



Bicyclostreptins are radical SAM enzyme-modified peptides with unique cyclization motifs

Leah B. Bushin¹, Brett C. Covington¹, Kenzie A. Clark¹, Alessio Caruso¹ and Mohammad R. Seyedsayamdost^{1,2}✉

Microbial natural products comprise diverse architectures that are generated by equally diverse biosynthetic strategies. In peptide natural products, amino acid sidechains are frequently used as sites of modification to generate macrocyclic motifs. Backbone amide groups, among the most stable of biological moieties, are rarely used for this purpose. Here we report the discovery and biosynthesis of bicyclostreptins—peptide natural products from *Streptococcus* spp. with an unprecedented structural motif consisting of a macrocyclic β -ether and a heterocyclic sp^3 - sp^3 linkage between a backbone amide nitrogen and an adjacent α -carbon. Both reactions are installed, in that order, by two radical S-adenosylmethionine (RaS) metalloenzymes. Bicyclostreptins are produced at nM concentrations and are potent growth regulation agents in *Streptococcus thermophilus*. Our results add a distinct and unusual chemotype to the growing family of ribosomal peptide natural products, expand the already impressive catalytic scope of RaS enzymes, and provide avenues for further biological studies in human-associated streptococci.

Microbial natural products have been indispensable as a source of therapeutic leads and a driver of innovation in diverse scientific disciplines^{1–4}. Until the turn of the century, natural product discovery was largely guided by bioactivity-based, so-called ‘grind-and-find’ approaches⁵. Since then, advances in DNA sequencing technologies and the development of bioinformatic methods have provided a paradigm shift from activity-first to gene-first search strategies^{6,7}. Although powerful and successful, one of the major limitations of the gene-first approach is the bias toward what constitutes a biosynthetic gene. Because prior precedent informs the selection of biosynthetic gene clusters (BGCs) that are pursued, the gene-first approach has often yielded new compounds within already known compound families. Despite some notable exceptions⁸, the discovery of entirely new natural product chemotypes using the approach has lagged behind.

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are especially well-suited for the gene-first approach^{9–11}. In RiPP biosynthesis, a precursor peptide is synthesized by the ribosome, modified by generally a small number of tailoring enzymes, and usually trimmed at the final step to deliver the mature natural product¹¹. Importantly, the substrate for the modification enzyme is genetically encoded, thereby providing a key first step in piecing together the biosynthetic pathway of the final product. To minimize the chances of finding known RiPP compound families, we recently conducted a bioinformatic search wherein we identified genetic loci that encode a RiPP precursor peptide and a RaS enzyme adjacent to a quorum sensing (QS) operon^{12,13}. We reasoned that the selection of RaS enzymes—the most versatile enzyme superfamily with over 500,000 largely unstudied members^{14–16}—would provide a high likelihood of undiscovered biosynthetic chemistry and that regulation by QS, which controls important microbial community behaviors such as virulence factor production¹⁷, would ensure physiological relevance. The operons observed were then organized into a sequence similarity network based on the sequence of the precursor peptide (Fig. 1a)^{12,18}. The search revealed ~600 RaS-RiPP gene

clusters in streptococci alone, suggesting that this genus, which was not known for its prowess as a secondary metabolite producer, is indeed a deep source of potentially novel natural products¹².

Our subsequent analysis validated this conclusion, as we found several novel metalloenzyme-mediated transformations, including formation of carbon–carbon and carbon–heteroatom crosslinks at unactivated positions (Fig. 1a)^{19–21}. Likewise, examination of the mature natural products revealed unexpected structures and biological activities, including streptosactin²², tryglysin A²³ and streptide (Fig. 1b)²⁴. Together, these and other discoveries, such as identification of darobactin A (Fig. 1b)²⁵, polytheonamide²⁶ and xenorceptide²⁷, demonstrate that RaS-RiPPs form a broad, architecturally diverse family of natural products ripe for further exploration. The RaS-RiPPs uncovered so far are usually cyclized by a single enzyme. Biosynthetic operons that encode two or more RaS enzymes could give rise to yet more complex peptide scaffolds. Three such operons that are so far uncharacterized can be observed in our original network; they contain conserved HGH, SSH and GRC motifs in their precursor peptides (Fig. 1a)¹².

In the current article, we examine one of these clusters that codes for two RaS enzymes and a possible RiPP recognition element (RRE)²⁸. We show that both RaS enzymes catalyze novel post-translational modifications in a distinct order, resulting in the formation of a uniquely fused bicyclic scaffold containing a β -ether-linked macrocycle adjacent to an imidazolidine-4-one group. This heterocycle is formed by crosslinking the backbone amide nitrogen to an unactivated carbon—the first instance of radical chemistry toward activation of an amide bond in RiPP biosynthesis. Knowledge of the reactions of the RaS enzymes allowed us to isolate the mature natural product, which we term bicyclostreptin, from culture supernatants of the probiotic *S. thermophilus*. The strategy proved essential as bicyclostreptin is produced at nM concentrations and is difficult to identify otherwise. Bicyclostreptins harbor narrow-spectrum, bacteriostatic activity against select *S. thermophilus* strains, including the producing host.

¹Department of Chemistry, Princeton University, Princeton, NJ, USA. ²Department of Molecular Biology, Princeton University, Princeton, NJ, USA.

✉e-mail: mrseyed@princeton.edu

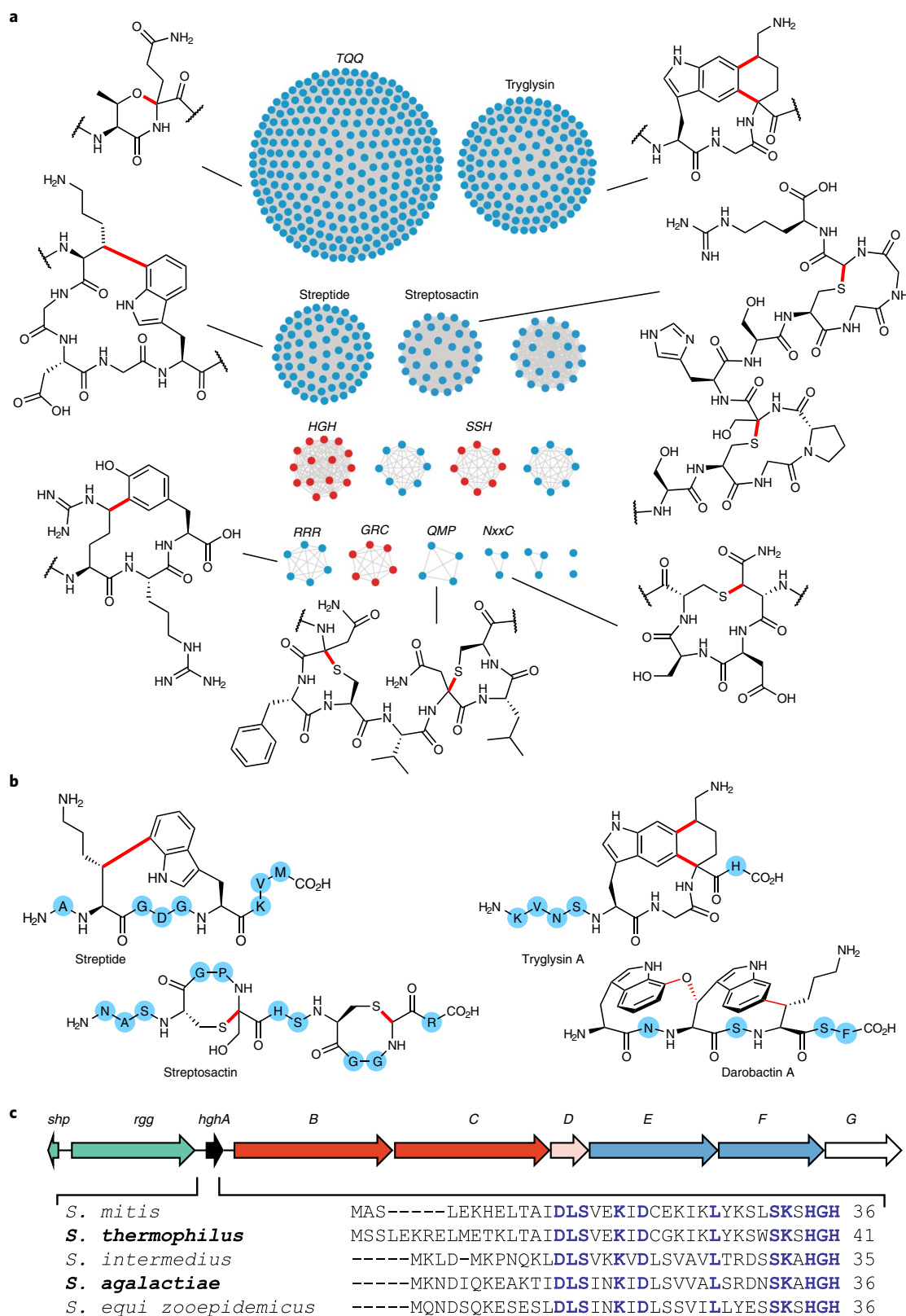


Fig. 1 | RaS-RiPP network in streptococci. a, Sequence similarity network of RaS-RiPPs with each node representing a unique BGC. Bonds installed by RaS enzymes are shown in red. Subfamilies are named based on conserved motifs in the precursor peptide, except streptide, streptosactin and tryglysin, for which the mature products have been identified. Those that encode multiple RaS enzymes are shown in red. **b**, Structures of mature RaS-RiPPs. Blue spheres represent unmodified amino acids with the indicated one-letter code. The biosynthesis of darobactin A awaits further investigation. **c**, The *hgh* gene cluster is controlled by an *shp/rgg* QS operon (green) and codes for a precursor peptide (HghA), two RaS enzymes (HghBC), an RRE (HghD), two protease transporters (HghEF) and a hypothetical protein (HghG). HghA sequences from five different streptococci are shown, with conserved residues highlighted in blue.

Macrocyclization via β -etherification and backbone alteration to install a carbon–nitrogen sp^3 – sp^3 bond provide new paradigms for post-translational modifications. More broadly, the results provide further impetus for targeting RiPP gene clusters that encode multiple RaS enzymes.

Results

The HGH subfamily. Several gene clusters in our network, including *hgh*, *grc* and *ssh*, encode multiple RaS enzymes, suggesting a more complex mature product¹². We focused on the *hgh* locus, which codes for a precursor peptide (HghA), two RaS enzymes (HghB and HghC), a possible discrete RRE (HghD), two transporters (HghE and HghF) and a hypothetical membrane-associated protein (HghG). An *shp/rgg* QS operon is located directly upstream (Fig. 1c). In our network, the *hgh* locus occurs in species that constitute the normal human flora, including the probiotic *S. thermophilus*, the commensal strain *Streptococcus intermedius*, as well as opportunistic pathogens *Streptococcus agalactiae*, *Streptococcus mitis* and *Streptococcus equi*. Each species encodes a distinct precursor peptide as shown (Fig. 1c), but all carry the conserved HGH C-terminal motif, for which the subfamily is named¹². We focused on *hgh* from *S. thermophilus* JIM 8232, isolated from milk, and *S. agalactiae* MRI Z1-218, isolated from bovine mastitis^{29,30}.

HghB and HghC are active RaS enzymes in vivo. We initially chose an in vivo strategy to characterize the reactions carried out by the RaS enzymes HghB and HghC. The precursor peptide HghA was expressed heterologously in *Escherichia coli* with an N-terminal hexaHis maltose-binding protein (6HMBP) purification tag. An HRV3C cleavage site was placed between the tag and *hghA* (Supplementary Tables 1–3). Also encoded on the vector along with *hghA* were *hghB*, *hghC*, *hghB* and *hghC*, or neither. The possible RRE, *hghD*, was included whenever *hghB* or *hghC* were present. Upon expression, the reaction was allowed to proceed within the confines of the *E. coli* cytosol, after which the modified peptide was purified. The tag was removed with HRV3C proteolysis and the structure of the product was elucidated with high-resolution mass spectrometry (HR-MS), tandem HR-MS (HR-MS/MS), and 1D/2D nuclear magnetic resonance (NMR) spectroscopy.

The construct with *S. agalactiae* HghA alone gave the expected unmodified, linear *m/z* after cleavage with the protease ($[M+5H]^{5+}_{\text{obs}}$ 761.40495, $[M+5H]^{5+}_{\text{exp}}$ 761.39980; Fig. 2a and Supplementary Table 4). The HghA *m/z* did not change upon co-expression with HghBD, implying that HghB does not modify the linear precursor (Fig. 2b). Co-expression with HghCD, however, yielded a product that was 2 Da lighter than substrate, suggesting that a modification had occurred (Fig. 2c). In the presence of HghBCD—that is, both RaS enzymes and the possible RRE—the peptide product was 4 Da lighter (Fig. 2d and Supplementary Table 4). These results were mirrored with the *S. thermophilus* systems (Supplementary Table 5), suggesting that, in both cases, HghC modifies HghA first, thereby generating an intermediate that is subsequently acted upon by HghB.

Initial insights regarding the sites of modification were provided by HR-MS/MS (Extended Data Fig. 1a). All b-ions generated in the collision-induced dissociation of the *S. agalactiae* HghCD product appeared unmodified except for those C-terminal to His34. Concurrently, y-ions N-terminal to His34 were 2 Da lighter than the same fragments derived from the unmodified substrate (Extended Data Fig. 1a and Supplementary Tables 6 and 7). The pattern implicated His34 within the $H^{34}G^{35}H^{36}$ motif in the modification installed by the RaS enzyme HghC. With the HghBCD product, y-ions N-terminal to the HGH motif were now 4 Da lighter, suggesting that HghB may further modify the conserved region (Extended Data Fig. 1a and Supplementary Table 8). HR-MS/MS carried out with the *S. thermophilus* products provided equivalent results, with

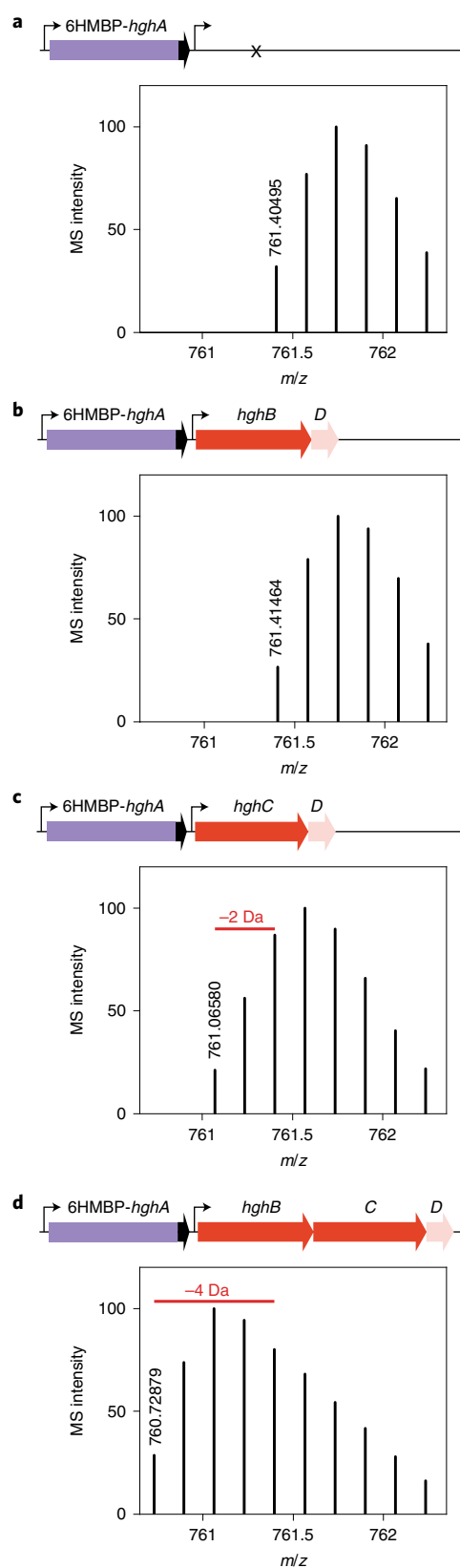


Fig. 2 | HR-MS analysis of heterologous expression constructs indicated.

Shown are data for HghA after co-expression with no other enzymes as control (a), with HghBD (b), with HghCD (c) or with HghBCD (d). The construct is shown above each spectrum, which is zoomed in on the $[M+5H]^{5+}$ ion. Co-expression with HghCD and HghBCD gives product that is –2 Da and –4 Da relative to substrate, respectively.

modifications localized to residues H³⁹G⁴⁰H⁴¹ (Supplementary Tables 9 and 10).

To simplify structural elucidation by NMR, we sought to reduce the length of each product by removing the unmodified portions of HghA. Digestion of *S. agalactiae* and *S. thermophilus* HRV3C-cleaved HghCD and HghBCD products with trypsin yielded C-terminal 8mer peptides. These truncated peptides maintained the same mass losses and exhibited fragmentation patterns analogous to those of their full-length counterparts (Supplementary Tables 11–14). Surprisingly, proteolysis did not occur at the final Lys residue for either product, but it did take place for the unmodified HghA peptides (Extended Data Fig. 1b). This difference suggested that a residue N-terminal to Lys32 in *S. agalactiae* (and Lys37 in *S. thermophilus* HghA) may be modified and that structural changes imposed by the modification may interfere with trypsinolysis.

Structural elucidation of the products of HghB and HghC.

To elucidate the product of HghC from *S. agalactiae*, the reactions were carried out on a large scale, and the cleaved 8mer product was isolated and subsequently analyzed by 1D/2D NMR spectroscopy. The corresponding, unmodified linear 8mer peptide (H₂N-DNSKAHGH-CO₂H) was prepared by solid-phase peptide synthesis (SPPS) to aid in comparative spectral analysis (Supplementary Table 15 and Supplementary Figs. 1 and 2). A DEPT-edited heteronuclear single quantum coherence (HSQC) spectrum of the HghCD product revealed a new methine group at 71.8 ppm (¹³C) and 4.97 ppm (¹H) as well as disappearance of one of the β-methylene ¹Hs of His6 (Fig. 3, Supplementary Figs. 3 and 4 and Supplementary Table 16), implying that the β-methylene had been converted to a methine. Its sizable ¹³C downfield shift from 26.2 ppm to 71.8 ppm suggested that the β-carbon is engaged in a heteroatom bond. The remaining β-proton of His6 showed HMBC correlations to Cα, C4 and C5 of the imidazole ring and to a carbon with a chemical shift of 67.2 ppm (Supplementary Figs. 3 and 4), which we traced to the β-position of Ser3 (corresponding to Ser31 in Extended Data Fig. 1b), establishing a covalent linkage between the two sidechains (Fig. 3). More specifically, the β-carbon of His6 participates in an ether bond with the hydroxyl-oxygen of Ser3. The entrapment of Lys4 within the macrocycle likely explains the lack of trypsin cleavage at this position. The Ser3 residue is entirely conserved among HghA precursor peptides (Fig. 1c).

We next turned to the structure of the HghBCD product from *S. agalactiae*. Correlations related to the β-ether crosslink were still present, indicating that the modification installed by HghCD remained intact (Supplementary Figs. 5 and 6). Peaks associated with the methylene of Gly7, however, were absent, suggestive of a reaction at the Gly7 α-carbon (Fig. 3). A new peak in the HSQC spectrum at 69.8/5.70 ppm pointed to conversion of the Gly7 α-methylene to a heteroatom-linked methine. An unusual TOCSY correlation between the α-¹Hs of His6 and Gly7 established a second covalent linkage between the two residues in addition to the existing amide bond (Supplementary Figs. 5 and 6). Notably, the His6 amide-¹H and its 2D correlations were absent, consistent with a connection between the amide nitrogen and the Gly α-carbon to form an unusual imidazolidine-4-one. The observed ¹H and ¹³C shifts were consistent with this motif (Supplementary Table 16). Through-space ROESY correlations between the Ala5 amide-¹H and α-¹H to the remaining Gly7 α-¹H further substantiated the presence of the new carbon–nitrogen bond (Fig. 3).

NMR data collected on the HghBCD product from *S. thermophilus* were analogous to those described and consistent with the crosslinks observed with the *S. agalactiae* operon (Supplementary Fig. 7). Together, the data indicate that HghC installs an ether crosslink between the sidechains of Ser31 and His34 (Ser36 and His39 in *S. thermophilus*), giving rise to a 4-residue, 13-membered macrocycle. A β-ether crosslink is a novel reaction for the RaS enzyme

superfamily and, thereby, expands its already impressive reaction scope^{14–16}. The mono-crosslinked product serves as the substrate for HghB, which installs an unprecedented imidazolidine-4-one within the peptide backbone. The HghB-catalyzed heterocyclization via the backbone nitrogen is unprecedented; as an unusual *sp*³–*sp*³ crosslink, it highlights the unique catalytic repertoire of RaS enzymes in activating both an amide-¹H and an amino acid α-¹H.

HghC and HghB are active RaS enzymes in vitro. Having analyzed the reactions of HghC and HghB in vivo, we next sought to reconstitute their activities in vitro. Both enzymes and HghD were expressed in *E. coli* as N-terminal His₆-tagged constructs and isolated via metal affinity chromatography, under anaerobic conditions in the case of HghC and HghB. HghA was expressed with a 6HMBP-tag, and the tag was removed via proteolysis. The purified enzymes exhibited UV-vis absorption features—a broad absorption band centered at 400 nm and a shoulder at 320 nm—indicative of the presence of [4Fe-4S]²⁺ cluster(s) (Fig. 4a,b). Assessment of Fe and labile S²⁻ revealed 4.2 ± 0.1 Fe and 5.7 ± 0.1 S²⁻ per HghC protomer and 4.1 ± 0.1 Fe and 5.8 ± 0.1 S²⁻ per HghB protomer, consistent with the UV-vis spectra. Notably, in the absence of substrate (or HghD), both enzymes generated 5'-dA, when supplied with cofactor SAM (Extended Data Fig. 2). This futile SAM cleavage reaction is diagnostic for RaS enzymes and indicates that both enzymes are competent in homolytic, reductive cleavage of the cofactor.

When HghC was supplemented with HghA, HghD, reductant and SAM, we observed time-dependent formation of a –2-Da product and formation of 5'-dA (Fig. 4c, Extended Data Fig. 3 and Supplementary Table 17). The product was not observed in the absence of HghC, HghA, SAM or reductant. Tandem HR-MS data on this product were entirely consistent with the macrocyclic β-ether linkage between Ser31 and His34 with the fragmentation pattern mirroring that of the product observed in vivo (Fig. 4e, Extended Data Fig. 1a and Supplementary Table 18). When the reaction was supplemented with HghB, we observed formation of the –4-Da product, relative to HghA (Fig. 4d). Again, the HR-MS and HR-MS/MS data mirrored those obtained for the product in vivo (Fig. 4f, Extended Data Fig. 1a and Supplementary Tables 17 and 19). HghB and HghC are, thus, active RaS enzymes that catalyze novel reactions, the mechanisms of which can be investigated in the future. Interestingly, in the absence of HghD, both enzymes still modified their substrates, and the product yields were unaffected (Fig. 4c,d). The size of HghD suggested that it could act as an RRE²⁸. However, upon closer analysis, it does not show substantial sequence homology to PqqD, the archetypal RRE domain (Extended Data Fig. 4); the role of HghD remains to be determined.

Discovery of the *hgh* mature product bicyclostreptin. With knowledge of the modifications installed by HghB and HghC, we attempted to identify the mature product of the *hgh* operon from culture supernatants. We first focused on *S. thermophilus* as we previously discovered two other RaS-RiPPs under the regulation of *shp/rgg* operons from this species^{22,24}. Based on the assumption that doubly crosslinked HghA would be proteolyzed at a residue upstream of modified Ser36, a list of possible mature product *m/z* values were assembled computationally. Supernatants from *S. thermophilus* cultures were cleared by solid-phase extraction and subjected to HR-MS analysis. Metabolites with appropriate masses were then characterized by HR-MS/MS. Two of the analytes with sequences corresponding to the C-terminal 7 (H₂N-WSKSHGH-CO₂H) or 8 (H₂N-SWSKSHGH-CO₂H) residues, produced in a 1:2.5 ratio, exhibited fragmentation patterns consistent with modifications installed by HghBCD (Supplementary Tables 20 and 21). The abundance of the peptides, however, was prohibitively low to confirm their identities via isolation and NMR-based structural elucidation. Fortunately, with the 8mer candidate also being the product

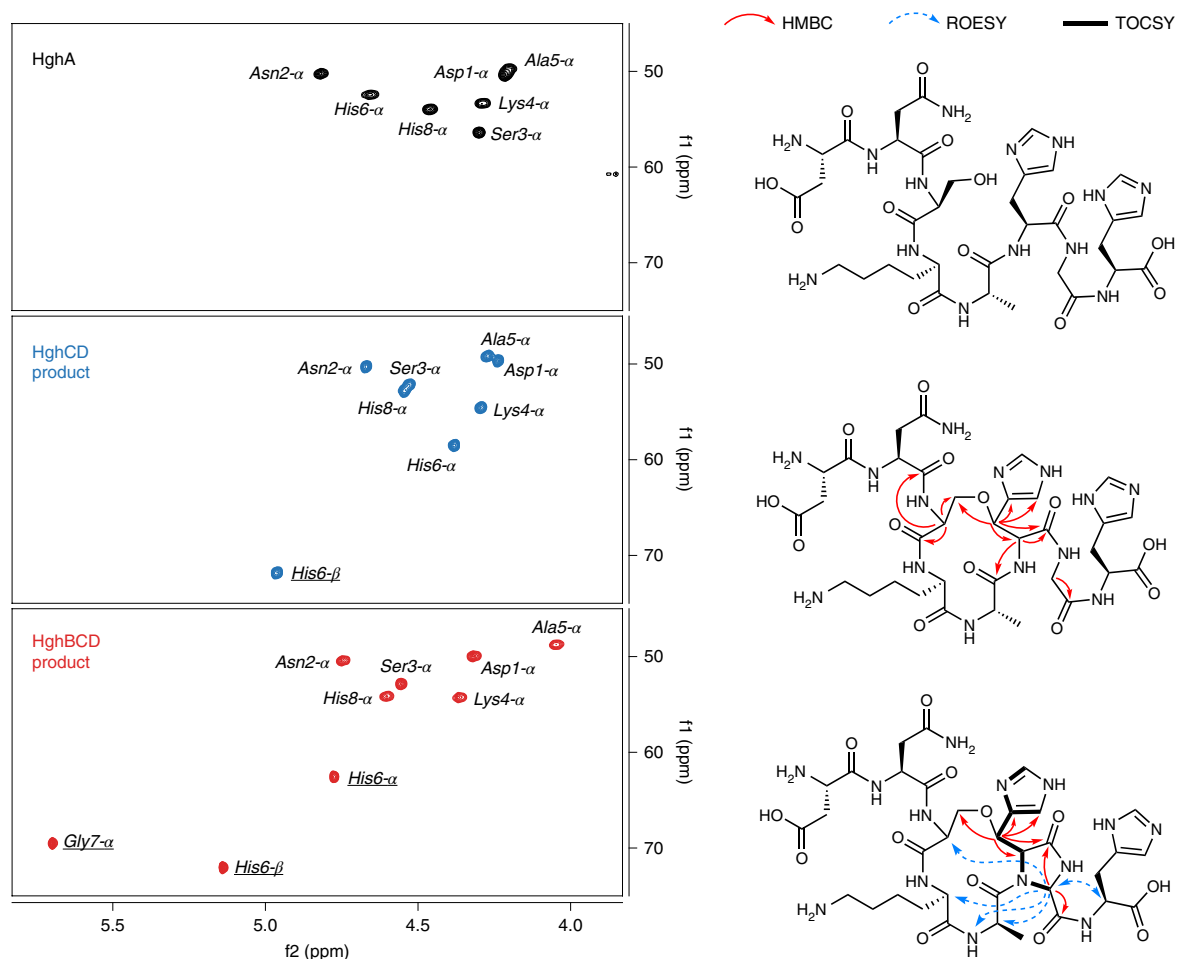


Fig. 3 | Structural elucidation of the products of *S. agalactiae* HghBD and HghBCD. Left, comparative analysis of HSQC spectra of synthetic linear C-terminal 8mer (top), HghCD product (middle), and HghBCD product (bottom). All observed crosspeaks are labeled; key correlations are underlined. Right, relevant NMR correlations used to solve the structure of each 8mer. Absolute configurations of the two newly formed chiral centers remain to be determined.

of trypsin cleavage, a synthetic standard was already at hand. The HR-MS/MS data of the trypsin-cleaved 8mer HghBCD product perfectly matched that of the putative authentic product, and the two compounds exhibited identical chromatographic behavior, as a sample containing a one-to-one mixture of the two compounds gave a single chromatographic peak (Fig. 5a). These results establish the bicyclic 8mer, which we named bicyclostreptin A (1), as one of the products of the *S. thermophilus* *hgh* cluster (Fig. 5b). The less abundant 7mer, bicyclostreptin B (2), is likely the result of an alternative proteolytic route. One possibility is that each of the transporters in the *hgh* cluster, HghE and HghF, cleaves at a distinct residue, generating two unique products.

Using the strategy outlined above, we also searched supernatants of *S. agalactiae* MRI Z1-218 for the corresponding orthologous product but did not find any analytes matching the predicted mass, suggesting that the mature product is not synthesized or is below our limit of detection under these conditions. Based on the sequence alignment of the precursor peptides (Fig. 1c), we suspect that the *S. agalactiae* *hgh* cluster generates an 8mer mature peptide, bicyclostreptin C (3) (Fig. 5b), with the crosslinks identified above. Overall, identification of the mature natural products points to a biosynthetic pathway in which two successive RaS enzyme-catalyzed reactions build the bicyclostreptin scaffold, with the precursor peptide being acted upon by HghC to generate the monocyclic β -etherified product, followed by installation of the

unusual imidazolidine-4-one heterocycle by HghB. Finally, the proteases HghE and/or HghF produce the 7mer and 8mer mature products (Extended Data Fig. 5).

Quantification of bicyclostreptin A and B using a standard curve gave very low bulk concentrations in the nM range (Fig. 5c). The low titers provide a rationale for the low isolation yields, which precluded de novo structural elucidation of the authentic product. They are produced in a narrow time window during mid-exponential growth phase. Although we did not detect any membrane-associated or intracellular bicyclostreptins, it remains possible that bicyclostreptins are imported but beyond our lower limit of detection. Alternatively, they could be degraded and, therefore, undetectable beyond mid-exponential phase. Interestingly, in *S. thermophilus*, this growth phase is associated with production of competence stimulating peptides along with possible fratricidal molecules, which kill non-competent sibling cells^{22,31–34}. These features provided some clues toward bicyclostreptin's biological activity, which we explored next.

Growth-regulatory effects by bicyclostreptins. The biological properties of streptosactin and tryglysin have been surprising, both revealing narrow-spectrum, growth-inhibitory activities^{22,23}. With these features and the limited supply of bicyclostreptins in mind, we first tested the activity of bicyclostreptin C against three *S. thermophilus* strains, three *S. agalactiae* strains and several

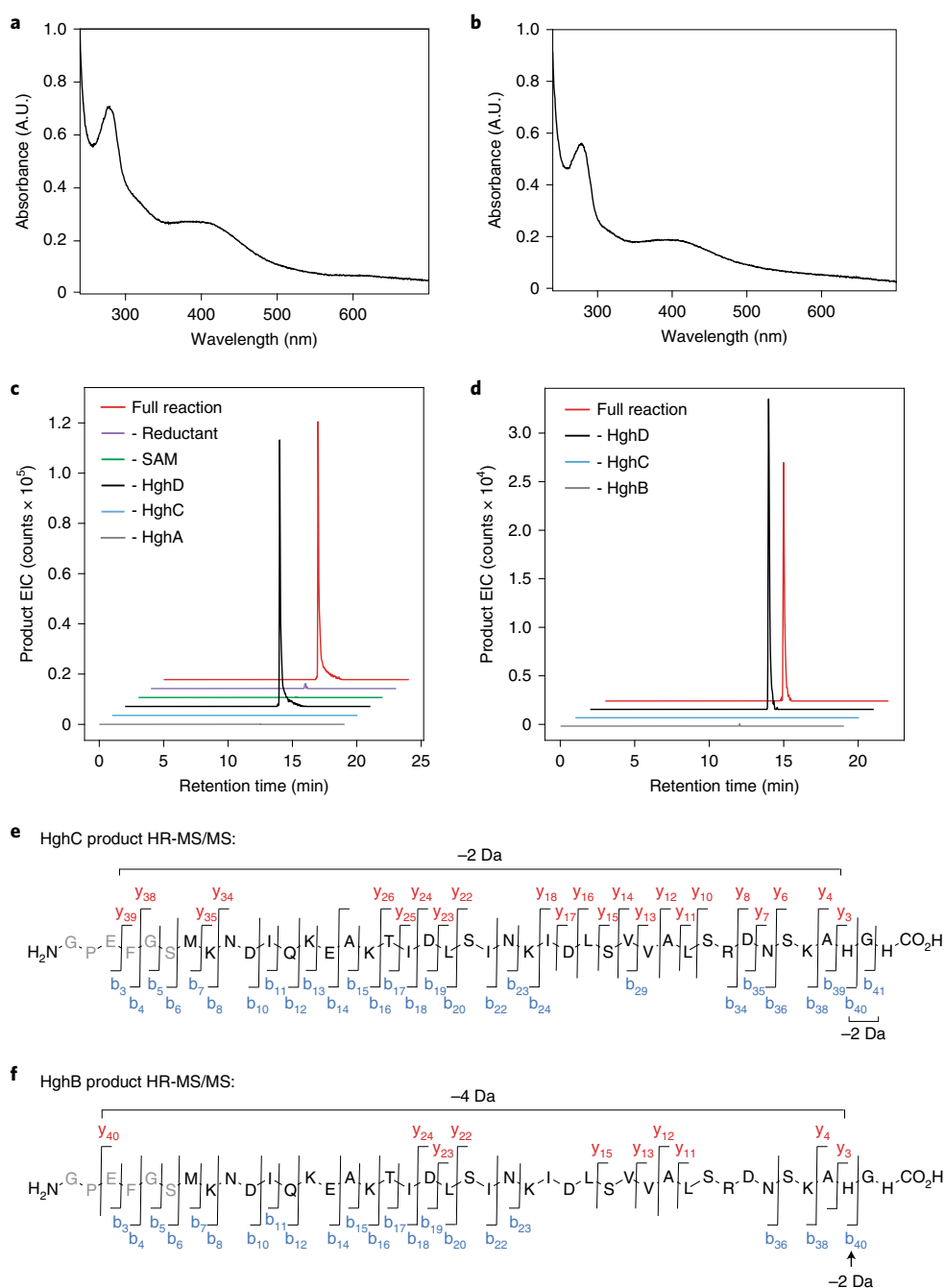


Fig. 4 | Characterization of HghC and HghB in vitro. **a,b**, UV-vis absorption spectra of HghC (**a**) and HghB (**b**) after recombinant expression and purification in *E. coli*. **c,d**, Enzymatic activity assays of HghC (**c**) and HghB (**d**) in vitro via HPLC-Qtof-MS analysis. Shown are extracted ion chromatograms for the -2-Da product of HghC (**c**) and the -4-Da product of HghB (**d**). Traces are offset in both axes for clarity and color-coded as indicated; HghD is not required for either reaction. **e,f**, HR-MS/MS analysis of the products of HghC (**e**) and HghB (**f**). Detected fragments are indicated; the patterns observed match those for the products of HghC and HghB in vivo (Extended Data Fig. 1).

common opportunistic pathogens (Fig. 5d and Extended Data Fig. 6). Growth of opportunistic pathogens, *S. agalactiae* strains and *S. thermophilus* LMD-9 was insensitive up to a concentration of 74.3 μ M bicyclostreptin C. *S. thermophilus* strains JIM 8232 and LMG 18311, however, were mildly growth-inhibited, with bicyclostreptin C exhibiting a half-maximal inhibitory concentration (IC₅₀) of 9.5 μ M against strain JIM 8232 (ref. 29), which encodes the *hgh* cluster, and 13.1 μ M against strain LMG 18311, which does not encode *hgh* (Extended Data Fig. 6). Neither the linear 8mer nor the monocyclic precursor of bicyclostreptin C had an

effect, emphasizing the importance of the imidazolidine-4-one for bioactivity (Fig. 5d). To further explore the observed inhibitory activity, strain JIM 8232 was treated with 1 μ M or 10 μ M of bicyclostreptin C, and its growth was monitored over time (Extended Data Fig. 7). Treatment with 10 μ M caused a growth defect. The optical density of the culture did, however, recover—an observation that is consistent with destruction of the toxin or emergence of resistant strains or may indicate outgrowth of a small subpopulation of producers that express immunity gene(s) and can eventually take over the culture.

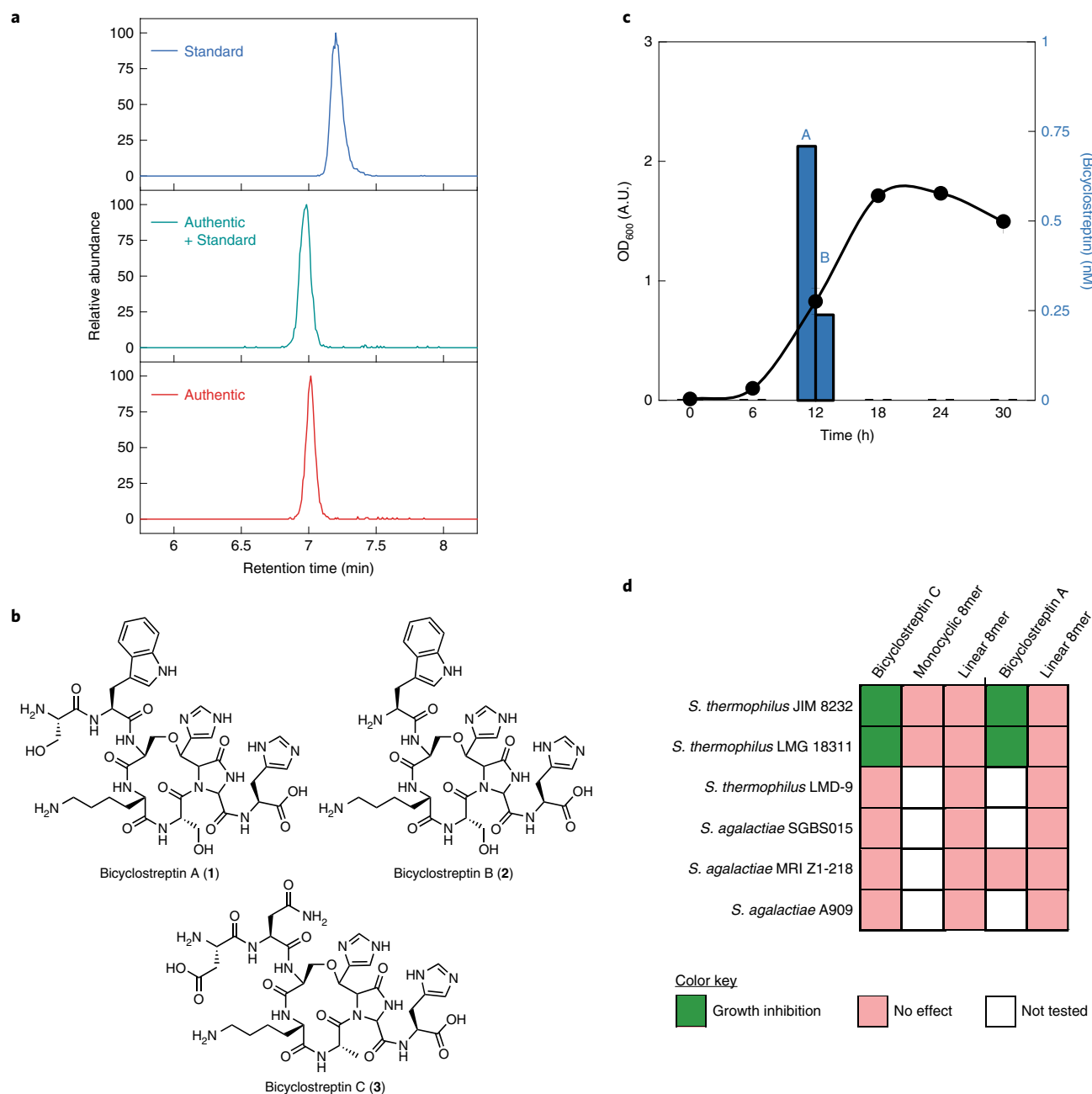


Fig. 5 | Characterization of the mature *hgh* products from *S. thermophilus* JIM 8232 and *S. agalactiae* MRI Z1-218. **a, HPLC-Qtof-MS analysis of authentic and synthetic products. Shown are extracted ion chromatographs for heterologously produced 8mer product (blue, top), authentic 8mer product isolated from *S. thermophilus* JIM 8232 (red, bottom), and a 1:1 co-injection of these two samples (green, middle), which co-elute. **b**, Structures of bicyclostreptin A and B from *S. thermophilus* JIM 8232 and the proposed mature product from *S. agalactiae* MRI Z1-218, bicyclostreptin C. **c**, Detection of bicyclostreptin A and B (blue bars) as a function of *S. thermophilus* JIM 8232 optical density (black circles). The averages of two independent biological measurements are shown. **d**, Effect of bicyclostreptin C, its –2-Da monocyclic precursor or the linear 8mer or bicyclostreptin A and its linear 8mer precursor on the growth of strains indicated. Only the mature products exhibit growth-inhibitory activity (green) against select strains (Extended Data Fig. 6). A.U., absorbance units.**

The isolation of bicyclostreptin A from the trypsin cleavage reaction was low-yielding as we experienced degradation during the purification process. From 12 L of 6HMBPHghABCD-expressing *E. coli* cultures, we obtained sufficient material to test the three strains shown (Extended Data Fig. 6). We observed potent growth inhibition against the producer *S. thermophilus* JIM 8232 and *S. thermophilus* LMG 18311 with minimal inhibitory concentrations (that is, the concentration at which no visible growth was detected) of $<0.54 \mu\text{M}$ and $<0.27 \mu\text{M}$, respectively (Extended Data Fig. 6). By contrast, *S. thermophilus* growth was insensitive to even high

concentrations of the linear 8mer peptide. *S. agalactiae* was not affected by bicyclostreptin A. Given that bicyclostreptin A targets self or near-kin strains, it could be involved in an allostatic process, as has previously been described for extracellular protein toxins in *Streptococcus pneumoniae*^{32–35}. Alternatively, bicyclostreptin A could represent a growth-curbing signal, used by streptococci to regulate growth as population density increases. Growth control mechanisms are varied and common in bacteria and prepare cells for challenges associated with survival in stationary phase^{36–38}. The apparent species-specific activity could be explained with the limited

distribution of the signal receptors required to detect bicyclostreptin A. The identification of bicyclostreptins sets the stage for detailed bioactivity studies to elucidate their roles on the physiology of *S. thermophilus* and other strains.

Discussion

We expand the RiPP natural product family with the addition of a new class—the first that incorporates a crosslink at the backbone amide nitrogen via radical chemistry. A set of diverse RaS enzyme-modified sidechain modifications involving carbon–heteroatom and carbon–carbon linkages as well as α -carbon epimerization and tyramine excision have already been documented^{19–22,27,39–48}. Moreover, backbone modifications to incorporate N-methyl groups or insert azole modifications are well known. However, the involvement of the amide nitrogen in a radical-mediated crosslink is novel and sets the stage for future studies regarding the underlying enzymatic mechanism, as HghB would need to activate both the Gly-H α and the stable amide N-H to establish an sp^3 – sp^3 linkage. The β -ether connection in bicyclostreptins is also a new modification for RaS enzymes and expands the single exemplar of an α -ether-bearing heterocycle that was previously identified²⁰.

Streptococci are host-associated bacteria with characteristically small genomes. They have not been thought of as prolific producers of secondary metabolites. Our search for RaS-RiPP BGCs uncovered at least 600 such BGCs in 2,875 streptococcal genomes, and follow-up investigations have revealed streptococci, against expectation, as a productive source of new and unusual RiPPs. *S. thermophilus* JIM 8232, in particular, boasts a surprisingly spectacular ‘RiPP-ome’. It encodes the *str*, *ggg*, *hgh* and *kgr* gene clusters. The products of the first two clusters, streptide and streptosactin, have already been reported (Fig. 1b); the identification of two bicyclostreptins brings the tally of RiPPs from this strain to four. Studies with the *kgr* locus are underway. The RiPPs examined thus far from *S. thermophilus* exhibit uncommon architectures and a narrow spectrum of antimicrobial and/or signaling activity. They target the producing host and highly related species, and they likely represent intra-species signals with which *S. thermophilus* communicates. These pathways can be intercepted and/or exploited by competing species. The receptors and targets of *S. thermophilus* RiPP signals will be the topic of future studies, and the significance of the observed narrow-spectrum activity remains to be established, especially with regard to the fitness of *S. thermophilus* populations and the consequences for polymicrobial communities in the microbiome. Aside from a source of novel enzymatic chemistry, the network in Fig. 1a is a reservoir of unexplored microbiome natural products. Further investigations using the biosynthesis-guided approach used in the current study promise to unveil new metabolites, even low-abundance ones such as bicyclostreptins, thereby paving the way for detailed functional and mode-of-action studies.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41589-022-01090-8>.

Received: 10 October 2021; Accepted: 21 June 2022;
Published online: 11 August 2022

References

- Bentley, R. Microbial secondary metabolites play important roles in medicine; prospects for discovery of new drugs. *Perspect. Biol. Med.* **40**, 364–394 (1997).
- Clardy, J. & Walsh, C. Lessons from natural molecules. *Nature* **432**, 829–837 (2004).
- Demain, A. L. & Sanchez, S. Microbial drug discovery: 80 years of progress. *J. Antibiot.* **62**, 5–16 (2009).
- Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770–803 (2020).
- Miller, S. J. & Clardy, J. Natural products: beyond grind and find. *Nat. Chem.* **1**, 261–263 (2009).
- Ziemert, N., Alanjary, M. & Weber, T. The evolution of genome mining in microbes—a review. *Nat. Prod. Rep.* **33**, 988–1005 (2016).
- Katz, L. & Baltz, R. H. Natural product discovery: past, present, and future. *J. Ind. Microbiol. Biotechnol.* **43**, 155–176 (2016).
- Ting, C. P. et al. Use of a scaffold peptide in the biosynthesis of amino acid-derived natural products. *Science* **365**, 280–284 (2019).
- Velásquez, J. E. & van der Donk, W. A. Genome mining for ribosomally synthesized natural products. *Curr. Opin. Chem. Biol.* **15**, 11–21 (2011).
- Russell, A. H. & Truman, A. W. Genome mining strategies for ribosomally synthesised and post-translationally modified peptides. *Comput. Struct. Biotechnol. J.* **18**, 1838–1851 (2020).
- Montalban-Lopez, M. et al. New developments in RiPP discovery, enzymology and engineering. *Nat. Prod. Rev.* **38**, 130–239 (2021).
- Bushin, L. B., Clark, K. A., Pelczar, I. & Seyedsayamdost, M. R. Charting an unexplored streptococcal biosynthetic landscape reveals a unique peptide cyclization motif. *J. Am. Chem. Soc.* **140**, 17674–17684 (2018).
- Fleuchot, B. et al. Rgg proteins associated with internalized small hydrophobic peptides: a new quorum-sensing mechanism in streptococci. *Mol. Microbiol.* **80**, 1102–1119 (2011).
- Frey, P. A., Hegeman, A. D. & Ruzicka, F. J. The radical SAM superfamily. *Crit. Rev. Biochem. Mol. Biol.* **43**, 63–88 (2008).
- Landgraf, B. J., McCarthy, E. L. & Booker, S. J. Radical S-adenosylmethionine enzymes in human health and disease. *Annu. Rev. Biochem.* **85**, 485–514 (2016).
- Broderick, J. B., Duffus, B. R., Duschene, K. S. & Shepard, E. M. Radical S-adenosylmethionine enzymes. *Chem. Rev.* **114**, 4229–4317 (2014).
- Bassler, B. L. & Losick, R. Bacterially speaking. *Cell* **125**, 237–246 (2006).
- Gerlt, J. A. Genomic enzymology: web tools for leveraging protein family sequence–function space and genome context to discover novel functions. *Biochemistry* **56**, 4293–4308 (2017).
- Caruso, A., Bushin, L. B., Clark, K. A., Martinie, R. J. & Seyedsayamdost, M. R. Radical approach to enzymatic β -thioether bond formation. *J. Am. Chem. Soc.* **141**, 990–997 (2019).
- Clark, K. A., Bushin, L. B. & Seyedsayamdost, M. R. Aliphatic ether bond formation expands the scope of radical SAM enzymes in natural product biosynthesis. *J. Am. Chem. Soc.* **141**, 10610–10615 (2019).
- Caruso, A., Martinie, R. J., Bushin, L. B. & Seyedsayamdost, M. R. Macrocyclization via an arginine-tyrosine crosslink broadens the reaction scope of radical S-adenosylmethionine enzymes. *J. Am. Chem. Soc.* **141**, 16610–16614 (2019).
- Bushin, L. B., Covington, B. C., Rued, B. E., Federle, M. J. & Seyedsayamdost, M. R. Discovery and biosynthesis of streptosactin, a sactipeptide with an alternative topology encoded by commensal bacteria in the human microbiome. *J. Am. Chem. Soc.* **142**, 16265–16275 (2020).
- Rued, B. E. et al. Quorum sensing in *Streptococcus mutans* regulates production of tryglysin, a novel RaS-RiPP antimicrobial compound. *mBio* **12**, e02688–20 (2021).
- Schramma, K. R., Bushin, L. B. & Seyedsayamdost, M. R. Structure and biosynthesis of a macrocyclic peptide containing an unprecedented lysine-to-tryptophan crosslink. *Nat. Chem.* **7**, 431–437 (2015).
- Imai, Y. et al. A new antibiotic selectively kills Gram-negative pathogens. *Nature* **576**, 459–464 (2019).
- Freeman, M. F. et al. Metagenome mining reveals polytheonamides as posttranslationally modified ribosomal peptides. *Science* **338**, 387–390 (2012).
- Nguyen, T. Q. N. et al. Post-translational formation of strained cyclophanes in bacteria. *Nat. Chem.* **12**, 1042–1053 (2020).
- Burkhart, B. J., Hudson, G. A., Dunbar, K. L. & Mitchell, D. A. A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat. Chem. Biol.* **11**, 564–570 (2015).
- Delorme, C. et al. Complete genome sequence of the pigmented *Streptococcus thermophilus* strain JIM8232. *J. Bacteriol.* **193**, 5581–5582 (2011).
- Sørensen, U. B. S., Wiegand, I. C., Boes, J. & Farre, M. The distribution of clones of *Streptococcus agalactiae* (group B streptococci) among herds/persons and dairy cows demonstrates lack of host specificity for some lineages. *Vet. Microbiol.* **235**, 71–79 (2019).
- Fontaine, L. et al. Novel pheromone quorum-sensing system controls the development of natural competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. *J. Bacteriol.* **192**, 1444–1454 (2010).
- Claverys, J.-P. & Håvarstein, L. S. Cannibalism and fratricide: mechanisms and reasons d'être. *Nat. Rev. Microbiol.* **5**, 219–229 (2007).

33. Claverys, J.-P., Martin, B. & Håvarstein, L. S. Competence-induced fratricide in streptococci. *Mol. Microbiol.* **64**, 1423–1433 (2007).
34. Prozorov, A. A. & Danilenko, V. N. Altolysis in bacteria. *Microbiology* **80**, 1–9 (2011).
35. Guiral, S., Mitchell, T. J., Martin, B. & Claverys, J.-P. Competence-programmed predation of noncompetent cells in the human pathogen *Streptococcus pneumoniae*: genetic requirements. *Proc. Natl Acad. Sci. USA* **102**, 8710–8715 (2005).
36. Lewis, K. Persister cells, dormancy, and infectious disease. *Nat. Rev. Microbiol.* **5**, 48–56 (2007).
37. Rittershaus, E. S. C., Baek, S.-H. & Sasseti, C. M. The normalcy of dormancy: common themes in microbial quiescence. *Cell Host Microbe* **13**, 643–651 (2013).
38. Hayes, C. S. & Low, D. A. Signals of growth regulation in bacteria. *Curr. Opin. Microbiol.* **12**, 667–673 (2009).
39. Flühe, L. et al. The radical SAM enzyme AlbA catalyzes thioether bond formation in subtilisin A. *Nat. Chem. Biol.* **8**, 350–357 (2012).
40. Huo, L., Rachid, S., Stadler, M., Wenzel, S. C. & Müller, R. Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. *Chem. Biol.* **19**, 1278–1287 (2012).
41. Barr, I. et al. Demonstration that the radical S-adenosylmethionine (SAM) enzyme PqqE catalyzes *de novo* carbon-carbon cross-linking within a peptide substrate PqqA in the presence of the peptide chaperone PqqD. *J. Biol. Chem.* **291**, 8877–8884 (2016).
42. Khaliullin, B., Ayikpoe, R., Tuttle, M. & Latham, J. A. Mechanistic elucidation of the mycofactacin-biosynthetic radical S-adenosylmethionine protein, MftC. *J. Biol. Chem.* **292**, 13022–13033 (2017).
43. Benjdia, A., Guillot, A., Ruffié, P., Leprince, J. & Berteau, O. Post-translational modification of ribosomally synthesized peptides by a radical SAM epimerase in *Bacillus subtilis*. *Nat. Chem.* **9**, 698–707 (2017).
44. Parent, A. et al. The B₁₂-radical SAM enzyme PoyC catalyzes valine C_β-methylation during polytheonamide biosynthesis. *J. Am. Chem. Soc.* **138**, 15515–15518 (2016).
45. Grell, T. A. J. et al. Structural and spectroscopic analyses of the sporulation killing factor biosynthetic enzyme SkfB, a bacterial adomet radical sactisynthase. *J. Biol. Chem.* **293**, 17349–17361 (2018).
46. Morinaka, B. I. et al. Natural noncanonical protein splicing yields products with diverse β -amino acid residues. *Science* **359**, 779–782 (2018).
47. Ayikpoe, R. S. & Latham, J. A. MftD catalyzes the formation of a biologically active redox center in the biosynthesis of the ribosomally synthesized and post-translationally modified redox cofactor mycofactocin. *J. Am. Chem. Soc.* **141**, 13582–13591 (2019).
48. Hudson, G. A. et al. Bioinformatic mapping of radical S-adenosylmethionine-dependent ribosomally synthesized and post-translationally modified peptides identifies new C α , C β , and C γ -linked thioether-containing peptides. *J. Am. Chem. Soc.* **141**, 8228–8238 (2019).

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2022

Methods

Materials. All chemicals and material were purchased from Sigma-Aldrich or Thermo Fisher Scientific unless otherwise specified. Restriction enzymes T4 DNA Ligase, Q5 High-Fidelity DNA Polymerase, Shrimp Alkaline Phosphatase, HiFi Assembly MasterMix and Trypsin-Ultra (Mass Spectrometry grade) were purchased from New England Biolabs (NEB). DNA oligos were purchased from Sigma-Aldrich. Codon-optimized synthetic gene fragments were obtained from GENEWIZ. *S. thermophilus* JIM 8232 was kindly provided by Christine Delorme (INRA, Micalis Institut)²⁹. *S. agalactiae* SGBS015 was kindly provided by Samuel Shelbourne (The University of Texas MD Anderson Cancer Center)⁴⁹. *S. agalactiae* MRI Z1-218 (LMG 26552)³⁰ was purchased from the BCCM/LMG Bacteria Collection.

General procedures. High-performance liquid chromatography (HPLC)-coupled HR-MS and HR-MS/MS were carried out on an Agilent 6540 quadrupole time-of-flight (Qtof) MS instrument, consisting of an automated liquid sampler, a 1260 Infinity Series HPLC system, a diode array detector, a JetStream ESI source and the 6540 Series Qtof. HPLC purifications were carried out on Agilent 1260 Infinity Series HPLC systems equipped with a temperature-controlled column compartment, a diode array detector, an automated fraction collector and an automated liquid sampler. The mobile phases for all HPLC methods were water (solvent A) and MeCN (solvent B), both containing 0.1% formic acid (FA). Specific columns and gradients are detailed in the relevant sections below. Agilent instruments were controlled with Agilent MassHunter software.

NMR spectra were acquired at Princeton University Department of Chemistry facilities. NMR spectra were collected in D₂O or 90% H₂O/10% D₂O in the QCI cryoprobe of a Bruker Avance III HD 800-MHz NMR spectrometer or in the TCI cryoprobe of a Bruker Avance III 500-MHz NMR spectrometer, which were controlled with Bruker Topsin software. 1D/2D NMR data were analyzed with MestReNova software, and figures were generated in Adobe Illustrator.

General molecular biology/cloning procedures. Genomic DNA from *S. thermophilus* JIM 8232 was isolated using the Wizard Genomic DNA Purification Kit (Promega). PCR reactions were performed with Q5 High Fidelity DNA polymerase (NEB) in FailSafe 2x PreMix Buffer D (Epicenter) in a CFX96 thermocycler (Bio-Rad). PCR products and restriction enzyme-digested products were purified using the Qiagen PCR Purification Kit and the Qiagen Gel Extraction Kit, respectively, according to the manufacturer's protocols. Linearized vectors were treated with Shrimp Alkaline Phosphatase (rSAP) before gel extraction. Ligation reactions were performed with T4 DNA Ligase and DNA assemblies with HiFi DNA Assembly Master Mix, according to instructions. Chemically competent *E. coli* DH5α cells were transformed with ligation or HiFi mixtures by heat-shock and plated onto LB agar containing the appropriate antibiotic. Assembled plasmids were confirmed by Sanger sequencing.

Construction of pRSFDuet-1_6HMBP*hghA*_JIM8232. Plasmid pRSFDuet-1_6HMBP reported previously²² was digested with BamHI and PstI to allow fusion of *hghA* to the 3' end of the 6HMBP coding sequence. An *hghA* gene fragment with appropriate overlap regions was generated by PCR amplification of *S. thermophilus* JIM8232 genomic DNA using primers *hghA*_JIM8232_F and *hghA*_JIM8232_R (Supplementary Table 3). The PCR product was combined with the linearized plasmid in a HiFi reaction.

Construction of pRSFDuet-1_6HMBP*hghA*_hghBCD_JIM8232. Plasmid pRSFDuet-1_6HMBP*hghA*_JIM8232 (see above) was digested with NdeI and XhoI to allow for insertion of gene fragments into MCS2. A DNA region extending from the start codon of *hghB* to the stop codon of *hghD* was PCR-amplified from genomic DNA using primers *hghBCD*_JIM8232_F (NdeI) and *hghBCD*_JIM8232_R (XhoI). The ~3-kb PCR product was subsequently digested with NdeI and XhoI and ligated to linearized pRSFDuet-1_6HMBP*hghA*.

Construction of pRSFDuet-1_6HMBP*hghA*_hghBD_JIM8232. An *hghB* gene fragment with 5' and 3' overlap regions was generated by PCR-amplifying genomic DNA with primers *hghBCD*_JIM8232_F (NdeI) and *hghBD*_JIM8232_R. An *hghD* gene fragment with 5' and 3' overlap regions was generated by PCR-amplifying genomic DNA with primers *hghBD*_JIM8232_F and *hghBCD*_JIM8232_R (XhoI). The *hghB* and *hghD* fragments were mixed in a three-way HiFi DNA assembly reaction with NdeI/XhoI-digested pRSFDuet-1_6HMBP*hghA*_JIM8232.

Construction of pRSFDuet-1_6HMBP*hghA*_hghCD_JIM8232. The DNA region from the start codon of *hghC* to the stop codon of *hghD* was PCR-amplified from genomic DNA using primers *hghCD*_JIM8232_F and *hghBCD*_JIM8232_R (XhoI). The PCR product, which contained appropriate overlap regions at its 5' and 3' ends, was then joined with NdeI/XhoI-digested pRSFDuet-1_6HMBP*hghA*_JIM8232 in a HiFi assembly.

Construction of pRSFDuet-1_6HMBP*hghA*_MRIZ1218. The *hghA* gene was obtained as a codon-optimized synthetic DNA fragment for *E. coli* expression. Overlap regions for insertion between BamHI and PstI of pRSFDuet-1_6HMBP were appended to the 5' and 3' ends of the *hghA* gene by PCR-amplifying the

synthetic fragment with primers *hghA*_MRIZ1218_F and *hghA*_MRIZ1218_R. The PCR product was then joined with linearized pRSFDuet-1_6HMBP in a HiFi assembly, fusing *hghA* to the 3' end of the 6HMBP coding sequence.

Construction of pRSFDuet-1_6HMBP*hghA*_hghBCD_MRIZ1218. The *hghB*, *hghC* and *hghD* genes were obtained as synthetic DNA fragments, codon-optimized for expression in *E. coli*. Each gene was PCR-amplified from its synthetic fragment with primers containing 5' overhangs, providing fragments with appropriate overlap regions for HiFi assembly. The *hghB* fragment was obtained using primers *hghB*_MRIZ1218_F and *hghBC*_MRIZ1218_R, the *hghC* fragment using primers *hghBC*_MRIZ1218_F and *hghCD*_MRIZ1218_R and the *hghD* fragment using primers *hghCD*_MRIZ1218_F and *hghD*_MRIZ1218_R. Primers were designed to include synthetic ribosome binding site (RBS) sequences in intergenic regions between *hghB* and *hghC* as well as between *hghC* and *hghD*. The 3' fragments and pRSFDuet-1_6HMBP*hghA*_MRIZ1218, which had been linearized by digestion with NdeI and XhoI, were then joined in a four-way HiFi assembly.

Construction of pRSFDuet-1_6HMBP*hghA*_hghBD_MRIZ1218. The *hghB* synthetic fragment was PCR-amplified using primers *hghB*_MRIZ1218_F and *hghBD*_MRIZ1218_R to produce an *hghB* fragment with appropriate overlap regions, including an RBS-containing sequence downstream. Another fragment with a complementary RBS-containing sequence upstream of *hghD* and an appropriate 3' overlap region was PCR-amplified from the *hghD* synthetic fragment using primers *hghBD*_MRIZ1218_F and *hghD*_MRIZ1218_R. The two fragments were then combined with NdeI/XhoI-digested pRSFDuet-1_6HMBP*hghA*_MRIZ1218 in a three-way HiFi assembly.

Construction of pRSFDuet-1_6HMBP*hghA*_hghCD_MRIZ1218. The same *hghD* fragment used in the assembly of pRSFDuet-1_6HMBP*hghA*_hghBCD_MRIZ1218 was combined with a new *hghC* fragment and NdeI/XhoI-digested pRSFDuet-1_6HMBP*hghA*_MRIZ1218 in a three-way HiFi assembly. The new *hghC* fragment was obtained by PCR-amplifying the *hghC* synthetic fragment with primers *hghC*_MRIZ1218_F and *hghCD*_MRIZ1218_R, which provided the appropriate 5' overlap region and the 3' RBS-containing intergenic region.

Generation of *E. coli* BL21(DE3) strains for in vivo activity assays. The plasmids above were constructed for in vivo activity assays. Each pRSFDuet-1-based plasmid, along with pDB1282 (*isc* operon), was transformed into chemically competent *E. coli* BL21(DE3) by heat-shock and plated onto LB agar supplemented with 100 µg ml⁻¹ of ampicillin (Amp) and 50 µg ml⁻¹ of kanamycin (Kan). For simplicity, each strain is referred to by the *hgh* gene(s) in its plasmid; for example, *E. coli* BL21(DE3) carrying pRSFDuet-1_6HMBP*hghA*_hghBD is called strain ABD.

Analytical-scale growth of *E. coli* BL21(DE3) strains, purification of 6HMBP*hghA* fusion proteins and MS analysis. For testing the *S. thermophilus* JIM 8232 *hgh* system or the *S. agalactiae* MRI Z1-218 *hgh* system, the four strains (A, ABD, ACD and ABCD) were grown side by side. Strain A, which expresses only the 6HMBP-*hghA* gene fusion, functioned as a control. Each strain was streaked out from frozen glycerol stocks onto LB agar supplemented with 100 µg ml⁻¹ of Amp and 50 µg ml⁻¹ of Kan. A 15-ml culture tube containing 5 ml of LB (+50 µg ml⁻¹ of Kan and 100 µg ml⁻¹ of Amp) was inoculated with a single colony and grown overnight at 37 °C/200 r.p.m. The next morning, 0.5 ml of overnight culture was used to inoculate a 125-ml flask containing 50 ml of Terrific Broth (Kan/Amp). This 50-ml culture was grown at 37 °C/200 r.p.m. to an optical density (at 600 nm, OD₆₀₀) of 0.4, at which point it was supplemented with 0.05 mM FeCl₃ and 0.05% arabinose to induce expression of the *isc* operon. The culture was grown to OD₆₀₀ ≈ 0.8, at which point it was supplemented with 0.5 mM IPTG to induce expression of the *hgh* gene(s). Growth was continued at 37 °C/200 r.p.m. for 18 hours. Cells were harvested by centrifugation (30,000g, 30 minutes, 4 °C). A typical yield was 100–200 mg of cell paste per 50 ml of culture.

The 6HMBP*hghA* fusion proteins were purified from cell paste in the same manner as 6HMBP*GggA* fusion proteins in previously published protocols²². In brief, clarified cell lysate was subjected to Ni affinity chromatography using a HisPur Ni-NTA Spin column (0.2-ml resin bed volume, Thermo Fisher Scientific). Eluate containing 6HMBP*hghA* protein was then subjected to gel filtration using a PD-10 desalting column (GE Healthcare).

6HMBP tags were subsequently removed from isolated fusion proteins to facilitate MS analysis. Proteolysis with HRV3C was carried out in cleavage buffer (50 mM Tris-HCl, 150 mM NaCl, pH 8) for 16 hours at 4 °C and quenched by incubation at 95 °C for a few minutes. Heat-quenched reactions were centrifuged at 18,000g for 3 minutes to remove precipitate. Clarified and filtered samples were then analyzed by HPLC-Qtof-MS using a Phenomenex Jupiter C18 300-Å column (5 µm, 4.6 mm × 100 mm) operating at 0.6 ml min⁻¹. Elution was performed with an isocratic step of 8% B for 5 minutes, followed by gradient steps of 8–68% B over 15 minutes and 68–100% B over 3 minutes.

Large-scale growth of *E. coli* BL21(DE3) strains for structural elucidation of HghA modified in vivo. A 250-ml flask containing 100 ml of LB (+50 µg ml⁻¹ of Kan and 100 µg ml⁻¹ of Amp) was inoculated with a single colony of the desired

strain. The cells were grown overnight at 37°C/200 r.p.m. and used to inoculate 8 × 0.8 L of TB-Kan/Amp in 2.0-L flask at 1% dilution. Cultures were grown continuously at 37°C/200 r.p.m. Arabinose and FeCl₃ were added to the culture at OD₆₀₀ ≈ 0.4 at final concentrations of 0.05% mM and 0.05 mM, respectively. IPTG was added to the culture at OD₆₀₀ ≈ 0.8 at a final concentration of 0.5 mM. After 18 hours of growth post-induction with IPTG, cells were harvested by centrifugation (15,000g, 30 minutes, 4°C) and frozen at -80°C. A typical yield was 5–7 g of cell paste per liter of culture.

Large-scale purification and HRV3C cleavage of 6HMBP_{HghA} fusion proteins. Large-scale purification of 6HMBP_{HghA} fusion proteins and subsequent 6HMBP tag removal using HRV3C protease was carried out as reported previously for 6HMBP_{GggA}-R22 (ref. 27).

Purification of HRV3C-cleaved HghA peptides. Each HRV3C reaction mixture from above was heat-quenched at 95°C and clarified by centrifugation and filtration through a 0.22-μm syringe filter. The liberated HghA peptide was then isolated from the solution by HPLC using a semi-preparative Phenomenex Jupiter 300-Å C18 column (5 μm, 10 × 250 mm) operating at 2.5 ml min⁻¹. The solution (900 μl per run) was injected onto the column and resolved using 10% B for 5 minutes and a gradient of 10–70% B over 20 minutes. *S. thermophilus* HghA eluted from ~19 to 20 minutes in 52–55% B, and *S. agalactiae* HghA eluted from ~18 to 19 minutes in 49–52% B. Fractions containing HghA were pooled, dried in vacuo to remove MeCN and lyophilized to remove residual H₂O.

Trypsin cleavage of isolated HRV3C-cleaved HghA peptides. Isolated HRV3C-cleaved HghA from above was treated with Trypsin-ultra Mass Spectrometry Grade (NEB) according to the manufacturer's instructions. In brief, lyophilized material was resuspended in cleavage buffer containing 20 mM CaCl₂ (1 ml per ~5 mg of peptide). The solution was then supplemented with trypsin (2.5 μg ml⁻¹) and incubated at 37°C overnight.

Purification of trypsin-cleaved HghA 8mer peptides. The 8mer HghCD or HghBCD product was isolated from the reaction mixture above by HPLC using an analytical Phenomenex Luna Omega Polar C18 100-Å column (5 μm, 100 × 4.6 mm) operating at 0.5 ml min⁻¹. Elution of *S. agalactiae* products (H₂N-DNSKAHGH-CO₂H) was performed with an isocratic step of 100% A for 5 minutes. The peptide eluted at ~3 minutes. The elution method for *S. thermophilus* products (H₂N-SWSKSHGH-CO₂H) consisted of an isocratic step of 0% B for 5 minutes, followed by a gradient of 0–17.5% B over 7 minutes. The peptide eluted at ~11.5 minutes.

Synthesis of linear 8mer peptides. Linear peptide standards were prepared by Fmoc-based SPPS on a preloaded H-His(Trt)-HMPB-ChemMatrix resin using a Liberty Blue automated peptide synthesizer equipped with a Discover microwave module (CEM). The deprotection solution consisted of 10% piperazine (w/v) in a 10:90 solution of EtOH:NMP (N-methylpyrrolidine) supplemented with 0.1 M HOBT (1-hydroxybenzotriazole). The activator solution consisted of 0.5 M DIC (N,N'-diisopropylcarbodiimide) in DMF and activator base solution of 1.0 M Oxyma with 0.1 M DIPEA in DMF. A typical coupling cycle used 5 equiv of amino acid and 5 equiv of coupling reagent. Synthesis was performed on a 100-μmol scale. Upon completion of the synthesis, the resin was removed from the reaction vessel and transferred to an Econo-Pac column (Bio-Rad). The resin was washed several times with DMF, followed by DCM, and dried thoroughly under vacuum. The peptide was cleaved from the resin by incubation with freshly prepared cleavage cocktail (5 ml per 100 mg of resin) consisting of 92.5% TFA, 2.5% H₂O, 2.5% TIS (triisopropylsilane) and 2.5% DODT (2,2'-(ethylenedioxy)diethanethiol). The reaction was stirred for 3 hours at room temperature. The mixture was drained from the reaction tube, and the resin was rinsed several times with TFA. The filtrate and rinses were combined and subsequently concentrated by evaporation of TFA under a stream of N₂. The peptide was then precipitated by addition of 10 volumes of ice-cold diethyl ether and isolated by centrifugation (4,000g, 10 minutes, 4°C). The ether was poured off, and the peptide was dried overnight in a fume hood.

Purification of linear 8mer peptides. Crude peptides were dissolved in H₂O (+0.1% FA) and purified by semi-preparative HPLC using a Phenomenex Synergi Fusion-RP column (4 μm, 250 × 10 mm) operating at 2.5 ml min⁻¹. Elution of the *S. agalactiae* linear 8mer (H₂N-DNSKAHGH-CO₂H) was carried out with an isocratic method of 100% A for 5 minutes. The *S. thermophilus* linear 8mer (H₂N-SWSKSHGH-CO₂H) was eluted with a gradient method consisting of 0% B for 5 minutes, followed by 0–8% B over 8 minutes. The linear standards were structurally characterized by HPLC HR-MS, HPLC HR-MS/MS and 1D/2D NMR.

Construction of pET-28b(+)_hghB_MRIZ128 for in vitro activity assays. The *hghB* synthetic fragment was PCR-amplified using primers *hghB*_MRIZ1218_F (NdeI) and *hghB*_MRIZ1218_R (XhoI). The PCR product was subsequently digested with NdeI and XhoI and ligated to plasmid pET-28b(+), which had been digested with the same restriction enzymes.

Construction of pET-28b(+)_hghC_MRIZ128 for in vitro activity assays. The *hghC* synthetic fragment was PCR-amplified using primers *hghC*_MRIZ1218_F (NdeI) and *hghC*_MRIZ1218_R (XhoI). The PCR product was subsequently digested with NdeI and XhoI and ligated to NdeI/XhoI-digested pET-28b(+).

Construction of pET-28b(+)_hghD_MRIZ128 for in vitro activity assays. The *hghD* synthetic fragment was PCR-amplified using primers *hghD*_MRIZ1218_F (NdeI) and *hghD*_MRIZ1218_R (XhoI). The PCR product was subsequently digested with NdeI and XhoI and ligated to NdeI/XhoI-digested pET-28b(+).

Expression and purification of His₆-HghC for in vitro activity assays. Expression of His₆-HghC was found to be best using modified conditions from our standard, previously published RaS enzyme expression protocol⁵⁰. Next, 4 × 4.0 L flasks each containing 2.5 L of M9 media (supplemented with 2 mM Mg(SO₄)₂, 0.1 mM CaCl₂, 20 mM glucose, 50 μg ml⁻¹ of Kan and 100 μg ml⁻¹ of Amp) were inoculated with 1% of a culture grown overnight from a single colony of *E. coli* BL21(DE3) cells carrying the pET-28b(+)_hghC and pDB1282 plasmids. Cells were grown at 37°C/180 r.p.m. to an OD₆₀₀ of ~0.4 and were supplemented with 0.2% (w/v) L-(+)-arabinose. They were grown to an OD₆₀₀ of ~0.8, cooled to 18°C and supplemented with 50 μM FeCl₃ before induction with 100 μM IPTG. Cells were grown overnight at 18°C/110 r.p.m. and harvested by centrifugation after 18 hours. A typical yield was 2.3 g of pellet per liter of culture.

Purification of His₆-HghC was carried out anaerobically using previously published procedures without modifications⁵. Reconstitution with Fe^{II} and Na₂S was not necessary to obtain active enzyme. Purified His₆-HghC was quantified spectrophotometrically using a calculated extinction coefficient of 43,320 M⁻¹ cm⁻¹. A yield of 1 mg of pure His₆-HghC per gram of cell paste was typical.

Expression, purification and reconstitution of His₆-HghB for in vitro activity assays. Expression, purification and reconstitution of His₆-HghB was carried out using previously published procedures without modifications⁵⁰. Purified His₆-HghB was quantified spectrophotometrically using a calculated extinction coefficient of 55,810 M⁻¹ cm⁻¹. A yield of 1 g of cell pellet per liter of culture and 1.3 mg of pure His₆-HghB per gram of cell paste was typical.

Expression and purification of His₆-HghD for in vitro activity assays. A 250-ml Erlenmeyer flask containing 100 ml of LB supplemented with Kan (50 μg ml⁻¹) was inoculated with a single colony of *E. coli* BL21(DE3) cells carrying pET-28b(+)_hghD. The culture was grown for 6 hours at 37°C/200 r.p.m. and used to inoculate 4 × 4.0 L flasks each containing 1.5 L of LB-Kan at 1% dilution. This culture was grown at 37°C/200 r.p.m. to OD₆₀₀ ≈ 0.9, cooled in an ice bath and supplemented with a final concentration of 0.1 mM IPTG to induce expression. Cultures were grown at 18°C/200 r.p.m. for 17 hours. Cells were then harvested by centrifugation (15,000g, 30 minutes, 4°C) and frozen at -80°C. The yield was ~3.6 g cells per liter of culture.

Purification was carried out at 4°C in a cold room. Wet cell paste (~22 g) was resuspended in lysis buffer (5 ml g⁻¹ of cell paste), which consisted of 50 mM HEPES, 300 mM KCl, 25 mM imidazole, 15% glycerol, pH 7.5 (+2 mM TCEP). Phenylmethanesulfonyl fluoride (PMSF) was added to a final concentration of 0.1 mM, and Protease Inhibitor Cocktail was added to a concentration of 10 μl ml⁻¹. Lysozyme and DNaseI were added to final concentrations of 1 mg ml⁻¹ and 0.1 mg ml⁻¹ of lysis buffer, respectively, and the suspension was stirred for 30 minutes at 4°C. The cells were then incubated on ice and subsequently sonicated for a total of 4 minutes in 15-seconds-on/15-seconds-off cycles at a 30% power rating. The cells were allowed to rest for 4 minutes, and the cycle was repeated. The crude cell lysate was transferred to centrifuge tubes, and cell debris was pelleted by centrifugation (35,000g, 65 minutes, 4°C). The supernatant was then loaded onto a nickel affinity resin (5 ml), which had been equilibrated with 10 column volumes (CV) of lysis buffer. The column was washed with an additional 5 CV of lysis buffer, followed by 5 CV of wash buffer consisting of 50 mM HEPES, 300 mM KCl, 25 mM imidazole, 15% glycerol, pH 7.5 (+2 mM TCEP). His₆-HghD was eluted with 1.5–2 CV of elution buffer consisting of 50 mM HEPES, 300 mM KCl, 300 mM imidazole, 15% glycerol, pH 7.5 (+2 mM TCEP). The purified protein was subsequently desalted on a Sephadex G-25 column (~50 ml, d = 1.25 cm, l = 25 cm), which had been equilibrated in G-25 buffer (100 mM HEPES, 300 mM KCl, 10% glycerol, pH 7.5 (+5 mM DTT)). Fractions containing His₆-HghD, as judged by the Bradford assay, were pooled, aliquoted and stored at -80°C. A yield of 3 mg of pure His₆-HghD per gram of cell paste was typical. The concentration of His₆-HghD was determined spectrophotometrically using a calculated extinction coefficient of 11,460 M⁻¹ cm⁻¹.

Expression and purification of HghA for in vitro activity assays. 6HMBP-HghA was expressed from *E. coli* BL21(DE3) carrying pRSFDuet-1_6HMBP_{HghA}_MRIZ1218 (Strain A) exactly as described above but in 4 × 1.5 L flasks. A typical yield was 4 g of cell paste per liter of culture.

Purification of 6HMBP-HghA was carried out exactly as described for purification of His₆-HghD. The concentration of the 6HMBP-HghA peptide was determined spectrophotometrically using a calculated extinction coefficient of 66,350 M⁻¹ cm⁻¹. The yield was 5.9 mg of protein per gram of pellet. After

purification, the 6HMBP tag was removed using HRV3C protease as described above. The HRV3C reaction mixture was heat-quenched at 95 °C and clarified by centrifugation and filtration through a 0.22- μ m syringe filter. Its concentration was then determined spectroscopically.

Enzymatic activity assays. Enzymatic assays were performed in an inert atmosphere in an MBraun glovebox. Reactions were typically carried out in 0.5-ml microcentrifuge tubes on a 50- μ l scale with final concentrations of 5 mM DTT, 0.2 mM titanium citrate, 0.2 mM HghA, 5 μ M HghB, 2 μ M HghC, 10 μ M HghD and 1 mM SAM. Reactions were incubated overnight. At the end of the incubation period, the tubes were taken out of the glovebox and heated at 95 °C for 3 minutes. Precipitated protein was removed by centrifugation. Samples were then diluted with an equal volume H₂O (+0.1% FA) and directly analyzed by HPLC-Qtof-MS using a Phenomenex Kinetex C18 100-Å column (2.6 μ m, 4.6 mm $\times$ 100 mm) operating at 0.5 ml min⁻¹. Elution was performed with an isocratic step of 5% B for 5 minutes, followed by gradient steps of 5–70% B over 7 minutes and 70–100% B over 1 minute (Extended Data Fig. 2b). In separate reactions, individual components were omitted to assess their effects on product formation. Samples for the analysis of 5'-dA formation in the absence of substrate (Extended Data Fig. 2a) were analyzed on an analytical Phenomenex Synergi Fusion-RP 80-Å column (4 μ m, 4.6 mm $\times$ 100 mm) operating at 0.5 ml min⁻¹. Elution was performed with an isocratic step of 0% B for 5 minutes, followed by gradient steps of 0–30% B over 9 minutes and 30–100% B over 3 minutes.

Identification of natural product from *S. thermophilus* JIM 8232. Growth of *S. thermophilus* and analysis of supernatants was performed according to previously published procedures used for the discovery of streptosactin²². Bicyclostreptin was detected in culture supernatants but not in cell pellets or intracellular fractions.

Microtiter bioactivity assays. *Streptococcus* strains were streaked out from frozen glycerol stocks onto THY/BHI agar plates and incubated overnight at 37 °C under 5% CO₂ atmosphere. Multiple colonies from the agar plate were directly suspended in liquid CDM. The bacterial suspension was then diluted with CDM to an initial OD₆₀₀ of 0.05–0.1. Bicyclostreptin was dissolved in H₂O (1.28 mg ml⁻¹) and diluted ten-fold in CDM to 128 μ g ml⁻¹. A 96-well plate was then set up following the broth microdilution protocol described by Wiegand et al.⁵¹. This setup provided the following range of bicyclostreptin final concentrations: 0, 0.25, 0.5, 1.0, 2.0, 4.0, 8.0, 16.0, 32.0 and 64.0 μ g ml⁻¹. Due to limited material for testing the susceptibility of *S. thermophilus* LMG18311 to bicyclostreptin A, the stock solution was diluted 20-fold in CDM to 64 μ g ml⁻¹. This gave the following range of final concentrations: 0, 0.125, 0.25, 0.5, 1.0, 2.0, 4.0, 8.0, 16.0 and 32.0 μ g ml⁻¹. Next, 96-well plates were incubated at 37 °C, 5% CO₂ for 20 hours. Final OD₆₀₀ values were measured using a BioTek plate reader that was controlled with BioTek Gen5 software. Susceptibility testing with other strains (*E. coli*, *P. aeruginosa*, *S. aureus* and *E. faecalis*) was performed in the same manner except bacteria were grown using Mueller Hinton agar/broth at 37 °C without CO₂. All liquid media were pre-warmed before use. IC₅₀ values were calculated using Dose Response analysis in OriginLab.

Growth inhibition assay in culture tubes. Bacterial strains were streaked out from frozen glycerol stocks onto THY/BHI agar plates and incubated overnight at 37 °C, 5% CO₂. The next day, colonies from the plate were directly suspended in pre-warmed liquid CDM. The bacterial resuspension was then used to inoculate 4 ml of pre-warmed CDM in a 5-ml culture tube to OD₆₀₀ of 0.05–0.1. The bicyclostreptin stock concentration was added to the appropriate samples. Cultures

were then incubated at 37 °C, 5% CO₂ for the duration of the time course. At each time point, the tube was removed from the incubator, and the OD₆₀₀ was measured using a GENESYS 30 Vis Spectrophotometer equipped with a test tube adapter and then returned to the incubator.

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

Data availability

Plasmids and strains used in this study are described in Supplementary Tables 1 and 2. All oligonucleotides are shown in Supplementary Table 3. The sequences of codon-optimized gene fragments are provided in Supplementary Note 1. Other relevant data supporting the findings of this study are available within the paper and the supplementary material. Raw NMR data used to elucidate natural product structures are available from the corresponding author upon reasonable request.

References

- Flores, A. R. et al. Sequence type 1 group B *Streptococcus*, an emerging cause of invasive disease in adults, evolves by small genetic changes. *Proc. Natl Acad. Sci. USA* **112**, 6431–6436 (2015).
- Bushin, L. B. & Seyedsayamdost, M. R. Guidelines for determining the structures of radical SAM enzyme-catalyzed modifications in the biosynthesis of RiPP natural products. *Methods Enzymol.* **606**, 439–460 (2018).
- Wiegand, I., Hilpert, K. & Hancock, R. E. W. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protoc.* **3**, 163–175 (2008).

Acknowledgements

We thank B. A. Johnson for technical assistance with isolation of bicyclostreptin A as well as the Eli Lilly Edward C. Taylor Fellowship in Chemistry (to K.A.C.) and the National Science Foundation (NSF GRFP to L.B.B. and NSF CAREER Award to M.R.S.) for financial support.

Author contributions

L.B.B. and M.R.S. conceived of the study. L.B.B. carried out in vivo characterization and structural elucidation of the reactions of all enzymes. B.C.C. identified the mature natural products from the native organism. K.A.C. conducted in vitro analysis of the radical SAM enzymes. L.B.B., B.C.C. and A.L. carried out bioactivity assays. L.B.B. and M.R.S. wrote the manuscript, with contributions from all authors.

Competing interests

The authors declare no competing interests.

Additional information

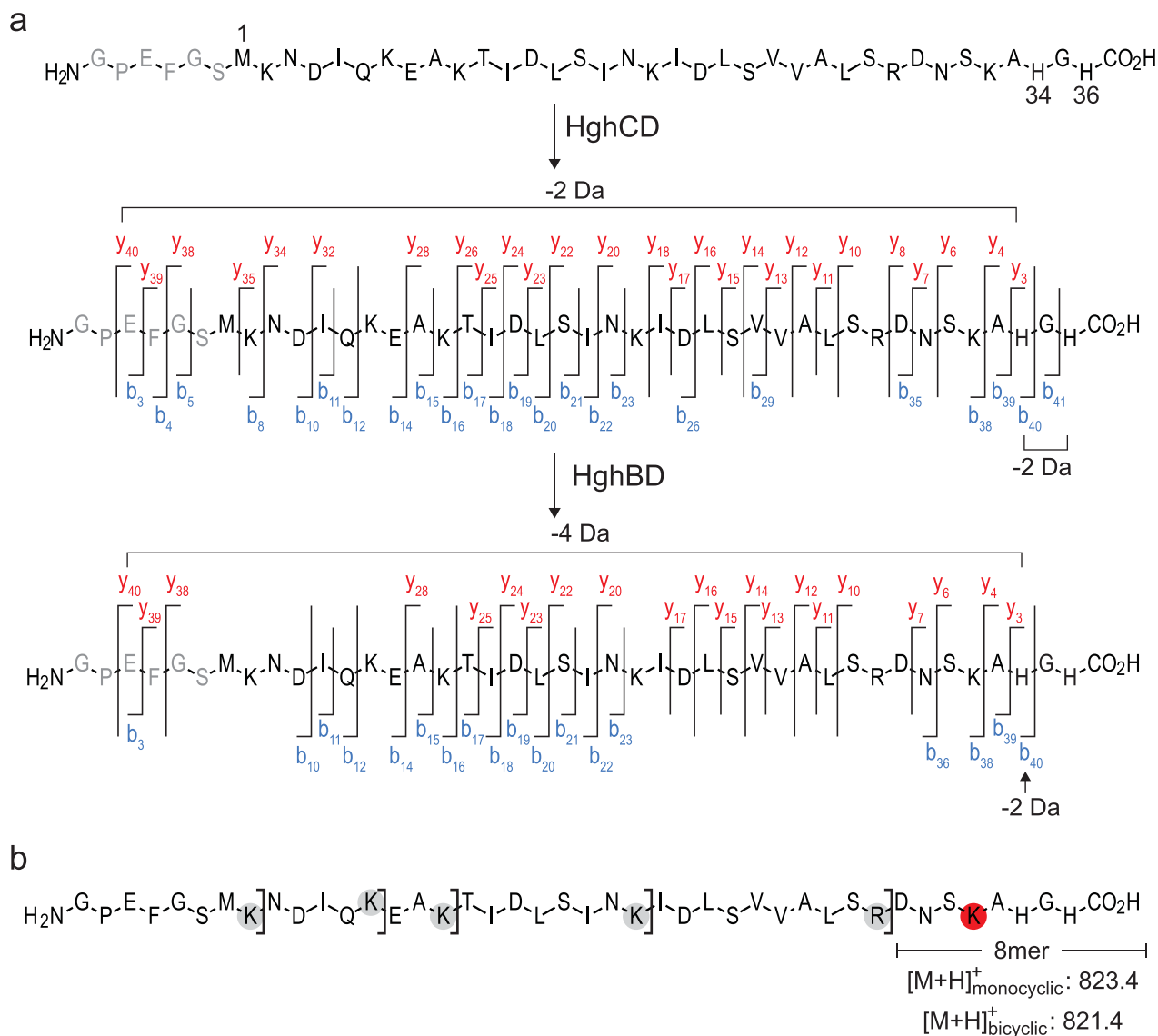
Extended data is available for this paper at <https://doi.org/10.1038/s41589-022-01090-8>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41589-022-01090-8>.

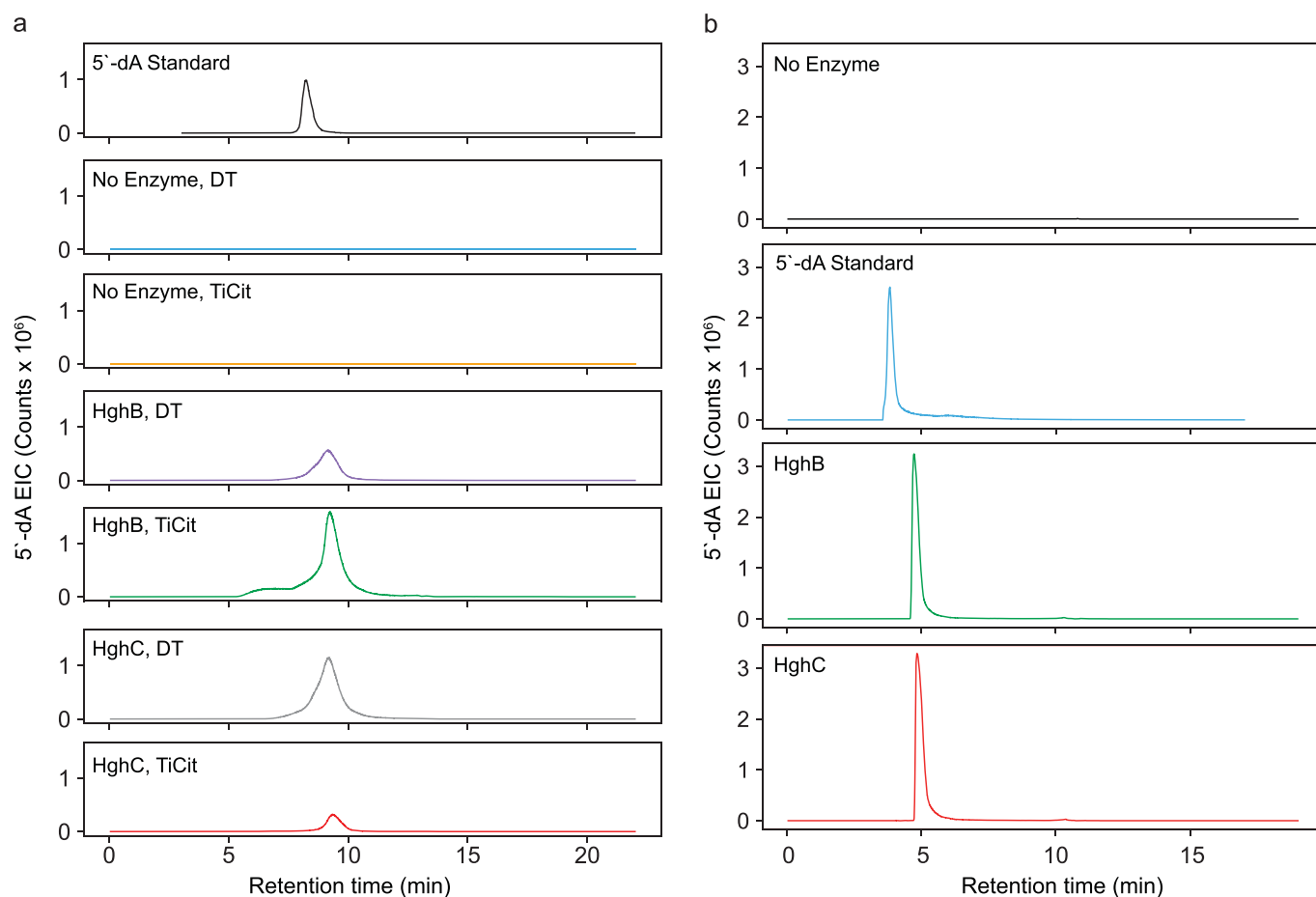
Correspondence and requests for materials should be addressed to Mohammad R. Seyedsayamdost.

Peer review information *Nature Chemical Biology* thanks the anonymous reviewers for their contribution to the peer review of this work.

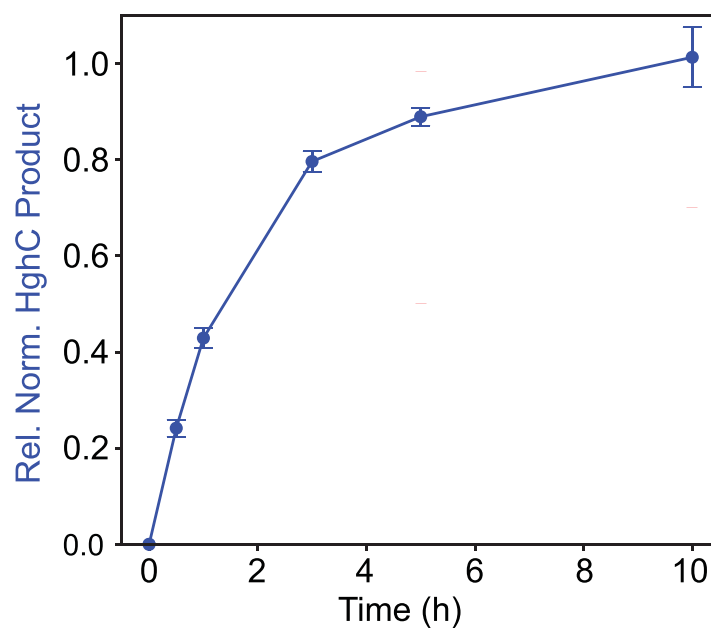
Reprints and permissions information is available at www.nature.com/reprints.



Extended Data Fig. 1 | HR-MS/MS analysis of *S. agalactiae* HghCD and HghBCD reaction products. (a) All y-ions N-terminal to His34 and b-ions C-terminal to His34 for the HghCD product are -2 Da relative to substrate. The same y-ions from the HghBCD product are -4 Da, while b40 remains 2 Da lighter. This pattern points to modifications at His34 within the 'HGH' motif. **(b)** Trypsinolysis of the HghBD or HghBCD products gives the fragments shown. The C-terminal 8mer captures both -2 Da (HghBD) and -4 Da (HghBCD) products. Cleavage does not occur at the Lys residue in red.



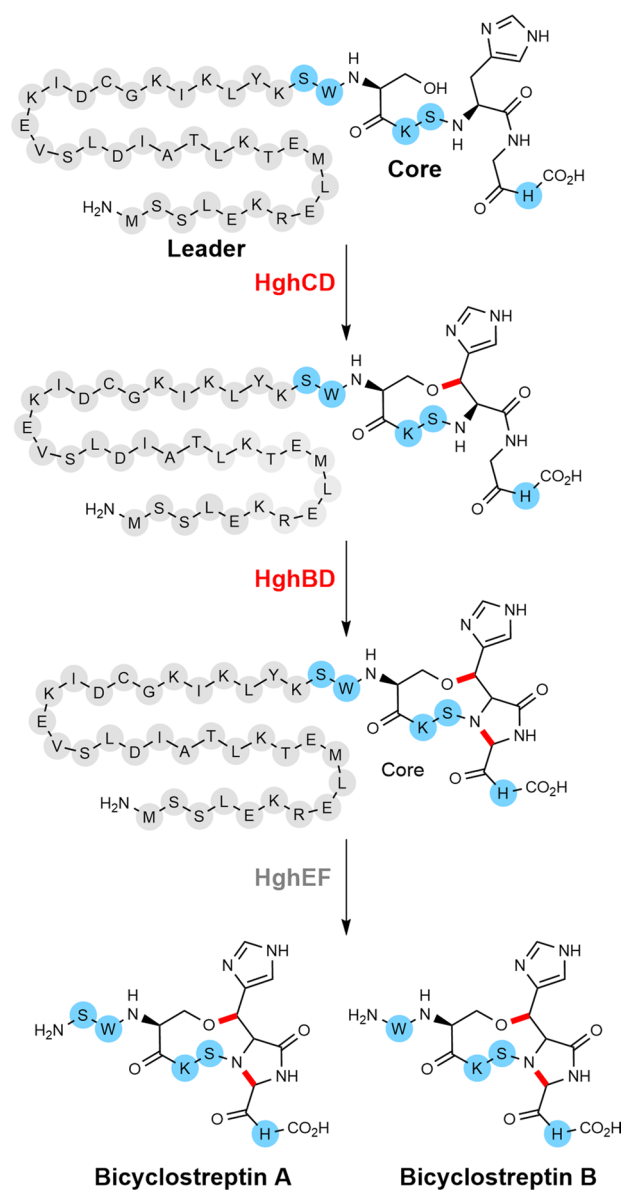
Extended Data Fig. 2 | HghC and HghB are competent in reductive cleavage of SAM to generate 5'-dA. (a) Reaction of HghC or HghB with SAM in the absence of HghA. Shown are HR-MS extracted ion chromatograms for 5'-dA after an overnight reaction (18 h). The cleaved 5'-dA product is seen with both reductants but not in the absence of enzyme. (b) Detection of 5'-dA after the reaction of HghC or HghB with substrate and SAM after 18 h. Note, different elution programs were carried out in panels a and b, thus explaining the different retention times for 5'-dA (see Methods).



Extended Data Fig. 3 | Time-dependent formation of the HghC product as measured by HPLC-Qtof-MS extracted ion chromatography. The average of three independent samples are shown; bars represent standard error.

PqqD	---MQKTSIVAFRRGYRLQWEAAQESHVILYPEGMAKLN---ETAAAILVDGRRDVAA	54
WgkC	---MKEHFVIKGRDFIVN---KVADEYIGYDRLDLEYYSFDEIGAEILYCISKNFSLDK	54
HghD-Sagal	-MQISKNWFIREFPNFQT-----EQNTILFNKLNGLYEVTEVYYDFIDSLRSG---TD	50
HghD-Stherm	MHINNKVWFIRFEPNFQG-----DKKTILYNKSSGELYEVSEIYFHFLESIRNK---KS	51
	.: .: .: .: .. * : . : * : :	
PqqD	IIAMLNERF-----PEAGGVDDDDVIE---FLQIACQQKWITCREPE	92
WgkC	IVELLKQDYEVSEDECKQAIISFLEETPILHIIYANLVKSDIYQLK-----PFREK-	106
HghD-Sagal	FQTLLLRKYKIKKEDVFEIENTILKNLISLGVINE-----	86
HghD-Stherm	CQDALKQRYNTENIDIEKIIQQIKENLLNLGVILND-----	87
	* . : : :	

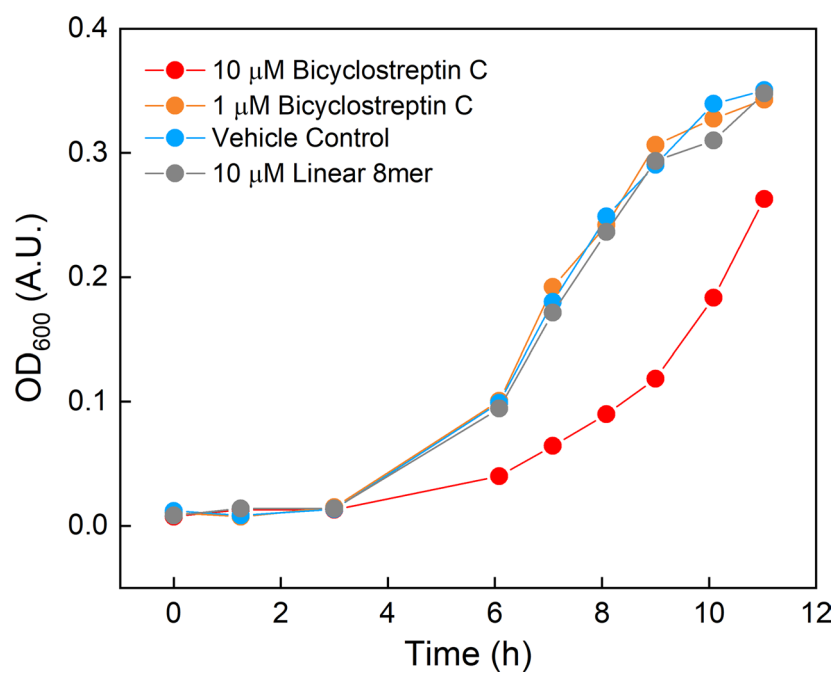
Extended Data Fig. 4 | Sequence alignment of HghD from *S. agalactiae* or *S. thermophilus* with PqqD and WgkC. No substantial homology is observed.



Extended Data Fig. 5 | Biosynthetic pathway of bicyclostreptins A and B from *S. thermophilus* JIM 8232. Blue spheres represent unmodified amino acids in the core region with the indicated one-letter code. Gray spheres are amino acids in the leader sequence.

	Bicyclostreptin C	Monocyclic 8mer	Linear 8mer	Bicyclostreptin A	Linear 8mer
<i>Streptococcus thermophilus</i> JIM8232	9.5	>74.2	>74.0	<0.54	>69.2
<i>Streptococcus thermophilus</i> LMG18311	13.1	>74.2	>74.0	<0.27	>69.2
<i>Streptococcus thermophilus</i> LMD-9	>74.		>74.0		>69.2
<i>Streptococcus agalactiae</i> SGBS015	>74.3		>74.0		>69.2
<i>Streptococcus agalactiae</i> MRI Z1-218	>74.3		>74.0	>69.5	>69.2
<i>Streptococcus agalactiae</i> A909	>74.3		>74.0		>69.2
<i>Escherichia coli</i> K12	>74.3				>69.2
<i>Pseudomonas aeruginosa</i> PAO1	>74.3				>69.2
<i>Staphylococcus aureus</i> Newman	>74.3				>69.2
<i>Enterococcus faecalis</i> OG1RF	>74.3				>69.2

Extended Data Fig. 6 | Summary of growth inhibition assays of bicyclostreptins A and C and the variants shown against select bacteria. Shown are MICs in μM (black and red) or IC_{50} values in μM (blue), which were calculated using the averages of two independent biological samples.



Extended Data Fig. 7 | Effect of bicyclostreptin C on the growth of *S. thermophilus* JIM 8232. The averages of two biologically independent samples are shown.

Reporting Summary

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

Statistics

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

n/a Confirmed

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

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Agilent MassHunter Workstation Data Acquisition 10.0
Agilent OpenLab CDS 2.6
BioTek Synergy/H1 microplate reader Gen5 2.08
Bruker Biospin Topspin 2.1.6

Data analysis

Agilent MassHunter Workstation Qualitative Analysis 10.0
MestReNova x64 14.1
OriginPro 2020
Microsoft 365 2205
PerkinElmer ChemOffice Suite 2019
Adobe Illustrator 26.0
JGI IMG (<https://img.jgi.doe.gov>)

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

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

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

Plasmids and strains used in this study are described in Supplementary Tables 1 and 2. All oligonucleotides are shown in Supplementary Table 3. The sequences of codon-optimized gene fragments are provided in Supplementary Note 1. Other relevant data supporting the findings of this study are available within the paper and the supplementary material. Raw NMR data used to elucidate natural product structures are available from the corresponding author upon request.

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size	All experiments were carried out in at least two independent biological replicates. This is standard in the field and is justified by the low error and high degree of reproducibility observed. This
Data exclusions	No data were excluded from any analyses.
Replication	All experimental results could be reliably replicated. All experiments were conducted independently at least twice.
Randomization	No experimental groups were used in this study, and randomization was therefore not relevant.
Blinding	Group allocation was not relevant to the experiments in this study.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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