ARTICLES
https://doi.org/10.1038/s41564-021-01013-8
nature microbiology

Identification of structurally diverse menaquinone-binding antibiotics with in vivo activity against multidrug-resistant pathogens

Lei Li[1], Bimal Koirala[1], Yozen Hernandez[1], Logan W. MacIntyre[1], Melinda A. Ternei[1], Riccardo Russo[2] and Sean F. Brady[1]✉

The emergence of multidrug-resistant bacteria poses a threat to global health and necessitates the development of additional in vivo active antibiotics with diverse modes of action. Directly targeting menaquinone (MK), which plays an important role in bacterial electron transport, is an appealing, yet underexplored, mode of action due to a dearth of MK-binding molecules. Here we combine sequence-based metagenomic mining with a motif search of bioinformatically predicted natural product structures to identify six biosynthetic gene clusters that we predicted encode MK-binding antibiotics (MBAs). Their predicted products (MBA1–6) were rapidly accessed using a synthetic bioinformatic natural product approach, which relies on bioinformatic structure prediction followed by chemical synthesis. Among these six structurally diverse MBAs, four make up two new MBA structural families. The most potent member of each new family (MBA3, MBA6) proved effective at treating methicillin-resistant *Staphylococcus aureus* infection in a murine peritonitis-sepsis model. The only conserved feature present in all MBAs is the sequence 'GXLXXXW', which we propose represents a minimum MK-binding motif. Notably, we found that a subset of MBAs were active against *Mycobacterium tuberculosis* both in vitro and in macrophages. Our findings suggest that naturally occurring MBAs are a structurally diverse and untapped class of mechanistically interesting, in vivo active antibiotics.

Antimicrobial resistance represents a major and growing healthcare problem and contributes annually to approximately 700,000 deaths worldwide[1,2]. The widespread emergence of multidrug-resistant (MDR) pathogens necessitates the development of in vivo active antibiotics that differ in mode of action from those currently in clinical use[3–5]. In most anaerobic and Gram-positive bacteria, menaquinone (MK) plays an important role in electron transport[6]. Humans are not capable of producing MK, making it an appealing target for antibiotic development[7]. Historically, inhibition of MK biosynthesis by synthetic small molecules has been the predominant mode of action explored to develop mechanistically new antibiotics[7–9]. Three closely related non-ribosomal peptide synthetase (NRPS)-derived bacterial cyclic lipopeptides (lysocin E, WAP-8294A2 and WBP-29479A1) have been shown to kill bacteria by binding directly to MK to induce membrane disruption and rapid cell lysis (Fig. 1)[10–12]. Recently, Santiago et al.[13] reported that lysocin E also binds to Lipid II, a precursor for bacterial cell wall synthesis. Although WAP-8294A2 (lotilibcin) progressed to phase I clinical trials, a dearth of additional chemical entities that can kill bacteria through MK binding has hindered the successful therapeutic development of this mechanistically interesting class of antibiotics[14].

The search for additional bacterially produced MK-binding antibiotics (MBAs) is limited by the fact that most of the biosynthetic diversity in the global microbiome is functionally inaccessible. This is because of our inability to culture most bacteria and because only a small subset of biosynthetic gene clusters (BGCs) found in cultured bacteria is expressed in laboratory fermentation studies[15,16]. While these factors limit direct functional screening for additional MBAs, next-generation sequencing methods are revealing high numbers of previously inaccessible bacterial BGCs from both cultured bacteria and diverse metagenomes[17–19]. The genetic information contained in a BGC has historically been decoded using biological processes (that is, transcription, translation and biosynthesis). This paradigm is limited by the fact that most BGCs are natively silent in laboratory fermentation studies and even the best activation strategies can only activate a small fraction of BGCs in native or heterologous organisms[15]. The increasing accuracy of structural predictions derived from the bioinformatic analysis of BGCs presents the alternative possibility of accessing the metabolite encoded by a BGC using total chemical synthesis of its bioinformatically predicted product (that is, a synthetic bioinformatic natural product (syn-BNP))[20,21].

In this study, we took two orthogonal bioinformatic approaches to guide the discovery of bacterial BGCs that encode structurally new classes of MBAs (Fig. 1). First, in a culture-independent approach, sequence homology was used to identify predicted MBA BGCs from diverse soil metagenomes. A detailed analysis of BGCs that arose from this study led us to propose a minimal MK-binding motif (GXLXXXW). Next, in what we believe is a unique search strategy, we screened a large database of bioinformatically predicted natural product structures for this proposed minimal MK-binding motif to identify additional potential MBA BGCs in sequenced bacterial genomes. Total chemical synthesis of the structures predicted to arise from the BGCs we identified resulted in six structurally diverse MBAs. All six MBAs were broadly active against MDR Gram-positive pathogens. Notably, we showed that a subset of MBAs is active against MDR *Mycobacterium tuberculosis* both in vitro and in a macrophage assay, defining MK-binding as an anti-*M. tuberculosis* mode of action. Among the antibiotics we identified, four fell into two new structural classes. Antibiotics from both

[1]Laboratory of Genetically Encoded Small Molecules, The Rockefeller University, New York, NY, USA. [2]Department of Medicine, Center for Emerging and Re-emerging Pathogens, Rutgers University-New Jersey Medical School, Newark, NJ, USA. ✉e-mail: sbrady@rockefeller.edu

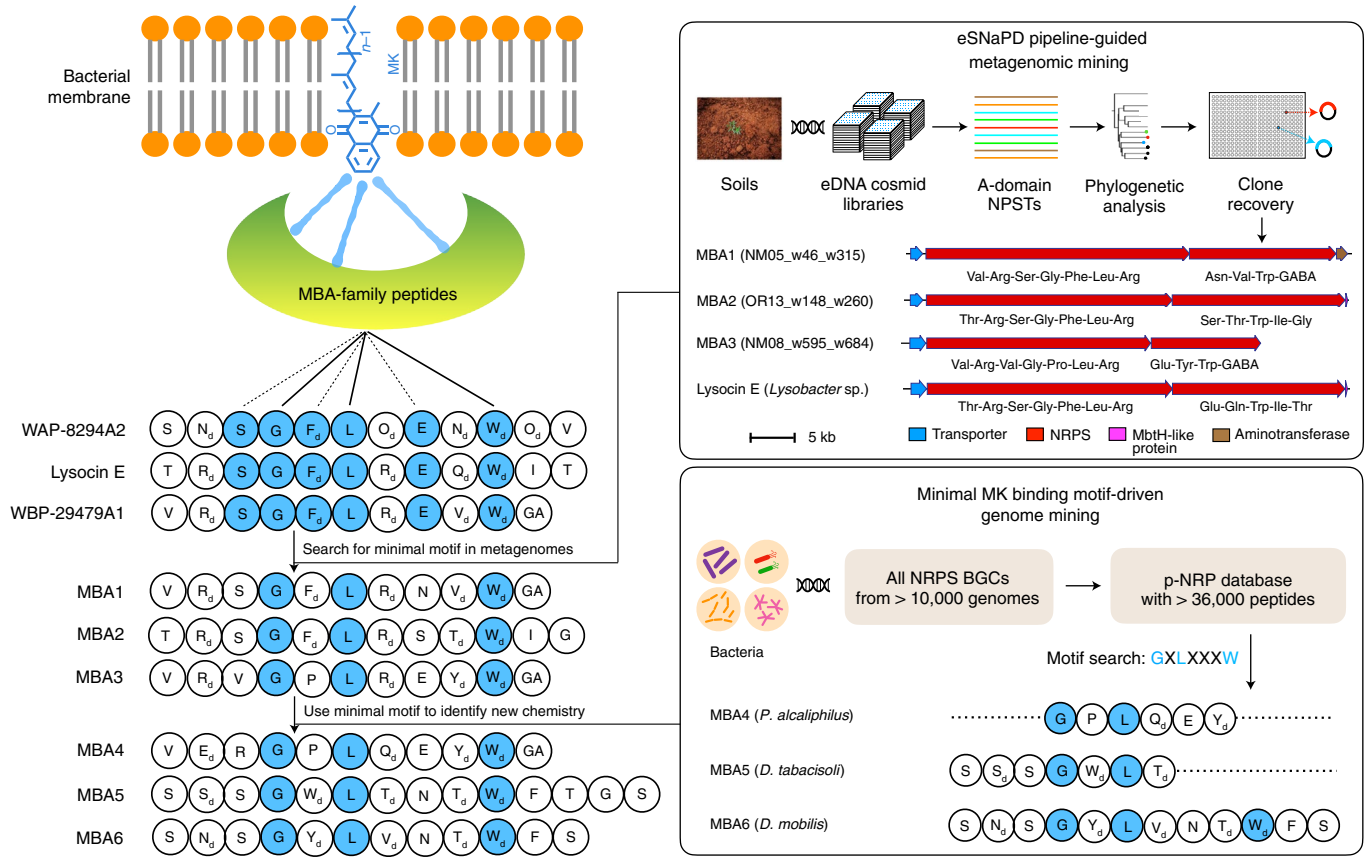
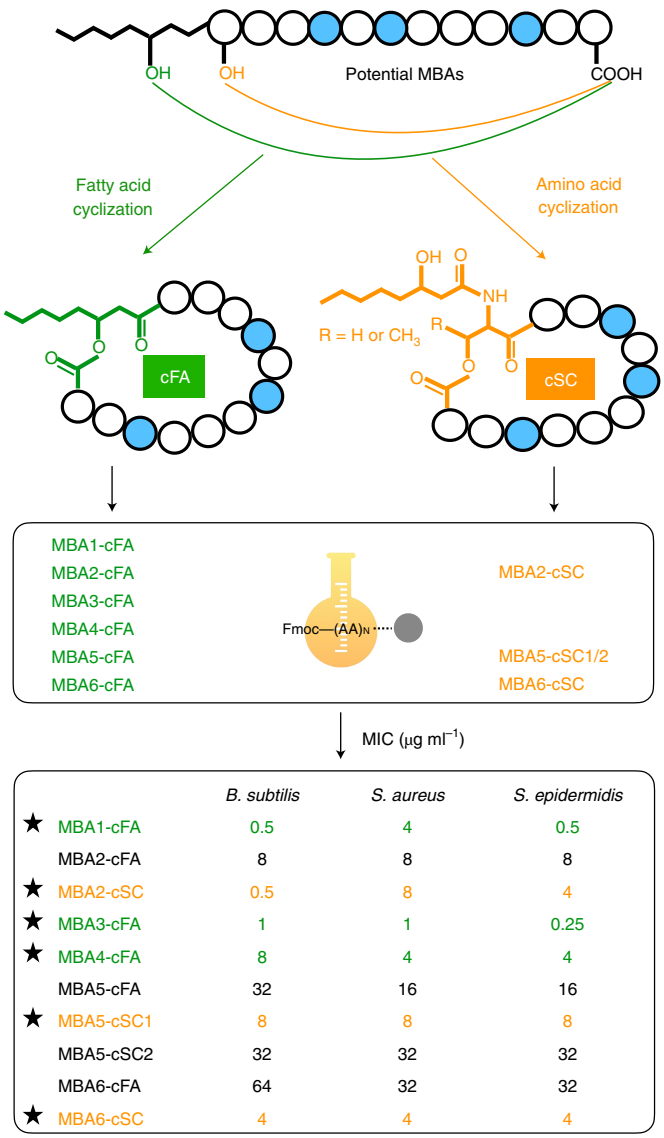


Fig. 1 | Identification of BGCs predicted to encode new MBAs. Using a sequence-based soil metagenome BGC discovery pipeline, three BGCs were identified that showed high A-domain sequence identity and similar overall gene organization to known MBA BGCs. We predicted that each of these encoded a new MBA. Using the conserved GXLXXXW motif that we detected in structurally diverse MBAs to search an in-house database of p-NRP structures, we identified three additional BGCs that we predicted would encode new MBAs. The GXLXXXW motif we found in all MBAs is predicted to represent the minimal MK-binding motif that is necessary for the antibacterial activity of this underexplored and structurally diverse class of natural antibiotics.

new structural classes proved effective against methicillin-resistant *Staphylococcus aureus* (MRSA) in a murine peritonitis-sepsis model, thus providing two new MBAs for use in the development of antibiotics with different modes of action and activity against MDR pathogens. We believe the general approach presented in this study of searching a database of structures bioinformatically sequenced to identify BGCs that encode molecules with specific desired features is broadly applicable to the search for bioactive small molecules.

Results

Identification of BGCs predicted to encode MBAs. *Metagenomic search for BGCs predicted to encode new MBAs.* We began our search for BGCs that might encode MBAs by looking at the NRPS adenylation (A)-domain sequence data generated from soil metagenomic libraries. As part of our ongoing soil metagenome-guided natural product discovery programme, we created a collection of saturating cosmid-based soil metagenomic libraries for use in targeted BGC discovery studies[22–25]. The construction and screening of these libraries for A-domain sequences associated with a specific natural product family have been described in detail previously[22,25]. Briefly, for each library, DNA extracted directly from soil (environmental DNA (eDNA)) was used to construct $>2\times10^7$ unique cosmid clones in *Escherichia coli*. Each library was arrayed as a set of sub-pools containing approximately 5,000–25,000 unique cosmids each. DNA from each sub-pool was used as template in a PCR reaction with a unique set of barcoded A-domain degenerate primers. The resulting amplicons were sequenced using Illumina technology to generate a database of A-domain sequences (Natural Product Sequence Tags (NPSTs)) that can be used to track BGCs of interest to specific library sub-pools. Using the environmental Surveyor of Natural Product Diversity (eSNaPD) software[26], each A-domain from a BGC that encoded one of the three known MBAs was compared to all of the library-derived A-domain NPSTs. Known MBAs shared six positionally conserved amino acids: L-Ser-3; Gly-4; D-Phe-5; L-Leu-6; L-Glu-8; and D-Trp-10 (Fig. 1). NPSTs that returned low e-values ($<10^{-12}$ to $<10^{-60}$) for the A-domains that installed one of these six conserved residues were used to generate six A-domain-specific phylogenetic trees (Extended Data Fig. 1). eDNA cosmid clones containing BGCs associated with A-domains that fell into the same or a closely related clade as an A-domain from known MBA BGCs were recovered from the appropriate library sub-pools. Fully sequenced and annotated eDNA-derived NRPS BGCs were analysed for the potential to encode MBA-like peptides. The linear peptide encoded by each eDNA-derived BGC was predicted using the ten amino acid residues that line each A-domain substrate binding pocket (A-domain signature sequence)[27,28]. Based on this analysis, no predicted peptides contained all six residues that were conserved among known MBAs. However, in three cases where the eDNA-derived BGC showed a similar gene organization to that seen in known MBA BGCs, the predicted peptide products shared some sequence similarity to known MBAs (Extended Data Fig. 2), leading us to explore the possibility that the structures predicted to arise from these BGCs might be MBAs. As shown in the



		B. subtilis	S. aureus	S. epidermidis
★	MBA1-cFA	0.5	4	0.5
	MBA2-cFA	8	8	8
★	MBA2-cSC	0.5	8	4
★	MBA3-cFA	1	1	0.25
★	MBA4-cFA	8	4	4
	MBA5-cFA	32	16	16
★	MBA5-cSC1	8	8	8
	MBA5-cSC2	32	32	32
	MBA6-cFA	64	32	32
★	MBA6-cSC	4	4	4

Fig. 2 | Synthesis and antibacterial activity of syn-BNPs based on BGCs predicted to encode MBAs. The (*R*)-3-hydroxy-octanoic acid-derivatized linear peptides that are predicted to be encoded by potential MBA BGCs were cyclized through either the hydroxyl group of the fatty acid (cFA) or through a nucleophilic amino acid side chain (cSC). When the first amino acid was predicted to contain a nucleophilic side chain (that is, a serine or threonine) both the cFA and cSC analogues were synthesized (MBA2, MBA5 and MBA6). In the case of MBA5-cSC2, the serine at position 2 was also used for cyclization. If the first amino acid of the peptide did not contain a nucleophilic side chain, only a cFA derivative was produced (MBA1, MBA3 and MBA4). syn-BNPs marked with an asterisk were the most active to arise from each BGC and were assumed to be the 'naturally cyclized' versions of the potential MBAs. All MIC assays were done in duplicate ($n = 2$).

bioactivity analysis presented below, each of these predicted peptides in fact represents a new MBA (Fig. 2).

GXLXXXW motif search of sequenced genomes for BGCs predicted to encode MBAs. In a second round of screening, we turned to sequenced bacterial genomes to see if we could identify BGCs that might encode additional MBAs. For this study, we analysed BGCs from approximately 10,000 bacterial genomes. The A-domain substrate binding pockets from NRPS BGCs in these genomes were compared to a manually curated list of A-domain signature sequences from characterized BGCs (Methods). Based on the A-domain substrate predictions that arose from this analysis, we generated a database of linear peptides that were predicted to arise from >36,000 NRPS BGCs (Fig. 1 and Supplementary Fig. 1). Our initial search failed to identify any BGCs that might encode a peptide that contains the same six amino acid pattern seen in previously characterized MBAs. By including metagenome-derived MBA sequences in this study, we reduced the conserved residues observed across MBAs from the initial six residues to a proposed minimal MK-binding motif of just three residues (Gly-X-Leu-X-X-X-Trp, GXLXXXW; Fig. 1). A search of our database of predicted NRPS-derived linear peptides with the simpler GXLXXXW motif identified one BGC in the genome of *Dyella mobilis* suggesting that it might encode a potential MBA (Fig. 1 and Extended Data Fig. 3). Since NRPS BGCs are often truncated in genomes assembled using short-read sequence data, we also searched our database for subsequences of the GXLXXXW motif. GXL was found in 22 peptides and LXXXW was only found in the peptide from *D. mobilis* (Supplementary Table 1). A manual examination the remaining amino acids in these 22 peptides identified 2 additional examples where peptides shared additional amino acids with newly predicted potential MBAs. One example was from a truncated BGC in *Dyella tabacisoli* and the other was found in partially sequenced NRPS BGCs from four different *Paracoccus* strains (Extended Data Fig. 3 and Supplementary Table 1). To access the complete non-ribosomal peptide (NRP) sequence encoded by each of these BGCs, we resequenced the genomes of *D. tabacisoli* KCTC 62035 and *Paracoccus alcaliphilus* ATCC 51199. In both cases, A-domain analysis of resequenced BGCs predicted that they would encode GXLXXXW-containing peptides that might represent MBAs (Fig. 1).

Prediction and synthesis of ten potential MBA BGC products. In total, we identified six BGCs that we predicted might encode MBAs. Traditional natural product discovery methods that rely on biological systems to decode the information present in a BGC are hindered by a lack of BGC expression in laboratory fermentation studies and the time-consuming process of isolating and structurally characterizing molecules from bacterial fermentation broths. In cases where the final product of a BGC can be bioinformatically predicted with confidence, total chemical synthesis of this product can provide an alternative and potentially more straightforward means of accessing the metabolite encoded by the BGC[20,21]. Each potential MBA BGC contains two large NRPS genes with a condensation start domain that is predicted to initiate NRPS biosynthesis with a fatty acid[29]. As described above, A-domain substrate specificity analysis allowed us to predict with high confidence the amino acid incorporated by every A-domain found in these BGCs (Supplementary Fig. 2). With the exception of the MBA6 BGC, no BGCs were predicted to encode tailoring enzymes (Fig. 1 and Extended Data Fig. 3). The dioxygenase encoded by the MBA6 BGC was expected to be involved in the hydroxylation of the asparagine incorporated as the second amino acid in the peptide[30]. In the known MBA WAP-8294A2, the hydroxyl group on the asparagine at the same position (D-OHAsn-2) is not required for antibacterial activity[11,31]. Therefore, the unmodified linear lipopeptide produced by each NRPS system was predicted to be the direct precursor to the final biologically active cyclic lipopeptide produced by each BGC. Since we felt that the product of each potential MBA BGC could be bioinformatically predicted with confidence, and in each case the linear NRPS-derived peptide appeared to be the direct precursor to the final cyclic peptide product, we thought that total chemical synthesis was likely to be the most straightforward means of accessing the bioactive metabolites encoded by the six potential MBA BGCs we identified.

Two key structural features that cannot be bioinformatically predicted with high confidence from the primary sequence of these BGCs are the exact lipid used to initiate biosynthesis and the mode of peptide cyclization. Distinct from many other cyclic lipodepsipeptides[32], changes in the fatty acid tail of known MBAs did not result in substantial differences in antibacterial activity[10,33]. (*R*)-3-hydroxy-octanoic acid, which is found in two of the three known MBAs (WAP-8294A and lysocins) was used in the synthesis of all predicted BGC products[10,33]. The *N*-acylated linear peptide corresponding to the bioinformatically predicted product of each potential MBA BGC was generated by fluorenylmethyloxycarbonyl protecting group (Fmoc)-based solid-phase peptide synthesis (SPPS). (*R*)-3-hydroxy-octanoic acid-derivatized linear peptides can either be cyclized through the β-hydroxyl of the fatty acid (cyclized through fatty acid (cFA)) or through a nucleophilic amino acid side chain (cyclized through side chain (cSC)) (Fig. 2). When no nucleophilic side chain was present in the peptide, only a cFA derivative was produced from the linear peptide (MBA1, MBA3 and MBA4). However, when the first amino acid was predicted to contain a nucleophilic side chain (that is, a serine or threonine), both cFA and cSC analogues were synthesized from the linear peptide (MBA2, MBA5 and MBA6). In the case of MBA5, which was predicted to contain amino acids with a nucleophilic side chain at the first two positions, two distinct cSC peptides were generated. In total, we generated ten syn-BNPs for bioactivity screening (Fig. 2). For use as controls, we synthesized (*R*)-3-hydroxy-octanoic acid analogues of the three known MBAs (Extended Data Fig. 4a, WAP-SA1, lysocin and WBP-A2).

Antimicrobial spectrum. MBAs are expected to have Gram-positive antibacterial activity because MK plays an important role in the electron transport system of Gram-positive bacteria[6,7]. We initially tested each syn-BNP against three Gram-positive bacteria (*Bacillus subtilis, S. aureus* and *Staphylococcus epidermidis*) to determine which had antibacterial activity (Fig. 2 and Supplementary Table 2). In all cases where we synthesized differentially cyclized structures (MBA2, MBA5 and MBA6) from a single linear peptide, cyclization through the serine or threonine at the first position showed the most potent antibiotic activity. We believe that these more active structures are the likely products of these BGCs; therefore, they were used in all of the subsequent studies. We synthesized 'unnaturally cyclized' versions of two known MBA analogues (lysocin and WAP-SA1); in both cases, these structures showed a comparable decrease or loss of antibacterial activity (Extended Data Fig. 4). In the case of syn-BNPs predicted from the MBA1, MBA3 and MBA4 BGCs, the only possible cyclization pattern was through the 3-hydroxy of the *N*-terminal fatty acid. All three compounds showed good Gram-positive antibacterial activity (Fig. 2). Ultimately, this analysis resulted in six new and structurally diverse cyclic lipopeptide antibiotics (MBA1–6) that are only linked by a shared GXLXXXW motif, which we hypothesized was associated with MK binding as a mode of action.

All six syn-BNPs were broadly active against Gram-positive bacteria (Supplementary Table 2). Against a panel of *S. aureus* strains resistant to diverse clinically relevant antibiotics, the MBA MICs ranged from 0.25 to 8 μg ml^{-1}. MBA3, which was predicted from a metagenome-derived BGC, was the most potent antibiotic with MICs against the *S. aureus* strain ranging from 0.25 to 2 μg ml^{-1}. A subset of these syn-BNPs, MBA3 in particular, were active against *M. tuberculosis* (Supplementary Table 2). Lysocin E is active against *Mycobacterium smegmatis*[34], but to the best of our knowledge, no MBA has been reported to have anti-*M. tuberculosis* activity. To begin to explore the relevance of MK to the activity of these antibiotics in more detail, we determined their MICs against a collection of *Enterococcus* and *Streptococcus* spp. that either natively produce or do not produce MK. The syn-BNPs were active against the two MK-producing strains we tested (*Enterococcus casseliflavus* and *Streptococcus cremoris*) but were inactive against all six MK-deficient *Enterococcus* and *Streptococcus* strains we tested (Supplementary Table 3)[35,36]. In contrast, all four known Lipid II binders (lysobactin, nisin, ramoplanin and vancomycin) we tested were active against all eight *Enterococcus* or *Streptococcus* strains (Supplementary Table 3)[37,38]. In broader bioactivity screening, no syn-BNP showed activity against any wild-type (WT) Gram-negative bacteria or fungi that were tested (Supplementary Table 2). Against outer membrane-permeabilized *E. coli* (*E. coli* BAS849), syn-BNP MICs ranged from 8 to 32 μg ml^{-1} (Supplementary Table 4)[39]. In Gram-negative bacteria, MK is produced under both aerobic and anaerobic growth conditions; however, it is produced at higher levels in anaerobic growth conditions because of its key role in respiration[40,41]. Under anaerobic growth conditions, syn-BNP MICs against *E. coli* BAS849 were two- to four-fold lower than under aerobic conditions (Supplementary Table 4). The absence of activity against *E. coli*, and likely other Gram-negative bacteria, appears to be due to the inability of syn-BNPs to cross the outer membrane. Collectively, this spectrum of activity data provided our first experimental evidence that all six new antibiotics were likely MBAs.

Mode-of-action studies. Known MBAs cause rapid cell death due to membrane lysis[10,11]. Therefore, we tested each active syn-BNP for the ability to lyse *S. aureus*. As shown in Fig. 3a, when added to *S. aureus* cultures each syn-BNP antibiotic caused a rapid decrease in the number of viable cells. Membrane depolarization and cell lytic activities by all six syn-BNPs were confirmed using 3,3′-dipropylthiadicarboncyanine iodide ($DiSC_3(5)$) and SYTOX fluorescence assays, respectively (Fig. 3b and Extended Data Fig. 5a).

The relevance of MK to the antibacterial activity of each syn-BNP was examined in four ways: (1) we assayed for the ability of MK to suppress antibiotic activity; (2) we tested for antibiotic activity against MK biosynthesis knockout strains; (3) we raised resistant mutants to each antibiotic; and (4) we accessed the binding of each active syn-BNP to MK directly using isothermal titration calorimetry (ITC). When MK was added to the assay medium, the minimum inhibitory concentration (MIC) of each syn-BNP against *S. aureus* increased in a dose-dependent manner. The related structure, ubiquinone (UQ), had no effect on the antibiotic activity of any syn-BNP (Fig. 3c). We used two different *S. aureus* MK biosynthesis knockout strains (Δ*menA* and Δ*menB)* to test whether antibiotic activity was dependent on native production of MK. In Gram-positive bacteria, MK is used as an electron donor in respiration. Although in the absence of MK *S. aureus* cannot respire, they can survive by generating ATP from substrate phosphorylation. Both the *menA* and *menB* deletion strains are viable but they have small colony variant (SCV) phenotypes because they can only generate ATP from substrate phosphorylation[42]. All six syn-BNPs were inactive (MIC > 64 μg ml^{-1}) against both *S. aureus* strains (Fig. 3d). Furthermore, we selected *S. aureus* (USA300) mutants that could grow on 4× the MIC of each antibiotic. At 4× each syn-BNP's MIC resistant mutants appeared at a frequency of 0.7–5 × 10^{-6} for all syn-BNPs (Extended Data Fig. 5b). For each antibiotic, we sequenced multiple representative MK-resistant mutants. In all cases, the resistant strains contained a point mutation in an MK biosynthesis gene and in almost all cases no other mutations were detected in the genome (Supplementary Table 5). Finally, using ITC, an exothermic response was detected when liposomes (1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC):1,2-dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol (DOPG) = 1:1, mol:mol) containing MK were added into each syn-BNP. The calculated dissociation constants between MK and these antibiotics ranged from 0.09 to 0.30 μM (Extended Data Fig. 6). In contrast, no syn-BNP induced an exothermic reaction when

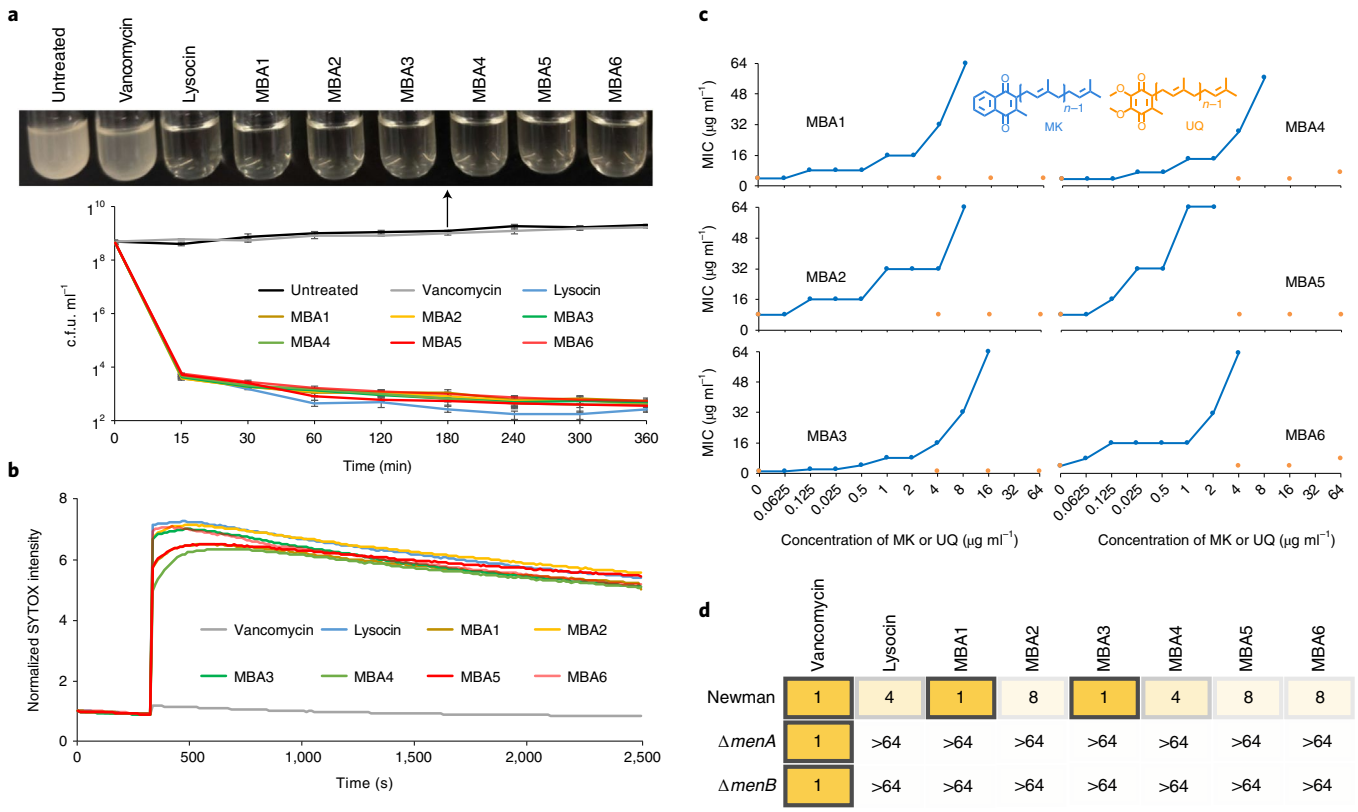


Fig. 3 | Bactericidal effects and mode-of-action analysis of six new MBAs. a, Bactericidal activity of MBA1–6 against *S. aureus* USA300. Cultures were incubated with each antibiotic at 2× its MIC. The number of viable cells was counted ($n=3$). Cultures were plated at defined times points to determine c.f.u. per ml. **b**, The effect of MBA1–6 on *S. aureus* membrane lysis was determined using a SYTOX fluorescence assay. **c**, The *S. aureus* antibacterial activity of MBA1–6 was determined in the presence of different concentrations of MK (blue) or UQ (orange) ($n=2$). **d**, The MICs (μg ml^{-1}) of MBA1–6 against *S. aureus* mutants deficient in MK biosynthesis (Δ*menA* or Δ*menB*) ($n=2$). Lysocin and vancomycin were included as positive (MK-binding) and negative (non-MK-binding) antibiotic controls, respectively. The different shades of yellow represent MBA potency, with darker colour indicating higher activity.

UQ instead of MK was included in the liposomes (Supplementary Fig. 3). We also tested the interaction between MK and syn-BNP derivatives with cyclization modes that showed reduced antibacterial activity (Fig. 2). All of these showed higher dissociation constants than their more active counterparts. For example, the calculated dissociation constant for the cFA version of MBA6, MBA6-cFA, was 45.5 μM (Extended Data Fig. 7a). This was 175-fold higher than the dissociation constant observed for the interaction between MK and the active cSC MBA6. The strong correlation we observed between dissociation constant and antibiotic potency suggests a direct link between the binding of these antibiotics to MK and cell killing (Extended Data Fig. 7b). While the GXLXXXW motif containing the natural product lysocin E has been reported to bind Lipid II[13], in the ITC experiments with liposomes containing Lipid II we did not observe an antibiotic-specific binding between syn-BNP MBAs and Lipid II (Supplementary Discussion and Supplementary Figs. 4 and 5). Collectively, these data provided multiple distinct lines of evidence that the GXLXXXW motif-containing syn-BNPs specifically bound MK and their antibacterial activity was dependent on MK.

Anti-*M. tuberculosis* activity and mode-of-action analysis for MBAs. Tuberculosis is one of the deadliest infectious diseases in the world. Anti-*M. tuberculosis* agents with new modes of action are urgently needed due to the rapid emergence of MDR and extensively drug-resistant *M. tuberculosis* mutants[43,44]. Although enzymes in the MK biosynthesis pathway have been explored as potential anti-*M. tuberculosis* targets[45–47], to the best of our knowledge MK binding has not been tested as a mode of action for an anti-*M. tuberculosis* agent. We assayed MBAs 1 through 6 against a panel of *M. tuberculosis* strains that included WT H37Rv, two mutants that can be studied using biosafety level 2 containment (mc^2 6206 and mc^2 7901) and four MDR strains (800, 4557, 10571 and 116) (Fig. 4a). All MBAs, with the exception of MBA5 and MBA6, were active against this panel of *M. tuberculosis* strains (MIC ≤ 10 μg ml^{-1}). MBA3 was the most potent anti-*M. tuberculosis* compound among MBAs, with an MIC as low as 0.078 μg ml^{-1} against MDR *M. tuberculosis*.

We sought to confirm the mode of action of this class of antibiotics against *M. tuberculosis*. As seen with *S. aureus*, when MK was added to the assay medium the MIC of MBAs against *M. tuberculosis* increased in a dose-dependent manner (Fig. 4b). UQ, which is structurally related and involved in electron transport in mammalian mitochondria, had no effect on the anti-*M. tuberculosis* activity of MBAs (Fig. 4b). As seen with *S. aureus*, MBAs caused an increase in $DiSC_3(5)$ fluorescence indicating that they also induced membrane depolarization in *M. tuberculosis* (Fig. 4c). Taken together, these data suggest that in *M. tuberculosis*, the antibacterial activity of MBAs is MK-dependent membrane disruption.

Macrophages play a central role in recognizing and destroying invading pathogens. *M. tuberculosis*, however, is capable of replicating and surviving in macrophages[48]. In vitro macrophage infection models, which provide a window into the interaction of host and pathogen, have proved to be good initial predictors of drug efficacy against *M. tuberculosis*. We examined the anti-*M. tuberculosis* activity of MBAs in a murine macrophage model. In this assay, J774A.1 mouse macrophages infected with *M. tuberculosis* harbour-

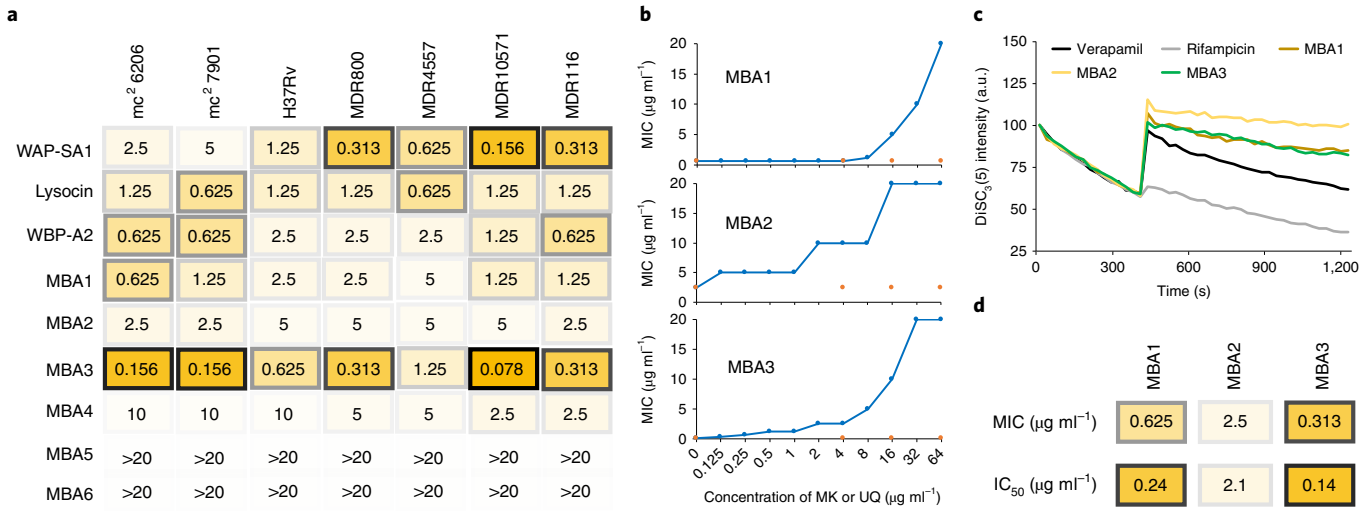


Fig. 4 | Anti-*M. tuberculosis* activity and mode-of-action analysis of new MBAs. a, MIC (μg ml^{-1}) values of all known and new MBAs against a panel of *M. tuberculosis* strains, including 2 biosafety level 2 strains (mc^2 6206 and mc^2 7901), a WT strain (H37Rv) and 4 MDR clinical isolates ($n=2$). Of the four MDR strains, two are resistant to rifampicin (800 and 4557), one is resistant to ethambutol/isoniazid/rifampicin/streptomycin (10571) and the other is resistant to ethambutol/isoniazid/para-aminosalicylic acid (116). **b**, Effects of MK (blue) or UQ (orange) on the antibacterial activity of MBA1–3 against *M. tuberculosis* mc^2 6206 ($n=2$). **c**, The ability of *M. tuberculosis*-active MBAs to permeabilize the *M. tuberculosis* membrane was examined using a $DiSC_3(5)$ fluorescence assay. Verapamil and rifampicin were used as positive and negative depolarization controls, respectively. a.u., arbitrary units. **d**, Activity of the three *M. tuberculosis*-active MBAs against *M. tuberculosis* mc^2 6206/mLux (MIC, $n=2$) and *M. tuberculosis* mc^2 6206/mLux in a macrophage infection assay (IC_{50}, $n=2$). The different shades of yellow represent MBA potency, with the darker colour indicating higher activity.

ing the mLux plasmid were treated with each *M. tuberculosis*-active MBA; then, residual bacterial cell viability inside the macrophages was determined by luminescence measurements. The three *M. tuberculosis*-active MBAs we tested inhibited *M. tuberculosis* growth in macrophages with a half-maximal inhibitory concentration (IC_{50}) ranging from 0.14 to 2.1 μg ml^{-1} (Fig. 4d). As seen in other assays, MBA3 was the most potent MBA in this assay. Considering the suggested correlation between activity in macrophages and potential activity in vivo[49], in the future it will be interesting to explore the in vivo anti-*M. tuberculosis* activity of MBA3.

Two new MBA structural families. Although the six MBAs we identified share a conserved GXLXXXW sequence that we propose is important for MK binding, they exhibit substantial differences in overall peptide sequence as well as different modes of cyclization and antimicrobial potency. To visualize the structural relationships among these antibiotics more easily, linear MBA peptide sequences were aligned and a phylogenetic tree was generated (Fig. 5a). This tree contains three distinct clades, one of which consists of known (lysocin E and WBP-29479A1) and new MBAs (MBA1 and MBA2), while the other two clades only contain the MBAs identified in this study (Fig. 5a). MBA1 and MBA2 are closely related to WBP-29479A1 and lysocin E, respectively. They differ from these known structures by only one amino acid in the case of MBA1 and three amino acids in the case of MBA2 (Fig. 5b). The four remaining antibiotics identified in this study make up two new MBA structural families. The first new family consists of MBA3 and MBA4 (Fig. 5c). Both structures were cyclized between the hydroxyl on (*R*)-3-hydroxy-octanoic acid and C-terminal GABA to form an undeca-lipodepsipeptide. Beyond the conserved minimal MK binding sequence, these two structures share a distinct L-Pro-5 and D-Tyr-9. The L-Pro-5 replaces the D-*N*-Me-aromatic amino acid that appears in all other known or new MBAs between G and L in the conserved GXLXXXW motif. Proline, like *N*-methylated amino acids, can introduce discrete conformations into cyclic peptides[50], suggesting that these two types of amino acids may play similar roles in MBAs. If we include this observation in our definition of a minimal MK binding motif, it would restrict the first X in the motif to being either an *N*-methyl-aromatic amino acid (*N*MeAAA) or proline (G(*N*MeAAA/P)LXXXW).

The second new family of MBA antibiotics consists of MBA5 and MBA6 (Fig. 5d). In addition to the conserved GXLXXXW motif, these two peptides share a serine, serine, asparagine, threonine and phenylalanine at positions 1, 3, 8, 9 and 11, respectively. Both peptides are cyclized using the serine at the first position in the linear peptide. However, they differ by the size of the resulting macrocycle. MBA6 contains 12 amino acids, while MBA5 contains 14 amino acids, making it the largest MBA characterized to date. Unlike other MBAs, MBA5 and MBA6 do not contain any positively charged amino acids. Positively charged residues found in known MBAs and other classes of lipopeptide antibiotics have been proposed to interact with the anionic polar head groups of the bacterial membrane and help induce rapid bacteriolysis[51,52]. Therefore, the absence of cationic residues in MBA5 and MBA6 is a key distinguishing feature of this new subclass of MBAs.

To explore the importance of the individual residues of the proposed minimal MK-binding motif, we synthesized analogues of the most potent syn-BNP MBA, MBA3, where the three conserved amino acids (Gly-4, L-Leu-6 and D-Trp-10) were individually replaced with L-Ala or D-Ala. The resulting three analogues with one amino acid change each were assayed for both antibacterial activity and MK binding (Extended Data Fig. 8). Changing L-Leu-6 to L-Ala decreased the antibiotic activity of MBA3 by 16–32-fold and its MK-binding affinity by 365-fold, while changing Gly-4 to L-Ala or D-Trp-10 to D-Ala completely abrogated its antibacterial activity (MIC > 64 μg ml^{-1}) as well as its ability to interact with MK. Therefore, all three conserved residues of the proposed minimal MK-binding motif are critical for both potent antibacterial activity and high-affinity MK binding. How the G(*N*MeAAA/P)LXXXW sequence that is conserved across MBAs interacts with MK is to be determined. However, it is likely that the indole of tryptophan interacts with the quinone from MK[53] and that the *N*-methyl aro-

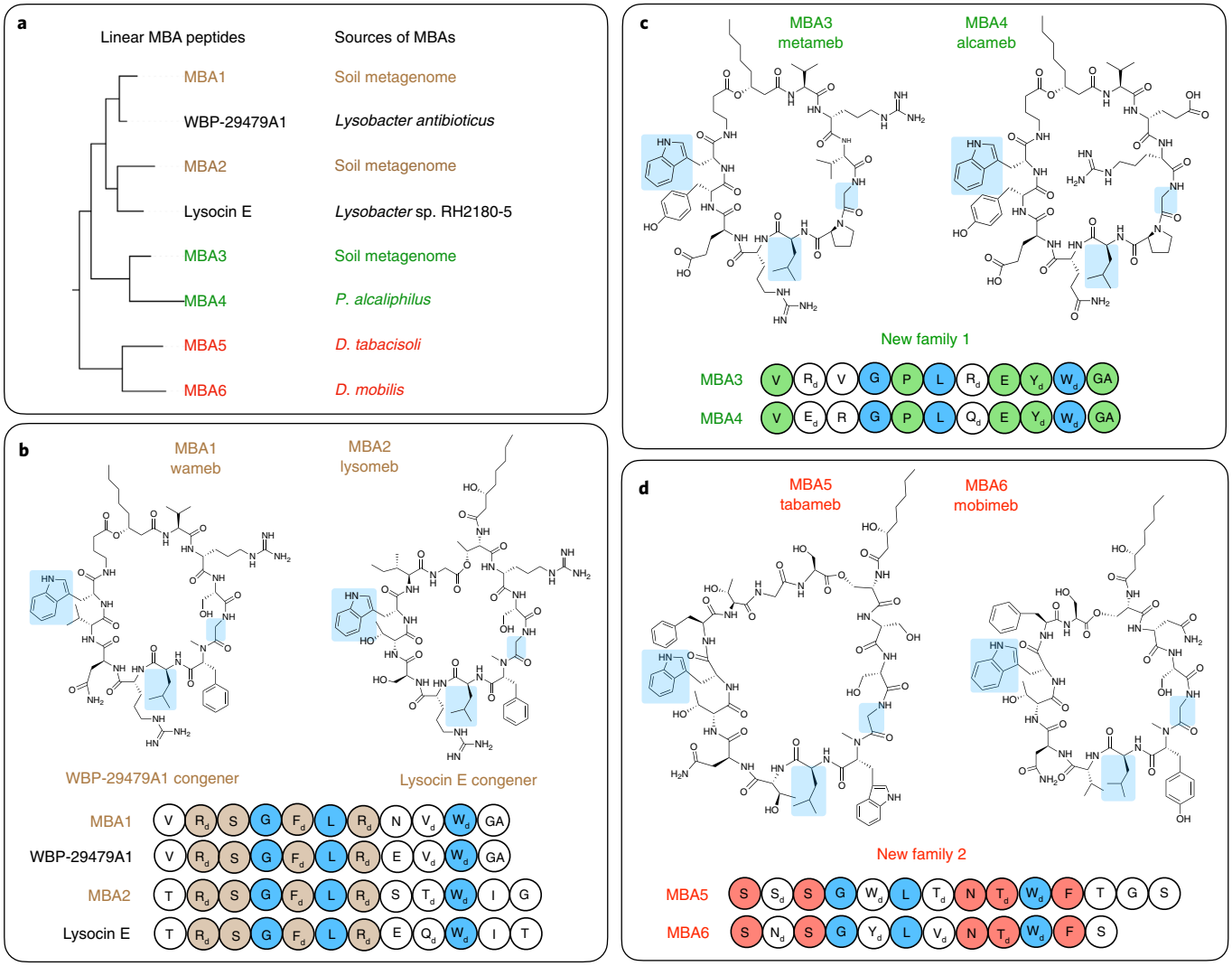


Fig. 5 | Six new MBAs grouped by structural family. a, Phylogenetic tree of linear MBA peptide sequences. The branches on the tree are labelled with the name of the MBA and the source of its BGC. **b**–**d**, In this study, we identified congeners of two known MBAs (**b**) and two new MBA structural families (new family 1 (**c**) and new family 2 (**d**)). All MBAs share the conserved GXLXXXW motif (blue) that is predicted to be the minimal sequence associated with MK binding as a mode of action. The conserved residues within each MBA family are highlighted. In accordance with the long-standing tradition of giving bioactive natural products trivial names, we have given these six syn-BNPs the following names: wameb (MBA1, WBP-29479A1-like menaquinone-binding antibiotic), lysomeb (MBA2, lysocin E-like menaquinone-binding antibiotic), metameb (MBA3, metagenome menaquinone-binding antibiotic), alcameb (MBA4, *P. alcaliphilus* menaquinone-binding antibiotic), tabameb (MBA5, *D. tabacisoli* menaquinone-binding antibiotic) and mobimeb (MBA6, *D. mobilis* menaquinone-binding antibiotic). The blue residues represent building blocks that are conserved across all MBAs. The brown, green and red residues represent building blockings that are conserved across the known MBA family (**b**), the first new MBA family (**c**) or the second new MBA family (**d**), respectively.

matic amino acid and the proline induce similar cyclic peptide conformations that help create an MK binding pocket. The hydrophobic glycine and leucine residues and the hydrophobic lipid tail seen in all MBAs are likely important for interacting with either the hydrophobic polyprenyl tail of MK or the lipid biolayer in bacterial membranes.

Additional sources of MBAs. Known MBAs are produced by the genus *Lysobacter*[12,30,54]. Interestingly, both new MBA families were inspired by BGCs that are found in bacteria from different taxa (Fig. 5a and Supplementary Fig. 6). While the BGC for MBA3 was cloned from a soil metagenome and therefore its source is unknown, the BGC for MBA4 is found in the genome of *P. alcaliphilus*. The BGCs for MBA5 and MBA6 were found in the genomes of *D. tabacisoli* and *D. mobilis*, respectively (Fig. 5a and Supplementary Fig. 6). *Paracoccus* and *Dyella* are genera of Proteobacteria that have not traditionally been part of bacterial natural product discovery programmes[55,56]. *P. alcaliphilus* is an alkaliphilic, facultative, methanol-utilizing bacterium, while the genus *Dyella* was only first described in 2005 (refs. [57,58]). As outlined in the metagenomic library screening studies described above, the L-Leu-6 A-domain is highly conserved across MBA BGCs. In fact, when eDNA A-domain NPSTs were compared to the known MBA BGC L-Leu-6 A-domain, all NPSTs that returned e-values $\leq 10^{-45}$ were found to arise from an MBA BGC (Extended Data Fig. 1). To explore MBA BGC diversity in the environment more extensively, we screened an archived collection of A-domain NPSTs generated from diverse soil metagenomes for MBA BGC-like L-Leu-6 A-domain sequences[22]. One in every 25 soils we screened (80 out of 2,000) contained a unique MBA L-Leu-6-like A-domain NPST (that is, e-value $\leq 10^{-45}$). Therefore,

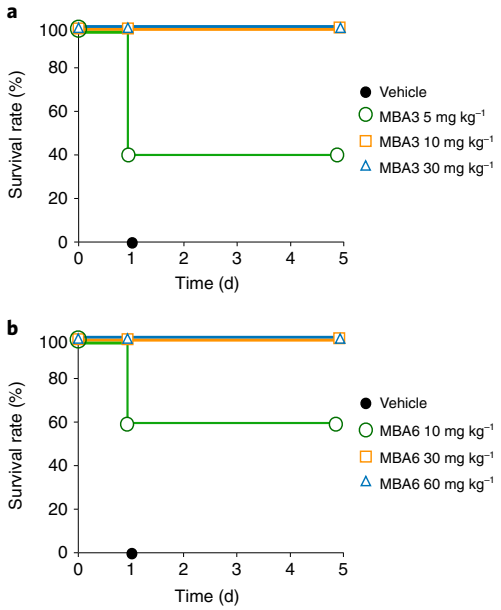


Fig. 6 | MBA3 and MBA6 are effective against *S. aureus* infections in mice. a,**b**, Either MBA3 (**a**) or MBA6 (**b**) was injected subcutaneously 1 h after intraperitoneal administration of *S. aureus* COL into mice ($n = 5$); 30% Solutol was used as the vehicle.

our studies not only identified additional cultured sources of structurally diverse MBAs, but also suggest that there is a high number of as yet uncharacterized MBA BGCs in soil microbiomes.

Activity of MBA3 and MBA6 in a murine peritonitis-sepsis model. Using a mouse peritonitis-sepsis model we examined the in vivo efficacy of MBA3 and MBA6, which were the most potent analogues of the two new MBA families we identified in this study. As shown in Fig. 6, treatment of a methicillin-resistant *S. aureus* COL strain infection with either MBA3 (5, 10 and 30 mg kg^{-1}) or MBA6 (10, 30 and 60 mg kg^{-1}) dramatically decreased the mortality of infected mice compared to treatment with vehicle alone (30% Solutol). A minimal dose of 10 and 30 mg kg^{-1} for MBA3 and MBA6, respectively, was required for 100% survival. Collectively, our results suggest that MBA3 and MBA6 could be valuable candidates for the development of therapeutics for treating MDR *S. aureus* infections.

Discussion

In this study, we combined sequence-based metagenomic mining with a pattern search of bioinformatically generated natural product structures to identify BGCs that we predicted would encode MBAs. Synthesis of the structures bioinformatically predicted to arise from these BGCs produced six MK-binding cyclic lipodepsipeptides with diverse structures, anti-*M. tuberculosis* activity and potent in vivo anti-*S. aureus* activity. Although MBAs exhibited substantial differences in peptide sequence and different cyclization modes, they shared a conserved GXLXXXW sequence that we propose is a minimal MK binding motif. Notably, the most potent members of the two new MBA structural families we identified, MBA3 (metameb) and MBA6 (mobimeb), proved effective at treating methicillin-resistant *S. aureus* infections in a murine peritonitis-sepsis model. Our discovery of new MBA structural families suggests that MBAs are a diverse and still underexplored class of naturally occurring antibiotics. Both the new MBA structures reported in this study and the MBA search tools developed in this study should prove useful in ultimately identifying a member of this mechanistically interesting class of antibiotics that can be successfully brought through therapeutic development into clinical use.

MBA resistance can arise from mutations in MK or haem biosynthesis[10,13]. In both cases, these mutants show an SCV phenotype[59]. In addition, both MK- and haem-deficient mutants have shown reduced virulence in animal models[10,59]. When growth-compensatory mutants were directly selected using a *menB* point mutant background, all growth-compensatory mutants showed increased MK production[60]. When growth-compensatory mutants were selected by serial passage using either MK- or haem-deficient strains, approximately 96% of the mutants showed restored MK levels[61]. These studies suggest that while growth-compensatory mutations will undoubtedly arise, most of the mutants are likely to restore susceptibility to MBAs. When the mutation frequency is taken together with decreased virulence by these SCVs and the restoration of MK production in most growth-compensatory mutants, we believe that MBAs are likely to show a lower clinically relevant resistance rate than suggested by the resistance frequency seen in the laboratory. With that said, this will clearly be something that will need to be explored in more detail in the clinical development process.

Although we have previously used metagenomic mining methods to identify BGCs of interest[22,23], we believe that identifying molecules with a desired mode of action by searching for a substructure among a large dataset of bioinformatically predicted natural products has not been described previously. The success of this approach for identifying new classes of MBAs suggests that bioinformatic structure prediction algorithms have developed to the extent that direct structure, or substructure, searches of collections of bioinformatically predicted natural product structures now represent an alternative generalizable approach to screen sequenced BGCs for the potential production of bioactive small molecules.

Methods

Chemical reagents, consumables and instruments. All reagents were purchased from commercial sources and used without further purification. Preloaded 2-chlorotrityl resin for peptide synthesis was purchased from Matrix Innovation. Reagents for SPPS, dichloromethane (DCM), *N*,*N*-diisopropylethylamine (DIPEA), 4-dimethylaminopyridine (DMAP), *N*,*N*-dimethylformamide (DMF), *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HATU), hexafluoroisopropanol (HFIP), (7-azabenzotriazol-1-yloxy) tripyrrolidinophosphonium hexafluorophosphate (PyAOP) and trifluoroacetic acid (TFA) were purchased from P3 BioSystems. Standard *N*-Fmoc amino acid building blocks were purchased from P3 BioSystems and Chem-Impex International. (*R*)-3-hydroxy-octanic acid and Fmoc-*N*-Me-D-Phe-OH were purchased from Enamine. Fmoc-*N*-Me-D-Trp(Boc)-OH was purchased from Alabiochem. Fmoc-*N*-Me-D-Tyr(tBu)-OH was purchased from 1Click Chemistry. Fmoc-GABA-OH was purchased from Sigma-Aldrich. All solvents used for chromatography were high-performance liquid chromatography (HPLC) grade or higher. Thiazolyl blue tetrazolium bromide (MTT) and type II mucin from porcine stomach were purchased from Sigma-Aldrich. The fluorescent dyes SYTOX Green and $DiSC_3(5)$ were purchased from Thermo Fisher Scientific and the assay results were recorded using a Tecan Infinite M Nano$^+$ plate reader. DOPC and DOPG were purchased from Avanti Polar Lipids. Menaquinone-4 (MK4) and ubiquinone-10 (UQ10) were purchased from Sigma-Aldrich and Gram-positive Lys-Lipid II was purchased from Antimicrobial Discovery Solutions Ltd.

For all liquid chromatography, solvent A = H_2O (0.1% v/v formic acid) and solvent B = CH_3CN (0.1% v/v formic acid). Ultra-performance liquid chromatography (UPLC)–low-resolution mass spectrometry (LRMS) data were acquired on a Waters ACQUITY system equipped with quadrupole Dalton and photodiode-array detection detectors, a Phenomenex Synergi Fusion-RP 80 Å column (2.0 mm × 50 mm, 4 µm) and controlled by the Waters MassLynx v.4.1 software. The following chromatographic conditions were used for UPLC–LRMS: 5% B from 0 to 0.9 min; 5–95% B from 0.9 to 4.5 min; 95% B from 4.5 to 5.0 min; 95–5% B from 5.0 to 5.4 min; and 5% B from 5.4 to 6 min (flow rate of 0.6 ml min^{-1} and 10 µl injection volume). HPLC–high-resolution mass spectrometry (HRMS) data were acquired on a SCIEX ExionLC HPLC coupled to an X500R QTOF mass spectrometer, equipped with a Phenomenex Kinetex PS C18 100 Å column (2.1 mm × 50 mm, 2.6 µm) and controlled by the SCIEX OS v.2.1 software. The following chromatographic conditions were used for UPLC–HRMS: 5% B from 0 to 1.0 min; 5–95% B from 1.0 to 10.0 min; 95% B from 10.0 to 12.5 min; 95–5% B from 12.5 to 13.5 min; and 5% B from 13.5 to 17.0 min (flow rate of 0.4 ml min^{-1}

and 1 µl injection volume). Peptide purification was performed using an Agilent 1200 Series HPLC with ultraviolet detection and equipped with an XBridge Prep C18 130 Å column (10 × 150 mm, 5 µm). ^{1}H and ^{13}C nuclear magnetic resonance (NMR) spectra were acquired on a Bruker Avance DMX 600 MHz spectrometer equipped with cryogenic probes (The Rockefeller University). All spectra were recorded at 25 or 50 °C in dimethyl sulfoxide (DMSO)-d_6. Chemical shift values are reported in parts per million (ppm) and referenced to residual solvent signals: 2.50 ppm (^{1}H) and 39.52 ppm (^{13}C).

Identification of metagenomic MBA BGCs. Previously archived soil eDNA cosmid libraries were probed to recover BGCs predicted to encode new MBAs. Construction, PCR screening with barcoded A-domain degenerate primers, amplicon sequencing and read processing for these cosmid libraries have been described in detail previously[22]. Using A-domain sequences from the three known MBAs as references, the eSNaPD v.2 software was used to identify similar sequences among A-domain amplicon sequences generated from archived metagenomic libraries. The library well locations of hits found in this analysis were identified using the barcode parsing functionality of the eSNaPD v.2 software. Clones associated with select eDNA A-domain hits were then recovered from the appropriate library wells using a previously described dilution PCR strategy[22]. Recovered cosmids were sequenced using a MiSeq Reagent Nano Kit V2 on a MiSeq sequencer (Illumina) and the resulting reads were assembled into contigs using Newbler v.2.6 (Roche). Assembled complete and partial BGCs were analysed using antiSMASH v.5.1.2 and our manual in-house NRP predictor to predict the substrate specificity of each A-domain. In this analysis, building blocks were predicted by comparing the Stachelhaus code of predicted A-domains to that of A-domains from known NRPs.

Screening 2,000 soils to explore more potential MBAs. eDNA was extracted from approximately 2,000 ecologically and geographically diverse soil samples; NPSTs of soil metagenomes were generated using a previously established pipeline[22]. These NPSTs were then searched using the eSNaPD pipeline against the manually curated L-Leu-6 sequences from the three known and six new MBA BGCs. A-domain amplicons that matched MBA L-Leu-6 at an e-value $\leq 10^{-45}$ were considered as hits. A multiple sequence alignment of all qualifying hit sequences was generated using the multiple sequence comparison by using multiple sequence alignment algorithm MUSCLE in MacVector v.17.0.5; the resulting alignment file was used to generate a maximum likelihood tree in MacVector v.17.0.5.

Construction of an in-house predicted NRP database. GenBank files for 38,933 NRP BGCs representing 10,858 complete bacterial genome assemblies were retrieved from the antiSMASH database. Our predicted NRP (p-NRP) database was constructed from the BGCs by synthesizing data from five A-domain prediction resources: antiSMASH database[62]; Norine amino acid database[63]; and A-domain substrate predictions from MIBiG v.2[64], SANDPUMA[65] and our own NRP BGC analyses (Supplementary Fig. 1). We analysed all BGCs using antiSMASH v.5.1.2 (ref. [28]) and the parallel job execution tool GNU Parallel v.3[66]. After removing duplicate entries, we obtained a total of 36,957 NRPS antiSMASH v.5 regions. We downloaded the entirety of the 'monomer' dataset from Norine, which represents known A-domain substrates and other known amino acids, and we used it as the basis for our own A-domain 'substrates' table. We extended this table with the addition of auxiliary columns, such as the substrate's charge, as well as normalizing the shorthand names of some substrates to follow convention in the NRP BGC community more closely. Lastly, we incorporated predictions for Stachelhaus codes using MIBiG, which stores curated information about well-characterized BGCs, and SANDPUMA, which maintains a table of known A-domain-to-substrate specificities. The A-domain substrate specificity codes obtained from our manual curation of characterized NRPs were also included. For any substrates not represented in the Norine database, a search was conducted in the PubChem database and the relevant data were imported[67], including the PubChem ID, to ensure uniqueness. Substrates that did not appear in PubChem were still included in the database; however, other uniquely identifying information, such as International Union of Pure and Applied Chemistry name and simplified molecular-input line-entry system string, were relied on to prevent a substance being erroneously represented multiple times. When no substrate specificity code was identified for an A-domain, A-domain substrate specificity was predicted using the NRPSpredictor2 'small cluster prediction'[68]. Finally, based on this collection of curated substrate specificity codes, a table of linear peptide sequence predictions for each NPRS BGC was generated. The starting point of a p-NRP sequence was determined either by the presence of a condensation starter domain or the presence of an A-domain with no immediately preceding condensation domain. The end of the peptide sequence was defined by the presence of a thioesterase domain. Ultimately, for each BGC the order of the monomers predicted in our A-domain analysis followed the order of A-domains predicted by antiSMASH for that BGC. To identify potential MBA BGCs, the resulting p-NRP database was searched for the full GXLXXXW motif and two partial motifs: GXL and LXXXW.

Peptide synthesis. *SPPS.* All peptides were synthesized using standard Fmoc-based SPPS methods on 2-chlorotrityl chloride resin using commercially available Fmoc-protected amino acids. Peptide synthesis started from the conserved leucine seen at the sixth position of each peptide. Preloaded leucine on 2-cholorotrityl resin (0.3 g, 0.552 mmol g^{-1}) was swollen in DCM for 20 min at room temperature then drained and washed with DMF (3 ml, 3×). Coupling of individual amino acids was carried out by using Fmoc-protected amino acids (2 equivalent relative to resin loading) mixed with HATU (2 equivalent) and DIPEA (2 equivalent) in DMF (5 ml). Coupling reactions were carried out for 1 h with occasional swirling then washed with DMF (3 ml, 3×). Fmoc deprotection was done using 20% piperidine in DMF (3 ml) for 7 min and repeated twice. The resin was washed with DMF (3 ml, 5×) and then coupled with a subsequent amino acid.

Ester bond formation. Ester bonds were formed either between the hydroxyl group on the *N*-terminal fatty acid or an amino acid-associated hydroxyl group and the *C*-terminal carboxyl group of the peptide. To carry out ester bond formation, the resin was mixed with amino acid (20 equivalent), DIPEA (40 equivalent), benzoyl chloride (20 equivalent) and DMAP (0.8 equivalent) in 10 ml of DCM and gently shaken for 72 h. After ester bond formation, the remaining amino acids were coupled as described above.

Peptide cyclization. Peptides were cleaved from the resin by treatment with 20% HFIP in DCM for 2 h. After air drying overnight, the cleaved linear peptide was cyclized without purification using PyAOP (8 equivalent) and DIPEA (30 equivalent) in DMF (50 ml). After 2 h, DCM (100 ml) was added and washed repeatedly with 1% formic acid in water (5 ml, 10×). The extracted cyclic peptide was air-dried overnight.

Final cleavage. Air-dried cyclic peptide was dissolved in 3 ml of cleavage cocktail (95% v/v TFA, 2.5% v/v triisopropylsilane and 2.5% v/v water) for 1.5 h. A cold mixture of diethyl ether:hexane (1:1) was then added and kept at −20 °C for 10 min to precipitate the peptide. Peptide pellets were collected by centrifuging (2,500*g*) for 5 min, redissolved in 5 ml of methanol and dried under vacuum overnight.

Peptide purification by HPLC. Crude cyclic lipopeptides were purified on a XBridge Prep C18 HPLC column using a dual solvent system (A/B: water/acetonitrile supplemented with 0.1% v/v formic acid). All peptides were eluted using a linear gradient from 30 to 50% gradient of B. Peptide purity was confirmed by UPLC. The identity of each peptide was confirmed by HRMS (Supplementary Fig. 7) and tandem mass spectrometry (Supplementary Figs. 8–13) analyses. ^{1}H and ^{13}C NMR spectra were recorded for each peptide (Supplementary Figs. 14–19). Pure peptides were dissolved in DMSO at 6.4 mg ml^{-1} for MIC measurements and mode-of-action studies.

Antimicrobial assays against Gram-positive and Gram-negative bacteria and yeast pathogens. All antimicrobial assays were run in 96-well microtitre plates using a broth microdilution method. Diluted overnight cultures were used in all assays. For yeast strains, overnight cultures were diluted 2,000-fold in yeast extract peptone dextrose broth. For *Enterococcus faecium* and *S. aureus*, overnight cultures were diluted 1,000- and 10,000-fold in lysogeny broth (LB), respectively. For *Streptococcus* strains, overnight cultures were diluted 5,000-fold in brain heart infusion broth. For other bacteria, overnight cultures were diluted 5,000-fold in LB. Then, 100 µl of each diluted culture was mixed with 100 µl of LB containing a syn-BNP at twofold serial dilutions across a 96-well microtitre plate row. The final concentration of each compound ranged from 64 to 0.25 µg ml^{-1}. Plates were incubated at 37 °C (bacteria) or 30 °C (yeast) for 16 h. The lowest concentration that inhibited visible microbial growth was recorded as the MIC. All MIC assays were done in duplicate ($n = 2$).

Antibacterial assay against *M. smegmatis* mc^2 155. *M. smegmatis* mc^2 155 was shaken in 7H9 broth (supplemented with 0.2% glucose, 0.2% glycerol and 0.05% tyloxapol) for 48 h at 37 °C. The culture was then diluted to an OD_{600} of 0.005 and 100 µl was added to 100 µl of 7H9 broth containing each syn-BNP at twofold serial dilutions across a 96-well plate row. The final concentration of each compound ranged from 64 to 0.25 µg ml^{-1}. Plates were incubated for 48 h at 37 °C and then 30 µl of alamarBlue Cell Viability Reagent (Thermo Fisher Scientific) was added. After an additional 24 h incubation, wells that remained blue by visual inspection were deemed to contain inhibitory concentrations of each antibiotic. All MIC assays were done in duplicate ($n = 2$).

Antibacterial assay against *M. tuberculosis*. *M. tuberculosis* mc^2 6206, mc^2 7901, WT H37Rv and 4 MDR strains (116, 800, 4557 and 10571) were passaged in 7H9 broth (supplemented with oleic acid-albumin-dextrose-catalase, 0.2% glycerol and 0.02% tyloxapol) at 37 °C to an OD_{600} of 0.5–0.7. The culture was then diluted to an OD_{600} of 0.005 and 100 µl of the diluted culture was distributed in 96-well plates. Then, 100 µl of 7H9 broth containing each syn-BNP at twofold serial dilutions across a 96-well plate row was added to give final test concentrations ranging from 20 to 0.039 µg ml^{-1}. Plates were then incubated at 37 °C and 5% CO_2. After incubation for 6 d, 30 µl of alamarBlue Cell Viability Reagent was added, cultures were incubated for another 24 h and absorbance was read at 570 nm and normalized to 600 nM. All MIC assays were done in duplicate ($n = 2$).

Murine macrophage infection. The activity of each MBA against intracellular *M. tuberculosis* was determined by infecting J774A.1 mouse macrophages (catalogue no. 91051511; Sigma-Aldrich) with *M. tuberculosis* mc^2 6206 harbouring the mLux plasmid (*M. tuberculosis* mc^2 6206/mLux). Macrophages were initially suspended at a concentration of 4–5 × 10^5 cells ml^{-1} in DMEM (Sigma-Aldrich) supplemented with FBS (10%) and L-glutamate (2 mM). Flat bottom 96-well white plates were seeded with 100 µl of the macrophage suspension and incubated overnight to allow cells to adhere to the plates. *M. tuberculosis* mc^2 6206/mLux was grown to mid-log phase (OD_{600} = 0.5–0.7). *M. tuberculosis* cultures were then spun down, washed with PBS and resuspended in DMEM containing 10% FBS, L-glutamate (2 mM), pantothenic acid (50 µg ml^{-1}) and leucine (50 µg ml^{-1}). Assay plates were inoculated with 100 µl of *M. tuberculosis* mc^2 6206/mLux at a multiplicity of infection of 1:10 and incubated for 4 h to allow *M. tuberculosis* to infect the macrophages. Adhered macrophages were then washed twice with 100 µl of PBS and then 100 µl of DMEM was added to each well. After a 1 h incubation plates were washed twice with 100 µl of PBS. Then, 100 µl of each MBA serial diluted in DMEM (from 64 to 0.0004 µg ml^{-1}) was added to the assay plates. Plates were then incubated at 37 °C for 72 h. Residual *M. tuberculosis* cell viability inside macrophages was determined by luminescence measurement on a Spark multimode microplate reader (Tecan). Dose–response curves were generated by non-linear regression in Prism 8 (GraphPad Software) and plotted as the logarithm of concentration versus normalized percentage cell viability. The antibiotic concentration that led to 50% cell viability (IC_{50}) was determined from the dose–response curves. Each treatment was carried out in duplicate (n = 2).

Membrane lysis assay. Membrane lysis assays were done in 384-well black microtitre plates. An overnight culture of *S. aureus* USA300 was collected by centrifugation and resuspended in PBS to give an OD_{600} of 0.5. SYTOX Green (5 mM, 1 µl) was added to the cell suspension (2.5 ml), which was then incubated in the dark at room temperature for 10 min. Fluorescence intensity of the mixture was recorded continually at 2 s intervals (excitation/emission 488/523 nm). When the signal stabilized the appropriate amount of each antibiotic (6.4 mg ml^{-1} DMSO stock solutions) to give 2× its MIC was added and immediately mixed by manual pipetting. Vancomycin and lysocin were used as negative and positive controls, respectively. Data were presented as the relative intensity with regard to the average fluorescence signal before the addition of the MBA. All assays were done in duplicate (n = 2). A representative fluorescence recording for each antibiotic is shown in Fig. 3b.

Membrane depolarization assay. Membrane depolarization assays were done in 384-well black microtitre plates. An overnight culture of *S. aureus* USA300 was collected by centrifugation and resuspended in PBS to give an OD_{600} of 0.5; 100 µl of this cell suspension and 20 µM of $DiSC_3(5)$ (50 µl) were added to 300 µl of PBS and then incubated in the dark at room temperature for 15 min. KCl (2 M, 50 µl) was added and the culture was incubated for another 15 min. Fluorescence intensity of the mixture was recorded continually at 2 s intervals (excitation/emission 643/675 nm). When the signal stabilized the appropriate amount of each MBA (6.4 mg ml^{-1} DMSO stock solutions) to give 2× its MIC was added and immediately mixed by manual pipetting. Vancomycin and lysocin were used as negative and positive controls, respectively. Data were presented as the relative intensity with regard to the average fluorescence signal before the addition of the MBA. All assays were done in duplicate (n = 2). A representative fluorescence recording for each antibiotic is shown in Extended Data Fig. 5a.

Antibiotic-resistant mutant selection. A single *S. aureus* USA300 colony was inoculated into LB and grown overnight at 37 °C. A portion of the overnight culture containing approximately 10^9 cells was diluted (1/10× or 1/40×) into LB containing each antibiotic at 4× its MIC. The resulting mixtures were distributed into microtitre plates at 5 µl per well. After incubating statically at 37 °C overnight, colonies that appeared were transferred into fresh LB agar plates. The MICs of 4–8 individual colonies were then determined using the microtitre dilution method described above. DNA was extracted from cultures of colonies that showed an elevated MIC relative to WT and the resulting DNA was sequenced using the MiSeq Reagent Kit V3 (catalogue no. MS-102-3003; Illumina). Single-nucleotide polymorphisms for each new MBA were identified using Snippy v.4.6.0 (https://github.com/tseemann/snippy) by mapping MiSeq reads to the reference genome of *S. aureus* USA300_FPR3757 (Reference Sequence database assembly accession no. GCF_000013465.1) (Extended Data Fig. 5b).

ITC. For MK and UQ binding, a 1:1 mixture of DOPC and DOPG containing either 1.25 mol per cent MK4 or UQ10 was dissolved in chloroform. A lipid film was generated by drying this material under a stream of nitrogen followed by 2 h of vacuum drying. The resulting film was hydrated using 10 mM of HEPES (pH 7.5) with 100 mM of NaCl to give a final total lipid concentration of 5 mM. Using an Avanti Mini Extruder, this suspension was passed through a 100 nm polycarbonate filter 10 times. For ITC, the sample cell was filled with 400 µl of 25 µM of MBA prepared in 10 mM of HEPES buffer. The syringe (150 µl) was loaded with a 5 mM lipid suspension with 1.25 mol per cent MK4 or UQ10. For Lipid II binding, 10 mM of DOPC containing 1 mol per cent Lipid II was dissolved in chloroform. The resulting film was hydrated using 50 mM of Tris (pH 7.5) with 100 mM of NaCl and passed through a 100 nm polycarbonate filter 10 times. For ITC, the sample cell was filled with 400 µl of 25 µM of an MBA prepared in 50 mM of Tris buffer. The syringe (150 µl) was loaded with a 10 mM DOPC suspension with 1 mol per cent Lipid II. Data were collected by using an Auto-iTC200 (Malvern Panalytical) and processed by the web-based application AFFINImeter for ITC using the 'one binding site' model.

MK extraction and identification in *S. aureus*. MK extraction was performed using a previously reported lysozyme-chloroform-methanol extraction method[69]. Cultures of the *menA* deletion mutant (Δ*menA*), the *hemB* transposon insertion mutant (tn:*hemB*) and *S. aureus* Newman and USA300 were grown overnight in LB. Cultures of the *menA* and *hemB* mutants were adjusted to the same OD_{600} as the *S. aureus* Newman and USA300 cultures. Cells from 20 ml of each density-normalized culture were collected by centrifugation. The resulting cell pellets were suspended in 50 ml of 10 mM of Tris-HCl buffer (pH 7.4) containing 5 mg of lysozyme and then incubated at 37 °C for 1 h. This mixture was then centrifuged for 10 min at 2,500*g* to collect the lysozyme-treated cells; 10 ml of chloroform/methanol (2:1, v/v) was added to the cell pellets to extract MK. This extraction process was repeated three times. The chloroform/methanol extracts were combined and evaporated under vacuum. The dried material was suspended in 50 µl of chloroform/methanol (2:1, v/v) for analysis by thin-layer chromatography (TLC). MK extracts were spotted on LuxPlate TLC silica gel 60 F_{254} (Merck Millipore) plates and plates were developed in a mixture of hexane and diethyl ether (85:15, v/v). MK was visualized by ultraviolet exposure and the plates were photographed. Finally, MK bands were collected from the TLC plates and eluted using isopropanol. Isopropanol-eluted MK was analysed by HPLC–HRMS and MK4 was used as a standard.

***M. tuberculosis* membrane depolarization assay.** *M. tuberculosis* membrane depolarization assays were done in a 384-well black microtitre plate. *M. tuberculosis* mc^2 6206 was grown to exponential phase (OD_{600} = 0.5–0.7), washed twice with HG buffer (5 mM of HEPES and 5 mM of glucose, pH 7.2) and resuspended in HG buffer (OD_{600} = 0.1). $DiSC_3(5)$ (4 µM) was added to the cell suspension and incubated in the dark at room temperature for 2 h. The fluorescence intensity of the mixture was recorded continually at 2 s intervals (excitation/emission 622/670 nm). After fluorescence intensity had stabilized, each antibiotic (6.4 mg ml^{-1} DMSO stock) was added to give a final concertation of 2× its *M. tuberculosis* MIC and immediately mixed by manual pipetting. Rifampicin and verapamil were used as the negative and positive controls, respectively. Data were presented as the relative intensity with regard to the average fluorescence signal before the addition of each MBA. All assays were done in duplicate (n = 2). A representative fluorescence recording for each antibiotic is shown in Fig. 4c.

Cytotoxicity assessment. HEK293 human cells (catalogue no. CRL-1573; ATCC) were grown at 37 °C in a 5% CO_2 atmosphere in DMEM supplemented with FCS (10% v/v), L-glutamate (2 mM), penicillin (10 U ml^{-1}) and streptomycin (10 U ml^{-1}). HEK293 cells were seeded into 96-well flat bottom microtitre plates (target density of 2,500 cells per well) and incubated in DMEM at 37 °C for 24 h to allow cells to adhere. The DMEM medium was then removed by aspiration and replaced with 100 µl of fresh DMEM medium containing each antibiotic at 10 serially diluted concentrations ranging from 32 to 0.0625 µg ml^{-1}. After 48 h at 37 °C, the DMEM medium was removed and 100 µl of an MTT solution (10 µl of 5 mg ml^{-1} MTT in PBS premixed with 100 µl of DMEM) was added to each well. After 4 h at 37 °C, the solution was aspirated from each well. Precipitated formazan crystals were dissolved by adding 100 µl of solubilization solution (40% DMF, 16% SDS and 2% glacial acetic acid in H_2O). OD_{570} readings were used to calculate relative growth (%) based on the positive (2 µM taxol) and negative (DMSO) controls. Dose–response curves were generated by non-linear regression in Prism v.8. Cytotoxicity assays were performed in triplicate (n = 3).

Mouse peritonitis-sepsis model. Female outbred Swiss Webster mice (6 weeks old) were used in all experiments. Mice were housed on a 12 h light cycle (7:00 on and 19:00 off) with a temperature of 21 °C and a relative humidity of 30%. *S. aureus* COL was grown in Mueller–Hinton broth at 37 °C overnight and diluted with 7% type II porcine stomach mucin supplemented with 0.2 mM of $FeNH_4$ citrate. Cultures were diluted to provide a challenge inoculum of approximately 5 × 10^8 colony-forming units (c.f.u.) in 0.2 ml; 0.2 ml of the challenge inoculum was administered via intraperitoneal injection. Thirty-five mice were randomly grouped into 5 per cohort (n = 5) and each cohort was given a single dose of either vehicle (30% Solutol), MBA3 at 5, 10 or 30 mg kg^{-1} or MBA6 at 10, 30 or 60 mg kg^{-1} 1 h after infection via subcutaneous injection.

Ethics statement. The Rockefeller University Animal Care and Use Committee approved all animal procedures under protocol no. 19032. Mice were maintained in accordance with the American Association for Accreditation of Laboratory Care criteria.

Statistics and reproducibility. Unless stated otherwise, three independent biological replicates ($n=3$) were used for each measurement. Data were reported as mean values ± s.e.m. No statistical method was used to predetermine sample size but our sample sizes are similar to those reported in previous publications[9,10,20,22]. No data were excluded from the analyses. With the exception of the mouse experiments, experiments were not randomized. Investigators were not blinded to allocation during the experiments and outcome assessment.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The biosynthetic gene clusters for MBA1–6 have been deposited in GenBank under accession nos. MZ146900 and MZ146905, respectively. All other data are available in the main text or as supplementary information. The following five publicly available resources were used in this study: antiSMASH (https://antismash-db.secondarymetabolites.org/); MIBiG (https://mibig.secondarymetabolites.org/); Norine (https://bioinfo.lifl.fr/norine/); SANDPUMA (https://bitbucket.org/chevrm/sandpuma); and PubChem databases (https://pubchem.ncbi.nlm.nih.gov/). Source data are provided with this paper.

Received: 24 May 2021; Accepted: 29 October 2021;
Published online: 23 December 2021

References

1. *Tackling Drug-resistant Infections Globally: Final Report and Recommendations. The Review on Antimicrobial Resistance* (Wellcome Trust and UK Government, 2016).
2. de Kraker, M. E. A., Stewardson, A. J. & Harbarth, S. Will 10 million people die a year due to antimicrobial resistance by 2050? *PLoS Med.* **13**, e1002184 (2016).
3. Brown, E. D. & Wright, G. D. Antibacterial drug discovery in the resistance era. *Nature* **529**, 336–343 (2016).
4. Niu, G. & Li, W. Next-generation drug discovery to combat antimicrobial resistance. *Trends Biochem. Sci.* **44**, 961–972 (2019).
5. Lewis, K. The science of antibiotic discovery. *Cell* **181**, 29–45 (2020).
6. Johnston, J. M. & Bulloch, E. M. Advances in menaquinone biosynthesis: sublocalisation and allosteric regulation. *Curr. Opin. Struct. Biol.* **65**, 33–41 (2020).
7. Boersch, M., Rudrawar, S., Grant, G. & Zunk, M. Menaquinone biosynthesis inhibition: a review of advancements toward a new antibiotic mechanism. *RSC Adv.* **8**, 5099–5105 (2018).
8. Paudel, A., Hamamoto, H., Panthee, S. & Sekimizu, K. Menaquinone as a potential target of antibacterial agents. *Drug Discov. Ther.* **10**, 123–128 (2016).
9. Le, P. et al. Repurposing human kinase inhibitors to create an antibiotic active against drug-resistant *Staphylococcus aureus*, persisters and biofilms. *Nat. Chem.* **12**, 145–158 (2020).
10. Hamamoto, H. et al. Lysocin E is a new antibiotic that targets menaquinone in the bacterial membrane. *Nat. Chem. Biol.* **11**, 127–133 (2015).
11. Itoh, H. et al. Total synthesis and biological mode of action of WAP-8294A2: a menaquinone-targeting antibiotic. *J. Org. Chem.* **83**, 6924–6935 (2018).
12. Sang, M. et al. Identification of an anti-MRSA cyclic lipodepsipeptide, WBP-29479A1, by genome mining of *Lysobacter antibioticus*. *Org. Lett.* **21**, 6432–6436 (2019).
13. Santiago, M. et al. Genome-wide mutant profiling predicts the mechanism of a Lipid II binding antibiotic. *Nat. Chem. Biol.* **14**, 601–608 (2018).
14. Butler, M. S., Blaskovich, M. A. & Cooper, M. A. Antibiotics in the clinical pipeline in 2013. *J. Antibiot. (Tokyo)* **66**, 571–591 (2013).
15. Rutledge, P. J. & Challis, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat. Rev. Microbiol.* **13**, 509–523 (2015).
16. Crits-Christoph, A., Diamond, S., Butterfield, C. N., Thomas, B. C. & Banfield, J. F. Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis. *Nature* **558**, 440–444 (2018).
17. Libis, V. et al. Uncovering the biosynthetic potential of rare metagenomic DNA using co-occurrence network analysis of targeted sequences. *Nat. Commun.* **10**, 3848 (2019).
18. Kalkreuter, E., Pan, G., Cepeda, A. J. & Shen, B. Targeting bacterial genomes for natural product discovery. *Trends Pharmacol. Sci.* **41**, 13–26 (2020).
19. Nayfach, S. et al. A genomic catalog of Earth's microbiomes. *Nat. Biotechnol.* **39**, 499–509 (2021).
20. Chu, J. et al. Discovery of MRSA active antibiotics using primary sequence from the human microbiome. *Nat. Chem. Biol.* **12**, 1004–1006 (2016).
21. Chu, J., Vila-Farres, X. & Brady, S. F. Bioactive synthetic-bioinformatic natural product cyclic peptides inspired by nonribosomal peptide synthetase gene clusters from the human microbiome. *J. Am. Chem. Soc.* **141**, 15737–15741 (2019).
22. Hover, B. M. et al. Culture-independent discovery of the malacidins as calcium-dependent antibiotics with activity against multidrug-resistant Gram-positive pathogens. *Nat. Microbiol.* **3**, 415–422 (2018).
23. Peek, J. et al. Rifamycin congeners kanglemycins are active against rifampicin-resistant bacteria via a distinct mechanism. *Nat. Commun.* **9**, 4147 (2018).
24. Charlop-Powers, Z. et al. Global biogeographic sampling of bacterial secondary metabolism. *eLife* **4**, e05048 (2015).
25. Lemetre, C. et al. Bacterial natural product biosynthetic domain composition in soil correlates with changes in latitude on a continent-wide scale. *Proc. Natl Acad. Sci. USA* **114**, 11615–11620 (2017).
26. Reddy, B. V., Milshteyn, A., Charlop-Powers, Z. & Brady, S. F. eSNaPD: a versatile, web-based bioinformatics platform for surveying and mining natural product biosynthetic diversity from metagenomes. *Chem. Biol.* **21**, 1023–1033 (2014).
27. Stachelhaus, T., Mootz, H. D. & Marahiel, M. A. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* **6**, 493–505 (1999).
28. Blin, K. et al. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **47**, W81–W87 (2019).
29. Chen, H. T., Olson, A. S., Su, W., Dussault, P. H. & Du, L. Fatty acyl incorporation in the biosynthesis of WAP-8294A, a group of potent anti-MRSA cyclic lipodepsipeptides. *RCS Adv.* **5**, 105753–105759 (2015).
30. Zhang, W. et al. Identification and characterization of the anti-methicillin-resistant *Staphylococcus aureus* WAP-8294A2 biosynthetic gene cluster from *Lysobacter enzymogenes* OH11. *Antimicrob. Agents Chemother.* **55**, 5581–5589 (2011).
31. Chen, D., Tian, L., Po, K. H. L., Chen, S. & Li, X. Total synthesis and a systematic structure-activity relationship study of WAP-8294A2. *Bioorg. Med. Chem.* **28**, 115677 (2020).
32. Bionda, N., Pitteloud, J.-P. & Cudic, P. Cyclic lipodepsipeptides: a new class of antibacterial agents in the battle against resistant bacteria. *Future Med. Chem.* **5**, 1311–1330 (2013).
33. Kato, A. et al. A new anti-MRSA antibiotic complex, WAP-8294A. I. Taxonomy, isolation and biological activities. *J. Antibiot. (Tokyo)* **51**, 929–935 (1998).
34. Yagi, A. et al. Anti-*Mycobacterium* activity of microbial peptides in a silkworm infection model with *Mycobacterium smegmatis*. *J. Antibiot. (Tokyo)* **70**, 685–690 (2017).
35. Collins, M. D. & Jones, D. The distribution of isoprenoid quinones in streptococci of serological groups D and N. *J. Gen. Microbiol.* **114**, 27–33 (1979).
36. Huycke, M. M. et al. Extracellular superoxide production by *Enterococcus faecalis* requires demethylmenaquinone and is attenuated by functional terminal quinol oxidases. *Mol. Microbiol.* **42**, 729–740 (2001).
37. Müller, A., Klöckner, A. & Schneider, T. Targeting a cell wall biosynthesis hot spot. *Nat. Prod. Rep.* **34**, 909–932 (2017).
38. Malin, J. J. & de Leeuw, E. Therapeutic compounds targeting Lipid II for antibacterial purposes. *Infect. Drug Resist.* **12**, 2613–2625 (2019).
39. Sampson, B. A., Misra, R. & Benson, S. A. Identification and characterization of a new gene of *Escherichia coli* K-12 involved in outer membrane permeability. *Genetics* **122**, 491–501 (1989).
40. Meganathan, R. Biosynthesis of menaquinone (vitamin K_2) and ubiquinone (coenzyme Q): a perspective on enzymatic mechanisms. *Vitam. Horm.* **61**, 173–218 (2001).
41. Meganathan, R. & Kwon, O. Biosynthesis of menaquinone (vitamin K2) and ubiquinone (coenzyme Q). *EcoSal Plus* https://doi.org/10.1128/ecosalplus.3.6.3.3 (2009).
42. Wakeman, C. A. et al. Menaquinone biosynthesis potentiates haem toxicity in *Staphylococcus aureus*. *Mol. Microbiol.* **86**, 1376–1392 (2012).
43. Saravanan, M. et al. Review on emergence of drug-resistant tuberculosis (MDR & XDR-TB) and its molecular diagnosis in Ethiopia. *Microb. Pathog.* **117**, 237–242 (2018).
44. *Global Tuberculosis Report 2019* (World Health Organization, 2019).
45. Libardo, M. D. J., Boshoff, H. I. & Barry, C. E.3rd The present state of the tuberculosis drug development pipeline. *Curr. Opin. Pharmacol.* **42**, 81–94 (2018).
46. Wellington, S. & Hung, D. T. The expanding diversity of *Mycobacterium tuberculosis* drug targets. *ACS Infect. Dis.* **4**, 696–714 (2018).
47. Berube, B. J. et al. Novel MenA inhibitors are bactericidal against *Mycobacterium tuberculosis* and synergize with electron transport chain inhibitors. *Antimicrob. Agents Chemother.* **63**, e02661-18 (2019).
48. Pieters, J. *Mycobacterium tuberculosis* and the macrophage: maintaining a balance. *Cell Host Microbe* **3**, 399–407 (2008).
49. Rastogi, N., Labrousse, V. & Goh, K. S. In vitro activities of fourteen antimicrobial agents against drug susceptible and resistant clinical isolates of *Mycobacterium tuberculosis* and comparative intracellular activities against the virulent H37Rv strain in human macrophages. *Curr. Microbiol.* **33**, 167–175 (1996).

50. Laufer, B., Chatterjee, J., Frank, A. O. & Kessler, H. Can *N*-methylated amino acids serve as substitutes for prolines in conformational design of cyclic pentapeptides? *J. Pept. Sci.* **15**, 141–146 (2009).
51. Stark, M., Liu, L.-P. & Deber, C. M. Cationic hydrophobic peptides with antimicrobial activity. *Antimicrob. Agents Chemother.* **46**, 3585–3590 (2002).
52. Kaji, T. et al. Total synthesis and functional evaluation of fourteen derivatives of lysocin E: importance of cationic, hydrophobic, and aromatic moieties for antibacterial activity. *Chemistry* **22**, 16912–16919 (2016).
53. Kaupp, M. The function of photosystem I. Quantum chemical insight into the role of tryptophan–quinone interactions. *Biochemistry* **41**, 2895–2900 (2002).
54. Yu, L. et al. Identification of the biosynthetic gene cluster for the anti-MRSA lysocins through gene cluster activation using strong promoters of housekeeping genes and production of new analogs in *Lysobacter* sp. 3655. *ACS Synth. Biol.* **9**, 1989–1997 (2020).
55. Masschelein, J., Jenner, M. & Challis, G. L. Antibiotics from Gram-negative bacteria: a comprehensive overview and selected biosynthetic highlights. *Nat. Prod. Rep.* **34**, 712–783 (2017).
56. Liu, Y., Ding, S., Shen, J. & Zhu, K. Nonribosomal antibacterial peptides that target multidrug-resistant bacteria. *Nat. Prod. Rep.* **36**, 573–592 (2019).
57. Urakami, T., Tamaoka, J., Suzuki, K.-I. & Komagata, K. *Paracoccus alcaliphilus* sp. nov., an alkaliphilic and facultatively methylotrophic bacterium. *Int. J. Syst. Bacteriol.* **39**, 116–121 (1989).
58. Xie, C.-H. & Yokota, A. *Dyella japonica* gen. nov., sp. nov., a γ-proteobacterium isolated from soil. *Int. J. Syst. Evol. Microbiol.* **55**, 753–756 (2005).
59. Proctor, R. A. et al. Small colony variants: a pathogenic form of bacteria that facilitates persistent and recurrent infections. *Nat. Rev. Microbiol.* **4**, 295–305 (2006).
60. Lannergård, J. et al. Identification of the genetic basis for clinical menadione-auxotrophic small-colony variant isolates of *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **52**, 4017–4022 (2008).
61. Cao, S., Huseby, D. L., Brandis, G. & Hughes, D. Alternative evolutionary pathways for drug-resistant small colony variant mutants in *Staphylococcus aureus*. *mBio* **8**, e00358-17 (2017).
62. Blin, K., Shaw, S., Kautsar, S. A., Medema, M. H. & Weber, T. The antiSMASH database version 3: increased taxonomic coverage and new query features for modular enzymes. *Nucleic Acids Res.* **49**, D639–D643 (2021).
63. Flissi, A. et al. Norine: update of the nonribosomal peptide resource. *Nucleic Acids Res.* **48**, D465–D469 (2020).
64. Kautsar, S. A. et al. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* **48**, D454–D458 (2020).
65. Chevrette, M. G., Aicheler, F., Kohlbacher, O., Currie, C. R. & Medema, M. H. SANDPUMA: ensemble predictions of nonribosomal peptide chemistry reveal biosynthetic diversity across Actinobacteria. *Bioinformatics* **33**, 3202–3210 (2017).
66. Samdani, A. & Vetrivel, U. POAP: a GNU parallel based multithreaded pipeline of open babel and AutoDock suite for boosted high throughput virtual screening. *Comput. Biol. Chem.* **74**, 39–48 (2018).
67. Kim, S. et al. PubChem in 2021: new data content and improved web interfaces. *Nucleic Acids Res.* **49**, D1388–D1395 (2021).
68. Röttig, M. et al. NRPSpredictor2––a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* **39**, W362–W367 (2011).
69. Xie, F, Pei, S., Lin, X., Tian, Y. & Zhang, G. A rapid and efficient method for the extraction and identification of menaquinones from *Actinomycetes* in wet biomass. *BMC Microbiol.* **21**, 175 (2021).

Acknowledgements

We thank the High Throughput Screening and Resource Center at The Rockefeller University for their assistance with the ITC experiments. We thank J. Rock at The Rockefeller University for *M. smegmatis* mc^2 155. We thank W. Jacobs Jr at Albert Einstein College of Medicine for *M. tuberculosis* mc^2 6206, mc^2 7901 and mc^2 6206 with the mLux plasmid. We thank V. Fischetti at The Rockefeller University for *S. aureus* Newman, D712 and O315. We thank Y. Q. Xiong at the David Geffen School of Medicine at University of California Los Angeles for *S. aureus* A215 and SA684. We thank E. Skaar at Vanderbilt University Medical Center for *S. aureus* Δ*menA* and Δ*menB*. We thank J. Peek and J. Burian at The Rockefeller University for careful proofreading. This work was supported by the National Institutes of Health (grant nos. 1U19AI142731 and 5R35GM122559 to S.F.B.).

Author contributions

L.L. designed the overall experimental plan for the manuscript, performed most of the experiments presented and wrote the manuscript. B.K. performed the peptide synthesis and assisted with the mouse experiments. Y.H. performed the bioinformatics analysis. L.W.M. contributed to the NMR data analysis. M.A.T. performed the MiSeq sequencing. R.R. performed the antibacterial assay against the WT and MDR *M. tuberculosis* strains. S.F.B. contributed to project management, designed the overall experimental plan for the manuscript and wrote the manuscript. All authors edited the manuscript and approved the final submission.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41564-021-01013-8.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41564-021-01013-8.

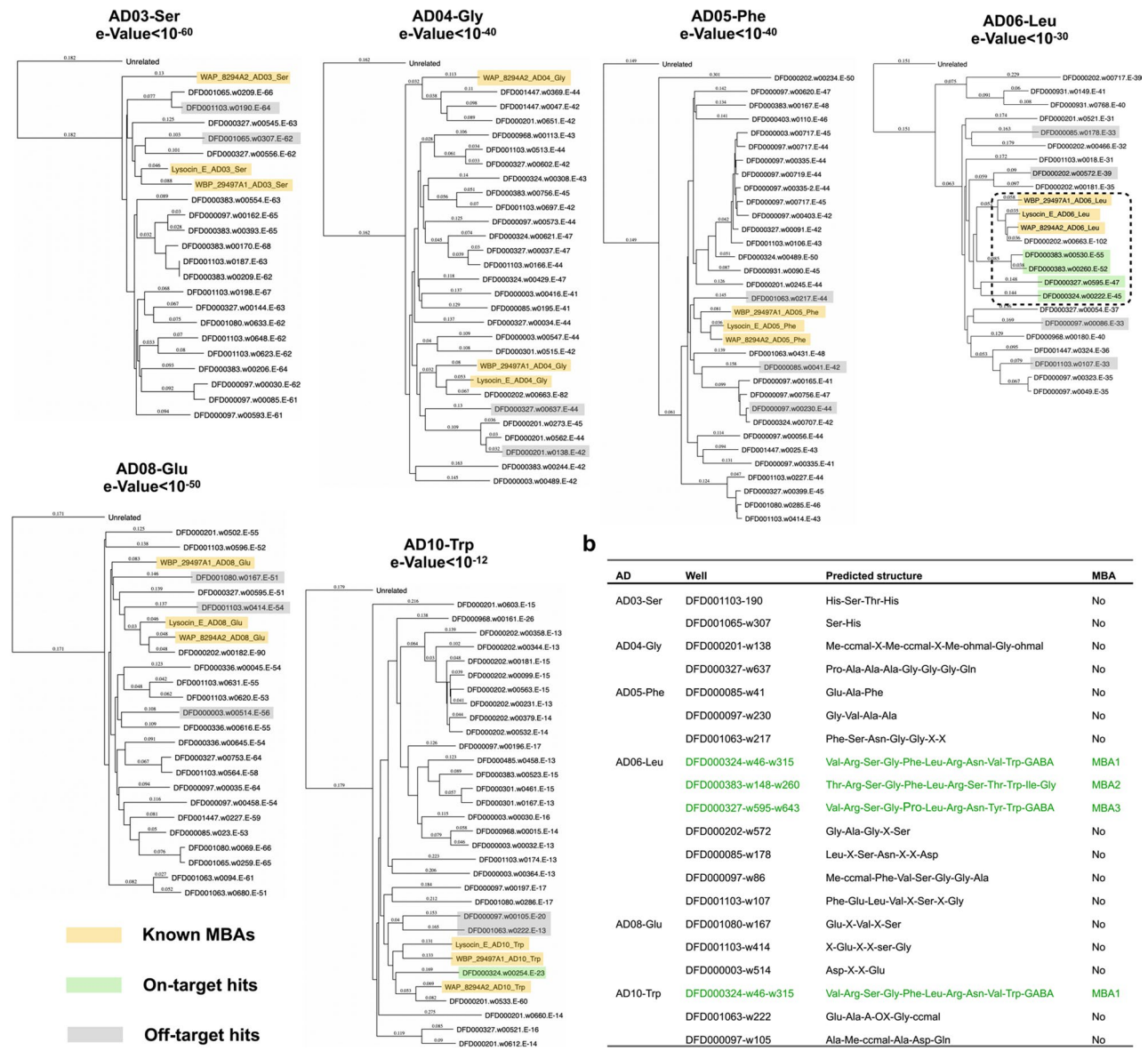
Correspondence and requests for materials should be addressed to Sean F. Brady.

Peer review information *Nature Microbiology* thanks Michael R. Barbachyn, Stephan A. Sieber and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

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

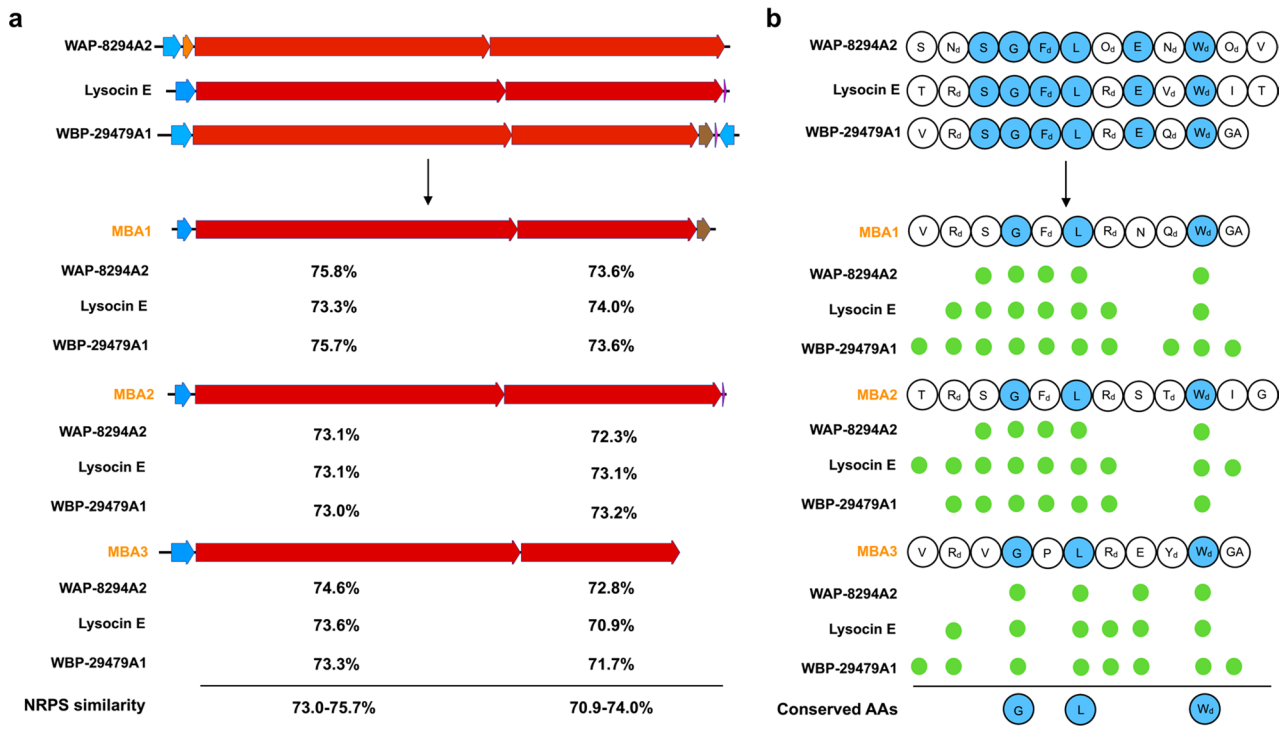
a



b

AD	Well	Predicted structure	MBA
AD03-Ser	DFD001103-190	His-Ser-Thr-His	No
	DFD001065-w307	Ser-His	No
AD04-Gly	DFD000201-w138	Me-ccmal-X-Me-ccmal-X-Me-ohmal-Gly-ohmal	No
	DFD000327-w637	Pro-Ala-Ala-Ala-Gly-Gly-Gly-Gln	No
AD05-Phe	DFD000085-w41	Glu-Ala-Phe	No
	DFD000097-w230	Gly-Val-Ala-Ala	No
	DFD001063-w217	Phe-Ser-Asn-Gly-Gly-X-X	No
AD06-Leu	DFD000324-w46-w315	Val-Arg-Ser-Gly-Phe-Leu-Arg-Asn-Val-Trp-GABA	MBA1
	DFD000383-w148-w260	Thr-Arg-Ser-Gly-Phe-Leu-Arg-Ser-Thr-Trp-Ile-Gly	MBA2
	DFD000327-w595-w643	Val-Arg-Ser-Gly-Pro-Leu-Arg-Asn-Tyr-Trp-GABA	MBA3
	DFD000202-w572	Gly-Ala-Gly-X-Ser	No
	DFD000085-w178	Leu-X-Ser-Asn-X-X-Asp	No
	DFD000097-w86	Me-ccmal-Phe-Val-Ser-Gly-Gly-Ala	No
	DFD001103-w107	Phe-Glu-Leu-Val-X-Ser-X-Gly	No
AD08-Glu	DFD001080-w167	Glu-X-Val-X-Ser	No
	DFD001103-w414	X-Glu-X-X-ser-Gly	No
	DFD000003-w514	Asp-X-X-Glu	No
AD10-Trp	DFD000324-w46-w315	Val-Arg-Ser-Gly-Phe-Leu-Arg-Asn-Val-Trp-GABA	MBA1
	DFD001063-w222	Glu-Ala-A-OX-Gly-ccmal	No
	DFD000097-w105	Ala-Me-ccmal-Ala-Asp-Gln	No

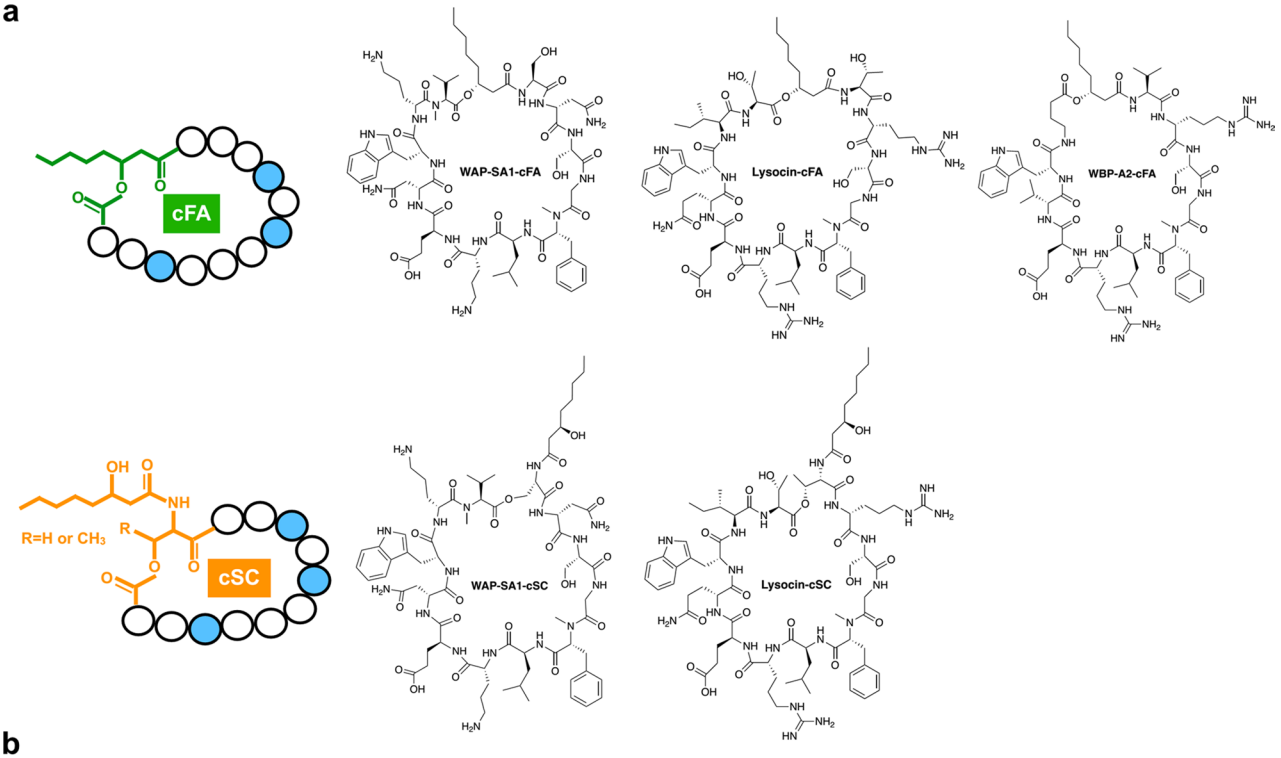
Extended Data Fig. 1 | Phylogenetic analysis of eSNaPD hits from six conserved A-domains found in the BGCs of the three known MBAs. Phylogenetic analysis of eSNaPD hits (**a**) and predicted peptide sequences of recovered clones (**b**). All the hits at an e-value $\leq 10^{-45}$ from A-domain analysis of L-Leu-6 encoded new MBAs and formed a separate, well-defined clade with A-domains of three known MBAs, which suggested that L-Leu-6 in the proposed minimal MK-binding motif were encoded by the most highly conserved A domain in MBA-family peptides.



Extended Data Fig. 2 | Three potential MBA BGCs from eSNaPD-guided soil metagenomic mining. Comparison of NRPS gene organization (**a**) as well as amino acid substrates (**b**) between the three known MBA BGCs and the three potential MBA BGCs we cloned from soil metagenomes. The blue residues represent building blocks that are conserved across all MBAs. The green circles represent residues that are shared between known and potential MBAs.

Extended Data Fig. 3 | Predicted MBA peptide sequences identified in a motif search of the p-NRP database (a) and the BGCs associated with these predicted peptides (b). The blue residues represent building blocks that are conserved across all MBAs.

a

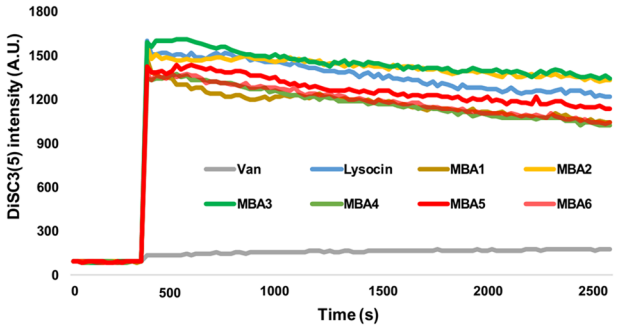


b

Peptides	Cyclization	Natural cyclization site	*B. subtilis*	*S. aureus*	*S. epidemidis*	*M. tuberculosis*
WAP-SA1	**cFA**	**Yes**	**2**	**4**	**4**	**1.25**
	cSC	No	>64	>64	>64	>20
Lysocin	**cSC**	**Yes**	**0.5**	**4**	**2**	**1.25**
	cFA	No	32	32	32	>20
WBP-A2	**cFA**	**Yes**	**0.5**	**4**	**0.5**	**2.5**

Extended Data Fig. 4 | The structures (a) and anti-bacterial activities (b) of the *N*-acylated peptides associated with known MBAs cyclized in two different ways. The (*R*)-3-hydroxy-octanoic acid analogs of lysocin E, WBP-29479A1 and the deoxy version of WAP-8294A1 shown here were synthesized in this study. *B. subtilis* 168 1A1, *S. aureus* USA300, *S. epidemidis* RP62A and *M. tuberculosis* H37Rv were used as tested strains. The blue residues represent building blocks that are conserved across all MBAs.

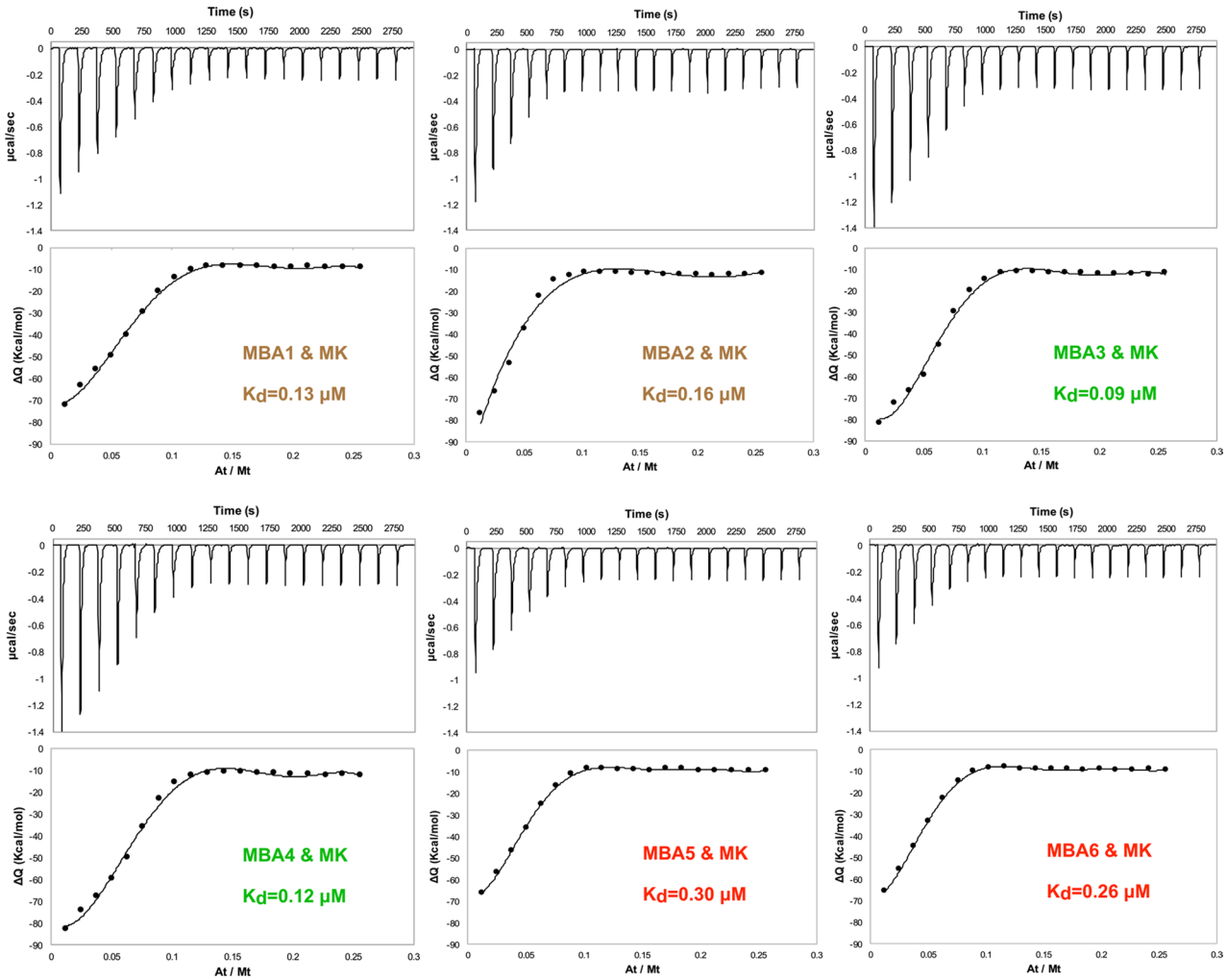
a



b

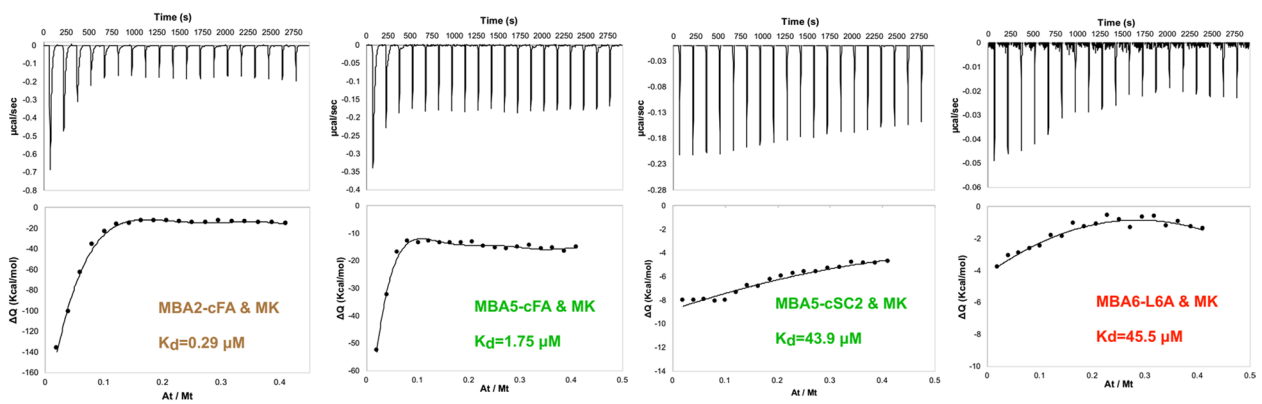
Compounds	Treating concentration	Resistant frequency
MBA1	16 μg/mL	2.9×10^{-6}
MBA2	32 μg/mL	2.1×10^{-6}
MBA3	4 μg/mL	5.0×10^{-6}
MBA4	16 μg/mL	2.9×10^{-6}
MBA5	32 μg/mL	0.7×10^{-6}
MBA6	16 μg/mL	2.1×10^{-6}

Extended Data Fig. 5 | Membrane depolarization activity and resistance frequency of MBAs 1 through 6. a, The effect of each MBA on *S. aureus* membrane potential was measured using 3,3′-Dipropylthiadicarbocyanine iodide [DiSC3(5)]. Vancomycin (Van) and lysocin were used as the negative and positive controls, respectively. **b**, Resistance frequency of MBAs 1 through 6 against *S. aureus* USA300 in the presence of 4x the MIC of each antibiotic.



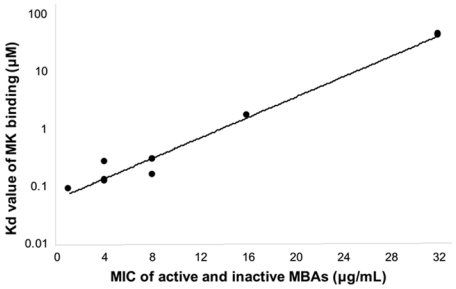
Extended Data Fig. 6 | Isothermal titration of 1:1 (mol/mol) DOPC:DOPG vesicles containing MK into each MBA.

a



b

Compound	MIC (μg/mL)	Kd (μM)
MBA3	1	0.09
MBA1	4	0.13
MBA4	4	0.12
MBA6	4	0.26
MBA2	8	0.16
MBA5	8	0.30
MBA2-cFA	8	0.29
MBA5-cFA	16	1.75
MBA5-cSC2	32	43.9
MBA6-cFA	32	45.5

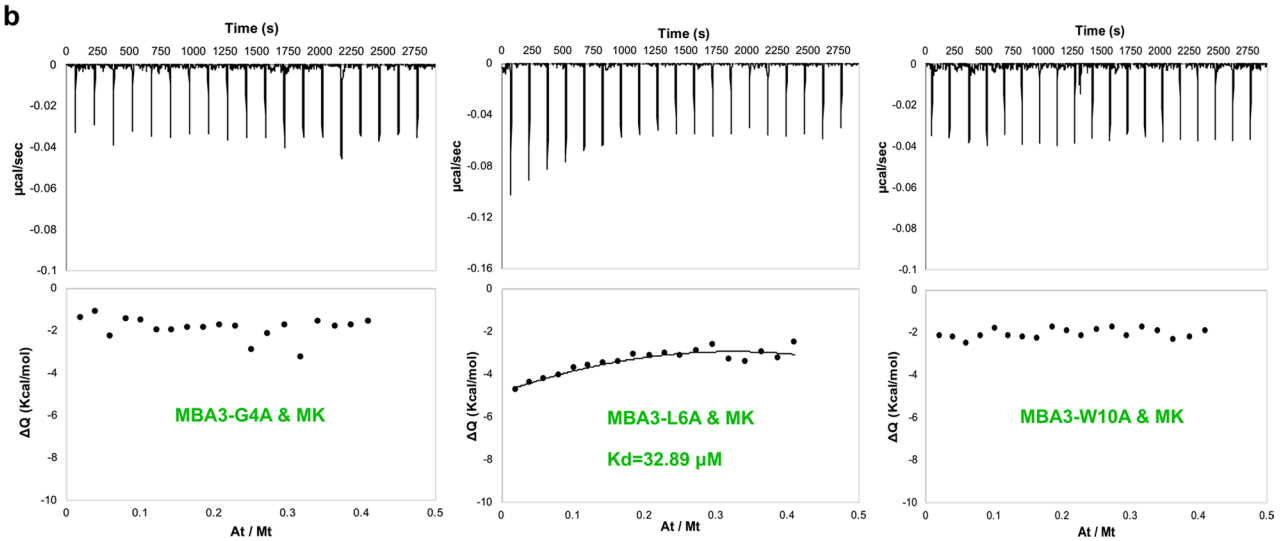


Extended Data Fig. 7 | Correlation between antibiotic activity and MK binding affinity for active or inactive syn-BNP MBAs. a, Isothermal titration of 1:1 (mol/mol) DOPC:DOPG vesicles containing MK into the four additional syn-BNPs we generated in Fig. 2. **b**, Comparison of Kd values and MICs against *S. aureus* USA300 for all syn-BNP MBAs in Fig. 2.

a

Compound	*S. aureus* Newman	*S. aureus* USA300	*S. aureus* COL
MBA3	1	1	1
MBA3-G4A	>64	>64	>64
MBA3-L6A	32	16	32
MBA3-W10A	>64	>64	>64

b



Extended Data Fig. 8 | Antibiotic activity and MK binding of MBA3 with single point mutations in the proposed minimal MK-binding motif. MIC in µg/mL, highest concentration tested was 64 µg/mL.

nature research

Corresponding author(s): Sean Brady

Last updated by author(s): Oct 27, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Cosmid or genome sequencing data was collected using an Illumina Miseq sequencer. NMR spectra were acquired on a Bruker Avance DMX 600 MHz spectrometer equipped with cryogenic probes. Murine macrophage infection data was collected using a Spark multimode microplate reader. Membrane depolarization/lysis and cytotoxicity data was collected using a Tecan Infinite M Nano+ plate reader. ITC data was collected using an Auto-iTC200 microcalorimeter.
Data analysis	eSNaPD 2.0 was used for metagenomic mining of biosynthetic gene clusters. Newbler 2.6 was used to assemble Illumina sequencing reads. MacVector 17.0.5 was used for analyzing biosynthetic gene clusters and building phylogenetic tree. antiSMASH v5.1.2 was used for ORF detection. GraphPad Prism v8 was used to generate dose response curves. SNIPPY (https://github.com/tseemann/snippy) were used to identify Single-nucleotide polymorphisms (SNPs) for resistant mutants. The web-based application Affinimeter ITC (https://www.affinimeter.com/site/itc/) was used to process ITC data based on the “one binding site” model.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Biosynthetic gene clusters for MBAs 1 through 6 have been deposited in GenBank under Accession numbers MZ146900 to MZ146905, respectively. All other data are available in the main text or supplementary information. Source data are provided with this paper. The following five publicly available resources were used in this study: antiSMASH database (https://antismash-db.secondarymetabolites.org/), MIBiG (https://mibig.secondarymetabolites.org/), NORINE database (https://bioinfo.lifl.fr/norine/), SANDPUMA (https://bitbucket.org/chevrm/sandpuma) and PubChem database (https://pubchem.ncbi.nlm.nih.gov/).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We selected the number of independent biological replicates based on our previous experience with equivalent experiments, feasibility and level of variability between replicates. Sample size was indicated as "n" in the corresponding figure legends. No statistical methods were used to determine sample size.
Data exclusions	No data was excluded from analyses reported in this manuscript.
Replication	All attempts at replication were successful. The number of independent biological replicates for each experiment are provided in the Figure legends.
Randomization	For animal studies, Female Swiss Webster (6 week old) mice were assigned randomly to the experimental and control groups. With the exception of the mouse experiments, experiments were not randomized.
Blinding	For all experiments in this study no blinding was necessary as the magnitude of the effect observed preclude inadvertent human influence over the interpretation of the results.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	The human cell line HEK293 was purchased from ATCC (CRL-1573) and the mouse cell line J774A.1 was purchased from Sigma-Aldrich (91051511).

Authentication	The human cell HEK293 and mouse cell J774A.1 were authenticated externally by commercial vendor using STR profiling approach. The two cell lines were then inspected visually in-house.
Mycoplasma contamination	Cell lines were verified to be free of mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Female Swiss Webster (6 week old) mice were used in this study.
Wild animals	This study did not include wild animals.
Field-collected samples	This study did not include field-collected samples.
Ethics oversight	The Rockefeller University Animal Care and Use Committee approved all animal procedures under the protocol 19032. Mice were maintained in accordance with American Association for Accreditation of Laboratory Care criteria.

Note that full information on the approval of the study protocol must also be provided in the manuscript.