Discovery of Three 22-Membered Macrolides by Deciphering the Streamlined Genome of Mangrove-Derived Streptomyces sp. HM190

Yanghui Ye1, Nusratgul Anwar2, Xuming Mao2, Shihua Wu2, Cen Yan1, Zhe Zhao2, Ran Zhang2, Yanfang Nie2, Jianwei Zhang2, Jidong Wang4* and Min Wu1*

1 Ocean College, Zhejiang University, Hangzhou, China, 2 College of Life Sciences, Zhejiang University, Hangzhou, China, 3 Cardiovascular Health Department, AstraZeneca Trading Co., Ltd., Wuxi, China, 4 Key Laboratory of Vector Biology and Pathogen Control of Zhejiang Province, College of Life Science, Huzhou University, Huzhou, China

Strain HM190, a moderate halophile, was isolated from rhizosphere soil of the mangrove Kandelia obovata in Fugong village, China. The 16S ribosomal RNA (rRNA) gene sequence and the results of phylogenetic analysis revealed that strain HM190 belonged to the genus Streptomyces and had the highest sequence similarity of 99.79% to Streptomyces heilongjiangensis NEAU-W2T. The complete genome of strain HM190 comprised 7,762,826 bp in a linear chromosome with 71.97% G + C content. According to antiSMASH analysis, a total of 30 biosynthetic gene clusters (BGCs) were predicted to be involved in secondary metabolism, 12 of which were responsible for the production of polyketide- and non-ribosomal peptide-derived secondary metabolites. Gene cluster 5 was responsible for macrolide biosynthesis in a strain-specific 126,331-bp genomic island belonging to the left-arm region. Combined genomics–metabolomics analysis led to the discovery of three 22-membered macrolides (compounds 1–3). Their structures were elucidated by using spectroscopic techniques including high-resolution electrospray ionization mass spectroscopy (HRESIMS) and nuclear magnetic resonance (NMR). The absolute configurations of compounds 1–3 were determined by the X-ray single crystal diffraction and NMR data analysis. All three compounds displayed moderate cytotoxic activities toward tumor cell lines HepG2, A549, and HCT116.

Keywords: Streptomyces sp. HM190, biosynthetic gene clusters, 22-membered macrolides, structure elucidation, cytotoxic activities

INTRODUCTION

Mangrove ecosystems cover about 60–75% of the world’s tropical and subtropical coastlines and grow in saline coastal sediment habitats at transition zones with ocean, fresh water, and land (Holguin et al., 2001). The tidal action in these areas causes large changes in temperature, oxygen, and salinity levels during the day, which leads to the formation of unique microbial communities (Kathiresan and Bingham, 2001). The phylogenetic diversity of mangrove microbial communities...
has been well established based on 16S ribosomal RNA (rRNA) analytical approaches, including PCR cloning (Li et al., 2011), denaturing gradient gel electrophoresis (Tian et al., 2008), and pyrosequencing (Santos et al., 2011). Recently, various mangrove-derived microorganisms have attracted increasing attention in the drug discovery field, especially those of the genus *Streptomyces*, as important sources of biologically active compounds with diverse structures (Xu, 2011; Thatoi et al., 2013; Blunt et al., 2015).

The genus *Streptomyces* belongs to the family *Streptomycetaceae*, the order *Streptomyces*, and the class *Actinobacteria*, and was originally described in Waksman and Henrici (1943). At the time of writing, 854 *Streptomyces* species with validly published names had been proposed according to List of Prokaryotic names with Standing in Nomenclature1. Some of these *Streptomyces* strains have been isolated from a wide range of marine habitats, including mangroves (Yan et al., 2010), marine sediments (Zhao et al., 2009), seawater (Zhu et al., 2011), sponges (Huang et al., 2016), algae (Girão et al., 2019), and corals (Alfredo et al., 2017). *Streptomyces* constitutes an important source of new natural products because its strains contain a large number of biosynthetic gene clusters (BGCs) associated with the production of multiple secondary metabolites (Huang et al., 2014; Liu et al., 2018). Since the first complete genome of *Streptomyces coelicolor* A3(2) was sequenced and reported in Bentley et al. (2002), increasing numbers of *Streptomyces* genomes have been sequenced and deposited in public databases in recent years (Studholme, 2016), which has led to an increase in genomics–metabolomics studies of this genus (Xu et al., 2019). However, most of these genomes were draft genomes. The number of complete genomes of *Streptomyces* has been limited because of their high G + C content, which results in shorter reads and much higher error rates during genome sequencing (Ioanna et al., 2012; Zhong et al., 2013).

During our work focused on discovering microorganism diversity from mangrove sample, Strain HM190, as a novel species of the genus *Streptomyces*, was isolated from rhizosphere soil of the mangrove *Kandelia obovata* of Fugong village (117° 57’ N 24° 24’ E), in Zhangzhou, China, which was placed into clean plastic buckets and stored at 4°C until use. Serially diluted (10-fold dilutions each) samples were made and spread on the five selective isolation media by the traditional dilution-plating method: HV (humic acid-vitamin agar medium), M7 (glicerine-peptone agar medium), GS (Gauze’s modified medium no. 1), ISP2 (ISP medium no. 2), and MA (marine agar 2216 medium). All of the media were supplemented with 25 mg/L of nalidixic acid and 50 mg/L of nystatin. After 7 days of incubation at 30°C, a white-pigmented colony was picked from MA plates and named as HM190. After repeated plate streaking on the same medium, pure strains were obtained from individual colonies and preserved at −80°C as suspension with 25% (v/v) glycerol for further use (Ye et al., 2019).

For phylogenetic studies, the isolate was grown in MB (marine broth 2216 medium) for 7 days at 30°C. The 16S rRNA gene sequence was amplified by using universal primers 27F [5’-AGAGTTTGATCCCTGCTAG-3’) and 1492R [5’-ACGGTACCTTGTAGCACT-3’) (Anzai et al., 2000). Amplification reactions were prepared in a 25 µL final reaction volume containing 12.5 µL of PCR SuperMix, 11 µL of distilled water, 0.5 µL of each primer, and 0.5 µL of extracted DNA template. PCR was performed under the following conditions: 30 cycles of 94°C/5 min, 40°C/30 s, 55°C/30 s, 72°C/75 s and a final extension of 72°C/10 min. PCR products were ligated to vector pMD 19-T (TaKaRa) and cloned into *Escherichia coli* DH5a for sequencing, and the almost-complete sequence of the 16S rRNA gene sequence was obtained. The obtained sequence was assembled with DNASTAR SeqMan (LaserGene, Madison, WI, United States). Further, the 16S rRNA gene

---

1http://www.bacterio.net/streptomyces.html

### MATERIALS AND METHODS

#### General Experimental Procedures

Silica gel plates HSGF254 (Yantai Chemical Industry Research Institute, Yantai, China) were used for thin-layer chromatography (TLC). Column chromatography was performed with commercial silica gel (QingDao HaiYang Chemical Group Co., QingDao, China, 200-300 mesh). High-performance liquid chromatography (HPLC) separation was performed on semipreparative HPLC (Agilent 1100, Zorbax SB-C18, 5 µm, 250 × 9.4 mm inner diameter; 1.5 mL/min; 210 nm; Agilent, Palo Alto, CA, United States). NMR spectra were measured with a Bruker DRX-400 spectrometer (Bruker, Rheinstetten, Germany). The HRESIMS spectra were taken on a Q-TOF Micro LC-MS-MS mass spectrometer (Waters Co., Milford, MA, United States). Optical rotation was measured on a Perkin-Elmer 341 polarimeter (Perkin-Elmer, Suzhou, China). IR spectra were recorded on a Nicolet Magna FT-IR 750 spectrometer (Nicolet, Tokyo, Japan). UV spectra were recorded on a Varian CARY 300 BIO spectrophotometer (Varian, Cary, NC, United States).

#### Isolation and Identification of Strain HM190

Strain HM190 was isolated from a soil sample collected from rhizosphere soil of the mangrove *K. obovata* of Fugong village (117°57’ N 24°24’ E), in Zhangzhou, China, which was placed into clean plastic buckets and stored at 4°C until use. Serially diluted (10-fold dilutions each) samples were made and spread on the five selective isolation media by the traditional dilution-plating method: HV (humic acid-vitamin agar medium), M7 (glicerine-peptone agar medium), GS (Gauze’s modified medium no. 1), ISP2 (ISP medium no. 2), and MA (marine agar 2216 medium). All of the media were supplemented with 25 mg/L of nalidixic acid and 50 mg/L of nystatin. After 7 days of incubation at 30°C, a white-pigmented colony was picked from MA plates and named as HM190. After repeated plate streaking on the same medium, pure strains were obtained from individual colonies and preserved at −80°C as suspension with 25% (v/v) glycerol for further use (Ye et al., 2019).
sequence was analyzed by submitting to the NCBI and the EzTaxon-e server (Altschul et al., 1997; Yoon et al., 2017). The multiple sequences were aligned with Clustal W (Thompson et al., 1994). Phylogenetic trees were constructed with the MEGA 5.0 software package (Molecular Evolutionary Genetics Analysis, version 5.0) (Tamura et al., 2011) by using neighbor-joining (Saitou and Nei, 1987), minimum-evolution (Rzhetsky and Nei, 1992), and maximum-likelihood (Felsenstein, 1981). Bootstrap analysis (1000 replicates) was used to evaluate the trees topology. Kimura two-parameter model was used for phylogeny construction and evolutionary distances analysis (Kimura, 1981).

The GenBank/EMBL/DDBJ accession number for the 16S rRNA gene sequence of strain HM190 is MN897722.

**Genome Sequencing and Annotation**

For DNA isolation, strain HM190 was inoculated into MB medium and grown at 30°C with shaking (200 r/min) for 7 days. High-quality genomic DNA was extracted using the Bacteria Genomic DNA Extraction Kit (DongSheng Biotech) according to the manufacturer's instructions. The genome was sequenced using the PacBio RS II platform and the Illumina HiSeq 4000 platform (Beijing Genomics Institute) (163-fold). Four SMRT cells Zero-Mode Wave guide arrays of sequencing were used by the PacBio platform to generate the sub-reads set. The PacBio sub-reads (length < 1 kb) were removed. To improve the accuracy of the genome sequence, the GATK4 and the SOAP (SOAP2, SOAPsnp, SOAPindel) tool packages were used to make single-base corrections. To confirm the presence of any plasmid, the filtered Illumina reads were mapped using the SOAP to the bacterial plasmid database. The gene prediction was performed by the glimmer 3 with hidden Markov models. The open reading frames (ORFs) were annotated by the Rapid Annotation using Subsystem Technology (RAST) server online (Overbeek et al., 2014). The Clusters of Orthologous Groups (COG) database was used for general function annotation. For the identification of secondary metabolism gene clusters, the antiSMASH 5.0 program was used (Blin et al., 2019). Whole Genome Shotgun project of *Streptomyces* sp. HM190 has been deposited at DDBJ/ENA/GenBank under the accession CP047318.

**Fermentation, Extraction, and Isolation of Strain HM190**

Strain HM190 was inoculated into a 1000-mL flask containing 250 mL of the seed culture medium consisting of glucose (0.4%), yeast extract (0.4%), malt extract (1.0%), and CaCO$_3$ (0.2%), with a pH of 7.2. After incubated at 30°C for 3 days on a rotary shaker operating at 250 r/min, the 5.0% seed culture broth was added to a 50 L fermenter (containing 30 L of fermentation medium) incubation for 7 days at 30°C. The fermentation medium component was the same as the isolation medium of MB (marine broth 2216 medium).

The final 30 L of fermentation broth was centrifuged to separate supernatant and bottomed mycelial cake. The mycelial cake was subsequently extracted with MeOH (3 L). The supernatant passed through a Diaion HP-20 resin column (Mitsubishi Chemical Co., Ltd., Tokyo, Japan) and eluted with 95% EtOH. The MeOH extract and the EtOH eluent were concentrated under reduced pressure at 55°C to yield a crude extract (30 g). The crude extract was chromatographed on a silica gel column (Qingdao Haiyang Chemical Group, Qingdao, China; 200-300 mesh) eluted with CHCl$_3$/MeOH (100:0, 98:2, 95:5, 90:10, 80:20, 70:30, 60:40, and 50:50, v/v) to give six fractions (Fr.1-6) based on the TLC profiles. After the Fr.5 was concentrated under reduced pressure at 53°C and dried in vacuo, the material was subjected to a Sephadex LH-20 gel column (GE Healthcare, Glies, United Kingdom) and eluted with CHCl$_3$/MeOH (1:1, v/v), to yield two fractions (Fr.5-1 and Fr.5-2) by using TLC detection. The Fr.5-1 was further chromatographed on a silica gel column and eluted with CHCl$_3$/MeOH (100:0, 98:2, 95:5, 90:10, 85:15, 80:20, 75:25, and 70:30 v/v) to give four fractions (Fr.5-1-1–Fr.5-1-4). The Fr.5-1-1 was analyzed and purified by semi-preparative reversed-phase HPLC and eluted with CH$_3$CN/H$_2$O (60:40, v/v) to get compound 1 (t$R$...
28.5 min, 15.2 mg). The Fr.5-1-3 was subjected to a Sephadex LH-20 column eluted with CHCl₃/MeOH (1:1, v/v) to give three fractions (Fr.5-1-3a, Fr.5-1-3b, and Fr.5-1-3c). The Fr.5-1-3c was further separated by semi-preparative reversed-phase HPLC and eluted with CH₂CN/H₂O (48:52, v/v) to get compounds 2 (tr 37.0 min, 9.8 mg) and 3 (tr 45.0 min, 5.8 mg), respectively.

Compound 1: white powder; [α]D²⁵ = −4 (c 0.05, EtOH); IR (KBr) νmax 3441.9, 2931.3, 1701.1, 1647.0, 1419.2, 1392.8, and 1287.2 cm⁻¹; UV (EtOH) λmax (log ε) 202 (3.93), 217 (3.82) nm; ¹H (400 MHz) and ¹³C NMR (100 MHz) data shown in Table 1; positive HRESIMS m/z 765.4359 [M + Na]⁺, (calcd. for C₉₉H₈₈NaO₁₃, 765.4396).

Compound 2: white powder; [α]D²⁵ = −6 (c 0.05, EtOH); IR (KBr) νmax 3400, 2900, 1700, 1640, 1450, 1380, 1260, 1180, 1080, and 980 cm⁻¹; UV (MeOH) λmax 223 nm (ε 8,900); ¹H (400 MHz) and ¹³C NMR (100 MHz) data shown in Table 1; positive HRESIMS m/z 763.4546 [M + Na]⁺, (calcd. for C₉₉H₈₆NaO₁₃, 763.4603).

**Table 1** The ¹H and ¹³C NMR data for compounds 1–3 (in ppm).

| Position | ¹H (CDCl₃) | ²H (CDCl₃) | ³H (CD₂OD) |
|----------|------------|------------|-------------|
|          | δc (J in Hz) | δc (J in Hz) | δc (J in Hz) |
| 1        | 164.9, C    | 165.1, C    | 166.9, C    |
| 2        | 119.5, CH   | 119.5, CH   | 120.0, CH   |
| 3        | 149.9, CH   | 151.1, CH   | 153.8, CH   |
| 4        | 75.1, C     | 75.5, C     | 76.8, C     |
| 5        | 79.3, CH    | 80.1, CH    | 81.2, CH    |
| 6        | 35.3, CH    | 39.5, CH    | 39.8, CH    |
| 7        | 77.2, CH    | 77.7, CH    | 77.8, CH    |
| 8        | 65.9, CH    | 35.1, CH₂    | 36.2, CH₂    |
| 9        | 73.0, CH    | 78.4, CH    | 78.4, CH    |
| 10       | 77.1, C     | 75.3, C     | 75.7, C     |
| 11       | 36.8, CH₂   | 37.5, CH₂   | 38.7, CH₂   |
| 12       | 21.9, CH₂   | 22.5, CH₂   | 23.8, CH₂   |
| 13       | 32.4, CH₂   | 32.9, CH₂   | 35.3, CH₂   |
| 14       | 132.5, CH   | 132.1, CH   | 126.2, CH   |
| 15       | 132.7, CH   | 133.0, CH   | 129.7, CH   |
| 16       | 51.1, CH    | 50.9, CH    | 110.0, C    |
| 17       | 96.9, C     | 97.1, C     | 148.8, C    |
| 18       | 39.7, CH₂   | 39.7, CH₂   | 32.8, CH₂   |
| 19       | 66.1, CH    | 66.4, CH    | 68.9, CH    |
| 20       | 34.3, CH    | 34.7, CH    | 34.2, CH    |
| 21       | 69.5, CH    | 69.2, CH    | 71.9, CH    |
| 22       | 35.3, CH₂   | 35.4, CH₂   | 36.8, CH₂   |
| 23       | 98.2, C     | 98.4, C     | 99.4, C     |
| 24       | 34.2, CH₂   | 34.2, CH₂   | 30.7, CH₂   |
| 25       | 19.3, CH₂   | 27.0, CH₂   | 27.8, CH₂   |
| 26       | 30.8, CH₂   | 30.8, CH₂   | 30.8, CH₂   |
| 27       | 72.3, CH    | 67.5, CH    | 69.8, CH    |
| 28       | 43.7, CH₂   | 41.6, CH₂   | 41.8, CH₂   |
| 29       | 68.3, CH    | 68.6, CH    | 71.4, CH    |
| 30       | 29.9, CH₂   | 30.3, CH₂   | 32.1, CH₂   |
| 31       | 10.2, CH₂   | 10.4, CH₂   | 10.5, CH₃   |
| 32       | 66.2, CH₂   | 66.3, CH₂   | 72.4, CH₂   |
| 33       | 30.5, CH    | 30.6, CH    | 28.6, CH₃   |
| 34       | 33.7, CH₂   | 34.3, CH₂   | 30.8, CH₂   |
| 35       | 28.1, CH₃   | 27.4, CH₃   | 27.4, CH₃   |
| 36       | 5.6, CH₃    | 5.2, CH₃    | 6.0, CH₃    |
| 37       | 22.4, CH₃   | 19.9, CH₃   | 20.9, CH₃   |
| 38       | 6.4, CH₃    | 6.3, CH₃    | 5.6, CH₃    |
| 39       | 16.9, CH₃   | 11.4, CH₃   | 11.7, CH₃   |
| 40       | 17.1, CH₃   | 17.1, CH₃   | 17.6, CH₃   |

---

₃¹H (400 MHz), ³¹C (100 MHz); ²¹H (600 MHz), ²¹C (150 MHz).
Compound 3: white powder; [α] = −5 (c 0.05, EtOH); IR (KBr) νmax 3450, 2950, 1720, 1660, 1460, 1390, 1280, 1180, 1100, 980 cm−1; UV (MeOH) λmax 247 nm (ε 13,500); 1H (600 MHz) and 13C NMR (150 MHz) data shown in Table 1; positive HRESIMS m/z 745.4437 [M + Na]+, (calcd. for C39H66NaO13, 745.4497). The IR and UV data of compounds 2 and 3 from Akira H. et al. X-Ray Crystallography
Small seed crystals of compound 1 were produced by controlling evaporation of solutions at room temperature over 2 weeks in CHCl3/MeOH (60:40, v/v). Re-crystallization at 4°C to induce the production of high-quality crystals for X-ray diffraction was conducted for 20 days. A suitable crystal was selected and mounted on a Bruker D8 Venture diffractometer. The crystal was kept at 170.02 K during data collection. The structure was solved with the ShelXT (Sheldrick, 2015) structure solution program using Intrinsic Phasing. Calculations were made with the ShelXL (Kratzert et al., 2015) refinement package using least squares minimization as implementation in Olex2 (Dolomanov et al., 2009). Crystallographic data (excluding structure factor) for compound 1 has been deposited with the Cambridge Crystallographic Data Center under the deposition number CCDC 1990958.

Biological Assay
The antimicrobial activities of compounds 1–3 against pathogenic bacteria Klebsiella pneumoniae, methicillin-resistant Staphylococcus aureus, and pathogenic fungus Candida albicans were detected with the minimum inhibitory concentrations (MICs) method recommended by the Clinical and Laboratory Standards Institute (Wang et al., 2019). Amphotericin B (an antifungal antibiotic) and gentamicin (an antibacterial antibiotic) were used as a positive control.

The cytotoxic activities of compounds 1–3 were investigated against the human colon tumor cell line HCT116, the human lung carcinoma cell line A549, and the hepatocellular carcinoma cell line HepG2 in vitro by the CCK8 (cell counting kit-8) colorimetric method. The cell lines were incubated in a 5% CO2 incubator at 37°C for 4 h under Dulbecco's Modified Eagle's Medium (DMEM) solution containing 10% calf serum. The adherent cells of the logarithmic growth stage were digested and seeded in a 96-well culture plate at a density of 1 × 104 cells per/well. Then the test samples and controls were added to the medium and cultivated for 48 h. Further the CCK8 (DOJINDO, Kumamoto, Japan) reagent was added to the medium and cultured for 3 h. Cell viability was detected by the absorbance at 450 nm using a SpectraMax M5 microplate reader (Molecular Devices Inc., Sunnyvale, CA, United States) (Wang et al., 2009). Doxorubicin was used as a positive control. Cell solution (dimethylsulfoxide) was tested as a negative control. The inhibitory rate of cell proliferation was expressed as IC50 value.

RESULTS AND DISCUSSION
16S rRNA Gene Sequence and Phylogenetic Analysis
PCR was used to determine the 16S rRNA gene sequence (1451 bp, NCBI GenBank accession number: MN897722) of strain HM190. The results of the analysis indicated that the isolate belonged to the genus Streptomyces and showed the highest sequence similarities to Streptomyces heilongjiangensis NEAU-W2T (99.79%) and Streptomyces neyagawaensis NRRL B-3092T (99.59%). In the phylogenetic analysis based on a neighbor-joining tree (Figure 2), strain HM190 fell within the cluster of the genus Streptomyces and formed a coherent clade with S. heilongjiangensis NEAU-W2T. The clade had firm bootstrap support and represented an independent lineage. Similar results were obtained by using the maximum-parsimony and maximum-likelihood trees (Supplementary Figures S26, S27).
Genome Sequence and AntiSMASH Analysis

The whole genome sequence of *Streptomyces* sp. HM190 was assembled using the PacBio RSII and Illumina HiSeq400 platforms. High-quality clean data of size 1266 Mb with a total genome size of 9.10 Mb were generated and assembled into a linear chromosome. The complete genome consisted of 776,282 bp composing a linear chromosome with 71.97% G + C content. No plasmid was identified. A total of 6971 protein-coding genes were predicted. Of these, 4728 genes
(67.8%) were annotated by querying the COG database. A map of the chromosome and the COG functional categories of strain HM190 are shown in Figure 3. For further accurate secondary metabolism BGC (SMBGC) mining analysis, 30 SMBGCs of strain HM190 were proposed using antiSMASH 5.0 (Blin et al., 2019), occupying 15.4% of the chromosome (Supplementary Table S2).

Twelve of the 30 BGCs were predicted to be responsible for the production of PKS and non-ribosome peptide- (NRPS) derived secondary metabolites, including one type I PKS (cluster 5), one type II PKS (cluster 14), two type III PKS (clusters 2, 16), two NRPS (clusters 3, 21), and six hybrid BGCs (clusters 1, 4, 6, 10, 15, 23), which possessed genes encoding more than one type of biosynthetic enzyme. Cluster 5 encoded for a type I PKS that is probably responsible for the biosynthesis of a 20/21-membered macrolide of apoptolidin. A sequence similarity of 74.5% was obtained between cluster 5 and the Nocardiopsis sp. FU 40 apoptolidin gene cluster (NCBI accession number: JF819834) based on BLASTN analysis. According to Du et al. (2011), the apoptolidin gene cluster of Nocardiopsis sp. FU 40 produces 10 apoptolidin and isoapoptolidin compounds, and the biosynthetic pathway was confirmed. Cluster 14 was probably involved in the production of a hiroshidine-like compound, based on its high similarity to the type II PKS enzyme from the hiroshidine BGC from Streptomyces hiroshimensis (Moon et al., 2019). Gene clusters 2 and 16 encoded, respectively, for type III PKSs that resembled the alkylresorcinol and germicidin PKSs from Streptomyces spp. (Funabashi et al., 2008; Becerril et al., 2018). Cluster 3 encoded for an NRPS that probably synthesized scabichelin (Kodani et al., 2013). Cluster 21 was
similar to the *Salinispora tropica* CBM-440 gene cluster associated with biosynthesis of the anticancer agent salinosporamide A (Fenical et al., 2009). Based on its high similarity to the recently discovered naphthrydindomycin BGC from *Streptomyces lusitanus*, cluster 1 was probably involved in the production of a naphthrydindomycin-like compound (Pu et al., 2013). Gene clusters 4 and 15 were predicted to be responsible for biosynthesis of gaudimycin- and glycosyl ester-type compounds, respectively (Kallio et al., 2008; Wang et al., 2015). Cluster 6 encoded for a type II PKS that probably synthesized granaticin (Ichinose et al., 2009). Cluster 18 was probably involved in the production of gaudimycin- and glycosyl ester-type compounds, respectively (Ichinose et al., 2009). Cluster 1 was probably involved in the production of a naphthrydindomycin-like compound (Pu et al., 2013). Gene clusters 11 and 25 (cluster 13 showed high similarity to bottromycin A2 (Nakagawa et al., 1989). Compound 1 had six methyl groups, whereas phthoramycin contained seven methyl groups. Comparisons of the NMR data between 1 and phthoramycin indicated that 1 differed from phthoramycin by the absence of a methyl substituent at C-26 and the presence of a hydroxy group at C-8.

The molecular formula of compound 2 was established as C_{40}H_{50}O_{12} (seven degrees of unsaturation) by the HRESIMS data (m/z 763.4437 [M + Na]^+). The molecular formula of compound 2 revealed that it had the same weight as the known kaimonolide A (Akira et al., 1989). Based on detailed NMR analysis of the crude extract, the structure of compound 2 was found to show the same planar structure with kaimonolide A (Akira et al., 1989).

**Structure Elucidation**

Based on the COG categorization of the genome of strain HM190, there were a considerable number of genes related to secondary metabolism biosynthesis, transport, and catabolism (Figure 3). Together with the antiSMASH analysis, strain HM190 appeared to be an isolate with prolific potential for the production of various new and unique secondary metabolites. In order to assess this, the strain was grown for scale-up culture in 30 L of fermentation medium for 7 days at 30°C. Results of HPLC analysis of the crude extract are shown in Figure 4. Compounds 1–3 were purified and subjected to detailed structural analysis.

Table 2 shows the crystal data and structural refinement for compound 1.

**Table 2 | Crystal data and structural refinement for compound 1.**

| Empirical formula | C_{39}H_{50}O_{13} |
|-------------------|-------------------|
| Formula weight    | 742.91            |
| Temperature (K)    | 170.02            |
| Crystal system     | Monoclinic        |
| Space group        | P21               |
| Unit cell dimensions | a = 9.9665(11) Å, a = 90° |
|                    | b = 21.014(2) Å, b = 96.321(3) |
|                    | c = 21.960(2) Å, c = 90° |
| Volume (Å^3)       | 4569.3(8)         |
| Z                  | 4                 |
| Calculated density (Mg/m³) | 1.080       |
| Absorption coefficient (mm⁻¹) | 0.422    |
| F(000)             | 1616.0            |
| Crystal size (mm³) | 0.06 × 0.06 × 0.05 |
| Wavelength (Å)     | 1.34139           |
| Theta range for data collection | 7.32–109.99° |
| Index ranges       | -12 ≤ h ≤ 12, -25 ≤ k ≤ 25, -26 ≤ l ≤ 25 |
| Reflections collected | 42,745            |
| Independent reflections | 17,092 [R(int) = 0.0893, R(sigma) = 0.1157] |
| Data/restraints/parameters | 17,092/18,964        |
| Goodness-of-fit on F^2 | 0.997              |
| Final R indices   | R₁ = 0.0856, wR₂ = 0.1426 |
|                   | R₁ = 0.1294, wR₂ = 0.1700 |
| Largest diff. peak and hole (e Å⁻³) | 0.23 and −0.25 |
| Flack parameter   | −0.11(15)         |
spectral data for 2 and 3 are shown in Table 1. Comparisons of the molecular formulas and NMR data of 2 and 3 indicated that 3 was a dehydrated product of 2, resulting from the formation of a double bond at C-16/C-17. It should be noted that during our studies, compound 2 would have gradually changed to 3 in CDCl₃ solution (Akira et al., 1989). Compounds 2 and 3 also closely resembled phthoramycin and cytovaricin B, which have been reported in Streptomyces spp. (Sakurai et al., 1983; Nakagawa et al., 1989).

**Absolute Configurations of Compounds 1–3**

To define the correct absolute configuration of compound 1, a single crystal of compound 1 was prepared for X-ray single-crystal diffraction. Crystals of compound 1 (C₃₀H₆₆O₁₃, CCDC 1990958) were almost colorless and monoclinic with typical dimensions 0.06 × 0.06 × 0.05 mm³. Compound 1 crystallized in space group P2₁ with unit cell parameters a = 9.9665(11), b = 21.014(2), c = 21.950(2) Å, and Z = 4 molecules per unit cell. The refined structure of compound 1 had final R index [I ≥ 2σ(I)] values of R₁ = 0.0656, wR₂ = 0.1426, and final R index [all data] values of R₁ = 0.1294, wR₂ = 0.1700. Detailed crystal data and structure refinements are shown in Table 2. The quality of the crystal was high enough for absolute configuration analysis to be carried out (Palmer and Potter, 2008). The absolute structure Flack parameter was −0.11(15), which was equal to 0.0 with standard deviation of 15 (Flack, 1983), confirming that the structure of 1 had been assigned the correct absolute configuration. Figure 6 shows the ORTEP diagram of compound 1 according to the X-ray data.
Compound 2 was dehydrated at C16/C-17 to produce compound 3. Thus, we determined the absolute configurations of compounds 2 and 3 simultaneously. Based on the similarity of the NMR spectra of compounds 1–3, including chemical shifts and peak multiplicities (Table 1), the multiple stereocenters of 2 and 3 had the same absolute configuration as in compound 1, although within differences at C-26 and C-27, because of the addition of a methyl substituent at C-26. According to Bilyk et al. in 2019, the absolute configurations of 22- to 26-membered macrolides compounds were confirmed by the data of NMR spectra, single-crystal X-ray crystallography, and elegant total synthesis of representative. Detailed NMR data comparison revealed a striking structural similarity of compounds 2, 3 to kaimonolide A (Akira et al., 1989; Bilyk et al., 2019) and cytovaricin B (Sakurai et al., 1983; Yamashita et al., 1997). Finally, the absolute configurations of compounds 2 and 3 were established as shown in Figure 1.

Biosynthetic Pathway Analysis

In strain HM190, the predicted BGC of macrolide compounds 1–3 contained 15 individual ORFs, including eight type I PKS genes, three cytochrome P450 genes, one crotonyl-CoA reductase (CCR) gene, one transcriptional regulator gene, and two hypothetical protein genes (Du et al., 2011; Bilyk et al., 2019). PKS genes S1–S8 were proposed to be responsible for the biosynthesis of the polyketide core of compounds 1–3. A suggested module and domain organization and a proposed model for the PKS template assembly of the macrolide backbone are shown in Figure 7. PKS gene S2 encoded the putative protein possessing the probable initiating module. The next 13 extension modules were proposed to be encoded by genes S3–S8 and were ordered according to the predicted linear arrangement of the required domains. Finally, the terminating module protein was identified to be encoded by gene S8 by its terminal punctuation with a thioesterase domain. One additional PKS gene S1 containing a complete module sequence “KS-AT-KR-ER-DH-ACP” and CCR apparently encoded a free-standing isobutylmalonyl-ACP that assembled to the macrolide backbone of module 7 in gene S5. After the core skeletons were constructed, the structural diversity of both macrocyclization and spiroketalization was mainly due to the tailoring reactions of cytochrome P450 genes to complete the hydroxylation reactions.
were established by the X-ray single-crystal diffraction and NMR data analysis. The absolute configurations of compounds 1–3 were determined based on the HRESIMS data of 7,762,826 bp with 71.97% G+C content. Based on the results of antiSMASH analysis, a total of 30 gene clusters were predicted to be involved in the biosynthesis of secondary metabolites; 12 of the 30 BGCs were responsible for the production of PKS- and NRPS-derived secondary metabolites. Several other secondary metabolites were also predicted by their BGCs. Gene cluster 5 was predicted to be involved in the production of an apoptolidin-like macrolide based on its high similarity of 74.5% to the type I PKS enzyme from the macrolide BGC of Nocardiopsis sp. HM190. The whole-genome shotgun project of Streptomyces sp. HM190 has been deposited at DDBJ/ENA/GenBank under the accession CP047318.

### DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the GenBank/EMBL/DDBJ and accession number for the 16S rRNA gene sequence of strain HM190 is MN897722. Whole Genome Shotgun project of Streptomyces sp. HM190 has been deposited at DDBJ/ENA/GenBank under the accession CP047318.

### ETHICS STATEMENT

All studies of this article were processed with the standard of biosecurity and institutional safety procedures.

### AUTHOR CONTRIBUTIONS

YY conducted and performed the experiments and prepared the manuscript. JW and MW designed the experiments. XM and SW determined the compounds structures and the biosynthetic pathway analysis. NA performed the biological activities experiments. ZZ and RZ undertook the genome analysis. CY and YN performed the microbiology experiments. JZ revised the manuscript. All authors contributed to the article and approved the submitted version.

### FUNDING

This work was supported by the Science and Technology Basic Resources Investigation Program of China (2017FY100300), the Sea Mountain Environmental Management Plan and APEI Construction Project Research (DY135-E2-2-05), and the Key Research and Development Program of Zhejiang Province (Grant No. 2020C02028).

### ACKNOWLEDGMENTS

The authors thank Yu Liu (Life Science Institute of Zhejiang University) for performing NMR spectrometry.

### SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.01464/full#supplementary-material
The following are available online: the HRESIMS and NMR spectra (Supplementary Figures S1–S25) of compounds 1–3, the maximum-parsimony and maximum-likelihood phylogenetic trees (Supplementary Figures S26, S27) of strain HM190, the cluster of organelous groups (COG) classification (Supplementary Table S1) and the antiSMASH-predicted BGCs (Supplementary Table S2) of strain HM190, and the crystal data (Supplementary Tables S3–S5) of compound 1.

REFERENCES

Akira, H., Hideki, O., Toshitaka, K., Mitsuru, N., Hiroshi, H., and Akira, I. (1989). Structure of kaimonolide A, a novel macrolide plant growth inhibitor from a Streptomyces strain. Agr. Biol. Chem. 53, 2831–2833. doi: 10.1271/bbb1961.53.2831

Alfredo, B., Sarmiento, V. A., Miguel, O., Pérez, V. I., Martín, J., Nuria, D. P., et al. (2017). Lobophorin K, a new natural product with cytotoxic activity produced by Streptomyces sp. M-207 associated with the deep-sea coral Lophelia pertusa. Mar. Drugs 15:144. doi: 10.3390/md15050144

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., and Lipman, D. I. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402. doi: 10.1093/37.25.3389

Anzai, Y., Kim, H., Park, J. Y., Wakabayashi, H., and Oyaizu, H. (2000). Phylogenetic affiliation of the Pseudomonads based on 16S rRNA sequence. Int. J. Syst. Evol. Microbiol. 50, 1563–1589. doi: 10.1099/00218998.50.4-5.1563

Ikeda, H., Murakami, O., Sugawara, Y., and Sakai, M. (2007). The role of sediment microorganisms in the productivity, conservation, and rehabilitation of mangrove ecosystems: an overview. Biol. Fert. Soils 42, 339–341. doi: 10.1007/s00172-007-0247-6

Hwang, K. S., Kim, H. U., Charusantis, P., Palsson, B. O., and Lee, S. Y. (2014). Systems biology and biotechnology of Streptomyces species for the production of secondary metabolites. Biotechnol. Adv. 32, 255–268. doi: 10.1016/j.biotechadv.2013.10.008

Ichinose, K., Bedford, D. J., Tornus, D., Bechthold, A., and Hopwood, D. A. (1998). The ganatic biosynthetic gene cluster of Streptomyces violaceoruber TU22: sequence analysis and expression in a heterologous host. Cell Chem. Biol. 5, 647–659. doi: 10.1017/S01074-S521/98/90292-7

Ioanna, P., Konstantinos, L., Jakob, J., Chen, I. A., Tatyana, S., Bahador, N., et al. (2012). The genomes online database (GOLD) v. 4: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res. 40, 571–579. doi: 10.1093/nar/gkr1100

Kallo, P., Liu, Z., Mantslā, P., Niemi, J., and Metsä-Ketelä, M. (2008). Sequential action of two flavoenzymes, PGae and PGAm, in angucycline biosynthesis: chemoenzymatic synthesis of gaudimycin. C. R. Chim. 11, 157–166. doi: 10.1016/j.crci.2007.12.011

Kathiresan, K. and Bingham, B. L. (2001). Biology of mangroves and mangrove ecosystems. Adv. Mar. Biol. 40, 81–251. doi: 10.1016/S0065-6881(01)00003-4

Kimura, M. (1981). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120. doi: 10.1007/BF01731581

Kodani, S., Bizc, J., Song, L., Deeth, R. J., Ohnishi-Kameyama, M., Yoshida, M., et al. (2013). Structure and biosynthesis of scabichelin, a novel tris-hydroxamate siderophore produced by the plant pathogen Streptomyces scabies 82/77. Org. Biomol. Chem. 11, 6466–6494. doi: 10.1039/C3OB34368B

Kratzelt, D., Holstein, J. L., and Kroissing, I. (2015). DSR: enhanced modelling and refinement of disordered structures with Shelxl. J. Appl. Crystallogr. 48, 933–938. doi: 10.1107/S0021889815005580

Li, M., Hong, Y. G., Cao, H. L., and Gu, J. D. (2011). Mangrove trees affect the community structure and distribution of anamnosa bacteria at an anthropogenic-polluted mangrove in the pearl river delta reflected by 16s rRNA and hydrazine oxidoreductase (hox) encoding gene analyses. Ecotoxicology 20, 1780–1790. doi: 10.1007/s10646-011-0711-4

Liu, R., Deng, Z., and Liu, T. (2018). Streptomyces species: ideal chassis for natural product discovery and overproduction. Metab. Eng. 50, 74–84. doi: 10.1016/j.ymben.2018.05.015

Moon, K., Xu, F., Zhang, C., and Seyedsayamdost, M. R. (2019). Bioactivity-HITES unveils cryptic antibiotics encoded in actinomycete bacteria. ACS Chem. Biol. 14, 767–774. doi: 10.1021/acschembio.9b00049

Nakagawa, A., Miura, S., Imai, H., Imamura, N., and Omura, S. (1989). Structure and biosynthesis of a new antifungal antibiotic, phthoramycin. J. Antibiot. 42, 679–683. doi: 10.1093/jaubio/42.9.679

Overbeek, R., Olson, R., Busch, G. D., Olsen, G. J., Davis, J. J., Diz, T., et al. (2014). The SEED and the Rapid Annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res. 42, 206–214.

Palmer, R. A., and Potter, B. S. (2008). X-ray structures and absolute configurations of the antibiotics oligomycins A, B and C: inhibitors of ATP synthase. J. Chem. Crystallogr. 38, 243–253. doi: 10.1016/S0021-8898(07)-93505-1

Pu, J. Y., Peng, C., Tang, M. C., Zhang, Y., Guo, J. P., Song, L. Q., et al. (2013). Naphthyridinomycin biosynthesis revealing the use of leader peptide to guide nonribosomal peptide assembly. Org. Lett. 15, 3674–3677. doi: 10.1021/ol401549y

Rzhetsky, A., and Nei, M. (1992). Statistical properties of the ordinary least-squares, generalized least-squares, and minimum-evolution methods of phylogenetic inference. J. Mol. Evol. 34, 365–375. doi: 10.1007/BF01731581

Saitou, N., and Nei, M. (1987). The neighbour-joining method: a new method for inferring phylogenies. J. Mol. Evol. 26, 406–425. doi: 10.1007/BF00160427

Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., et al. (2019). AntiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res. 47, 81–87. doi: 10.1093/nar/gkz310

Smirnova, T. A., and Deeth, R. J. (2013). Structure and biosynthesis of scabichelin, a novel tris-hydroxamate siderophore produced by the plant pathogen Streptomyces scabies 82/77. Org. Biomol. Chem. 11, 6466–6494. doi: 10.1039/C3OB34368B

Ye et al. Combined Genomics–Metabolomics Analysis

Ye et al. Combined Genomics–Metabolomics Analysis
