



**HAL**  
open science

# Integrated Metabolomic, Molecular Networking, and Genome Mining Analyses Uncover Novel Angucyclines From *Streptomyces* sp. RO-S4 Strain Isolated From Bejaia Bay, Algeria

Rima Ouchene, Didier Stien, Juliette Segret, Mouloud Kecha, Alice M S Rodrigues, Carole Veckerlé, Marcelino T Suzuki

## ► To cite this version:

Rima Ouchene, Didier Stien, Juliette Segret, Mouloud Kecha, Alice M S Rodrigues, et al.. Integrated Metabolomic, Molecular Networking, and Genome Mining Analyses Uncover Novel Angucyclines From *Streptomyces* sp. RO-S4 Strain Isolated From Bejaia Bay, Algeria. *Frontiers in Microbiology*, 2022, 13, pp.906161. 10.3389/fmicb.2022.906161 . hal-03746686v2

**HAL Id: hal-03746686**

**<https://hal.science/hal-03746686v2>**

Submitted on 5 Aug 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Integrated Metabolomic, Molecular Networking, and Genome Mining Analyses Uncover Novel Angucyclines From *Streptomyces* sp. RO-S4 Strain Isolated From Bejaia Bay, Algeria

Rima Ouchene<sup>1,2</sup>, Didier Stien<sup>2\*</sup>, Juliette Segret<sup>2</sup>, Mouloud Kecha<sup>1</sup>, Alice M. S. Rodrigues<sup>2</sup>, Carole Veckerlé<sup>2</sup> and Marcelino T. Suzuki<sup>2\*</sup>

<sup>1</sup> Laboratoire de Microbiologie Appliquée (LMA), Faculté des Sciences de la Nature et de la Vie, Université de Bejaia, Bejaia, Algeria, <sup>2</sup> Sorbonne Université, CNRS, Laboratoire de Biodiversité et Biotechnologies Microbiennes, LBBM, F-66650, Banyuls-sur-mer, France

## OPEN ACCESS

### Edited by:

Imen Nouioui,  
German Collection of Microorganisms  
and Cell Cultures GmbH  
(DSMZ), Germany

### Reviewed by:

Harald Gross,  
University of Tübingen, Germany  
XueHong Zhang,  
Shanghai Jiao Tong University, China

### \*Correspondence:

Didier Stien  
didier.stien@cnrs.fr  
Marcelino T. Suzuki  
suzuki@obs-banyuls.fr

### Specialty section:

This article was submitted to  
Systems Microbiology,  
a section of the journal  
Frontiers in Microbiology

Received: 28 March 2022

Accepted: 16 May 2022

Published: 23 June 2022

### Citation:

Ouchene R, Stien D, Segret J,  
Kecha M, Rodrigues AMS, Veckerlé C  
and Suzuki MT (2022) Integrated  
Metabolomic, Molecular Networking,  
and Genome Mining Analyses  
Uncover Novel Angucyclines From  
*Streptomyces* sp. RO-S4 Strain  
Isolated From Bejaia Bay, Algeria.  
Front. Microbiol. 13:906161.  
doi: 10.3389/fmicb.2022.906161

Multi-omic approaches have recently made big strides toward the effective exploration of microorganisms, accelerating the discovery of new bioactive compounds. We combined metabolomic, molecular networking, and genomic-based approaches to investigate the metabolic potential of the *Streptomyces* sp. RO-S4 strain isolated from the polluted waters of Bejaia Bay in Algeria. Antagonistic assays against *methicillin-resistant Staphylococcus aureus* with RO-S4 organic extracts showed an inhibition zone of 20 mm by using the agar diffusion method, and its minimum inhibitory concentration was 16  $\mu$ g/ml. A molecular network was created using GNPS and annotated through the comparison of MS/MS spectra against several databases. The predominant compounds in the RO-S4 extract belonged to the angucycline family. Three compounds were annotated as known metabolites, while all the others were putatively new to Science. Notably, all compounds had fridamycin-like aglycones, and several of them had a lactonized D ring analogous to that of urdamycin L. The whole genome of *Streptomyces* RO-S4 was sequenced to identify the biosynthetic gene cluster (BGC) linked to these angucyclines, which yielded a draft genome of 7,497,846 bp with 72.4% G+C content. Subsequently, a genome mining analysis revealed 19 putative biosynthetic gene clusters, including a grincamycin-like BGC with high similarity to that of *Streptomyces* sp. CZN-748, that was previously reported to also produce mostly open fridamycin-like aglycones. As the ring-opening process leading to these compounds is still not defined, we performed a comparative analysis with other angucycline BGCs and advanced some hypotheses to explain the ring-opening and lactonization, possibly linked to the uncoupling between the activity of *GcnE* and *GcnM* homologs in the RO-S4 strain. The combination of metabolomic and genomic approaches greatly improved the interpretation of the metabolic potential of the RO-S4 strain.

**Keywords:** marine *Streptomyces*, antibacterial activity, MRSA, metabolomic analysis, molecular networking, genome mining

## INTRODUCTION

The emergence of novel mechanisms of antimicrobial resistance is increasing and spreading worldwide, posing a challenge to mankind. The World Health Organization has stated that antibiotic resistance will be one of the biggest threats to human health in the future (Prestinaci et al., 2015). Multidrug-resistant organisms have become common not only in hospital settings but also in the wide community settings, suggesting that reservoirs of antibiotic-resistant bacteria are present outside hospitals, as reviewed by Munita and Arias (2016). This antibiotic resistance crisis has been attributed to the overuse and inappropriate use of these drugs and the lack of antimicrobial drug development by the pharmaceutical industry due to reduced economic incentives and difficult regulatory requirements (Ventola, 2015; Chokshi et al., 2019). Methicillin-resistant *Staphylococcus aureus* (MRSA) is the most common cause of nosocomial infections as it is very capable of developing antibiotic resistance (Chambers and DeLeo, 2009; Barbosa et al., 2020). Many challenges are faced by laboratories and clinicians in the diagnosis and treatment of MRSA infections, some of which were highlighted by Edwards et al. (2014). It is thus clear that the search for new bioactive compounds to combat antimicrobial resistance is a research priority.

Marine environments represent a largely unexplored source for the isolation of new microorganisms (Bhatnagar and Kim, 2010). They display a unique combination of environmental conditions and organisms with distinct metabolic capabilities to adapt and thrive (Bose et al., 2015; Stien, 2020). A large number of bioactive compounds have been isolated from marine organisms (Lane and Moore, 2011; Blunt et al., 2012; Jose and Jha, 2016), particularly Actinobacteria, which have been the main source of natural products in the past (Hassan and Shaikh, 2017; Lewis, 2020; Jose et al., 2021). Among the latter, the *Streptomyces* genus is well-known for its ability to produce a wide range of bioactive metabolites as well as antibacterial, anticancer, antifungal, antiparasitic, and immunosuppressive agents (Baltz, 2008; Olano et al., 2009), representing the most prolific source of bioactive metabolites that have been approved for clinical use, notably as antibiotics (Bérdy, 2005).

Traditionally, activity-guided fractionation of metabolite extracts, followed by purification and characterization of metabolites, has commonly been used for natural product research, but this approach often leads to the isolation of already known molecules. More recently, significant developments in genetics, genomics, and data analysis have greatly changed natural product research, leading to a new era in the emerging field of systems biology. Consequently, new avenues were opened for the discovery of novel compounds from actinomycetes (e.g., Dettmer et al., 2007; Gomez-Escribano et al., 2016; Jose and Jha, 2016; Cornell et al., 2018; Hu et al., 2020; Jose et al., 2021). Interestingly, metabolomics and genomics approaches have proven to be efficient and promising tools for defining phenotypes in a dynamic context, with the potential to reduce rediscovery rates (Khoomrung et al., 2017; Pye et al., 2017; van Der Hooff et al., 2020), and several tools have been

designed for this purpose, as reported by Caesar et al. (2021). These approaches have been successfully applied to study the chemical diversity of marine bacteria and to uncover novel bioactive molecules (Macintyre et al., 2014; Schneider et al., 2018; Ye et al., 2020), despite the challenges encountered due to the complexity of biological matrices (Liu and Locasale, 2017). More recently, molecular networking, a tandem mass spectrometry (MS/MS) data organizational approach, has been introduced in the field of drug discovery (Wang et al., 2016), and the combination of system analyses involving multi-omics data and genome-scale, metabolic network models has greatly contributed to exploring bioactive Actinobacteria (Palazzotto et al., 2019) and has great potential to accelerate natural product discovery.

Here, we investigated the secreted metabolome of *Streptomyces* sp. RO-S4 in the quest for novel antimicrobial compounds against MRSA. For this purpose, we used a metabolomic approach based on ultra-performance high-resolution tandem mass spectrometry (UPLC-HRMS/MS), followed by the creation of a molecular network using the Global Natural Product Social Molecular Networking (GNPS) analysis. These analyses were combined with a genomic analysis to refine and further annotate the structural hypothesis generated, and conversely, to understand the biosynthesis of the major angucyclines produced by this strain.

## MATERIALS AND METHODS

### Isolation of the RO-S4 Strain

The RO-S4 strain was isolated from polluted seawater that was collected from the coastline of Bejaia City at S3 point (36°43'55.2 "N5°04'37.9"E) by Ouchene et al. (2022). It was isolated after filtration onto a 0.45 µm pore size membrane filter (Hirsch and Christensen, 1983) and laid onto solid starch casein agar medium (SCA, also named M2 medium by Jensen et al., 1991).

### Molecular Identification of the RO-S4 Isolate

The initial molecular identification of the RO-S4 strain was based on the 16S rRNA gene sequence. Genomic DNA was extracted from the grown strain using the Wizard® Genomic DNA Purification Kit (Promega, USA) according to the manufacturer's instructions. PCR and sequencing were realized as previously described (Fagervold et al., 2013), utilizing universal primers recommended for bacteria 27Fmod: 5'AGRGTGGATCMTGGCTCAG 3' and 1492Rmod: 5'TACGGYTACCTTGTTAYGACTT 3'. The PCR product was purified with a purification kit (Promega, USA), and then sequenced by the dideoxy termination reaction using an AB3130 DNAxl sequencer. The obtained 16S rRNA sequence was identified by comparison to the EZBioCloud database (<https://www.ezbiocloud.net/>) recommended by Yoon and co-workers (Yoon et al., 2017). The 16S rRNA gene sequence of strain RO-S4 was deposited in the GenBank database under the accession number (MW448345).

## Whole Genome Sequencing and Assembly of *Streptomyces* sp. RO-S4 Strain

Genomic DNA was isolated from 50 ml of RO-S4 grown in starch casein broth for 12 days at 28 °C with shaking (150 rpm/min). The DNA was extracted using the Bacteria Genomic DNA Extraction Kit (Promega, United States) according to the manufacturer's instructions. Illumina whole genome sequencing was performed by the Genotoul facility in Toulouse, France. Briefly, libraries (Truseq nano HT, Illumina) were constructed using 200 ng of purified DNA and sequenced on a Novaseq 6000 sequencer (Illumina), generating 93 million paired 150 base pair (bp) reads. The entire dataset was assembled using SPAdes (v3.14.0) (Prjibelski et al., 2020) with the option "careful". The assembly was manually curated to remove contigs with low (<500) coverage and low (<55%) G+C content. Since a gene cluster of interest was truncated in a contig, we manually extended it using blast searches against raw reads and re-assembly of reads and of a downstream contig using the *gap4* tool of the Staden package (<http://staden.sourceforge.net>). This final assembly was auto-annotated using PROKKA v. 1.14.6 (Seemann, 2014) using *Streptomyces* sp. Vc74B-19 protein descriptions and the annotation was manually curated prior to final submission to 1) add the ORF corresponding to *GcnM* in *Streptomyces* sp. LR32, 2) remove partial rRNA genes and possible adapters in contig ends. The genomic sequence (JAJQKZ000000000) was compared to that of close strains based on the 16S rRNA gene analysis described above, relatives (type strains) and genomes in the NCBI GenBank, based on top *blastp* hits using the five housekeeping genes described by Antony-Babu and coworkers (Antony-Babu et al., 2017). ANI values were calculated using the OrthoANI tool available through the EZbiocloud server (<https://www.ezbiocloud.net/tools/orthoani>). Further genome-based taxonomic identification using estimated DNA-DNA hybridization was performed using the TYGS server (Meier-Kolthoff and Goker, 2019).

## Antimicrobial Assays

The antibacterial potential of the RO-S4 isolate was evaluated against *methicillin-resistant S. aureus* (MRSA) ATCC 43300 by the agar diffusion method (Shomura et al., 1979). The agar cylinders (8 mm in diameter) of the RO-S4 strain (starch casein agar medium, incubation for 14 days at 28°C) were inserted into Mueller Hinton plates previously seeded with the targeted bacterium at 10<sup>7</sup> CFUs/ml. The plates were placed for 2 h at 4°C to allow metabolites to diffuse prior to MRSA growth, and antibacterial activity was estimated by measuring the inhibition zone around the agar disc after incubation of the plates for 24 h at 37 °C.

## Culture Strain and the Production of raw Extract

The production of bioactive compounds by the selected strain was carried out by agar surface fermentation, according to Nkanga and Hagedorn (1978). Briefly, the RO-S4 isolate was initially grown on an SCA medium. After 14 days, the mycelium layers were peeled off and extracted overnight in ethyl acetate

(EtOAc), covering the entire surface, and the ethyl acetate extract was drawn. Subsequently, the organic extract was concentrated under a vacuum with a rotary evaporator at 40°C and then stored at -80°C until further analysis. A control, uninoculated medium, was extracted with the same protocol.

## The Minimum Inhibitory Concentration of the RO-S4 Crude Extract

The minimal inhibitory concentration (MIC) of the RO-S4 ethyl acetate extract was evaluated against the methicillin-resistant *S. aureus* strain ATCC 43300 by the broth microdilution method according to the Clinical and Laboratory Standards Institute (CLSI) guidelines and as described by Wiegand et al. (2008). The assays were performed in serial dilutions (in triplicate) in 96-well plates. Briefly, the EtOAc extract was diluted in DMSO and tested at different concentrations ranging from 256 to 0.5 µg/ml. The targeted bacterium culture was prepared in Mueller Hinton broth at 2•10<sup>5</sup> CFUs/ml. Afterward, 10 µl of the test bacterial culture was pipetted into each well. The last column (column 12) with no inoculum served as a sterility control, while wells that were not treated with the crude extract served as a negative control (column 11). The final volume of each well was adjusted to 100 µl. The microplate was shaken gently, then incubated for 24 h at 37°C. Inhibition was evaluated as well, where the growth medium appeared clear, indicating that the test extract prevented the growth or killed the bacteria.

## UHPLC-HRMS Profiling

The protocol for high-resolution Full MS data dependent MS<sup>2</sup> analyses was adapted from previous reports (Stien et al., 2019, 2020). Briefly, the UHPLC system was an Ultimate 3000 (Thermo Fisher Scientific) coupled to a Q-Exactive Focus Orbitrap detector and equipped with a Phenomenex Luna Omega polar C-18 column 150 mm × 2.1 mm, 1.6 µm. Here, crude bacterial and culture medium (starch casein broth) extracts were dissolved in MeOH at a concentration of 1.5 mg/ml. Pure methanol injections were used as blanks for metabolomics. In HPLC, the solvent system was a mixture of water (solution A) with increasing proportions of acetonitrile (solution B), both solvents modified with 0.1% formic acid. Here, the gradient was as follows: 5% of solution B for 5 min before injection, then from 1 to 12 min, a linear increase of B up to 100%, followed by 100% of solution B for 8 min.

## MS/MS Molecular Networking Analysis and Spectra Annotation

The molecular network was constructed using the Global Natural Product Social Molecular Networking (GNPS) platform available at (<https://gnps.ucsd.edu>) as recommended by Wang and collaborators (Wang et al., 2016), using the molecular networking (MN) tool. The MS<sup>2</sup> data of the crude extract, solvent (blank), and culture media were converted from the RAW into mzXML file format using the Proteowizard MSConvert tool version (3.0.20104), then uploaded to GNPS. For MN construction, the precursor ion mass tolerance was set at 0.005 Da and the MS<sup>2</sup> fragment ion tolerance was set at 0.01 Da. A network was created where edges were filtered to have a

cosine score above 0.77 and 11 or more matched peaks. The maximum size of a molecular family was set at 85. The MS<sup>2</sup> spectra in the network were searched against the 'GNPS spectra library'. All matches between network and library spectra were required to have a score above 0.7 and at least 6 matched peaks. The visualization of the molecular network was performed in Cytoscape (3.8.0), which allowed its visualization as a network of nodes and edges (Otasek et al., 2019). Redundancies and adducts were cleared manually. In **Figure 2**, node numbers are consensus parent masses, node size is linked to the relative molecular ion intensity based on the peak area measured from the extracted ion chromatogram. Peak areas were measured automatically with the FreeStyle Genesis algorithm, sometimes modified manually if found unfitting, and then pasted manually into the Cytoscape table. This information was also used to create pie charts in which each portion represents the relative peak area of different isomers included in the same node (each GNPS node is a cluster of MS<sup>2</sup> spectra that may come from different isomeric protonated molecular ions). For nodes gathering protonated molecular ions and in-source fragments of higher molecular weight compounds, only protonated molecular ion integrations were included for peak area information. Any ion present in the network solely due to in-source fragmentation was given the arbitrary extracted ion intensity of 0 and was white in the network. Furthermore, spectra of interest were manually annotated using different databases and tools, including Sirius (Dührkop et al., 2019), MetFrag (Ruttkies et al., 2016), available at (<https://msbi.ipb-halle.de/MetFrag/>), Pubchem, Sci-Finder, and Mass Bank of North America (MoNA, <https://mona.fiehnlab.ucdavis.edu/>). Detailed spectral data are provided in supporting information (**Supplementary Figures S1–S75**), and the raw spectral files are available (see Data Availability).

## Prediction of Secondary Metabolite Biosynthetic Gene Clusters in the RO-S4 Genome

The sequenced genome (DNA sequence as input) of the RO-S4 strain was analyzed for the prediction of secondary metabolites and biosynthetic gene clusters (BGCs) using the genome mining tool antiSMASH version 6.0.1 (Blin et al., 2021) available through (<https://antismash.secondarymetabolites.org>) using the “relaxed” option. Synteny plots of the RO-S4 angucycline BGC were performed using Clinker v 0.0.21 (Gilchrist and Chooi, 2021). We also submitted the PROKKA annotated GenBank formatted files using different annotation parameters to antiSMASH. All results and descriptions of the annotation pipelines are available (see Data Availability below).

## Phylogenetic Analysis of *UrdE* and *UrdM* Homologs

Phylogenetic analysis of *UrdE* and *UrdM* homologs, enzymes involved in the oxidation of angucyclines (Rix et al., 2002) in RO-S4, and in the BGCs of other angucyclines and the ring-opening Baeyer-Villiger monooxygenase *MtmOIV* in the mithramycin BGC was performed using SeaView Version 4.6 (Gouy et al., 2010) and MEGA v.11 (Tamura et al., 2021). Briefly, amino-acid

sequences were aligned using muscle and a mask was created using the gblocks options “allow smaller blocks positions” and “allow gap positions” followed by manual curation of the mask, and a tree was generated by maximum likelihood with the LG+F model and gamma distribution using MEGA v.11 based on 359 homologous positions. A neighbor-joining tree was constructed using the JTT substitution model. The robustness of both trees was evaluated by bootstrap analysis using 100 replicates. The raw alignments, the regions used for reconstruction, and the results of model testing are available (see Data Availability).

## RESULTS

### Initial Identification, and Antimicrobial Assays

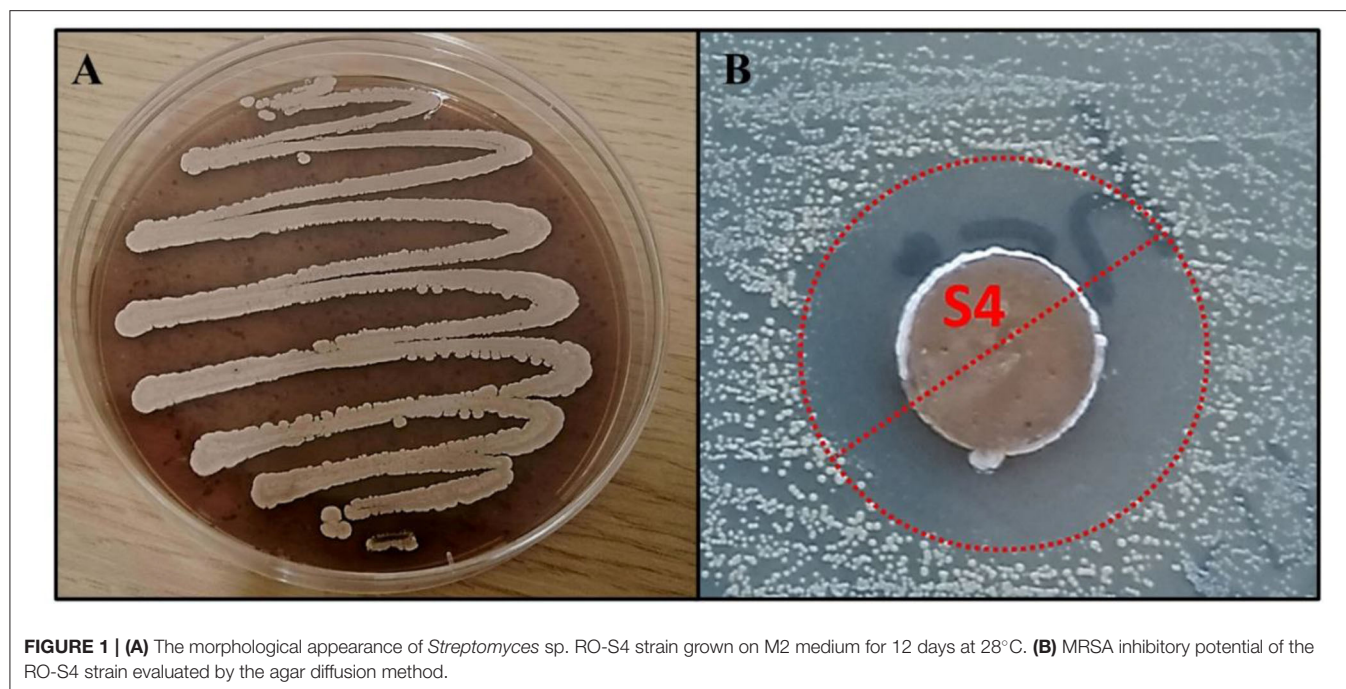
The RO-S4 strain was isolated from the Bay of Bejaia City in Algeria. It grew well on the SCA medium, showing substrate growth typical for *Streptomyces* strains, with a brown powdery aspect and producing a dark-brown pigment (**Figure 1A**). Antimicrobial activity was first evaluated against the methicillin-resistant *S. aureus* ATCC 43300 strain by the agar diffusion method. It exhibited antagonistic activity against this bacterium with an inhibition zone estimated at 20 mm (**Figure 1B**). A MIC value of 16 µg/ml was observed against the methicillin-resistant *Staphylococcus aureus* ATCC 43300 strain. The 16S rRNA gene sequence indicated that the strain belongs to the *Streptomyces* genus, with 99.79% identity to *Streptomyces albogriseolus* NRRL B-1305<sup>T</sup>.

### Whole Genome Sequencing

The whole genome of *Streptomyces* sp. RO-S4 was sequenced using the Illumina Novaseq technology. The complete genome consisted of 7,497,846 bp in 129 contigs with an NC50 of 117,347 bp and a 72.4% G+C content. We were unable to obtain a full-length rRNA in our assembly. The closest genome to *Streptomyces* sp. RO-S4 sequenced from a type strain was that of *S. albogriseolus* JCM 4616<sup>T</sup> [assembly GCA\_014650475; 90.94% ANI; Estimated % DNA-DNA hybridization d<sub>0</sub> 59.2 (55.6–62.8), d<sub>4</sub> 41.7 (39.2–44.2) and d<sub>6</sub> 55.7 (52.6–58.9); 0.06 difference in G+C content]. The genome-based tree places *Streptomyces* sp. RO-S4 clearly exterior to the clade containing *S. albogriseolus* JCM 4616<sup>T</sup>. Overall, these values indicate strain RO-S4 might represent a new *Streptomyces* species.

### Untargeted Metabolomic Analysis and Molecular Networking

The metabolomic profile of the active ethyl acetate (EtOAc) crude extract was investigated using UPLC-HRMS. An examination of the MS and collision-induced MS/MS (MS<sup>2</sup> hereafter) spectra of the main constituents of the mixture indicated that most metabolite molecular ions fragmented to yield a product at *m/z* 487.1600 corresponding to the formula C<sub>25</sub>H<sub>27</sub>O<sub>10</sub><sup>+</sup> (calcd. 487.1599) (**Figure 2**). The same ion was also detected as a protonated molecular ion corresponding to compound **1** (**Table 1**). Compound **1** was annotated either as aquayamycin or as fridamycin A or B (**Figure 4**) by various dereplication tools. A comparison with experimental spectra from the MoNA



**FIGURE 1 | (A)** The morphological appearance of *Streptomyces* sp. RO-S4 strain grown on M2 medium for 12 days at 28°C. **(B)** MRSA inhibitory potential of the RO-S4 strain evaluated by the agar diffusion method.

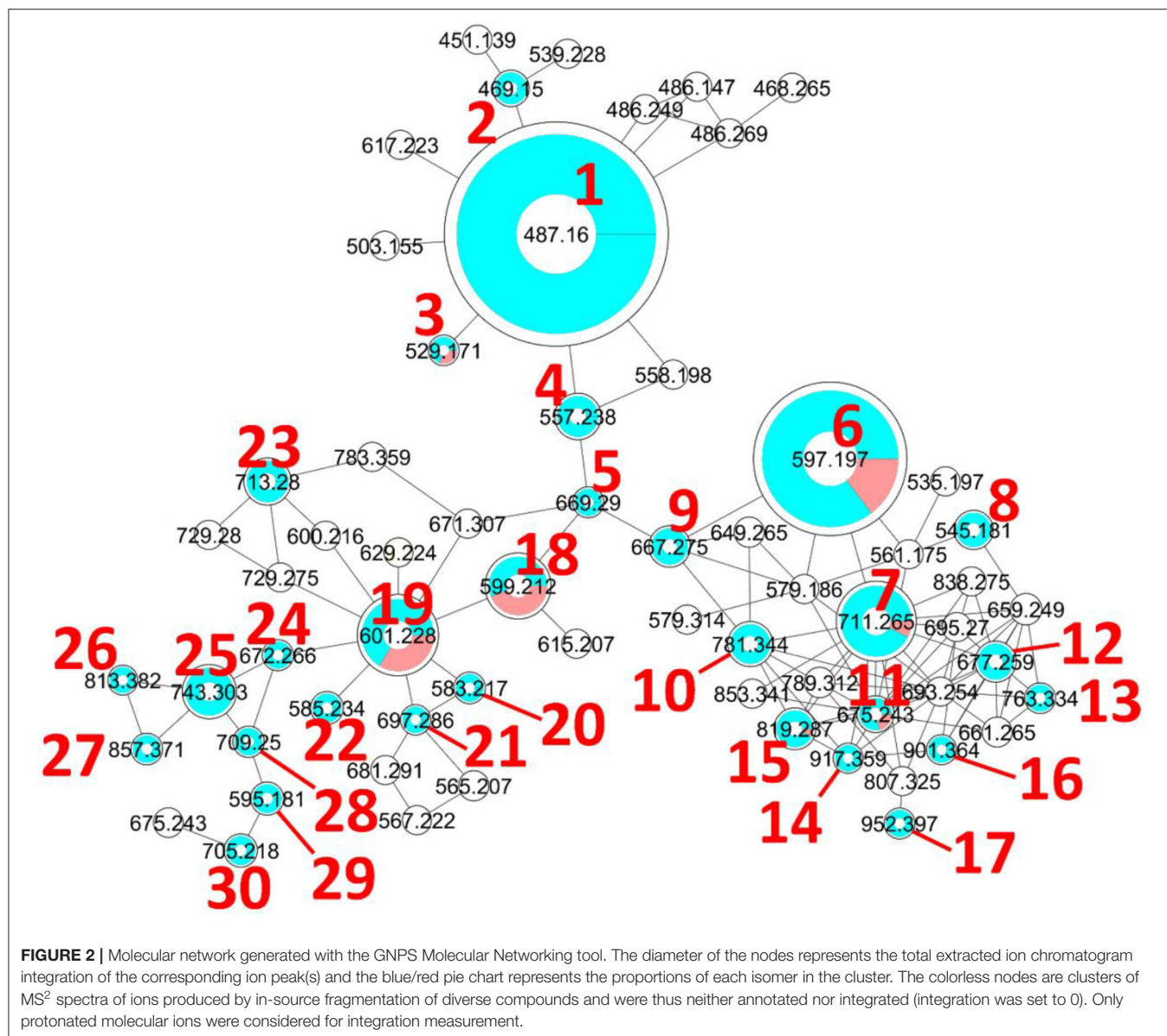
database confirmed that compound **1** was fridamycin A or its diastereomer, fridamycin B. The MS<sup>2</sup> spectra of fridamycin A and aquayamycin are very similar (**Supplementary Figure S3**). Nevertheless, two fragment ions are diagnostic. These are the ions at  $m/z$  347.09 and 427.14, the relative intensities of which are very low in fridamycin A when compared to those of aquayamycin. Fridamycin A has been previously shown to have anti-MRSA as well as antidiabetic activities (Yoon et al., 2019; Sabido et al., 2020).

As mentioned above, the  $[1+H]^+$  ion was also produced as a fragment resulting from the in-source fragmentation of many metabolites in the profile. The MS<sup>2</sup> spectra of all ions lead to a 487.1599 fragment clustered in the same node in the molecular network (MN). The MS<sup>2</sup> spectrum obtained for the protonated molecular ion of compound **1** was compared to all other MS<sup>2</sup> spectra of equimassic ions found whenever other more complex metabolites were fragmented in the source. The fragmentation patterns were all similar, confirming that many metabolites in the MN were fridamycin A or B analogs. It was deduced that the strain biosynthesizes the central fridamycin core and then adds various substituents to generate its diverse products. Notably, the pentacyclic aquayamycin subunit was not detected in any of the annotated metabolites. This observation was supported by the presence of a biosynthetic gene cluster very close to that of *Streptomyces* sp. CNZ-748 (Shang et al., 2021), which also produces a majority of fridamycin-like compounds.

The parameters for MN were set to construct the best representative network containing all fridamycin analogs (**Figure 2**). The MN was constituted of three groups of ions. In the first group, the annotation was propagated from **1** as follows. Compound **2** is a dehydrofridamycin based on its

molecular formula and MS<sup>2</sup> spectrum. The molecular formula of compound **3** was C<sub>27</sub>H<sub>28</sub>O<sub>11</sub>, which might be annotated as an acetyl-fridamycin A or B, while the position of the acetyl group could not be inferred from MS<sup>2</sup>. Compound **4**, whose molecular formula was C<sub>30</sub>H<sub>36</sub>O<sub>10</sub> ( $m/z$  for  $[M+H]^+$  669.2903, calcd. 669.2905) was a fridamycin bearing a C<sub>5</sub>H<sub>10</sub> substituent. A fridamycin isopentyl ester was thought to be a reasonable putative structure based on the biosynthetic considerations below. Compound **5** could not be annotated more precisely than just with its molecular formula, but its MS<sup>2</sup> spectrum is also one of a fridamycin analog. These considerations indicated that **5** was new to Science.

In the second part of the fridamycin MN, compound **6a** protonated molecular ion was at  $m/z$  597.1973, corresponding to the formula C<sub>31</sub>H<sub>33</sub>O<sub>12</sub><sup>+</sup> (calcd.  $m/z$  597.1967). The formula and fragmentation pattern were consistent with the annotation of **6a** as fridamycin D, previously reported to have antibacterial activity (Maskey et al., 2003). Annotation as fridamycin D was also supported by Sirius. An isomer of **6a** (**6b**) was also detected at a different retention time in smaller relative proportions in the strain's metabolomic profile. Another major constituent of the strain metabolome was compound **7a**, whose protonated molecular ion at  $m/z$  711.2645 corresponded to the formula C<sub>37</sub>H<sub>43</sub>O<sub>14</sub><sup>+</sup> (calcd. 711.2647). The molecular formula indicated that angucycline **7a** may be vineomycin C, a potent inhibitor of inducible nitric oxide synthase (Alvi et al., 2000). It was annotated as vineomycin C by Sirius as well, although with a 57% CSI finger ID score. In the collision-induced MS<sup>2</sup> spectrum of **7a** protonated molecular ion, the rhodinose (or its distereoisomer amicetose) and the aculose oxonium ions were present at  $m/z$  115.0756 and 111.0443 (**Supplementary Table S1**),



respectively, while the presence of the fridamycin A aglycone was ascertained based on the common fridamycin A fragment ions (**Supplementary Figure S17**). Nonetheless, the first fragmentation steps in the MS<sup>2</sup> spectrum suggested that the sugar sequence might be different from that of vineomycin C. The protonated molecular ion lost both water ( $m/z$  for C<sub>37</sub>H<sub>41</sub>O<sub>13</sub><sup>+</sup> 693.2258) and a C<sub>6</sub>H<sub>12</sub>O<sub>3</sub> fragment ( $m/z$  for C<sub>31</sub>H<sub>31</sub>O<sub>11</sub><sup>+</sup> 597.1938), which could only be assigned to the rhodiose/amicetose moiety. Hence, the rhodiose could not be placed between the aglycone and the aculose, as in vineomycin C, and compound **7a** must be considered new to Science unless the published structure of vineomycin C requires revision. Possible annotations are reported in **Figure 3**. These alternative proposals could account for the preferential fragmentation observed in the MS<sup>2</sup> spectrum. The proximity of **7a** with fridamycin D (**6a**) in

the MN suggested that the most probable annotation might be the one in which aculose underwent a Michael addition, as in **6a** (**Figure 3**).

A minor isomer of **7b** also appeared in the MN under the same cluster. Based on the same considerations as for **7a**, **7b** could not be annotated as vineomycin C either. It was therefore new to Science. Compound **8**, a protonated molecular ion with an  $m/z$  of 545.1808 was found next to compounds **6** and **7**, corresponding to the formula C<sub>31</sub>H<sub>29</sub>O<sub>9</sub><sup>+</sup> (calcd. 545.1806). The formula suggested that **8** might be marangucycline B (Song et al., 2015). Sirius also annotated this compound as marangucycline B, although with 58% confidence only. Marangucycline B is an analog of fridamycin D with a modified aglycone presenting weak antibacterial activity but strong *in vitro* cytotoxicity activity against cancer cell lines (Song et al., 2015). The presence of an

**TABLE 1** | Annotated angucyclines in the metabolomic profile of strain RO-S4.

	Exp. <i>m/z</i>	Ion type	Molecular formula	Th. <i>m/z</i>	# <sup>a</sup>	t <sub>R</sub> (min)	Annotation <sup>b</sup>
1	487.1600	[M+H] <sup>+</sup>	C <sub>25</sub> H <sub>26</sub> O <sub>10</sub>	487.1599	1	8.59	Fridamycin A or B ( <b>Figure 4</b> )
2 <sup>c</sup>	469.1492	[M+H] <sup>+</sup>	C <sub>25</sub> H <sub>24</sub> O <sub>9</sub>	469.1493	1	8.48	[1-H <sub>2</sub> O]
3 <sup>c</sup>	529.1708	[M+H] <sup>+</sup>	C <sub>27</sub> H <sub>29</sub> O <sub>11</sub>	529.1704	1	8.58	Acetyl-fridamycin A or B
4 <sup>c</sup>	557.2383	[M+H] <sup>+</sup>	C <sub>30</sub> H <sub>36</sub> O <sub>10</sub>	557.2381	1	11.69	See <b>Figure 4</b>
5 <sup>c</sup>	669.2903	[M+H] <sup>+</sup>	C <sub>36</sub> H <sub>44</sub> O <sub>12</sub>	669.2906	1	13.08	n.a.
6	a: 597.1973; b: 597.1969	[M+H] <sup>+</sup>	C <sub>31</sub> H <sub>32</sub> O <sub>12</sub>	597.1967	2	a: 11.01; b: 12.57	a: Fridamycin D ( <b>Figure 4</b> )
7 <sup>c</sup>	a: 711.2650; b: 711.2656	[M+H] <sup>+</sup>	C <sub>37</sub> H <sub>42</sub> O <sub>14</sub>	711.2647	2	a: 11.87; b: 11.17	See <b>Figure 4</b>
8 <sup>c</sup>	545.1808	[M+H] <sup>+</sup>	C <sub>31</sub> H <sub>28</sub> O <sub>9</sub>	545.1806	1	14.21	n.a.
9 <sup>c</sup>	667.2751	[M+H] <sup>+</sup>	C <sub>36</sub> H <sub>42</sub> O <sub>12</sub>	667.2749	1	13.66	See <b>Figure 4</b>
10 <sup>c</sup>	781.3436	[M+H] <sup>+</sup>	C <sub>42</sub> H <sub>52</sub> O <sub>14</sub>	781.3430	1	14.31	See <b>Figure 4</b>
11 <sup>c</sup>	a: 675.2438; b: 675.2436	[M+H] <sup>+</sup>	C <sub>37</sub> H <sub>38</sub> O <sub>12</sub>	675.2436	2	a: 13.74; b: 13.10	See <b>Figure 4</b>
12 <sup>c</sup>	677.2597	[M+H] <sup>+</sup>	C <sub>37</sub> H <sub>40</sub> O <sub>12</sub>	677.2593	1	11.02	See <b>Figure 4</b>
13 <sup>c</sup>	763.3328	[M+H] <sup>+</sup>	C <sub>42</sub> H <sub>50</sub> O <sub>13</sub>	763.3324	1	12.88	n.a.
14 <sup>c</sup>	917.3601	[M+H] <sup>+</sup>	C <sub>49</sub> H <sub>56</sub> O <sub>17</sub>	917.3590	1	12.03	See <b>Figure 4</b>
15 <sup>c</sup>	a: 819.2863; b: 819.2869	[M+H] <sup>+</sup>	C <sub>43</sub> H <sub>46</sub> O <sub>16</sub>	819.2859	2	a: 12.64; b: 12.95	n.a.
16 <sup>c</sup>	901.3647	[M+H] <sup>+</sup>	C <sub>49</sub> H <sub>56</sub> O <sub>16</sub>	901.3641	1	12.94	See <b>Figure 4</b>
17	952.3957	[M+NH <sub>4</sub> ] <sup>+</sup>	C <sub>49</sub> H <sub>58</sub> O <sub>18</sub>	952.3961	1	12.27	Vineomycin B2 ( <b>Figure 4</b> )
18 <sup>c</sup>	a: 599.2124; b: 599.2125	[M+H] <sup>+</sup>	C <sub>31</sub> H <sub>34</sub> O <sub>12</sub>	599.2123	2	a: 10.11; b: 10.53	See <b>Figure 4</b>
19 <sup>c</sup>	a: 601.2284; b: 601.2288	[M+H] <sup>+</sup>	C <sub>31</sub> H <sub>36</sub> O <sub>12</sub>	601.2279	2	a: 10.06; b: 10.95	See <b>Figure 4</b>
20 <sup>c</sup>	583.2170	[M+H] <sup>+</sup>	C <sub>31</sub> H <sub>34</sub> O <sub>11</sub>	583.2174	1	8.62	[19a-H <sub>2</sub> O]
21 <sup>c</sup>	697.2861	[M+H] <sup>+</sup>	C <sub>37</sub> H <sub>44</sub> O <sub>13</sub>	697.2855	1	9.32	See <b>Figure 4</b>
22 <sup>c</sup>	585.2333	[M+H] <sup>+</sup>	C <sub>31</sub> H <sub>36</sub> O <sub>11</sub>	585.2330	1	10.87	n.a.
23	713.2782	[M+H] <sup>+</sup>	C <sub>37</sub> H <sub>44</sub> O <sub>14</sub>	713.2804	1	11.88	n.a.
24 <sup>c</sup>	672.2655	[M+H] <sup>+</sup>	C <sub>34</sub> H <sub>41</sub> NO <sub>13</sub>	672.2652	1	9.70	See <b>Figure 4</b>
25 <sup>c</sup>	743.3031	[M+H] <sup>+</sup>	C <sub>37</sub> H <sub>46</sub> N <sub>2</sub> O <sub>14</sub>	743.3022	1	9.44	See <b>Figure 4</b>
26 <sup>c</sup>	813.3816	[M+H] <sup>+</sup>	C <sub>42</sub> H <sub>56</sub> N <sub>2</sub> O <sub>14</sub>	813.3804	1	11.55	n.a.
27 <sup>c</sup>	857.3709	[M+H] <sup>+</sup>	C <sub>43</sub> H <sub>56</sub> N <sub>2</sub> O <sub>16</sub>	857.3703	1	10.24	See <b>Figure 4</b>
28 <sup>c</sup>	709.2490	[M+H] <sup>+</sup>	C <sub>37</sub> H <sub>40</sub> O <sub>14</sub>	709.2491	1	10.95	n.a.
29 <sup>c</sup>	595.1808	[M+H] <sup>+</sup>	C <sub>31</sub> H <sub>30</sub> O <sub>12</sub>	595.1810	1	9.47	n.a.
30 <sup>c</sup>	705.2182	[M+H] <sup>+</sup>	C <sub>37</sub> H <sub>36</sub> O <sub>14</sub>	705.2178	1	11.83	n.a.

<sup>a</sup>Number of isomers detected in the MN cluster.<sup>b</sup>Proposed structures are annotations based on literature data and automatic and manual analysis of MS<sup>2</sup> spectra.

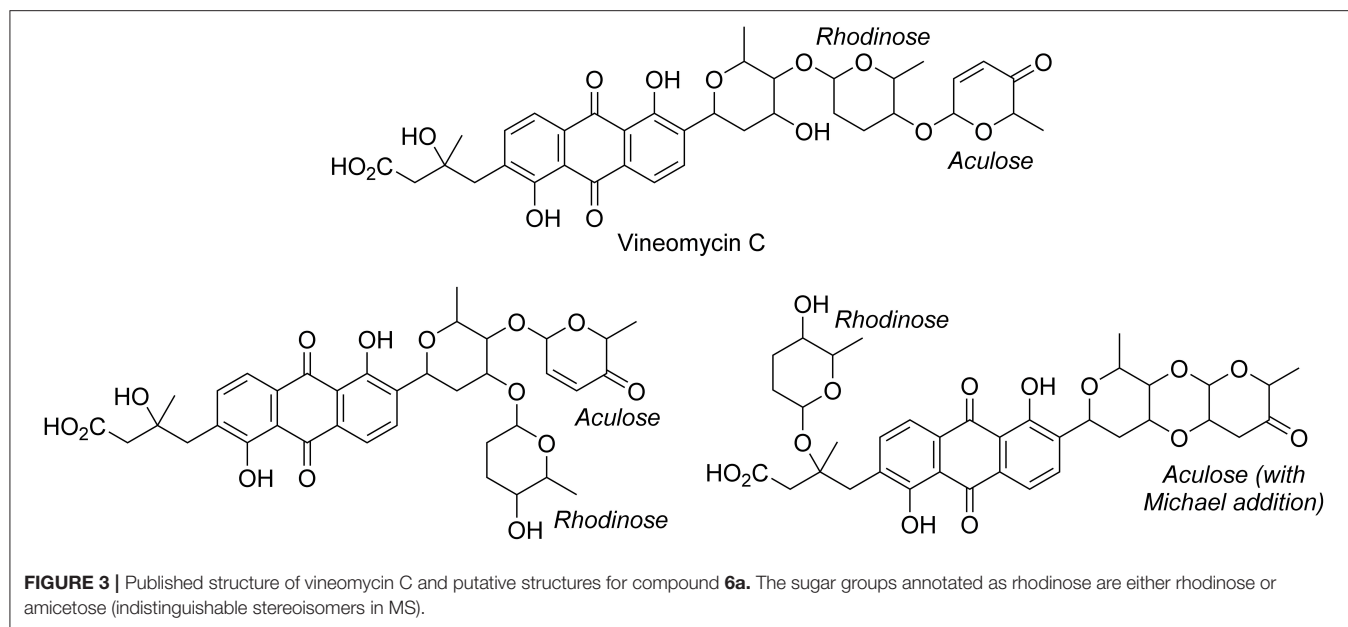
n.a., not annotated.

<sup>c</sup>Novel compound.

aculose subunit has been confirmed by MS<sup>2</sup> (*m/z* 111.0442), but the fragmentation pattern was not attributable to marangucycline B. Hence, while this compound is probably not marangucycline B it could not be annotated further. Angucycline **9** was also new to Science. Its protonated molecular ion at *m/z* 667.2753 corresponded to the formula C<sub>36</sub>H<sub>43</sub>O<sub>12</sub><sup>+</sup> (calcd. 667.2749). In the MS<sup>2</sup> spectrum, the fragmentation from the protonated molecular ion to the product at *m/z* 579.1851 corresponded to the loss of a neutral fragment the formula of which was C<sub>5</sub>H<sub>12</sub>O (**Supplementary Figure S23**). This group has been annotated above as an isopentanol, probably linked to the carboxylic acid

moiety. A carbon monoxide loss was also detected from *m/z* 561.1750 to 533.1812, indicating that the carboxylic acid side chain of the aglycone should be present. The MS<sup>2</sup> spectrum also showed a hydrated aculose oxonium ion at *m/z* 129.0549 and its dehydrated form at *m/z* 111.0443. For all these reasons, compound **9** was annotated as shown in **Table 1**. The protonated molecular ion of compound **10** was found at *m/z* 781.3438, a mass that corresponds to the formula C<sub>42</sub>H<sub>53</sub>O<sub>14</sub><sup>+</sup> (calcd. 781.3430). An isopentanol, a rhodinoside/amicetose, and an aculose were pointed out in the fragmentation spectrum of the parent ion. Rhodinoside and aculose oxonium ions were also detected at *m/z* 115.0754 and



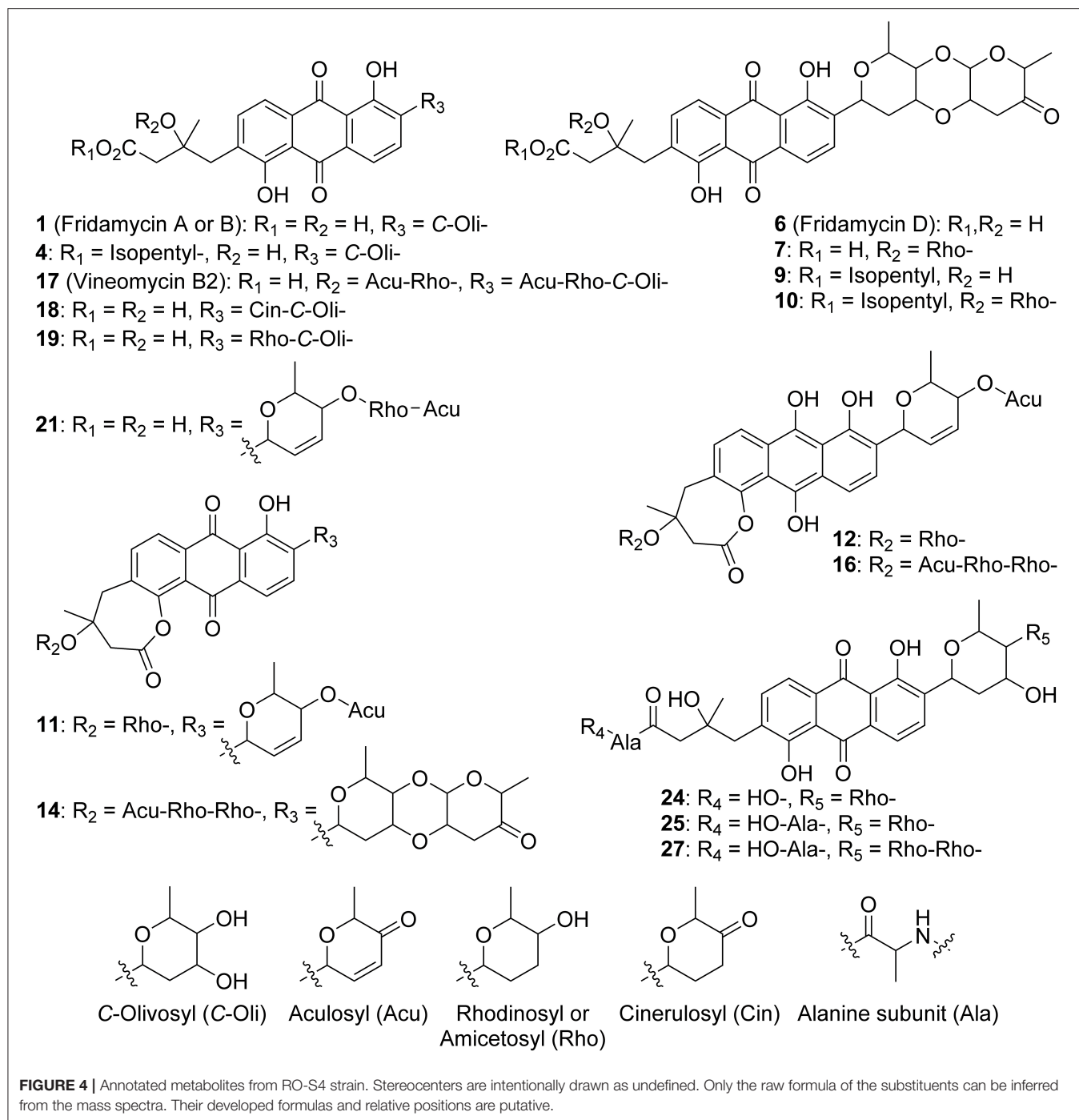


111.0442, respectively, showing that **10** should be annotated as the isopentanol ester of **7a/b** (Figure 4).

The formula of compound **11a** protonated molecular ion was found to be  $C_{42}H_{53}O_{14}^+$  (exp.  $m/z$  675.2438, calcd. 675.2436). This corresponded to the grincamycin H molecular formula (Zhu et al., 2017). Nonetheless, the  $MS^2$  spectrum indicated that **11a** successively lost rhodnose/amicetose and aculose, and therefore could not be annotated as grincamycin H. The mass of the protonated aglycone after oses fragmentation was  $m/z$  451.1388, corresponding to a doubly dehydrated fridamycin A. We hypothesized that one hydroxyl group of the olivose side chain might be dehydrated and that the second  $H_2O$  loss might be explained by ring closure of the lactone as in urdamycin L (Rix et al., 2003), a possibility supported by the analysis of the biosynthetic gene cluster discussed below. Compound **11b** was presumably a minor diastereoisomer of **11a** due to the high similarity of their respective  $MS^2$  spectra. Compound **12** protonated molecular ion at  $m/z$  677.2597 indicated the formula  $C_{37}H_{41}O_{12}^+$  (calcd. 677.2593). The  $MS^2$  fragmentation pattern showed the successive losses of rhodnose/amicetose and aculose. The masses of the fragments generated by cleavage of the aglycone were all shifted by 2 Da relative to those of compound **11a**, indicating that the two additional hydrogens were in the center of the aglycone. Thus, it could be deduced that **12** was likely the hydroquinone form of **11a**. Compound **13** is a minor metabolite and could not be annotated with reasonable confidence. The molecular formula of compound **14** was  $C_{49}H_{56}O_{17}$  (exp.  $m/z$  917.3601 ( $[M+H]^+$ ), calcd. 917.3590). The  $MS^2$  spectrum indicated a loss of an aculosyl-rhodnose (or aculosyl-amicetose) moiety. Then the product at  $m/z$  675.2443 lost the neutral group  $C_6H_8O$  (a dehydro-rhodnose), indicating that the aculosyl-rhodnose moiety was linked to another rhodnose. Then the aglycone ion at  $m/z$  451.1398 was produced by the loss of another aculose. The molecular weight of this

aglycone ion (fridamycin A–2  $H_2O$ ) along with the proximity of other lactonic aglycones in the MN spoke in favor of **14** also being a lactone derivative, as reported in Table 1. The fragmentation spectrum of compound **15** was not very clear and **15** could not be annotated with enough confidence. At  $m/z$  901.3653, compound **16** protonated molecular ion indicated the formula  $C_{49}H_{57}O_{16}^+$  (calcd. 901.3641). In  $MS^2$ , the ion  $16+H^+$  lost an aculosyl-rhodnose group to give the product at  $m/z$  659.2499, which then lost dehydro-rhodnose. Further dehydration produced an ion at  $m/z$  545.1790 which again lost aculose to yield the aglycone ion at  $m/z$  435.1471. This aglycone was doubly dehydrated compared to fridamycin A, indicating that the C-olivoyl group in **16** may be dehydrated. The structure proposed in Table 1 appeared to be a reasonable hypothesis for **16**. The molecular formula of compound **17** is  $C_{49}H_{58}O_{18}$  (exp.  $m/z$  952.3960 for  $[M+NH_4]^+$ , calcd. 952.3961). This molecular formula and sodium adduct fragmentation pattern, in which two successive aculosyl-rhodnose losses were recorded, were compatible with the annotation of **17** as vineomycin B2, previously reported to have antibacterial activity (Omura et al., 1977; Imamura et al., 1982).

In the third part of the fridamycin MN, compounds **18a** and **18b** were annotated as isomers with the molecular formula  $C_{31}H_{34}O_{12}$  (exp.  $m/z$  599.2124 for  $[M+H]^+$ , calcd. 599.2123). Both isomers fragmented extensively in the ESI source to produce the fridamycin A protonated molecular ion  $[1+H]^+$  losing  $C_6H_8O_2$ , i.e., a dehydro-cinerulose A moiety. The cinerulose A oxonium ion was also detected in  $MS^2$ , while the aglycone fragmentation was very similar to what was recorded for  $[1+H]^+$ . Angucyclins **18a** and **18b** were therefore annotated as cinerulosyl-fridamycin A or B. The position of the cinerulosyl side chain was not determined and may not be identical for both isomers. Compounds **19a** and **19b** were annotated as isomers with the molecular formula



$C_{31}H_{36}O_{12}$  (exp.  $m/z$  601.2284 for  $[M+H]^+$ , calcd. 601.2279). Both isomers also fragmented extensively in the ESI source to produce the fridamycin A protonated molecular ion  $[1+H]^+$  losing  $C_6H_{10}O_2$ , i.e., dehydro-rhodinose/amicetose subunit, the corresponding oxonium of which was also present in  $MS^2$ . As mentioned above, the  $MS^2$  spectrum of the aglycone protonated ion formed by in-source fragmentation was identical to the one of  $[1+H]^+$ , therefore confirming that **19a** or **b** should not be annotated as grincamycin L (Bao et al., 2018). Instead,

**19a** and **19b** were annotated as rhodinosyl-and/or amicetosyl-fridamycin A/B and should be considered new in the field. Angucycline **19a** was one of the major constituents in the profile of the strain. Compound **20** molecular formula was  $C_{31}H_{34}O_{11}$  (exp.  $m/z$  583.2170 for  $[M+H]^+$ , calcd. 583.2174). Its  $MS^2$  spectrum showed a rhodinosyl/amicetosyl subunit and a dehydrated protonated aglycone. Therefore, compound **20** was annotated as a dehydro-**19a**. Compound **21** molecular formula was  $C_{37}H_{45}O_{13}$  (exp.  $m/z$  697.2861 for  $[M+H]^+$ , calcd.

697.2855). Its MS<sup>2</sup> spectrum revealed the successive loss of two rhodnose/amicetose subunits, yielding a dehydrated protonated aglycone. It was thus annotated as shown in **Table 1**. Compound **22** molecular formula was C<sub>31</sub>H<sub>37</sub>O<sub>11</sub> (exp. *m/z* 585.2333 for [M+H]<sup>+</sup>, calcd. 585.2330). Its MS<sup>2</sup> spectrum showed the loss of one rhodnose/amicetose subunit, yielding a protonated deoxy-aglycone. The structure of the aglycone could not be readily inferred from the MS<sup>2</sup> spectrum and compound **22** could not be annotated further. Compound **23**'s molecular formula was C<sub>37</sub>H<sub>44</sub>O<sub>14</sub> (exp. *m/z* 713.2782 for [M+H]<sup>+</sup>, calcd. 713.2804). A rhodnose/amicetose was visible in MS, but the MS<sup>2</sup> spectrum was impure and further annotation was not possible. Compound **24** molecular formula was C<sub>34</sub>H<sub>41</sub>NO<sub>13</sub> (exp. *m/z* 672.2655 for [M+H]<sup>+</sup>, calcd. 672.2652). Its MS<sup>2</sup> spectrum showed the loss of 1 rhodnose/amicetose subunit, yielding a product at *m/z* 558.1983 (C<sub>28</sub>H<sub>32</sub>NO<sub>11</sub><sup>+</sup>), which then successively lost water, CH<sub>2</sub>O<sub>2</sub> (formic acid or H<sub>2</sub>O+CO), and C<sub>2</sub>H<sub>5</sub>N to generate a protonated didehydro-fridamycin A at *m/z* 451.1393. This fragmentation pattern was compatible with **24** being an alanine amide of **19**, as shown in **Table 1**. The presence of an alanine subunit was also supported by the fragment ion at *m/z* 90.0550, corresponding to an alaninium ion (Tan et al., 2010). Compound **25** molecular formula was C<sub>37</sub>H<sub>46</sub>N<sub>2</sub>O<sub>14</sub> (exp. *m/z* 743.3026 for [M+H]<sup>+</sup>, calcd. 743.3022). In-source fragmentation indicated the successive loss of a dehydro-rhodnose/amicetose group and a C<sub>6</sub>H<sub>10</sub>N<sub>2</sub>O<sub>2</sub> subunit. The MS<sup>2</sup> spectrum highlighted the loss of one dehydro-rhodnose/amicetose and two alanine subunits to generate the protonated didehydro-fridamycin A at *m/z* 451.1393 (see **Supplementary Figure S65**). Overall, compound **25** could be annotated with high confidence as shown in **Table 1**. Neighboring minor compound **26** in the MN could not be annotated. Compound **27** was an analog of **25** with one rhodnose/amicetose subunit more, the presence of which could be ascertained by examination of both in-source fragmentation scheme and the collision-induced MS<sup>2</sup> spectrum. Compound **28** molecular formula was C<sub>37</sub>H<sub>40</sub>O<sub>14</sub> (exp. *m/z* 709.2490 for [M+H]<sup>+</sup>, calcd. 709.2491). The MS<sup>2</sup> spectrum clearly indicated that the protonated molecular ion lost both dehydro-rhodnose/amicetose and a C<sub>6</sub>H<sub>8</sub>O<sub>4</sub> neutral fragment, therefore confirming that this compound was new and should not be annotated as saprolmycin B (Nakagawa et al., 2012). However, the annotation remained ambiguous and **28** were not annotated further. Compounds **29** and **30** were analogs of **28**; **29** did not have the rhodnose/amicetose subunit, while **30** had an aculose in place of the rhodnose/amicetose. All the annotated compounds' MS<sup>2</sup> spectra were provided in the supporting information (**Supplementary Figures S1–S75**), and a list of sugars potentially linked to the angucyclins is provided in **Supplementary Table S1**.

## Secondary Metabolite Biosynthetic Gene Clusters of *Streptomyces* sp. RO-S4

The genome of the RO-S4 strain was analyzed by the antibiotics and Secondary Metabolite shell (antiSMASH) to determine its putative biosynthetic capabilities. A total of 19 putative biosynthetic gene clusters (BGCs) were annotated (**Supplementary Table S2**), including three

types of polyketide synthases [Type 1 (T1PKS), Type 2 (T2PKS), and Type 3 (T3PKS)] polyketide synthases, class I Lanthipeptides, a Lasso peptide, a Ribosomally Synthesized and Post-Translationally Modified Peptide (RiPP), an Ectoine, a Terpene, Phenazines, and a Butyrolactone BGC. In addition, four hybrid clusters were recovered that were composed of 1). T2PKS, Oligosaccharide Phenazine, Siderophore, 2). RiPP-like, betalactam, Terpene, 3). Two hybrid Non-ribosomal Peptide Synthase (NRPS), and T1PKS. The analysis showed that 17 out of the 19 identified BGCs showed high content similarity with known BGCs, five of which (3, 5, 7, 13, and 17) showed 100% content similarity with known BGCs. Two clusters (Cluster 10 and 16) were annotated as orphan BGCs for which no homologous gene clusters could be identified, suggesting that they could be responsible for the biosynthesis of novel natural products or natural products with no characterized BGCs. Many of these clusters are known to encode genes linked to the production of biologically active natural compounds, such as antibiotics. Notably, we have recovered a T2PKS BGC very similar to those linked to the biosynthesis of angucycline compounds, consistent with the metabolomic analysis.

## Description of the Angucycline Biosynthetic Gene Cluster

AntiSMASH analysis using an unannotated genomic DNA sequence revealed that Cluster 2 in Contig RO-S4\_2 of the assembly displayed a high synteny (> 65% common genes) to that of BGCs linked to several angucycline compounds, such as grincamycin (97%), saprolmycin E (83%), saquayamycin A (75%), landomycin A (71%), and saquayamycin Z (67%). All these compounds share a common tetracyclic angular benz[a]anthraquinone aglycone. Due to the predominance of tricyclic aglycones ("open" aglycone(s) hereafter) among the major metabolites of the RO-S4 strain, we performed an in-depth analysis of Cluster 2.

Cluster 2 contains genes putatively involved in the biosynthesis and modification of the aglycone core. A typical set of genes responsible for an angucycline core assembly named "minimal PKS" has been identified, supporting the synthesis of angucycline-like molecules by its cluster. This included three genes: a ketoacyl synthase  $\alpha$  (LRR80\_00487), a ketoacyl synthase  $\beta$ /chain length factor CLF (LRR80\_00488), and an acyl carrier protein (ACP) (LRR80\_004989). Two possible cyclase genes (LRR80\_00486 and LRR80\_00491) were also annotated, which are likely responsible for the polyketide chain cyclization into the benz[a]anthracene structure. In addition, this cluster harbors two genes encoding oxygenase enzymes (LRR80\_00485 and LRR80\_00492) probably involved in the modification of the aglycone and possibly in the lactonization and opening of the angular aglycone cycle (see discussion below). (Keto)reductase-coding genes, including LRR80\_00490 and LRR80\_00470, were annotated to exhibit a high degree of sequence similarity to known enzymes involved in the modification of aromatic polyketides. Three genes (LRR80\_00495, LRR80\_00496, and LRR80\_00498) likely associated with the glycosylation steps showed high similarity to genes coding for glycosyltransferases

(GTs) in other angucyclines. All the annotated genes involved in the BGC of Cluster 2 and their homologs are listed in **Supplementary Table S3**.

The closest BGCs to RO-S4 Cluster-2 are those of the grincamycin-producing *Streptomyces lusitanus* SCSIO LR32 (*Gcn* LR32) (Zhang et al., 2013) and *Streptomyces* sp. CZN-748 (*Gcn* CZN-748) (Shang et al., 2021) graciously provided by the authors. We performed a synteny analysis comparing the three BGCs and notably, when PROKKA (Seemann, 2014)-annotations were used for RO-S4 and CZN-748, the *GcnM* ORF was absent in these strains, as previously described by Shang and co-workers (Shang et al., 2021). In contrast, the BGC annotated by AntiSMASH from genomic sequences identified these ORFs. This difference could be explained by a possible tRNA<sub>ala</sub> in the region coding the *GcnM* orthologs (**Supplementary Figure S76**). In addition, the comparison between the BGC of LR32 identified two missing genes (*GcnU* and *GcnT*) in the BGC of RO\_S4\_2, whereas the region downstream of *GcnS8* was not present in the available BGC of strain CZN-748 (**Figure 5**). Cluster 2 of RO-S4 showed near complete synteny and a higher average amino acid identity (99.8% for 28 common ORFs) to the grincamycin BGC of CZN-748, which fits the observation that both strains produce a majority of “open” aglycone angucyclines, whereas LR32 (93.2% average amino acid identity for 28 common ORFs) produces primarily tetracyclic angucyclines, as previously noted by Shang et al. (2021).

## DISCUSSION

Here we report the use of a combined genomic-metabolomic approach to investigate the antagonistic potential of the *Streptomyces* sp. RO-S4 strain isolated from a polluted marine environment. Based on 16S rRNA gene sequencing and genomic analysis, the strain belongs to the genus *Streptomyces*, likely meriting to be considered the type strain of a new species. RO-S4 extracts show inhibitory activity against the methicillin-resistant *Staphylococcus aureus* ATCC 43300 strain with a MIC of 16 μg/ml.

Metabolomic analyses of the crude extract produced by the RO-S4 strain using mass-spectrometry-based molecular networking revealed diverse angucycline derivatives as dominant products, which have mostly (but not exclusively) been linked to the *Streptomyces* genus (Hu et al., 2016; Peng et al., 2018). Angucyclines represent the largest group of Type 2 PKS natural products produced by actinobacteria, and they show diverse pharmacological activities including cytotoxicity, antitumor, antibacterial, and antiviral properties (Rohr and Thiericke, 1992; Kharel et al., 2012). Two of the predominant ions (fridamycin A and D) have previously shown anti-MRSA activity (Sabido et al., 2020), albeit with a higher MIC than raw extracts of *Streptomyces* sp. RO-S4, but also cytotoxic activity (e.g., Bao et al., 2018), like many other angucyclines. Combined with the lower ion integrated peak sizes of the novel molecules, these factors guided our choice of not further purifying and testing novel molecules in this study. *In silico* analysis or synthesis will probably be

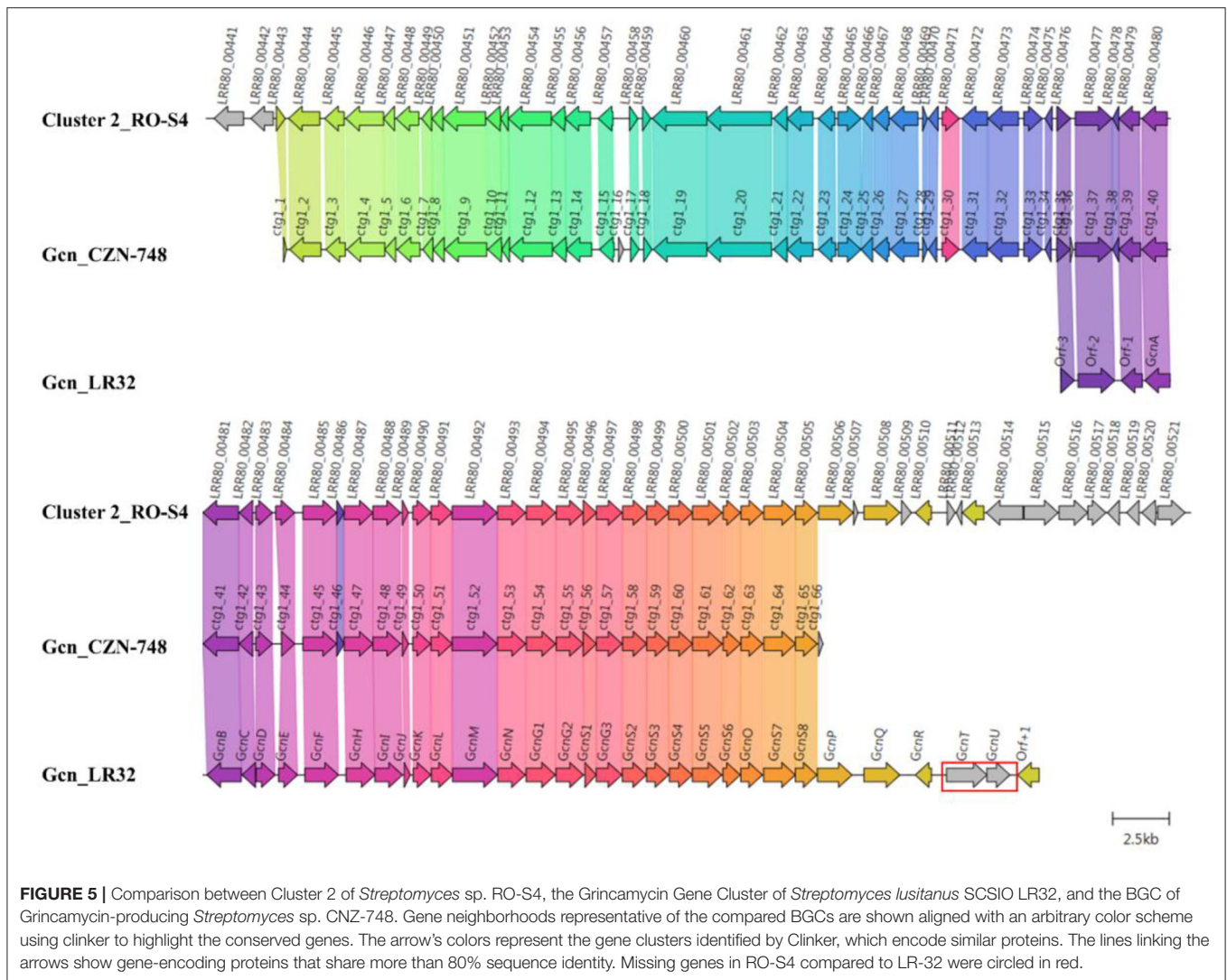
better strategies to evaluate the biotechnological potential of these molecules in the future.

We have demonstrated that many of the compounds identified by our untargeted metabolomic analysis are novel to Science, and this high diversity of novel molecules can be explained by the ability of HRMS to highlight minor compounds even though some of the major metabolites are also new to Science. Our annotation of the RO-S4 angucycline-like compounds was further supported by genomic analysis.

The combination of metabolomic and genome analysis provides some interesting insights into the biosynthesis of angucyclines. Cluster 2 shows high content, synteny, and sequence similarity to previously described BGCs of grincamycin-like products and is certainly responsible for producing the structures predicted by the MN analysis. Since the RO-S4 strain and *Streptomyces* CNZ-748 produce primarily tricyclic glycosylated structures, while LR32 and many other strains produce tetracyclic angucyclines, we were interested in possible enzymatic processes that could lead to the biosynthesis of these tricyclic aglycons (Shang et al., 2021). Several studies have indicated that oxygenase complexes are required for cyclic C-C bond cleavage, in particular Baeyer-Villiger type oxygenases (Fürst et al., 2019). In the case of RO-S4 products, we hypothesized that this reaction would take place by oxidation of the C12b–C1 single bond of the aglycone (an UWM6-like molecule), prior to or after glycosylation, and subsequent hydrolysis of the lactone (**Figure 6**). We have therefore focused the analysis on putative Baeyer-Villiger mono-oxidases (BVMOs) in Cluster 2 of the RO-S4 strain.

The observation that—as in the case of the grincamycin BGC of strain CZN-748—annotation with PROKKA (Seemann, 2014) failed to identify an ORF downstream of the T2PKS synthases and the cyclase putatively involved in the generation of the angular cycle [the region coding for *GcnM* in *S. lusitanus* LR32 (Zhang et al., 2013)], suggested that this ORF could be involved in the ring opening process. Since AntiSMASH identified an ORF both in RO-S4 (LRR80\_00492) and CNZ-748 (CTG1-52) when a genomic sequence was used as the query, we determined that PROKKA failed to annotate the ORF since its Aragorn (Laslett and Canback, 2004) step identified a putative tRNA<sub>ala</sub> (**Supplementary Figure S76**) in its complementary strand. However, the facts that the putative tRNA is in the reverse strand, that it contains mismatches in the side hairpins, and that four more canonical tRNA<sub>ala</sub> are coded in the genome, suggest that this tRNA<sub>ala</sub> could have been misidentified. We performed an RNA fold analysis (Gruber et al., 2008) that identified a highly probable and low-entropy hairpin loop that could potentially affect the translation of this ORF (**Supplementary Figure S77**) and possibly decreases the production of the coded protein.

The LRR80\_00492 ORF codes a hybrid FAD-dependent oxidase-reductase (*GcnM* in the grincamycin BGC) homologous to *UrdM*, that has been linked to C12b hydroxylation in the biosynthesis of urdamycins by *S. fradiae* TŪ 2717 (Faust et al., 2000; Rix et al., 2003). Furthermore, a mutant with an in-frame deletion of the reductase domain of *UrdM* produced small amounts of urdamycin-L, a product containing oxygen



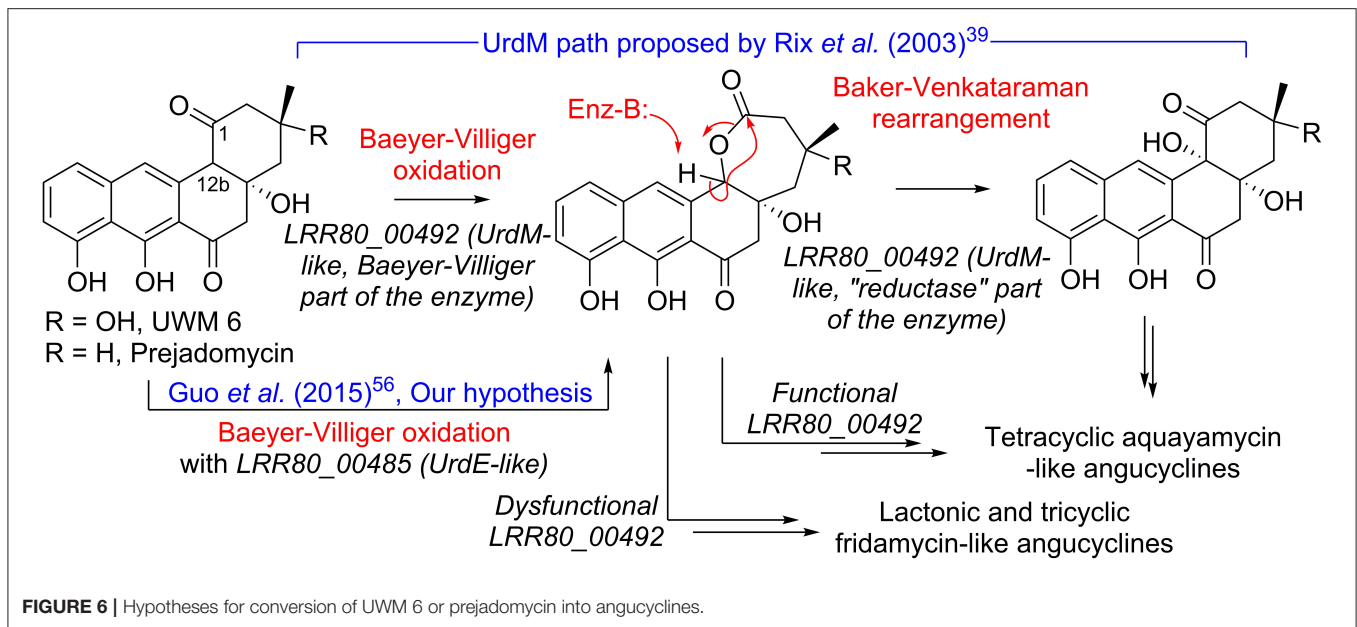
**FIGURE 5** | Comparison between Cluster 2 of *Streptomyces* sp. RO-S4, the Grincamycin Gene Cluster of *Streptomyces lusitanus* SCSIO LR32, and the BGC of Grincamycin-producing *Streptomyces* sp. CNZ-748. Gene neighborhoods representative of the compared BGCs are shown aligned with an arbitrary color scheme using Clinker to highlight the conserved genes. The arrow's colors represent the gene clusters identified by Clinker, which encode similar proteins. The lines linking the arrows show gene-encoding proteins that share more than 80% sequence identity. Missing genes in RO-S4 compared to LR-32 were circled in red.

between C12b and C1, leading to the hypothesis that *UrdM* is involved in the C12b–C1 bond oxidation and subsequent lactone Baker-Venkatarman rearrangement leading to the tetracyclic skeleton of aquayamycin-like angucyclines (**Figure 6**) (Rix et al., 2003). However, in this model, low levels or absence of LRR80\_00492 due to the secondary structure described above would not lead to fridamycin-like aglycones as those in RO-S4 (**Table 1**), and another BVMO should be responsible for the oxidation of the C12b–C1 bond in UWM 6 (or other intermediates) leading to compounds **11**, **12** and **16** and fridamycin-like aglycones.

The BGC of RO-S4 and of all grincamycin-producing strains contain a second FAD-dependent putative BVMO product of LRR80\_00485 (*GcnE* in the grincamycin BGC) that is homologous to FAD-dependent monooxygenases involved in angucycline modifications (e.g., *UrdE*, *PgaE*, *BexE*, *CabE*; **Supplementary Figures S78, S79**). In earlier studies, *UrdE* was hypothesized to directly hydroxylate different positions of the aglycon (C6, C12, C12b) (Decker and Haag, 1995; Faust et al.,

2000; Rix et al., 2002) in urdamycin biosynthesis, but more recent evidence have suggested that its homolog *PgaE* might also oxidize the C12b–C1 bond of prejadomycin leading to the tricyclic aglycones of gaudimycins D and E (Guo et al., 2015). *In vitro* assays using enzymes heterologously expressed in *E. coli* also showed that *PgaE/CabE* oxidizes UWM6 and is dependent on the *PgaM<sub>red</sub>* homolog *CabV* to complete the hydroxylation of UWM6 at C12b (Kallio et al., 2011).

Since in the RO-S4 BGC there are two possible FAD-dependent mono-oxidases that could be involved in the oxidation of the C12b–C1 bond and subsequent ring opening of RO-S4 and CNZ-748, we attempted to compare the sequences of both ORFs to different enzymes oxidizing analogous cyclic compounds, including *MtmOIV*, shown to perform Baeyer–Villiger oxidation and ring opening of premithramycin B to mithramycin. *Blastp* analyses indicated that the LRR80\_00492 (*UrdM* like)/*MtmOIV* alignment was shorter and had a lower overall score, but a higher number of identical positions, whereas the LRR80\_00485 (*UrdE*-like)/*MtmOIV* alignment was longer and had a higher



total score, but with fewer identical positions. Phylogenetic analyses including *UrdE* homologs and the oxidase portion of *UrdM* homologs separated these oxidases into two groups, with maximum likelihood and distance methods showing that LRR80\_00485 was slightly closer to *MtmOIV* than LRR80\_00492, and in a subclade including *PgaE* (Supplementary Figures S78, S79).

In aggregate, these results lead to different possibilities that could be tested in the future using genetic modifications of the different ORFs in the angucycline BGC of the RO-S4 strain or CNZ-748. Hypothesis (1): the product of LRR80\_00492 would be solely responsible for the oxidation of the C12b–C1 bond. This hypothesis is supported by the prediction that this enzyme has several AAs unique to RO-S4 and CNZ-748. On the other hand, since the reductase portion of the enzyme is present, one would expect that tetracyclic angucyclines with a hydroxylated C12b would be produced. Hypothesis (2): the most likely hypothesis based on our work on the RO-S4 strain is that the LRR80\_00492 ORF is inactive due to the presence of a tRNA<sub>ala</sub> in the coding region or that its translation is affected by secondary structure, in which case LRR80\_00485 would generate the lactone via Baeyer-Villiger oxidation, possibly allowing for a later opening of the ring. Other alternative hypotheses that could be related to the ring opening are Hypothesis (3): that the LRR80\_00492 ORF would be partly transcribed due to the mRNA secondary structure, which would allow its oxidase portion to be transcribed but not the reductase, much like the case with the *UrdM* partial knockout mutant that produces urdamycin L (Rix et al., 2003) and is responsible for C12b–C1 single bond oxidation and Hypothesis (4): another BVMO enzyme not in the BGC could be responsible for ring opening. The genome of RO-S4 codes for another enzyme with a slightly higher *blastp* score when queried with *MtmOIV*. That ORF is present in a hybrid NRPS-T1PKS hybrid BGC similar to that of polyoxypeptin A BGC (Du et al.,

2014). However, the function of that homolog (ORF4) in the polyoxypeptin A BGC has not yet been described.

In addition to these oxygenases, three genes are presumed to code for glycosyltransferase (GT) enzymes, which show similarities to known GTs involved in angucycline biosynthesis. The first, RO-LRR80\_00494, shows high homology to *GcnG1* [96.98% (Zhang et al., 2013)], *sqnG1* [88.14% (Salem et al., 2017)], *sprGT1* [86.74% (Kawasaki et al., 2016)], and *SchS10* [81.63% (Basnet et al., 2006)]. The second (LRR80\_00495) is closely related to *GcnG2* [92.46% (Zhang et al., 2013)], *sprGT2* [83.92% (Kawasaki et al., 2016)], *sqnG2* [81.91% (Salem et al., 2017)], and *schS9* [74.74% (Basnet et al., 2006)]. The third GT (LRR80\_00497) is most similar to enzymes involved in the glycosylation of D-olivose at the C9 position of several angucycline-like molecules [ex. *SchS7* in the Sch-47554 biosynthetic gene cluster (Basnet et al., 2006); *SprGT3* in saprolmycin biosynthesis (Kawasaki et al., 2016); *UrdGT2* in urdamycin biosynthesis in *Streptomyces fradiae* Tü 2717 (Mittler et al., 2007), among others], supporting well the predicted structures that are shown in Table 1, all of which are predicted to have been glycosylated with an olivose at the C9-position. *SchS9* and *SchS10* are thought to be O-glycosyltransferases involved in the biosynthesis of Sch-47554 (Basnet et al., 2006). Genetic studies using heterologous expression and targeted gene disruption have shown that *SchS7* attaches D-amictose at C-9 and *SchS9* further extends the saccharide chain, while *SchS10* attaches L-aculose at the C-3 position (Fidan et al., 2018). The *SqnGT1-G3* are glycotransferases involved in the biosynthesis of saquayamycin A in *Streptomyces* sp. KY40-1. According to genetic experimentation, *sqnG2* was identified as catalyzing both O- and C-glycosylations (Salem et al., 2017).

Hence, based on similarities between known GTs and the sugars annotated by metabolomic analysis, we hypothesize that a D-olivose is added to C9 by LRR80\_00498 and

further O-glycosylated with rhodnose/amicetose by one of the three glycosylases, and in some cases (ex. compound 9), experiencing a Michael addition, or further glycosylation as in compounds 19 and 27 (and in previously described angucyclines). Interestingly, all compounds glycosylated at position C3 in RO-S4 appear to have a rhodnose/amicetose as the sugar moiety, as do all grincamycins. It is worth noting that the composition and varying lengths of the oligosaccharide chains in angucyclines have a considerable impact on their biological potential, as previously reported by Elshahawi et al. (2015).

In addition to the rare lactonized D cycle, we also detected other modifications that, to the best of our knowledge, have so far not been shown for fridamycin-like molecules, including alanyl amidation(s) (compounds 25–28). Amide modifications have been previously shown in fridamycin G (Myronovsky et al., 2016), and fridamycin I produced by *Actinokineospora spheciospongiae* (Tawfike et al., 2019). Fridamycin G contains an ethanolamine moiety and was produced heterologously, and the amidation was attributed to a process linked to the host since the BGC source (*S. cyanogenus* S136) does not produce it. Fridamycin I contains a benzylamine moiety, the biosynthetic origin of which was not discussed. As compounds 25–28 contain multiple alanine moieties, we speculate that fridamycin-like molecules could undergo peptide-like elongation with aminoacids (alanine or L-p-hydroxyphenylglycine in the case of fridamycin I), perhaps via an NRPS in a manner analogous to what has been described in the biosynthesis of actinomycin-D (Keller et al., 2010). The presence of an *hpgT* homolog—the enzyme linked to L-p-hydroxyphenylglycine production in actinobacteria (Hubbard et al., 2000)—in the *Actinokineospora spheciospongiae*, supports this hypothesis.

## CONCLUSION

Our combination of untargeted UPLC-HRMS/MS metabolomic, molecular networking, and genomic analyses generated many structural and biosynthetic hypotheses for targeted structural determination and genetic manipulation, possibly using recently developed gene editing approaches (e.g., Cobb et al., 2015; Tong et al., 2020). This approach does not substitute traditional isolation-NMR structural analyses and genetic manipulations, which ultimately will be needed to confirm these hypotheses. However, as it yields structural information of many minor compounds linked to BGCs in a relatively fast manner, it

can streamline and accelerate pipelines of discovery of new drugs, biosynthetic pathways, and enzymes, and hopefully inspire the discovery of novel antibiotics effective against multi-resistant microorganisms.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: National Center for Biotechnology (NCBI) GenBank, <https://www.ncbi.nlm.nih.gov/genbank/>, JAJQKZ000000000, NCBI Sequence Read Archive (SRA), <https://www.ncbi.nlm.nih.gov/sra/>, SRR17084181 and GitHub, [https://github.com/suzumar/ROS4\\_manus](https://github.com/suzumar/ROS4_manus).

## AUTHOR CONTRIBUTIONS

RO, DS, and MS conceived and designed the experiments, wrote the original draft, and prepared tables and figures. RO and MK collected biological material. RO performed the experiments and wrote the manuscript. AR performed the UHPLC-HRMS analysis. CV performed the DNA extraction. RO, DS, and JS performed molecular networking and metabolomics analysis. RO and MS performed the genomic analysis. MS and DS made critical revisions to the final version. All authors read and approved the final version.

## FUNDING

This work was funded by recurrent funds of the CNRS and Sorbonne University attributed to the LBBM laboratory.

## ACKNOWLEDGMENTS

We thank the Bio2Mar platform (<http://bio2mar.obs-banyuls.fr>) for providing technical support and access to instrumentation. This work benefited from access to the Observatoire Océanologique de Banyuls-sur-Mer, an EMBRC-France and EMBRC-ERIC site.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.906161/full#supplementary-material>

## REFERENCES

- Alvi, A. A., Baker, D. D., Stienecker, V., Hosken, M., and Nair, B. G. (2000). Identification of inhibitors of inducible nitric oxide synthase from microbial extracts. *J. Antibiot.* 53, 496–501. doi: 10.7164/antibiotics.53.496
- Antony-Babu, S., Stien, D., Eparvier, V., Parrot, D., Tomasi, S., and Suzuki, M. T. (2017). Multiple *Streptomyces* species with distinct secondary metabolomes have identical 16S rRNA gene sequences. *Sci. Rep.* 7, 11089. doi: 10.1038/s41598-017-11363-1
- Baltz, R. (2008). Renaissance in antibacterial discovery from actinomycetes. *Curr Opin Pharmacol* 8, 557–563. doi: 10.1016/j.coph.2008.04.008
- Bao, J., He, F., Li, Y., Fang, L., Wang, K., Song, J., et al. (2018). Cytotoxic antibiotic angucyclines and actinomycins from the *Streptomyces* sp. XZHG99T. *J. Antibiot.* 71, 1018–1024. doi: 10.1038/s41429-018-0096-1
- Barbosa, F., Pinto, E., Kijjoa, A., Pinto, M., and Sousa, E. (2020). Targeting antimicrobial drug resistance with marine natural products. *Int. J. Antimicrob.* 56, 1–29. doi: 10.1016/j.ijantimicag.2020.106005

- Basnet, D. B., Oh, T. J., Vu, T. T. H., Sthapit, B., Liou, K., Lee, H. C., et al. (2006). Angucyclines Sch 47554 and Sch 47555 from *Streptomyces* sp. SCC-2136: cloning, Sequencing, and Characterization. *Mol. Cells* 9, 154–162.
- Bérdy, J. (2005). Bioactive microbial metabolites: a personal view. *J. Antibiot.* 58, 1–26. doi: 10.1038/ja.2005.1
- Bhatnagar, I., and Kim, S. K. (2010). Immense essence of excellence: marine microbial bioactive compounds. *Mar. Drugs* 8, 2673–2701. doi: 10.3390/md8102673
- Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., Van Wezel, G. P., Medema, M. H., et al. (2021). AntiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res.* 49, 29–35. doi: 10.1093/nar/gkab335
- Blunt, J. W., Copp, B. R., Keyzers, R. A., Munro, M. H. G., and Prinsep, M. R. (2012). Marine natural products. *Nat. Prod. Rep.* 30, 144–222. doi: 10.1039/C2NP00090C
- Bose, U., Hewavitharana, A. K., Ng, Y. K., Shaw, P. N., Fuerst, J. A., and Hodson, M. P. (2015). LC-MS-Based metabolomics study of marine bacterial secondary metabolite and antibiotic production in *Salinispora arenicola*. *Mar. Drugs* 13, 249–266. doi: 10.3390/md13010249
- Caesar, L. K., Montaser, R., Keller, N. P., and Kelleher, N. L. (2021). Metabolomics and genomics in natural products research: complementary tools for targeting new chemical entities. *Nat. Prod. Rep.* 38, 2041–2065. doi: 10.1039/D1NP00036E
- Chambers, H. F., and DeLeo, F. R. (2009). Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat. Rev. Microbiol.* 7, 629–641. doi: 10.1038/nrmicro2200
- Chokshi, A., Sifri, Z., Cennimo, D., and Horng, H. (2019). Global contributors to antibiotic resistance. *J. Global. Infect. Dis.* 11, 36–42. doi: 10.4103/jgid.jgid\_110\_18
- Cobb, R. E., Wang, Y., and Zhao, H. (2015). High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/CAS system. *ACS Synth. Biol.* 4, 723–728. doi: 10.1021/sb500351f
- Cornell, R. C., Marasini, D., and Mohamed, F. K. (2018). Molecular characterization of plasmids harbored by actinomycetes isolated from the great salt plains of oklahoma using PFGE and next generation whole genome sequencing. *Front. Microbiol.* 9, 1–13. doi: 10.3389/fmicb.2018.02282
- Decker, H., and Haag, S. (1995). Cloning and characterization of a polyketide synthase gene from *Streptomyces fradiae* Tü2717, which carries the genes for biosynthesis of the angucycline antibiotic urdamycin A and a gene probably involved in its oxygenation. *J. Bacteriol.* 177, 6126–6136. doi: 10.1128/jb.177.21.6126-6136.1995
- Dettmer, K., Aronov, P. A., and Hammock, B. D. (2007). Mass spectrometry-based metabolomics. *Mass. Spectrom. Rev.* 26, 51–78. doi: 10.1002/mas.20108
- Du, Y., Wang, Y., Huang, T., Tao, M., Deng, Z., and Lin, S. (2014). Identification and characterization of the biosynthetic gene cluster of polyoxypeptin A, a potent apoptosis inducer. *BMC Microbiol.* 14, 30. doi: 10.1186/1471-2180-14-30
- Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A. A., Melnik, A. V., Meusel, M., et al. (2019). SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat. Methods* 16, 299–302. doi: 10.1038/s41592-019-0344-8
- Edwards, B., Andini, R., Esposito, S., Grossi, P., Lew, D., Mazzei, T., et al. (2014). Treatment options for methicillin-resistant *Staphylococcus aureus* (MRSA) infection: where are we now? *J. Glob. Antimicrob. Resist.* 2, 133–140. doi: 10.1016/j.jgar.2014.03.009
- Elshahawi, S. I., Shaaban, K. K., Kharel, M. K., and Thorson, J. S. (2015). A comprehensive review of glycosylated bacterial natural products. *Chem Soc. Rev.* 44, 7591–7697. doi: 10.1039/C4CS00426D
- Fagerbold, S. K., Urios, L., Intertaglia, L., Batailler, N., Lebaron, P., and Suzuki, M. T. (2013). *Pleionea mediterranea* gen. nov., sp. nov., a gammaproteobacterium isolated from coastal seawater. *Int. J. Syst. Evol.* 63, 2700–2705. doi: 10.1099/ijs.0.045575-0
- Faust, B., Hoffmeister, D., Weitnauer, G., Westrich, L., Haag, S., Schneider, P., et al. (2000). Two new tailoring enzymes, a glycosyltransferase and an oxygenase, involved in biosynthesis of the angucycline antibiotic urdamycin A in *Streptomyces fradiae* Tü2717. *Microbiology* 146, 147–154. doi: 10.1099/00221287-146-1-147
- Fidan, O., Yan, R., Gladstone, G., Zhou, T., Zhu, D., and Zhan, J., et al. (2018). New insights into the glycosylation steps in the biosynthesis of Sch47554 and Sch47555. *Chem. Bio. Chem.* 19, 1424–1432. doi: 10.1002/cbic.201800105
- Fürst, M. J. L. J., Gran-Scheuch, A., Aalbers, F. S., and Fraaije, M. W. (2019). Baeyer–Villiger monooxygenases: tunable oxidative biocatalysts. *ACS Catal.* 9, 11207–11241. doi: 10.1021/acscatal.9b03396
- Gilchrist, C. L. M., and Chooi, Y. H. (2021). Clinker & clustermap. js: Automatic generation of gene cluster comparison figures. *Bioinformatics* 37, 2473–2475. doi: 10.1093/bioinformatics/btab007
- Gomez-Escribano, J. P., Alt, S., and Bibb, M. (2016). Next generation sequencing of actinobacteria for the discovery of novel natural products. *Mar. Drugs* 14, 78. doi: 10.3390/md14040078
- Gouy, M., Guindon, S., and Gascuel, O. (2010). SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* 27, 221–224. doi: 10.1093/molbev/msp259
- Gruber, A. R., Lorenz, R., Bernhart, S., Neubock, R., and Hofacker, I. L. (2008). The vienna RNA suite. *Nucleic Acids Res.* 36, 70–74. doi: 10.1093/nar/gkn188
- Guo, F., Xiang, S., Li, L., Wang, B., Rajasärkkä, J., Gröndahl-Yli-Hannuksela, K., et al. (2015). Targeted activation of silent natural product biosynthesis pathways by reporter-guided mutant selection. *Metab Eng* 28, 134–142. dx. doi: 10.1016/j.ymben.2014.12.006
- Hassan, S. S., and Shaikh, A. L. (2017). Marine actinobacteria as a drug treasure house. *Biomed. Pharmacother.* 87, 46–57. doi: 10.1016/j.biopha.2016.12.086
- Hirsch, C. F., and Christensen, D. L. (1983). Novel method for selective isolation of Actinomycetes. *Appl. Environ. Microbiol.* 46, 925–929. doi: 10.1128/aem.46.4.925-929.1983
- Hu, D., Sun, C., Jin, T., Fan, G., Mok, K. M., Li, K., et al. (2020). Exploring the potential of antibiotic production from rare actinobacteria by whole-genome sequencing and guided MS/MS analysis. *Front. Microbiol.* 11, 1–12. doi: 10.3389/fmicb.2020.01540
- Hu, Z., Qin, L., Wang, Q., Ding, W., Chen, Z., and Ma, Z. (2016). Angucycline antibiotics and its derivatives from marine-derived actinomycete *Streptomyces* sp. A6H. *Nat. Prod. Res.* 30, 2551–2558. doi: 10.1080/14786419.2015.1120730
- Hubbard, B. K., Thomas, M. G., and Walsh, C. T. (2000). Biosynthesis of Lp-hydroxyphenylglycine, a non-proteinogenic amino acid constituent of peptide antibiotics. *Chem. Biol.* 7, 931–941. doi: 10.1016/S1074-5521(00)00043-0
- Imamura, N., Kakinuma, K., Ikekawa, N., Tanaka, H., and Omura, S. (1982). Biosynthesis of vineomycins A1 and B2. *J. Antibiot.* 35, 602–608. doi: 10.7164/antibiotics.35.602
- Jensen, P. R., Dwight, R., and Fenical, W. (1991). Distribution of actinomycetes in near-shore tropical marine sediments. *Appl. Environ. Microbiol.* 57, 1102–1108. doi: 10.1128/aem.57.4.1102-1108.1991
- Jose, P. A., and Jha, B. (2016). New dimensions of research on actinomycetes: quest for next generation antibiotics. *Front. Microbiol.* 7, 1295. doi: 10.3389/fmicb.2016.01295
- Jose, P. A., Maharshi, A., and Jha, B. (2021). Actinobacteria in natural products research: Progress and prospects. *Microbiol. Res* 246, 126708. doi: 10.1016/j.micres.2021.126708
- Kallio, P., Patrikainen, P., Suomela, J. P., Mäntsälä, P., Metsä-Ketelä, M., and Niemi, J. (2011). Flavoprotein hydroxylase PgaE catalyzes two consecutive oxygen-dependent tailoring reactions in angucycline biosynthesis. *Biochemistry* 50, 5535–5543. dx. doi: 10.1021/bi200600k
- Kawasaki, T., Moriyama, A., Nakagawa, K., and Imamura, N. (2016). Cloning and identification of saprolmycin biosynthetic gene cluster from *Streptomyces* sp. TK08046. *Biosci. Biotechnol. Biochem.* 80, 2144–2150. doi: 10.1080/09168451.2016.1196574
- Keller, U., Lang, M., Crnovcic, I., Pfennig, F., and Schauwecker, F. (2010). The actinomycin biosynthetic gene cluster of *Streptomyces chrysomallus*: a genetic hall of mirrors for synthesis of a molecule with mirror symmetry. *J. Bacteriol.* 192, 2583–2595. doi: 10.1128/JB.01526-09
- Kharel, M. K., Pahari, P., Shepherd, M. D., Tibrewal, N., Nybo, S. E., Shaaban, K. A., et al. (2012). Angucyclines: biosynthesis, mode-of-action, new natural products, and synthesis. *Nat. Prod. Rep.* 29, 264–325. doi: 10.1039/C1NP00068C
- Khoomrung, S., Wanichtharak, K., Nookaew, I., Thamsersang, O., Seubnooch, P., Laohapand, T., et al. (2017). Metabolomics and integrative omics for the development of Thai traditional medicine. *Front. Pharmacol.* 8, 474. doi: 10.3389/fphar.2017.00474
- Lane, A. L., and Moore, B. S. (2011). A sea of biosynthesis: marine natural products meet the molecular age. *Nat. Prod. Rep.* 28, 411–428. doi: 10.1039/C0NP90032J



- Laslett, D., and Canback, B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* 32, 11–16. doi: 10.1093/nar/gkh152
- Lewis, K. (2020). The science of antibiotic discovery. *Cell.* 181, 29–45. doi: 10.1016/j.cell.2020.02.056
- Liu, X., and Locasale, J. W. (2017). Metabolomics: a primer. *Trends Biochem. Sci.* 42, 274–284. doi: 10.1016/j.tibs.2017.01.004
- Macintyre, L., Zhang, T., Viegelmann, C., Martinez, I. J., Cheng, C., Dowdells, C., et al. (2014). Metabolomic tools for secondary metabolite discovery from marine microbial symbionts. *Mar. Drugs.* 12, 3416–3448. doi: 10.3390/md12063416
- Maskey, R. P., Helmke, E., and Laatsch, H. (2003). Himalomycin A and B: isolation and structure elucidation of new fridamycin type antibiotics from a marine *Streptomyces* isolate. *J. Antibiot.* 56, 942–949. doi: 10.7164/antibiotics.56.942
- Meier-Kolthoff, J., and Goker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Communications.* 10. doi: 10.1038/s41467-019-10210-3
- Mittler, M., Bechthold, A., and Schulz, G. E. (2007). Structure and action of the C–C bond-forming glycosyltransferase UrdGT2 involved in the biosynthesis of the antibiotic urdamycin. *J. Mol. Biol.* 372, 67–76. doi: 10.1016/j.jmb.2007.06.005
- Munita, M., and Arias, C. A. (2016). Mechanisms of antibiotic resistance. *Microbiol Spectr.* 4, 1–37. doi: 10.1128/microbiolspec.VMBF-0016-2015
- Myronovskiy, M., Brötz, E., Rosenkränzer, B., Manderscheid, N., Tokovenko, B., Rebets, Y., et al. (2016). Generation of new compounds through unbalanced transcription of landomycin A cluster. *Appl. Microbiol. Biotechnol.* 100, 9175–9186. doi: 10.1007/s00253-016-7721-3
- Nakagawa, K., Hara, C., Tokuyama, S., Takada, K., and Imamura, I. (2012). Saprolomycins A–E, new angucycline antibiotics active against *Saprolegnia parasitica*. *J. Antibiot.* 65, 599–607. doi: 10.1038/ja.2012.86
- Nkanga, E. J., and Hagedorn, C. (1978). Detection of antibiotic-producing *Streptomyces* inhabiting forest soils. *Antimicrob. Agents Chemother.* 14, 51–59. doi: 10.1128/AAC.14.1.51
- Olano, C., Méndez, C., and Salas, A. (2009). Antitumor compounds from marine actinomycetes. *Mar. Drugs.* 7, 210–248. doi: 10.3390/md7020210
- Omura, S., Tanaka, H., Oiwa, R., Awaya, J., Masuma, R., and Tanaka, K. (1977). New antitumor antibiotics, OS-4742 A1, A2, B1 and B2 produced by a strain of *Streptomyces*. *J. Antibiot.* 30, 908–916. doi: 10.7164/antibiotics.30.908
- Otasek, D., Morris, J. H., Bouças, Pico, A. R., and Demchak, D. (2019). Cytoscape automation: empowering workflow-based network analysis. *Genome. Biol.* 20, 1–15. doi: 10.1186/s13059-019-1758-4
- Ouchene, R., Intertaglia, L., Zaatout, N., Kecha, M., and Suzuki, M. T. (2022). Selective isolation, antimicrobial screening, and phylogenetic diversity of marine actinomycetes derived from the Coast of Bejaia City (Algeria), a polluted and microbiologically unexplored environment. *J. Appl. Microbiol.* 132, 2870–2882. doi: 10.1111/jam.15415
- Palazzotto, E., Tong, Y., Lee, S. Y., and Weber, T. (2019). Synthetic biology and metabolic engineering of actinomycetes for natural product discovery. *Biotechnol Adv.* 37, 1–15. doi: 10.1016/j.biotechadv.2019.03.005
- Peng, A., Qu, X., Liu, F., Li, X., Li, E., and Xie, W. (2018). Angucycline glycosides from an intertidal sediment strain *streptomyces* sp. and their cytotoxic activity against hepatoma carcinoma cells. *Mar Drugs.* 16, 1–10. doi: 10.3390/md16120470
- Prestinaci, F., Pezzotti, P., and Pantosti, A. (2015). Antimicrobial resistance: a global multifaceted phenomenon. *Pathog. Glob. Health.* 109, 309–318. doi: 10.1179/204773215Y.0000000030
- Pribelski, A., Antipov, D., Meleshko, D., Lapidus, A., and Korobeynikov, A. (2020). Using SPAdes de novo assembler. *Curr Protoc Bioinform* 70, 1–29. doi: 10.1002/cpbi.102
- Pye, C. R., Bertin, M. J., Lokey, R. S., Gerwick, W. H., and Linington, R. G. (2017). Retrospective analysis of natural products provides insights for future discovery trends. *PNAS* 114, 5601–5606. doi: 10.1073/pnas.1614680114
- Rix, U., Fischer, C., Remsing, L. L., and Rohr, J. (2002). Modification of post-pks tailoring steps through combinatorial biosynthesis. *Nat. Prod. Rep.* 19, 542–580. doi: 10.1039/b103920m
- Rix, U., Remsing, L. L., Hoffmeister, D., Bechthold, A., and Rohr, J. (2003). Urdamycin L: a novel metabolic shunt product that provides evidence for the role of the urdM gene in the urdamycin A biosynthetic pathway of *Streptomyces Fradiae* TŪ 2717. *Chem. Bio. Chem.* 4, 109–911. doi: 10.1002/cbic.200390002
- Rohr, J., and Thiericke, R. (1992). Angucycline group antibiotics. *Nat Prod Rep* 9, 103–137. doi: 10.1039/np9920900103
- Ruttkies, C., Schymanski, E. L., Wolf, W., Hollender, J., and Neumann, S. (2016). MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J. Cheminform.* 8, 1–16. doi: 10.1186/s13321-016-0115-9
- Sabido, E., Tenebro, C., Suarez, A., Ong, S., Trono, D., Amago, D., et al. (2020). Marine sediment-derived *Streptomyces* strain produces angucycline antibiotics against multidrug-resistant *Staphylococcus aureus* harboring SCCmec Type 1 gene. *J. Marine Sci. Eng.* 8, 734. doi: 10.3390/jmse8100734
- Salem, S. M., Weidenbach, S., and Rohr, J. (2017). Two cooperative glycosyltransferases are responsible for the sugar diversity of saquayamycins isolated from *Streptomyces* sp. KY 40-1. *ACS Chem. Biol.* 12, 2529–2534. doi: 10.1021/acscchembio.7b00453
- Schneider, O., Simic, N., Aachmann, F. L., Rückert, C., Kristiansen, K. A., Kalinowski, J., et al. (2018). Genome mining of *Streptomyces* sp. YIM 130001 isolated from lichen affords new thiopeptide antibiotic. *Front. Microbiol.* 9, 1–12. doi: 10.3389/fmicb.2018.03139
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinform.* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Shang, Z., Ferris, Z. E., Sweeney, D., Chase, A. B., Yuan, C., Hui, Y., et al. (2021). Grincamycins P–T: Rearranged Angucyclines from the Marine Sediment-Derived *Streptomyces* sp. CNZ-748 Inhibit Cell Lines of the Rare Cancer Pseudomyxoma Peritonei. *J. Nat. Prod.* 84, 1638–1648. doi: 10.1021/acs.jnatprod.1c00179
- Shomura, T., Yoshida, J., Amano, S., Kojima, M., Inouye, S., and Niida, T. (1979). Studies on Actinomycetales producing antibiotics only on agar culture. I. Screening, taxonomy and morphology-productivity relationship of *Streptomyces halstedii*, strain SF-1993. *J. Antibiot.* 32, 427–435. doi: 10.7164/antibiotics.32.427
- Song, Y., Liu, G., Li, J., Huang, H., Zhang, X., Zhang, H., et al. (2015). Cytotoxic and antibacterial angucycline- and prodigiosin- analogues from the deep-sea derived *Streptomyces* sp. SCSIO 11594. *Mar. Drugs.* 13, 1304–1316. doi: 10.3390/md13031304
- Stien, D. (2020). Marine microbial diversity as a source of bioactive natural products. *Mar. Drugs.* 18, 1–5. doi: 10.3390/md18040215
- Stien, D., Clergeaud, F., Rodrigues, A. M., Lebaron, K., Pillot, R., Romans, P., et al. (2019). Metabolomics reveal that octocrylene accumulates in *pocillopora damicornis* tissues as fatty acid conjugates and triggers coral cell mitochondrial dysfunction. *Anal. Chem.* 91, 990–995. doi: 10.1021/acs.analchem.8b04187
- Stien, D., Suzuki, M., Rodrigues, A., Yvin, M., Clergeaud, F., Thorel, E., et al. (2020). A unique approach to monitor stress in coral exposed to emerging pollutants. *Sci. Rep.* 10, 1–11. doi: 10.1038/s41598-020-66117-3
- Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11 Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* 38, 3022–3027. doi: 10.1093/molbev/msab120
- Tan, B., O'Dell, D. K., Yu, Y. W., Monn, M. F., Burstein, S., and Walker, J. M. (2010). Identification of endogenous acyl amino acids based on a targeted lipidomics approach. *J. Lipid. Res.* 51, 112–219. doi: 10.1194/jlr.M900198-JLR200
- Tawfike, A., Attia, E. Z., Desoukey, S. Y., Hajjar, D., Makki, A. A., Schupp, P. J., et al. (2019). New bioactive metabolites from the elicited marine sponge-derived bacterium *Actinokineospora Sphaciospongiae* sp. Nov. *AMB Expr.* 9, 1–9. doi: 10.1186/s13568-018-0730-0
- Tong, Y., Whitford, C. M., Blin, K., Jørgensen, T. S., Weber, T., and Lee, S. Y. (2020). CRISPR–Cas9, CRISPRi and CRISPR–BEST-mediated genetic manipulation in streptomycetes. *Nat. Protoc.* 15, 2470–2502. doi: 10.1038/s41596-020-0339-z
- van Der Hoof, J. J., Mohimani, H., Bauermeister, A., Dorrestein, P. C., Duncan, K. R., and Medema, M. H., et al. (2020). Linking genomics and metabolomics to chart specialized metabolic diversity. *Chem. Soc. Rev.* 49, 3297–3314. doi: 10.1039/D0CS00162G
- Ventola, C. L. (2015). The antibiotic resistance crisis. *Phar. Therapeut.* 40, 277–283.
- Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., et al. (2016). Sharing and community curation of mass spectrometry data with

- Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* 34, 828–837. doi: 10.1038/nbt.3597
- Wiegand, I., Hilpert, K., and Hancock, R. E. (2008). Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protocols.* 3, 163–175. doi: 10.1038/nprot.2007.521
- Ye, Y., Anwar, N., Mao, X., Wu, S., Yan, C., Zhao, Z., et al. (2020). Discovery of three 22-membered macrolides by deciphering the streamlined genome of mangrove-derived *Streptomyces* sp. HM190. *Front. Microbiol.* 11, 1–13. doi: 10.3389/fmicb.2020.01464
- Yoon, S. H., Ha, S. M., Kwon, S., Lim, J., Kim, Y., Seo, H., et al. (2017). Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol.* 67, 1613–1617. doi: 10.1099/ijsem.0.001755
- Yoon, S. Y., Lee, S. R., Hwang, J. Y., Benndorf, R., Beemelmans, C., Chung, S. J., et al. (2019). Fridamycin A, a microbial natural product, stimulates glucose uptake without inducing adipogenesis. *Nutrients.* 11, 765. doi: 10.3390/nu11040765
- Zhang, Y., Huang, H., Chen, Q., Luo, M., Sun, A., Song, Y., et al. (2013). Identification of the grincamycin gene cluster unveils divergent roles for GcnQ in different hosts, tailoring the L-rhodinose moiety. *Org Lett.* 15, 3254–3257. doi: 10.1021/ol401253p
- Zhu, X., Duan, Y., Cui, Z., Wang, Z., Li, Z., Zhang, Y., et al. (2017). Cytotoxic rearranged angucycline glycosides from deep sea-derived *Streptomyces lusitanus* SCSIO LR32. *J. Antibiot.* 70, 819–822. doi: 10.1038/ja.2017.17

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Ouchene, Stien, Segret, Kecha, Rodrigues, Veckerlé and Suzuki. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.