Supplementary information

Computational identification of a systemic antibiotic for Gram-negative bacteria

In the format provided by the authors and unedited

Supplementary Information

Supplementary Notes

Cryo-EM MicroED

The compound crystallized in space group C2 with one molecule in the asymmetric unit, and the structure was solved by ab initio direct methods in SHELXT, followed by refinement in SHELXL⁶². The crystal packing is dominated by antiparallel β-sheet arrangements between adjacent molecules (**Extended Data Figure 3C**), with the polypeptide strands related by a crystallographic $2₁$ screw axis (parallel to the b axis of the crystal), and the directions of the strands oriented normal to this direction.

A critical question in the diffraction analysis was establishing the orientation of the His⁶ sidechain, as the crosslink with Tyr⁸ Cβ could involve either the side chain Nε2 or Cε2 atom. In view of the similarities of the electron scattering factors for these two elements, it was not possible to unambiguously distinguish these two possibilities through the crystallographic analysis. The assignment of the His 6 sidechain orientation supporting the role of Nε2 in the crosslink was based on comparison of the sidechain bond angles to those observed in small molecule structures in the Cambridge Structure Database⁸⁰. In this study, the extracyclic Cβ-Cv-Cδ and Cβ-Cγ-Nδ bond angles in the different protonated forms of the histidine side chain were found to be 130° and 123°, respectively; in the refined dynobactin A structure, these are 127.4° and 122.5˚, respectively. Flipping the ring orientation (corresponding to participation of the Cε atom in the crosslink) would interchange the Nδ and Cδ atoms and hence would be inconsistent with the small molecule bond angle observations.

The planes of the two macrocyclic rings are approximately perpendicular to one another. The polypeptide torsion angles of dynobactin are in the β-strand region of the Ramachandran plot, with the exceptions of S³ and N⁴ in the left- and right-handed α-helix regions, respectively. The Cα-Cα spacings between alternate residues in the extended conformation of dynobactin (residues 5-10) average ~6.5 Å, which is close to that observed in the darobactin structure (6.5 Å for residues 1-7); the most significant difference is the compressed Cα-Cα spacing between residues H 6 and Y 8 of dynobactin (5.95 Å) reflecting the smaller crosslink involving the H 6 sidechain relative to the longer crosslinks involving the W¹ and W³ sidechains in darobactin. In both dynobactin A and darobactin A, these crosslinks function to stabilize a pre-ordered βstrand conformation that can bind with high affinity to an exposed β-strand of the target protein. The final model contains the full dynobactin A decapeptide and 8 water molecules. As the net charge on dynobactin A is plausibly $+1$ (with a protonated N-terminal nitrogen and Arg⁹ sidechain, and a negatively charged C-terminal carboxylate), electroneutrality requires a negatively charged counterion that is likely distributed among the water molecules or disordered.

Secondary Structural Confirmations, Marfey's Analysis

We also confirmed the structure of dynobactin A using two well-established approaches for natural product small molecules. First, dynobactin A was acid-hydrolyzed to break all peptide bonds and to liberate individual amino acids. Following this, Marfey's reagent was used to derivatize, separate, and identify these residues⁸¹. In this case, other N-C connections were also broken by the hydrolysis, liberating tyrosine from histidine, making a C-C closure in the second cyclophane ring extremely unlikely. This showed that both the amino acids and their chirality was consistent with the proposed microED structure (Extended Data Figure 4).

Secondary Structural Confirmations, NMR

In addition, nuclear magnetic resonance studies (NMR) provided a structural assignment (Supplementary Table 4, Extended Data Figure 4) by the detailed 1D (Extended Data Figure 4A-B) and 2D NMR analysis including COSY, ROESY, TOCSY, ¹H-¹³C HSQC, and ¹H-¹³C HMBC in both DMSO- d_6 and D₂O solvents (**Extended Data Figure 4C**). In the DMSO- d_6

sample, 15 spin systems were identified from COSY and TOCSY spectra, including 10 from the amino acid backbone, and 2 from the tryptophan side chain, two overlapping spin systems from the tyrosine side chain and one from the phenylalanine side chain. The α- and β-carbon and proton chemical shifts in each amino acid residue were identified by phase-sensitive ¹H-¹³C HSQC experiments. Among the 10 amino acid backbone spin systems, eight showed α-protons connecting to a methylene, two showed α-protons connected to another methine. The D_2O sample provided additional ¹H-¹³C HMBC correlations and the key 2D correlations are summarized in (**Extended Data Figure 4D**). The three aromatic side chains of tryptophan, tyrosine and phenylalanine were readily established by COSY and HMBC correlations. H-5 from the tryptophan side chain showed $3J$ HMBC correlations to C-3 connecting the tryptophan side chain to the backbone spin system. Both H-13 and H-14 had an HMBC correlation to two carbonyl carbons C-12 and C15, establishing an asparagine moiety. H-2, H-3 and H-13 shared the same HMBC correlations to the carbonyl carbon C-1, connecting the tryptophan moiety and the asparagine moiety. Two spin systems H-17 and H-18, H-35 and H-36 had similar carbon and proton chemical shifts, and they were assigned as two serine side chains based on their characteristic β-carbon chemical shifts. Both H-17 and H-18 shared an HMBC correlation to C-16, while H-35 had an HMBC correlation to C-28. A second asparagine was identified based on HMBC correlations from H-20 to two carbonyl carbons C-19 and C-22; H-21 also has ²J correlations to C-22. Interestingly, the β-carbon in the spin system was a methine instead of a methylene group based on phase sensitive HSQC signals, which suggested a substitution at the C-21 position. The HMBC $3J$ correlations between H-7 and H-9 to C-21, and the $2J$ HMBC correlation from H-21 to C-8 suggested the connectivity between C-8 and C-21, connecting the β-carbon of the asparagine to the side chain of tryptophan. The valine spin system was characterized by two methane signals; H-26 and C-27 shared a COSY correlation to a methine proton, C-25. The α-proton showed an HMBC correlation to C-19, connecting the valine moiety to the substituted asparagine. Both H-24 and H-25 shared an HMBC correlation to the carbonyl carbon C-23. The histidine moiety was characterized by two aromatic methine groups at positions 33 and 32, where H-33 had an HMBC correlation to C-32. The tyrosine side chain was connected to its backbone, shown by HMBC correlations from H-39 to C-41 and C-40, as well as from H-38 to C-40. The α-proton of tyrosine had HMBC correlations to two carbonyl carbons C-34 and C-37. Interestingly, the β position of tyrosine showed a methine group, instead of a methylene group, as evidenced by the phase-sensitive HSQC, suggesting a substitution at the β position. The key HMBC correlation from H-39 to C-32 established the connectivity between C-39 and the histidine side chain. In particular, the characteristic ¹³C NMR chemical shift of C-39 (δ 63.4) strongly suggested carbon-nitrogen linkage. In addition, ROESY correlations between H-32 and H-38, as well as between H-32 and H-41 further confirmed the macrocyclic connection between tyrosine β-carbon, C-39, and the histidine side chain. Due to the HSQC correlation of both C-33 and C-32, C-39 must connect to a histidine side chain nitrogen atom. The assignment between C-39 and the nitrogen between C-32 and C-33 was based on cryo-EM microED 3D structure, indicating connection to an atom in the ε position on the histidine side chain, here meaning Nε2. The spin system consisting of C-47, C-48, C-49, and C-50 was assigned as arginine based on the characteristic methylene proton and carbon chemical shifts at C-50. In the end, HMBC correlations between H-54 to H-55 and H-56 established the phenylalanine moiety, and an HMBC correlation was observed between H-54 to the terminal carbonyl carbon C-52. The two serine side chain regions had limited HMBC correlations and the assignment of the backbone spin systems could be switched. The last backbone spin system consisting of C-29 and C-30 has to be assigned to histidine to complete the full structure, although no HMBC correlations between the side chain and the backbone was observed.

Supplementary Figure 1. Cryo-electron microscopy structure of the BAMdynobactin A complex.

(A) Data processing flowchart to generate the cryo-EM map (see Materials & Methods). (B) Representative electron micrograph of BAM-dynobactin A (C) Examples of 2D classes from cryoSPARC. (D) Viewing direction distribution plot for the final 3D reconstruction. (E) Fourier shell correlation (FSC) curves for unmasked, spherical, loose, and tight masks, and corrected FSC curve for the final reconstruction, yielding a gold standard FSC resolution of 3.6 Å. (F) Local resolution variations of the EM reconstruction. POTRA domains 1 and 2 are at a local resolution below 5 Å and are visualized only at a lower contour level. (G) Directional FSC plot (red; mean \pm SD) and histogram of per angle FSC (blue). The grey dotted line shows the FSC of 0.143, indicating a resolution of 3.75 Å. (H) Overview of the cryo-EM reconstruction of the BAM-dynobactin A complex. BAM is shown in ribbon representation and the coulomb potential map as slate blue mesh. Close-Up views showing the map around selected atoms in stick representation of (I) the BamA barrel and (J) dynobactin A.

Supplementary Figure 2. Inhibition of BAM-mediated OMP folding by darobactins and dynobactin A.

Fluorescence emission of BAM-mediated OmpT activity in presence of (A) darobactin A, (B) darobactin B, and (C) dynobactin A at variable compound concentration as indicated. From these data, the initial reaction rates were determined and plotted against the compound concentration to determine the IC₅₀ values with a 95% confidence interval (Extended Data Figure 6E). (D) IC₅₀ value and 95% confidence intervals, as determined by the BAM- mediated OmpT folding assay. (E) Significance test for shared IC_{50} for a pair of datasets implemented in GraphPad Prism with an extra sum-of-squares F-test. In both cases, the IC₅₀ shows a significant difference between the datasets. Exact p-values for dynobactin A versus darobactin A and B were reported as 0.000000074 and 0.000000000000013, respectively.

Supplementa**ry** Tables

Supplementa**ry** Table 1. Identified putative RTEs and Propeptides.

Classifies identified RiPPs into clades based upon SPASM motif. Includes NCBI accessions, an example source organism, and partial propeptide sequence. Predicted RiPP core sequence is bolded.

Supplementa**ry** Table 2. Rapid compound identification mass table.

List of observed masses present in active wells from fractionated supernatant are listed for each separation condition. Masses conserved across all conditions are indicated in yellow highlight.

Compound

Supplementa**ry** Table 3. MICs for dominant antimicrobial metabolites from P. australis

a Values at parentheses are for the outer shell

Supplementa**ry** Table 4. Cryo-EM microED data collection and refinement statistics for dynobactin A.

^aChemical shifts can be exchangeable.

Supplementary Table 5. ¹H and ¹³C NMR (900/225 MHz) chemical shift in D₂O.

Supplementa**ry** Table 6. Cryo-EM data collection and refinement statistics of BAM-dynobactin A complex.

Supplementa**ry** Table 7. X-ray diffraction data and refinement statistics of BamA-β/dynobactin A complex.

a not defined, because association followed a bimodal kinetic

Supplementa**ry** Table 8. SPR measurements of BamA and ligands.