP, 1 U hexokinase, 1 U phosphoglucose isomerase, 10 µM Rif15a, 10 µM Rif15b, 200 µM R-S, and 0.5 mM ThDP were mixed and incubated at 28 °C for 16 h. To prepare the <sup>13</sup>C-labeled rifamycin B, [39-<sup>13</sup>C] R-L was first prepared as described above; after 8 h, 2 µM Rif16, 20 µM seFdx, 10 µM seFdxR, and 1 mM NADPH were added into the reaction mixture, and the reaction continued for 16 h at 28 °C. Notably, the one-pot reaction was unsuccessful because NADPH would reduce R-S to R-SV, thus preventing the Rif15 catalyzed reaction.

**Isolation and purification of R-L and R-B.** The scaled-up Rif15 (for R-L) and Rif16 (for R-B) reactions (20–100 mL) were respectively extracted by the same volume of ethyl acetate five times. Then, the extracts were dried under nitrogen flow, and re-dissolved in methanol. The purification of R-L or R-B was performed using semi-preparative HPLC (Waters XBridge™ Prep C18 5 µm, 10 × 250 mm) with a linear gradient of 40–80% acetonitrile in ddH<sub>2</sub>O + 0.1% trifluoroacetic acid over 20 min, and then 100% acetonitrile for 5 min at a flow rate of 2.5 mL/min. The collected fractions containing R-L or R-B were combined. The solvents were removed using a Rotavapor R-3 rotary evaporator (Buchi, Switzerland) and N<sub>2</sub> blowing. Finally, 10.0 mg R-L and 3.0 mg R-B with >98.0% purity were obtained.

**LC-HRMS analysis.** The LC-HRMS analysis was performed on a Waters symmetry column (4.6 × 150 mm, RP18) using the negative-mode electrospray ionization with a linear gradient of 10–100% acetonitrile in ddH<sub>2</sub>O with 0.1% formic acid over 20 min, and followed by 100% acetonitrile for 5 min at a flow rate of 0.5 mL/min. The high resolution mass spectra were recorded on a Dionex Ultimate 3000 coupled to a Bruker Maxis Q-TOF.

**NMR analysis.** The <sup>13</sup>C NMR spectra of <sup>13</sup>C-labeled glucose, G-6-P, and F-6-P were acquired using D<sub>2</sub>O as solvent. The <sup>1</sup>H, <sup>13</sup>C and 2D NMR spectra of rifamycin derivatives were obtained using CDCl<sub>3</sub> or MeOD as solvent on a Bruker Avance III 600 MHz spectrometer with a 5 mm TCI cryoprobe.

**Crystallization and structure determination.** The crystal of Rif16 alone and the complex crystal of Rif16 and R-L were obtained at 16 °C by hanging drop vapor diffusion. The native Rif16 crystal screen droplets consisted of a 1:1 (v/v) protein at 16 mg/mL and the well solution of 100 mM Bis-Tris, pH 6.5, 200 mM magnesium chloride hexahydrate, and 25% PEG3350. Rif16 and R-L were mixed at a molar ratio of 1:5 and the co-crystallization was carried out in 200 mM magnesium chloride hexahydrate, 100 mM Tris, pH 8.5, 30% PEG4000. Crystals appeared after 1 week and were ready for data collection in 20 days. The crystals were flash-cooled in liquid nitrogen. The diffraction data were collected at 100 K under the synchrotron radiation at beamline BL19U1 of the Shanghai Synchrotron Radiation Facility (SSRF). The data sets were integrated and scaled with the HKL3000 package<sup>48</sup>. The structure is determined by molecular replacement with the structure of CYP105A (PDB accession code 4OQS) as the initial search model with the program Phaser. The programs Refmac5 and Coot9 were used for the refinement and model building<sup>49–52</sup>. Ramachandran plots were generated with Coot9. The statistics for data processing and structure refinement are shown in Supplementary Table 5. The coordinates were deposited to Protein Data Bank and the PDB ID codes are 5YSM and 5YSW for substrate-free Rif16 and R-L-bound Rif16, respectively. Figures were prepared using PYMOL (<http://www.pymol.org>).

**Data availability.** Data that support the findings of this study have been deposited in Protein Data Bank with the PDB ID codes 5YSM and 5YSW. Data referenced in this study are available in GenBank with the accession codes AMED\_0651, AMED\_0652, and AMED\_0653. All other relevant data are available from the corresponding authors.

Received: 3 April 2018 Accepted: 12 May 2018

Published online: 14 June 2018

## References

- Riva, S. & Silvestri, L. G. Rifamycins: a general view. *Ann. Rev. Microbiol.* **26**, 199–224 (1972).
- Floss, H. G. & Yu, T. W. Rifamycin-mode of action, resistance, and biosynthesis. *Chem. Rev.* **105**, 621–632 (2005).
- Calvori, C., Frontali, L., Leoni, L. & Tecce, G. Effect of rifamycin on protein synthesis. *Nature* **207**, 417–418 (1965).
- August, P. R. et al. Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. *Chem. Biol.* **5**, 69–79 (1998).
- Schupp, T., Toupet, C., Engel, N. & Goff, S. Cloning and sequence analysis of the putative rifamycin polyketide synthase gene cluster from *Amycolatopsis mediterranei*. *FEMS Microbiol. Lett.* **159**, 201–207 (1998).
- Tang, L., Yoon, Y. J., Choi, C. Y. & Hutchinson, C. R. Characterization of the enzymatic domains in the modular polyketide synthase involved in rifamycin B biosynthesis by *Amycolatopsis mediterranei*. *Gene* **216**, 255–265 (1998).
- White, R. J., Martinelli, E. & Lancini, G. Ansamycin biogenesis: studies on a novel rifamycin isolated from a mutant strain of *Nocardia mediterranei*. *Proc. Natl. Acad. Sci. USA* **71**, 3260–3264 (1974).
- Floss, H. G. From ergot to ansamycins-45 years in biosynthesis. *J. Nat. Prod.* **69**, 158–169 (2006).
- Xu, J., Wan, E., Kim, C. J., Floss, H. G. & Mahmud, T. Identification of tailoring genes involved in the modification of the polyketide backbone of rifamycin B by *Amycolatopsis mediterranei* S699. *Microbiology* **151**, 2515–2528 (2005).
- Xiong, Y., Wu, X. & Mahmud, T. A homologue of the *Mycobacterium tuberculosis* PapA5 Protein, Rif-Orf20, is an acetyltransferase involved in the biosynthesis of antitubercular drug rifamycin B by *Amycolatopsis mediterranei* S699. *Chembiochem* **6**, 834–837 (2005).
- Oppolzer, W., Prelog, V. & Sensi, P. The composition of rifamycin B and related rifamycins. *Experientia* **20**, 336–339 (1964).
- Funayama, S. & Cordell, G. A. in *Studies in Natural Products Chemistry* Vol. 23 (ed. Rahman Atta, u. r.) 51–106 (Elsevier, 2000).
- Lancini, G. C., Gallo, G. G., Sartori, G. & Sensi, P. Isolation and structure of rifamycin L and its biogenetic relationship with other rifamycins. *J. Antibiot. (Tokyo)* **22**, 369–377 (1969).
- Sugawara, S., Karasawa, K., Watanabe, M. & Hidaka, T. Production of rifamycin O by *Streptomyces* 4107 A2. *J. Antibiot.* **17**, 29–32 (1964).
- Ghisalba, O., Roos, R., Schupp, T. & Nuesch, J. Transformation of rifamycin S into rifamycins B and L. A revision of the current biosynthetic hypothesis. *J. Antibiot.* **35**, 74–80 (1982).
- Zhao, W. et al. Complete genome sequence of the rifamycin SV-producing *Amycolatopsis mediterranei* U32 revealed its genetic characteristics in phylogeny and metabolism. *Cell Res.* **20**, 1096–1108 (2010).
- Yuan, H. et al. Two genes, *rif15* and *rif16*, of the rifamycin biosynthetic gene cluster in *Amycolatopsis mediterranei* likely encode a transketolase and a P450 monooxygenase, respectively, both essential for the conversion of rifamycin SV into B. *Acta Biochim. Biophys. Sin.* **43**, 948–956 (2011).
- Zhang, X. & Li, S. Expansion of chemical space for natural products by uncommon P450 reactions. *Nat. Prod. Rep.* **34**, 1061–1089 (2017).
- Guengerich, F. P. Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chem. Res. Toxicol.* **14**, 611–650 (2001).
- Kochetov, G. A. & Solovjeva, O. N. Structure and functioning mechanism of transketolase. *Biochim. Biophys. Acta* **1844**, 1608–1618 (2014).
- Nelson, D. R. The cytochrome P450 homepage. *Hum. Genomics* **4**, 59–65 (2009).
- Ma, L. et al. Reconstitution of the *in vitro* activity of the cyclosporine-specific P450 hydroxylase from *Sebekia benihana* and development of a heterologous whole-cell biotransformation system. *Appl. Environ. Microbiol.* **81**, 6268–6275 (2015).
- Scrutton, M. C. Divalent metal ion catalysis of the oxidation of rifamycin SV to rifamycin S. *FEBS Lett.* **78**, 216–220 (1977).
- Lindqvist, Y., Schneider, G., Ermler, U. & Sundström, M. Three-dimensional structure of transketolase, a thiamine diphosphate dependent enzyme, at 2.5 Å resolution. *EMBO J.* **11**, 2373–2379 (1992).
- Kochetov, G. A. Structure of the active center of transketolase\*. *Ann. N. Y. Acad. Sci.* **378**, 306–311 (1982).
- Dampier, M. F., Chen, C. W. & Whitlock, H. W. Jr. Substituent effects on the solution conformation of rifamycin S. *J. Am. Chem. Soc.* **98**, 7064–7069 (1976).
- Gotoh, O. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J. Biol. Chem.* **267**, 83–90 (1992).
- Rudolf, J. D., Chang, C. Y., Ma, M. & Shen, B. Cytochromes P450 for natural product biosynthesis in *Streptomyces*: sequence, structure, and function. *Nat. Prod. Rep.* **34**, 1141–1172 (2017).
- Poulos, T. L., Finzel, B. C. & Howard, A. J. High-resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.* **195**, 687–700 (1987).
- Sherman, D. H. et al. The structural basis for substrate anchoring, active site selectivity, and product formation by P450 PikC from *Streptomyces venezuelae*. *J. Biol. Chem.* **281**, 26289–26297 (2006).
- Tang, M. C., Zou, Y., Watanabe, K., Walsh, C. T. & Tang, Y. Oxidative cyclization in natural product biosynthesis. *Chem. Rev.* **117**, 5226–5333 (2017).
- Grandner, J. M., Cacho, R. A., Tang, Y. & Houk, K. N. Mechanism of the P450-catalyzed oxidative cyclization in the biosynthesis of griseofulvin. *ACS Catal.* **6**, 4506–4511 (2016).
- Seong, B. L., Son, H. J., Mheen, T. I. & Han, M. H. Microbial transformation of rifamycin B: a new synthetic approach to rifamycin derivatives. *J. Antibiot.* **36**, 1402–1404 (1983).
- Han, M. H., Seong, B. L., Son, H. J. & Mheen, T. I. Rifamycin B oxidase from *Monocillium* spp., a new type of diphenol oxidase. *FEBS Lett.* **151**, 36–40 (1983).
- Liu, Z. et al. Reduction of quinones by NADH catalyzed by organoiridium complexes. *Angew. Chem. Int. Ed.* **52**, 4194–4197 (2013).
- Peng, C. et al. Hijacking a hydroxyethyl unit from a central metabolic ketose into a nonribosomal peptide assembly line. *Proc. Natl. Acad. Sci. USA* **109**, 8540–8545 (2012).
- Craig, J. W. & Brady, S. F. Discovery of a metagenome-derived enzyme that produces branched-chain acyl-(acyl-carrier-protein)s from branched-chain alpha-keto acids. *Chembiochem* **12**, 1849–1853 (2011).
- Wilson, M. C., Gulder, T. A. M., Mahmud, T. & Moore, B. S. Shared biosynthesis of the saliniketals and rifamycins in *Salinispora arenicola* is controlled by the *sare1259*-encoded cytochrome P450. *J. Am. Chem. Soc.* **132**, 12757–12765 (2010).
- Saxena, A., Kumari, R., Mukherjee, U., Singh, P. & Lal, R. Draft genome sequence of the rifamycin producer *Amycolatopsis rifamycinica* DSM 46095. *Genome Announc.* **2**, e00662-14 (2014).
- Huang, H. et al. *Micromonospora rifamycinica* sp. nov., a novel actinomycete from mangrove sediment. *Int. J. Syst. Evol. Microbiol.* **58**, 17–20 (2008).
- Promnuan, Y., Kudo, T., Ohkuma, M. & Chantawannakul, P. *Actinomadura apis* sp. nov., isolated from a honey bee (*Apis mellifera*) hive, and the reclassification of *Actinomadura cremea* subsp. *rifamycinii* Gauze et al. 1987 as *Actinomadura rifamycinii* (Gauze et al. 1987) sp. nov., comb. nov. *Int. J. Syst. Evol. Microbiol.* **61**, 2271–2277 (2011).
- Du, L. et al. Characterization of a unique pathway for 4-cresol catabolism initiated by phosphorylation in *Corynebacterium glutamicum*. *J. Biol. Chem.* **291**, 6583–6594 (2016).
- Omura, T. & Sato, R. The carbon monoxide-binding pigment of liver microsomes II. Solubilization, purification, and properties. *J. Biol. Chem.* **239**, 2379–2385 (1964).
- Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).
- Li, S., Podust, L. M. & Sherman, D. H. Engineering and analysis of a self-sufficient biosynthetic cytochrome P450 PikC fused to the RhFRED reductase domain. *J. Am. Chem. Soc.* **129**, 12940–12941 (2007).
- Tsuboi, K. K., Estrada, J. & Hudson, P. B. Enzymes of the human erythrocyte. IV. Phosphoglucose isomerase, purification and properties. *J. Biol. Chem.* **231**, 19–29 (1958).
- Hauck, T., Landmann, C., Bruhlmann, F. & Schwab, W. Formation of 5-methyl-4-hydroxy-3[2H]-furanone in cytosolic extracts obtained from *Zygosaccharomyces rouxii*. *J. Agric. Food Chem.* **51**, 1410–1414 (2003).
- Otwinowski, Z. & Minor, W. *Methods in Enzymology*. 276, 307–326 (Academic Press, 1997).
- McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D.* **53**, 240–255 (1997).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of *Coot*. *Acta Crystallogr. D.* **66**, 486–501 (2010).
- Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D.* **50**, 760–763 (1994).

## Acknowledgements

This work was supported by the Shandong Provincial Natural Science Foundation (ZR2017ZB0207 to W.Z. and S.L.), the National Natural Science Foundation of China (grants 81741155, 31422002, and 21472204 to S.L., 31600036 to F.Q., 21572243 to Y.X., 31430004 and 31670058 to G.Z.), Chinese Academy of Sciences (grants QYZDB-SSW-SMC042 to S.L. and XDPB0402 to Y.X.), and the Science and Technology Commission of

Shanghai Municipality (grant 15JC1400402 to Y.X.). We thank the staff of BL19U1 beamline at the Shanghai Synchrotron Radiation Facility for assistance during X-ray diffraction data collection. We are also grateful to Dr. Gong Chen at Nankai University for helpful discussions about reaction mechanisms.

### Author contributions

F.Q., C.L., J.W., G.-L.T., Y.X., G.Z., and S.L. designed research; F.Q., C.L., F.L., X.Z., J.W., W.Z., Z.F., and W.L. performed research; F.Q., C.L., F.L., X.Z., J.W., G.-L.T., Y.X., G.Z., and S.L. analyzed results; F.Q., C.L., F.L., X.Z., Y.X., G.Z., and S.L. wrote the manuscript; All authors read and approved the manuscript.

### Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-018-04772-x>.

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

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**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