Structure and Biosynthesis of the Antibiotic Bottromycin D

Yanpeng Hou,†§ Ma. Diarey B. Tianero,‡§ Jason C. Kwan,‡ Thomas P. Wyche,‡ Cole R. Michel,† Gregory A. Ellis,† Emmanuel Vazquez-Rivera,‡ Doug R. Braun,‡ Warren E. Rose,† Eric W. Schmidt,‡ and Tim S. Bugni*†

Pharmaceutical Sciences Division, University of Wisconsin—Madison, Madison, Wisconsin 53705, United States, and Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112, United States
tbugni@pharmacy.wisc.edu

Received August 15, 2012

ABSTRACT

Drug resistant infectious diseases are quickly becoming a global health crisis. While Streptomyces spp. have been a major source of antibiotics over the past 50 years, efficient methods are needed to identify new antibiotics and greatly improve the rate of discovery. LCMS-based metabolomics were applied to analyze extracts of 50 Streptomyces spp. Using this methodology, we discovered bottromycin D and used whole genome sequencing to determine its biosynthesis by a ribosomal pathway.

Drug resistant infectious diseases are quickly becoming a global health crisis. For example, one resistant organism, methicillin-resistant Staphylococcus aureus (MRSA), causes more deaths in the United States each year than HIV/AIDS, Parkinson’s disease, and emphysema combined.1 Drug resistant infections are now the second leading cause of death worldwide.1 Historically, a major source of antibiotics has been soil-derived bacteria, primarily from the genus Streptomyces, but this genus has been extensively studied for more than 80 years, meaning that the frequency of rediscovery of known compounds is high.

Efficient methods to address this problem could greatly assist with the discovery of new antibiotics. Therefore, we have employed UHPLC/HRMS-based metabolomics for strain selection/dereplication and to efficiently discover novel small molecules from Streptomyces and other actinomycetes.2 Similar methodology has been applied to strains of Myxococcus spp. with success.3,4

In particular, we have used principal component analysis (PCA) as a metabolomics tool to effectively, efficiently, and statistically analyze LCMS profiles from bacteria to identify those producing novel small molecules. We recently used the methodology to aid in the discovery of the microtermolides isolated from an insect-associated Streptomyces sp.5 Herein, we report the discovery of a new bottromycin A2 analog, bottromycin D (1), using LCMS-PCA to analyze 50 marine ascidian-derived Streptomyces spp. as outlined in the Supporting Information (SI). The producing strain WMMB272 was quickly identified as having unique metabolites using LCMS-PCA, leading to the quick discovery of this new antibiotic. Given the

1 Boucher, H. W.; Talbot, G. H.; Bradley, J. S.; Edwards, J. E.; Gilbert, D.; Rice, L. B.; Scheld, M.; Spellberg, B.; Bartlett, J. Clin. Infect. Dis. 2009, 48, 1–12.
2 Hou, Y.; Braun, D. R.; Michel, C. R.; Klassen, J. L.; Adnani, N.; Wyche, T. P.; Bugni, T. S. Anal. Chem. 2012, 84, 4277–4283.
3 Krug, D.; Zurek, G.; Revermann, O.; Vos, M.; Velicer, G. J.; Müller, R. Appl. Environ. Microbiol. 2008, 74, 3058–3068.
4 Krug, D.; Zurek, G.; Schneider, B.; Garcia, R.; Müller, R. Anal. Chim. Acta 2008, 624, 97–106.
5 Carr, G.; Poulsen, M.; Klassen, J. L.; Hou, Y.; Wyche, T. P.; Bugni, T. S.; Currie, C. R.; Clardy, J. Org. Lett. 2012, 14, 2822–2825.
interest in this class of antibiotics, we subsequently sequenced the genome of WMMB272, discovering a highly novel ribosomal biosynthetic pathway, reported herein.

The structure of bottromycin A2 was recently confirmed through total synthesis and showed potential as an effective antibiotic against MRSA and VRE. More recently, it was shown that bottromycin A2 is rapidly hydrolyzed to the carboxylic acid, likely by esterases, rendering it inactive. Semisynthesis to replace the ester with a ketone led to a potent and stable analog with improved pharmacological properties. The new analog showed in vivo efficacy in a mouse model of MRSA and was orally active, albeit at a high dose. Therefore, new analogs of this class represent important potential therapeutics to help combat drug resistant infectious disease. Upon the discovery of bottromycin D (I) in the current report, we sequenced the genome of the producing organism and identified the biosynthetic pathway, creating an opportunity to improve pharmacokinetics via biosynthetic engineering. Herein, we present the discovery of a new anti-MRSA natural product as well as its biosynthesis.

Strain WMMB272 was isolated from the ascidian, *Didemnum psammathode*, collected in the Florida Keys. Analysis of the 16S rRNA gene sequence revealed that the strain was most closely related to (99% 16S identity) *Streptomyces sulphureus* NRRL B-1627. Using LCMS-PCA on 50 *Streptomyces* spp., two major classes of putative novel natural products were quickly identified (see Figure S1). The first was a rare class of pyrindipyrones discovered independently in our lab just preceding the publication of the structures by Fenical and co-workers. Two additional compounds were identified from analysis by PCA: one at t<sub>R</sub> 5.7 min and m/z 781.4072 and another at t<sub>R</sub> 6.5 min and m/z 795.4228, which we termed bottromycin D (I). A query of Antibase indicated that these were new natural products, and subsequent purification of bottromycin D (I) confirmed that it was predominantly responsible for the activity against MRSA seen from the WMMB272 extract. On the basis of the accurate mass, the compound with m/z 781.4072 contained one less CH<sub>2</sub> than bottromycin D (I), but was only a minor component. The initial isolation of I was from a solid agar medium, but our yield was only approximately 10 μg, estimated using NMR. Therefore, we optimized fermentation conditions and greatly improved the yield (to nearly a 10 mg/L isolation yield from liquid media), which facilitated structure determination.

The structure determination was achieved using standard NMR methods but was complicated due to the presence of a 1:1 ratio of conformers in CDCl<sub>3</sub>. Changing the solvent to 2:1 CD<sub>3</sub>CN/DMSO resulted in three major conformers. To rule out the possibility of a diastereomer, a combination of using low temperature NMR and changing the ratio of CD<sub>3</sub>CN/DMSO were used to investigate conformer populations. The ratio of the major conformers changed at low temperature in a manner inconsistent with the presence of a diastereomer (see SI).

In order to determine the biosynthetic origin of bottromycins as antibiotics is the lability of the side chain ester to hydrolysis in vivo, leading to inactive metabolites.

Synthetic approaches have attempted to circumvent this problem, but rational biosynthetic engineering provides a complementary approach. The biosynthesis of these compounds was unknown, but two biosynthetic routes seemed likely. Based upon structural considerations, bottromycin D (I) could be synthesized via the nonribosomal thioester mechanism, or it could be part of the growing family of ribosomally synthesized and post-translationally modified peptides (RiPPs). In RiPPs, a precursor peptide is first ribosomally synthesized and then highly tailored to yield the final natural product. In particular, the RiPP route is interesting because simple modification of the codons could potentially lead to libraries of derivatives. In order to determine the biosynthetic origin of bottromycins, we sequenced the genome of WMMB272 using Illumina HiSeq and performed bioinformatic analysis. BLAST analysis did not yield any candidate nonribosomal pathways. Instead, the precise amino acid sequence encoding bottromycin D (GPAAVFDC) was found within the context of a short precursor peptide, BstA (Figure 1).
was flanked by genes encoding putative post-translational modifying enzymes. This region was used to search GenBank, leading to the identification of a homologous cluster in the genome of *Streptomyces scabies*. Indeed, the *S. scabies* putative bottromycin cluster contained a precursor peptide that was nearly identical to bottromycin D (1), with one major exception: the precursor peptide encoded the sequence GPVVVFDC, indicating an Ala3→Val substitution in comparison to BstA. This sequence is identical to that of the previously described compound, bottromycin A2. It thus seemed likely that we had correctly identified the bottromycin gene cluster, which we named bst.

To further demonstrate that these genes were responsible for bottromycin production, we engineered derivatives of the precursor peptide. bstA was cloned into the *E. coli*–*Streptomyces* shuttle vector, pSET152, to generate pSET-bstA. This sequence was mutated to generate a series of derivatives. In particular, we constructed derivative pSET-bstA1, containing an Ala3→Val substitution. This vector mimicked the natural bottromycin A2 peptide sequence found in *S. scabies*, but was not found in the wild-type genome of strain WMMB272. Vectors were transformed into strain WMMB272 by conjugation. Fermentation of the pSET-bstA1 mutant led to production of authentic bottromycin A2, as demonstrated by high-resolution mass spectrometry (Figure S3). By contrast, bottromycin A2 was not found in fermentation of the wild-type strain, nor was it found in the pSET-bstA mutant strain. These results verify that bst encodes the pathway to bottromycin D (1). The pSET vector was inserted into a different (ectopic) site on the chromosome, and not into the bst pathway. Therefore, all strains produced

### Table 1. Genes in Putative Bottromycin D (1) Gene Cluster in WMMB272 and *S. scabies*

| Genes | % identity | Proposed function | Closest homologue (% identity/accession) |
|-------|------------|------------------|-----------------------------------------|
| WMMB272 | S. scabies | Amino acid | Nucleotide |
| bstK   | scab_566711 | 71.8 | 72.2 | MFS-family/H⁺-antiporter | MFS transporter [**Microlunatus phosphororus**] (43/YP_00457499) |
| bstB   | scab_566701 | 66.7 | 71.1 | O-methyltransferase | O-methyltransferase domain-containing protein [**Frankia symbiont of Datisca glomerata**] (50/YP_004583867) |
| bstC   | scab_56691  | 76.0 | 73.0 | Radical SAM dependent enzyme | Radical SAM domain protein [**Bacillus sp.**] 1NL3[3E] (33/ZP_06600892) |
| bstA   | scab_56681  | 81.8 | 82.6 | Precursor peptide | Hypothetical protein yber0001_29470 [**Yersinia bercovieri** ATCC 43970] (21/ZP_06627890) |
| bstD   | scab_56671  | 63.5 | 70.0 | YcaO protein | Hypothetical protein [**Streptosporangium roseum** DSM 43021] (29/ZP_003341786) |
| bstE   | scab_56661  | 73.1 | 74.4 | YcaO protein | Radical SAM domain-containing protein [**Salinispora tropica** CCB-440] (38/ZP_001159120) |
| bstF   | scab_56651  | 77.8 | 78.2 | Radical SAM dependent enzyme | Radical SAM domain-containing protein [**Rhodococcus equi**] (39/ZP_0815955) |
| bstG   | scab_56641  | 69.2 | 74.1 | α–β hydrolase | Amidohydrolase [**Alkaliphilus oremlandii** OhILAs] (26/ZP_00154298) |
| bstH   | scab_56631  | 65.2 | 69.5 | Amidohydrolase | Cytochrome P450 Thermomonospora curvata DSM 43183 (37/ZP_081530966) |
| bstI   | scab_56621  | 74.8 | 78.0 | Cytochrome P450 | Radical SAM domain protein [**Bacillus sp.**] 1NL3[3E] (35/ZP_06600892) |

Figure 1. Gene map of the bst cluster in WMMB272 and alignment of the precursor peptide sequence (BstA) with scab_56681. Mutations evaluated in this study are highlighted.

---

(11) Bignell, D. R.; Seipke, R. F.; Huguet-Tapia, J. C.; Chambers, A. H.; Parry, R. J.; Loria, R. *Mol. Plant Microbe. Interact.* 2010, 23, 161–175.

(12) Bierman, M.; Logan, R.; O’Brien, K.; Seno, E. T.; Rao, R. N.; Schoner, B. E. *Gene* 1992, 116, 43–9.
bottromycin D (1), even in the case where bottromycin A₂ was also produced.

The genome of WMMB272 as well as PCR and sequence alignments showed that bst is highly similar to and syntenic with its complement in S. scabies (Table 1). The predicted enzymes in the cluster were (~66–78%) identical to those in S. scabies, and the predicted precursor peptides were 82% identical in the amino acid sequence. Based upon the identified genes, a rational biosynthetic pathway can be proposed. However, further work is required to confirm that the identified cluster contains all of the necessary genes for biosynthesis. For example, a YcaO-like protein could form the heterocyclic thiazole residue. Several radical enzymes in the cluster were proposed. However, further work is required to confirm this. The predicted precursor peptides were 82% identical in the amino acid sequence. Based upon the identified genes, a rational biosynthetic pathway can be proposed. However, further work is required to confirm that the identified cluster contains all of the necessary genes for biosynthesis. For example, a YcaO-like protein could form the heterocyclic thiazole residue. Several radical enzymes in the cluster were proposed. However, further work is required to confirm that the identified cluster contains all of the necessary genes for biosynthesis. For example, a YcaO-like protein could form the heterocyclic thiazole residue. Several radical enzymes in the cluster were proposed. However, further work is required to confirm that the identified cluster contains all of the necessary genes for biosynthesis. For example, a YcaO-like protein could form the heterocyclic thiazole residue.

In summary, a new analog of bottromycin, herein named bottromycin D (1), was discovered, along with its biosynthetic pathway. To the best of our knowledge, this is the first analog that has an amino acid substitution in the heterocyclic ring. The Val₃→Ala substitution appears to make bottromycin D (1) more conformationally flexible, perhaps due to decreased steric interactions in the three-dimensional structure, which has recently been determined. Importantly, the activity of bottromycin D (1) was on par with other analogs, indicating that some substitutions to the heterocyclic ring may be tolerated. Overall, the rapid discovery of bottromycin D (1) demonstrates the power of LCMS-based metabolomics coupled with PCA to identify novel antibiotics and antibiotic analogs, while its production by a RiPP biosynthetic cluster indicates the importance of this biosynthetic route for natural products.

Acknowledgment. This work was supported by funding from the University of Wisconsin—Madison School of Pharmacy, the Graduate School at the University of Wisconsin, and the UW College of Agriculture and Life Sciences. This work was also funded by the NIH, NIGMS Grant R01 GM092009. We would like to thank the Analytical Instrumentation Center at the University of Wisconsin—Madison for the facilities to acquire spectroscopic data. This study made use of the National Magnetic Resonance Facility at Madison, which is supported by NIH Grants P41RR02301 (BRTP/NCRR) and P41GM66326 (NIGMS). Additional equipment was purchased with funds from the University of Wisconsin, the NIH (RR02781, RR08438), the NSF (DMB-8415048, OIA-9977486, BIR-9214394), and the USDA. We would like to thank D. Demaria for assistance with collection and R. Van Wagoner for acquiring HR-LCMS data on the transformants.

Supporting Information Available. Experimental details, details surrounding PCR primers used, nucleotide sequences, tables of NMR assignments, and NMR spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

The authors declare no competing financial interest.