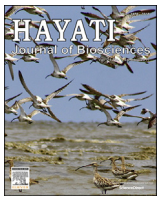
HAYATI Journal of Biosciences 24 (2017) 6–15

HOSTED BY

Contents lists available at ScienceDirect

HAYATI Journal of Biosciences

journal homepage: http://www.journals.elsevier.com/hayati-journal-of-biosciences



Original research article

# Metagenomic Survey of Potential Symbiotic Bacteria and Polyketide Synthase Genes in an Indonesian Marine Sponge

Nia M. Kurnia,[1] Agustinus R. Uria,[2,6*] Yudi Kusnadi,[2] Lidia Dinawati,[3] Dewi S. Zilda,[2] Tri A. Hadi,[4] Siswa Setyahadi,[5] Feliatra Felix[3]

[1] *Graduate School of Pharmacy, Pancasila University, Jakarta, Indonesia.*
[2] *Laboratory of Marine Biotechnology, Research Center for Marine and Fisheries Product Processing and Biotechnology, Ministry of Marine Affairs and Fisheries, Jalan KS Tubun Petamburan VI, Jakarta 10260, Indonesia.*
[3] *Faculty of Fisheries and Marine Science, University of Riau, Pekanbaru, Indonesia.*
[4] *Research Center for Oceanography, Indonesian Institute of Sciences (LIPI), Indonesia.*
[5] *Center for Bioindustrial Technology, Agency for the Assessment and Application of Technology (BPPT), Jakarta, Indonesia.*

ARTICLE INFO

*Article history:*
Received 9 October 2016
Received in revised form
12 April 2017
Accepted 26 April 2017
Available online 29 May 2017

*KEYWORDS:*
"*Candidatus* Entotheonella sp.",
marine sponge,
microbiome,
polyketide synthase,
*Rhabdastrella* sp.

ABSTRACT

There has been emerging evidence that the bacteria associated with marine sponges are the key producers of many complex bioactive compounds. The as-yet uncultured candidate bacterial genus "*Candidatus* Entotheonella" of the marine sponge *Theonella swinhoei* from Japan have recently been recognized as the source of numerous pharmacologically relevant polyketides and modified peptides, as previously reported by the Piel group (Wilson *et al.* 2014). This work reported the presence of "*Candidatus* Entotheonella sp." in the highly complex microbiome of an Indonesian marine sponge from Kapoposang Island, South Sulawesi. We further identified the Kapoposang sponge specimen used in this work as *Rhabdastrella* sp. based on the integrated morphological, histological, and cytochrome oxidase subunit I (COI) gene analyses. To detect the polyketide biosynthetic machinery called type I polyketide synthase (PKS) in this Indonesian *Rhabdastrella* sp., we amplified and cloned the ketosynthase-encoding DNA regions of approximately 700 bp from the uncultured sponge's microbiome. Further sequencing and analysis of several randomly chosen clones indicated that all of them are mostly likely involved in the biosynthesis of methyl-branched fatty acids. However, employing a PKS-targeting primer designed in this work led to the isolation of four positive clones. BlastX search and subsequent phylogenetic analysis showed that one of the positive clones, designed as RGK32, displayed high homology with ketosynthase domains of many type I PKS systems and may belong to the subclass *cis*-AT PKS group.


## 1. Introduction

Polyketides are a large class of highly diverse natural products, which include many important pharmaceuticals, agrochemicals, and veterinary agents (Cane and Walsh 1999). The enormous diversity of most polyketides are built from simple carboxylic acid monomers through a number of programmed events catalyzed by polyketide synthases (PKSs) (Hertweck 2009; Llewellyn and Spencer 2007). Three types of bacterial PKSs known to date are type I, II, and III. Type I PKSs are multifunctional enzymes that are organized into modules, in which each module harbors catalytic domains that mediate one cycle of polyketide elongation and modification. A module minimally consists of three domains: an acyltransferase (AT) domain that chooses the appropriate acyl-CoA building block, an acyl carrier protein (ACP) domain that serves as an anchor for the building block, and a ketosynthase (KS) domain that catalyzes the polyketide chain elongation. Optional domains perform various functional modifications on the β-position (Piel 2010; Rawlings 2001; Staunton and Weissman 2001). In principle, type I PKS systems are similar to type I fatty acid synthases (FASs) that use simple precursors such as acetyl-CoA and malonyl-CoA monomers. FASs are different from type I PKSs in that they consist of a single module that is

* Corresponding author.
*E-mail addresses:* agustinus.uria@gmail.com, agustinus.uria@kkp.go.id (A.R. Uria).
Peer review under responsibility of Institut Pertanian Bogor.
[6] Present address: Faculty of Pharmaceutical Sciences, Hokkaido University, Kita 12, Nishi 6, Kita-ku, Sapporo 060-0812, Japan.

http://dx.doi.org/10.1016/j.hjb.2017.04.004

used iteratively and rigidly adhered-to specificities of component enzymes to ensure the faithful production of a saturated long-chain fatty acid (Smith and Tsai 2007). Since the PKS module architecture and function corresponds to the resulting intermediates, a polyketide structure can be predicted from PKS domain sequences. This correspondence is known as the colinearity principles that are usually applied for a subclass of modular type I PKSs called "*cis*-AT PKSs" with an AT domain integrated in each module (Hertweck 2009; Piel 2010). In contrast to *cis*-AT PKSs, the subclass *trans*-AT PKSs lack integrated AT domains and perform unusual enzymatic features to generate diverse complex polyketides. In *trans*-AT PKS systems, a free-standing AT domain acts in *trans* to load acyl building blocks into the assembly line (Nguyen *et al.* 2008).

Marine sponges are a rich source of natural products that include polyketides and modified peptides. There has recently been emerging evidence that many sponge-derived polyketides and peptides are actually produced by the associated symbiotic bacteria (Piel 2010; Uria and Piel 2009). The associated microbial consortia can account for up to 60% of the sponge total biomass and consist of hundreds to thousands of bacterial species (Taylor *et al.* 2007). However, the limitation to cultivate the vast majority (>99%) of these bacterial symbionts (Amann *et al.* 1995; Hugenholtz *et al.* 1998; Webster and Hill 2001) has hampered the attempts to access their biotechnological potential. Cultivation-independent approaches, popularly called metagenome mining, have enabled cloning of biosynthetic gene clusters of interest within diverse genetic mixtures of sponge symbiotic systems providing convincing proof about the bacterial origin of sponge-derived polyketides and modified peptides, as exemplified by onnamide A (Nguyen *et al.* 2008; Piel *et al.* 2004), psymberin (Fisch *et al.* 2009), polytheonamides (Freeman *et al.* 2012), and misakinolide (Ueoka *et al.* 2015). Subsequent heterologous expression of gene clusters in easily culturable bacteria could generate sustainable and large-scale supply of sponge-derived drug candidates for drug development (Piel 2004; Uria and Piel 2009).

Identifying PKS systems in the microbiomes of marine sponges is extremely challenging, because sponge-associated microbial assemblages are highly complex (Scheuermayer *et al.* 2006; Taylor *et al.* 2007) that may contain hundreds of individual genomes with numerous homologous genes from diverse pathways (Schirmer *et al.* 2005). Recent single-cell and metagenomic studies on microbiome inhabiting the Japanese sponge *Theonella swinhoei* have revealed that "*Candidatus* Entotheonella sp." are the true producers of many polyketides and modified peptides (Ueoka *et al.* 2015; Wilson *et al.* 2014). Furthermore, *Entotheonella* symbionts were found to be widely distributed in taxonomically diverse sponge species from distant geographical regions (Wakimoto *et al.* 2014; Wilson *et al* 2014). Such previous studies strongly suggest that different *Entotheonella* variants, either in different sponge species or within the same sponge species from geographically different locations, may produce different biologically active compounds encoded on their genomes.

Some sponges belonging to the class Demospongiae (*e.g. Theonella swinhoei*, *Discodermia dissoluta*) are known to harbor "*Candidatus* Entotheonella sp.", as-yet uncultivable filamentous bacteria living outside the sponge cells mostly in the mesohyl part (Brück *et al.* 2008; Schmidt *et al.* 2000; Wilson *et al.* 2014). Their large unusual morphology and extracellular occurrence made easy to isolate or enrich them by simple mechanical separation (Bewley *et al.* 1996). In this preliminary study, we reported identification of "*Candidatus* Entotheonella sp." in an Indonesian marine sponge from Kapoposang Island, South Sulawesi. A targeting primer pair previously reported by Wilson *et al* (2014) was applied in this work to detect such as-yet uncultivable symbiont in the complex sponge's microbiome. We subsequently reported the presence of a PKS system in a highly complex sponge-microbe symbiotic interaction. The taxonomic status of the Indonesian sponge specimen was investigated in this work by integrated morphological, histological, and DNA analyses. Finding of KS-encoding sequences that belong to PKS systems is expected to become a basis for isolating entire gene clusters encoding the biosynthesis of rare novel pharmacology-relevant polyketides.

## 2. Materials and Methods

### 2.1. Sponge collection and microbiome preparation

A sponge specimen was collected from reefs at the depth range of 5–10 m in Kapoposang Island, Indonesia on September 2015 through Coral Triangle Initiative (CTI) program. Seawater (1 L) was sterilized by filtration on 0.45-μm and 0.22-μm membranes. To prepare *Entotheonella*-harboring microbiome, the sponge sample (10 gram) was sliced into small pieces and stored in 70% ethanol diluted with such sterile seawater. A small sample piece (2.5 grams) was squeezed with 10 mL of sterile seawater in a 50-ml falcon tube, mixed by vortex, and let at room temperature for a few minutes. The resulting supernatant was transferred to five 2-ml eppendorf vials and centrifuged at 500 rpm for 10 minutes using Microfuge 222R (Beckman Coulter). The microbial cell pellet in each vial was resuspended with 1 mL of $ddH_2O$ and stored at −20°C for being used later in PCR detection of *Entotheonella* and cloning of KS-encoding fragments. The resulting precipitate (microbial cell-free sponge tissue) was rinsed 3× with $ddH_2O$ for being used in DNA barcoding.

### 2.2. Sponge identification

Sponge identification in this work involved morphological observation, histological analysis, and DNA barcoding. Histological analysis of the sponge tissue was conducted using light microscopy (Hooper 2003) which consisted of spicule preparation and section preparation. Small fragments of the tissue from inner and outer sides were placed into a 10-ml bottle. A small portion of bleaching agent containing sodium hypochlorite was added to the fragment and waited for a short period to dissolve the organic tissues, leaving only mineral skeleton (spicules). Then, it was washed by replacing the bleaching agent with water for several times. Next, clean spicule suspensions were sucked and pipetted onto an object glass. Finally, it was covered using cover glass carefully to keep the spicule in their original shapes. Sponge was cut in a cross section and then sliced about 1-mm thick. Then the thin slice was placed onto an object glass. To make it observable, several small drops of water were added to the slice and then covered with the cover glass. For sponge identification based on DNA barcoding, genomic DNA was prepared from the squeezed sponge tissue using CTAB method (Piel *et al.* 2004) and then used as the template for PCR amplification using two primer pairs targeting the 5′ end of the mitochondrial cytochrome oxidase subunit 1 (COI): dgLCO1490 (5′-GGTCAACAAATCATAAAGAYATYGG-3′) and dgHCO2198 (5′-TAAACTTCAGGGTGACCAAARAAYCA-3′) (Meyer *et al.* 2005). The PCR master mix consisted of 12.5-μL of 2× KAPA Taq Extra HotStart ReadyMix with dye containing 2-mM $MgCl_2$, 1.25-μL DMSO, 1.5-μL of 10-mM dgLCO1490 primer, 1.5-μL of 10-mM dgHCO2198, 1-μL of sponge DNA, and 7.25-μL $ddH_2O$. The PCR program was set up at 35 cycles consisting of denaturation at 95°C for 3 minutes, denaturation at 95°C for 30 seconds, annealing at 46.2°C for 30 seconds, elongation at 72°C for 1 minute, and final elongation at 72°C for 1 minute. The target PCR product of approximately 659 bp was separated on the

1% agarose gel, purified from the gel, and subjected to DNA sequencing.

### 2.3. *Entotheonella* detection

The presence of "*Candidatus* Entotheonella sp." in the sponge's microbiome was detected using the primer pair 735F (5′-GYAT-TAAGCCKYGGAAACKGT-3′) and 1290R (5′-GCCCRGCWY-VACCCGGTA-3′) (Wilson *et al.* 2014). The PCR mixture was set up for five reactions on ice. Each PCR reaction comprised 12.5 μL of 2× Maxima Hot Start Green PCR Master Mix (Thermo Scientific) and 1 μL of each primer (100 mM). Subsequently, 1 μL of the cell suspension mentioned above was added into individual PCR reactions, except for the negative control (no DNA template was added). The PCR program was set up for 35 cycles on T Professional Thermocycler (Biometra), consisting of pre-denaturation at 95°C for 5 minutes, denaturation at 95°C for 1 minute, annealing at 55.9°C for 1 minute 15 seconds, extension at 72°C for 1 minute, and final extension at 72°C for 7 minutes. PCR product was visualized on 1% agarose gel stained with ethidium bromide. Target product of approximately 650 bp was purified from the gel using GeneJET Gel Extraction and DNA Cleanup Micro Kit (Thermo Scientific) and automatically sequenced.

### 2.4. Cloning of DNA regions encoding ketosynthase (KS) domains

Partial KS-encoding DNA regions were PCR-amplified from the bacterial cell fraction mentioned above using the primer set KSDPQQF (5′-MGNGARGCNNWNSMNATGGAYCCNCARCANMG-3′) and KSHGTGR (5′-GGRTCNCCNARNSWNGTNCCNGTNCCRTG-3′) designed on the basis of the KS motifs DPQQ and HGTG as reported previously by Piel (2002). As the positive control, one PCR reaction was added with the genomic DNA of *Streptomyces* sp. 677. The PCR composition and condition are the same as described in the Section 2.3 above, except the annealing temperature was set up at 60.1°C. The PCR products were separated on 1% agarose gel in the electrophoresis (Power Pac Basic, Bio-Rad). Target fragments of approximately 700 bp were extracted from the gel using GeneJET Gel Extraction and DNA Cleanup Micro Kit (Thermo Scientific). The extracted DNA solution was concentrated to 7 μL by evaporator, followed with ligation with pGEM-T Easy (Promega). The ligation product was transformed into electrocompetent cells of *Escherichia coli* DH5α using the electroporator Micropulser (BioRad), added with 1-mL LB and incubated at 37°C for 1 hour in a shaking incubator. The transformed cell suspension was spread on LB plates containing 100 μg/mL ampicillin previously supplemented with 40 μL of 0.1-M isopropyl β-D-1-thiogalactopyranoside (IPTG) and 40 μL of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal). After overnight incubation at 37°C, the resulting white colonies were picked up, transferred to LB agar plates, 96-well microplates, and tubes containing 5-ml LB, followed by overnight incubation. The 5-ml cultures were individually transferred to 1.5-ml eppendorf tubes and subjected to preparation of plasmids using GeneJet Plasmid Miniprep Kit (Thermo Scientific) and the standard mini-preparation protocol described by Sambrook and Russel (2001). To check the presence of inserts, the plasmid samples were individually cut with *Eco*RI. Plasmids harboring inserts were then individually digested with *Bam*HI to know the restriction pattern. Recombinant plasmids with unique restriction patterns were subjected to DNA sequencing.

### 2.5. Screening of *E. coli* clones harboring *trans*-AT ketosynthase

To screen *E. coli* clones harboring KS fragments of *trans*-AT PKS, a single degenerate primer was designed in this work based on the motif ANRVSYYCDF that is unique and highly conserved among *trans*-AT KS sequences. This degenerate primer, designated as transKSANRF (5′-GCNAAYMGNGTNWSNTA-3′), was coupled with the reverse primer KSHGTGR mentioned above to generate a target PCR product of approximately 500 bp. Of the approximately 200 clones growing on plates, 68 clones on two plates were individually screened by PCR using such *trans*-AT PKS-targeting primer. To simplify the screening process, 30 colonies on each plate were pooled into five pools, in which each pool contained six colonies mixed well in 30-μL dd$H_2O$. Aliquot (1 μL) from each pool was subjected to PCR amplification. Subsequently, colonies from positive pools were individually screened by PCR. Positive colonies were each inoculated into 5-ml LB supplemented with 100 μg/mL ampicillin. The recombinant plasmids were recovered from the overnight cultures and digested with *Eco*RI to check the presence of inserts. The clone inserts (around 700 bp) were sequenced using T7 primer.

### 2.6. DNA sequencing and phylogenetic analysis

DNA sequencing was conducted using the BigDye Terminator Cycle sequencing kit on an ABI Prism 3700 DNA analyzer (Applied Biosystems, USA) according to the dideoxy mediated chain termination method (Sanger *et al.* 1977). The primer T7 (5′-GTAA-TACGACTCACTATAGGG-3′) that recognizes the cloning vector was used to sequence the insert of individual positive clones. PCR products were individually sequenced using the same primers employed for their PCR amplification. The resulting bacterial DNA signature (partial 16S rRNA gene) sequence was subjected to BLASTn analysis to predict its taxonomic affiliation. Sponge DNA signature (partial CO1 gene) sequence obtained in this work was compared with sequences in the NCBI porifera barcoding database. All of the KS-encoding DNA sequences found in this work were translated into amino acid sequences using the Web-based translation tool ExPASy (http://www.expasy.org/tools/dna.html), and subsequently aligned with representative KS sequences from fatty acid, *cis*-AT, and *trans*-AT polyketide pathways retrieved from the GenBank (http://www.ncbi.nlm.nih.gov/Entrez/) in BioEdit (Hall 1999) using ClustalW program (Thompson *et al.* 1994). The alignment result was transferred to Molecular Evolutionary Genetic Analysis, version 7 (MEGA7) software (Kumar *et al.* 2016), and a phylogenetic tree was constructed using the neighbor-joining method (Saitou and Nei 1987), which was evaluated based on Felsenstein's bootstrap test (Felsenstein 1985). Boot-strapping values inferred from 1000 replicates were set up on each tree branch to estimate the reliability of phylogenetic reconstruction. Based on this phylogenetic tree, the affiliation of KS sequences was predicted.

## 3. Results

"*Candidatus* Entotheonella sp." has been proposed as a new candidate phylum called "Tectomicrobia" to reflect their uncultured status and capability to produce bioactive compounds (Wilson *et al.* 2014). We reported here our studies on identifying "*Candidatus* Entotheonella sp." and PKS systems in highly complex bacterial consortia of a marine sponge collected from Kapoposang Island, South Sulawesi. These include detection of *Entotheonella*, identification of sponge specimen, cloning, and analysis of KS-encoding regions in PKS genes.

### 3.1. Identifying *Entotheonella* in the uncultured microbiome of a marine sponge

To identify *Entotheonella* in the sponge-bacteria associations, we initially separated microbial cells from the sponge tissue by simple squeezing in filter-sterilized seawater. The microbial cell suspension obtained was subjected to the PCR amplification to detect the

presence of *Entotheonella* using a targeting primer pair described by Wilson *et al* (2014). PCR product of approximately 650 bp was separated on 1% agarose gel, purified, and subjected to DNA sequencing. The sequencing result (deposited with accession number KX120255.1 (https://www.ncbi.nlm.nih.gov/nuccore/KX120255) in the GenBank) indicated its high homology (97% identity) with the 16S rRNA gene of "*Candidatus* Entotheonella sp." clone TSY1 (accession no. KF926817.1), an environmental taxon proposed as candidate phylum "Tectomicrobia" previously reported by Wilson *et al* (2014).

### 3.2. Sponge identification by the integrated morphological, histological, and COI gene analyses

Due to the importance of the sponge specimen investigated in this work from the medical point of view, this specimen was subsequently identified at the species level by integrating morphological annotation, histological analysis, and DNA barcoding. As shown in Figure 1A, its morphology was irregularly massive globular with small osculum lying on the upper surface. The specimen turned dark brown on collection. The exterior color was dark brown and the interior was yellow in alcohol (Figure 1B). Histological analysis showed that megascleres were oxeas with short point and triane which is rare (Figure 1C), and micloscleres were tylaster, oxyspheraster, and spheraster (Figure 1D). These morphological and histological characters indicated its identity as *Rhabdastrella globostellata*. The presence of triane and tylaster and the color in alcohol distinguished it from *Rhodopteriana distincta*.

DNA barcoding used in this work relies on the use of degenerate primers to amplify partial mitochondrial COI gene (Meyer *et al.* 2005). When using total sponge metagenomic DNA as the PCR template, we did not get any PCR product at any amplification condition. To overcome the difficulty in amplifying COI gene from this sponge, we initially removed microbial cells from the sponge tissue through simple squeezing followed with centrifugation and rinsing. We then extracted DNA from the sponge tissue that was relatively free of microbial cells. Using this sponge DNA as the PCR template, we successfully amplified a target PCR product of approximately 659 bp in length at the annealing temperature of 46.2°C using the COI gen-targeting primer set dglCO1490 and dghCO2198. Further sequencing of this PCR product and sequence analysis by BlastX against the NCBI Porifera database showed that the amplified region belongs to a partial COI gene that shared homology (51% identity and 66% similarity) with that of *Rhabdastrella globostellata* (Accession number ALD10374.1). Based on the integrated morphological, tissue, and DNA analyses, we propose the taxonomic status of the sponge specimen studied in this work as *Rhabdastrella* sp.

### 3.3. Cloning of ketosynthase-encoding DNA regions

The presence of PKS system in *Rhabdastrella* sp. was identified in this work through PCR amplification, cloning and sequencing of the KS domain-encoding DNA regions (Piel 2002). Metagenomic survey of PKS diversity in the sponge's microbiome was based on KS sequences amplified using universal KS-targeting primers (Piel 2002) to generate PCR product of approximately 700 bp in size (Figure 2A). This target PCR product was then purified from the gel and cloned into *E. coli*. Subsequently, 12 of the resulting transformed colonies growing on the selective media were randomly selected from the growing plate and subjected to insert sequencing (Figure 2B).

BLASTx analysis (Altschul *et al.* 1990) indicated that the protein sequences of the 11 clones exhibited high similarities (70%—80%) to the KS domains in a single SupA module (Table 1) (Fieseler *et al.* 2007). This was supported by phylogenetic analysis showing their affiliation into the clade of animal type I PKS-like FASs (Fieseler *et al.* 2007; Schweizer and Hofmann 2004), suggesting the possible involvement of the isolated KS sequences in fatty acid biosynthesis. To specifically target type I PKS systems, we subsequently designed a targeting degenerate primer that specifically recognizes KS domains that belong to type I PKSs and then employed this primer to screen the PCR clone library. To do this, we initially aligned all of the KS sequences of FA pathway obtained in this work with several representative KS sequences known to belong to *cis*-AT and *trans*-AT PKSs retrieved from the GenBank database. The sequence alignment result enabled us to identify a highly conserved motif that might be unique in KS domains of *trans*-AT PKS systems. This unique *trans*-AT KS motif was then used a basis to generate a targeting degenerate primer (Figure 3). We then coupled this targeting primer with one of the existing universal KS to screen our PCR clone library. The reason for selecting a unique motif of KS sequences that belong to *trans*-AT PKS as the target region is due to the significant importance of these unusual PKS systems from medical point of view (Piel 2010). Some notable

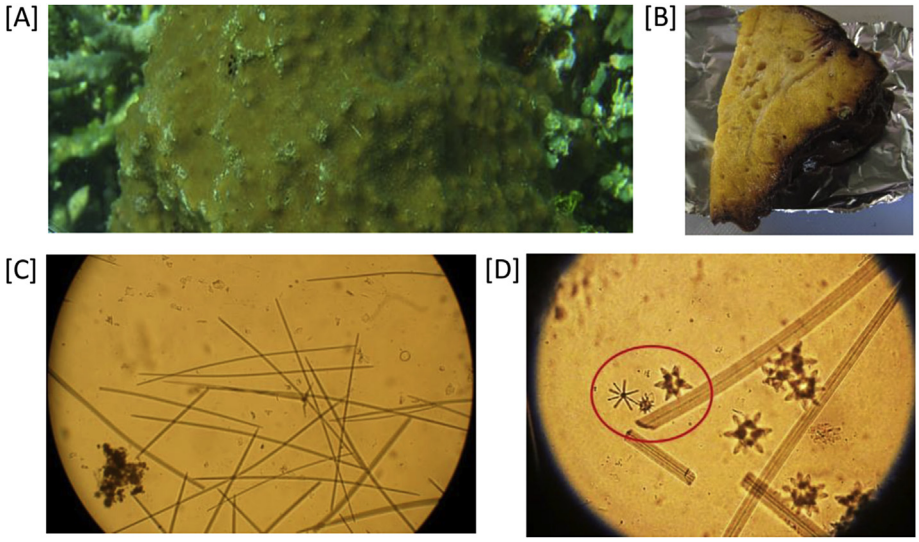



**Figure 1.** Identification of a sponge specimen investigated in this work. Morphology of the specimen underwater [A] and upon collection [B]. Histological analysis showing megascleres [C] and microscleres [D].

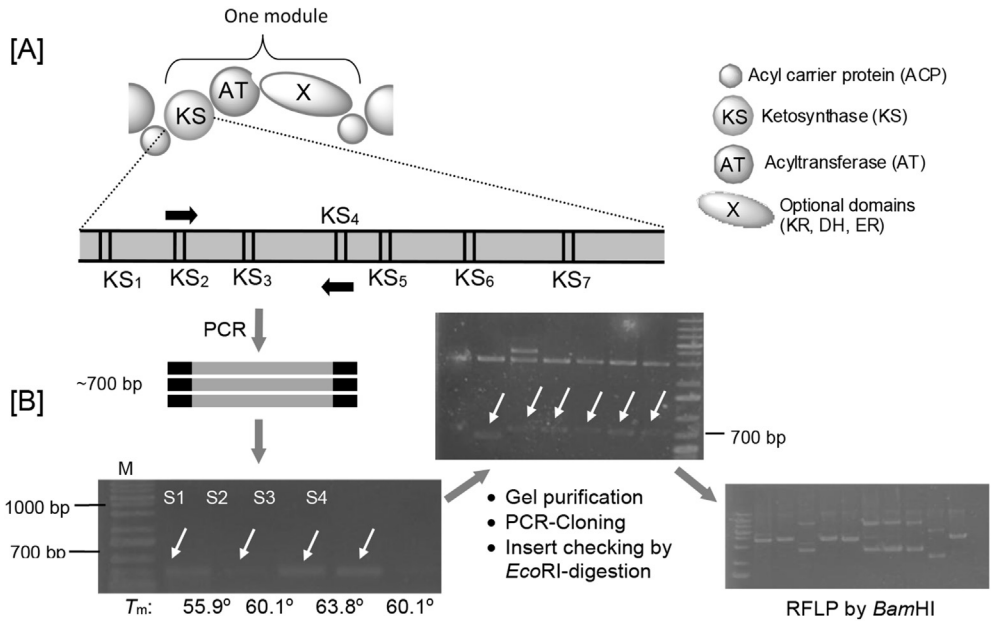



**Figure 2.** PCR-cloning of KS sequences from the uncultivated filamentous bacterial fraction of the Kapoposang marine sponge, *Rhabdastrella* sp. **[A]** PCR detection of 700-bp DNA fragments at three different $T_m$s. Highly conserved motifs of KS domains: $KS_1$ = EPIAIV, $KS_2$ = DPQQR, $KS_3$ = CSSS, $KS_4$ = HGTGTxLGD, $KS_5$ = NIGH, $KS_6$ = GxGGxNAHVILEE, $KS_7$ = TYPFARER. **[B]** The target PCR product was cloned into *Escherichia coli* using pGEM-T Easy (Promega). Some resulting white clones were subjected to digestion with *Eco*RI to check the presence of inserts and *Bam*HI to know the difference of restriction pattern among clones containing an insert. Clones with unique restriction pattern (12 clones) were sequenced (shown with white arrows). M = marker, S1–S3 = Samples of microbial cells at different $T_m$S.

Table 1. Predicted function of KS sequences cloned from the sponge *Rhabdastrella* sp. in this work

| Clone | AA | Protein homolog (BLASTX) | | | I/S (%) | Group |
|---|---|---|---|---|---|---|
| | | Homolog | Organism | Accession no. | | |
| RGK19 | 717 | SupA | Uncultured bacterial symbiont of *Aplysina aerophoba* | ABE03915.1 | 82/72 | FAS |
| RGK23 | 706 | SupA | Uncultured bacterial symbiont of *A. aerophoba* | ABE03915.1 | 81/71 | FAS |
| RGK24 | 717 | SupA | Uncultured bacterial symbiont of *Theonella swinhoei* | ABE03935.1 | 95/87 | FAS |
| RGK33 | 715 | SA1_PKSC | Uncultured bacterial symbiont of *Discodermia dissoluta* | AAY00027.1 | 67/58 | FAS |
| RGK43 | 717 | SupA | Bacterial symbiont of *T. swinhoei* | ABE03935.1 | 94/87 | FAS |
| RGK55 | 769 | SupA | Bacterial symbiont of *T. swinhoei* | ABE03935.1 | 74/67 | FAS |
| RGK60 | 717 | SupA | Bacterial symbiont of *T. swinhoei* | ABE03935.1 | 92/85 | FAS |
| RGK158 | 717 | SupA | Bacterial symbiont of *A. aerophoba* | ABE03915.1 | 83/66 | FAS |
| RGK161 | 713 | SupA | Bacterial symbiont of *T. swinhoei* | ABE03935.1 | 94/87 | FAS |
| RGK163 | 717 | SupA | Bacterial symbiont of *T. swinhoei* | ABE03935.1 | 92/85 | FAS |

Note: AA = amino acids; I/S = identity/similarity; FAS = fatty acid synthases; SupA is a single module (KS-AT-DH-MT-ER-KR-ACP-KS-AT) that generates methyl-branched fatty acids (Hochmuth *et al.* 2010).

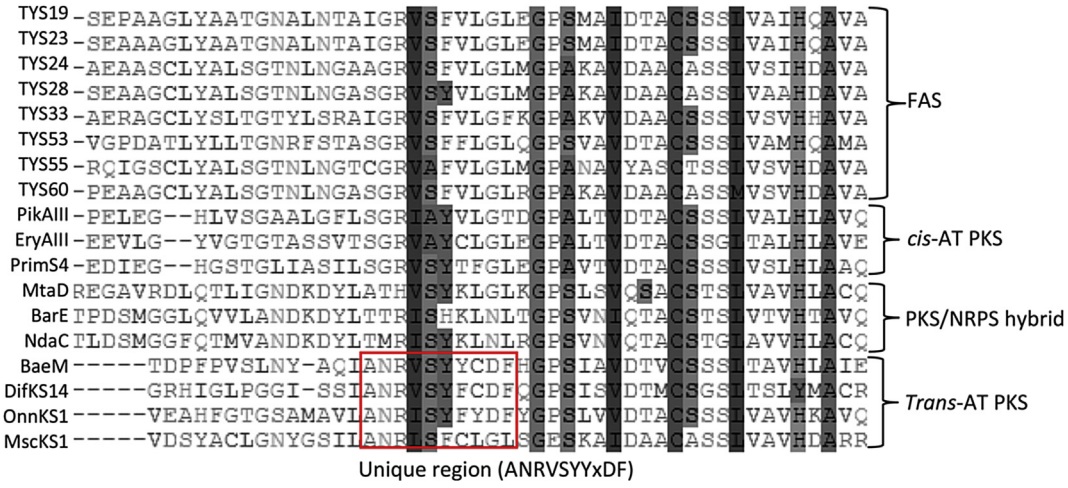



**Figure 3.** Designing a primer targeting ketosynthase (KS)-encoding DNA regions of *trans*-AT PKS. KS sequences putatively involved in the production of methyl-branched fatty acids included TYS19, TYS23, TYS24, TYS28, TYS33, TYS53, and TYS55. KS sequences retrieved from GenBank are those in the PKS pathways of pikromycin (PikAIII), erythromycin A (EryAIII), pimaricin (PimS4), myxothiazol (MtaD), barbamide (BarE), nodularin (NdaC), bacillaene (BaeKS11), difficidin (DifKS14), onnamide (OnnKS11), and misakinolide A (MscKS1). KS = ketosynthase; PKS = polyketide synthase.

examples of pharmacologically important sponge-derived compounds in which their biosynthesis are directed by *trans*-AT PKS systems include onnamides (Nguyen *et al.* 2008; Piel *et al.* 2004), psymberin (Fisch *et al.* 2009), and misakinolide A (Ueoka *et al.* 2015).

We screened more than 300 clones in our PCR-amplicon library by combining colony pooling and whole-cell PCR analysis. We isolated four positive single colonies based on PCR detection using the *trans*-AT targeting primer. The recombinant plasmids from individual positive clone were recovered, analyzed by enzymatic restriction, and subjected to DNA sequencing. Sequence BLASTx analysis indicated that one of the four, designated as RGK32 shared high similarities with KS domains likely associated with many type I PKSs (64%–71% identity), suggesting its possible role in directing the biosynthesis of a pharmacologically relevant polyketide. However, BLASTx analysis of RGK32 showed no high similarity with KS sequences from any characterized polyketide pathways. We therefore analyzed RGK32 further by PSI-BLAST. It was found that RGK32 shares high similarity (76%–80%) and identity (63%–69%) with CrpB of the cryptophycin PKS/NRPS pathway in *Microcystis aeruginosa* PCC 9809, PpsA of the phthiocerol PKS pathway in *Mycobacterium tuberculosis*, and MicA protein of microginin synthase in *Planktothrix prolifica* NIVA-CYA 98 (Rounge *et al.* 2009).

To specify the affiliation of RGK32 obtained in this work, we performed phylogenetic analysis by including KS-representatives from *cis*-AT PKS, *trans*-AT PKS, and FAS biosynthetic pathways. Eleven different KS sequences (RGK19, RGK33, RGK24, RGK28, RGK33, RGK43, RGK53, and RGK58) obtained in this work were also included in the phylogenetic analysis. The results indicated the affiliation of such seven KS sequences into the same clade as those known to belong to characterized FAS modules such as SA1_PKSA, SA1_PKSB, and SupA. Interestingly, RGK32 is clustered into the *cis*-AT PKS modules such as StiG, JerC, EryA3, PikA3, and PimS4 (Figure 5). These results suggest that we have obtained a KS sequence that may involve in the biosynthesis of novel pharmacologically relevant polyketide.

### 3.4. Accession numbers

The partial 16S rRNA gene sequence obtained in this work using an *Entotheonella*-targeting primer set was deposited in the GenBank with an accession number KX120255 (https://www.ncbi.nlm.nih.gov/nuccore/KX120255). Partial mitochondrial COI gene obtained in this work has been submitted to the GenBank with an accession number KY552667 (https://www.ncbi.nlm.nih.gov/nuccore/KY552667). All of the KS sequence data obtained in this work were deposited in the GenBank database under accession numbers KX099902 to KX099925 (https://www.ncbi.nlm.nih.gov/nuccore/?term=A.R.+Uria).

## 4. Discussion

Sponges are well known as chemically rich sources of polyketides or hybrids with peptides (Blunt *et al.* 2007; Fusetani and Matsunaga 1993). There has been emerging evidence that sponge-derived polyketides and modified peptides are synthesized by the sponge-associated symbiotic bacteria. This is best exemplified by onnamide A, keramamides, and misakinolide A that have previously proved to be produced by "*Candidatus* Entotheonella sp.", an uncultivable as-of-yet filamentous symbiont of the Japanese sponge *T. swinhoei* (Piel *et al.* 2004; Ueoka *et al.* 2015; Wilson

```
LnmJ4      MDPQQRLFLQTAWRVFEDAGYRPADLAGAPCGLFVGVATHDYDDLLKENGVAVQAHTATG
PpsA       MDPQQRLLLEVVHEALESAGIPADTLAETRTGVFAGASAGDYAQLGASDLSQVDAWYGTG
MicA       LDPQQRLLLEVAWEALENAAIVPETLVGSKSGVFIGISDVDYHRLAYQSPPNLTAYVGTG
CrpB       MDPQQRLLLEVSWEALENACIAPETLAGSQTGVFVGISSDDHARLLSKDNESIGTYYGTG
RGK32      MDPQHRLLLEVAWEAFEDAGIPPETWSGAQTGVYVGISSSDYAQLQSAQK--TSVYSGTG
SA1_PKSA   LDPQQRLLLETSWQALEDAGIPPVLLRGSRTGVYAGITSIDYAELIASEQNGTGLYVATG

LnmJ4      IAHSVLANRVSYLFDLNGPSEAVDTACSSSLVAIHRALRAIQDGECELAVAGGVNVILTP
PpsA       GSISIIANRVSYFFDFRGPSVTIDTACSSSLVAIHLACQSLRTGDSDVALAAGVNLLLSP
MicA       NSTSIAANRLSYVFDLRGPSLALDTACSSSLVAVHLACQSLQNQESNLCLVGGVNLILSP
CrpB       NAFCVAANRLSYFLDFHGPSLAIDTACSSSLVAVHEACKSLTDGECHLALAAGVNLLLSP
RGK32      NAHSIAANRLSYVFDFRGPSVAVDTACSSSLVAVHLATQSLQSGECDQAIAAGVNLLLSP
SA1_PKSA   NSPSTAIGRVAFTLGLEGPAMAVDTACSSSLVAVHQAVASLQRGETDLALAGGVNAILTP

LnmJ4      GLLESFTQSGMLSPDGRCKTFDADADGYVRGEGVGAVLLKPLARAEADGDHIYAVVKGTA
PpsA       AGTRSLDQADAMSKTGQCHAFDAAADGFVRGEGCGVAVLKRLSDAQRDGDRILAVIRGSA
MicA       ETTIVFSQARMMAPDSRCKTFDASADGYVRSEGCGMVVLKRLRDAIEDGDRVLAVIKGSA
CrpB       QLTINFSKAGMLAADGRCKTFDESANGYVRGEGCGVVILKRLEKAIQDGDRIYAIIRGSA
RGK32      ALTETFQQAGMLAPDGHCKTFDADADGYVRGEGCGVVILKRLSLAQRDGDKIWGVIRGTA
SA1_PKSA   TLTESFASGGMLAPDGRCKTFDAAADGYVRGEGCGMVVLKRLADAEADGDRIWGVIRGTA

LnmJ4      VNHGGRSNSLTAPNPESQARVVAAAVREAGVEPDTITYIEAHGTGT
PpsA       VNQDGRSNGLMAPNPSAQMAVLRAAYATAGVDPREVDYVEAHGTGT
MicA       VNQDGLSNGLTAPNGPAQQAVIRQALENAQVEPAQISYIEAHGTGT
CrpB       VNQDGHSNGLTAPNKQAQQAVIKKALAKAQVSPKDISYVEAHGTGT
RGK32      INQDGRSNGLTAPNSLAQQAVIKAAIAKAGIQPSQVSYVEAHGTGT
SA1_PKSA   VNQDGASAGLTVPNGPAQERVIAEALARAGIQPAEVDYLEAHGTGT
```

**Figure 4.** Sequence analysis of RGK32 (gray highlight) identified in the PCR-amplicon library generated from *Rhabdastrella sp.* microbiome. Multiple alignment between RGK32 and KS sequences from known polyketide pathways. NaPDos analysis (Ziemert *et al.* 2012) showed that RGK32 shared 57% identity with the LnmJ4 KS sequence of leinamycin pathway. PSI-BLAST analysis showed that RGK32 shared 69/80% (I/S) with CrpB in *Microcystis aeruginosa* PCC 9809 (acc. no. CCI28338.1), 65/76% with PpsA in *Mycobacterium tuberculosis* (acc. no. CKP05251.1), 65/78% with MicA protein in *Planktothrix prolifica* NIVA-CYA 98 (acc. no. CAQ48259.1), 63/78% with SA1_PKSA in uncultured bacterial symbiont of *Discodermia dissoluta* (acc. no. AAY00025.1). Highly conserved motifs of KS domains are indicated by yellow highlight. A unique region used as a basis for identifying KS sequences that belong to *trans*-AT PKS are highlighted with green color. PCR = polymerase chain reaction; KS = keto synthase.

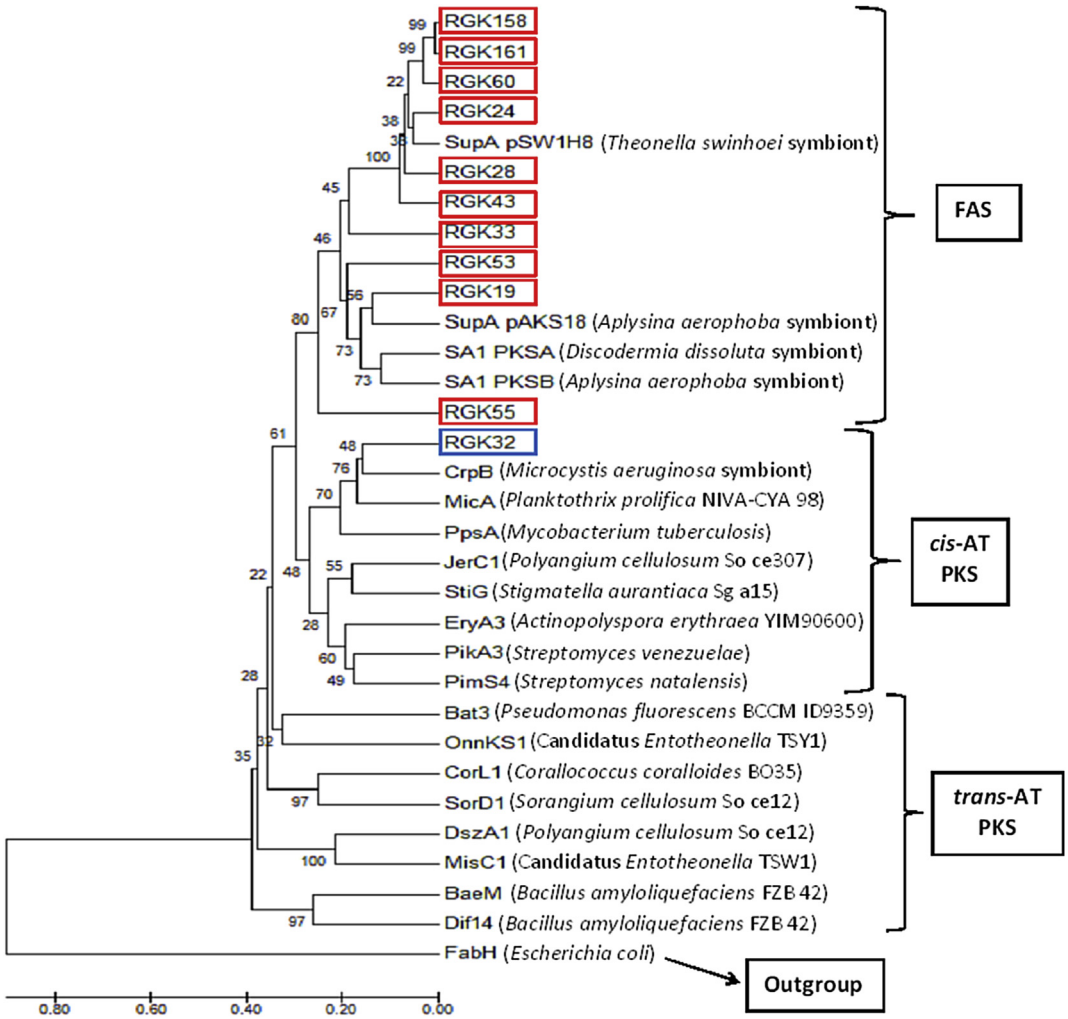



**Figure 5.** Phylogenetic analysis of RGK32 along with other partial KS amplicons (RGKs 19, 24, 28, 33, 43, 53, 55, 60, 158, and 161) cloned from the metagenomic DNA of *Rhabdastrella* sp. This phylogenetic analysis was performed in MEGA7 program (Kumar *et al.* 2016) using the neighbor-joining method (Saitou and Nei 1987). The bootstrap percentages are shown next to the branches (Felsenstein 1985). The boostrap consensus tree was inferred from 1000 replicates. DszA1 of disorazoles pathway, MisC1 of misakinolide A pathway, Bat3 of batumin pathway, BaeM of bacillaene, Dif14 of difficidin pathway, CorL of corallopyronin A pathway, SorD of sorangicin pathway, StiG of stigmatellin pathway, JerC of jerangolid pathway, EryA3 of erythromycin, PikA3 of pikromycin, and PimS4 of pimaricin pathway. The FAS component FabH from *Escherichia coli* was the outgroup. FAS = fatty acid synthase; AT = acyltransferase; PKS = polyketide synthase.

*et al.* 2014). Using *Entotheonella*-targeting primers (Wilson *et al.* 2014), we detected the presence of "*Candidatus* Entotheonella sp." in an Indonesian marine sponge, suggesting their wide distribution in taxonomically diverse sponge species (Wilson *et al* 2014). "*Candidatus* Entotheonella sp." is new bacterial taxon previously proven as the true producer of almost all bioactive polyketides and peptides known from the Japanese *T. swinhoei* chemotype Y (Wilson *et al.* 2014). The presence of another *Entotheonella* variant in this Indonesian sponge specimen has motivated us to determine the sponge taxonomic status and to identify the PKS genes.

We subsequently determined the taxonomic status of the Indonesian sponge harboring *Entotheonella* through the integrated morphological, histological analyses and DNA barcoding. DNA barcoding has emerged as a tool to speed sponge identification up to the species level (Wörheide and Erpenbeck 2007) since the sponge barcoding project (www.spongebarcoding.org) was established (Wörheide *et al.* 2008). Mitochondrial DNA was used as the basis in sponge identification due to its very divergent patterns of sequences across among different demosponges (Redmond *et al.* 2011). Sponge DNA barcoding in this work was directed to amplify and sequence mitochondrial CO1 gene. However, attempts to amplify CO1 gene from the total sponge DNA failed at any PCR conditions, probably due to the complexity of sponge metagenome consisting of genomes derived from highly diverse microorganisms in association with this sponge. Therefore, preferential amplification of CO1 sequences from non-target microorganisms represents a main bottleneck in sponge DNA barcoding (Siddall *et al.* 2009). For example, DNA barcoding of 96 sponge families reported by Vagas *et al* (2012) showed that sequences from non-target organisms in 40% of the sequenced samples. To tackle the difficulty in the PCR amplification of CO1 gene, we initially removed microbial cells from the sponge tissue sample and subsequently extracted DNA from sponge tissue that is relatively free of microbial cells. Using this sponge DNA as the PCR template, we observed a target PCR product of approximately 659 bp on 1% agarose gel. Further sequencing and analysis of the 659-bp amplicon showed its high similarity (51% identity and 66% similarity) with that of *Rhabdastrella globostellata* (accession number ALD10374.1). This data highlighted the importance of removing the associated microbiome before PCR amplification to increase the success rate of identifying COI gene in the sponge samples. Combining morphological and histological analyses and DNA barcoding strongly suggests that the

identity of the sponge specimen investigated in this work as *Rhabdastrella* sp.

The extraordinary metabolic potential of "*Candidatus* Entotheonella sp." as reported previously by Wilson *et al* (2014) has motivated us to identify PKS genes in other sponge species from Indonesia. However, identification for polyketide biosynthetic pathways in sponges is extremely challenging due to the high complexity of the associated microbial assemblages. Bacterial biomass may be in the range of $6.4 \times 10^8$ to $1.5 \times 10^9$ bacterial cells just in 1 mL of sponge extract (Scheuermayer *et al.* 2006) and consists of hundreds to thousands of microbial species (Taylor *et al.* 2007). The high microbial complexity in a sponge system suggests that the symbiotic assemblage contain numerous homologous genes from diverse pathways (Piel 2010). Relatively low percentage of genes directing polyketide biosynthesis in comparison with those involved in fatty acid biosynthesis makes identification of PKSs become more difficult. In our recent study, we have made attempts to identify the polyketide biosynthetic systems in *R. globostellata* containing *Entotheonella*. This sponge species typically contains isomalabaricane-type compounds, a member of triterpenes exhibiting significant inhibition of the growth of tumor cells (Fouad *et al.* 2006; Lv *et al.* 2004) and are potentially used as chemotaxonomic markers of *Rhabdastrella* (Tasdemir *et al.* 2002). To the best of our knowledge, there are no reports on any polyketide isolated from *Rhabdastrella* suggesting that identifying the polyketide biosynthetic machinery called type I PKS in the sponge species may provide a basis for the discovery of new polyketide biosynthetic pathway(s).

PCR identification of type I PKS systems in our present study was based on the PKS KS domain. A universal KS primer pair designed based on the KS conserved motifs DPQQ and HGTGT (Piel 2002) was employed to obtain a target PCR product of approximately 700 bp. Subsequent cloning of the correct-sized PCR product followed with clone sequencing and phylogenetic analysis led to the isolation of partial KS domains that are highly similar to those in a single SupA module (Fieseler *et al.* 2007). The term Sup referred to "sponge symbiont ubiquitous PKS," an unusual type of small FAS-like PKS present in the sponge symbiont group "Poribacteria" (Siegl *et al.* 2011). Sup module is comprised of KS, AT, DH, MT, ER, KR, and ACP often with additional KS and inactive AT domains in the C-terminus (Hochmuth *et al.* 2010). The biosynthetic product of this module is methyl-branched FAs that are a result of the full reduction of the polyketide backbone by KR-DH-ER domains followed by methylation by MT domain (Fieseler *et al.* 2007; Hochmuth *et al.* 2010). Functional study of *sup* genes previously conducted by Hochmuth *et al* (2010) revealed that SupA uses malonyl-CoA units as building blocks to generate its product methyl-branched FASs. The chain lengths of sponge FAs encoded on *sup* genes are usually in the range of 14 to 24 carbon atoms with methyl branches at all positions between C2 and C17 (exception at the positions 4, 5, and 13) (Hochmuth *et al.* 2010).

An interesting question arises about the biotechnological potential of methyl-branched FAs synthesized by Sup proteins because many lipids of marine sponges, including saturated FAs, exhibited antimicrobial activities (Desbois 2012; Mishra *et al.* 2015). Among various linear-chain saturated FAs, myristic acid exhibits the highest bactericidal activity against the highly virulent pathogenic strain *Mycobacterium bovis* and the avirulent strain *M. tuberculosis* (Kondo and Kanal 1977). Sponge-derived FAs, particularly mid-chain-branched FAs (MBFAs), have often been proposed to be produced by symbiotic bacteria (Thiel *et al.* 1999). This was supported by the analyses of *sup* genes which shows that members of the candidate phylum "Poribacteria" are the important producers of sponge MBFAs (Hochmuth *et al.* 2010; Siegl *et al.* 2011). Sponge-derived saturated FAs can potentially show interesting biological activities and therefore, it might be interesting to express *sup* genes and subsequently test the recombinantly produced FAs for antimicrobial activities. Based on the biotechnological potential of FAs described above, information on the KS sequences obtained in this work can potentially be used as a basis for isolating the corresponding FAS gene clusters. Another potential application is that FA biosynthesis integrated with chain termination enzymes can potentially be used to make a wide variety of free long-chain FA molecules with great potential for being used as FA-based biofuels (Lennen and Pfleger 2013).

The results above suggest that it is extremely challenging to identify KS sequences that belong to type I PKS systems in highly complex microbiome of a marine sponge. The difficulty to identify type I PKSs in sponges was also previously reported for *Psammocinia aff. Bulbosa* and *Cacospongia mycofijiensis*. The sequencing of 81 KS amplicons and 276 KS amplicons from both sponge species, respectively, indicated that none of the sequenced amplicons predicted to involve in the biosynthesis of complex polyketides (Fieseler *et al.* 2007; Kim and Fuerst 2006; Schirmer *et al.* 2005). To encounter the challenge in this work, we screened the PCR-amplicon library of approximately 300 clones generated in this work using a targeting PCR primer pair (Figure 3). This screening led to the isolation of four positive clones. BlastX analysis indicated that one of them, designated as RGK32, shared high similarity with many KS sequences from many type I PKSs.

Based on PSI-BLAST analysis, RGK32 shared high similarity (76%–80%) and identity (63%–69%) with CrpB of the cryptophycin PKS/NRPS pathway in *M. aeruginosa* PCC 9809, PpsA of the phthiocerol PKS pathway in *M. tuberculosis*, and MicA protein of microginin synthase in *P. prolifica* NIVA-CYA 98 (Rounge *et al.* 2009). The CrpB protein consists of two *cis*-AT PKS modules (KS-AT-KR-ACP-KS-AT-DH-MT-KR-ACP) that accepts a phenylactetate-like starter unit in the biosynthetic formation of the cryptophycin's fragment A (Magarvey *et al.* 2006). The PpsA protein consists of a single *cis*-AT PKS (PCP-KS-AT-KR-ACP) as a part of phthiocerol pathway that accepted a long fatty acid of *n*-$C_{16}$-$C_{28}$ as the precursor (Gokhale *et al.* 2007). MicA protein contains a *cis*-AT PKS module (KS-AT-ACP) that accepted the precursor octanoic acid in microginin biosynthesis (Kremer 2006). The high similarity of RGK32 with KS sequences from modules with integrated AT suggests its possible classification into *cis*-AT PKS system. Further multiple alignments, as shown in Figure 4, indicated that RGK32 contains the unique motif ANRVSYYxDF that is also present in the KS sequences from *cis*-AT PKS modules, as exemplified by CrpB, PpsA, and MicA. However, this unique motif is absent in the KS sequences from FAS modules, as exemplified by SA1_PKSA (Figure 4). This suggests that the unique ANRVSYYxDF motif is not exclusively present in *trans*-AT PKS systems due to its presence in *cis*-AT PKS systems as well. Although the data suggest that RGK32 obtained in this work may belong to a type I PKS; however, at this point, it is still not clear whether it belongs to *cis*-AT PKS or *trans*-AT PKS. Therefore, further phylogenetic analysis should be carried out to confirm its affiliation.

Our phylogenetic analysis shown in Figure 5 clearly classified RGK32 into *cis*-AT PKS system, suggesting the importance of this RGK32 sequence information as a basis for the discovery of a PKS gene cluster encoding a novel *cis*-AT PKS pathway. Based on this RGK32 sequence, specific primers could then be designed and subsequently used to screen a large-insert metagenomic library constructed from the sponge's microbiome. We are currently working on generating a microbial metagenomic library from *R. globostellata*. Therefore, finding of this KS designated as RGK32 is very important because it provides a basis for the discovery of a novel PKS pathway that could potentially be activated in an easily

culturable bacterium to generate a pharmaceutically relevant polyketide in sustainable way.

## Acknowledgement

This work was conducted at the Research Center for Marine and Fisheries Product Processing and Biotechnology (P3DSPB) Jakarta, and financially supported by the Ministry of Marine Affairs and Fisheries, Republic of Indonesia (MMAF) granted to P3DSPB. The authors are grateful to Nurrahmi D. Fajarningsih, Indra H. Januar, Sri Iswani and Nurmaya Legistari for their help during the sponge sample collection and processing. The authors would like to acknowledge Professor Jörn Piel (ETH Zürich) for valuable discussion and Professor Ekowati Chasanah (MMAF) for advice and support. The authors are thankful to Benget Simanjuntak for laboratory assistance.

## References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–10.
Amann RI, Ludwig W, Schleifer K-H. 1995. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol Rev* 59(1): 143–69.
Bewley CA, Holland ND, Faulkner DJ. 1996. Two classes of metabolites from *Theonella swinhoei* are localized in distinct populations of bacterial symbionts. *Experientia* 52:716–22.
Blunt JW, Copp BR, Hu W, Munro MHG, Northcote PT, Prinsep MR. 2007. Marine natural products. *Nat Prod Rep* 24:31–86.
Brück WM, Sennett SH, Pomponi SA, Willenz P, McCarthy PJ. 2008. Identification of the bacterial symbiont *Entotheonella* sp. in the mesohyl of the marine sponge *Discodermia* sp. *ISSM J* 2(3):335–9.
Cane DE, Walsh CT. 1999. The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. *Chem Biol* 6(12):R319–25.
Desbois AP. 2012. Potential applications of antimicrobial fatty acids in medicine, agriculture and other industries. *Recent Pat Antiinfect Drug Discov* 7:111–22.
Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–91.
Fieseler L, Hentschel U, Grozdanov L, Schirmer A, Wen G, Platzer M, Hrvatin S, Butzke D, Zimmermann K, Piel J. 2007. Widespread occurrence and genomic context of unusually small polyketide synthase genes in microbial consortia associated with marine sponges. *Appl Environ Microbiol* 73:2144–55.
Fisch KM, Gurgui C, Heyke N, van der Sar SA, Anderson SA, Webb VL, Taudien S, Platzer M, Rubio BK, Robinson SJ, Crews P, Piel J. 2009. Polyketide assembly lines of uncultivated sponge symbionts from structure-based gene targeting. *Nat Chem Biol* 5(7):494–501.
Fouad M, Edrada RA, Ebel R, Wray V, Müller WEG, Lin WH, Proksch P. 2006. Cytotoxic isomalabaricane triterpenes from the marine sponge *Rhabdastrella globostellata*. *J Nat Prod* 69(2):211–8.
Freeman MF, Gurgui C, Helf MJ, Morinaka BI, Uria AR, Oldham NJ, Sahl H-G, Matsunaga S, Piel J. 2012. Metagenome mining reveals polytheonamides as posttranslationally modified ribosomal peptides. *Science* 338:387–90.
Fusetani N, Matsunaga S. 1993. Bioactive sponge peptides. *Chem Rev* 93:1793–806.
Gokhale RS, Saxena P, Chopra T, Mohanty D. 2007. Versatile polyketide enzymatic machinery for the biosynthesis of complex mycobacterial lipids. *Nat Prod Rep* 24:267–77.
Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis for Windows 95/98/NT. *Nucl Acids Symp Ser* 41:95–8.
Hertweck C. 2009. The biosynthetic logic of polyketide diversity. *Angew Chem Int Ed* 48:4688–716.
Hochmuth T, Niederkrüger H, Gernert C, Siegl A, Taudien S, Platzer M, Crews P, Hentschel U, Piel J. 2010. Linking chemical and microbial diversity in marine sponges: possible role for poribacteria as producers of methyl-branched fatty acids. *ChemBioChem* 11:2572–8.
Hooper JNA. 2003. Sponguide: Guide to Sponge Collection and Identification. Brisbane: Queensland Museum, Brisbane. pp. 26.
Hugenholtz P, Goebel BM, Pace NR. 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol* 180(18): 4765–74.
Kim TK, Fuerst JA. 2006. Diversity of polyketide synthase genes from bacteria associated with the marine sponge *Pseudoceratina clavata*: culture-dependent and culture-independent approaches. *Environ Microbiol* 8:1460–70.
Kondo E, Kanal K. 1977. The relationship between the chemical structure of fatty acids and their mycobactericidal activity. *J Med Sci Biol* 30(4):171–8.
Kremer D. 2006. Microginin producing proteins and nucleic acids encoding a microginin gene cluster as well as methods for creating microginins. European Patent EP20060818955.
Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33(7):1870–4.
Lennen RM, Pfleger BF. 2013. Microbial production of fatty acid-derived fuels and chemicals. *Curr Opin Biotechnol* 24:1044–53.
Llewellyn NM, Spencer JB. 2007. Enzymes line up for assembly. *Nature* 448(16): 755–6.
Lv F, Deng ZW, Li J, Fu HZ, Soest RWM, Proksch P, Lin WH. 2004. Isomalabaricane-type compounds from the marine sponge *Rhabdastrella* aff. distincta. *J Nat Prod* 67:2033–6.
Magarvey NA, Beck ZQ, Golakoti T, Ding Y, Huber U, Hemscheidt TK, Abelson D, Moore RE, Sherman DH. 2006. Biosynthetic characterization and chemoenzymatic assembly of the cryptophycins, potent anticancer agents from cyanobionts. *ACS Chem Biol* 1(12):766–79.
Meyer C, Geller J, Paulay G. 2005. Fine scale endemism on coral reefs: archipelagic differentiation in turbinid gastropods. *Evolution* 59:113–25.
Mishra PM, Sree A, Panda PK. 2015. Fatty acids of marine sponges. In: Kim Se-Kwon (Ed.). *Springer Handbook of Marine Biotechnology*. pp. 851–68.
Nguyen T, Ishida K, Jenke-Kodama H, Dittmann E, Gurgui C, Hochmuth T, Taudien S, Platzer M, Hertweck C, Piel J. 2008. Exploiting the mosaic structure of *trans*-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nat Biotechnol* 26(2):225–33.
Piel J. 2002. A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. *Proc Natl Acad Sci USA* 99: 14002–7.
Piel J. 2004. Metabolites from symbiotic bacteria. *Nat Prod Rep* 21:519–38.
Piel J, Hui D, Wen G, Butzke D, Platzer M, Fusetani N, Matsunaga S. 2004. Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*. *Proc Natl Acad Sci USA* 101:16222–7.
Piel J. 2010. Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat Prod Rep* 27:996–1047.
Rawlings BJ. 2001. Type I polyketide biosynthesis in bacteria (Part B). *Nat Prod Rep* 18:231–81.
Redmond NE, Raleigh J, van Soest RWM, Kelly M, Travers SAA, Bradshaw B, Vartia S, Stephens KM, McCormack GP. 2011. Phylogenetic relationships of the marine *Haplosclerida* (Phylum Porifera): employing ribosomal (28S rRNA) and mitochondrial (cox1, nad1) gene sequence data. *PLoS ONE* 6(9), e24344.
Rounge TB, Rohrlack T, Nederbragt AJ, Kristensen T, Jakobsen KS. 2009. A genome-wide analysis of nonribosomal peptide synthetase gene clusters and their peptides in a *Planktothrix rubescens* strain. *BMC Genom* 10:396.
Saitou N, Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–25.
Sambrook J, Russell DW. 2001. Molecular Cloning: a Laboratory Manual, 3rd edition. New York: Cold Spring Harbor Laboratory Press, New York.
Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74:5463–7.
Scheuermayer M, Pimentel-Elardo S, Fieseler L, Grozdanov L, Hentschel U. 2006. Microorganisms of sponges: phylogenetic diversity and biotechnological potential. In: Proksch Peter, Müller Werner EG (Eds.). *Frontiers in Marine Biotechnology*.
Schirmer A, Gadkari R, Reeves CD, Ibrahim F, De Long EF, Hutchinson CR, Reid RC. 2005. Metagenomic analysis reveals diverse polyketide synthase gene clusters in microorganisms associated with the marine sponge *Discodermia dissoluta*. *Appl Environ Microbiol* 71:4840–9.
Schmidt EW, Obraztsova AY, Davidson SK, Faulkner DJ, Haygood MG. 2000. Identification of the antifungal peptide-containing symbiont of the marine sponge *Theonella swinhoei* as a novel δ-proteobacterium, "Candidatus *Entotheonella palauensis*". *Mar Biol* 136(6):969–77.
Schweizer E, Hofmann J. 2004. Microbial type I fatty acid synthases (FAS): major players in a network of cellular FAS systems. *Microbiol Mol Biol Rev* 68: 501–17.
Siddall ME, Fontanella FM, Watson SC, Kvist S, Erseus C. 2009. Barcoding bamboozled by bacteria: convergence to metazoan mitochondrial primer targets by marine microbes. *Syst Biol* 58:445–51.
Siegl A, Kamke J, Hochmuth T, Piel J, Richter M, Liang C, Dandekar T, Hentschel U. 2011. Single-cell genomics reveals the lifestyle of *Poribacteria*, a candidate phylum symbiotically associated with marine sponges. *The ISME J* 5:61–70.
Smith S, Tsai SC. 2007. The type I fatty acid and polyketide synthases: a tale of two megasynthases. *Nat Prod Rep* 24(5):1041–72.
Staunton J, Weissman KJ. 2001. Polyketide biosynthesis: a millennium review. *Nat Prod Rep* 18:380–416.
Tasdemir D, Mangalindan GC, Concepción GP, Verbitski SM, Rabindran S, Miranda M, Greenstein M, Hooper JNA, Harper MK, Ireland CM. 2002. Bioactive isomalabaricane triterpenes from the marine sponge *Rhabdastrella globostellata*. *J Nat Prod* 65:210–4.
Taylor MW, Radax R, Steger D, Wagner M. 2007. Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. *Microbiol Mol Biol Rev* 71:295–347.
Thiel V, Jenisch A, Wörheide G, Löwenberg A, Reitner J, Michaelis W. 1999. Mid-chain branched alkanoic acids from 'living fossil' demosponges: a link to ancient sedimentary lipids? *Org Geochem* 30:1–14.
Thompson JD, Higgins DG, Gibson TJ. 1994. Clustal W: improving the sensitivity of progressive multiple alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–80.
Ueoka R, Uria AR, Reiter S, Mori T, Karbaum P, Peters EE, Helfrich EJN, Morinaka BI, Gugger M, Takeyama H, Matsunaga S, Piel J. 2015. Metabolic and evolutionary origin of actin-biding polyketides from diverse organisms. *Nature Chem Biol* 11:705–12.
Uria AR, Piel J. 2009. Cultivation-independent approaches to investigate the chemistry of marine symbiotic bacteria. *Phytochem Rev* 8:401–14.

Vargas S, Schuster A, Sacher K, Büttner G, Schätzle S, Läuchli B, Hall K, Hooper JNA, Erpenbeck D, Wörheide G. 2012. Barcoding sponges: an overview based on comprehensive sampling. *PloS ONE* 7(7):e39345.

Wakimoto T, Egami Y, Nakashima Y, Wakimoto Y, Mori T, Awakawa T, Ito T, Kenmoku H, Asakawa Y, Piel J, Abe I. 2014. Calyculin biogenesis from a pyrophosphate protoxin produced by a sponge symbiont. *Nat Chem Biol* 10:648–55.

Webster NS, Hill RT. 2001. The culturable microbial community of the Great Barrier Reef sponge *Rhopaloeides odorabile* is dominated by an α-Proteobacterium. *Mar Biol (Berl)* 138:843–51.

Wilson MC, Mori T, Ruckert Ch, Uria AR, Helf MJ, Takada K, Gernert Ch, Steffens UAE, Heycke N, Schmitt S, Rinke C, Helfrich EJN, Brachmann AO, Gurgui C, Wakimoto T, Kracht M, Crüsemann M, Hentschel U, Abe I, Matsunaga S, Kalinowski J, Takeyama H, Piel J. 2014. An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature* 506: 58–62.

Wörheide G, Erpenbeck D. 2007. DNA taxonomy of sponges-progress and perspectives. *J Mar Biol Ass* 87:1–6.

Wörheide G, Erpenbeck D, Menke C. 2008. The Sponge Barcoding Project: aiding in the identification and description of poriferan taxa. In: *Porifera Research: Biodiversity, Innovation & Sustainability*, : *Museu Nacional de Rio de Janiero Book Series*. pp. 123–8.

Ziemert N, Podell S, Penn K, Badger JH, Allen E, Jensen PR. 2012. The Natural Product Domain Seeker NaPDoS: a phylogeny based bioinformatic tool to classify secondary metabolite gene diversity. *PLoS One* 7(3), e34064.