HHS Public Access

Author manuscript

Nat Microbiol. Author manuscript; available in PMC 2022 June 23.

Published in final edited form as:

Nat Microbiol. 2022 January ; 7(1): 120–131. doi:10.1038/s41564-021-01013-8.

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], **Sean F. Brady**[1]

[1]Laboratory of Genetically Encoded Small Molecules, The Rockefeller University, 1230 York Avenue, New York, NY 10065, United States

[2]Department of Medicine, Center for Emerging and Re-emerging Pathogens, Rutgers University-New Jersey Medical School, Newark, NJ 07103, United States

Abstract

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 (MOA). Directly targeting menaquinone (MK), which plays an important role in bacterial electron transport, is an appealing, yet underexplored, MOA 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 (syn-BNP) approach, which relies on bioinformatic structure prediction followed by chemical synthesis. Among these six structurally diverse MBAs, four makeup 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 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.

Corresponding author: Sean F. Brady, **Contact**: Laboratory of Genetically Encoded Small Molecules, The Rockefeller University, 1230 York Avenue, New York, NY 10065, **Phone:** 212-327-8280, **Fax:** 212-327-8281, sbrady@rockefeller.edu.

Author Contributions

L.L. designed the overall experimental plan for the manuscript, performed the majority of the experiments presented and wrote the manuscript. B.K. performed the peptide synthesis and assisted the mouse experiments. Y.H. performed the bioinformatic 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 wild-type 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 financial interests.

Introduction

Antimicrobial resistance presents a major and growing healthcare problem and contributes annually to ~700,000 deaths around the world.[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 that are currently in clinical use.[3–5] In the majority of 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 novel antibiotics.[7–9] Three closely related nonribosomal 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. reported that lysocin E also binds to lipid II, a precursor for bacterial cell wall synthesis.[13] Although WAP-8294A2 (lotilibcin) progressed to phase I clinical trials, a dearth of additional chemical entities that can induce antibiosis 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 remains functionally inaccessible. This is due both to our inability to culture most bacteria and to the fact that 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 large 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 (*i.e.,* 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 (*i.e.*, a synthetic bioinformatic natural product, syn-BNP).[20, 21]

Here we took two orthogonal bioinformatic approaches to guide the discovery of bacterial BGCs that encode structurally novel 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 that 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* (*Mtb*) both *in vitro* and in a macrophage assay, defining MK-binding as an anti-*Mtb* mode of action. Among the antibiotics we identified, four fell into two new structural classes. Antibiotics from both 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 here 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 that are predicted to encode new MBAs: We began our search for BGCs that might encode MBAs by looking at NRPS adenylation (A)-domain sequence data generated from soil metagenomic libraries. As part of our ongoing soil metagenome-guided natural product discovery program, we have 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 before.[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 ~5–25,000 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 (*i.e.,* 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 package,[26] each A-domain from a BGC that encodes 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 install 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 analyzed for the potential to encode MBA-like peptides. The linear peptide encoded by each eDNA derived BGC was predicted using the 10 amino acid residues that line each A-domain substrate binding pocket (*e.g.,* A-domain signature sequence).[27, 28] Based on this analysis no predicted peptides contained all six residues that were conserved among known MBAs. In three cases however, 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 bioactivity analysis presented below each of these predicted peptides does in fact represent a new MBA (Fig. 2).

GXLXXXW motif search of sequenced genomes for BGCs that are 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 analyzed BGCs from ~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 (see 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 that is seen in previously characterized MBAs. By including metagenome derived MBA sequences in this study, we were able to reduce 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). As NRPS BGCs are often truncated in genomes assembled using short read sequence data (*e.g.*, Illumina sequencing), we also searched our database for sub-sequences 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 two additional examples where peptides shared additional amino acids with newly predicted 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 re-sequenced the genomes of *D. tabacisoli* KCTC 62035 and *Paracoccus alcaliphilus* ATCC 51199. In both cases A-domain analysis of re-sequenced BGCs predicted they would encode GXLXXXW containing peptides that might represent MBAs (Fig. 1).

Prediction and synthesis of 10 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, as well as 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 Asn incorporated as the second amino acid in the peptide.[30] In the known MBA WAP-8294A2, the hydroxyl group on the Asn at the same position ($_D$-OHAsn-2) is not required for anti-bacterial activity.[11, 31] The unmodified linear lipopeptide produced by each NRPS system was therefore predicted to be the direct precursor to the final biologically active cyclic lipopeptide produced by each BGC. As we felt 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 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 significant differences in anti-bacterial 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 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 (fatty acid cyclized, cFA) or through a nucleophilic amino acid side chain (side chain cyclized, cSC) (Fig. 2). When no nucleophilic side chain was present in the peptide, only a fatty acid cyclized 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 (*i.e.,* a serine or threonine) both fatty acid cyclized and amino acid side chain cyclized analogs 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 amino side chain cyclized peptides were generated. In total we generated 10 synthetic bioinformatic natural products (syn-BNPs) for bioactivity screening (Fig. 2). For use as controls, we synthesized (*R*)-3-hydroxy-octanoic acid analogs of the three known MBAs (Extended Data Fig. 4a, WAP-SA1, lysocin and WBP-A2).

Anti-microbial spectrum

MK binding antibiotics 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 a small number of Gram-positive bacteria (*e.g.*, *Bacillus subtilis, Staphylococcus 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, and they were therefore used in all of the subsequent studies. We also synthesized “unnaturally cyclized” versions of two known MBA analogs (lysocin

and WAP-SA1) and in both cases these structures showed a comparable decrease or loss of anti-bacterial 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 antibiosis (Fig. 2). Ultimately, this analysis resulted in six new and structurally diverse cyclic lipopeptide antibiotics (MBA1 through 6) that are only linked by a shared GXLXXXW motif that we hypothesized was associated with MK binding as a mode of action.

All six syn-BNPs were broadly active against Gram-positive bacteria. Against a panel of *S. aureus* strains that are resistant to diverse clinically relevant antibiotics, the MBA MICs ranged from 0.25 to 8 μg/mL. MBA3, which was predicted from a metagenome derived BGC, was the most potent antibiotic with the MICs against *S. aureus* strain ranging from 0.25 to 2 μg/mL. A subset of these syn-BNPs, MBA3 in particular, were active against *Mtb* (Supplementary Table 2). Lysocin E has been reported to be active against *Mycobacterium smegmatis*,[34] but to the best of our knowledge, no MBA has been reported to have anti-*Mtb* 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 (*e.g., Enterococcus casseliflavus* and *Streptococcus cremoris*) but were inactive against all six MK-deficient *Enterococcus* and *Strepococcus* 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 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 (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 2 to 4-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, these 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] We therefore 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 each syn-BNP's antibiosis activity was examined in four ways: 1) we assayed for MK's ability to suppress antibiosis, 2) we tested for antibiosis 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 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 antibiosis 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 *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) against both *S. aureus* strains (Fig. 3d). Furthermore, we selected *S. aureus* (USA300) mutants that could grow on 4x the MIC of each antibiotic. At 4x each syn-BNP's MIC resistant mutants appeared at a frequency of $0.7 – 5 \times 10^{-6}$ 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 a MK biosynthesis gene and in almost all cases no other mutations were detected in the genome (Supplementary Table 5). Finally, by 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 0.09–0.30 μM (Extended Data Fig. 6). In contrast, no syn-BNP induced an exothermic reaction when 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 to their more active counter parts. For example, the calculated dissociation constant for the fatty acid cyclized version of MBA6, MBA6-cFA, was 45.5 μM (Extended Data Fig. 7a). This was 175-fold lower than the dissociation constant observed for the interaction between MK and the active, side chain cyclized, 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 natural product lysocin E has been reported to bind lipid II,[13] in ITC experiments with liposomes containing lipid II we did not observe an antibiotic specific binding between syn-BNP MBAs and lipid II (see Supplementary discussion and Supplementary Fig. 4 and 5). Collectively, these data provided multiple distinct lines of evidence that "GXLXXXW" motif containing syn-BNPs specifically bound MK and their antibacterial activity was dependent on MK.

Anti-*Mtb* activity and mode of action analysis for MBAs

Tuberculosis remains one of the deadliest infectious diseases in the world. Anti-*Mtb* agents with novel modes-of-action are urgently needed due to the rapid emergence of MDR and extensively drug-resistant *Mtb* mutants.[43, 44] Although enzymes in the MK biosynthesis pathway have been explored as potential anti-*Mtb* targets,[45–47] to the best of our knowledge MK binding has not been tested as a mode of action for an anti-*Mtb* agent. We assayed MBAs 1 through 6 against a panel of *Mtb* strains that included wild-type H37Rv, two

mutants that can be studied using BSL2 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 *Mtb* strains (MIC ≤ 10 μg/mL). MBA3 was the most potent anti-*Mtb* compound among MBAs, with an MIC as low as 0.078 μg/mL against MDR *Mtb*.

We sought to confirm the mode of action of the class of antibiotics against *Mtb*. As seen with *S. aureus* when MK was added to the assay media the MIC of MBAs against *Mtb* 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-*Mtb* 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 *Mtb* (Fig. 4c). Taken together, these data suggest that in *Mtb* MBA's antibiosis activity remains MK dependent membrane disruption.

Macrophages play a central role in recognizing and destroying invading pathogens. *Mtb,* 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 *Mtb*. We examined the anti-*Mtb* activity of MBAs in a murine macrophage model. In this assay, J774A.1 mouse macrophages infected with *Mtb* harboring the mLux plasmid were treated with each *Mtb* active MBA, and then residual bacterial cell viability inside the macrophages was determined by luminescence measurements. The three *Mtb* active MBAs we tested all inhibited *Mtb* growth in macrophages with IC_{50} ranging from 0.14 to 2.1 μg/mL (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*,[4] in the future it will be interesting to explore the *in vivo* anti-*Mtb* activity of MBA3.

Two new MBA structural families

Although six MBAs we identified share a conserved GXLXXXW sequence that we propose is important for MK-binding, they exhibit significant differences in overall peptide sequence as well as different modes of cyclization and anti-microbial potency. To more easily visualize structural relationships among these antibiotics, linear MBA peptide sequences were aligned and a phylogenetic tree was generated (Fig. 5a). This tree contains three distinct clades, one of which is composed of known (lysocin E and WBP-29479A1) and new MBAs (MBA1 and MBA2), while the other two clades only contain 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 building block 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. One of which consists of MBA3 and MBA4 (Fig. 5c). Both structures were cyclized between the hydroxyl on (*R*)-3-hydroxy-octanoic acid and the C-terminal GABA to form undeca-lipodepsipeptide. Beyond the conserved minimal MK binding sequence, these two structures share a distinct $_L$-Pro-5 and a $_D$-Tyr-9. The $_L$-Pro-5 replaces the $_D$-*N*-Me-aromatic amino acid that appears in all other known or new MBAs between the G and L in the

conserved GXLXXXW motif. Proline like *N*-methylated amino acids can introduce discrete conformations into cyclic peptides,[50] suggesting 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 Ser, Ser, Asn, Thr and Phe 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 6 do not contain any positively charged amino acids. Positively charged residues found in known MBAs as well as 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] The absence of cationic residues in MBA5 and MBA6 is therefore 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 analogs 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 analogs 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 to 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) as well as its ability to interact with MK. All three conserved residues of the proposed minimal MK-binding motif are therefore 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 remains to be determined. However, it is likely that the indole of Trp interacts with the quinone from MK,[53] and that the *N*-methyl aromatic amino acid and the proline induce similar cyclic peptide conformations that help create a MK binding pocket. The hydrophobic Gly and Leu residues as well as 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 we don't know its source, the BGC for MBA4 is found in the genome of *P. alcaliphilus*. 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 programs.[55, 56] *P. alcaliphilus* is an alkaliphilic facultative methanol-utilizing bacterium,

while the genus *Dyella* was only first described in 2005.[57, 58] As outlined in 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 known MBA BGC $_{L}$-Leu-6 A-domain, all NPSTs that returned e-values ≤ 10^{-45} were found to arise from an MBA BGC (Extended Data Fig. 1). To more extensively explore MBA BGC diversity in the environment, 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 (*i.e.,* e-value ≤ 10^{-45}). Our studies therefore not only identified additional cultured sources of structurally diverse MBAs, but also suggest that there is a large 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 analogs of the two new MBA families we identified in this study. As shown in Fig. 6, treatment of a methicillin-resistant *S. aureus* COL infection with either MBA3 (5, 10 and 30 mg/kg) or MBA6 (10, 30 and 60 mg/kg) dramatically decreased the mortality of infected mice compared to treatment with vehicle alone (30% solutol). A minimal dose of 10 and 30 mg/kg 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-*Mtb* activity and potent *in vivo* anti-*S. aureus* activity. Although MBAs exhibit significant differences in peptide sequence and different cyclization modes, they share 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 here, as well as 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 heme biosynthesis.[10, 13] In both cases, these mutants show a small colony variant (SCV) phenotype.[59] In addition, both MK and heme deficient mutants have been found to show 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-deficient or heme-deficient strains ~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 the majority of 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 novel classes of MBAs suggests that bioinformatic structure prediction algorithms have developed to the extent that direct structure, or sub-structure, searches of collections of bioinformatically predicted natural product structures now represents 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. Pre-loaded 2-chlorotrityl resin for peptide synthesis was purchased from Matrix Innovation, Inc (Quebec, Canada). Reagents for solid-phase peptide synthesis (SPPS), DCM (dichloromethane), DIPEA (*N,N*-diisopropylethylamine), DMAP (4-dimethylaminopyridine), DMF (dimethylformamide), HATU (O-(7-azabenzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate), HFIP (hexafluoroisopropanol), PyAOP ((7-azabenzotriazol-1-yloxy) tripyrrolidinophosphonium hexafluorophosphate) and TFA (trifluoroacetic acid), were purchasd from P3 BioSystems (Louisville, KY). Standard *N*-Fmoc amino acid building blocks were purchased from P3 BioSystems and Chem-Impex International (Wood Dale, IL). (*R*)-3-hydroxy-octanic acid and Fmoc-*N*-Me-D-Phe-OH were purchased from Enamine (Monmouth, NJ). Fmoc-*N*-Me-D-Trp(Boc)-OH was purchased from Alabiochem (Suzhou, China). Fmoc-*N*-Me-D-Tyr(tBu)-OH was purchased from 1 ClickChemistry (Kendall Park, NJ). Fmoc-GABA-OH was purchased from Sigma-Aldrich (St. Louis, Missouri). All solvents used for chromatography were HPLC grade or higher. MTT (thiazolyl blue tetrazolium bromide) and Type II mucin from porcine stomach were purchased from Sigma-Aldrich. Fluorescent dyes SYTOX Green and $DISC_3(5)$ (3,3'-dipropylthiadicarbocyanine Iodide) were purchased from ThermoFisher Scientific (Waltham, MA), and the assay results were recorded using a Tecan Infinite M Nano$^+$ plate reader (Morrisville, pNC). 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) and 1,2-dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol) (DOPG) were purchased from Avanti Polar Lipids (Alabaster, Alabama). 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 (Coventry, UK).

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). UPLC-LRMS data were acquired on a Waters Acquity system equipped with QDa and PDA detectors, a Phenomenex Synergi Fusion-RP 80 Å column (2.0 × 50 mm, 4 μm) and controlled by Waters MassLynx software. The following chromatographic conditions were used for UPLC-LRMS: 5% B from 0.0 to 0.9 min, 5% to 95% B from 0.9 to 4.5 min, 95% B from 4.5 to 5.0 min, 95% to 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 and 10 μL injection volume). HPLC-HRMS data were acquired on a SCIEX ExionLC HPLC coupled to an X500R QTOF mass spectrometer, equipped with a Phenomonex Kinetex PS C18 100 Å column (2.1 × 50 mm, 2.6 μm) and controlled by SCIEXOS software. The following chromatographic conditions were used for UPLC-HRMS: 5% B from 0.0 to 1.0 min, 5% to 95% B from 1.0 to 10.0 min, 95% B from 10.0 to 12.5 min, 95% to 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 and 1 μL injection volume). Peptide purification was performed using an Agilent 1200 Series HPLC with UV detection and equipped with an XBridge Prep C18 130 Å column (10 × 150 mm, 5 μm). ^{1}H and ^{13}C NMR spectra were acquired on a Bruker Avance DMX 600 MHz spectrometer equipped with cryogenic probes (The Rockefeller University, New York). All spectra were recorded at 25 or 50 °C in DMSO-d_6. Chemical shift values are reported in 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 novel 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 (environmental Surveyor of Natural Product Diversity) software package 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 software package. 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 2.6 (Roche). Assembled complete and partial BGCs were analyzed using antiSMASH v5.1.2 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-domain to that of A-domains from known NRPs.

Screening ~2000 soils to explore more potential MBAs.

eDNA was extracted from ~2000 ecologically and geographically diverse soil samples and Natural Product Sequence Tags (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 ≤ 10^{-45} were considered as hits. A multiple sequence alignment of all qualifying hit sequences was generated using MUSCLE, and the resulting alignment file was used to generate a maximum likelihood tree in MacVector 17.0.5.

Construction of an in-house predicted non-ribosomal peptide (p-NRP) database.

GenBank files for 38,933 NRP BGCs representing 10,858 complete bacterial genome assemblies were retrieved from the antiSMASH database. Our p-NRP database was constructed from the BGCs by synthesizing data from five A-domain prediction resources: antiSMASH database,[62] the NORINE amino acid database,[63] A-domain substrate predictions from MIBiG,[64] SANDPUMA[65] and our own NRP BGC analyses (Supplementary Fig. 1). We analyzed all BGCs using antiSMASH v5.1.228 and the parallel job execution tool GNU Parallel.[66] After removing duplicate entries, we obtained a total of 36,957 NRPS antiSMASH v5 regions. We downloaded the entirety of the "monomer" dataset from NORINE, which represents known A-domain substrates as well as 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 more closely in the NRP BGC community. 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. 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 was imported,[67] including the PubChem ID to ensure uniqueness. Substrates which did not appear in PubChem were still included in the database; however, other uniquely identifying information, such as IUPAC name and SMILES string, were relied upon to prevent a substance being erroneously represented multiple times. When no substrate specificity code was identified for an A-domain, the 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 predicted NRP sequence was determined either by the presence of a condensation starter (Cs) domain, or the presence of A-domain with no immediately preceding condensation (C) domain. The end of the peptide sequence was defined by the presence of a thioesterase (TE) 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 as well as two partial motifs: GXL and LXXXW.

Peptide synthesis.

Solid phase peptide synthesis (SPPS): All peptides were synthesized using standard Fmoc-based solid-phase peptide synthesis methods on 2-chlorotrityl chloride resin using commercially available Fmoc-protected amino acids. Peptide synthesis started from the

conserved Leu seen at 6th position of each peptide. Pre-loaded Leu on 2-cholorotrityl resin (0.3 g, 0.552 mmol/g) was swollen in DCM for 20 min at room temperature then drained and washed with DMF (3 mL, 3x). Coupling of individual amino acids was carried out by using Fmoc-protected amino acids (2 equiv., relative to resin loading) mixed with HATU (2 equiv.) and DIPEA (2 equiv.) in DMF (5 mL). Coupling reactions were carried out for 1 h with occasional swirling then washed with DMF (3 mL, 3x). 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, 5x) 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 equiv.), DIPEA (40 equiv.), benzoyl chloride (20 equiv.) and DMAP (0.8 equiv.) in 10 mL DCM and gently shaken for 72 h. After the ester bond formation, 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 equiv.) and DIPEA (30 equiv.) in DMF (50 mL). After 2 h, DCM (100 mL) was added and washed repeatedly with 1% formic acid in water (5 mL, 10x). The extracted cyclic peptide was air dried overnight.

Final cleavage: Air dried cyclic peptide was dissolved in 3 mL 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 in −20 °C for 10 min to precipitate the peptide. Peptide pellets were harvested by centrifuging (2500xg) for 5 min, re-dissolved in 5 mL methanol and dried *in vacuo* 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 MS (Supplementary Fig. 8–13) analyses. ^{1}H NMR and ^{13}C NMR spectra were recorded for each peptide (Supplementary Fig. 14–19). Pure peptides were dissolved in DMSO at 6.4 mg/mL for MIC measurements and mode of action studies.

Antimicrobial assays against Gram-positive bacteria, Gram-negative bacteria and yeast pathogens.

All antimicrobial assays were run in 96-well microtiter plates using a broth micro-dilution method. Diluted overnight cultures were used in all assays. For yeast strains, overnight cultures were diluted 2000-fold in YPD broth. For *Enterococcus faecium* and *Staphylococcus aureus*, overnight cultures were diluted 1000- and 10,000-fold in LB broth, respectively. For *Streptococcus* strains, overnight cultures were diluted 5,000-fold in Brain Heart Infusion (BHI) broth. For other bacteria, overnight cultures were diluted 5,000-fold in

LB broth. 100 μL of each diluted culture was mixed with 100 μL of LB broth containing a syn-BNP at 2-fold serial dilutions across a 96-well microtiter plate row. The final concentration of each compound ranged from 64 to 0.25 μg/mL. 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 minimum inhibition concentration (MIC). All MIC assays were done in duplicate (n = 2).

Antibacterial assay against *Mycobacterium 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 2-fold serial dilutions across a 96-well plate row. The final concentration of each compound ranged from 64 to 0.25 μg/mL. Plates were incubated for 48 h at 37 °C and then 30 μL of Alamar Blue cell viability reagent (Thermo Fisher Scientific) was added. After an additional 24 h incubation, the 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 *Mycobacterium tuberculosis*.

Mtb BSL2 mc^2 6206, BSL2 mc^2 7901, wild-type H37Rv and four multidrug-resistant 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 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. 100 μL of 7H9 broth containing each syn-BNP at 2-fold serial dilutions across a 96-well plate row was added to give final test concentrations ranging from 20 to 0.039 μg/mL. The plates were then incubated at 37 °C and 5% CO_2. After incubation for 6 days, 30 μL of Alamar Blue cell viability reagent was added, the cultures were incubated for another 24 h, and the 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 *Mtb* was determined by infecting J774A.1 mouse macrophages (Sigma-Aldrich, 91051511) with *Mtb* mc^2 6206 harboring the mLux plasmid (*Mtb* mc^2 6206/mLux). Macrophages were initially suspended at a concentration of 4–5 × 10^5 cells/mL in Dulbecco's Modified Eagle Medium (DMEM, Sigma-Aldrich) supplemented with Fetal Bovine Serum (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. *Mtb* mc^2 6206/mLux was grown to mid-log phase (OD_{600}=0.5–0.7). *Mtb* cultures were then spun down, washed with Phosphate-Buffered Saline (PBS), and resuspended in DMEM containing 10% FBS, L-glutamate (2 mM), pantothenic acid (50 μg/mL) and leucine (50 μg/mL). The assay plates were inoculated with 100 μL of *Mtb* mc^2 6206/mLux at a multiplicity of infection of 1:10 and incubated for 4 h to allow *Mtb* to infect the macrophages. The adhered macrophages were then washed twice with 100 μL 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. 100 μL of each MBA serial diluted in DMEM (from 64 to 0.0004 μg/mL) was added to the assay plates. Plates

were then incubated at 37 °C for 72 h. Residual *Mtb* 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 GraphPad Prism v8 and plotted as the logarithm of concentration vs normalized percent 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 microtiter 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 (Ex/Em 488/523 nm). When the signal stabilized the appropriate amount of each antibiotic (6.4 mg/mL DMSO stock solutions) to give 2x 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 respect to the average fluorescence signal prior to 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 lysis assays were done in 384-well black microtiter 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 $DiSC_3(5)$ (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 then added and incubated for another 15 min. Fluorescence intensity of the mixture was recorded continually at 2 s intervals (Ex/Em 643/675 nm). When the signal stabilized the appropriate amount of each MBA (6.4 mg/mL DMSO stock solutions) to give 2x 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 respect to the average fluorescence signal prior to 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 medium and grown overnight at 37 °C. A portion of the overnight culture containing approximately 10^9 cells was diluted (1/10x or 1/40x) into LB containing each antibiotic at 4x its MIC. The resulting mixtures were distributed into microtiter 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 microtiter dilution method described above. DNA was extracted from cultures of colonies that showed an elevated MIC relative to the wild-type, and the resulting DNA was sequenced using MiSeq Reagent Kit V3 (MS-102−3003, Illumina). Single-nucleotide polymorphisms (SNPs) for each new MBA were identified using SNIPPY (http://github.com/tseemann/snippy) by mapping MiSeq reads

to the reference genome of *S. aureus* USA300_FPR3757 (RefSeq assembly accession: GCF_000013465.1) (Extended Data Fig. 5b).

Isothermal Titration Calorimetry (ITC).

For MK and UQ binding, a 1:1 mixture of 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) and 1,2-dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol) (DOPG) containing either 1.25 mol % 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 HEPES (pH 7.5) with 100 mM 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 MBA prepared in 10 mM HEPES buffer. The syringe (150 μL) was loaded with a 5 mM lipid suspension with 1.25 mol % MK4 or UQ10. For lipid II binding, 10 mM DOPC containing 1 mol % lipid II was dissolved in chloroform. The resulting film was hydrated using 50 mM Tris (pH 7.5) with 100 mM NaCl and passed through a 100 nm polycarbonate filter 10 times. For ITC the sample cell was filled with 400 μL of 25 μM an MBA prepared in 50 mM Tris buffer. The syringe (150 μL) was loaded with a 10 mM DOPC suspension with 1 mol % lipid II. Data were collected by using an Auto-iTC200 (Malvern Panalytical) and processed by the web-based application Affinimeter 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*), *S. aureus* Newman and USA300 were grown overnight in LB liquid media. 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 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 2,500xg 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 TLC LuxPlate silica gel 60 F_{254} (Millipore) plates and the plates were developed in a mixture of hexane and diethyl ether (85:15, v/v). MK was visualized by UV exposure, and the plates were photographed. Finally, MK bands were collected from the TLC plates and eluted using isopropanol. Isopropanol-eluted MK was analyzed by HPLC-HRMS and MK4 was used as a standard.

Mtb membrane depolarization assay.

Mtb membrane depolarization assays were done in a 384-well black microtiter plate. *Mtb* mc^2 6206 was grown to exponential phase (OD_{600}=0.5–0.7), washed twice with HG buffer (5 mM HEPES and 5 mM 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. Fluorescence intensity of the mixture was recorded continually at 2 s intervals (Ex/Em 622/670 nm). After the fluorescence intensity stabilized each antibiotic (6.4 mg/mL DMSO stock) was added to give a final concertation of 2x its *Mtb* 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 respect to the average fluorescence signal prior to 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 (ATCC, CRL-1573) were grown at 37 °C in a 5% CO_2 atmosphere in Dulbecco's modified Eagle medium (DMEM) supplemented with fetal bovine serum (10% *v/v*), L-glutamate (2 mM), penicillin (10 units/mL) and streptomycin (10 units/mL). HEK293 cells were seeded into 96-well flat bottom microtiter plates (target density of 2500 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 uL of fresh DMEM medium containing each antibiotic at 10 serially diluted concentrations ranging from 32 to 0.0625 ug/mL. After 48 h at 37 °C the DMEM media was removed and 100 uL of a MTT solution (10 uL of 5 mg/mL MTT in PBS premixed with 100 uL of DMEM) was added into each well. After 4 h at 37 °C the solution was aspirated from each well. Precipitated formazan crystals were dissolved by addition of 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 GraphPad Prism v8. Cytotoxicity assays were performed in triplicate (n = 3).

Mouse peritonitis-sepsis model.

Female outbred Swiss Webster mice (6 week old) were used in all experiments. The housing room for mice was on a twelve hour light cycle (7 am on and 7 pm off) with temperature around 21 °C and humidity around 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 $FeNH_4$-citrate. Cultures were diluted to provide a challenge inoculum of ~5×10^8 CFU in 0.2 mL. 0.2 mL of the challenge inoculum was administered via intraperitoneal injection. 35 mice were randomly grouped into five 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 or MBA6 at 10, 30 or 60 mg/kg 1 h after infection via subcutaneous injection.

Ethics statement.

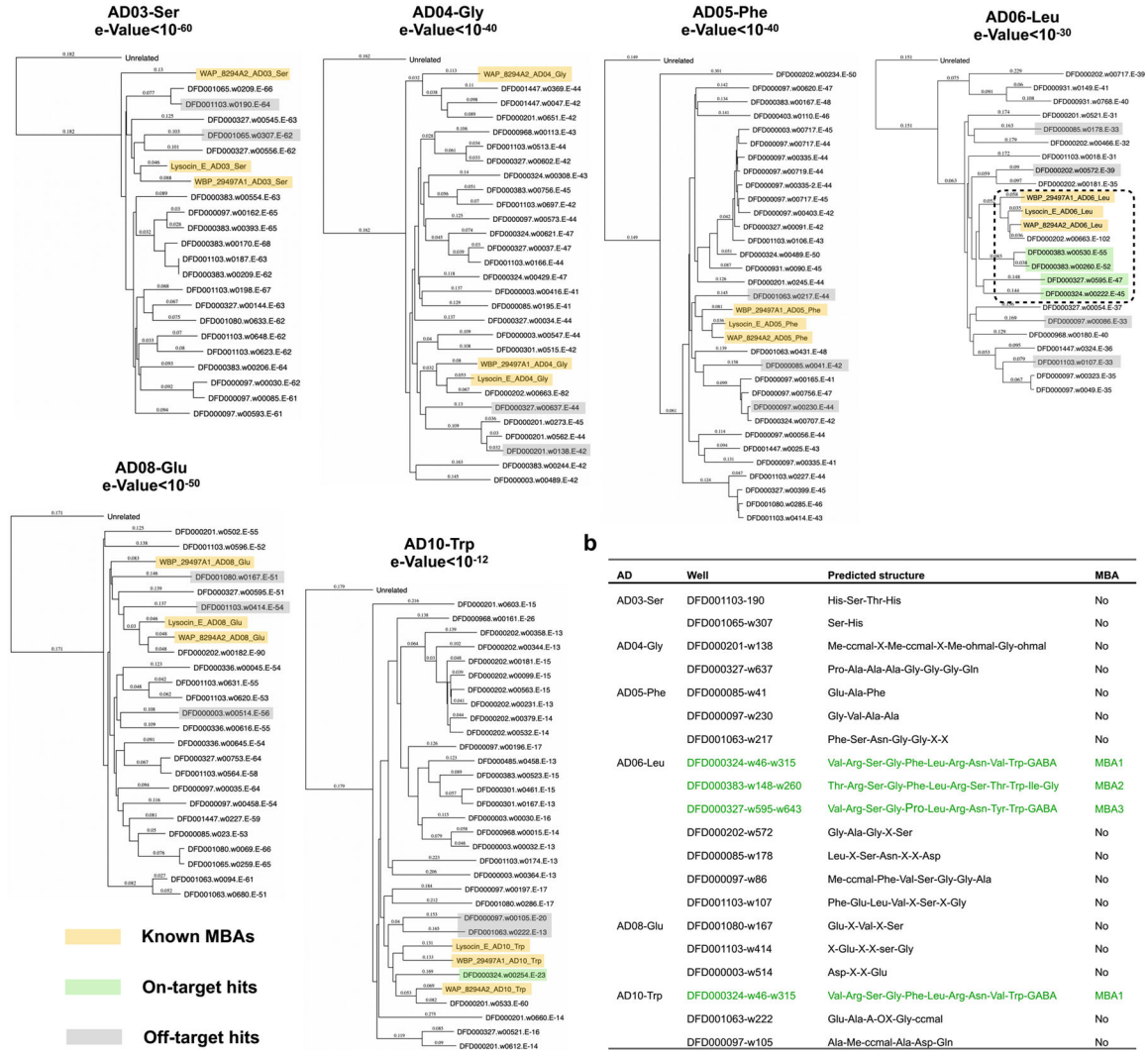
The Rockefeller University Animal Care and Use Committee approved all animal procedures under the protocol 19032. Mice were maintained in accordance with American Association for Accreditation of Laboratory Care criteria.

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. The Investigators were not blinded to allocation during experiments and outcome assessment.

Extended Data

a



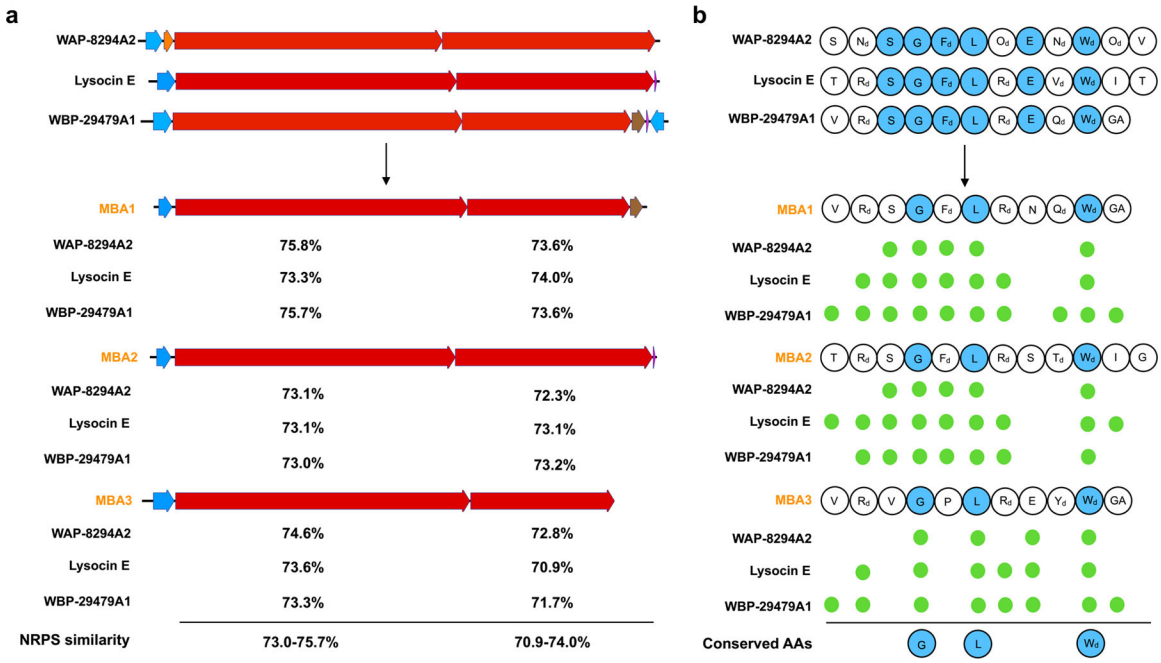
b

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

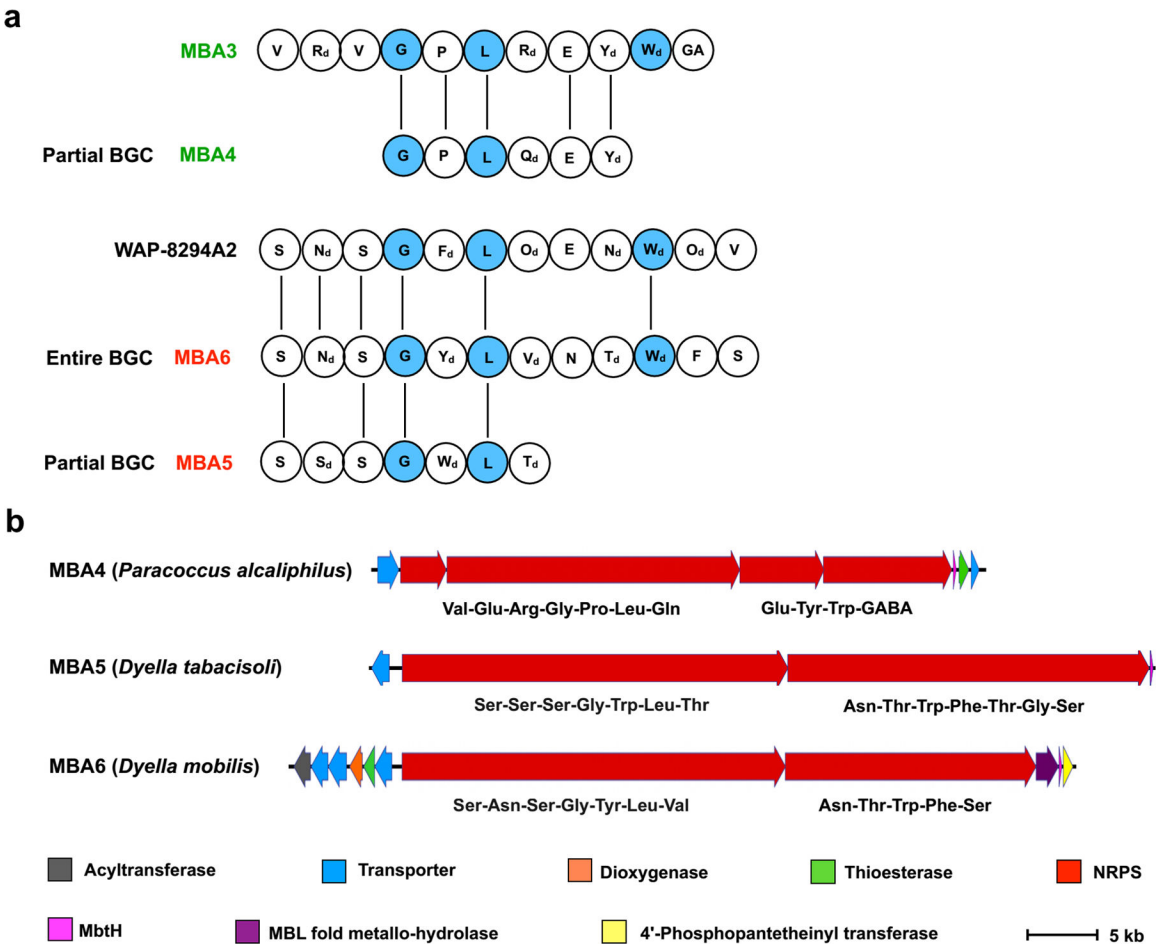
Phylogenetic analysis of eSNaPD hits (**a**) and predicted peptide sequences of recovered clones (**b**). All the hits at an e-value ≤10^{-45} from A-domain analysis of L-Leu-6 encoded new MBAs and formed a separate, well-defined clade with A-domains of three known

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

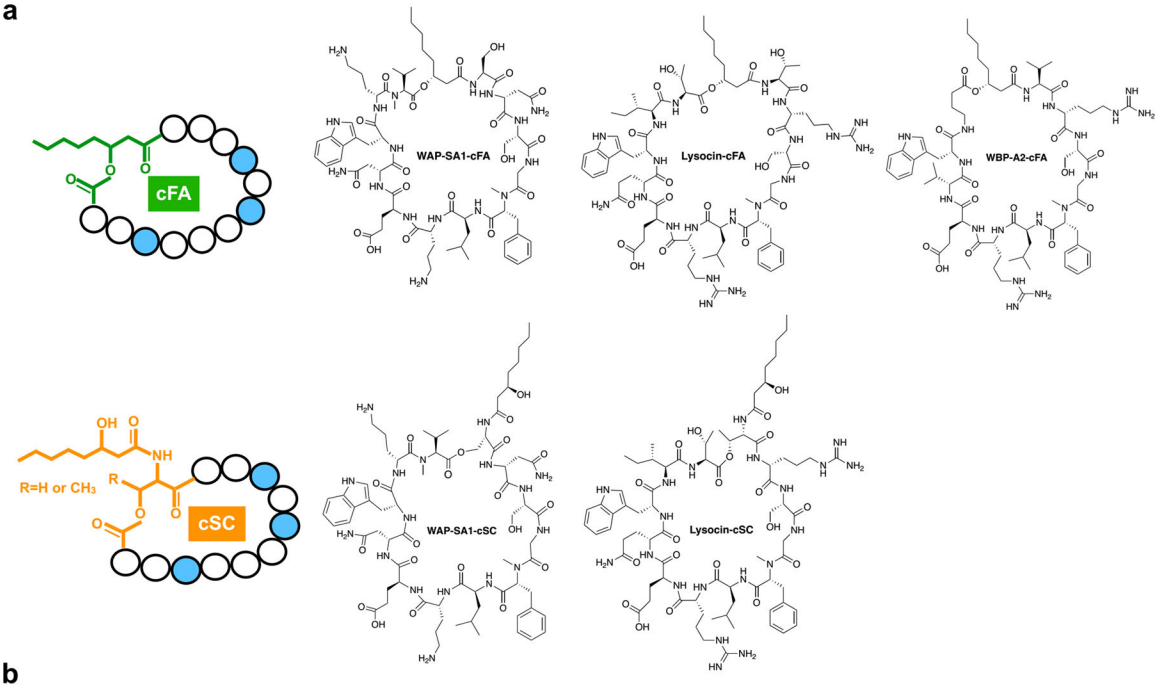
Comparison of NRPS gene organization (**a**) as well as amino acid substrates (**b**) between the three known MBA BGCs and the three potential MBA BGCs we cloned from soil metagenomes. The blue residues represent building blocks that are conserved across all MBAs. The green circles represent residues 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



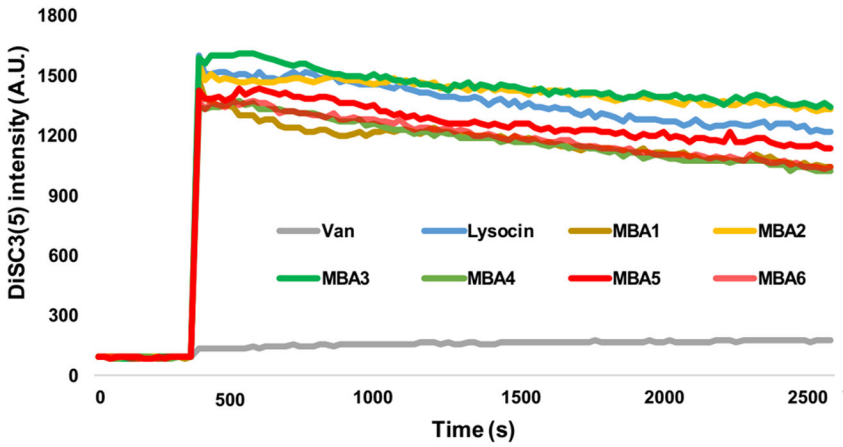
b

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

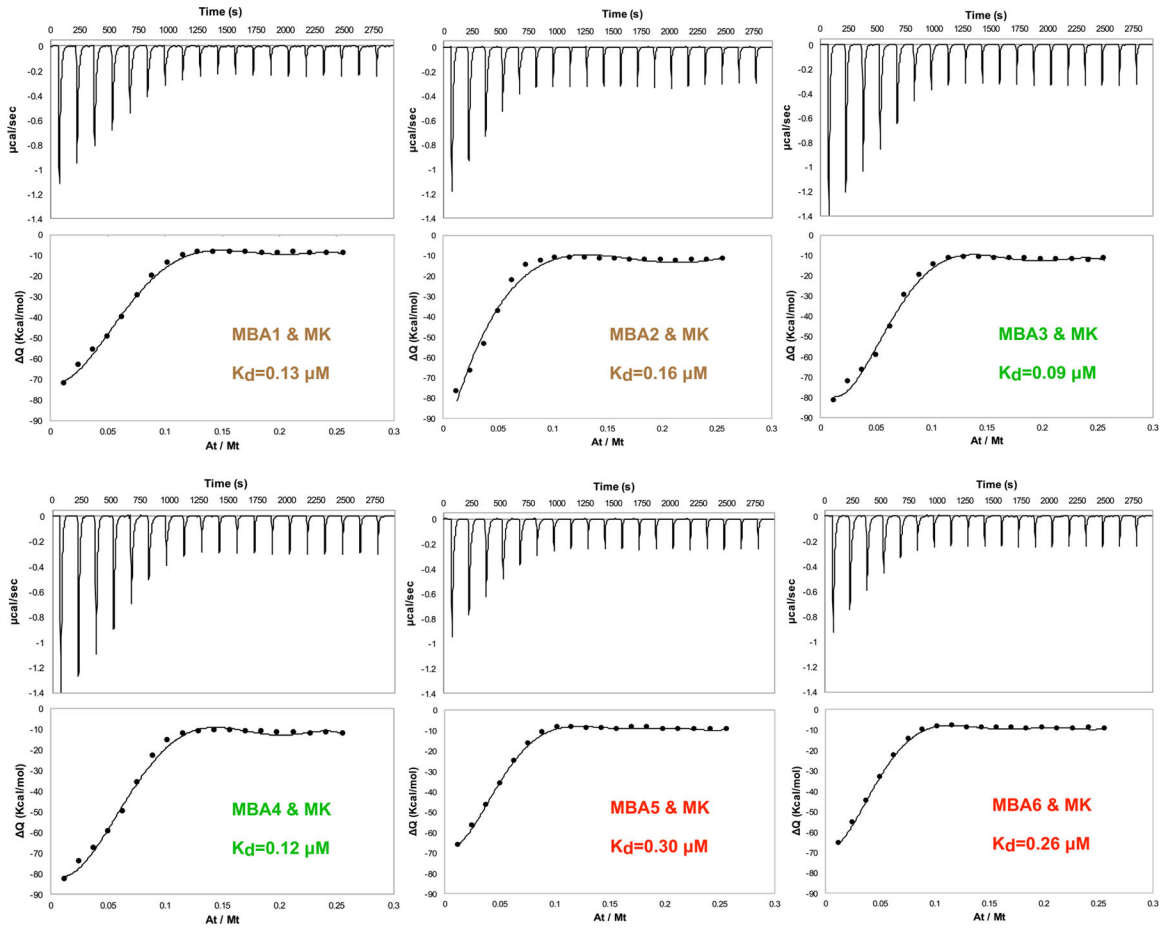


b

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}

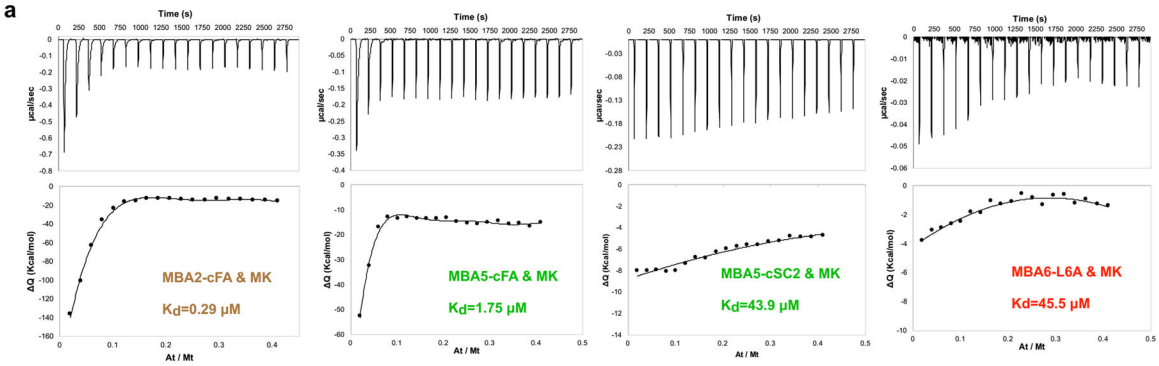
Extended Data Fig. 5. Membrane depolarization activity and resistance frequency of MBAs 1 through 6

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.



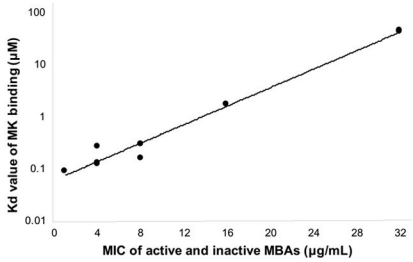
Extended Data Fig. 6.

Isothermal titration of 1:1 (mol/mol) DOPC:DOPG vesicles containing MK into each MBA



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



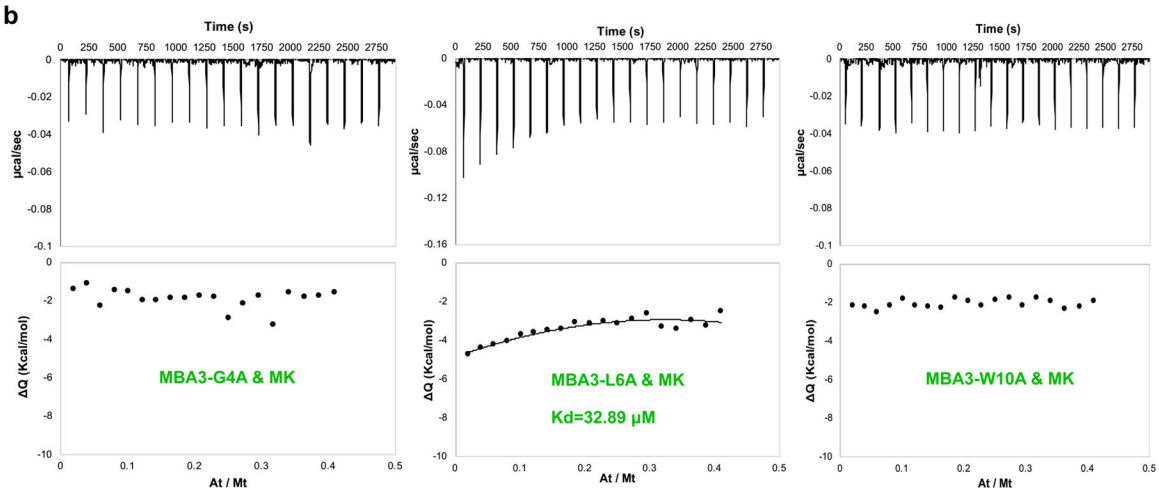
Extended Data Fig. 7. Correlation between antibiotic activity and MK binding affinity for active or inactive syn-BNP MBAs

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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the High Throughput Screening and Resource Center at The Rockefeller University for their assistance with ITC experiments. We thank Jeremy Rock at The Rockefeller University for *Mycobacterium smegmatis* mc^2 155. We thank William 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 Vincent Fischetti at The Rockefeller University for *S. aureus* Newman, D712 and O315. We thank Yan Q. Xiong at the David Geffen School of Medicine at UCLA for *S. aureus* A215 and SA684. We thank Eric Skaar at Vanderbilt University Medical Center for *S. aureus* Δ*menA* and Δ*menB*. We thank James Peek and Jan Burian at The Rockefeller University for careful proofreading. This work was supported by the National Institutes of Health (grant 1U19AI142731 S.F.B. and 5R35GM122559 S.F.B.).

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/).

References

1. Wellcome Trust and UK Government. Review on antimicrobial resistance – tackling drug-resistant infections globally: Final report and recommendations Wellcome Trust and UK Government (2016).
2. de Kraker MEA, Stewardson AJ & Harbarth S Will 10 million people die a year due to antimicrobial resistance by 2050? PLoS Med 13, e1002184 (2016). [PubMed: 27898664]
3. Brown ED & Wright GD Antibacterial drug discovery in the resistance era. Nature 529, 336–343 (2016). [PubMed: 26791724]
4. Niu GQ & Li WL Next-generation drug discovery to combat antimicrobial resistance. Trends Biochem. Sci 44, 961–972 (2019). [PubMed: 31256981]
5. Lewis K The science of antibiotic discovery. Cell 181, 29–45 (2020). [PubMed: 32197064]
6. Johnston JM & Bulloch EM Advances in menaquinone biosynthesis: sublocalisation and allosteric regulation. Curr. Opin. Struct. Biol 65, 33–41 (2020). [PubMed: 32634692]
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). [PubMed: 27431268]
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). [PubMed: 31844194]
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). [PubMed: 25485686]
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). [PubMed: 29019678]
12. Sang ML et al. Identification of an anti-MRSA cyclic lipodepsipeptide, WBP-29479A1, by genome mining of Lysobacter antibioticus. Org. Lett 21, 6432–6436 (2019). [PubMed: 31386380]
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). [PubMed: 29662210]
14. Butler MS, Blaskovich MA & Cooper MA Antibiotics in the clinical pipeline in 2013. J. Antibiot 66, 571–591 (2013).
15. Rutledge PJ & Challis GL Discovery of microbial natural products by activation of silent biosynthetic gene clusters. Nat. Rev. Microbiol 13, 509–523 (2015). [PubMed: 26119570]
16. Crits-Christoph A, Diamond S, Butterfield CN, Thomas BC & Banfield JF Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis. Nature 558, 440–444 (2018). [PubMed: 29899444]
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). [PubMed: 31451725]
18. Kalkreuter E, Pan GH, Cepeda AJ & Shen B Targeting bacterial genomes for natural product discovery. Trends Pharmacol. Sci 41, 13–26 (2020). [PubMed: 31822352]
19. Nayfach S et al. A genomic catalog of Earth's microbiomes. Nat. Biotechnol 39, 499–509 (2020). [PubMed: 33169036]
20. Chu J et al. Discovery of MRSA active antibiotics using primary sequence from the human microbiome. Nat. Chem. Biol 12, 1004–1006 (2016). [PubMed: 27748750]
21. Chu J, Vila-Farres X & Brady SF 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). [PubMed: 31545899]

22. Hover BM 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). [PubMed: 29434326]
23. Peek J et al. Rifamycin congeners kanglemycins are active against rifampicin-resistant bacteria via a distinct mechanism. Nat. Commun 9, 4147 (2018). [PubMed: 30297823]
24. Charlop-Powers Z et al. Global biogeographic sampling of bacterial secondary metabolism. eLife 4, e05048 (2015). [PubMed: 25599565]
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. U.S.A 114, 11615–11620 (2017). [PubMed: 29078342]
26. Reddy BV, Milshteyn A, Charlop-Powers Z & Brady SF eSNaPD: a versatile, web-based bioinformatics platform for surveying and mining natural product biosynthetic diversity from metagenomes. Chem. Biol 21, 1023–1033 (2014). [PubMed: 25065533]
27. Stachelhaus T, Mootz HD & Marahiel MA The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem. Biol 6, 493–505 (1999). [PubMed: 10421756]
28. Blin K et al. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res 47, W81–W87 (2019). [PubMed: 31032519]
29. Chen HT, Olson AS, Su W, Dussault PH & Du LC 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). [PubMed: 21930890]
31. Chen DL, Tian LW, Po KHL, Chen S & Li XC Total synthesis and a systematic structure-activity relationship study of WAP-8294A2. Bioorg. Med. Chem 28 (2020).
32. Bionda N, Pitteloud JP & Cudic P Cyclic lipodepsipeptides: a new class of antibacterial agents in the battle against resistant bacteria. Future Med. Chem 5, 1311–1330 (2013). [PubMed: 23859209]
33. Kato A et al. A new anti-MRSA antibiotic complex, WAP-8294A - I. Taxonomy, isolation and biological activities. J. Antibiot 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 70, 685–690 (2017).
35. Collins MD & Jones D The distribution of isoprenoid quinones in streptococci of serological groups D and N. J. Gen. Microbiol 114, 27–33 (1979). [PubMed: 118232]
36. Huycke MM et al. Extracellular superoxide production by Enterococcus faecalis requires demethylmenaquinone and is attenuated by functional terminal quinol oxidases. Mol. Microbiol 42, 729–740 (2001). [PubMed: 11722738]
37. Müller A, Klöckner A & Schneider T Targeting a cell wall biosynthesis hot spot. Nat. Prod. Rep 34, 909–932 (2017). [PubMed: 28675405]
38. Malin JJ & de Leeuw E Therapeutic compounds targeting Lipid II for antibacterial purposes. Infect. Drug Resist 12, 2613–2625 (2019). [PubMed: 31692545]
39. Sampson BA, Misra R & Benson SA Identification and characterization of a new gene of Escherichia coli K-12 involved in outer membrane permeability. Genetics 122, 491–501 (1989). [PubMed: 2547691]
40. Meganathan R Biosynthesis of menaquinone (vitamin K_2) and ubiquinone (coenzyme Q): A perspective on enzymatic mechanisms. Vitam. Horm 61, 173–218 (2001). [PubMed: 11153266]
41. Meganathan R & Kwon O Biosynthesis of menaquinone (vitamin K_2) and ubiquinone (coenzyme Q). EcoSal Plus 3, 10.1128/ecosalplus.3.6.3.3 (2009).
42. Wakeman CA et al. Menaquinone biosynthesis potentiates haem toxicity in Staphylococcus aureus. Mol. Microbiol 86, 1376–1392 (2012). [PubMed: 23043465]
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). [PubMed: 29486274]
44. WHO. Global Tuberculosis Report 2019 World Health Organization (2019).

45. Libardo MDJ, Boshoff HIM & Barry CE The present state of the tuberculosis drug development pipeline. Curr. Opin. Pharmacol 42, 81–94 (2018). [PubMed: 30144650]
46. Wellington S & Hung DT The expanding diversity of Mycobacterium tuberculosis drug targets. ACS Infect. Dis 4, 696–714 (2018). [PubMed: 29412643]
47. Berube BJ et al. Novel MenA inhibitors are bactericidal against Mycobacterium tuberculosis and synergize with electron transport chain inhibitors. Antimicrob. Agents Chemother 63, e02661–18 (2019). [PubMed: 30962346]
48. Pieters J Mycobacterium tuberculosis and the macrophage: maintaining a balance. Cell Host & Microbe 3, 399–407 (2008). [PubMed: 18541216]
49. Rastogi N, Labrousse V & Goh KS *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). [PubMed: 8672093]
50. Laufer B, Chatterjee J, Frank AO & 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). [PubMed: 18985637]
51. Stark M, Liu LP & Deber CM Cationic hydrophobic peptides with antimicrobial activity. Antimicrob. Agents Chemother 46, 3585–3590 (2002). [PubMed: 12384369]
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). [PubMed: 27739191]
53. Kaupp M The function of photosystem I. Quantum chemical insight into the role of tryptophan-quinone interactions. Biochemistry 41, 2895–2900 (2002). [PubMed: 11863427]
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). [PubMed: 32610008]
55. Masschelein J, Jenner M & Challis GL Antibiotics from Gram-negative bacteria: a comprehensive overview and selected biosynthetic highlights. Nat. Prod. Rep 34, 712–783 (2017). [PubMed: 28650032]
56. Liu Y, Ding SY, Shen JZ & Zhu K Nonribosomal antibacterial peptides that target multidrug-resistant bacteria. Nat. Prod. Rep 36, 573–592 (2019). [PubMed: 30324212]
57. Urakami T, Tamaoka J, Suzuki KI & Komagata K *Paracoccus alcaliphilus* sp. nov., an alkaliphilic and facultatively methylotrophic bacterium. Int. J. Syst. Bacteriol 39, 116–121 (1989).
58. Xie CH & Yokota A Dyella japonica gen. nov., sp. nov., a gamma-proteobacterium isolated from soil. Int. J. Syst. Evol. Microbiol 55, 753–756 (2005). [PubMed: 15774657]
59. Proctor RA et al. Small colony variants: a pathogenic form of bacteria that facilitates persistent and recurrent infections. Nat. Rev. Microbiol 4, 295–305 (2006). [PubMed: 16541137]
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). [PubMed: 18779359]
61. Cao S, Huseby DL, Brandis G & Hughes D Alternative evolutionary pathways for drug-resistant small colony variant mutants in Staphylococcus aureus. mBio 8, e00358–17 (2017). [PubMed: 28634236]
62. Blin K, Shaw S, Kautsar SA, Medema MH & Weber T The antiSMASH database version 3: increased taxonomic coverage and new query features for modular enzymes. Nucleic Acids Res 49, D639–D643 (2021). [PubMed: 33152079]
63. Flissi A et al. Norine: update of the nonribosomal peptide resource. Nucleic Acids Res 48, D465–D469 (2020). [PubMed: 31691799]
64. Kautsar SA et al. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. Nucleic Acids Res 48, D454–D458 (2020). [PubMed: 31612915]
65. Chevrette MG, Aicheler F, Kohlbacher O, Currie CR & Medema MH SANDPUMA: ensemble predictions of nonribosomal peptide chemistry reveal biosynthetic diversity across Actinobacteria. Bioinformatics 33, 3202–3210 (2017). [PubMed: 28633438]

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). [PubMed: 29533817]

67. Kim S et al. PubChem in 2021: new data content and improved web interfaces. Nucleic Acids Res 49, D1388–D1395 (2021). [PubMed: 33151290]

68. Rottig M et al. NRPSpredictor2--a web server for predicting NRPS adenylation domain specificity. Nucleic Acids Res 39, W362–W367 (2011). [PubMed: 21558170]

69. Xie FQ, Pei SX, Lin XH, Tian Y & Zhang GY A rapid and efficient method for the extraction and identification of menaquinones from Actinomycetes in wet biomass. BMC Microbiol 21, 175 (2021). [PubMed: 34103006]

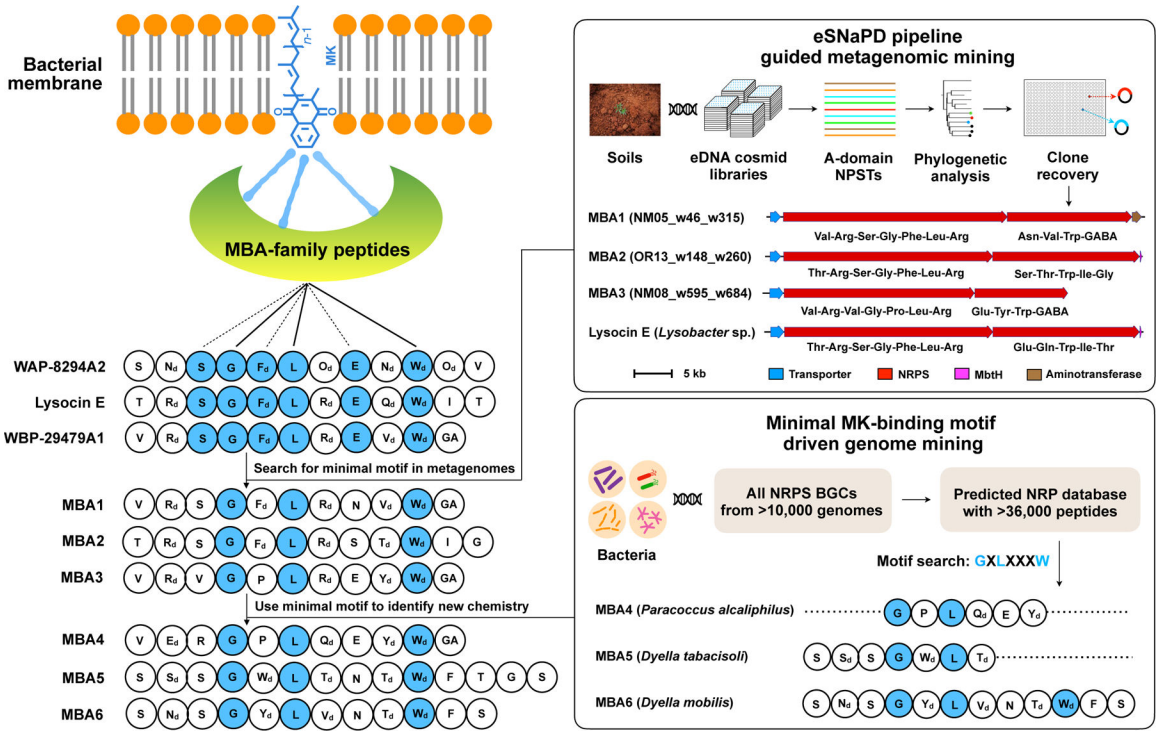
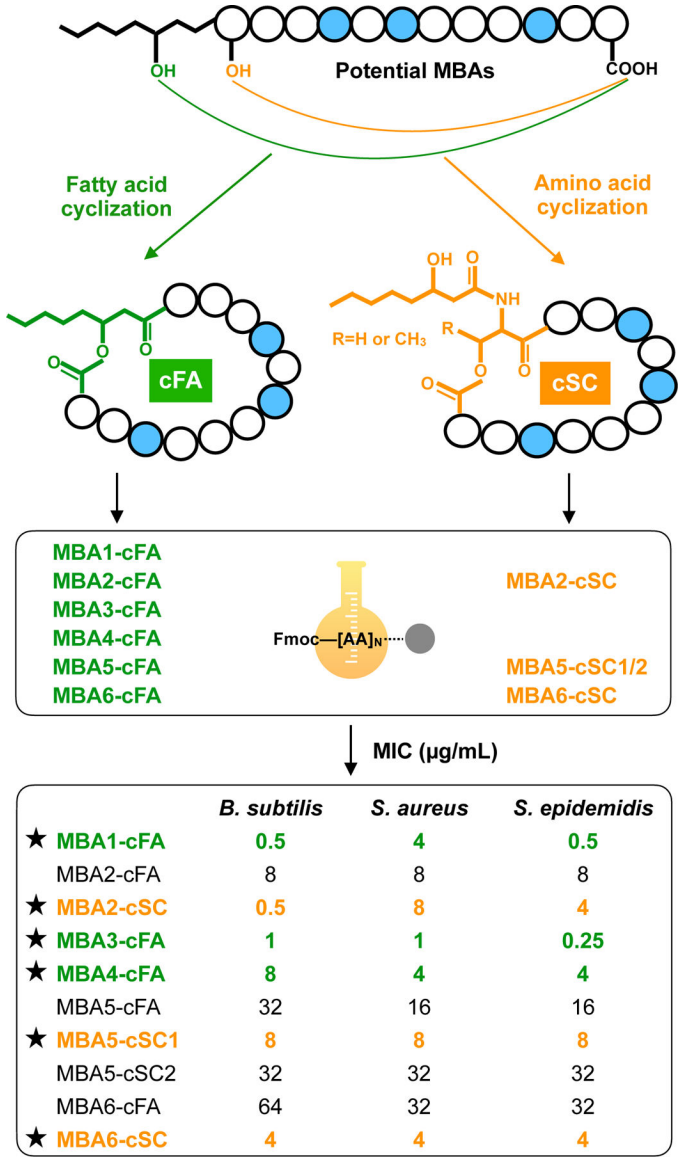


Fig. 1 |. Identification of BGCs predicted to encode new MBAs.

Using a sequence based soil metagenome BGC discovery pipeline, three BGCs were identified that show 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 predicted NRP (p-NRP) structures we identified three additional BGCs that we predicted would encode new MBAs. The GXLXXXW motif that 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.



	B. subtilis	*S. aureus*	*S. epidemidis*
★ 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 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 (*i.e.,* a serine or threonine) both the cFA and cSC analogs 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 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 assumed to be the “naturally cyclized” versions of the potential MBA. All MIC assays were done in duplicate (n=2).

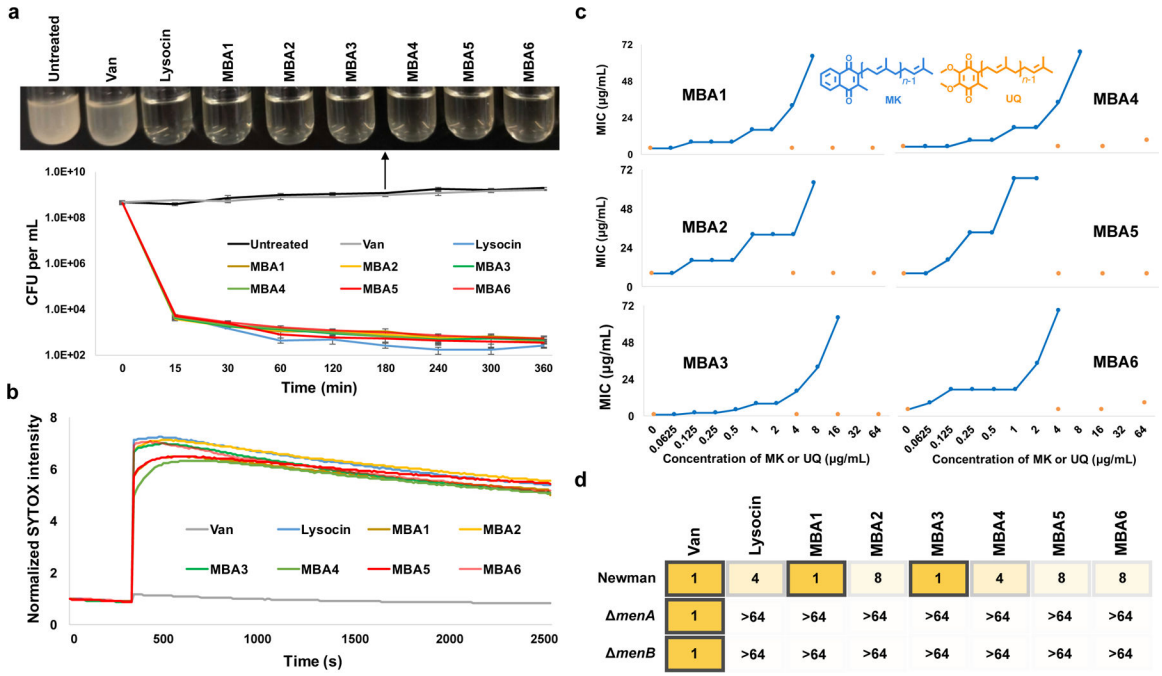


Fig. 3 |. Bactericidal effects and mode-of-action analysis of six new MBAs.

a, Bactericidal activity of MBA1–6 against *S. aureus* USA300. Cultures were incubated with each antibiotic at 2x its MIC. The number of viable cells was counted (n=3). Cultures were plated at defined times points to determine CFU per mL. **b**, Effect of MBA1–6 on *S. aureus* membrane lysis was determined using SYTOX fluorescence assay. **c**, The *S. aureus* antibacterial activity of MBA1–6 was determined in the presence of different concentrations of menaquinone (MK) (blue) or ubiquinone (UQ) (orange) (n=2). **d**, The MICs (μg/mL) of MBA1–6 against *S. aureus* mutants deficient in MK biosynthesis (Δ*menA* or Δ*menB)* (n=2). Lysocin and Van (vancomycin) were included as positive (MK binding) and negative (non-MK binding) antibiotic controls. The different shades of yellow represent MBA potency with darker color indicating higher activity.

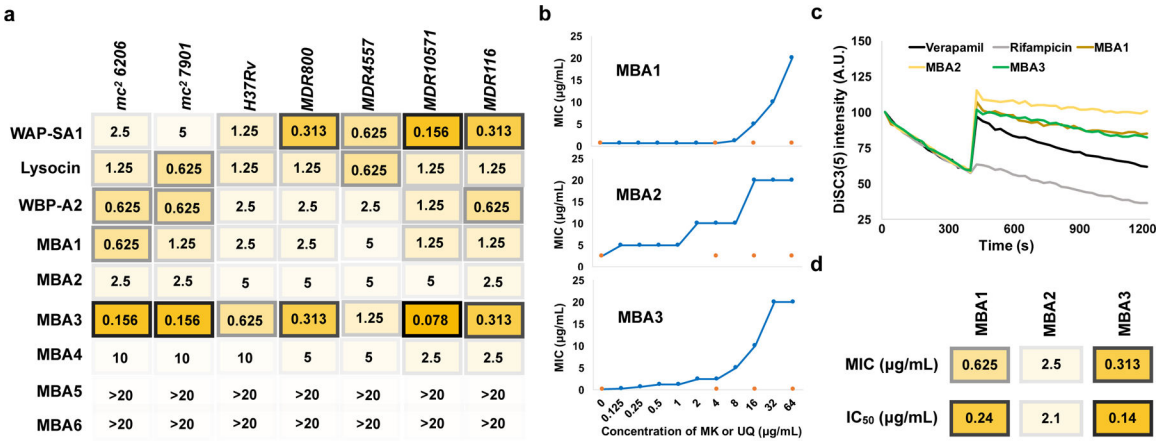


Fig 4. |. Anti-*Mtb* activity and mode of action analysis of new MBAs.

a, MIC (μg/mL) values of all known and new MBAs against a panel of *Mtb* strains, including two BSL2 strains (mc^2 6206 and mc^2 7901), a wild-type strain (H37Rv) and four multidrug-resistant (MDR) clinical isolates (n=2). The four MDR strains (800, 4557, 10571 and 116) are resistant to rifampicin, rifampicin, ethambutol/isoniazid/rifampicin/streptomycin and ethambutol/isoniazid/para-aminosalicylic acid, respectively. MIC in μg/mL. **b**, Effects of MK (blue) or UQ (orange) on the antibacterial activity of MBA1–3 against *Mtb* mc^2 6206 (n=2). **c**, The ability of *Mtb*-active MBAs to permeabilize the *Mtb* membrane was examined using a 3'-dipropylthiadicarboncyanine iodide [DiSC$_3$(5)] fluorescence assay. Verapamil and rifampicin were used as positive and negative depolarization controls, respectively. **d**, Activity of the three *Mtb*-active MBAs against *Mtb* mc^2 6206/mLux itself (MIC, n=2) and *Mtb* mc^2 6206/mLux in a macrophage infection assay (IC$_{50}$, n=2). The different shades of yellow represent MBA potency with darker color indicating higher activity.

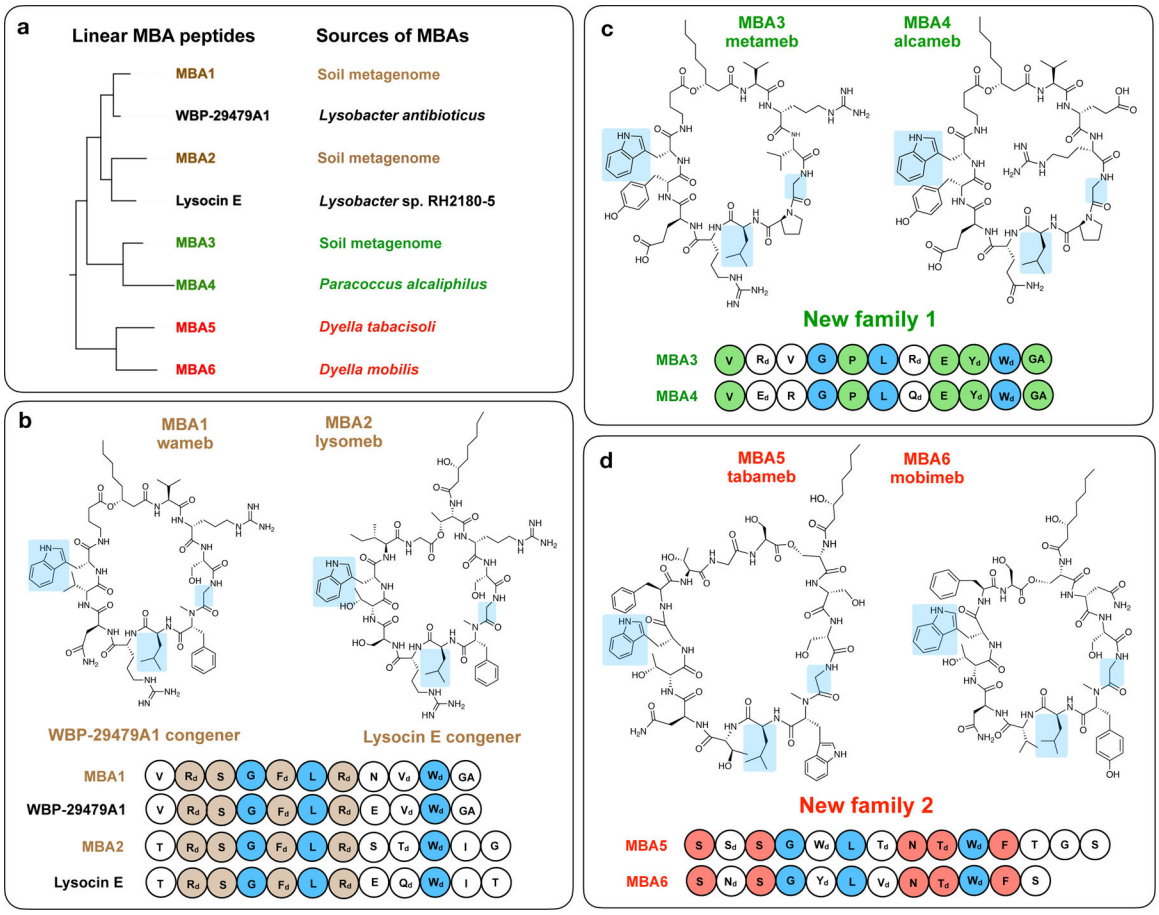


Fig. 5. |. Structures of six new MBAs grouped by structural family.

a, Phylogenetic tree of linear MBA peptide sequences. The branches on the tree are labeled with the name of the MBA and the source of its BGC. In this study we identified congeners of two known MBAs (**b**) as well as two new MBA structural families (**c**, **d**). All MBAs share the conserved GXLXXXW motif (blue) that is predicted to be the minimal sequence that is 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 acrosss the known MBA family (**b**), the first new MBA family (**c**) or the second new MBA family (**d**), respectively.

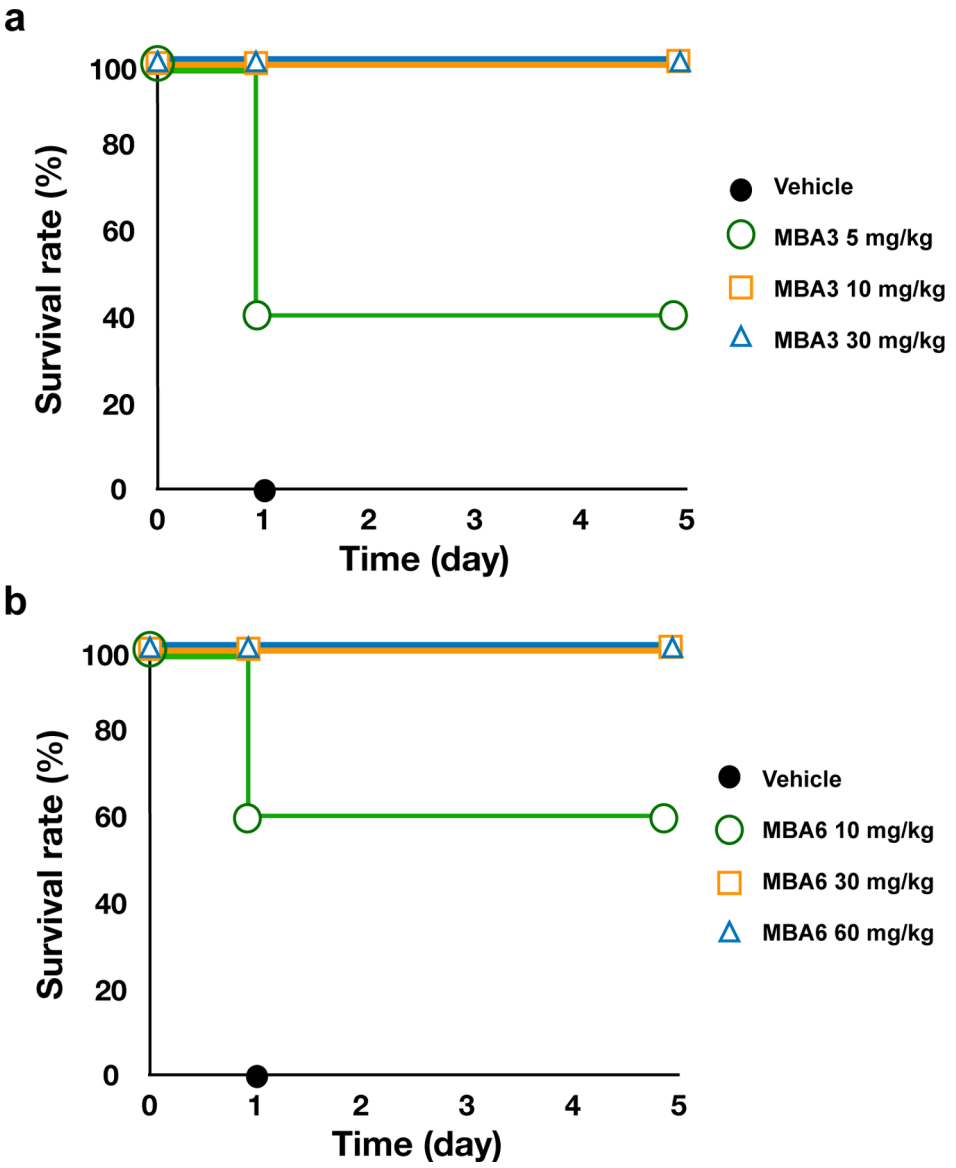


Fig 6. |. MBA3 (a) and MBA6 (b) are effective against *S. aureus* infections in mice.

Either MBA3 or MBA6 was subcutaneously injected 1 h after intraperitoneal administration of *S. aureus* COL into mice (n=5). 30% solutol was used as the vehicle.