Rifamycin congeners kanglemycins are active against rifampicin-resistant bacteria via a distinct mechanism

James Peek1, Mirjana Lilic2, Daniel Montiel1, Aleksandr Milshtein1, Ian Woodworth1, John B. Biggins1, Melinda A. Ternei1, Paula Y. Calle1, Michael Danziger1, Thulasi Warrier3, Kohta Saito3,4, Nathaniel Braffman2, Allison Fay5, Michael S. Glickman5, Seth A. Darst2, Elizabeth A. Campbell2 & Sean F. Brady1

Rifamycin antibiotics (Rifs) target bacterial RNA polymerases (RNAPs) and are widely used to treat infections including tuberculosis. The utility of these compounds is threatened by the increasing incidence of resistance (RifR). As resistance mechanisms found in clinical settings may also occur in natural environments, here we postulated that bacteria could have evolved to produce rifamycin congeners active against clinically relevant resistance phenotypes. We survey soil metagenomes and identify a tailoring enzyme-rich family of gene clusters encoding biosynthesis of rifamycin congeners (kanglemycins, Kangs) with potent in vivo and in vitro activity against the most common clinically relevant RifR mutations. Our structural and mechanistic analyses reveal the basis for Kang inhibition of RifR RNAP. Unlike Rifs, Kangs function through a mechanism that includes interfering with 5′-initiating substrate binding. Our results suggest that examining soil microbiomes for new analogues of clinically used antibiotics may uncover metabolites capable of circumventing clinically important resistance mechanisms.
semisynthetic derivatives of the bacterial natural product rifamycin (e.g., rifampicin or Rif) are components in the first-line treatment of tuberculosis and other gram-negative bacterial infections. As with many antibiotics, the clinical utility of these therapeutics has declined due to the increased incidence of antibiotic resistant bacterial pathogens. Resistance to the rifamycin family of antibiotics commonly occurs in clinical isolates as a result of point mutations in the antibiotic’s target, the DNA-dependent RNA polymerase (RNAP). These mutations, as well as other clinically relevant antibiotic resistance mechanisms, are also likely to be present in natural environments where they would have evolved in response to antibiotics produced by other bacteria. The search for biologically active bacterial natural products has frequently led to the discovery of families of structurally related antibiotics (congeners) that arise from evolutionarily related biosynthetic gene clusters. While these close analogues typically have the same molecular target, they often exhibit different biological activities, including differences in potency, spectrum of activity and activity against resistant bacteria. In this study, we postulated that competition between environmental microbes might have selected for the evolution of rifamycin congeners capable of circumventing common antibiotic resistance mechanisms, including those enriched in clinical settings, providing a source of new therapeutics to treat rifamycin resistant bacteria.

In an effort to understand natural rifamycin biosynthetic diversity we turned to the sequencing of soil metagenomes. Soils are believed to be a rich and underexplored reservoir of bacterial biosynthetic diversity, with each gram of soil containing thousands of previously unstudied bacterial species. The development of robust sequencing approaches for identifying biosynthetic gene clusters in complex microbiomes has made it possible to systematically explore soil ecosystems for gene cluster families of interest. We hypothesized that the most biosynthetically complex rifamycin-like gene clusters found in soil environments could represent nature’s most evolved responses to commonly encountered rifamycin resistance mechanisms. Our survey of soil metagenomes revealed a rich diversity in rifamycin biosynthesis. One family of gene clusters, which contained the largest collection of predicted tailoring genes, was of particular interest to us, as we expected it might encode for the most highly functionalized rifamycin congeners. We identified numerous examples of this gene cluster family in soil metagenomes as well as one example in the sequenced genome of a cultured bacterium. Hence, we report on the characterization of kanglemycin-like rifamycin congeners, kanglemycin (Kang) A, V1, and V2, that are encoded by a member of this tailoring enzyme-rich gene cluster family.

All three Kangs were more potent than Rif when assayed against bacteria carrying RNAP mutations corresponding to those commonly identified in Rif resistant (RifR) clinical isolates of Mycobacterium tuberculosis (Mtb), the causative agent of tuberculosis. Interestingly, Kang V1 and V2 exhibit their highest levels of activity against bacteria carrying different common RifR mutations. To study the inhibition of RNAP by Kang-like congeners, we solved X-ray co-crystal structures of Kang A in complex with both wild-type Mycobacterium smegmatis (Msm) RNAP and a RifR RNAP variant carrying the most commonly observed RifR mutation found in Mtb clinical isolates (Msm RNAP (S447L)). Our structural analysis revealed that Kang A binds the same site on RNAP as Rif. Additional interactions between the chemical groups unique to the Kangs (compared to Rif) and RNAP help to explain the ability of these compounds to inhibit RifR RNAP. The structural results combined with structure-guided biochemical studies indicate that the Kangs inhibit RNAP activity at a step of transcription upstream of the step inhibited by Rif. Taken together, these data suggest that the Kangs V1 and V2 represent potential lead structures for the development of therapeutics with activity against RifR bacteria.

### Results

#### Metagenomic survey of rifamycin biosynthetic diversity

The complexity of soil microbiomes limits the utility of shotgun sequencing as a tool for identifying biosynthetic gene clusters in soil metagenomes. Instead, PCR-based methods that use degenerate primers to target conserved natural product biosynthetic genes have been developed to study the biosynthetic gene cluster diversity present in an environmental sample, much in the same way that bacterial phylogenetic diversity is routinely evaluated through the analysis of PCR-amplified 16S genes. To assess the diversity of rifamycin-like gene clusters present in soil microbiomes, we used degenerate primers targeting the 3-amino-5-hydroxy benzoic acid (AHBA) synthase gene, which encodes the final step in AHBA biosynthesis (Fig. 1). AHBA is the universal precursor for the ansamycin family of natural products, including the rifamycins. The phylogenetic divergence of AHBA synthase genes correlates closely with the structural divergence of the metabolites encoded by the biosynthetic gene clusters from which an AHBA synthase gene arises, making it an information-rich target for identifying rifamycin-like gene clusters in metagenomes using PCR-based methods.

To identify metagenomes containing rifamycin-like biosynthetic gene clusters, environmental DNA (eDNA) isolated from a collection of approximately 1500 geographically and ecologically diverse soils was used as the template in PCR reactions with degenerate primers designed to amplify AHBA synthase genes. From each soil, the AHBA synthase amplicons were sequenced, and the sequences were used to determine the number of distinct AHBA synthase sequences present in each soil.

#### Rifamycin-like gene clusters from metagenomic libraries

In sequenced biosynthetic gene clusters that encode rifamycin family members, a variable region containing tailoring genes, responsible for generating most of the structural diversity seen in rifamycin congeners, resides directly downstream to the AHBA biosynthesis operon. To guide the isolation of eDNA cosmids containing tailoring genes, the seven newly constructed and two pre-existing soil eDNA libraries were screened with the same AHBA synthase degenerate primers that we used to screen crude eDNA extracts. We initially recovered 35 unique cosmids (i.e., primary clones) from sublibrary pools that yielded AHBA synthase sequences (Fig. 2a). Sequencing of these cosmids revealed that variations in the collections of predicted tailoring genes largely changed in concert with the phylogenetic divergence of the AHBA synthase genes.
Representative gene clusters associated with each major AHBA synthase clade were recovered in their entirety on sets of overlapping cosmid clones (Fig. 2a). Each collection of overlapping cosmids was sequenced, assembled into a single continuous stretch of DNA and annotated in silico to reveal an eDNA-derived rifamycin-like gene cluster (Fig. 2b). eDNA-derived gene clusters were predicted to encode a number of enzymes that have not previously been associated with rifamycin congener biosynthesis (e.g., N-acyltransferases, CoA-transferases, propionyl-CoA carboxylases, methylmalonyl-CoA mutases, and lanthionine synthetase-like enzymes) (Fig. 2b and Supplementary Figure 1A). A number of other tailoring genes found in these clusters are phylogenetically distinct from those found in known rifamycin-like gene clusters, suggesting they may differentially functionalize the rifamycin backbone. These genes are predicted to encode glycosyltransferases, methyltransferases, cytochrome P450s, oxidoreductases, and sugar biosynthesis enzymes (Supplementary Figures 2 and 3).

In most cases, the polyketide synthase (PKS) portion of each gene cluster is predicted to be functionally identical. However, a number of gene clusters with the most complex sets of tailoring genes were predicted to encode a change in the substrate specificity of the acyltransferase (AT) domain in the eighth PKS module (AT8*, Fig. 2b). These AT8* domains are predicted to use ethylmalonyl-CoA (Emal) as a substrate instead of methylmalonyl-CoA (Mmal)20,21, which would introduce a two-carbon branch into the rifamycin polyketide (PK) core (Supplementary Figure 1B). The combination of a potential change in the PK core structure and a complex collection of tailoring genes led us to prioritize this family of gene clusters for investigation. We hypothesized that these gene clusters would encode the most complex rifamycin congeners to have evolved to date and that this increased complexity might have evolved in response to common rifamycin resistance mechanisms. Based on AHBA synthase phylogeny, 13% of the rifamycin-like AHBA synthase sequences we amplified from soil environments are predicted to arise from this family of gene clusters (Fig. 2a, orange colored clades). While our screening suggests this is a common class of gene clusters in the environment, a search of all publicly available sequenced bacterial genomes only revealed one gene cluster that contains a similarly complex tailoring gene region and an AT8* domain. This previously uncharacterized gene cluster from Amycolatopsis vancoresmycina (Av) is identical in gene content and organization to the RifCon 10 gene cluster that we recovered from a soil eDNA library (Fig. 3a).

Highly functionalized congeners from an AT8* gene cluster. As an initial exploration of the tailoring gene-rich family of gene clusters that contain an AT8* domain, we looked for rifamycin congeners in ethyl acetate extracts from cultures of Av. While Av has never been reported to produce rifamycin-like UV spectra in culture broth extracts (Supplementary Figure 4). The structure of each metabolite was elucidated using a combination of high-resolution mass spectrometry (HRMS), 1D and 2D NMR and UV data. 13C-NMR, HRESIMS [calcd m/z for

**Fig. 1** The rifamycin biosynthetic gene cluster and the role of AHBA synthase. a The rifamycin gene cluster from Amycolatopsis mediterranei. b The reaction catalyzed by AHBA synthase and the structure of rifamycin SV (the product of the rifamycin gene cluster). The rifamycin SV structure is colored according to the genes responsible for producing its PK core (red), AHBA-derived substructure (green), and tailoring functionalities (black). The phylogenetic divergence of AHBA synthase genes from previously characterized gene clusters correlates with the different structural classes of ansamycins.
C_{50}H_{64}NO_{19} (M + H\textsuperscript{+}) 982.4073, found m/z 982.4025, and UV data for compound 1 were consistent with the structure of Kang A, a rifamycin congener originally characterized from *Amycolatopsis mediterranei* var. *kanglensis* and encoded by an uncharacterized gene cluster (Supplementary Figures 5–10, Supplementary Table 2)\(^{22,23}\). The most dramatic differences between Kang A and other rifamycins are that it contains a β-O-3,4-O,O’-methylene digitoxose deoxysugar (hereafter, K-sugar) substituent at C-27, an oxidized ethyl substituent in place of a methyl substituent at C-20, and a gem-dimethylsuccinic acid (K-acid) appended to the oxidized ethyl branch in the PK core (Fig. 3b).

The predicted molecular formula for the second metabolite, Kang V1 (2) [HRESIMS calcd m/z for C_{50}H_{65}NO_{19}Na (M + Na\textsuperscript{+}) 1006.4048, found m/z 1006.4006], suggested it was a reduced analogue of Kang A. A comparison of 1 and 2D NMR data from (1) and (2) allowed us to assign this difference to the reduction of the C-11 ketone to an alcohol [\(13C \delta 77.1, 1H \delta 5.49\) (1 H, s)] (Supplementary Figures 11–16, Supplementary Table 2). To the best of our knowledge, this C-11 reduction has only been seen in
one previously described rifamycin natural product congener, chaxamycin D<sup>24</sup>. In the case of the third metabolite, Kang V2 (3), HRMS data suggested it differed from Kang A (1) by the addition of a CH<sub>3</sub>O moiety [HRESIMS calcd m/z for C<sub>31</sub>H<sub>26</sub>NO<sub>19</sub> (M + H<sup>+</sup>) 996.4229, found m/z 996.4197]. The UV spectra of Kang V2 (3) supported the presence of a naphthohydroquinone moiety (λ<sub>max</sub> 302 nm) instead of the naphthoquinone (λ<sub>max</sub> 276) seen in (1) and (2) (Supplementary Figure 4). The naphthohydroquinone substructure was also supported by an HMBC correlation from H-3 to a carbon at δ 150.5 ppm (C-4; Supplementary Figures 17–22, Supplementary Table 2). In Kang A and V1, this carbon is significantly more deshielded (δ 184.9 and 188.8 ppm, respectively). The presence of the carbonyl at C-8 in Kang V2 was supported by an HMBC correlation from the C-14 methyl to C-8 (δ 191.7). The formation of a fourth ring on the naphthohydroquinone substructure through the addition of a highly deshielded methylene [13C δ 98.4, 1H δ 6.19 (1 H, d), δ 5.48 (1 H, d)] was defined by HMBC correlations from the new methylene protons to C-4 and C-11 of the naphthohydroquinone (Supplementary Figure 17). To the best of our knowledge, the fourth ring formed by the addition of the methylenedioxy bridge in Kang V2 (3) is not found in any reported rifamycin congeners.

Many of the new structural features found on the Kangs can be rationalized based on differences in gene content between the Kang (kng) gene cluster and other rifamycin family gene clusters (Fig. 3a, c, Supplementary Figures 23 and 24, Supplementary Table 3). In addition to the AT8*-containing kngD domain, the kng cluster contains a collection of deoxysugar biosynthesis genes (kng22, kng23, and kng27) and a glycosyltransferase gene (kng26) that we predict are involved in generating the K-sugar (Fig. 3a, Supplementary Figure 17). To the best of our knowledge, the fourth ring formed by the addition of the methylenedioxy bridge in Kang V2 (3) is not found in any reported rifamycin congeners.

In vitro transcription assay with radiolabeled nucleotides showing the activity of Rif and the Kangs at the concentrations indicated against wild-type and RifR<sup>Msm</sup> S447L RNAP. F, full-length transcript; A, abortive transcript.
Kangs are active against RifR RNAPs via a distinct mechanism. Kangs A, V1 and V2 are active as antibiotics against Gram-positive bacteria, including Staphylococcus aureus (Sau), Staphylococcus epidermidis, Listeria monocytogenes, and Mtb (Supplementary Table 4). Kangs V1 and V2 both show improved activity against Mtb (H37Rv; IC₉₀ 3.12 and 1.56 µM, respectively) compared to Kang A (12.5 µM). We were particularly interested in whether the complex structural features seen in the Kangs might impart improved activity against mutations in RNAP that confer RifR. Substitutions at just three RNApo amino acid positions, Mtb RNAP β subunit D441, H451, and S456 (corresponding to Msm/ E. coli [Eco] RNAP β subunit D432/D516, H442/H526, and S447/S531) account for the vast majority of mutations observed in RifR Mtb clinical isolates. The antibacterial activity of the Kangs against RifR RNAP mutants was assessed in vivo using a collection of Sau strains carrying RNAP point mutations and in vitro using purified wild-type and RifR (S447L) Msm RNAPs. The use of these models allowed us to explore the activity of the Kangs against mutations that correspond to the most commonly mutated sites in RifR Mtb, without necésitising the use of restrictive BSL3 assay conditions.

The Kangs are active against RifR Sau strains carrying RNAP mutations at sites corresponding to those commonly mutated in RifR Mtb clinical isolates (Fig. 3d). Kang V1 showed an ~80-fold lower MIC (0.069 µg mL⁻¹) than Rif (5.6 µg mL⁻¹) against a Sau RNAP βD471Y mutant strain. Kang V2 exhibited similarly potent activity (MIC 0.069 µg mL⁻¹) against a Sau strain carrying an RNAP βS486L mutation, which corresponds to the most commonly observed RifR mutation in Mtb clinical isolates (Mtb RNAP βS465L), appearing in ~40–80% of sequenced isolates from geographically diverse regions of the world. As with Mtb, the Sau RNAP βS468L mutation effectively abrogates antibacterial activity of Rif (MIC > 50 µg mL⁻¹). Remarkably, Kang V2 showed more potent activity against the Sau RNAP βS486L mutant than against the wild-type strain, suggesting it might have evolved in a niche where this variant is the dominant form of RNAP. Based on the results of our MIC assay, we predicted that treatment of wild-type cells with Kangs V1 and V2 could effectively suppress the development of two common RifR phenotypes. Indeed, in Sau we were not able to identify any Kang V2 resistant mutants that carried the βS486L mutation (Supplementary Figure 25) nor could we identify any βD471Y mutants that arose when cultures were treated with Kang V1. Each of these mutations occurred at a frequency of approximately 10% among RifR Sau colonies. Consistent with the results of our MIC assay, an H481Y mutation, which confers a high level of resistance to all of the compounds, was the predominant mutation that arose following exposure of Sau to either Rif or the Kangs. While mutations at H481 were the most common variants we sequenced in RifR mutant strains (~70%), the βS456L mutation (Sau βS486L) predominates in RifR Mtb clinical isolates.

To determine whether the activity of the Kangs against the Sau RNAP βS486L mutant could be generalized to mycobacterial RNAP carrying the equivalent mutation, we tested the in vitro activity of the Kangs against purified Msm RNAP using a run-off transcription assay (Fig. 3e). The Msm RNAP exhibits 91% sequence identity with Mtb RNAP at the amino acid level and shows complete conservation of residues in the Rif binding pocket. We found that the Kangs were all potent in vitro inhibitors of wild-type Msm RNAP, with comparable activity to Rif. While Rif was inactive against an Msm RNAP βS447L mutant (corresponding to Mtb/Sau RNAP β S447L/S465L), all three Kangs displayed potent activity against this mutant. In agreement with the results of our Sau MIC assays, Kang V2 showed the highest potency against the RifR Msm RNAP (Fig. 3e).

Kangs exhibit distinct mechanistic properties. Detailed analysis of the transcription assays suggested that the mechanism by which the Kangs inhibit RNAP differs from that of Rif. The effects of Rif on RNAP transcription activity at each stage of the transcription cycle have been probed extensively. Rif has little to no effect on promoter binding or open complex formation, but causes an increase in the apparent Kₘ for the initiating substrate NTPs binding in the enzyme i and i + 1 sites, thus affecting dinucleotide synthesis at lower NTP concentrations. Importantly, Rif does not affect RNAP catalysis itself (phosphodiester bond formation), the predominant effect of Rif is steric occlusion of the translocating nascent transcript after the formation of the first phosphodiester bond, resulting in the inhibition of the production of full-length transcript (F, Fig. 3e) but over-production of abortive dinucleotide transcripts (A, Fig. 3e). In contrast to the effect of Rif, the Kangs inhibited production of the full-length transcripts but also the abortive transcripts, suggesting that the Kangs inhibit a step of transcription preceding that of Rif—either substrate (DNA or initiating nucleotide) binding or phosphodiester bond catalysis itself.

Kang A and Rif share core interactions with RNAP. While Kang V1 and V2 showed the highest levels of activity against bacteria carrying specific clinically important RifR mutations, all three Kangs exhibit improved activity compared to Rif against RNAP variants carrying common mutations found in RifR Mtb clinical isolates. We speculated that the activity of the Kangs against RifR mutants and their potentially novel mechanism of inhibition could be related to the presence of the unique K-sugar and K-acid, which all three Kangs share. To explore this hypothesis, we examined a crystal structure of a mycobacterial RNAP complexed with Kang A, the parent compound in the Kang family, and compared it to a structure complexed with Rif. A more detailed examination of the interaction between each Kang congener and the specific RNAP mutant against which it is most potent will be the focus of future studies.

Kang A and Rif were soaked into crystals of an Msm RNAP transcription initiation complex (TIC). Both structures were phased by molecular replacement using the Msm TIC as a model and refined to 3.05 Å resolution (Fig. 4a–c, Supplementary Table 5). The structures of both antibiotics, including the K-sugar and K-acid moieties unique to the Kangs, as well as the RNAP β subunit interaction determinants for the antibiotics, were well-resolved (Supplementary Figure 26A and B). The tip of the σ-finger (a structural element of the σ subunit) also approaches each antibiotic and appears to make molecular contacts. However, because the σ-finger electron density is very weak (reflected in high atomic B-factors) and amino acid substitutions in σ that confer RifR have never been reported, the role of these interactions with Rif and Kang A remain to be established. We note that previous studies deleting the σ-finger suggested a role for this motif in binding to the Rif variant rifabutin but not to another variant, rifapentine, indicating that the significance of σ-finger/antibiotic interactions is dependent on the specific Rif variant.

The Rif/Msm RNAP interactions were similar to those described in previous structures (Fig. 4b). The Rif/Msm RNAP structure reveals a set of cation–π interactions that have not been noted previously. The conjugated double-bond system comprising C16–C19 of the PK backbone of Rif is approached from the RNAP side by the guanidino group of R445 in a geometry indicative of a cation–π interaction. The opposite face of the conjugated double-bond system is approached by the guanidino group of R604. We call this arrangement a cation–π sandwich (Supplementary Figure 26C). As expected, the Rif scaffold of Kang

ARTICLE NATURE COMMUNICATIONS | DOI: 10.1038/s41467-018-06587-2

NATURE COMMUNICATIONS | 6 (2018) 4147 | DOI: 10.1038/s41467-018-06587-2 | www.nature.com/naturecommunications
A binds in nearly the identical pocket and pose as Rif, and the interactions between the RNAP β subunit residues and the Kang A/Rif scaffold are nearly identical to Rif (Fig. 4c), including the cation-π sandwich (Supplementary Figure 26C and D).

**Structural basis of Kang inhibition of RifR RNAP.** In addition to nearly identical interactions between the RNAP β subunit and the PK backbone of either Rif or Kang A, the chemical moieties unique to Kang A establish new interactions (Fig. 4b, c and Fig. 5). The K-sugar interacts with two β subunit residues that do not contact Rif, R164 and T424. These residues correspond to *Mtb/Eco* RNAP β R173/R143 and T433/S508, respectively. To our knowledge, neither of these residues has ever been identified as conferring RifR when substituted. The K-acid also establishes an interaction with RNAP that does not occur with Rif, a salt bridge (4.4 Å) with the guanidino group of β R604 (Fig. 4b, c). We believe this interaction stabilizes Kang A binding in two ways, first by forming a favorable salt bridge between the negatively charged K-acid and the positively charged R604, but in addition the salt bridge rigidifies the side chain of R604, which may stabilize the cation-π interaction with the Kang A PK backbone (Supplementary Figure 26D).

We propose that the additional interactions with RNAP contributed by the unique Kang moieties (K-sugar and K-acid) stabilize the binding of the Kangs sufficiently to overcome the loss of interactions caused by the S447L substitution, leading to an IC₅₀ for the Kangs against this RifR RNAP that is at least two orders of magnitude lower than Rif (Fig. 3e). To test this hypothesis, we determined the structure of the RifR S447L RNAP in complex with Kang A and compared it to the structures of the wild-type enzyme bound to Rif and to Kang A. The structure was obtained similarly as described for the wild-type enzyme and was refined to 3.45 Å (Fig. 4d, Supplementary Table 5).

In the wild-type RNAP, S447(OG) forms a H-bond with Rif/ Kang A(O2) (Fig. 4b, c) and this favorable interaction is lost with the S447L substitution, leading to an IC₅₀ for the Kangs against this RifR RNAP that is at least two orders of magnitude lower than Rif (Fig. 3e). To test this hypothesis, we determined the structure of the RifR S447L RNAP in complex with Kang A and compared it to the structures of the wild-type enzyme bound to Rif and to Kang A. The structure was obtained similarly as described for the wild-type enzyme and was refined to 3.45 Å (Fig. 4d, Supplementary Table 5).

In the wild-type RNAP, S447(OG) forms a H-bond with Rif/ Kang A(O2) (Fig. 4b, c) and this favorable interaction is lost with the S447L substitution, leading to an IC₅₀ for the Kangs against this RifR RNAP that is at least two orders of magnitude lower than Rif (Fig. 3e). To test this hypothesis, we determined the structure of the RifR S447L RNAP in complex with Kang A and compared it to the structures of the wild-type enzyme bound to Rif and to Kang A. The structure was obtained similarly as described for the wild-type enzyme and was refined to 3.45 Å (Fig. 4d, Supplementary Table 5).

In the wild-type RNAP, S447(OG) forms a H-bond with Rif/ Kang A(O2) (Fig. 4b, c) and this favorable interaction is lost with the S447L substitution, leading to an IC₅₀ for the Kangs against this RifR RNAP that is at least two orders of magnitude lower than Rif (Fig. 3e). To test this hypothesis, we determined the structure of the RifR S447L RNAP in complex with Kang A and compared it to the structures of the wild-type enzyme bound to Rif and to Kang A. The structure was obtained similarly as described for the wild-type enzyme and was refined to 3.45 Å (Fig. 4d, Supplementary Table 5).

In the wild-type RNAP, S447(OG) forms a H-bond with Rif/ Kang A(O2) (Fig. 4b, c) and this favorable interaction is lost with the S447L substitution, leading to an IC₅₀ for the Kangs against this RifR RNAP that is at least two orders of magnitude lower than Rif (Fig. 3e). To test this hypothesis, we determined the structure of the RifR S447L RNAP in complex with Kang A and compared it to the structures of the wild-type enzyme bound to Rif and to Kang A. The structure was obtained similarly as described for the wild-type enzyme and was refined to 3.45 Å (Fig. 4d, Supplementary Table 5).
changes do not affect other RNAP/antibiotic contacts including, importantly, contacts with K-sugar and K-acid (Fig. 4d and Fig. 5).

Binding of Rif to the wild-type RNAP results in a buried surface area of 2,880 Å², while the binding of the RifR RNAP structure results in 3,240 Å². The additional chemical moieties of Kang A contribute an additional 360 Å² of buried surface area gained from the K-sugar and K-acid, supporting our hypothesis.

Structural basis for the Kang mechanism of action. Rif inhibits RNAP function by blocking RNA translocation and extension after formation of the first or second phosphodiester bond31,44,45, resulting in inhibition of full-length transcription production along with an increase of abortive products (Fig. 3e). By contrast, the Kangs inhibit the production of both full-length and abortive products (Fig. 3e), indicating that the Kangs inhibit transcription at a step earlier than Rif.

We next investigated substrate binding and phosphodiester bond formation. We modeled the positions of the first two nucleotide substrates occupying the i and i + 1 sites (the 5’- and 3’-initiating nucleotides, respectively) in an initiating complex by superimposing the structure of a T. thermophilus RNAP de novo initiation complex (4Q4Z)30 onto the Msm RNAP/Rif and Kang A structures (Fig. 6a, b). Rif did not clash sterically with the DNA or the NTP substrates, consistent with findings that Rif has only very small effects on the $K_m$ for initiating substrate31. The Rif piperazine moiety approaches the γ-phosphate of the modeled i site nucleotide ($i\text{NTP}$), and because the Rif piperazine N4 is positively charged and is poised within 3.6 Å from the closest oxygen in the modeled ($i\text{NTP}$ γ-phosphate, this interaction would not disfavor $i\text{NTP}$ binding.

In the modeled de novo initiation complex with Kang A, the pose of Kang A positioned the negatively charged carboxyl group of the K-acid very close (2.5 Å between the closest oxygen of each group) to the negatively charged $\gamma\text{-phosphate}$ of the $i\text{NTP}$ (Fig. 6b), suggesting that Kang A may increase the $K_m$ of the $i\text{NTP}$ by Coulombic repulsion. To test this hypothesis, we took advantage of RNAPs ability to efficiently initiate de novo with an NTP (β-phosphate 6.5 Å from K-acid) or an NMP (α-phosphate 8.0 Å from K-acid) as the 5’-initiating substrate ($i\text{NTP}$, $i\text{NDP}$) (Fig. 6e). As expected, Rif has an inhibitory effect on dinucleotide synthesis (Supplementary Figure 28C)41,42. However, relative to Rif, Kang A has a strong inhibitory effect on RNA dinucleotide synthesis when GTP serves as the 5’-initiating nucleotide, a weaker inhibitory effect with GDP, and no inhibitory effect with GMP (Fig. 6e). These results strongly support the hypothesis that Kang A interferes with binding of the $i\text{NTP}$ substrate via Coulombic repulsion between the K-acid and $i\text{NTP}$ γ-phosphate (Fig. 6).

Note that this mechanism for Kang A inhibition of initial phosphodiester bond formation does not preclude inhibition of RNA chain elongation by steric occlusion, the mechanism of action for Rif41,44,45. Maximal inhibition of pppGpU synthesis (at 1 μM antibiotic) by Kang A is about 75% (Fig. 6d), while inhibition of full-length transcripts in the run-off assay at the same Kang A concentration is essentially 100% (Fig. 3e), indicating that Kang A inhibits RNA chain synthesis via two mechanisms, inhibition of initial phosphodiester bond formation by interfering with binding of the $i\text{NTP}$ substrate, and blocking RNA chain elongation subsequent to formation of the first phosphodiester bond, the latter mechanism being in common with Rif.

Discussion. The potent activity of the Kangs against common mutations that confer RifR in clinical settings suggests that they may have arisen in response to prevalent resistance phenotypes in the producing bacterium’s natural environment. An examination of gene clusters we recovered from soil metagenomes provides potential insight into how these molecules could have evolved from an ancestral rifamycin-like gene cluster through a series of horizontal gene transfer events. The simplest $kng$ related gene cluster we identified (RifCon 12) contains additional biosynthetic genes, not seen in other rifamycin congener gene clusters, that we believe are required for the biosynthesis and transfer of the K-sugar (Fig. 7, genes highlighted in yellow) as well as the formation of the
methyleneoxy bridge on the naphthohydroquinone seen in Kang V2 (Fig. 7, gene highlighted in green). This gene cluster does not, however, contain any genes predicted to encode for the incorporation of either the Emal or K-acid moieties seen in the Kangs (Supplementary Table 6). It is possible that this simpler gene cluster arose from a rifamycin-like molecule can be rationalized with respect to our structural and mechanistic studies. The initial transcription factors, precursor biosynthetic genes, etc.) (Fig. 7, gene organization and accessory gene content (pumps, transduction factors, precursor biosynthetic genes, etc.) (Fig. 7, Supplementary Tables 7 and 8).

In our model, the different stages for the evolution of the Kangs from a simpler rifamycin-like molecule can be rationalized with respect to our structural and mechanistic studies. The initial acquisition of the K-sugar likely proved advantageous in providing additional contacts with RNAP that stabilize the binding of Kang A relative to the same condition with Rif, normalized by the results with no antibiotic. Kang A has a strong inhibitory effect with GTP as the 5'-initiating nucleotide (blue bars), a weaker effect with GDP (red bars), and no inhibitory effect with GMP (green bars). The error bars denote standard error of four measurements.
Fig. 7 Model for the evolution of a structurally complex Kang family molecule. A stepwise increase in the structural complexity of the antibiotic is envisioned to result from a series of horizontal gene transfer events. Genes acquired at each step are shown in boxes and are highlighted according to the structural feature they are predicted to encode.

with RNA transcripts longer than about 3 nucleotides\(^{41,45}\). The ability of RNAP to continuously synthesize 2–3 nucleotide abortive products in the presence of Rif (RNAP priming) likely increases the probability of producing a longer transcript if Rif transiently dissociates within the lifetime of the RNAP/promoter complex, which would block rebinding of Rif. Kang inhibition of initiating substrate binding, mediated by the K-acid, could minimize this priming mechanism. Thus, the novel inhibition mechanism may serve to further increase the potency of the Kangs, especially in the context of Rif\(^R\) mutations that decrease the lifetime of the antibiotic-bound state.

Efforts to improve rifamycin through semisynthesis have been most productive when focusing on modifications of the naphthohydroquinone\(^{52–54}\). Modification of this substructure has yielded the clinically used drugs rifampicin, rifapentine, rifabutin and rifaximin. Interestingly, in the case of the Kangs, evolution has led to the creation of biologically interesting congeners modified at three different positions, all of which have either been largely inaccessible or unproductive in semisynthesis studies.

While we do not know for certain that the evolution of the Kangs provides a selective advantage to the producing organisms in an ecological niche populated by Rif\(^R\) bacteria, their activity against this phenotype suggests this is likely to be the case. Competition between environmental bacteria may have provided strong evolutionary pressure to evolve antibiotic variants that are capable of circumventing common resistance mechanisms, including those that are prevalent in clinical settings. Large-scale metagenome sequencing methods, like those used here, allow for the systematic identification of the most complex gene clusters in known antibiotic families, which may represent highly evolved natural solutions to commonly encountered antibiotic resistance mechanisms. If this proves true across other families of gene clusters that encode antibiotics, a systematic examination of the global microbiome for new congeners of antibiotics in clinical use would likely uncover additional natural products capable of circumventing common clinically important antibiotic resistance mechanisms.

Methods

Screening soil for AHBA synthase gene sequences. eDNA was extracted from each soil sample using a modified DNA extraction protocol\(^{55,56}\). Brieﬂy, approximately 25 g of each soil was placed in a 50 mL falcon tube. 30 mL of lysis buffer (100 mM Tris-HCl, 100 mM ethylenediaminetetraacetic acid (EDTA), 1.5 M NaCl, 1% (w/v) cetyltrimethylammonium bromide, 2% (w/v) sodium dodecyl sulfate, pH 8.0) were added to each tube. After a 2-h incubation at 70 °C with gentle mixing by inversion in 15 min intervals, the tubes were spun down at 7000×g for 10 min at 4 °C. The supernatant was decanted into a clean tube and 0.6 volumes of isopropanol were added to precipitate DNA. Precipitated DNA was pelleted by centrifugation at 5000×g for 30 min at 4 °C. The pellet was washed with 70% ethanol and allowed to air-dry for several hours at room temperature. The dried DNA pellet was resuspended in 500 μL TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). The resulting crude eDNA samples were screened with degenerate primers targeting the AHBA synthase gene, rifK: (forward) 5′-CAGGCTTTCCGATCTGACCTCCG-3′ and (reverse) 5′-ACCGGGAACATSGCCATGTAGTG-3′\(^\text{57}\). These degenerate primers were appended with a collection of distinct 8 bp barcodes and 1–4 bp spacer sequences\(^\text{57}\) that were used to distinguish amplicons generated from each soil. All primers were also appended with adapters for Illumina sequencing: 5′-CTACAC GAGCTTTCCGATCT-3′ (forward primer adapter); 5′-CAGGCTTTCCGATCTGACCTCCG-3′ (reverse primer adapter); 5′-CCGCTTTCCGATCTGACCTCCG-3′ (reverse primer adapter). A typical eDNA PCR reaction contained 1 μL Thermopol master mix (10X stock, New England BioLabs Inc.), 0.1 μL rTaq polymerase (5 units ul\(^{-1}\) stock; Bulldog Bio), 0.5 μL of each primer (10 μM stock concentration), 2 μL of eDNA and 5.9 μL of water. A touchdown PCR protocol was used for all screening: 5 min at 95 °C, followed by 6 cycles of, 30 sec at 95 °C, 30 sec at 65 °C (1 °C/cycle) and 40 s at 72 °C, followed by 29 cycles of 30 s at 95 °C, 30 s at 58 °C and 40 s at 72 °C. PCR reactions were pooled and size selected by electrophoresis using an E-Gel (Invitrogen) prior to sequencing.
Sequencing of AHBA synthase gene amplicons. Sequencing of pooled amplicons was performed by Illumina MiSeq using 300 bp paired-end reads. The forward reads were trimmed to 240 bp, the reverse reads were trimmed to 175 bp from the 3′ end as the sequencing reads and combined with the reclustered centroid amplicon sequences. The combined sequences were aligned with MUSCLE v3.8.319 and a phylogenetic tree was constructed using FastTree v2.1.1060. The phylogenetic tree of AHBA amplicons from cosmids shows that sequences were aligned for sub-clades containing sequences more closely related to rifamycin AHBA synthase sequences than to AHBA synthase genes from other ansamycin gene clusters.

Library construction and screening for AHBA synthase genes. For construction of each metagenomic library, 500 g of soil was to remove large particulate matter and heated to 70 °C for 2 h in lysin buffer (100 mM Tris-HCl, 100 mM EDTA, 1.5 M NaCl, 1% (w/v) cetrimonium bromide, 2% (w/v) SDS, pH 8.0) for 1 h at 50 °C. DNA was then ethanol precipitated and resuspended in TE. High-molecular weight eDNA was gel purified, blunted, and ligated into pWEB::TNC (Epicenter). The ligation products were then packaged and ligated into pWEB::TNC (Epicenter). The ligation products were then packaged and ligated into pWEB::TNC (Epicenter). The ligation products were then packaged and ligated into pWEB::TNC (Epicenter). The ligation products were then packaged and ligated into pWEB::TNC (Epicenter). The ligation products were then packaged and ligated into pWEB::TNC (Epicenter). The ligation products were then packaged and ligated into pWEB::TNC (Epicenter). The ligation products were then packaged and ligated into pWEB::TNC (Epicenter).

NATURE COMMUNICATIONS | DOI: 10.1038/s41467-018-06587-2 | www.nature.com/naturecommunications

Ava fermentation. A spore stock of Ava (NRRL B-24208) was created from cultures grown on MS plates containing mannitol, testosterone, and glycerol. Glycerol was added to 5% v/v at 4°C and stored at 4°C. The resulting crude extract was fractionated by flash chromatography (RedSep RF, High Performance Gold 50 g HP C18 resin) using a linear gradient of 30–100% acetonitrile/water with 0.1% acetic acid over 30 min. The column elution was monitored by UV and fractions containing a strong absorbance at both 254 nm and 192 nm (EC100 Media NMR Core). HRMS analysis was performed at Memorial Sloan-Kettering Cancer Center using an LCT Premier XE system (Waters).

Structure elucidation of Kangs A, V1, and V2. NMR studies for Kang A and V1 were performed at room temperature on a Bruker Avance 600 MHz NMR equipped with a TCI triple resonance cryoprobe. Spectra for Kang A and V1 were collected in methylene chloride and deuterated methanol, respectively. Kang V2 exhibited broad peaks in all common NMR solvents when data were collected at room temperature. The peaks were considerably sharper at lower temperatures. Our final dataset for Kang V2 was collected in deuterated methanol at ~20 °C on a Bruker Avance HD 500 MHz NMR equipped with TCI cryoprobe with enhanced HD19 and C13 detection and low temperature (~40 °C) capabilities (Well Cornell Medicine NMR Core). HRMS analysis was performed on Memorial Sloan-Kettering Cancer Center using an LCT Premier XE system (Waters).
transcription assay. Recombinantly produced wild-type and S447L mutant DNA-dependent RNAP were purified from *M. smegmatis* MGM6209 strain expressing a chromosomal copy of *rpoC* with a C-terminal pHis-tag, and either wild-type *rpoB* gene or *rpoB* mutant allele (S447L). *M. smegmatis* cells were grown to late exponential phase and collected at the Bioexpression and Fermentation Facility at the University of Georgia. Cells were lysed in a French press (Avestin) in 50 mM Tris-HCl, pH 8, 1 mM EDTA, 5% (v/v) glycerol, 5 mM DTT, 1 mM protease inhibitor cocktail, and 1 mM phenylmethylsulfonyl fluoride, and RNAP was precipitated from the cleared lysate by polyethyleneimine (PEI) precipitation. The PEI precipitate was washed three times with 50 mM Tris-HCl, pH 8, 0.5 mM CaCl2, 1 mM EDTA, 5 mM DTT, and 5% (v/v) glycerol, then eluted three times with the same buffer but with 1 M NaCl. Protein was precipitated overnight with 35% (v/v) ammonium sulfate and resuspended in 20 mM Tris-HCl, pH 8, 5% (v/v) glycerol, 1 M NaCl, and 1 mM β-mercaptoethanol. Protein was loaded on a Ni2+–affinity column (HitTrap IMAC-HP, GE Healthcare Life Sciences) and eluted in 20 mM Tris-HCl, pH 8, 5% (v/v) glycerol, 0.5 M NaCl, and 0.25 M imidazole. Protein was diluted in 10 mM Tris-HCl, pH 8, 5% (v/v) glycerol, 0.1 mM EDTA, and 5 mM DTT to a final salt concentration of 0.1 M NaCl loaded on a Biorex (BioRad, Hercules, CA) ion exchange column, and eluted with a salt gradient (0–1.0 M NaCl). To generate crystals of the antibiotic to the RNAP, the mixtures were incubated at 37 °C for 5 min. Following incubation, 10 nM of AP3 promoter was added to each tube and the samples were incubated for an additional 15 min at 37 °C to allow formation of the RNAP open complex. Transcription was initiated by the addition of a nucleotide mixture consisting of 5 mM ATP, 200 μM UTP, 100 μM CTP, and 12.5 μM (0.3 μM) γ−32P-UTP. Each reaction was allowed to proceed for 15 min at 37 °C before the addition of 20 μl of stop buffer (0.5X TBE, pH 8.3, 8.5 μM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol). Reactions were then heated to 95 °C for 10 min and loaded onto a polyacrylamide gel (23% Acrylamide/Bis acrylamide (19:1), 6 Urea, 1X TBE, pH 8.3). Gels were run for 3 h at 500 V, then exposed on a phosphorimaging plate (GE Healthcare) for 12 hrs before being imaged using a Typhoon 9400 Variable Imager (Amersham Biosciences). Uncropped gel images are shown in Supplementary Figure 29.

**Cristallization of Msm RbpA/TIC.** Crystals of the Msm RbpA/TIC were prepared with purified Msm RbpA/A-holo and the upstream fork DNA. The DNA was assembled from synthetic GTTCTTGTCTTTGGTCAATC-3′ and 5′-AGGACAATTATTTACACTTTTGTCAACG-3′; Integrated DNA Technologies, Coralville, IA) by annealing in 10 mM Tris-HCl, pH 8, 1 mM EDTA, and 0.2 M NaCl. Protein was mixed in a 1:1 molar ratio with annealed upstream fork DNA for 15 min at room temperature to generate the Msm RbpA/TIC. Crystals were grown by hanging drop vapor diffusion by mixing 1 μl of Msm RbpA/TIC solution (11 mg ml−1 protein) with 1 μl of crystallization solution [0.1 M Bis-Tris, pH 6.0, 0.2 M LiSO4, 16% (v/v) polyethylene glycol 3350, 2.5% (v/v) ethylene glycol] and incubating over a well containing crystallization solution at 22 °C. Stabilization solutions [0.1 M Bis-Tris, pH 6.0, 0.2 M LiSO4, 20% (v/v) polyethylene glycol 3350, 2.5% (v/v) ethylene glycol, 1% (v/v) DMSO] with either 0.1 mM Rif or 0.1 mM Kmg were prepared by adding Rif or Kang A from 10 μM stock solutions in 100% DMSO. Fully grown crystals were transferred into stabilization solution and incubated for 4 h to allow antibiotic binding. The crystals were then cryo-protected by step-wise transfer (three steps) into 0.1 M Bis-Tris, pH 6.0, 0.2 M LiSO4, 22% (v/v) polyethylene glycol 3350, 20% (v/v) glycerol, 1% (v/v) DMSO and either 0.1 mM Rif or Kang A, and flash frozen by plunging into liquid nitrogen. Crystals of the S447L β RNAP variant of the Msm RbpA/TIC were grown and frozen as described for the wild-type RNAP.

**Data collection, structure determination, and refinement.** X-ray diffraction data were collected at the Argonne National Laboratory Advanced Photon Source (APS) NE-CAT beamline 24-ID-C (Rif/TIC) or the National Synchrotron Light Source II (NSLSII) AMX beamline 17-ID-D (Kang/A-TIC). Structural biology software was accessed through the SBGrid consortium. Data were integrated and scaled using HKL2000.

Starting with SWT1, the models were first improved by rigid body refinement of 20 individual mobile domains using PHENIX. The resulting models were improved by iterative cycles of manual building with COOT and refinement with PHENIX. Differences Fourier maps revealed excellent electron density for Rif or...
Kang A. Rif² or Kang A²³ crystal structures were easily modeled into the respective difference densities, but the Kang A structure required inversion of the coordinates through the origin (i.e., the deposited coordinates are the mirror image of the molecule). Further iterative cycles of building and refinement yielded the final models (Supplementary Table 5).

DNase I footprinting. AC50 promoter DNA with a 5′-end-labeled PCR primer (5′-GGCGCCTAGGGCTTTACCTGCTGATTCGAT-3′). The primer was initially labeled with 32P using substrate [y-32P] ATP and T4 polynucleotide kinase, followed by purification using a Nunc/Away nucleotide removal kit. The resulting PCR product was loaded on a nondenaturing 5% acrylamide gel and the DNA was eluted from the gel by the crush-soak method. DNaseI (New England Biolabs) was diluted to 200 U µl⁻¹ and kept on ice. Reactions (20 µl) were carried out in a 37 °C water bath with proteins using the following protocol: A-holoenzyme (50 nM) and increasing concentrations of Kang A or Rif, followed by 5 min incubation at 37 °C. The control reaction without antibiotics was also done. Formation of Rpo was allowed to proceed for 15 min. DNase I (200 U) was then added to the mixture and the reactions were incubated for an additional 2 min. The reactions were quenched by the addition of 100 µl of 0.5 M phenol, and 80 µl of a mixture of sodium acetate (375 mM) and EDTA (12.5 mM final), and 2 µl glycophase. The DNA was recovered in the aqueous layer, ethanol precipitated and washed. The air-dried pellet was resuspended in 2X loading buffer, heated at 95 °C for 1 min before being immediately loaded on an 8% polyacrylamide gel (19:1 acrylamide:bis-acrylamide). Abortive products were visualized by autoradiography and the band intensity was quantified using ImageJ16. Uncropped gel images are shown in Supplementary Figure 30.

De novo transcription initiation assay. Promoter DNA (+87 to +71 of Mifl5rrnAP3) was engineered with an engineered mutation, I3 T > A, was synthesized (GenScript) and placed into the pUC57 plasmid to generate pUC57-AP3 GU. Fragment +87 to +71 of pUC57-AP3 GU was PCR amplified. The promoter DNA fragment was subsequently separated on an agarose gel and gel purified (Qiagen). This promoter fragment served as the template for de novo initiation with GTP, GDP, or GMP as substrates (1 mM GTP, 2 mM GDP, or 4 mM GMP), plus 50 µM of unlabeled UTP + α32P-UTP. After 10 min of incubation, reactions were heated at 95 °C for 2 min and loaded on a 25% polyacrylamide gel (19:1 acrylamide:bis-acrylamide). Abortive products were visualized by phosphorylating and digitized using a Typhoon phosphorimager. Data were quantified using ImageJ.16 Uncropped gel images are shown in Supplementary Figure 30.

Data availability

The X-ray crystallographic coordinates and structure factor files have been deposited in the Protein Data Bank with accession codes 6CCV (Rif/TIC), 6CCE (Kang A/TIC), 6DCF (Kang A/S447L mutant). Metagenomic DNA sequences described in this manuscript have been deposited in GenBank with the accession numbers MH480516 to MH480581.

Received: 25 May 2018 Accepted: 29 August 2018 Published online: 08 October 2018

References

1. Chakraborty, S. & Rhee, K. Y. Tuberculosis drug development: history and evolution of the mechanism-based paradigm. Cold Spring Harb. Perspect. Med. 5, a021147 (2015).
2. Žuml, A. et al. The WHO 2014 Global tuberculosis report—further to go. Lancet Glob. Health 3, e10–e12 (2015).
3. Ramasawamy, S. & Musser, J. M. Molecular genetic basis of antimicrobial agent resistance in Mycobacterium tuberculosis: 1998 update. Tuber. Lung Dis. 79, 1–29 (1998).
4. Perron, G. G. et al. Functional characterization of bacteria isolated from ancient arctic soil exposes diverse resistance mechanisms to modern antibiotics. PLoS One 10, e0095533 (2015).
5. D’Costa, V. M. et al. Antibiotic resistance is ancient. Nature 477, 457–461 (2011).
6. Bhullar, K. et al. Antibiotic resistance is prevalent in an isolated cave microbiome. PLoS One 7, e34953 (2012).
7. Dawson, S., Malkinson, J. P., Paumier, D. & Searcy, M. Bistinctoral natural products with potential therapeutic applications: isolation, structure determination, synthetic and biological studies. Nat. Prod. Rep. 24, 109–126 (2007).
8. Kahne, D., Leimkuhler, C., Lu, W. & Walsh, C. Glycopeptide and lipopeptide antibiotics. Chem. Rev. 105, 425–448 (2005).
9. Strieker, M. & Marahiel, M. A. The structural diversity of acidic lipopeptide antibiotics. ChemBioChem 10, 607–616 (2009).
10. Roesch, L. F. et al. Pyrossynthetizing enzymes and contrasts soil microbial diversity. ISME J. 1, 283–290 (2007).
11. Daniel, R. The metagenomics of soil. Nat. Rev. Microbiol. 3, 470–478 (2005).
12. Milštejn, A., Schneider, J. S. & Brady, S. F. Mining the metabolome: identifying novel natural products from microbial communities. Chem. Biol. 21, 1211–1223 (2014).
13. Reddy, B. V. et al. Natural product biosynthetic gene diversity in geographically distinct soil microorganisms. Appl. Environ. Microbiol 78, 3744–3752 (2012).
14. Floss, H. G., Yu, T. W. & Arakawa, K. The biosynthesis of 3-amino-5-hydroxybenzoic acid (AHVA), the precursor of mC7N units in ansamycin and mitomycin antibiotics: a review. J. Antibiot. 64, 35–44 (2011).
15. Wang, H. X. et al. PCR screening reveals considerable unexploited biosynthetic potential of ansamycins and a mysterious family of AHVA-containing natural products in actinomycetes. J. Appl. Microbiol 115, 77–85 (2013).
16. Huito, Z. et al. PCR screening of 3-amino-5-hydroxybenzoic acid synthase gene leads to identification of ansamycins and AHVA-related antibiotic producers in Actinomycetes. J. Appl. Microbiol 106, 755–763 (2009).
17. Wood, S. A., Kirby, B. M., Goodwin, C. M., Le Roes, M. & Meyers, P. R. PCR screening reveals unexpected antibiotic biosynthetic potential in Amycolatopsis sp. strain UM16. J. Appl. Microbiol. 102, 245–253 (2007).
18. Zhu, N., Zhao, P. & Shen, Y. Selective isolation and ansamycin-targeted screening reveals unexpected antibiotic biosynthetic potential of the “maytansinoids-producing” arboresal Trewia nudiflora. Curr. Microbiol. 58, 87–94 (2009).
19. Everest, G. J. & Meyers, P. R. Evaluation of the antibiotic biosynthetic potential of the genus Amycolatopsis and description of Amycolatopsis ciri sp. nov., Amycolatopsis equina sp. nov. and Amycolatopsis hippodromi sp. nov. J. Appl. Microbiol. 111, 300–311 (2011).
20. Dunn, B. J. & Khosla, C. Engineering the acyltransferase substrate specificity of assembly line polyketide synthases. J. R. Soc. Interface 10, 20130297 (2013).
21. Minowa, Y., Araki, M. & Kanesha, M. Comprehensive analysis of distinctive polyketide and nonribosomal peptide structural motifs encoded in microbial genomes. J. Mol. Biol. 368, 1500–1517 (2007).
22. Bao, G. H. et al. Structure determination of kangelimycin A by X-ray diffraction method. Sci. Sin. B 30, 1181–1189 (1987).
23. Wang, N. J. et al. Isolation and structure of a new ansamycin antibiotic kangelimycin A from a Nocardia. J. Antibiot. 41, 264–267 (1988).
24. Rateb, M. E. et al. Chaxamycins A–D, bioactive ansamycins from a hyper-arid desert Streptomyces sp. J. Nat. Prod. 74, 1491–1499 (2011).
25. Bihlmaier, C. et al. Biosynthetic gene cluster for the polyenoyltetramic acid antibiotic mitomycin antibiotics: a review. Angew. Chem. Int. Ed. Engl. 44, 9814–9859 (2005).
26. Ilkevzawa, N. et al. Molecular cloning and characterization of CYP719, a methylenedioxy bridge-forming enzyme that belongs to a novel P450 family, from cultured Captis japonica cells. J. Biol. Chem. 278, 38587–38565 (2003).
27. Goldstein, B. P. Resistance to rifampicin: a review. J. Antibiot. 67, 625–630 (2014).
30. Srivastava, A., Degen, D., Ebright, Y. W. & Ebright, R. H. Frequency, spectrum, and nonzero fitness costs of resistance to mycoxonpr in Streptomyces aureus. Antimicrob. Agents Chemother. 56, 6250–6255 (2012).

31. Hubin, E. A. et al. Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA. Elife 6 e22520 (2017).

32. Tadese, M. et al. Drug resistance-conferring mutations in Mycobacterium tuberculosis cause detrimental infections in patients in South West Ethiopia. Int J. Mycobacteriol. 5, 185–191 (2016).

33. Thirumurugan, R. et al. Molecular analysis of rpoB gene mutations in rifampicin-resistant Mycobacterium tuberculosis isolates by multiple allele specific polymerase chain reaction in Puducherry, South India. J. Infect. Public Health 8, 619–625 (2015).

34. Saadu, A. et al. Molecular characterization of rifamp- and isoniazid-resistant Mycobacterium tuberculosis strains isolated in Poland. J. Clin. Microbiol. 42, 2425–2431 (2004).

35. Cavusoglu, C., Hilmioglu, S., Guneri, S. & Bilgic, A. Characterization of rpoB mutations in rifampicin-resistant clinical isolates of Mycobacterium tuberculosis from Turkey by DNA sequencing and line probe assay. J. Clin. Microbiol. 40, 4435–4438 (2002).

36. Brossier, F., Veziris, N., Truffot-Pernot, C., Jarlier, V. & Sougakoff, W. Performance of the genotype MTBDR line probe assay for detection of resistance to rifampicin and isoniazid in strains of Mycobacterium tuberculosis with low- and high-level resistance. J. Clin. Microbiol. 44, 3569–3564 (2006).

37. Sinha, R. et al. Prevalence of multirresistant tuberculosis and mutations in rpoB gene of Mycobacterium tuberculosis isolates from Pakistan state of India. J. Epidemiol. Glob. Health 7, 175–180 (2017).

38. Johansen, I. S., Lundgren, B., Sonnovskaja, A. & Thomsen, V. O. Direct detection of multidrug-resistant Mycobacterium tuberculosis in clinical specimens in low- and high-incidence countries by line probe assay. J. Clin. Microbiol. 41, 4477–4486 (2003).

39. Barnard, M., Albert, H., Coetzee, G., O’Mahony, M., M, E. Rapid molecular screening for multidrug-resistant tuberculosis in a high-volume public health laboratory in South Africa. Am. J. Respir. Crit. Care Med. 177, 787–792 (2008).

40. Muthaiah, M. et al. Prevalence of mutations in genes associated with secondary metabolism. Elife 2, e01456 (2013).

41. McClure, W. R. & Cech, C. L. On the mechanism of rifampicin inhibition of RNA synthesis. J. Biol. Chem. 253, 8949–8956 (1978).

42. Hinkle, D. C., Mangel, W. F. & Chamberlin, M. J. Studies of the binding of Escherichia coli RNA polymerase to DNA. IV. The effect of rifampicin on binding and on RNA chain initiation. J. Mol. Biol. 70, 209–220 (1972).

43. Pupov, D., Kuzin, I., Bass, I. & Kubalchinsky, A. Distinct functions of the bacterial RNA polymerase sigma subunit region 3.2 in RNA priming and promoter escape. Nucleic Acids Res. 42, 4494–4504 (2014).

44. Feklistov, A. et al. Rifamycins do not function by allosteric modulation of bacterial RNA polymerase. J. Biol. Chem. 289, 24549–24559 (2014).

45. Molodtsov, V., Scharf, N. T., Stefan, M. A., Garcia, G. A. & Murakami, K. S. Distinct functions of the bacterial rna polymerase. Mol. Microbiol. 67, 1094–1107 (2008).

46. Gonzalez-y-Merchand, J. A., Colston, M. J. & Cox, R. A. The RNA operons of Mycobacterium smegmatis and Mycobacterium tuberculosis: comparison of promoter elements and of neighbouring upstream genes. Microbiology 142 (Pt 3), 667–674 (1996).

47. Brufani, M., Cerrini, S., Fedeli, W. & Vacciago, A. Rifamycins: an insight into biological activity based on structural investigations. J. Mol. Biol. 87, 409–435 (1974).

48. Davis, E., Chen, J., Leon, K., Darst, S. A. & Campbell, E. A. Mycobacterial RNA polymerase forms unstable open promoter complexes that are stabilized by CarD. Nucleic Acids Res. 43, 443–445 (2015).

49. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9, 671–675 (2012).

Acknowledgments
We acknowledge the High Throughput Screening and Resource Center at the Rockefeller University and the Weill Cornell Medicine NMR Core for use of their facilities for NMR studies. We thank Carl Nathan (Weill Cornell Medicine) for his assistance with testing the activity of the kanglemycins against Mtb strains. We thank Hande Boyaci for help analyzing the data regarding the kanglemycin mechanism of inhibition. This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the NIHGMs from the NIH (P41 GM110714) to S.F.B and 1R35 GM122559 to S.F.B. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated by the DOE Office of Science by Brookhaven National Laboratory under contract No. DE-AC02-17CH03000. The Advanced Photon Source is a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract No. DE-AC02-06CH11357. The research used the 17–1 (1) (AMX) beamline at the National Synchrotron Light Source II, which is supported by the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE SC0012704. The use of the Rockefeller University Structural Biology Resource Center was made possible by NIH/NCRR 1S10RR027057. This project was supported by grants from the Bill and Melinda Gates Foundation (OPP117928) to S.F.B., as well as by NIH grants R01 CA108748 to M.S.G, R01 GM111810 to S.A., R01 GM114450 to E.A.C., 5U01 GM110714 to S.F.B and 1R35 GM122559 to S.F.B.

Author contributions
Conceptualization: S.A.D., E.A.C., and S.F.B; Investigation: J.P., M.L., D.M., A.M., J.W., J.B.B., M.A.T., P.Y.C., M.D., T.W., K.S., N.B., and A.F.; Analysis: J.P., M.L., D.M., S.!
A.D., E.A.C., and S.F.B.; Writing: S.A.D., E.A.C., S.F.B., J.P., A.M., and M.L.; Supervision: M.S.G., S.A.D., E.A.C., and S.F.B.; Funding acquisition, S.A.D., M.S.G., E.A.C., and S.F.B.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-06587-2.

Competing interests: S.F.B. founded and serves as a consultant for Lodo Therapeutics. The remaining 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.