Phylogeny-Informed Synthetic Biology Reveals Unprecedented Structural Novelty in Type V Glycopeptide Antibiotics

Min Xu, Wenliang Wang, Nicholas Waglechner, Elizabeth J. Culp, Allison K. Guitor, and Gerard D. Wright*

ABSTRACT: The rise and dissemination of glycopeptide antibiotic (GPA)-resistant pathogens in healthcare settings fuel efforts to discover GPAs that can overcome resistance. Members of the type V subclass of GPAs can evade common GPA resistance mechanisms and offer promise as new drug leads. We characterize five new type V GPAs—rimomycin-A/B/C and misaugamycin-A/B—discovered through a phylogeny-guided genome mining strategy coupled with heterologous production using our GPAHex synthetic biology platform. Rimomycin is a heptapeptide similar to kistamicin but includes an N-methyl-tyrosine at amino acid 6 (AA6) and substitutes 4-hydroxyphenylglycine for tyrosine and 3,5-dihydroxyphenylglycine at positions AA1 and AA3. Misaugamycin is characterized by an unprecedented N–C cross-link between AA2 and AA4 and unique N-terminal acylation by malonyl (misaugamycin-A) or 2-sulfoacetyl (misaugamycin-B) groups. We demonstrate that rimomycin-A/B/C and misaugamycin-A/B are potent antibiotics with activity against GPA-resistant clinical isolates and that the mode of action is consistent with the inhibition of cell division by the evasion of autolysin activity. These discoveries expand the chemical diversity of the type V GPAs, offer new chemical scaffolds for drug development, and demonstrate the application of the GPAHex platform in mining GPA chemical “dark matter”.

INTRODUCTION

Glycopeptide antibiotics (GPAs) such as vancomycin and teicoplanin are natural product drugs that are vital for treating serious infections caused by Gram-positive pathogens, particularly multi-drug-resistant enterococci and Staphylococcus aureus.1 GPAs have gradually lost their efficacy against life-threatening pathogens because of the development of resistance.2 Most GPAs inhibit bacterial growth through binding to the d-alanyl-d-alanine (d-Ala-d-Ala) terminus of the peptidoglycan (PG) stem peptide, thereby blocking the cross-linking and extension of the polymer.3 Replacing the terminal d-Ala-d-Ala dipeptide by d-Ala-d-lactate (d-Ala-d-Lac) through linking and extension of the polymer.4-6 However, the foreseeable development of resistance to these second-generation GPAs requires continued effort to discover and develop more effective GPAs and other Gram-positive-directed antibiotics.7

GPAs are grouped into five (I–V) structural subtypes.13 Type I GPAs such as vancomycin have aliphatic amino acids at AA1 and AA3. Type II–IV GPAs are composed entirely of aromatic amino acids. The chemical structures of GPAs are further diversified by postaglycone modifications such as methylation, glycosylation, sulfation, and acylation.14 Type I–IV GPAs bind d-Ala-d-Ala as described above to achieve their antibacterial activity. On the other hand, type V GPAs such as complestatin and kistamycin are characterized by the presence of a D-E biaryl ring system composed of tryptophan (Trp) and a central 4-hydroxyphenylglycine (Hpg), and corbomycin15 and GP673816 possess a nonapeptide scaffold distinct from that of the canonical type I–IV GPA heptapeptides (Figure 1A). Unlike type I–IV GPAs, type V GPAs lack postaglycone modifications including glycosylation. We recently reported that type V GPAs, including complestatin, corbomycin, and GP6738, exhibit a different mode of action (MOA) than d-Ala-d-Ala-binding type I–IV GPAs.15,16 Instead, type V GPAs bind PG and impair bacterial growth through the indirect inhibition of autolysins, which are essential PG hydrolases required to remodel the cell wall during elongation and division.15 Consequently, type V GPAs overcome the d-Ala-d-Lac GPA resistance and exhibit potent antibacterial activity against vancomycin-resistant enterococci.
Therefore, type V GPAs offer a promising new group of antibiotics for further discovery and development.

GPAs are synthesized by multimodular nonribosomal peptide synthetase (NRPS) systems (Figure 1B). The genes encoding the biosynthetic machinery of GPAs and other NPs are clustered in the chromosome in biosynthetic gene clusters (BGC). GPA BGCs possess genes required for peptide assembly, the amino acid precursor supply, scaffold cross-linking and modification, resistance, regulation, and transport. In addition to common amino acids, GPAs include several nonproteinogenic amino acid building blocks, including Hpg, 3,5-dihydroxyphenylglycine (Dpg), and \( \beta \)-hydroxytyrosine (\( \beta \)HT). The NRPS mega-enzymes assemble the peptide scaffold through collaboration among the adenylation (A), thiolation/peptidyl carrier protein (T/PCP), and condensation (C) domains. Amino acid building blocks are recognized and activated by the A domains and then loaded onto the adjacent T domain, which is then condensed to the downstream T-tethered amino acid residue via an amide bond by the following C domain. During the extension of the peptide chain, epimerization (E) domains and the methyltransferase (MT) domain may also be involved to alter the L-amino acid to the D-}

Figure 1. Structures of type V GPAs and the schematic biosynthetic pathway of complestatin. (A) Chemical structures of type V GPAs. Type V GPAs are classified by the presence of the conserved central Trp-Hpg-(m)Tyr cross-linked structural motif. Kistamicin has an extra A–O–B ring, and corbomycin has two extra B–C and G–O–H ring structures. Most GPAs are heptapeptides; however, corbomycin and GP6738 represent the only two known nonapeptide GPAs. (B) Complestatin biosynthetic pathway as an example of GPA biosynthesis. The core NRPS scaffold genes (comA-comD), MbtH gene (comE), Na+/H+ antiporter gene (comF), halogenase gene (comH), and P450 monooxygenase genes (comJ-comM) are shown. NRPS domains are labeled as A, adenylation; C, condensation; T, thiolation or peptidyl carrier protein; E, epimerization; MT, methyltransferase; X, Ox-recruiting domain; and TE, thioesterase. The halogenation catalyzed by ComH takes place when the amino acid building block is tethered to the T domain. The characteristic ring cross-linking reactions are performed by specific Oxy proteins as shown. Numbering of the rings on the GPA scaffolds is shown in bold font.
configuration and install N-methyl groups. The chlorination of T-tethered amino acid residues and the β-hydroxylation of T-tethered tyrosine can also occur during the elongation process. Multiple rounds of similar elongation steps result in the linear peptide scaffold of GPAs. Subsequently, the T-tethered GPA peptide scaffolds are cross-linked through sequential oxidative cyclization catalyzed by the P450 monooxygenases (OxyB > (OxyE) > OxyA > OxyC in type I–IV GPA biosynthesis) to generate the rigid 3D structure that is characteristic of GPAs. The Oxy proteins are recruited by a conserved penultimate Oxy-recruiting domain, X-domain, to catalyze the oxidative cyclization. One Oxy protein is typically responsible for the installation of one cross-link in GPA biosynthesis; however, OxyC from the kistamicin BGC was recently shown to perform dual oxidative cyclization in kistamicin biosynthesis. OxyC from the corbomycin BGC is also believed to mediate dual oxidative cyclization in corbomycin biosynthesis, suggesting that the Oxy proteins may represent a new diversification node for the generation of novel GPAs. The cross-linked GPA scaffolds are then released from the NRPS assembly line by terminal thioesterase (TE) domain-catalyzed hydrolysis. Subsequently, methylation, sulfation, glycosylation, and acylation may follow to further diversify the GPA scaffolds.

Our identification of new type V GPAs used a phylogenomic approach based on the analysis of the antibiotic BGCs, a strategy that revealed dozens of novel GPA BGCs. This GPA reservoir has not been explored mainly because of the "cryptic/silent" nature of the clusters, where either the organisms produce the compounds in a low yield or often do not express them at all under laboratory conditions. To access this untapped GPA "dark matter", we developed the GPAHex (glycopep-

Figure 2. In silico analysis of the GPA BGCs. (A) MinHash distance phylogeny computed with a mashtree. Nodes with bootstrap support 80 and above are indicated by red circles. BGC numbers (Supporting Information) identified from each actinobacterium are also used to label each strain. Most type V GPA BGCs are found in Streptomyces and are grouped on the top half of the tree. Streptomyces strains are labeled with different squares according to the type V GPA BGCs they possess: orange (Comp, complestatin), green (Corb, corbomycin), azure (GP6738), light gray (Rmo, rimomycin), dark gray (Misau, misaugamycin), light purple (Deca, decapetide type V GPA), and light blue (Dea, decapetide type V GPA). The black square represents kistamicin BGC. Corbomycin BGC containing strains split into two subclades and misaugamycin BGC containing strains are divided into three subclades in the tree. Class names are abbreviated as Streptosp. (Streptosporangiales) and Micros. (Micromonosporales). Strains from the class of Pseudonocardiales form two subclades in the tree. Species names are abbreviated as S. (Streptomyces), Amy. (Amycolatopsis), Non. (Nonomuraea), Herb. (Herbidospora), Actp. (Actinoplanes), Micr. (Micromonospora), Actm. (Actinomadura), Kib. (Kibdellosporangium), Sacm. (Saccharomonospora), Actk. (Actinokineospora), and Noc. (Nocardia). (B) Gene organization of rimomycin-related BGCs. Rimomycin BGCs show identical organization across the five Streptomyces strains. The domain organization of the NRPS assembly line is shown. Hpg biosynthetic genes are labeled as hmo, 4-hydroxymandelate oxidase; hmas, 4-hydroxymandelate synthase; hpgT, 4-hydroxyphenylglycine transaminase; and pdh, prephenate dehydrogenase. Predicted amino acid building blocks loaded by the adenylation domains are shown. (C) Gene organization of misaugamycin-related BGCs. Misaugamycin BGCs show almost identical organization across the six Streptomyces strains except for the presence of two additional putative transposase genes (misL and misM) in WAC00631, S. fradiae ATCC 19609, and S. fradiae olig 1-1. BGCs of the other type V GPAs are shown in Figures S1–S5. NRPS domains are labeled as in Figure 1B.
tide antibiotics heterologous expression) synthetic biology platform. Using GPAHex, two new cryptic GPAs, GP1416 and GP6738, were identified and characterized. Here we describe the discovery and characterization of five novel type V GPAs—rimomycin-A/B/C (named for its BGC that is widely distributed in Streptomyces rimosus strains) and misaugamycin-A/B (named for its Mississauga, Ontario, Canada geographical source)—using phylogeny-guided genome mining of GPA BGCs coupled with heterologous expression using the GPAHex platform, which we improve through the addition of new regulatory elements. We show that these new GPAs are consistent with the type V subclass but have unprecedented structures, cross-linking patterns, and an improved antimicrobial spectrum. This study further demonstrates the power of GPAHex in mining novel cryptic GPAs.

RESULTS AND DISCUSSION

Phylogeny-Guided Genome Mining of Type V GPA BGCs. From published genomes and our in-house sequenced genomes, we identified 116 candidate BGCs using the GPA fingerprint sequences (oxyB, oxyC, halI, dpgC, and oxyE) and BLASTp followed by antiSMASH analysis. All of the GPA BGCs were retrieved from organisms in the phylum Actinobacteria. Using the available whole genome sequences (n = 97), we constructed a multiple-species phylogeny tree using MinHash distances (Figure 2A). Unlike our previous tree built using concatenated single-copy TIGRFam sequences, MinHash distances use kmer information from the entire genome. Because most type V GPA BGCs are found in Streptomyces (kistamicin from Nonomuraea is the only exception), the associated strains are grouped in the top half of the tree and the D-Ala-D-Ala binding GPA BGCs are more widely distributed among other genera. Strains harboring more than one GPA BGC are indicated by multiple leaf nodes. Compared to the first description of the GPA BGCs species phylogeny tree, including 71 GPA BGCs, we have dramatically expanded (18 vs 42) the number of identified type V GPA BGCs. These type V GPA BGCs represent an untapped reservoir of novel type V GPA chemical entities.

Strains containing the same GPA BGCs are closely related and form subclades (Figure 2A). For example, isolates possessing complestatin and GP6738 BGCs are clustered. Similarly, strains with corbomycin BGCs are grouped, but these are further split into two subclades, indicating that corbomycin BGC may have undergone mobilization events. Besides the complestatin, corbomycin, and GP6738 BGCs, two more GPA BGCs shared by different subclades attracted our attention. The rimomycin and misaugamycin BGCs (Figures 2B,C) are highly conserved in the clustered strains, and their domain structures are closely related to the complestatin BGCs. However, the structure and function of GPAs encoded by these two BGC subgroups are unknown, encouraging us to explore these candidates further.

Beyond the known heptapeptide and nonapeptide scaffolds in the type V GPAs (Figures S1–S3), octapeptide and decapeptide scaffolds are also identified in our phylogenetic analysis (Figure 2A and Figures S4 and S5), indicating further chemical diversity in this subgroup of GPA.

GPAHex Production and Characterization of Rimomycin. The genome of WAC06783 was sequenced under our previous Illumina genome sequencing program...
WAC06783 using our standard protocols.15,33 Furthermore, biaryl ether cross-link may be shaped di-cross-link in rimomycin. However, the other Hpg4-(m)Tyr6 indicates that there should be a canonical Trp2-Hpg biaryl predicted molecular formulas C61H52N8O14 (A, calculated [M + 1155.3281, (C) 1189.2900 (Figure S8), correlating to the 1189.2902). Compared to the molecular formula of comple-

cermEp

Streptomyces coelicolor pGP6783 was mobilized into the GPAHex chassis strain, rimomycin should be chlorinated. The P450-OxyA rmo

OxyC

synthetic biology platform16 to produce rimomycin and hygromycin, and others. We therefore turned to our GPAHex cated by its intrinsic resistance to kanamycin, apramycin, WAC06783 was recalcitrant to genetic manipulation, compli-
in all three rimomycin analogs (A, B, and C;Figure S9).

We could not detect significant GPA production by strain WAC06783 using our standard protocols.15,33 Furthermore, WAC06783 was calcitrant to genetic manipulation, compi-
cated by its intrinsic resistance to kanamycin, apramycin, hygromycin, and others. We therefore turned to our GPAHex synthetic biology platform16 to produce rimomycin and characterize its structure (Figure 3A). A 70,290 bp DNA fragment covering the predicted rimomycin BGC in the chromosome of WAC06783 was cloned into the pCGW vector using transformation-associated recombination (TAR) in yeast,14 resulting in plasmid pGP6783 (Figure S7A,B). pGP6783 was mobilized into the GPAHex chassis strain, Streptomyces coelicolor M1154/pAMX4, through Escherichia coli–Streptomyces triparental mating35 for heterologous ex-

erpression. Comparative metabolic analysis of the high-performance liquid chromatography (HPLC) chromatograms identified a series of distinct peaks correlated to the introduction of pGP6783 (Figure 3B). High-resolution quadrupole time-of-flight mass spectrometry (HR-QTOF-MS) reveals three major signal mass values ([M + H]+, m/z) of (A) 1121.3679, (B) 1155.3281, (C) 1189.2900 (Figure S8), correlating to the predicted molecular formulas C61H52N8O14 (A, calculated [M + H]+): 1121.3681, C61H52N8O14Cl (B, calculated [M + H]+): 1155.3292, and C61H52N8O14Cl2 (C, calculated [M + H]+): 1189.2902. Compared to the molecular formula of comple-

statin, C61H52N8O14Cl2 rimomycin has one additional nitrogen and hydrogen atom but lacks an oxygen atom. According to the nitrogen rule, the HRMS data indicates that rimomycin may have one free amine instead of the α-keto group in the starter unit of complestatin. Because rimomycin and complestatin possess 61 carbons, there should be no acylation on the N-terminal Hpg. Substituting a ketone group with an amine group results in a net increase of three hydrogen atoms. However, rimomycin possesses only one more hydrogen atom than complestatin, consistent with an additional intramolecular cross-link.

To elucidate the chemical structure of rimomycin, an overproduction strain was constructed by introducing the constitutively expressed strR regulator-staQ17 driven by the ermEp* promoter in pJ10257 into S. coelicolor M1154/pAMX4/pGP6783. The overexpression of staQ from A47934 BGC results in a 2-fold increase in A and a total 4.2-fold increase in all three rimomycin analogs (A, B, and C; Figure S9). Rimomycin-A was purified, and the structure was determined by one- and two-dimensional nuclear magnetic resonance spectrometry (1D/2D NMR) analyses (Figures S10–S16 and Table S5). The structure of rimomycin-A is shown in Figure 3C. As expected from the HRMS data, no acyl modification was observed, indicating that the C-starter domain may function as a structural domain to initiate the NRPS assembly line instead of catalyzing amide bond formation between an acyl group and the N-terminal free amine on Hpg1. The predicted extra A–O–B biaryl ether cross-link was assigned according to the observation of the 1H–13C HMB correlations of H6 (δ 7.12 ppm)/D–C3 (δ 149.3 ppm) and C–H5 (δ 6.98 ppm)/D–C3 (δ 149.3 ppm), which was further confirmed by the observation of the NOSEY correlations of A–H6 (δ 4.94 ppm)/B–H3 (δ 6.88 ppm), A–H6 (δ 4.94 ppm)/B–H5 (δ 7.25 ppm), and A–H6 (δ 4.94 ppm)/B–H6 (δ 8.00 ppm) (Figure S15). The C–O D cross-link was assigned according to the observation of the 1H–13C HMB correlations of D–H6 (δ 7.18 ppm)/E–C6 (δ 134.8 ppm), E–H5 (δ 6.89 ppm)/D–C5 (δ 131.2 ppm), and E–H7 (δ 7.32 ppm)/D–C5 (δ 131.2 ppm) and was further confirmed by the observation of the NOSEY correlations of D–H6 (δ 5.18 ppm)/E–H4 (δ 7.50 ppm), D–H6 (δ 5.18 ppm)/E–H5 (δ 6.89 ppm), and D–H6 (δ 5.18 ppm)/E–H7 (δ 6.75 ppm) (Figure S15). Given that rimomycin-B and rimomycin-C are mono- and dichlorinated analogs of rimomycin-A, we first used tandem MS/MS to locate the chlorination sites on rimomycin-A. The MS2 production ion fragments of rimomycin-A/B/C, as shown in Figure S17, clearly show the chlorination sites on rimomycin-B (Hpg5) and rimomycin-C (Hpg3 and Hpg5). The structures of rimomycin-B/C were further confirmed through 1D/2D NMR spectroscopy as shown in Figures S18–S29 and Tables S6 and S7. Along with kistamicin,37 rimomycin is the second example of a type V GPa incorporating a 15-membered A-O-B ring linkage. Interestingly, we detected some production of rimomycin-A/B in the parental strain, WAC06783, after treating the crude extract with a 30% MeOH/H2O (v:v) wash, although we missed it in our initial analysis. A comparison of the crude extract of WAC06783 and the MeOH-treated samples showed that the rimomycin-A/B signal in the crude extract is masked by the massive production of oxytetracycline and rimocidin polyes (Figure S30). Accordingly, we were able to locate the BGCs for oxytetracycline38 and rimocidin39 in the genome of WAC06783 (Figure S31). This masking of small quantities of a novel antibiotic by the production of known antimicrobial compounds further demonstrates the value of targeted discovery platforms such as GPAHex.

GPAHex Production and Characterization of Misau
gamycin. The whole-genome sequence of WAC00631 was determined using a combination of Illumina and MinION nanopore platforms. Assembly using a hybrid approach generated a draft genome with 21 contigs, in which the misaugamycin BGC was identified on contig 4 using antiSMASH30 (Figure 2C and Table S4). Like rimomycin, the misaugamycin BGC shows extensive similarity to the comple-

statin BGC (93% according to antiSMASH). The prediction of the...
the A domain specificity reveals a heptapeptide scaffold of Hpg-Trp-Hpg-Hpg-Tyr-Hpg, identical to that of rimomycin and complestatin. Because there is no MT domain in module 6, there should be a Tyr6 in misaugamycin instead of a mTyr6 in rimomycin and complestatin. A starter C domain was also identified in misaugamycin BGC, indicating the presence of a putative N-terminal acylated Hpg1. The presence of the halogenase encoding gene (misH) predicts that misaugamycin is chlorinated. P450-OxyCmis forms a monophyletic clade with the OxyCs from GP6738 and corbomycin BGCs; however, P450-OxyAmis forms a monophyletic clade distinct from the OxyAs in complestatin, corbomycin, GP6738, and rimomycin.

**Figure 4.** Discovery of misaugamycin from WAC00631 using GPAHex. (A) Chemical structures of misaugamycin-A/B. The central Hpg in misaugamycin-A/B has been reduced to form the unique ocHeg residue. The key COSY, HMBC, and NOSEY correlations observed from the NMR spectra are labeled with bold lines, black/red curved arrows, and blue double-headed curved arrows, respectively. The \(^1H-^{15}N\) correlations (red arrows) observed in HMBC experiments are optimized for \(J_{HN} = 8\) Hz. Misaugamycin-A/B are differentiated by the substitution of the R group with carboxylic acid (A) or sulfonic acid (B) group. (B) 2D \(^1H-^{13}C\) and \(^1H-^{15}N\) HMBC spectra of the correlations between Trp2 and ocHeg4 residues in misaugamycin-A/B.
BGCs in the P450 phylogenetic tree (Figure S6). On the basis of the P450 phylogeny, we predict a Hpg4-Tyr6 biaryl ether cross-link in misaugamycin. Interestingly, a putative four-gene cassette misNOPQ encodes an acyl carrier protein (ACP), an acyl-CoA ligase, a phosphosulfolactate synthase, and a glyoxalase that appears to form one operon, surrounded by putative InsQ and ligase. We hypothesized that these genes are involved in the biosynthesis of the N-terminal acyl chain moiety of misaugamycin. However, the characteristic Trp2-Hpg4 biaryl ether cross-link in misaugamycin indicates that there should be an introduction of a novel acyl/sulfur modification. The predicted molecular formula of misaugamycin indicates that there should be an introduction of a novel acyl/sulfur modification. The predicted molecular formula of misaugamycin indicates that there should be an introduction of a novel acyl/sulfur modification.

Because misaugamycin production was not detectable from the fermentation of WAC00631, we applied our GPAHex synthetic biology platform to its heterologous production. A 78,889 bp DNA region covering the predicted misaugamycin BGC in the chromosome of WAC00631 was cloned into pCGW78,889 bp DNA region covering the predicted misaugamycin BGC in the chromosome of WAC00631 was cloned into pCGW (Figure S7C,D). The pGP631 plasmid was mobilized into the GPAHex chassis strain, C6 (Trp2)-C5 (Hpg4) biaryl cross-link is replaced by an N1 (Trp2)-C1 (Hpg4) cross-link. The previous C3 (Hpg4)-C4 (Tyr6) biaryl ether cross-link was transformed into a C3 (Hpg4)-O-C4 (Tyr6) ether cross-link. The C-O-D cross-link was assigned according to the observation of the key 1H-13C HMBC correlation of D-H3 (δ 4.32 ppm)/C-4 (δ 153.73 ppm) and 1H-13C HMBC correlation of D-H3 (δ 4.32 ppm)/C-4 (δ 153.73 ppm) and 1H-13C HMBC correlation of D-H3 (δ 4.32 ppm)/C-4 (δ 153.73 ppm) and 1H-13C HMBC correlation of D-H3 (δ 4.32 ppm)/C-4 (δ 153.73 ppm) and 1H-13C HMBC correlation of D-H3 (δ 4.32 ppm)/C-4 (δ 153.73 ppm) and 1H-13C HMBC correlation of D-H3 (δ 4.32 ppm)/C-4 (δ 153.73 ppm), which was further confirmed by the observation of NOSEY correlations of D-H2 (δ 2.02 ppm)/C-H3 (δ 6.43 ppm) and D-H2 (δ 2.02 ppm)/C-H5 (δ 6.98 ppm) (Figures S39 and S40).
The unprecedented N1-Trp2-C1(ocHeg4) cross-link in misaugamycin-A/B was assigned by the observation of the key $^{1}$H−$^{13}$C HMBC correlation of E-H2 ($\delta$ 6.60 ppm)/D-C1($\delta$ 63.89 ppm) (Figure 4B and Figures S38 and S46) and the NOSEY correlations of D-H5 ($\delta$ 6.40 ppm)/E-H2 ($\delta$ 6.60 ppm), D-H6 ($\delta$ 7.94 ppm)/E-H2 ($\delta$ 6.60 ppm), D-H2 ($\delta$ 3.48, 2.02 ppm)/E-H7 ($\delta$ 8.13 ppm), D-H3 ($\delta$ 4.32 ppm)/E-H7 ($\delta$ 8.13 ppm), and D-H$^\alpha$ ($\delta$ 5.50 ppm)/E-H7($\delta$ 8.13 ppm) (Figures S39 and S47). To confirm this unique N−C linkage, $^{1}$H−$^{15}$N heteronuclear single quantum correlation (HSQC) experiments were performed. As expected, seven $\alpha$-NH’s from the heptapeptide scaffold were identified, and the indole amine signal was missing from the HSQC spectra for both misaugamycin-A/B (Figures S40 and S48). The N−C linkage was further supported by observing the key $^{1}$H−$^{15}$N HMBC correlation of D-H2 ($\delta$ 2.02 ppm)/E-N1 ($\delta$ 138.5 ppm) for misaugamycin-A/B (Figure 4B and Figures S41 and S49).

Additionally, the N-terminal Hpg was acylated by malonyl (misaugamycin-A) and 2-sulfoacetyl (misaugamycin-B) groups, which are both unprecedented GPA modifications. Using Na$_2$SO$_3$ to supplement the fermentation culture of S. coelicolor M1154/pAMX4/pGP631/pIJ-staQ led to the favored production of sulfonated misaugamycin-B (Figure S32).

**Rimomycin and Misaugamycin-A/B Inhibit Autolysins.**

Type V GPAs show potent antibacterial activity against multiple Gram-positive bacteria, including multi-drug-resistant pathogens. These GPAs display a novel MOA by blocking autolysin action through binding to PG, distinct from conventional GPAs that bind to the terminal D-Ala-D-Ala on the PG stem peptide. Rimomycin-A/B/C and misaugamycin-A/B show broad-spectrum antibacterial activity against Gram-positive bacteria, including methicillin-resistant S. aureus (MRSA) and vancomycin-resistant enterococci (VREA and VREB) (Table 1). Interestingly, rimomycin-A/B/C show a broader antibacterial spectrum with additional antibacterial activity against the efflux pump and outer-membrane-compromised E. coli BW25113 $\Delta$bamB$\Delta$tolC, Mycobacterium smegmatis mc$^\text{2}155$, and Mycobacterium tuberculosis H37Ra, which has not been observed in other type V GPAs. Kistamicin possesses an A-O-B/C-O-D/D-E tricyclic scaffold identical to that of...
rimomycin; therefore, we purified this compound from the producer Nonomuraea sp. ATCC 55076. Kistamicin shows no activity against E. coli BW25113 ΔbamBΔtolC and M. smegmatis mc²155 (Table 1). The subtle changes in the peptide scaffolds in the reduction of the central Hpg4 to form OcHeg4. Ring-closed heptapeptide scaffolds through the unprecedented C-D-E (3). A typical C-O-D (1) > D-E (2) ring-closure sequence was proposed for misaugamycin biosynthesis. During the closure of the D-E ring through the unprecedented C-N linkage catalyzed by OxyA, in misaugamycin biosynthesis, the F420-dependent reductase, MisW, is proposed to take part in the reduction of the central Hpg4 to form OcHeg4. Ring-closed heptapeptide scaffolds are cleaved from the NRPS assembly line by the terminal TE domain to release rimomycin and misaugamycin.

![Figure 6. Proposed biosyntheses of (A) rimomycin and (B) misaugamycin. OxyA and OxyC in rimomycin and misaugamycin BGCs are recruited to the NRPS-tethered heptapeptide substrates by the conserved X-domain to mediate the oxidative cyclization. Because OxyC RMS is located in a distinct clade from OxyC RMS in the P450 phylogenetic tree, we proposed the following ring-closure sequence in rimomycin biosynthesis: C-O-D (1) > A-O-B (2) > D-E (3). A typical C-O-D (1) > D-E (2) ring-closure sequence was proposed for misaugamycin biosynthesis.](https://pubs.acs.org/doi/10.1021/acscentsci.1c01389) 

### CONCLUSIONS

GPAs are essential for treating infections caused by multi-drug-resistant Gram-positive pathogens. However, the emergence of GPA resistance threatens the efficacy of these antibiotics, so more effective GPAs are needed to help address the antibiotic resistance crisis. Although medicinal chemists have begun to address this problem by introducing second-generation GPAs, such as telavancin, dalbavancin, and oritavancin, into the clinic to combat VRE, there remains a need for novel GPAs with distinct MOAs to offer alternative strategies for the treatment of infections caused by multi-drug-resistant Gram-positive pathogens. Recently, our group characterized members of the poorly studied type V GPA subclass including complestatin, corbomycin, and GP6738 that inhibit bacterial cell division by binding to the PG and consequently blocking the activity of autolysins. This MOA is distinct from the d-Ala-d-Ala binding of other GPAs, enabling the type V GPAs to evade canonical GPA resistance. Expanding the type V GPA chemical diversity offers a promising direction for drug discovery and development.

The advances in genome sequencing have uncovered countless BGCs in the bacterial pan-genome, especially in Actinobacteria, that may encode novel antibiotics. Using a phylogeny-guided genome mining approach, we identified 116 GPA BGCs from public genomes and our in-house sequenced genomes. Many of these GPA clusters belong to the type V GPA BGCs from public genomes and our in-house sequenced genomes. Among them, rimomycin and misaugamycin were shared by several Streptomyces strains. Combined with the GPAHex synthetic biology platform, we captured the two BGCs for heterologous expression. Rimomycin production was increased by 8.5-fold in GPAHex and 37.5-fold when coupled with the overexpression of the StrR regulator compared to the parental strain. Similarly, we were able to express the cryptic/silent misaugamycin BGC in GPAHex and boost the production of misaugamycin-A/B when coupled with the overexpression of
the StrR regulator. As a master regulator of GPA biosynthesis, strR overexpression can significantly increase the output of GAPs.\textsuperscript{42–44} In our case, staQ, the strR from the A47934 BGC, was overexpressed instead of the BGC-associated strR, indicating that staQ can be applied as a general tool to promote GPA production.

Rimomycin-A/B/C and misaugamycin-A/B, new members of the type V GPA subclass, exhibit striking structural novelty. Compared to the typical C-O-D and D-E bicycloskeleton scaffold of the type V GPA,\textsuperscript{13} rimomycin-A/B/C possess an additional 15-membered A-O-B ring, a feature that is also observed in kistamicin.\textsuperscript{20} However, there are only two P450s present in both BGCs despite the presence of three cross-links. Given the difference between the biaryl ether and biaryl linkages, the OxyC protein is believed to catalyze the two biaryl ether cross-links, and the OxyA protein is thought to build the biaryl cross-link. \textit{In vivo} gene inactivation and \textit{in vitro} reconstitution studies in kistamicin biosynthesis support this hypothesis.\textsuperscript{20} It is proposed that the maturation of kistamicin biosynthesis undergoes a sequential cross-linking of C-O-D > D-E > A-O-B ring.\textsuperscript{20} However, OxyC\textsubscript{Kis} clusters more closely to the OxyC proteins from the d-Ala-d-Ala binding GPA BGCs instead of the OxyC proteins from the type V GPA BGCs in the P450 phylogenetic tree (Figure S6). Within the typical OxyC proteins from the type V GPA BGCs, OxyC\textsubscript{Kis} proteins form a monophyletic clade separate from ComJ and the OxyC proteins from the BGCs of corbomycin, GP6738, and misaugamycin. This may indicate that cross-linking in rimomycin biosynthesis is distinct from that of kistamicin (Figure 6A). Further studies are ongoing to elucidate the mechanism for the P450-catalyzed cross-linking in rimomycin-A/B/C biosynthesis.

Misaugamycin-A/B are distinct from all previously described GAPs, recruiting malonyl and 2-sulfoacetyl groups as the acyl tails and installing a reduced and tautomerized central Hpg4 and GPA, recruiting malonyl and 2-sulfoacetyl group (Figure S54). We propose that the terminal F\textsubscript{ACS}-dependent oxidoreductase coding gene that is conserved across the misaugamycin-like BGCs accounts for the reduction of Hpg4. The unique C–N linkage is predicted to be installed by the OxyC\textsubscript{mis} protein, which catalyzes the C–C biaryl linkage between Hpg4 and Trp2 in all of the other type V GAPs (Figure 6B). Although rimomycin-A/B/C and misaugamycin-A/B show distinct cross-link patterns, the X-domain-mediated recruitment of Oxs to perform the oxidative cyclization seems to be conserved in their biosynthesis as observed in other GAPs. The sequence alignment of OxyC\textsubscript{mis}/OxyC\textsubscript{main} and OxyA\textsubscript{mis}/OxyA\textsubscript{main} reveals the presence of the conserved characteristic X-domain recruitment fingerprint “PRRD” motif in the F-helix and the presence of characteristic X-domain interaction residues in the D-, E-, F-, and G-helices\textsuperscript{28} (Figure S55). These features suggest that the OxyC\textsubscript{mis}-catalyzed multiplex biaryl-ether cross-link and the OxyA\textsubscript{mis}-catalyzed C–N cross-link should all be mediated by the interaction between the Oxy proteins and the penultimate X-domain on the NRPS. Additional in-depth \textit{in vivo} and \textit{in vitro} studies are required to clarify the details of the biosynthetic mechanisms of these novel compounds.

Rimomycin-A/B/C and misaugamycin-A/B, like the previously characterized complestatin, corbomycin, and GP6738, interrupt the cell wall degradation steps essential for cell division by blocking the activity of autolysins.\textsuperscript{15,16} The conserved C-O-D and D-E dual-ring structure present in type V GAPs is the essential structural motif to exhibit their antibacterial activity. Reconstruction of the C-O-D and D-E dual-ring linkage through reduction of the central Hpg4 to ocHpg4 and the installation of the C–N-linked D-E ring in misaugamycin preserves its autolysins’ inhibition MOA. However, the more compressed 12-membered D-E ring in misaugamycin compared to the more common 16-membered D-E ring in complestatin, corbomycin, kistamicin, and rimomycin impaired its antibacterial activity. Alternatively, the unprecedented N-terminal acylation (2-sulfoacetyl) modification was introduced to compensate for the potency of misaugamycin. The more rigid 3D structures constructed by the installation of additional cross-links in the peptide scaffold of the type V GAPs (A-O-B cross-link in rimomycin/kistamicin and B-C/G-O-H cross-links in corbomycin) may increase the affinity of the GAPs for PG, contributing to improved antibacterial efficacy.\textsuperscript{16} Beyond the variations in the ring topology, chlorination modification can also improve the antibacterial activity in the type V GAPs. Interestingly, rimomycin-A/B/C show additional antimycobacterial activity and antibacterial activity against the efflux pump and the outer-membrane-compromised \textit{E. coli} BW25113 \textit{ΔbamBΔtolC} strain and improved efficacy against VRE compared to other type V GAPs. We note that rimomycin-A/B/C possess a free N-terminal amino group, providing a suitable site for semisynthesis differentiation to generate new derivatives with improved druglike properties.

The type V GAPs show a novel MOA with low resistance development, properties coveted in drug discovery to mitigate the antibiotic resistance crisis. Expanding this novel functional class of GAPs offers a promising avenue for the development of new drug leads. Phylogenetic analysis revealed dozens of novel GPA BGCs in bacterial genomes, especially those of \textit{Streptomyces}. As the growth of sequenced bacterial genomes continues, many more hidden natural product BGCs encoding unknown molecules will be brought to light. The combination of genomics and synthetic biology has the potential to accelerate innovations in the drug discovery pipeline.

\section*{ASSOCIATED CONTENT}

\subsection*{Supporting Information}

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscentsci.1c01389.

- Materials and methods, mass spectra, and 1D/2D NMR spectra of all compounds (DOCX)
- All information on the strains possessing GPA BGC identified in this study (XLSX)

\section*{AUTHOR INFORMATION}

\subsection*{Corresponding Author}

Gerard D. Wright — David Braley Center for Antibiotic Discovery, Michael G. DeGroote Institute for Infectious Disease Research, Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario L8N 3Z5, Canada; orcid.org/0000-0002-9129-7131; Email: wrightge@mcmaster.ca

\subsection*{Authors}

Min Xu — David Braley Center for Antibiotic Discovery, Michael G. DeGroote Institute for Infectious Disease Research, Department of Biochemistry and Biomedical Sciences,
Complete contact information is available at: https://pubs.acs.org/10.1021/acscentsci.1c01389

Author Contributions
M.X. and G.D.W. conceived the study and designed the experiments. W.W. performed the purification and structural elucidation of the compounds. N.W. performed the genome assembly of the Illumina genome sequencing data and the elucidation of the compounds. N.W. performed the genome assembly of WAC00631. M.X. performed all of the other experiments. M.X. and G.D.W. wrote the paper with input from all authors.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Bob Berno and Dr. Hilary A. Jenkins for help with the nitrogen NMR and Dr. Kalinka Koteva for help with the tandem MS/MS. We thank Dr. Michael Cook for testing the MIC against M. tuberculosis H37Ra. We also thank Haley Zubyk for suggesting the name misaugamycin. This research was funded by the Canadian Institutes of Health Research (grant FRN-148463), the Natural Sciences and Engineering Research Council (grant 237480), the Ontario Research Fund, and a Canada Research Chair to G.D.W. E.J.C. was supported by a CIHR Vanier Canada Graduate Scholarship. N.W. and A.K.G. were supported by a CIHR Canada Graduate Scholarship Doctoral Award.

■ REFERENCES

(1) Zeng, D.; Debabov, D.; Hartsell, T. L.; Cano, R. J.; Adams, S.; Schuyler, J. A.; McMillan, R.; Pace, J. L. Approved Glycopeptide Antibacterial Drugs: Mechanism of Action and Resistance. Cold Spring Harb Perspect. Med. 2016, 6 (12), a026989.
(2) Pootoolal, J.; Neu, J.; Wright, G. D. Glycopeptide antibiotic resistance. Annu. Rev. Pharmacol. Toxicol. 2002, 42, 381–408.
(3) Sheldrick, G. M.; Jones, P. G.; Kennard, O.; Williams, D. H.; Smith, G. A. Structure of vancomycin and its complex with acetyl-D-alanyl-D-alanine. Nature 1978, 271 (5642), 223–5.
(4) Bugg, T. D.; Dutka-Malen, S.; Arthur, M.; Courvalin, P.; Walsh, C. T. Identification of vancomycin resistance protein VanA as a D-alanine-D-alanine ligase of altered substrate specificity. Biochemistry 1991, 30 (8), 2017–21.
(5) Bugg, T. D.; Wright, G. D.; Dutka-Malen, S.; Arthur, M.; Courvalin, P.; Walsh, C. T. Molecular basis for vancomycin resistance in Enterococcus faecium BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins VanH and VanA. Biochemistry 1991, 30 (43), 10408–15.
(6) Wu, Z.; Wright, G. D.; Walsh, C. T. Overexpression, purification, and characterization of VanX, a D-β-Dipeptidase which is essential for vancomycin resistance in Enterococcus faecium BM4147. Biochemistry 1995, 34 (8), 2455–63.
(7) Cui, L.; Ma, X.; Sato, K.; Okuma, K.; Tenerove, F. C.; Mamiruza, E. M.; Gemmell, C. G.; Kim, M. N.; Ploy, M. C.; El-Sohb, N.; et al. Cell wall thickening is a common feature of vancomycin resistance in Staphylococcus aureus. J. Clin. Microbiol. 2003, 41 (1), 5–14.
(8) Cui, L.; Iwamoto, A.; Lian, J. Q.; Neoh, H. M.; Maruyama, T.; Horikawa, Y.; Hiramatsu, K. Novel mechanism of antibiotic resistance originating in vancomycin-intermediate Staphylococcus aureus. Anti-microb. Agents Chemother. 2006, 50 (2), 428–38.
(9) Malabarba, A.; Ciabatti, R.; Scotti, R.; Goldstein, B. P.; Ferrari, P.; Kurz, M.; Andreini, B. P.; Denaro, M. New semisynthetic glycopeptides MDL 63,246 and MDL 63,042, and other amide derivatives of antibiotic A-40,926 active against highly glycopeptide-resistant VanA enterococci. J. Antimicrob. Chemother. (Tokyo) 1995, 48 (8), 869–83.
(10) Cooper, R. D.; Snyder, N. J.; Zweifel, M. J.; Staszak, M. A.; Wilkie, S. C.; Nicas, T. I.; Mullen, D. L.; Butler, T. F.; Rodriguez, M. J.; Huff, B. E.; et al. Reductive alkylation of glycopeptide antibiotics: synthesis and antibacterial activity. J. Antimicrob. Chemother. (Tokyo) 1996, 49 (6), 575–81.
(11) Leadbetter, R. M.; Adams, S. M.; Bazzini, B.; Fatherree, P. R.; Karr, D. E.; Krause, K. M.; Lam, B. M.; Linsell, M. S.; Nodwell, M. B.; Pace, J. L.; et al. Hydrophobic vancomycin derivatives with improved ADME properties: discovery of telavancin (TD-6424). J. Antimicrob. Chemother. (Tokyo) 2004, 57 (5), 326–36.
(12) Blaskovich, M. A. T.; Hansford, K. A.; Butler, M. S.; Jia, Z.; Mark, A. E.; Cooper, M. A. Developments in Glycopeptide Antibiotics. ACS Infect. Dis. 2018, 4 (5), 715–735.
(13) Nicolaou, K. C.; Boddy, C. N.; Brase, S.; Winsinger, N. Chemistry, Biology, and Medicine of the Glycopeptide Antibiotics. Angew. Chem., Int. Ed. 1999, 38 (15), 2096–2152.
(14) Yim, G.; Thaker, M. N.; Koteva, K.; Wright, G. Glycopeptide antibiotic biosynthesis. J. Antibiot. (Tokyo) 2014, 67 (1), 31–41.
(15) Culp, E. J.; Waglechner, N.; Wang, W.; Fiebig-Comyn, A. A.; Hsu, Y. P.; Koteva, K.; Syanchta, D.; Coombes, B. K.; Van Nuenen, H. M.; Brun, Y. V.; et al. Evolution-guided discovery of antibiotics that inhibit peptidoglycan remodelling. Nature 2020, 578 (7796), 582–587.
(16) Xu, M.; Wang, W.; Waglechner, N.; Culp, E. J.; Siviter, A. K.; Wright, G. D. GPAHex-A synthetic biology platform for Type IV-V glycopeptide antibiotic production and discovery. Nat. Commun. 2020, 11 (1), 5232.
(17) van Wageningen, A. M.; Kirkpatrick, P. N.; Williams, D. H.; Harris, B. R.; Kershaw, J. K.; Lennard, N. J.; Jones, M.; Jones, S. J.; Solenberg, P. J. Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. Chem. & Biol. 1998, 5 (3), 155–62.
(18) Chiu, H. T.; Hubbard, B. K.; Shah, A. N.; Eide, J.; Fredenburg, R. A.; Walsh, C. T.; Khosla, C. Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster. Proc. Natl. Acad. Sci. U. S. A. 2001, 98 (15), 8548–53.
(19) Pootoolal, J.; Thomas, M. G.; Marshall, C. G.; Neu, J. M.; Hubbard, B. K.; Walsh, C. T.; Wright, G. D. Assembling the glycopeptide antibiotic scaffold: The biosynthesis of A47934 from Streptomyces toyoaensis NRRL15009. Proc. Natl. Acad. Sci. U. S. A. 2002, 99 (13), 8962–7.
(20) Greule, A.; Izore, T.; Ifime, D.; Tailhades, J.; Schoppet, M.; Zhao, Y.; Peschke, M.; Ahmed, I.; Kulik, A.; Adamek, M.; et al.
Kistamicin biosynthesis reveals the biosynthetic requirements for production of highly crosslinked glycopeptide antibiotics. Nat. Commun. 2019, 10 (1), 2613.

(21) Kittila, T.; Kittel, C.; Tailhades, J.; Butz, D.; Schoppe, M.; Buttner, A.; Goode, R. J. A.; Schittenhelm, R. B.; van Pee, K. H.; Sussmuth, R. D.; et al. Halogenation of glycopeptide antibiotics occurs at the amino acid level during non-ribosomal peptide synthesis. Chem. Sci. 2017, 8 (9), 5992–6004.

(22) Leng, D. J.; Greule, A.; Croy, M. J.; Tosin, M. Chemical probes reveal the timing of early chlorination in vancomycin biosynthesis. Chem. Commun. 2021, 57 (18), 2293–2296.

(23) Stinch, S.; Carrano, L.; Lazzarini, A.; Feroggio, M.; Grigoletto, A.; Sosio, M.; Donadio, S. A derivative of the glycopeptide A40926 produced by inactivation of the beta-hydroxylase gene in Nonomuraea. J. Bacteriol. 2006, 188 (2), 229–35.

(24) Forer, C. C.; Seyed Sayamdost, M. R. In Vitro Reconstitution of OxyC Activity Enables Total Chemoenzymatic Syntheses of Vancomycin Agylycone Variants. Angew. Chem., Int. Ed. 2018, 57 (27), 8048–8052.

(25) Tailhades, J.; Zhao, Y.; Ho, Y. T. C.; Greule, A.; Ahmed, I.; Schoppe, M.; Kulkarni, K.; Goode, R. J. A.; Schittenhelm, R. B.; De Voss, J. J.; et al. A Chemoenzymatic Approach to the Synthesis of Glycopeptide Antibiotic Analogues. Angew. Chem., Int. Ed. 2020, 59 (27), 10899–10903.

(26) Zhao, Y.; Goode, R. J. A.; Schittenhelm, R. B.; Tailhades, J.; Croy, M. J. Exploring the Tetracyclization of Teicoplanin Precursor Peptides through Chemoenzymatic Synthesis. J. Org. Chem. 2020, 85 (3), 1537–1547.

(27) Peschke, M.; Gussoni, M.; Sussmuth, R. D.; Croy, M. J. Understanding the crucial interactions between Cytochrome P450s and non-ribosomal peptide synthetases during glycopeptide antibiotic biosynthesis. Curr. Opin. Struct. Biol. 2016, 41, 46–53.

(28) Haslinger, K.; Peschke, M.; Briere, C.; Maximowitsch, E.; Croy, M. J. X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis. Nature 2015, 521 (7550), 105–9.

(29) Waglechner, N.; McArthur, A. G.; Wright, G. D. Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. Nat. Microbiol. 2019, 4 (11), 1862–1871.

(30) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res. 2019, 47 (W1), W81–W87.

(31) Katz, L. S.; Griswold, T.; Morrison, S. S.; Caravas, J. A.; Zhang, S.; den Bakker, H. C.; Deng, X.; Carleton, H. A. Mash: a rapid comparison of whole genome sequence files. J. Open Source Softw. 2019, 4 (44), 1762.

(32) Selengut, J. D.; Haft, D. H.; Davidsen, T.; Ganapathy, A.; Gwinn-Miller, S.; Nelson, W. C.; Richter, A. R.; White, O. TIGRFAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. Nucleic Acids Res. 2007, 35 (Database issue), D260–D264.

(33) Thaker, M. N.; Wang, W.; Spanogiannopoulos, P.; Waglechner, N.; King, A. M.; Medina, R.; Wright, G. D. Identifying producers of antibacterial compounds by screening for antibiotic resistance. Nat. Biotechnol. 2013, 31 (10), 922–7.

(34) Zhang, J. J.; Yamanaka, K.; Tang, X.; Moore, B. S. Direct cloning and heterologous expression of natural product biosynthetic gene clusters by transformation-associated recombination. Methods Enzymol. 2019, 621, 87–110.

(35) Kieser, T., et al. Practical Streptomyces Genetics, 2nd ed.; John Innes Foundation: Norwich, U.K., 2000.

(36) Hong, H. J.; Hutchings, M. I.; Hill, L. M.; Buttner, M. J. The role of the novel Fem protein VanK in vancomycin resistance in Streptomyces coelicolor. J. Biol. Chem. 2005, 280 (13), 13055–61.

(37) Naruse, N.; Tennyono, O.; Kobaru, S.; Hatori, M.; Tomita, K.; Hamagishi, Y.; Oki, T. New antiviral antibiotics, kistamicins A and B. I. Taxonomy, production, isolation, physico-chemical properties and biological activities. J. Antibiot. (Tokyo) 1993, 46 (12), 1804–11.