

# A Systematic Approach to Discover New Natural Product Scaffolds Using Database-Derived Relative Mass Spectral Defects and Molecular Networking

Hyo Moon Cho, Eleonora Boccia, Rahim Rajwani, Robert D. O'Connor, Helena Boshoff, Clifton Barry, III, Giuseppe Bifulco, and Carole A. Bewley\*



Cite This: *JACS Au* 2025, 5, 653–665



Read Online

ACCESS |

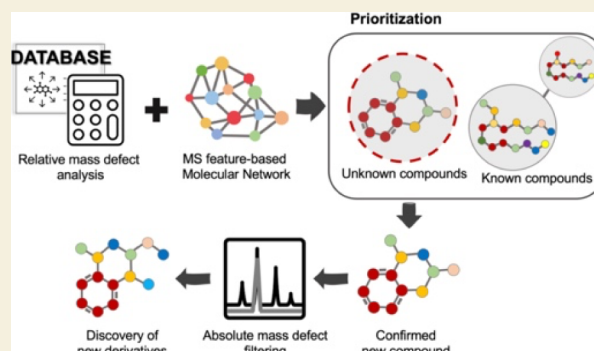
Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** Rapid advances in mass spectrometry (MS) data analysis have accelerated the identification of natural products from complex mixtures such as natural product extracts. However, limitations in MS data in metabolite libraries and dereplication strategies are still lacking for assigning structures to known compounds and searching for unidentified compounds. To overcome these limitations, we present an approach that combines molecular networking with MS database-derived mass defect analysis to preferentially discover new compounds with high structural novelty in the initial stage of a discovery workflow. Specifically, unknown metabolites or clusters generated from molecular networking are assigned to a compound class based on their relative mass defects (RMDs) calculated using open-source databases. If ancillary data such as ultraviolet and MS/MS spectra of the unknown clusters are incongruent with the RMD-assigned compound class, metabolites are considered to have a new skeleton that exhibits a large difference in RMD value due to structural changes. Here, we applied this RMD-assisted method to a desert-derived bacterial strain library and validated it through the discovery of brasiliencin A (**1**), a new 18-membered macrolide from *Nocardia brasiliensis*. A putative biosynthetic pathway of brasiliencin A was proposed through whole-genome sequence analysis, and an additional 29 analogs were detected using absolute mass defect filtering (AMDF) based on plausible biosynthetic products. This led to the isolation of three additional macrolides, brasiliencins B–D (**2–4**). The structures of the brasiliencins (**1–4**) were fully elucidated through spectroscopic data analysis and quantum chemical calculations including ROE distance and  $^{13}\text{C}$  NMR chemical shift calculations, and experimental and theoretical electronic circular dichroism (ECD). Brasiliencin A showed strong activity against *Mycobacterium smegmatis* and *Streptococcus australis* (MIC = 31.3 nM and 7.81  $\mu\text{M}$ , respectively) compared to brasiliencin B (MIC = 1000 nM and 62.5  $\mu\text{M}$ , respectively) that differs at a single stereocenter.

**KEYWORDS:** absolute mass defect filtering, brasiliencin, *Nocardia brasiliensis*, mass spectrometry, relative mass defect



## INTRODUCTION

Compounds derived from or inspired by nature comprise a majority of clinically used antibiotics, have been transformative in the development of chemotherapeutics, and continue to be a valuable source for drug and probe discovery. In recent years, mass spectrometry has become an essential bioinformatics and discovery tool that enables high accuracy and sensitive detection of metabolites through automated computer-controlled analyses of mass spectra.<sup>1,2</sup> The large amount of MS data in turn accelerated the creation of MS/MS spectral libraries and facilitated the development of molecular networking, a data analysis workflow that clusters and annotates mass spectral features on the basis of similarities between MS/MS spectra of reference compounds and ions of interest. Molecular networking is a powerful tool as it allows visualization of structural similarities between molecules from

complex natural product-derived extracts, and has greatly aided rational dereplication strategies.<sup>3</sup> The chemical diversity found in natural product-derived metabolites continues to attract researchers as estimates of the number of bacterial natural products alone is climbing to the billions.<sup>4</sup> However, to harness such molecular diversity, the current state of annotation strategies is lacking in part due to limitations in MS detection, the sizes and quality of spectral libraries, and inherent variability in liquid chromatography methods.<sup>5,6</sup> Furthermore,

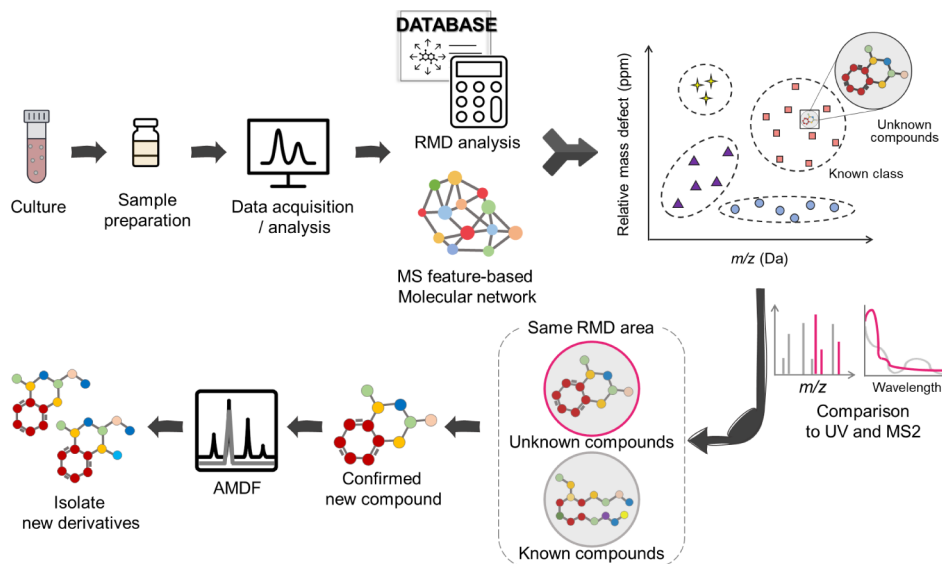
**Received:** September 24, 2024

**Revised:** December 28, 2024

**Accepted:** December 30, 2024

**Published:** January 16, 2025





**Figure 1.** Workflow for discovery of nonpredicted natural product classes using relative mass defect (RMD) filtering combined with molecular networking.

since molecular networking creates nodes and clusters on the basis of similarities between MS/MS spectra, information about the structural novelty of unannotated molecules or clusters is absent. As a result, despite its utility in dereplication workflows (identification of known compounds), most papers reporting new compound discovery use molecular networking to isolate unannotated clusters without prioritization.<sup>7–10</sup>

In this study we aimed to use the information provided from molecular networking and MS database mining, together with chemical classifications that can be estimated from relative mass defects in high-resolution MS data sets, to discover new chemical entities. The mass defect (MD) of an element or molecule corresponds to the difference between the nominal (exact) mass and the monoisotopic mass (based on the most abundant isotopes of each element) which differs for each element due to differences in nuclear binding energy.<sup>11</sup> Kendrick first used mass defects to facilitate the analysis of crude oil samples in the chemical and petroleum industries.<sup>12</sup> After that, absolute mass defect filtering (AMDF) was applied based on the advantage of effectively removing interfering ions and comprehensively capturing metabolites.<sup>13</sup> However, mass spectrometry has difficulty in unambiguously determining molecular formulas, especially for high molecular weight compounds, because of the increase in the number of possible elemental compositions within a given mass error range. To overcome this problem, an alternative approach was developed that calculates relative mass defects (RMDs). For the most abundant elements in organic compounds, carbon-12 (<sup>12</sup>C) has an exact mass of 12.0000 Da with zero mass defect,<sup>14</sup> while hydrogen, nitrogen and oxygen have absolute mass defects of +7.83, +3.07 and −5.09 ppm.<sup>11</sup> Taking advantage of the major contribution from hydrogen atoms in organic molecules, relative mass defect (RMD) analysis was developed to normalize the mass defect to the ionic mass. The RMD value in ppm is calculated by eq 1:

$$\frac{\text{mass defect}}{\text{measured monoisotopic mass}} \times 10^6 \text{ (ppm)} \quad (1)$$

Because each compound or natural product class has a characteristic hydrogen content, it is possible to use RMD

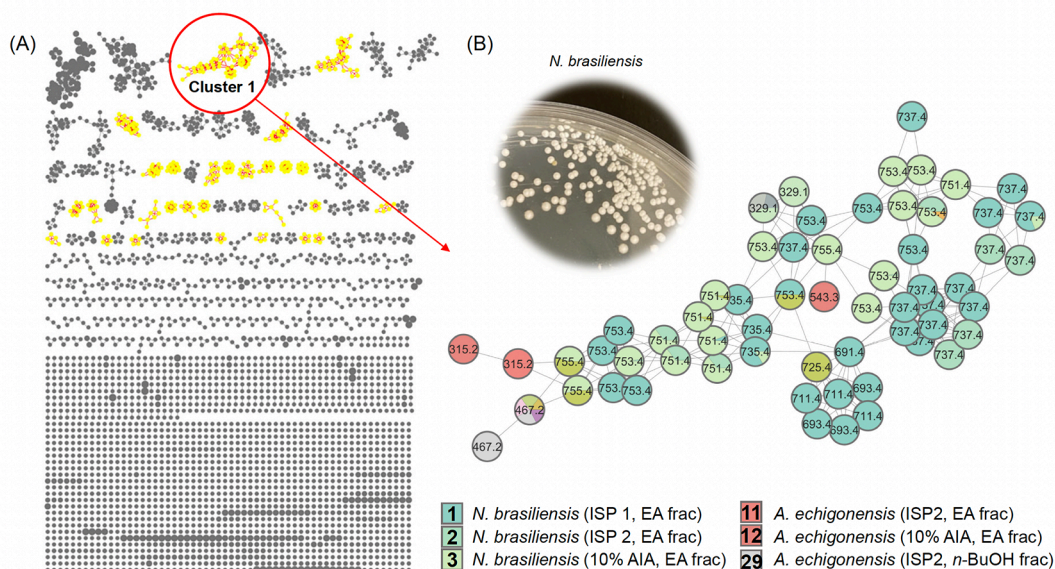
values to infer the class of an unknown compound.<sup>15</sup> RMD-filtering has been used in the analysis of herbal substances,<sup>16</sup> lipid analysis in fetal bovine serum (FBS),<sup>17</sup> and the assignment of terpene glycosides.<sup>15</sup>

To our knowledge, RMD filtering has not been used to rationally select unannotated features for new compound discovery. Here, we present a new discovery approach to preferentially discover new compounds with high structural novelty that combines molecular networking and database-derived RMD information (Figure 1). More specifically, by analyzing RMD values from natural product databases in combination with UV and MS/MS data, we can infer the compound class of unknown substances initially identified by molecular networking. Key, if there is no relationship between the known compounds of the inferred compound class and the unannotated cluster, the unannotated target cluster is assumed to have a new skeleton with large modifications, presenting a novel screening method for discovering structurally novel compounds. We validated this new method by applying it to extracts of a bacterial strain library isolated from desert soils and report the discovery of brasiliencin A. After the putative biosynthetic pathway of brasiliencins was deduced using whole-genome sequencing analysis, new derivatives were identified through AMDF based on biosynthetically plausible structures, and three additional new compounds (brasiliencins B–D) were purified, and their structures elucidated. The isolated new compounds exhibit diverse antibacterial activities and demonstrated strong and selective inhibitory activity against *Mycobacterium smegmatis*.

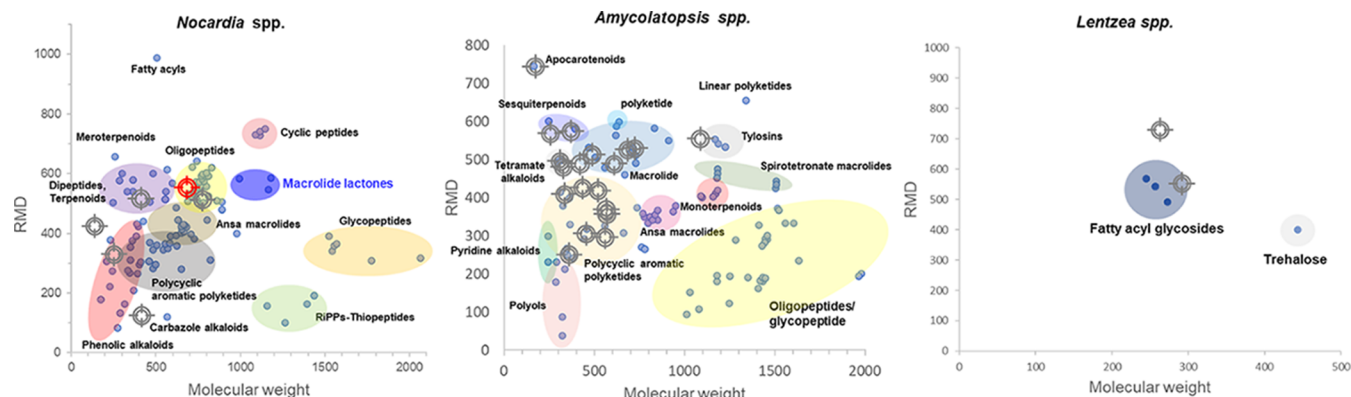
## RESULTS AND DISCUSSION

### Identification of New Compound Candidates through Relative Mass Defects and Molecular Networking

To test the approach of identifying new compound candidates using the combination of mass defect filtering and molecular networking, we analyzed the organic extracts of six actinobacteria (*Nocardia brasiliensis*, *N. cyriacigeorgica*, *Amycolatopsis japonica*, *A. echigonensis*, *A. lurida*, and *Lentzea violacea*). These strains were prioritized from an in-house actinobacterial library



**Figure 2.** Molecular network and node prioritization using RMD selection. (A) Molecular network generated from MS data of organic extracts of six bacterial strains from an in-house actinobacterial library. The 28 clusters shown in yellow represent metabolites that (i) were unannotated in the GNPS libraries, (ii) are only produced by single species, and (iii) contain five or more nodes. (B) Expansion of Cluster 1 corresponding to an unknown compound produced by *Nocardia brasiliensis* and having an average molecular weight of 700 Da and an average RMD of 557 ppm.



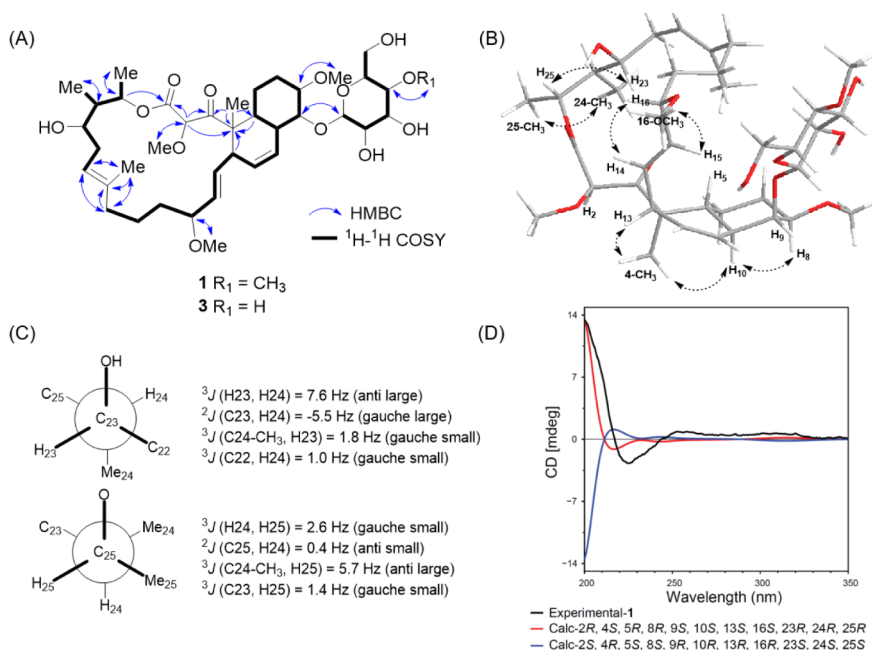
**Figure 3.** Plots of RMDs versus deprotonated molecular ion masses ( $m/z$  value) for compounds from *Nocardia* spp., *Amycolatopsis* spp., and *Lentzea* spp. Individual compound classes are indicated with ovals and were identified by NPClassifier.<sup>18</sup> Red and gray target symbols correspond respectively to candidate cluster 1 and all other candidate clusters selected from molecular networking.

based on their resistance to vancomycin and antibiotic activity of respective colonies in agar overlay assays. After culturing in three fermentation media (ISP1, ISP2 broth, 10% Actinomycete Isolation Agar) for 1 week, ethyl acetate and *n*-BuOH fractions were resuspended in MeOH and analyzed by ultrahigh-performance liquid chromatography-high resolution mass spectrometry (UHPLC-HRMS). Following data processing with MZmine 2, a molecular network was generated using the GNPS Web platform. A total of 3446 nodes and 456 clusters were generated, and 21 compound classes were annotated in 33 different clusters (Figure 2A). From the generated molecular networking, the clusters meeting three conditions were selected as prime candidates for discovering structurally new compounds. This included (i) clusters that were not annotated, (ii) clusters with five or more nodes that could suggest the presence of structural analogs, and (iii) clusters unique to a specific genus that could increase the chances of compound novelty. To determine the compound class of the unannotated clusters, we calculated the RMD for

each known compound using NPClassifier data, which includes both the compound class and the taxonomy of the producing organism, based on information from the Natural Products Atlas database.<sup>18</sup> The expected RMDs for known compounds were calculated using predicted  $m/z$  values obtained from the isotope distribution calculator (Agilent MassHunter Qualitative software version B.07.00) based on their given molecular formulas (see Experimental Methods for details). Finally, those curated RMD values were plotted as a function of molecular weight to give reference plots for each genus studied and various classes identified by NPClassifier were highlighted as shown in Figure 3. The average RMD values of the selected unannotated clusters were then marked as target indicators on the reference plot and prioritized for further study.

Among these, cluster 1, which was presumed to have a compound with a new structure and a high probability of containing many analogs, was selected for experimental verification (Figure 2B). Cluster 1 was suggested to contain





**Figure 4.** (A) Key HMBC (blue arrows) and  $^1\text{H}$ - $^1\text{H}$  COSY (bold lines) correlations for **1** and **3**. (B) Key ROESY correlations shown on an optimized structure of compound **1**. (C)  $J$ -based configurational analysis along the C23–C24 and C24–C25 axes of compound **1**. (D) Comparison of experimental ECD spectrum of **1** with calculated ECD spectra of the possible isomers. The simulated ECD spectra of **1** was shifted 20 nm and  $\sigma = 0.3 \text{ eV}$ .

oligopeptide-like compounds with an average molecular weight of 700 and an average RMD value of 557 ppm when compared with metabolites isolated from *Nocardia* species using the RMD vs molecular weight graph (Figure 3). However, neither peptide fragments nor amino acid sequences were observed in the MS/MS spectra, and the UV spectrum of the target peak lacked absorbance at 200–230 nm and 250–350 nm, wavelengths indicative of peptide amide bonds and aromatic amino acids, respectively (Figure S1). Therefore, Cluster 1 was unlikely to be an oligopeptide series, even though its RMD value corresponded to oligopeptides. Accordingly, we hypothesized that Cluster 1 was a different class and that a large change occurred in the skeleton, which may have significantly altered the RMD values of the compounds.

### Complete Structure Elucidation of Brasiliencin A

Brasiliencin A (**1**) was obtained as a white, acicular crystal with  $[\alpha] + 37.2$  ( $c$  0.1, MeOH). The molecular formula  $\text{C}_{39}\text{H}_{62}\text{O}_{13}$  was established based on the deprotonated molecular ion at  $m/z$  737.4124 (calcd for  $\text{C}_{39}\text{H}_{61}\text{O}_{13}$ , 737.4118,  $\Delta = +0.88 \text{ ppm}$ ) in the negative HRESIMS data. The  $^1\text{H}$  and  $^{13}\text{C}$  NMR data of **1** showed a total of 39 carbons, ascribable to four methyl groups, four methoxy groups, two carbonyl carbons including one ketone and one ester group, six olefinic carbons, seven methylene groups, 15 methines including nine that are oxygenated, and one quaternary carbon. The planar structure of **1** was determined to be a new 18-membered macrolide from  $^1\text{H}$ - $^1\text{H}$  COSY, multiplicity edited HSQC (mpHSQC), and  $^1\text{H}$ - $^{13}\text{C}$  HMBC spectra. The COSY and HMBC spectra revealed two spin systems from C-5 to C-19 and C-21 to C-25 on the macrolide ring (Figure 4A). Additional HMBC correlations from H-19 ( $\delta_{\text{H}}$  2.05 and 2.09) to C-20 ( $\delta_{\text{C}}$  136.1) and C-21 ( $\delta_{\text{C}}$  123.4), and from Me-20 ( $\delta_{\text{H}}$  1.65) to C-19 ( $\delta_{\text{C}}$  38.4), C-20, and C-21 connected the two fragments through the  $\text{sp}^2$  quaternary carbon C-20.

The sugar unit was determined to be 4'-O-methyl- $\beta$ -glucopyranose from COSY and HMBC correlations, and the large vicinal coupling constant ( $J = 8.2 \text{ Hz}$ ) of the anomeric proton at  $\delta_{\text{H}}$  4.42 (Table 1). To determine its configuration, the pyranose was isolated from the acid hydrolyzate (1 M HCl, 3 h) of a brasiliencins-rich fraction and optical rotations of the purified material established a D configuration ( $[\alpha]_{\text{D}}^{25} + 62.2$  ( $c$  0.1,  $\text{H}_2\text{O}$ )). An HMBC correlation from anomeric H-1' to C-9 ( $\delta_{\text{C}}$  75.3) connected the sugar to the 15-carbon spin system. HMBC correlations from Me-4 ( $\delta_{\text{H}}$  1.32) to the ketone carbon (C-3) at  $\delta_{\text{C}}$  206.2, a methyl-substituted quaternary carbon (C-4) at  $\delta_{\text{C}}$  53.0, and two methine carbons (C-5 and C-13) at  $\delta_{\text{C}}$  31.6 and 46.1, respectively, corresponded to an additional six-membered ring and the presence of a decalin ring system. In addition, the methoxy-substituted methine proton (H-2) at  $\delta_{\text{H}}$  4.59 showed an HMBC correlation to the ester ( $\delta_{\text{C}}$  168.5) and ketone groups connecting the C-3 ketone carbon to the decalin ring system at C-4. The HMBC correlation from H-25 of the terminal methine ( $\delta_{\text{H}}$  5.59) to the ester carbon (C-1) established the presence of a lactone and closed the 18-membered macrolide ring. HMBC correlations positioned the four methyl groups at  $\delta_{\text{H}}$  1.32, 1.65, 0.98, and 1.32 at carbons C-4, C-20, C-24 ( $\delta_{\text{C}}$  43.8), and C-25 ( $\delta_{\text{C}}$  72.5), and the three methoxy groups at  $\delta_{\text{H}}$  3.47, 3.39, and 3.16 at C-2 ( $\delta_{\text{C}}$  85.4), C-8 ( $\delta_{\text{C}}$  82.3), and C-16 ( $\delta_{\text{C}}$  80.2), respectively.

The relative stereochemistry of the decalin ring system was established from key ROESY correlations between Me-4/H-10, Me-4/H-13, and H-8/H-10; and the large  $^3J_{\text{H-H}}$  11 Hz coupling for H-5/H-10 and small coupling for H-9 ( $\delta_{\text{H}}$  4.23) and H-10; which indicate that Me-4, H-8, H-9, H-10, and H-13 are in the same plane and H-5 is in the opposite plane (Figures 4B and S46). The large  $^3J_{\text{HH}}$  of 15 Hz indicated an *E* geometry for C-14/C-15 ( $\delta_{\text{C}}$  130.0/138.1) and was supported by ROESY correlations between H-14 ( $\delta_{\text{H}}$  5.18)/H-16 ( $\delta_{\text{H}}$

Table 1. <sup>1</sup>H and <sup>13</sup>C NMR Data (600 MHz, CDCl<sub>3</sub>) for Brasiliencins A–D

no.	Brasiliencin A			Brasiliencin B			Brasiliencin C			Brasiliencin D		
	δ <sub>C</sub> , type	δ <sub>H</sub> , multiplicity (J in Hz)	δ <sub>C</sub> , type	δ <sub>H</sub> , multiplicity (J in Hz)	δ <sub>C</sub> , type	δ <sub>H</sub> , multiplicity (J in Hz)	δ <sub>C</sub> , type	δ <sub>H</sub> , multiplicity (J in Hz)	δ <sub>C</sub> , type	δ <sub>H</sub> , multiplicity (J in Hz)	δ <sub>C</sub> , type	δ <sub>H</sub> , multiplicity (J in Hz)
1	168.5, C		168.2, C		168.5, C		168.5, C		167.5, C		167.5, C	
2	85.4, CH	4.59 s	82.8, CH	4.45 s	85.4, CH	4.59 s	85.4, CH	4.59 s	82.3, CH	4.46 s	82.3, CH	4.46 s
3	206.2, C		205.0, C		206.2, C		206.2, C		203.0, C		203.0, C	
4	53.0, C		53.5, C		53.0, C		53.0, C		52.9, C		52.9, C	
5	31.6, CH	2.32 td (11.6, 2.7)	31.9, CH	2.34 <sup>a</sup>	31.6, CH	2.34 <sup>a</sup>	31.6, CH	2.31 <sup>a</sup>	31.7, CH	2.36 td (11.4, 2.9)	31.7, CH	2.36 td (11.4, 2.9)
6	24.6, CH <sub>2</sub>	0.83 m	24.7, CH <sub>2</sub>	0.90 <sup>a</sup>	24.6, CH <sub>2</sub>	0.90 <sup>a</sup>	24.6, CH <sub>2</sub>	0.81 m	24.7, CH <sub>2</sub>	0.85 m <sup>a</sup>	24.7, CH <sub>2</sub>	0.85 m <sup>a</sup>
		1.39 m <sup>a</sup>		1.47 m <sup>a</sup>		1.47 m <sup>a</sup>		1.39 m <sup>a</sup>		1.63 m <sup>a</sup>		1.63 m <sup>a</sup>
7	25.6, CH <sub>2</sub>	1.68 m <sup>a</sup>	25.5, CH <sub>2</sub>	1.68 m <sup>a</sup>	25.6, CH <sub>2</sub>	1.68 m <sup>a</sup>	25.6, CH <sub>2</sub>	1.68 m <sup>a</sup>	25.9, CH <sub>2</sub>	1.70 m <sup>a</sup>	25.9, CH <sub>2</sub>	1.70 m <sup>a</sup>
		1.86 m		1.88 m		1.88 m		1.86 m <sup>a</sup>		1.87 m		1.87 m
8	82.3, CH	3.14 m <sup>a</sup>	82.3, CH	3.17 m <sup>a</sup>	82.3, CH	3.17 m <sup>a</sup>	82.3, CH	3.15 m <sup>a</sup>	82.2, CH	3.16 m <sup>a</sup>	82.2, CH	3.16 m <sup>a</sup>
9	75.3, CH	4.23 br s	74.7, CH	4.25 br s	74.8, CH	4.25 br s	74.8, CH	4.24 s	75.8, CH	4.24 br s	75.8, CH	4.24 br s
10	42.7, CH	1.78 dt (11.2, 2.2)	43.1, CH	1.79 d (11.3)	42.7, CH	1.79 d (11.3)	42.7, CH	1.78 br d (11.4)	43.0, CH	1.82 d (11.1)	43.0, CH	1.82 d (11.1)
11	129.4, CH	5.59 m <sup>a</sup>	130.0, CH	5.62 m <sup>a</sup>	129.5, CH	5.62 m <sup>a</sup>	129.5, CH	5.59 (1H, m) <sup>a</sup>	129.2, CH	5.58 m <sup>a</sup>	129.2, CH	5.58 m <sup>a</sup>
12	127.0, CH	5.52 m <sup>a</sup>	126.4, CH	5.53 m <sup>a</sup>	126.8, CH	5.53 m <sup>a</sup>	126.8, CH	5.50 (1H, m) <sup>a</sup>	127.5, CH	5.57 m <sup>a</sup>	127.5, CH	5.57 m <sup>a</sup>
13	46.1, CH	3.02 t (5.3)	45.5, CH	2.80 t (5.8)	46.0, CH	2.80 t (5.8)	46.0, CH	3.03 t (5.7)	46.1, CH	2.84 t (5.9)	46.1, CH	2.84 t (5.9)
14	130.0, CH	5.18 dd (15.8, 6.2)	129.3, CH	5.17 dd (15.7, 6.1)	130.1, CH	5.17 dd (15.7, 6.1)	130.1, CH	5.18 dd (15.8, 6.2)	129.1, CH	5.41 dd (15.9, 6.9)	129.1, CH	5.41 dd (15.9, 6.9)
15	138.1, CH	5.52 m <sup>a</sup>	138.7, CH	5.57 dd (14.5, 6.7) <sup>a</sup>	138.1, CH	5.57 dd (14.5, 6.7) <sup>a</sup>	138.1, CH	5.53 (1H, m) <sup>a</sup>	136.2, CH	5.55 dd (15.5, 5.6)	136.2, CH	5.55 dd (15.5, 5.6)
16	80.2, CH	3.58 m	80.6, CH	3.52 m <sup>a</sup>	80.3, CH	3.52 m <sup>a</sup>	80.3, CH	3.58 (1H, m)	79.3, CH	3.68 m <sup>a</sup>	79.3, CH	3.68 m <sup>a</sup>
17	33.3, CH <sub>2</sub>	1.44 m <sup>a</sup>	34.5, CH <sub>2</sub>	1.47 m <sup>a</sup>	33.3, CH <sub>2</sub>	1.47 m <sup>a</sup>	33.3, CH <sub>2</sub>	1.44 (2H, m) <sup>a</sup>	33.0, CH <sub>2</sub>	1.30 m <sup>a</sup>	33.0, CH <sub>2</sub>	1.30 m <sup>a</sup>
										1.43 m		1.43 m
18	20.8, CH <sub>2</sub>	1.53 m	20.7, CH <sub>2</sub>	1.52 m <sup>a</sup>	20.8, CH <sub>2</sub>	1.52 m <sup>a</sup>	20.8, CH <sub>2</sub>	1.53 (1H, m)	22.1, CH <sub>2</sub>	1.54 m <sup>a</sup>	22.1, CH <sub>2</sub>	1.54 m <sup>a</sup>
		1.62 m <sup>a</sup>		1.64 m <sup>a</sup>		1.64 m <sup>a</sup>		1.63 m <sup>a</sup>		1.67 m <sup>a</sup>		1.67 m <sup>a</sup>
19	38.4, CH <sub>2</sub>	2.05 m <sup>a</sup>	39.4, CH <sub>2</sub>	2.02 m	38.4, CH <sub>2</sub>	2.02 m	38.4, CH <sub>2</sub>	2.05 m <sup>a</sup>	38.5, CH <sub>2</sub>	2.06 m	38.5, CH <sub>2</sub>	2.06 m
		2.09 m		2.11 m		2.11 m		2.09 m		2.13 m		2.13 m
20	136.1, C		136.0, C		136.1, C		136.1, C		139.5, C		139.5, C	
21	123.4, CH	5.37 t (7.7)	123.3, CH	5.45 t (7.8)	123.4, CH	5.45 t (7.8)	123.4, CH	5.37 t (7.7)	117.3, CH	5.35 t (7.6)	117.3, CH	5.35 t (7.6)
22	34.3, CH <sub>2</sub>	2.14 m <sup>a</sup>	35.0, CH <sub>2</sub>	2.14 m <sup>a</sup>	34.3, CH <sub>2</sub>	2.14 m <sup>a</sup>	34.3, CH <sub>2</sub>	2.15 m <sup>a</sup>	42.5, CH <sub>2</sub>	2.92 dd (17.6, 5.9)	42.5, CH <sub>2</sub>	2.92 dd (17.6, 5.9)
		2.26 ddd (14.2, 7.6, 2.4)		2.30 <sup>a</sup>		2.30 <sup>a</sup>		2.26 m <sup>a</sup>		3.40 m <sup>a</sup>		3.40 m <sup>a</sup>
23	74.1, CH	3.25 td (9.0, 2.4)	73.7, CH	3.21 <sup>a</sup>	74.1, CH	3.21 <sup>a</sup>	74.1, CH	3.25 <sup>a</sup>	211.9, C		211.9, C	
24	43.8, CH	1.62 m <sup>a</sup>	44.3, CH	1.61 m <sup>a</sup>	43.8, CH	1.61 m <sup>a</sup>	43.8, CH	1.62 m <sup>a</sup>	49.1, CH	3.01 p (7.2)	49.1, CH	3.01 p (7.2)
25	72.5, CH	5.59 m <sup>a</sup>	72.5, CH	5.64 <sup>a</sup>	72.5, CH	5.64 <sup>a</sup>	72.5, CH	5.59 m <sup>a</sup>	74.0, CH	5.27 m	74.0, CH	5.27 m
1'	101.8, CH	4.42 d (8.2)	101.6, CH	4.43 d (8.2)	101.8, CH	4.43 d (8.2)	101.8, CH	4.46 d (8.2)	102.1, CH	4.40 d (8.0)	102.1, CH	4.40 d (8.0)
2'	71.8, CH	3.44 t (9.1) <sup>a</sup>	71.5, CH	3.43 t (9.5) <sup>a</sup>	71.4, CH	3.44 <sup>a</sup>	71.4, CH	3.44 <sup>a</sup>	72.2, CH	3.43 m	72.2, CH	3.43 m
3'	76.6, CH	3.64 d (8.9) <sup>a</sup>	76.6, CH	3.64 t (8.9) <sup>a</sup>	71.4, CH	3.44 <sup>a</sup>	71.4, CH	3.44 <sup>a</sup>	76.6, CH	3.63 t (8.9) <sup>a</sup>	76.6, CH	3.63 t (8.9) <sup>a</sup>
4'	80.3, CH	3.08 t (9.1)	80.3, CH	3.08 t (9.1)	76.7, CH	3.08 t (9.1)	76.7, CH	3.53 <sup>a</sup>	80.3, CH	3.09 t (9.2)	80.3, CH	3.09 t (9.2)
5'	75.9, CH	3.41 m <sup>a</sup>	76.0, CH	3.41 m <sup>a</sup>	76.1, CH	3.41 m <sup>a</sup>	76.1, CH	3.46 <sup>a</sup>	75.9, CH	3.40 m <sup>a</sup>	75.9, CH	3.40 m <sup>a</sup>
6'	63.1, CH <sub>2</sub>	3.68 dd (11.4, 6.8)	63.1, CH <sub>2</sub>	3.68 dd (10.8, 6.2)	63.3, CH <sub>2</sub>	3.68 dd (10.8, 6.2)	63.3, CH <sub>2</sub>	3.70 <sup>b</sup>	63.1, CH <sub>2</sub>	3.68 <sup>a</sup>	63.1, CH <sub>2</sub>	3.68 <sup>a</sup>
		3.87 dd (11.4, 3.1)		3.88 dd (11.4, 3.0)		3.88 dd (11.4, 3.0)		3.90 <sup>b</sup>		3.87 dd (11.4, 3.1)		3.87 dd (11.4, 3.1)
4-CH <sub>3</sub>	16.1, CH <sub>3</sub>	1.32 s	16.1, CH <sub>3</sub>	1.35 s	16.1, CH <sub>3</sub>	1.35 s	16.1, CH <sub>3</sub>	1.33	15.8, CH <sub>3</sub>	1.30, s	15.8, CH <sub>3</sub>	1.30, s
20-CH <sub>3</sub>	15.6, CH <sub>3</sub>	1.65 s	15.6, CH <sub>3</sub>	1.64 s	15.6, CH <sub>3</sub>	1.64 s	15.6, CH <sub>3</sub>	1.65	15.8, CH <sub>3</sub>	1.65 s	15.8, CH <sub>3</sub>	1.65 s
24-CH <sub>3</sub>	10.6, CH <sub>3</sub>	0.98 d (6.8)	10.8, CH <sub>3</sub>	0.90 d (6.9) <sup>a</sup>	10.6, CH <sub>3</sub>	0.90 d (6.9) <sup>a</sup>	10.6, CH <sub>3</sub>	0.99 d (6.9)	13.8, CH <sub>3</sub>	1.16 d (7.0)	13.8, CH <sub>3</sub>	1.16 d (7.0)
25-CH <sub>3</sub>	17.4, CH <sub>3</sub>	1.31 d (6.6)	17.8, CH <sub>3</sub>	1.28 d (6.6)	17.4, CH <sub>3</sub>	1.28 d (6.6)	17.4, CH <sub>3</sub>	1.32 d (6.6)	18.4, CH <sub>3</sub>	1.25 d (6.5)	18.4, CH <sub>3</sub>	1.25 d (6.5)

Table 1. continued

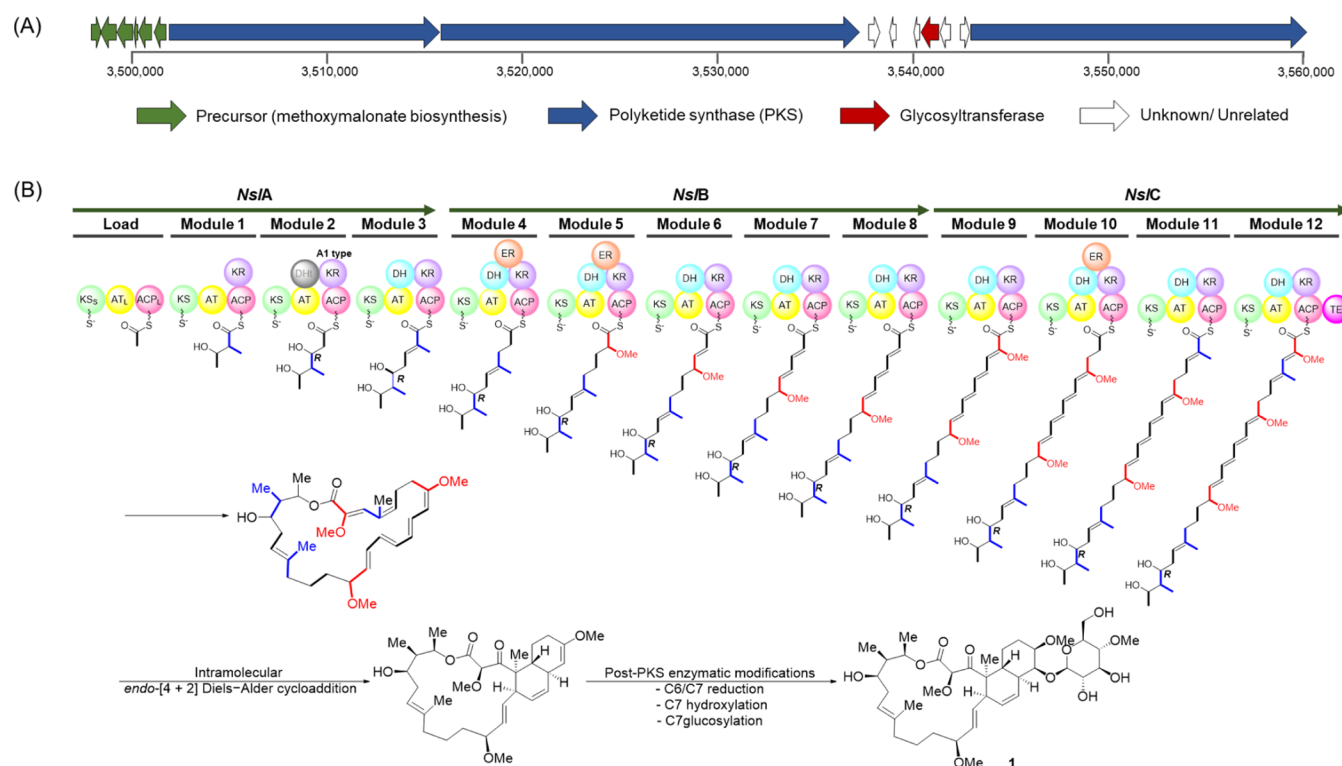
no.	Brasiliencin A		Brasiliencin B		Brasiliencin C		Brasiliencin D	
	$\delta_{\text{C}}$ type	$\delta_{\text{H}}$ multiplicity ( <i>J</i> in Hz)	$\delta_{\text{C}}$ type	$\delta_{\text{H}}$ multiplicity ( <i>J</i> in Hz)	$\delta_{\text{C}}$ type	$\delta_{\text{H}}$ multiplicity ( <i>J</i> in Hz)	$\delta_{\text{C}}$ type	$\delta_{\text{H}}$ multiplicity ( <i>J</i> in Hz)
2-OCH <sub>3</sub>	59.4, CH <sub>3</sub>	3.47 s	59.1, CH <sub>3</sub>	3.48 s	59.4, CH <sub>3</sub>	3.47 s	58.9, CH <sub>3</sub>	3.47 s
8-OCH <sub>3</sub>	56.8, CH <sub>3</sub>	3.39 s	56.8, CH <sub>3</sub>	3.39 s	56.8, CH <sub>3</sub>	3.39 s	56.7, CH <sub>3</sub>	3.39 s
16-OCH <sub>3</sub>	55.9, CH <sub>3</sub>	3.16 s	56.0, CH <sub>3</sub>	3.14 s	55.9, CH <sub>3</sub>	3.17 s	55.7, CH <sub>3</sub>	3.22 s
4'-OCH <sub>3</sub>	60.5, CH <sub>3</sub>	3.55 s	60.5, CH <sub>3</sub>	3.56 s			60.6, CH <sub>3</sub>	3.56 s

<sup>a</sup>Overlapped. <sup>b</sup>Impurities overlapped.

3.58), and H-15 ( $\delta_{\text{H}}$  5.52)/OMe-16 ( $\delta_{\text{H}}$  3.16). The  $\Delta^{20,21}$  olefinic bond was assigned as *E* based on ROESY correlations between H-19 ( $\delta_{\text{H}}$  2.05 and 2.09)/H-21 ( $\delta_{\text{H}}$  5.37) and Me-20 ( $\delta_{\text{H}}$  1.65)/H-22 ( $\delta_{\text{H}}$  2.14 and 2.26).

The relative configuration at C-23, C-24 and C-25 was assigned as 23*R*\*, 24*R*\*, 25*R*\* through the combination of *J*-based configurational analyses including proton–proton ( $^3J_{\text{H,H}}$ ) and carbon–proton coupling constants ( $^2J_{\text{C,H}}$  and  $^3J_{\text{C,H}}$ ) and ROESY correlations (Figures 4B,C and S46 and Table S3). For example, ROESY correlations between Me-24 ( $\delta_{\text{H}}$  0.98) and Me-25 ( $\delta_{\text{H}}$  1.31), and H-23 ( $\delta_{\text{H}}$  3.25) and H-25 ( $\delta_{\text{H}}$  5.59) place the two methyl groups and H-23/C-25 and H-25/C-23 in *gauche* orientations. We attempted to determine the absolute configuration of C-23 using the modified Mosher's method, but the compound degraded under multiple reaction conditions. Described below, the absolute configuration of C-23 could nevertheless be assigned as *R* based on the A1-type ketoreductase (KR) identified through amino acid sequence analysis of module 2 (Figures 5 and S49).<sup>19</sup> We propose the 23*R*, 24*R*, and 25*R* absolute stereochemistry considering the comprehensive analysis of ROESY data, calculated ROE intensities, and the biosynthesis.

The relative stereochemistry at C-2 was considered to be 2*R*\* from the key ROESY correlations (Figure S46), and was supported by the similarity between the measured and simulated ROE intensities from the internuclear distance using the program MSpin (Figure S12).<sup>20</sup> For example, an ROE between H-2 ( $\delta_{\text{H}}$  4.59) and H-13 ( $\delta_{\text{H}}$  3.02) was not observed suggesting their opposing orientations. To confirm the experimentally determined stereochemistry in **1**, we used a quantum mechanics (QM)/NMR approach to calculate two- and three-bond  $^1\text{H}$ – $^1\text{H}$  and  $^1\text{H}$ – $^{13}\text{C}$  coupling constants and  $^{13}\text{C}$  chemical shifts for **1**. To reduce the number of stereoisomers for consideration, we first assigned the relative configuration of the C-23/C-25 fragment by calculating experimental *J* values for the six different rotamers including three each for the *erythro* and *threo* configurations. The similarity between calculated and experimental *J* values,  $\Delta\bar{J}_{\text{calc}}$  (the sum of the differences divided by the number of considered data) indicates the relative configurations 23*R*\*, 24*R*\*, and 25*R*\* and establishes C-24/C-25 as *erythro* and H-23/C-25 as *threo* (Tables S6 and S7). With C-23, C-24 and C-25 stereochemistry fixed and the relative configuration of the decalin ring system determined, we calculated *J* couplings for the eight possible nonglycosylated macrolides. The stereoisomers for the ring system include 4*S*\*, 5*R*\*, 8*R*\*, 9*S*\*, 10*S*\*, 13*S*\* and 4*R*\*, 5*S*\*, 8*S*\*, 9*R*\*, 10*R*\*, 13*R*\* configurations, together with the two distinct configurations at C-2 and C-16. Each stereoisomer was subjected to Low Mode and Molecular Dynamics conformational searches and optimization followed by calculation of the  $^{13}\text{C}$  chemical shifts at the density functional theory (DFT) level. The mean absolute errors (MAEs) from the comparison between the calculated and experimental  $^{13}\text{C}$  chemical shifts (Table S8) gives the best match points to the relative configuration 2*R*, 4*S*, 5*R*, 8*R*, 9*S*, 10*S*, 11*Z*, 13*S*, 14*E*, 16*S*, 20*E*, 23*R*, 24*R*, 25*R* ( $^{13}\text{C}$  MAE = 1.82 ppm), also confirmed by the analysis of the ROE distances. Next, to compare theoretical and experimental ECD spectra of the enantiomers, we performed QM calculations at the TD-DFT MPW1PW91/6-31G(d,p) level. The experimental data showed highly similar Cotton effects in both sign and magnitude as the calculated ECD spectrum for the 2*R*, 4*S*, 5*R*, 8*R*, 9*S*, 10*S*, 11*Z*, 13*S*, 14*E*, 16*S*, 20*E*, 23*R*, 24*R*, 25*R*



**Figure 5.** Biosynthetic gene cluster (A) and the proposed biosynthetic pathway (B) of brasiliencin A.

enantiomer (Figure 4D), consistent with the absolute configuration determined from experimental data and biosynthetic considerations.

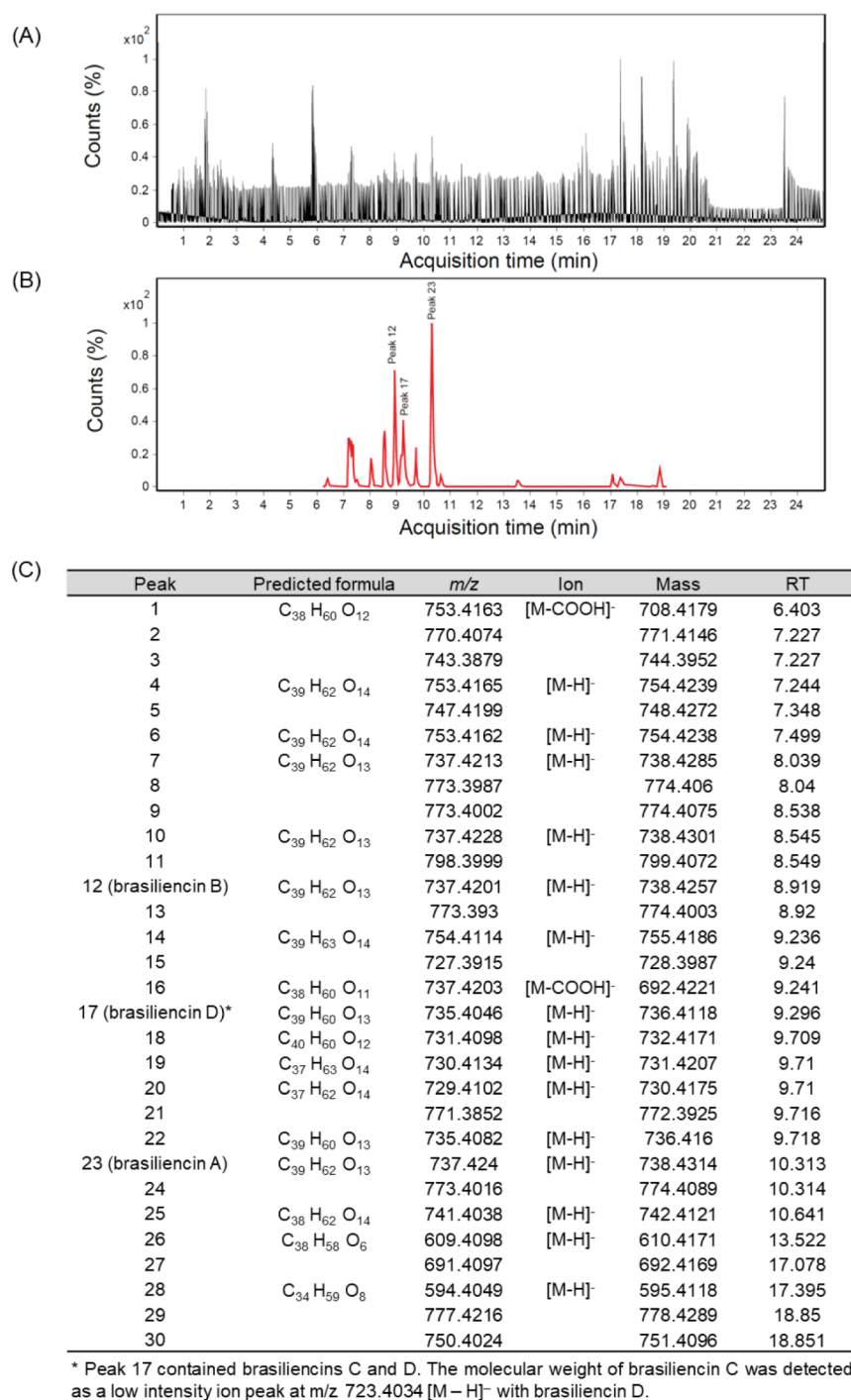
#### Analysis of the Proposed Brasiliencin Gene Cluster *nsI* in *Nocardia brasiliensis*

A long-read whole genome sequence of *Nocardia brasiliensis* was analyzed for the presence of Type I polyketide synthases using AntiSmash 7.<sup>21</sup> Only one region in the genome contained a polyketide synthase (PKS) having a sufficient number of modules for the biosynthesis of the macrolides and this region is proposed as the biosynthetic locus. The brasiliencin PKS, termed *nsI*, is a Type-1 modular PKS containing a loading module that includes a ketosynthase domain ( $KS^S$ ), an acyltransferase domain ( $AT^0$ ), and an acyl carrier protein domain ( $ACP^0$ ); and 12 extension modules (Figure 5). In the first module, the active site Cys 170 of the  $KS^S$  domain is substituted with a serine residue indicating the  $KS^S$  domain (Figure S47). A multiple sequence alignment of the  $AT^0$  domain suggests that malonyl-CoA may be the preferred substrate (Figure S48), but since the decarboxylase activity of  $KS^S$  has not been discovered in a previous study, unlike  $KS^Q$ , we speculate that the loose specificity of the loading AT may allow it to use acetyl-CoA as a starter unit, similar to nystatin biosynthesis.<sup>22,23</sup> Sequence alignments and bioinformatics of the AT domains revealed specificities for methoxymalonyl-CoA at modules 5, 9, and 12; methylmalonyl-CoA at modules 1, 3, and 11; and malonyl CoA at modules 2, 4, 6–8, and 10 (Figure S48), perfectly aligned with the experimentally determined structures. Sequence alignments for the ketoreductase (KR) domains identified an A1-type KR in module 2 based on the presence of a Trp residue N-terminal to the catalytic Tyr, B1-types in modules 4–12 based respectively on the presence of the “LDD” motif, and a C2-type KR in module 3 based on a catalytic Ser-to-Ala mutation and the

presence of a Gly residue four amino acids N-terminal to the catalytic Tyr.<sup>19</sup> The A1-type KR in module 2 is consistent with the R configuration at C-2 determined for compound 1 (Figure S49). The terminal thioesterase domain in module 12 is responsible for macrolide ring formation. Analysis of the BGC did not identify pericyclase or Diels–Alderase-type enzymes, suggesting the decalin ring system may be formed through spontaneous intramolecular [4 + 2] cycloaddition.<sup>24</sup> For example, the diene could approach the endo reface of the trans-dienophile to generate an [18,6,6]-tricyclic ring in which the trans-decalin is linked to the 18-membered ring with a *cis*-geometry (Figure 5B).

#### Discovery of New Derivatives Using Absolute Mass Defect Filtering

Absolute mass defect filtering (AMDF) was applied to the ethyl acetate extract of *Nocardia brasiliensis* to selectively discover new derivatives of brasiliencin A by removing unnecessary matrix interference in the complex total ion chromatogram (TIC).<sup>25,26</sup> For filtering, the precursor was set to a hub structure based on the proposed biosynthetic pathway, and the central molecular formula was set to  $C_{36}H_{56}O_{10}$  as the average value of the maximum ( $C_{40}H_{64}O_{14}$ ) and minimum ( $C_{32}H_{48}O_6$ ) elements for the range of possible analogs. This central molecular formula was predicted based on using the precursor structure as a motif and considering the possibilities of candidate structures containing zero or one sugar unit, methylation of hydroxyl groups, and intramolecular cyclizations that may occur through mechanisms other than Diels–Alder reactions or ester bond formation. Additionally, the mass defect tolerance was set to half the difference in mass defects between the largest and smallest molecules, and this value was applied as 0.0462 Da. To predict the molecular formula of the detected ion, the elemental composition of carbon was set from 32–40, the



**Figure 6.** Absolute mass defect (AMD) filtering identifies brasiliencin analogs. (A) Total ion chromatogram (TIC) of the ethyl acetate extract from *N. brasiliensis* and (B) the AMD filtered spectrum. (C) List of absolute mass defect filtered metabolites.

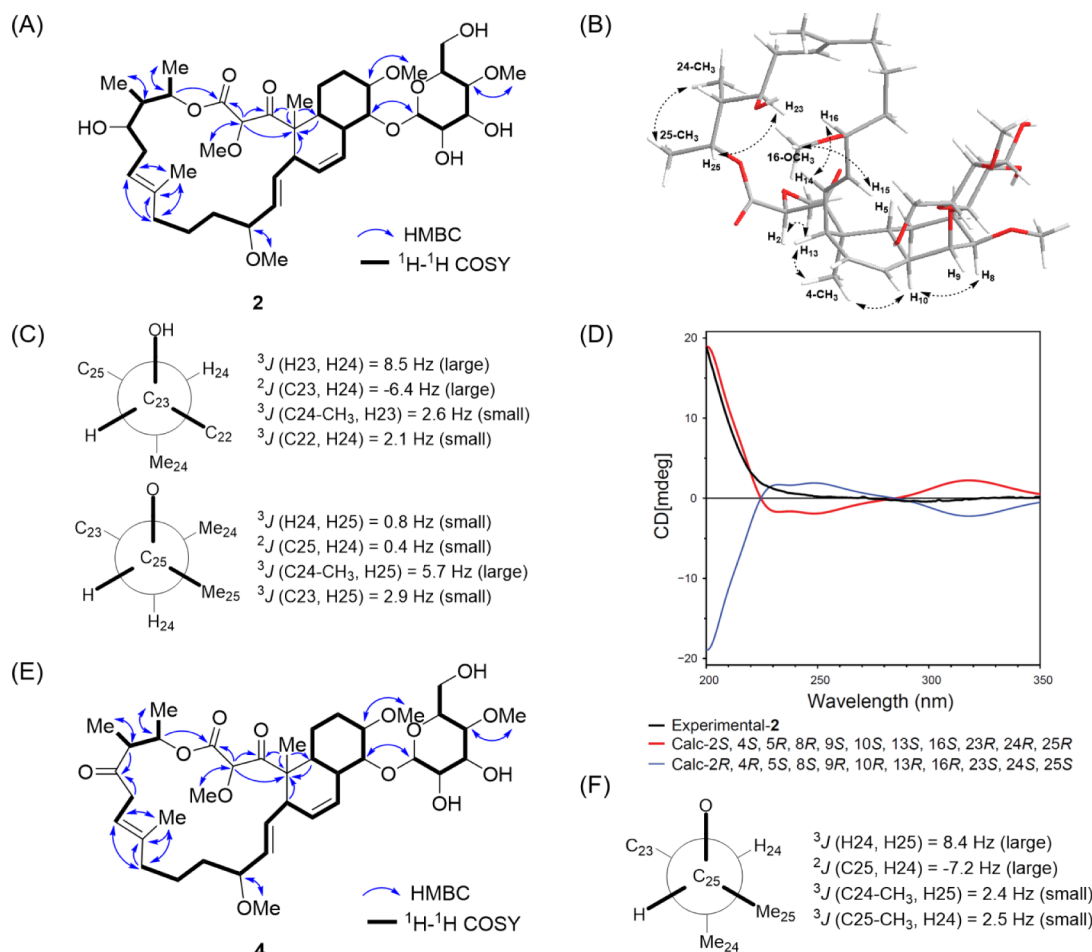
number of hydrogens from 48–64, and the number of oxygens from 6–14. Nitrogen and sulfur were not included among the elements. This filtering revealed twenty-nine additional and potentially new brasiliencins and the already isolated brasiliencin A that were detected but originally buried in the total ion chromatogram (Figure 6A,B). Among the newly identified compounds, the structural formula of 17 compounds could be predicted based on the detected ions (Figure 6C). The remaining 13 substances may be due to filtering errors, or correspond to metabolites containing S, N, Cl or Br that were excluded from the parameters for predicting structural

formulas but have a similar mass defect range. To verify the predicted results, peaks 12 and 17 were selectively purified, considering the yield of compounds. As a result, brasiliencin B was obtained from peak 12, and since peak 17 was a mixture of two compounds, it was purified again through a Waters X-Bridge BEH C18 reversed-phase column to isolate brasiliencins C and D.

#### Structure Elucidation of Brasiliencin Analogs 2–4

Brasiliencin B (2) was obtained as a white, amorphous solid with  $[\alpha]_D^{25} + 13.6$  ( $c$  0.03, MeOH). Its molecular formula  $C_{39}H_{62}O_{13}$  was identical to compound 1 based on the





**Figure 7.** Key HMBC (blue arrows),  $^1\text{H}$ – $^1\text{H}$  COSY (bold lines), and ROESY correlations (black dot arrows) shown on an optimized structure of **2** (A,B). (C)  $J$ -based configurational analysis along the C23–C24 and C24–C25 axes of **2**. (D) Comparison of experimental ECD spectrum of **2** with calculated ECD spectra of the possible isomers. The simulated ECD spectra of **2** was shifted 15 nm and  $\sigma = 0.32$  eV. (E) Key HMBC (blue arrows) and  $^1\text{H}$ – $^1\text{H}$  COSY (bold lines) correlations for **4**. (F)  $J$ -based configurational analysis along the C24–C25 axes of **4**.

deprotonated molecular ion at  $m/z$  737.4129 (calcd for  $\text{C}_{39}\text{H}_{61}\text{O}_{13}$ , 737.4118,  $\Delta = +1.58$  ppm) in the negative HRESIMS spectrum. Comparison of the  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra of compounds **1** and **2** were similar except for the chemical shifts of H-2 ( $\delta_{\text{H}}$  4.45), H-13 ( $\delta_{\text{H}}$  2.80), and C-2 ( $\delta_{\text{C}}$  82.8). Analysis of the same set of 2D NMR spectra described for compound **1** confirmed that **1** and **2** have the same planar structure (Figure 7A–C). However, a new correlation between H-2 and H-13 was observed in the ROESY spectrum suggesting a 2*S* configuration at C-2, and that **1** and **2** are diastereomers (Figures 7B and S46). In addition, the calculated quantitative ROEs between H-2 and H-13 were similar in magnitude to the experimental ROE of **2** with a 2*S* orientation and showed a clear difference from **1** that has the 2*R* configuration (Figures S18 and S23). Starting from the results obtained for compound **1** and considering their diastereoisomeric and epimeric nature, the TDDFT calculation of the ECD spectrum was carried out. Finally, compound **2** was determined to be a diastereomer of **1** with a 2*S*, 4*S*, 5*R*, 8*R*, 9*S*, 10*S*, 11*Z*, 13*S*, 14*E*, 16*S*, 20*E*, 23*R*, 24*R*, 25*R* configuration through comparison of calculated and experimental ECD spectra (Figure 7D).

Brasiliencin C (**3**) also was obtained as a white, amorphous solid with  $[\alpha]_{\text{D}}^{25} + 25.9$  ( $c$  0.03, MeOH). Its molecular formula  $\text{C}_{38}\text{H}_{60}\text{O}_{13}$  was assigned from the deprotonated molecular ion

at  $m/z$  723.3985 (calcd for  $\text{C}_{38}\text{H}_{59}\text{O}_{13}$ , 723.3961,  $\Delta = -3.24$  ppm) in the negative HRESIMS spectrum. Overall, the  $^1\text{H}$  and  $^{13}\text{C}$  spectra of compound **3** were nearly identical to the spectra of **1**, except for the absence of the methoxy group at C-3' on the sugar moiety. This led to an upfield shift in the  $^{13}\text{C}$  signals for C-3' and C-4', from 76.6 to 71.4 ppm and 80.3 ppm to 76.7 ppm, respectively (Table 1). Analysis of HSQC, HMBC, and COSY experiments confirmed the same planar structure (Figure 4A), and the presence of the same diagnostic ROESY correlations observed for **3** indicated that compounds **1** and **3** have the same configurations (Figures S29, S32 and S46). Finally, the experimental ECD spectrum of **3** showed high agreement with **1**, providing further support that the absolute configuration of compound **3** is 2*R*, 4*S*, 5*R*, 8*R*, 9*S*, 10*S*, 11*Z*, 13*S*, 14*E*, 16*S*, 20*E*, 22*R*, 23*R*, 24*R*, 25*R* (Figure S44), which is identical to compound **1**.

The final compound identified through AMDF and that was isolable from the *Nocardia* cultures was brasiliencin D (**4**), obtained as a white, amorphous solid with  $[\alpha]_{\text{D}}^{25} + 11.6$  ( $c$  0.03, MeOH). Its molecular formula was assigned as  $\text{C}_{39}\text{H}_{60}\text{O}_{13}$  from its deprotonated molecular ion at  $m/z$  735.3990 (calcd for  $\text{C}_{39}\text{H}_{59}\text{O}_{13}$ , 735.3961,  $\Delta = -3.91$  ppm) in the HRESIMS data. Comparison of the  $^{13}\text{C}$  chemical shifts of compound **4** with the shifts for **1**–**3** indicated that compound **4** shared the same *S* configuration at C-2 as brasiliencin B (Table 1), and

**Table 2. MICs of Brasiliencins A (1) and B (2) against *Mycobacterium smegmatis* MC<sup>2</sup> 155**

MIC $\mu\text{g/mL}$ ( $\mu\text{M}$ )					MIC $\mu\text{g/mL}$ (nM)	
vancomycin	chloramphenicol	ampicillin	rifampicin	ciprofloxacin	1	2
2.5 (1.7)	10 (30.9)	20 (57.2)	40 (48.6)	0.16 (12.1)	0.02 (31.3)	0.74 (1000)

**Table 3. MICs of Brasiliencins A (1) and B (2) against Pathogenic Human Strain**

	MIC $\mu\text{g/mL}$ ( $\mu\text{M}$ )		MIC $\mu\text{g/mL}$ ( $\mu\text{M}$ )	
	chloramphenicol	gentamicin	1	2
<i>Enterococcus faecium</i> (ATCC29212)	6.25 (19.3)		125	250
<i>Pseudomonas aeruginosa</i> (ATCC27853)		0.78 (1.6)	<250	<250
<i>Prevotella amnii</i> (DSM 23384)	0.78 (2.4)		<250	<250
<i>Neisseria lactamica</i> (ATCC23970)	0.78 (2.4)		125	125
<i>Streptococcus australis</i> (DSM 15627)	1.56 (4.8)		5.77 (7.81)	46.2 (62.5)
<i>Streptococcus mutans</i> (ATCC700641)	12.5 (38.7)		250	<250
<i>Shuttleworthia satelles</i> (DSM 14600)	1.56 (4.8)		62.5	125
<i>Dialister invisus</i> (DSM 15470)	1.56 (4.8)		62.5	125
<i>Olsenella uli</i> (DSM 7084)	3.91 (12.1)		62.5	125
<i>Prevotella buccae</i> (DSM 19025)	6.25 (19.3)		125	250

the hydroxyl group at C-23 had been replaced with a ketone group ( $\delta_{\text{C}}$  211.9). The chemical shifts at the adjacent C-22 and C-24 atoms were shifted downfield by 7.5 and 4.8 ppm compared to brasiliencin B, with C-22/H-22a,b resonating at  $\delta_{\text{C}}$  42.5 and  $\delta_{\text{H}}$  2.92 and 3.40; and methine C-24/H-24 at  $\delta_{\text{C}}$  49.1 and  $\delta_{\text{H}}$  3.01. Also, the HMBC correlations from H-22 and H-24 to the C-23 keto group confirmed this change (Figure 7E). A detailed analysis of the *J* couplings and ROESY correlations for **4** supported the same relative configurations throughout the macrolide and decalin ring system. *J*-based configurational analysis applied to the optimized structure suggested a gauche orientation for C-24/C-25 supported by the ROESY correlation between Me-24 ( $\delta_{\text{H}}$  1.16) and Me-25 ( $\delta_{\text{H}}$  1.25) shown in Figures 7F and S46 and Table S5. Together with their optical rotations and ECD spectra that are similar to compound **2**, the absolute configuration for compound **4** was assigned as 2*S*, 4*S*, 5*R*, 8*R*, 9*S*, 10*S*, 11*Z*, 13*S*, 14*E*, 16*S*, 20*E*, 22*R*, 24*R*, 25*R* (Figure S45).

### Antimicrobial Activity Is Structure Dependent

The brasiliencins were tested for antimicrobial activity against standard bacterial pathogens, *Mycobacterium smegmatis* and *M. tuberculosis*, along with a panel of oral bacteria (Tables 2 and 3). These included ten Gram-positive and Gram-negative pathogens, including seven different species of bacteria (*P. aeruginosa*, *E. faecium*, *P. amnii*, *N. lactamica*, *S. australis*, *S. mutans*, *S. satelles*, *D. invisus*, *O. uli*, and *P. buccae*) isolated from the human oral microbiome. The major compound brasiliencin A was extremely potent in inhibiting the growth of *M. smegmatis* MC<sup>2</sup>155 (ATCC 700084) with an MIC of 31.3 nM and inhibited the growth of the oral microbiome-derived *Streptococcus australis* at low micromolar concentrations (MIC of 7.8  $\mu\text{M}$ ). Interestingly, brasiliencin B (**2**) which differs from **1** in the stereochemistry at C-2 (2*S* vs 2*R* in compound **1**), shows a marked reduction (30- to 250-fold) in antibacterial activity, with MICs of **1** and 62.5  $\mu\text{M}$  toward *M. smegmatis* and *S. australis*, respectively. The brasiliencins were inactive (MIC > 100  $\mu\text{M}$ ) against the other bacteria.

## DISCUSSION AND CONCLUSIONS

Mass spectrometry has become an essential tool in natural products research. Besides being the primary method for

determining the molecular formula of novel and known compounds, curation of large databases of MS and MS/MS data and the development of molecular networking programs such as GNPS allow for untargeted and targeted MS analyses of entire metabolomes, a process that is invaluable to researchers studying natural products and specialized metabolomes. GNPS is now commonly used to detect the full suite of metabolites in a specific node or cluster of interest. In this work we developed a complementary yet distinct approach by calculating relative mass defects (RMDs) of a given MS data set, and used those RMD data to filter networks that for *unknown* compounds were either outliers in space in an RMD versus molecular weight plot, or whose accompanying data, such as UV/vis spectra or MS/MS profiles, did not fit in the overlapping bin for a compound class. Compound classes in the RMD plots were assigned using the structural classification tool NPClassifier which, unlike other chemical classification tools, can be correlated with taxonomy. When applying RMD filtering in a search for unique structure types, a node of interest had to meet several criteria. Those included the mismatch with the assigned chemical class (described directly above) and its occurrence in only a single species or closely related genus that we reasoned would eliminate primary or commonly produced metabolites. We demonstrated this workflow using six actinobacterial species from three genera; in addition to initial antimicrobial activity exhibited by the six strains, we reasoned that this workgroup would contain hundreds if not thousands of clusters, would contain a sufficiently diverse set of natural product classes for testing the method, and would generate unidentified metabolites with atypical RMD values. Seen in Figure 3, the approach identified twenty-eight nonclassifiable clusters; the one node that was investigated here resulted in the discovery of a new 18-membered macrolide with novel features relative to the referenced macrolides.<sup>27</sup>

Another key advantage of using mass defects for discovery is its application in absolute mass defect filtering (AMDF). Once a compound has been selected, whether for structure elucidation, biological activity, or other interests, AMDF can be used to identify related compounds present in the cluster of interest. There are two advantages to using AMDF: first, this filtering can be applied directly to the TIC spectrum, allowing

the location of peaks in the spectrum to be directly detected making it more efficient when isolation of additional compounds is the goal. Second, it can reduce the number of missed peaks in the MS data. This is because the data-dependent acquisition mode obtains MS/MS data by selecting only the several highest peaks in the MS1 spectrum (usually three). Thus, when the intensity of a peak is low, it is often missed in molecular networking because the MS2 result is not obtained. In comparison, AMDF has the advantage of being able to filter out even minor peaks because it uses only the MS1 results. Overall, RMD filtering for new compound discovery offers a promising new approach in natural products research. The macrolide ring system in the brasiliencins shares some structural features with tubelactomicin A, a 16-membered macrolide isolated from a *Nocardia* sp. that showed strong inhibition of “acid-fast” bacteria, or mycobacteria, and selectivity toward this group.<sup>28</sup> We observed similar results with antimicrobial screening with brasiliencin A, and to a lesser extent brasiliencin B, strongly inhibiting the growth of *M. smegmatis* and an oral *Streptococcus* strain. Studies are underway to determine the mechanism of inhibition of brasiliencin A on mycobacteria and its effect on fast versus slow growing strains.

## EXPERIMENTAL SECTION

General experimental procedures and antimicrobial screening methods are included in the [Supporting Information](#). Optical rotations and CD spectra were recorded on a JASCO *p*-2000 polarimeter (JASCO, Easton, MD, USA) and a Jasco J-815 CD spectrometer (Applied Photophysics, Leatherhead, Japan), respectively. IR spectra were obtained on an FT-IR Bruker Alpha II spectrometer (Bruker ALPHA-II FTIR, Germany). NMR spectra were recorded on Bruker Neo 500 or 600 MHz spectrometers equipped with triple resonance, *z*-gradient cryoprobes. Acquisition and processing of NMR spectra were performed with MestReNOVA 14.3.2 (Metrelab, Santiago de Compostela, Spain) and Topspin 4.1.4 (Bruker Biospin, Ettlingen, Germany).

### Genomic DNA Extraction and Analysis

*N. brasiliensis* was cultured in an LB medium for 1 week. After harvesting the cells, the genomic DNA was purified using Wizard genomic DNA purification kit (cat. A1120) following the manufacturer's protocol (Promega, Madison, WI). The quality of isolated genomic DNA was confirmed by 0.8% agarose gel electrophoresis (Invitrogen, Carlsbad, CA, USA) and the ratio of absorbance at 260 nm to that of 280 nm in the Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific). Finally, the 300  $\mu$ L of DNA quantification was determined to be 174 ng/ $\mu$ L using the PicoGreen assay on an Eppendorf biospectrometer fluorescence (Eppendorf, Hamburg, Germany). The bacterial whole genome sequencing performed on the Pacific Biosciences (PacBio) sequencing by NIH Intramural Sequencing Center (NISC).

### Preparation of Sample and UHPLC-qTOF-MS/MS Data Acquisition for Analysis

Six actinobacterial species isolated from desert soil were cultured in three different fermentation media (5 mL ISP1, ISP2 or 10% Actinomycete isolation agar) in 14 mL culture tubes for 7 days at 28 °C with shaking at 180 rpm. Media from each culture was partitioned three times using ethyl acetate (EtOAc) and *n*-butanol and the organic extracts were dried *in vacuo*. Samples were resuspended in HPLC grade MeOH (1 mg mL<sup>-1</sup>, filtered and analyzed by UHPLC-HRMS. Experimental details describing the chromatographic and MS and MS/MS details are provided in the [Supporting Information](#).

### Feature-Based Molecular MS/MS Network

After converting the raw data obtained from UHPLC into mzXML files through ProteoWizard, the converted files were preprocessed

using MZmine 2.<sup>29</sup> The processed data were generated a molecular network according to the online workflow of the GNPS platform using the following parameters: precursor and fragment ion mass tolerances 0.02 Da, minimum pairs cosine score 0.7, minimum matched fragment ion 6, library search minimum matched peak 4, score threshold 0.7. The obtained molecular networking was visualized by the Cytoscape 3.10.1. The generated molecular networking is available at the GNPS web platform (<http://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=a0d6883442bd4cd1b35a1f305217e86a>).

### Relative Mass Defects (RMDs) and RMD Plots as a Function of Molecular Ion Masses

The Relative Mass Defects (RMDs) used in the plots in [Figure 3](#) and throughout this paper were obtained as follows. Mass defect (MD) is defined as the difference between the monoisotopic mass and the nominal mass, and the RMD, which normalizes the MD to the ion mass, is calculated as shown above. The calculated values are converted to parts per million (ppm) to emphasize small variations in the values and facilitate comparison. We used the NPClassifier database to (i) search for all metabolites from the actinomycete genera *Nocardia*, *Amycolatopsis*, and *Lentzea*, and (ii) assign their compound classes, and those metabolites were used as the reference compounds. Based on the structural formula for each compound curated from NPAtlas and NPClassifier, the molecular ion theoretical mass was generated using the Isotope Distribution Calculator within MassHunter Qualitative software v B.07.00 (Agilent Technologies), and the RMDs of the reference compounds were calculated. For each genus, RMDs were then plotted as a function of molecular weight ([Figure 3](#)), and the areas corresponding to compound classes were highlighted in the reference plots. Finally, we used GNPS to analyze HRMS data of organic extracts from the three genera listed above and generate molecular networks. After calculating the RMDs of the compounds corresponding to selected clusters, the average RMD values of the clusters were overlaid on the reference RMD plot for comparison. Compounds whose RMD values did not correspond to the underlying compound class were prioritized for follow up and characterization.

### Bacterial Fermentation

*N. brasiliensis* was streaked to an ISP1 agar plate for 1 week, and one single colony was scraped and inoculated into 50 mL ISP1 media and cultured at 28 °C for 1 week on a shaking incubator at 180 rpm. Each 50 mL seed culture was transferred to eight 2.5 L flasks, each containing 1 L ISP1 medium and incubate for 1 week at 28 °C with 220 rpm rotation.

### Extraction and Isolation

After fermentation for 1 week, the culture was extracted three times with EtOAc. The EtOAc layer was concentrated to dryness *in vacuo*, and the residue was subjected to Sephadex LH-20 column chromatography using MeOH/CH<sub>2</sub>Cl<sub>2</sub> (1:1). Ten fractions were collected (EA1–10) and analyzed by LCMS for the presence of ions belonging to Cluster 1. Fraction EA2 that contained relevant ions was subjected to further chromatographic separations using preparative HPLC on an Agilent 1100 Series HPLC (Agilent Technologies Inc., CA, USA) and a YMC-Triart phenyl column (10  $\times$  250 mm, 5  $\mu$ m, YMC, Kyoto, Japan), eluting with 40% MeCN–H<sub>2</sub>O (containing 0.1% formic acid) at a flow rate of 3 mL/min to give two subfractions (EA2.1–2.2); compounds **1** (*t*<sub>R</sub> = 39 min, 2.5 mg) and **4** (*t*<sub>R</sub> = 33 min, 0.8 mg) were collected by UV detection at 205 or 254 nm. Fraction EA2.1 was purified by preparative HPLC using a Waters X-Bridge BEH C18 reversed-phase column (5  $\mu$ m; 10 mm  $\times$  150 mm) and eluted with 40% MeCN–H<sub>2</sub>O (containing 0.1% formic acid) to obtain compounds **2** (*t*<sub>R</sub> = 24 min, 1.8 mg) and **3** (*t*<sub>R</sub> = 35 min, 1 mg).

**Brasiliencin A (1).** White, acicular crystals; [ $\alpha$ ]<sub>D</sub><sup>20</sup> 37.2 (*c* 0.1, MeOH); UV (MeOH)  $\lambda_{\text{max}}$  (log  $\epsilon$ ) 200 (3.95) nm; IR  $\nu_{\text{max}}$  3412, 2931, 1735, 1700, 1596, 1381, 1355, 1292, 1110 cm<sup>-1</sup>; <sup>1</sup>H and <sup>13</sup>C NMR data, see [Table 1](#); HRESIMS *m/z* 737.4124 [*M* – H]<sup>–</sup> (calcd for C<sub>39</sub>H<sub>61</sub>O<sub>13</sub>, 737.4118).



**Brasiliencin B (2).** White, amorphous solid;  $[\alpha]_D^{20}$  13.6 ( $c$  0.03, MeOH); UV (MeOH)  $\lambda_{\max}$  (log  $\epsilon$ ) 200 (3.92) nm; IR  $\nu_{\max}$  3429, 2928, 1745, 1707, 1592, 1455, 1376, 1282, 1110  $\text{cm}^{-1}$ ;  $^1\text{H}$  and  $^{13}\text{C}$  NMR data, see Table 1; HRESIMS  $m/z$  737.4129  $[\text{M} - \text{H}]^-$  (calcd for  $\text{C}_{39}\text{H}_{61}\text{O}_{13}$ , 737.4118).

**Brasiliencin C (3).** White, amorphous solid;  $[\alpha]_D^{20}$  25.9 ( $c$  0.03, MeOH); UV (MeOH)  $\lambda_{\max}$  (log  $\epsilon$ ) 200 (2.92) nm; IR  $\nu_{\max}$  3398, 2923, 2858, 1599, 1458, 1368, 1162, 1082, 1032  $\text{cm}^{-1}$ ;  $^1\text{H}$  and  $^{13}\text{C}$  NMR data, see Table 1; HRESIMS  $m/z$  723.3985  $[\text{M} - \text{H}]^-$  (calcd for  $\text{C}_{38}\text{H}_{59}\text{O}_{13}$ , 723.3961).

**Brasiliencin D (4).** White, amorphous solid;  $[\alpha]_D^{20}$  11.6 ( $c$  0.03, MeOH); UV (MeOH)  $\lambda_{\max}$  (log  $\epsilon$ ) 218 (3.06), 275 (2.19) nm; IR  $\nu_{\max}$  3402, 2923, 2863, 1742, 1711, 1598, 1457, 1374, 1193, 1107, 1031  $\text{cm}^{-1}$ ;  $^1\text{H}$  and  $^{13}\text{C}$  NMR data, see Table 1; HRESIMS  $m/z$  735.3990  $[\text{M} - \text{H}]^-$  (calcd for  $\text{C}_{39}\text{H}_{59}\text{O}_{13}$ , 735.3961).

### Acid Hydrolysis and Sugar Analysis of 1–4

Fraction EA2 (40 mg), a mixture of compounds 1–4, was heated with 1 M HCl (200  $\mu\text{L}$ ) at 95  $^\circ\text{C}$  for 3 h, then left to cool to room temperature. Then, the reaction mixture was partitioned with EtOAc. The aqueous layer was neutralized with saturated  $\text{NaHCO}_3$  and concentrated to obtain an aqueous extract (8 mg). The resulting dried aqueous extract was purified by silica gel chromatography eluting with  $\text{CHCl}_3$ –MeOH (9:1) to obtain pure hydrolyzed sugar. The hydrolyzed sample was compared with an authentic standard of D-glucose by Si gel TLC developed in  $\text{CH}_2\text{Cl}_2$ –MeOH (8:2) to give the reported Rf of 0.2. The optical rotation of the hydrolyzed sugar sample was  $[\alpha]_D^{25} + 62.2$  ( $c$  0.1,  $\text{H}_2\text{O}$ ), consistent with the literature value of  $[\alpha]_D^{25} + 53.2$  ( $c$  0.1,  $\text{H}_2\text{O}$ ),<sup>30</sup> and confirmed the presence of D-glucose in compounds 1–4.

## ■ ASSOCIATED CONTENT

### Data Availability Statement

The genome sequence of *Nocardia brasiliensis* strain GA14-02 is available in Genbank, accession number ASM4269167v1.

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacsau.4c00889>.

NMR, HRMS, IR, and ECD spectra of brasiliencins A–D, detailed experimental procedures and antibiotic susceptibility testing, and tables of candidate strains and multiple sequence alignments used for assigning domain function (PDF)

Table of genes in the proposed BGC (XLSX)

## ■ AUTHOR INFORMATION

### Corresponding Author

**Carole A. Bewley** – Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States; [orcid.org/0000-0003-2203-0852](https://orcid.org/0000-0003-2203-0852); Email: [caroleb@nih.gov](mailto:caroleb@nih.gov)

### Authors

**Hyo Moon Cho** – Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States

**Eleonora Boccia** – Dipartimento di Farmacia, Università di Salerno, Fisciano (SA) 84084, Italy

**Rahim Rajwani** – Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney

Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States

**Robert D. O'Connor** – Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States

**Helena Boshoff** – Tuberculosis Research Section, Laboratory of Clinical Immunology and Microbiology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States; [orcid.org/0000-0002-4333-206X](https://orcid.org/0000-0002-4333-206X)

**Clifton Barry, III** – Tuberculosis Research Section, Laboratory of Clinical Immunology and Microbiology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States

**Giuseppe Bifulco** – Dipartimento di Farmacia, Università di Salerno, Fisciano (SA) 84084, Italy; [orcid.org/0000-0002-1788-5170](https://orcid.org/0000-0002-1788-5170)

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/jacsau.4c00889>

### Author Contributions

CRediT: **Hyo Moon Cho** conceptualization, data curation, formal analysis, investigation, methodology, validation, visualization, writing - original draft; **Eleonora Boccia** data curation, formal analysis; **Rahim Rajwani** formal analysis, writing - review & editing; **Robert D. O'Connor** data curation, formal analysis, methodology, writing - review & editing; **Helena I. M. Boshoff** investigation; **Clifton E. Barry III** funding acquisition; **Giuseppe Bifulco** data curation, formal analysis, software, supervision, validation, writing - review & editing; **Carole A. Bewley** conceptualization, formal analysis, funding acquisition, supervision, writing - original draft, writing - review & editing.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program of the National Institutes of Health (NIDDK).

## ■ REFERENCES

- (1) Münzenberg, G. Development of mass spectrometers from Thomson and Aston to present. *Int. J. Mass Spectrom.* **2013**, 349, 9–18.
- (2) Weber, T.; Kim, H. U. The secondary metabolite bioinformatics portal: Computational tools to facilitate synthetic biology of secondary metabolite production. *Synth. Syst. Biotechnol.* **2016**, 1 (2), 69–79.
- (3) Aron, A. T.; Gentry, E. C.; Mcphail, K. L.; Nothias, L. F.; Nothias-Espósito, M.; Bouslimani, A.; Petras, D.; Gauglitz, J. M.; Sikora, N.; Vargas, F.; et al. Reproducible molecular networking of untargated mass spectrometry data using GNPS. *Nat. Protoc.* **2020**, 15 (6), 1954–1991.
- (4) Kalkreuter, E.; Pan, G.; Cepeda, A. J.; Shen, B. Targeting Bacterial Genomes for Natural Product Discovery. *Trends Pharmacol. Sci.* **2020**, 41 (1), 13–26.
- (5) Yu, J. S.; Nothias, L.-F.; Wang, M.; Kim, D. H.; Dorrestein, P. C.; Kang, K. B.; Yoo, H. H. Tandem mass spectrometry molecular networking as a powerful and efficient tool for drug metabolism studies. *Anal. Chem.* **2022**, 94 (2), 1456–1464.
- (6) Clark, T. N.; Houriet, J.; Vidar, W. S.; Kellogg, J. J.; Todd, D. A.; Cech, N. B.; Lington, R. G. Interlaboratory Comparison of Untargeted Mass Spectrometry Data Uncovers Underlying Causes for Variability. *J. Nat. Prod.* **2021**, 84 (3), 824–835.



- (7) Nothias, L.-F.; Boutet-Mercey, S.; Cachet, X.; Torre, E. D. L.; Laboureur, L.; Gallard, J.-F.; Retailleau, P.; Brunelle, A.; Dorrestein, P. C.; Costa, J.; et al. Environmentally Friendly Procedure Based on Supercritical Fluid Chromatography and Tandem Mass Spectrometry Molecular Networking for the Discovery of Potent Antiviral Compounds from *Euphorbia semiperfoliata*. *J. Nat. Prod.* **2017**, *80* (10), 2620–2629.
- (8) Esposito, M. L.; Nothias, L.-F. L.; Retailleau, P.; Costa, J.; Roussi, F.; Neyts, J.; Leyssen, P.; Touboul, D.; Litaudon, M.; Paolini, J. Isolation of premyrsinane, myrsinane, and tiglane diterpenoids from *Euphorbia pithyusa* using a Chikungunya virus cell-based assay and analogue annotation by molecular networking. *J. Nat. Prod.* **2017**, *80* (7), 2051–2059.
- (9) Alcover, C. F.; Bernadat, G.; Kabran, F. A.; Le Pogam, P.; Leblanc, K.; Fox Ramos, A. E.; Gallard, J.-F. O.; Mouray, E.; Grellier, P.; Poupon, E.; et al. Molecular networking reveals serpentinine-related bisindole alkaloids from *Picralima nitida*, a previously well-investigated species. *J. Nat. Prod.* **2020**, *83* (4), 1207–1216.
- (10) Qu, B.; Liu, Y.; Shen, A.; Guo, Z.; Yu, L.; Liu, D.; Huang, F.; Peng, T.; Liang, X. Combining multidimensional chromatography-mass spectrometry and feature-based molecular networking methods for the systematic characterization of compounds in the supercritical fluid extract of *Tripterygium wilfordii* Hook F. *Analyst* **2022**, *148* (1), 61–73.
- (11) Sleno, L. The use of mass defect in modern mass spectrometry. *J. Mass Spectrom.* **2012**, *47* (2), 226–236.
- (12) Kendrick, E. A Mass Scale Based on  $\text{CH}_2 = 14.0000$  for High Resolution Mass Spectrometry of Organic Compounds. *Anal. Chem.* **1963**, *35* (13), 2146–2154.
- (13) Zhang, H.; Zhang, D.; Ray, K. A software filter to remove interference ions from drug metabolites in accurate mass liquid chromatography/mass spectrometric analyses. *J. Mass Spectrom.* **2003**, *38* (10), 1110–1112.
- (14) Cameron, A.; Wichers, E. Report of the international commission on atomic weights\*(1961). *J. Am. Chem. Soc.* **1962**, *84* (22), 4175–4197.
- (15) Ekanayaka, E. A. P.; Celiz, M. D.; Jones, A. D. Relative Mass Defect Filtering of Mass Spectra: A Path to Discovery of Plant Specialized Metabolites. *Plant Physiol.* **2015**, *167* (4), 1221–U1146.
- (16) Waldner, B. J.; Machalet, R.; Schönbichler, S.; Dittmer, M.; Rubner, M. M.; Intelmann, D. Fast Evaluation of Herbal Substance Class Composition by Relative Mass Defect Plots. *Anal. Chem.* **2020**, *92* (19), 12909–12916.
- (17) Stagliano, M. C.; Dekeyser, J. G.; Omiecinski, C. J.; Jones, A. D. Bioassay-directed fractionation for discovery of bioactive neutral lipids guided by relative mass defect filtering and multiplexed collision-induced dissociation. *Rapid Commun. Mass Spectrom.* **2010**, *24* (24), 3578–3584.
- (18) Kim, H. W.; Wang, M.; Leber, C. A.; Nothias, L. F.; Reher, R.; Kang, K. B.; Van Der Hooft, J. J. J.; Dorrestein, P. C.; Gerwick, W. H.; Cottrell, G. W. NPClassifier: A Deep Neural Network-Based Structural Classification Tool for Natural Products. *J. Nat. Prod.* **2021**, *84* (11), 2795–2807.
- (19) Keatinge-Clay, A. T. A tylosin ketoreductase reveals how chirality is determined in polyketides. *Chem. Biol.* **2007**, *14* (8), 898–908.
- (20) Lopes, A. B.; Miguez, E.; Kummerle, A. E.; Rumjanek, V. M.; Fraga, C. A.; Barreiro, E. J. Characterization of amide bond conformers for a novel heterocyclic template of N-acylhydrazone derivatives. *Molecules* **2013**, *18* (10), 11683–11704.
- (21) Blin, K.; Shaw, S.; Augustijn, H. E.; Reitz, Z. L.; Biermann, F.; Alanjary, M.; Fetter, A.; Terlouw, B. R.; Metcalf, W. W.; Helfrich, E. J. N.; Van Wezel, G. P.; Medema, M. H.; Weber, T. antiSMASH 7.0: New and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res.* **2023**, *51* (W1), W46–W50.
- (22) Brautaset, T.; Borgos, S. E. F.; Sletta, H.; Ellingsen, T. E.; Zotchev, S. B. Site-specific mutagenesis and domain substitutions in the loading module of the nystatin polyketide synthase, and their effects on nystatin biosynthesis in. *J. Biol. Chem.* **2003**, *278* (17), 14913–14919.
- (23) Brautaset, T.; Sekurova, O. N.; Sletta, H.; Ellingsen, T. E.; Ström, A. R.; Valla, S.; Zotchev, S. B. Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei* ATCC 11455: Analysis of the gene cluster and deduction of the biosynthetic pathway. *Chem. Biol.* **2000**, *7* (6), 395–403.
- (24) Gao, L.; Su, C.; Du, X.; Wang, R.; Chen, S.; Zhou, Y.; Liu, C.; Liu, X.; Tian, R.; Zhang, L.; Xie, K.; Chen, S.; Guo, Q.; Guo, L.; Hano, Y.; Shimazaki, M.; Minami, A.; Oikawa, H.; Huang, N.; Houk, K. N.; Huang, L.; Dai, J.; Lei, X. FAD-dependent enzyme-catalysed intermolecular [4 + 2] cycloaddition in natural product biosynthesis. *Nat. Chem.* **2020**, *12* (7), 620–628.
- (25) Zhang, H.; Zhang, D.; Ray, K.; Zhu, M. Mass defect filter technique and its applications to drug metabolite identification by high-resolution mass spectrometry. *J. Mass Spectrom.* **2009**, *44* (7), 999–1016.
- (26) Grabenauer, M.; Krol, W. L.; Wiley, J. L.; Thomas, B. F. Analysis of synthetic cannabinoids using high-resolution mass spectrometry and mass defect filtering: Implications for nontargeted screening of designer drugs. *Anal. Chem.* **2012**, *84* (13), 5574–5581.
- (27) Komatsu, K.; Tsuda, M.; Tanaka, Y.; Mikami, Y.; Kobayashi, J. I. Absolute stereochemistry of immunosuppressive macrolide brasilinolide A and its new congener brasilinolide C. *J. Org. Chem.* **2004**, *69* (5), 1535–1541.
- (28) Igarashi, M.; Hayashi, C.; Homma, Y.; Hattori, S.; Kinoshita, N.; Hamada, M.; Takeuchi, T. Tubelactomicin A, a novel 16-membered lactone antibiotic, from sp.: I. Taxonomy, production, isolation and biological properties. *J. Antibiot.* **2000**, *53* (10), 1096–1101.
- (29) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Oresic, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinf.* **2010**, *11*, 395.
- (30) Perrone, A.; Capasso, A.; Festa, M.; Kemertelidze, E.; Pizza, C.; Skhirtladze, A.; Piacente, S. Antiproliferative steroidal glycosides from *Digitalis ciliata*. *Fitoterapia* **2012**, *83* (3), 554–562.