Isolation and structure determination of a new antibacterial peptide pentaminomycin C from Streptomyces cacaoi subsp. cacaoi

Title: Isolation and structure determination of a new antibacterial peptide pentaminomycin C from *Streptomyces cacaoi* subsp. *cacaoi*

Running headline: structure determination of pentaminomycin C

Authors: Issara Kaweewan,[1] Hikaru Hemmi,[2] Hisayuki Komaki,[3] Shinya Kodani[1,4,5*]

Affiliations: [1]Graduate School of Science and Technology, Shizuoka University, Shizuoka, Japan, [2]Food Research Institute, National Agriculture and Food Research Organization (NARO), Ibaraki, Japan, [3]Biological Resource Center, National Institute of Technology and Evaluation (NBRC), Chiba, Japan, [4]Graduate School of Integrated Science and Technology, Shizuoka University, Shizuoka, Japan, [5]Academic Institute, Shizuoka University, Shizuoka, Japan

*To whom correspondence should be addressed: Shinya Kodani, College of Agriculture, Academic Institute, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka 422-8529 Japan, Tel/Fax; +81(54)238-5008, E-mail; kodani.shinya@shizuoka.ac.jp

Keywords: antibacterial, peptide, *Streptomyces cacaoi* subsp. *cacaoi*, NMR, NRPS

Abstract

A new antibacterial peptide named pentaminomycin C was isolated from an extract of *Streptomyces cacaoi* subsp. *cacaoi* NBRC 12748^{T}, along with a known peptide BE-18257A. Pentaminomycin C was determined to be a cyclic pentapeptide containing an unusual amino acid, *N*δ-hydroxyarginine (5-OHArg), by a combination of ESI-MS and NMR analyses. The structure of pentaminomycin C was determined to be cyclo(-L-Leu-D-Val-L-Trp-L-5-OHArg-D-Phe-). Pentaminomycin C exhibited antibacterial activities against Gram-positive bacteria including *Micrococcus luteus*, *Bacillus subtilis* and *Staphylococcus aureus*. The biosynthetic gene cluster for pentaminomycin C and BE-18257A was identified from the genome sequence data of *S. cacaoi* subsp. *cacaoi*.

Introduction

Many bioactive peptides have been isolated from microorganisms including bacteria and fungi.[1] Among them, several peptides, for example vancomycin, are used as therapeutic agents to treat bacterial infection. The peptide aglycone of vancomycin is biosynthesized by a nonribosomal peptide synthetase (NRPS) system which consists of seven modules on three NRPSs (VcmA, VcmB, and VcmC).[2, 3] Moreover, many commercial peptide antibiotics (chloramphenicol,[4] daptomycin,[5] teicoplanin,[6] actinomycin D,[7] bleomycins A2 and B2,[8] cyclosporine A[9]) have been reported to be biosynthesized by NRPS system.[10]

Cyclic peptides derived from microorganisms have been reported to exhibit a wide variety of biological activities. Malformins are a group of cyclic pentapeptides derived from fungus *Aspergillus niger* as inducers for malformations of bean plants and curvatures of corn roots.[11, 12] Among malformins, malformins A1 and C were reported to have potent cytotoxicity.[13, 14] Kakeromamide A is a cyclic pentapeptide isolated from the marine cyanobacterium, which possesses inducing activity for differentiation of neural stem cells into astrocytes in vitro.[15-17] The cyclic pentapeptides BE-18257A and B (**1** and **2** in Fig. 1) were isolated from *Streptomyces misakiensis* BA18257 as endothelium-derived vasoconstrictor factor binding inhibitors.[18, 19] Recently, cyclic

pentapeptides, pentaminomycins A and B (**3** and **4** in Fig. 1), which contain *N*δ-hydroxyarginine (5-OHArg), were isolated from *Streptomyces* sp. RK88-1441.[20] Pentaminomycin A showed an antimelanogenic activity against alpha-melanocyte stimulating hormone (α-MSH)-stimulated B16F10 melanoma cells.[20]

Based on this background, we performed chemical screening for new peptides in extracts of actinobacteria using ESI-MS and HPLC (Data not shown). As a result, we isolated a new antibacterial cyclic pentapeptide named pentaminomycin C (**5** in Fig. 1) from the MeOH extract of *Streptomyces cacaoi* subsp. *cacaoi* NBRC 12748^{T}, along with a known peptide BE-18257A (**1** in Fig. 1). Here we describe the isolation and structure determination of **5** from *S. cacaoi* subsp. *cacaoi*. In addition, we found the putative biosynthetic gene cluster containing two NRPSs for **1** and **5** in the genome data of *S. cacaoi* subsp. *cacaoi*.[21]

Results and Discussion

A chemical investigation of *Streptomyces cacaoi* subsp. *cacaoi* using HPLC and ESI-MS analyses was conducted. As a result, a new peptide named pentaminomycin C (**5** in Fig. 1) was detected by subjecting MeOH extract of *S. cacaoi* subsp. *cacaoi* to HPLC analysis (Figure S1). *S. cacaoi* subsp. *cacaoi* was cultured using 5L of modified ISP2 agar medium[22] and the extract was subjected to repeated HPLC separation to obtain **5**, along with a known peptide BE-18257A (**1** in Fig. 1).

The molecular formula of **5** was established as $C_{37}H_{51}N_9O_6$ by HR-ESI-MS, since an ion peak $[M+H]^+$ was observed at *m/z* 718.4010 (calculated *m/z* value: 718.4040). To determine the chemical structure of **5**, NMR analyses including ^1H, ^{13}C, DEPT-135, DQF-COSY, TOCSY, NOESY, HMBC, HSQC and N-H HSQC were performed. The assignment of the constituent five amino acids including leucine (Leu), valine (Val), tryptophan (Trp), *N*δ-hydroxyarginine (5-OHArg) and phenylalanine (Phe) was achieved using spin-system identification (Table 1). Regarding assignment of an unusual amino acid 5-OHArg, HMBC correlation was detected from proton (δH 10.52) to carbon of C=N bond (157.4). In addition, the chemical shifts of protons (δH 3.40) and carbon at δ position (δC 50.6) in 5-OHArg were shifted to downfield due to presence of hydroxy residue of δ-amine. The chemical shifts were similar to literature

values of 5-OHArg previous reported for the peptides pentaminomycins A and B.[20] The sequence of amino acids was determined by interpretation of HMBC data. The HMBC correlations from the α-proton and the amide proton to the same carbonyl carbon were observed (half ended arrows in Fig.2), indicating the sequence of 5-OHArg-Phe-Leu-Val-Trp. HMBC correlation from an α-proton (δH 4.27) to the carbonyl carbon (δC 171.9) in Trp was not observed. However a NOESY correlation was observed between α-proton of Trp (δH 4.27) and amide proton of 5-OHArg (δH 7.31) as shown by double ended arrow in Fig. 2, which indicated the connection between Trp and 5-OHArg. Pentaminomycin C was shown to have structure of cyclo(-Leu-Val-Trp-5-OHArg-Phe-). Similar ESI-MS and NMR analyses were performed on **1**. The chemical shifts were similar to reported literature values of BE-18257A[19] (Data not shown). In addition, we confirmed the planer structure of **1** by 2D NMR analyses in the same manner with **5**.

The absolute configurations of the amino acids (Leu, Val, Trp, 5-OHArg and Phe) in **5** were determined by modified Marfey's analysis.[23] To convert 5-OHArg to Arg, compound **5** (1.0 mg) was hydrolyzed in hydriodic acid (HI) according to a previous report.[24] To analyze Trp, compound **5** (1.0 mg) was hydrolyzed in 6N HCl containing 3% phenol to recover Trp.[25] Each hydrolysate was derivatized using L-FDLA, followed

98 by comparative analysis by HPLC using standard amino acids derivatized with L- or D-
99 FDLA. The results indicated that the absolute configurations of amino acids were L-Leu,
100 D-Val, L-Trp, L-5-OHArg and D-Phe. The structure of pentaminomycin C was therefore
101 determined to be cyclo(-L-Leu-D-Val-L-Trp-L-5-OHArg-D-Phe-). For the
102 stereochemistries of the constituent amino acids in **1**, the same method was applied
103 (Data not shown). We confirmed the stereochemistries of constituent amino acids
104 (L-Leu, D-Trp, D-Glu, L-Ala, D-Val) in **1**, which were identical with the previous
105 report.[19]

106 Compounds **1** and **5** were tested against *Bacillus subtilis*, *Micrococcus luteus*,
107 *Staphylococcus aureus*, *Escherichia coli*, and *Pseudomonas aeruginosa* to determine the
108 minimum inhibitory concentration (MIC), along with positive control compound
109 tetracycline (Table S1).[26] As a result, compound **5** showed antibacterial activities against
110 Gram-positive bacteria including *B. subtilis*, *M. luteus* and *S. aureus* with MIC of 16
111 μg/mL, but no activity for Gram-negative bacteria including *E. coli* and *P. aeruginosa* at
112 the concentration of 64 μg/mL (Table S1). On the other hand, compound **1** did not show
113 any antibacterial activity against any of the test bacteria at the concentration of 64
114 μg/mL.

115 To clarify the biosynthetic pathway, we searched the biosynthetic gene cluster for

116 pentaminomycin C and BE-18257A in draft genome sequence of *Streptomyces cacaoi*
117 subsp. *cacaoi* NRRL B-1220^{T} (MUBL01000001－MUBL0100981) isogenic to NBRC
118 12748^{T}. As pentaminomycin C and BE-18257A are composed of five amino acid
119 residues comprising two to three D-amino acids, we looked for NRPS genes harboring
120 five modules and comprising two to three epimerase (E) domains. Consequently, two
121 NRPS genes were found as shown in Fig. 3. The gene cluster includes regulatory genes,
122 transport-related genes and biosynthetic genes including *mtb*, putative hydrolase and
123 P450 genes in addition to two NRPS genes. The gene BZY53_RS10935 was considered
124 as the NRPS gene responsible for the synthesis of BE-18257A, because the second,
125 third and fifth modules include E domain which isomerizes L-amino acid to
126 D-configuration (Fig. 3). Our bioinformatic analysis using online analysis tools[27, 28]
127 suggested the substrates of the first, fourth and fifth adenylation (A) domains were Leu,
128 Ala and Val, respectively, which are in accordance with constituent amino acid residues
129 (L-Leu-D-Trp-D-Glu-L-Ala-D-Val) in BE-18257A as underlined (Fig. 3). Another NRPS
130 gene (BZY53_RS10880 to BZY53_RS01305) was considered responsible for
131 pentaminomycin C synthesis, because the second and fifth modules contain E domain,
132 and second, fourth and fifth A domains were predicted to incorporate Val, Arg and Phe
133 as the substrate, respectively, which account for the amino acid sequence of

pentaminomycin C (L-Leu-D-Val-L-Trp-L-Arg-D-Phe). P450 enzyme encoding gene(s) of BZY53_RS10990 and/or BZY53_RS10985 will hydroxylate the Arg residue to form 5-OHArg in pentaminomycin C. In NRPS pathways, elongated peptide chains are released from the peptidyl carrier protein (PCP) in the last module of NRPSs by the thioesterase (TE) domain next to the PCP, but no TE domains are present in both the NRPSs for pentaminomycin C and BE-18257A. Recently, putative hydrases, stand-alone enzymes belonging to the penicillin-binding family, such as SurE, MppK and DsaJ were shown to be involved in the release of peptide chains from the PCP and its macrocyclization to afford cyclic peptides.[29, 30] A putative hydrolase adjacent BZY53_RS10935 showed homology to these enzymes (about 40% similarities and 50% identities in amino acid sequence). As SurE has been reported as a trans-acting thioesterase cyclizing two distinct nonribosomal peptides encoded in a single gene cluster,[29, 30] the putative hydrolase (BZY53_RS10940) will also act to form two cyclic peptides pentaminomycin C and BE-18257A. Taken together, we herein propose the biosynthetic pathways of pentaminomycin C and BE-18257A in Fig. 4a and Fig. 4b, respectively. Each A domain converts each amino acid building block into aminoacyl adenylate and transfers them to the PCP within each module to form the corresponding aminoacyl thioester. E domain within each module epimerize α-carbon of Trp, Glu and

Val loaded on modules 2, 3 and 5 (BE-18257A) and Val and Phe on modules 2 and 5 (pentaminomycin C) to afford D-configuration. C domains catalyze for successive *N*-acylation to yield L, D, D, L, D- and L, D, L, L, D-pentapeptidyl thioesters attached to the PCPs of module 5, respectively. The Arg residue in pentaminomycin C is hydroxylated to be 5-OHArg by P450 enzymes. The putative hydrolase releases the two linear pentapeptide chains from the PCPs and then cyclizes them to yield pentaminomycin C and BE-18257A, respectively.

Genome-mining is often employed to search for novel secondary metabolites, for which accurate bioinformatic analysis of secondary metabolite-biosynthetic gene clusters is essential. Usually, a single gene cluster is believed to produce compounds containing the same backbone. However, the gene cluster that we reported here produces two types of peptides whose backbones are different. The methodology of bioinformatic to predict final product of NRPS system is based on relationship between amino acid sequence of substrate binding pocket in A domain and actual amino acid substrate. So far data accumulation is not enough to perfectly predict final product. Studies on structure determination of actual products coupled with identification of the corresponding biosynthetic gene clusters provide useful information to the research fields because accumulation of such knowledges is essential to bioinformatic analysis.

Experimental Section

Microbial strains

Bacterial strains including *Streptomyces cacaoi* subsp. *cacaoi* NBRC 12748^T, *Micrococcus luteus* NBRC 3333^T, *Pseudomonas aeruginosa* NBRC 12689^T, *Bacillus subtilis* subsp. *subtilis* NBRC 13719^T, *Staphylococcus aureus* subsp. *aureus* NBRC 100910^T and *Escherichia coli* NBRC 102203^T were obtained from the NBRC culture collection (NITE Biological Resource Center, Japan)

Isolation of peptides

Streptomyces cacaoi subsp. *cacaoi* NBRC 12748^T was cultured in 5L of modified ISP2 agar medium[22] (2 g leucine, 4 g yeast extract, 10 g malt extract, 4 g glucose, and 15 g agar in 1 L distilled water, adjusted pH 7.3) for 9 days at 30 °C. Bacterial cells were harvested and extracted with double volume of MeOH. The mixture was filtered through filter paper (Whatman No.1, GE Healthcare Life Science, UK), followed by centrifugation at 4000 rpm for 10 min to remove insoluble compounds. The extract was concentrated by rotary evaporation and compound **5** was isolated using HPLC (Navi C30-5 column; 4.6×250 mm; Wakopak, Wako Pure Chemical Industries, Tokyo, Japan). The HPLC condition was set as isocratic elution mode; 37% MeCN/water

188 containing 0.05% trifluoroacetic acid (TFA) at flow rate 1 mL/min and the UV detector
189 was set at the wavelength of 220 nm to obtain **5** (Retention time; 26.70 min). The
190 extract was repeatedly subjected to HPLC purification and freeze-drying to afford 3.0
191 mg of **5**, along with **1** (5.0 mg).

192 **ESI-MS analysis**

193 ESI-MS analyses were performed using a JEOL JMS-T100LP mass spectrometer. For
194 accurate MS analysis, reserpine was used as an internal standard.

195 **NMR analysis**

196 NMR sample was prepared by dissolving **5** in 500 μl of DMSO-d_6, 4:1. 1D ^{1}H, ^{13}C,
197 DEPT-135, and all 2D NMR spectra were obtained on Bruker Avance800 spectrometer
198 with quadrature detection following the previous report.[31]

199 **Modified Marfey's analysis**

200 The amino acid analyses were carried out in sealed vacuum hydrolysis tube. Compound
201 **5** (1.0 mg) was hydrolyzed in 500 μL of hydriodic acid (HI) at 155 °C for 24 h to
202 remove the hydroxy group of OHArg according to a previous report[24]. To recover Trp,
203 compound **5** (1.0 mg) was hydrolyzed in 500 μL of 6N HCl containing 3% phenol at
204 166 °C for 25 min[25]. After cooling to room temperature, the hydrolysate was evaporated
205 using rotary evaporator and the remaining solvent was completely dried under vacuum.

The hydrolysate was resuspended in 200 µL of water, followed by adding 10 µL of the solution of *N*α-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide (L-FDLA; Tokyo Chemical Industry Co., LTD, Tokyo, Japan) in acetone (10 µg/µL). The 1M $NaHCO_3$ solution (100 µL) was added to the hydrolysate and the mixture was incubated at 80°c for 3 min. The reaction mixture was cooled to room temperature and neutralized with 50 µL of 2N HCl. The 50% MeCN/water (1 mL) was added to the mixture before subjected to HPLC. For standard amino acids, each amino acid (1 mg) was derivatized with L-FDLA and D-FDLA in the same manner. Approximately 20 µL of each FDLA derivative was analyzed by HPLC (C18 column, 4.6×250 mm, Wakopak Handy ODS, Wako Pure Chemical Industries, Tokyo, Japan). DAD detector (MD-2018, JASCO, Tokyo, Japan) was used for detection of the amino acid derivatives accumulating the data of absorbance from 220 nm to 420 nm. HPLC analysis for all standard amino acids (Leu, Val, Trp, Arg and Phe) was performed using solvent A (distilled water containing 0.05% TFA) and solvent B (MeCN containing 0.05% TFA) at a flow rate of 1 mL/min in linear gradient mode from 0 min to 70 min, increasing percentage of solvent B from 25% to 60% (HPLC condition 1). The retention times (min) for L- and D-FDLA derivatized amino acids (HPLC condition 1) were as follows; L-Arg-D-FDLA (18.56 min), L-Arg-L-FDLA (20.25 min), L-Val-L-FDLA (40.68 min), L-Leu-L-FDLA (46.51 min),

L-Trp-L-FDLA (47.39 min), L-Phe-L-FDLA (48.66 min), L-Val-D-FDLA (53.76 min), L-Trp-D-FDLA (54.37 min), L-Phe-D-FDLA (59.30 min) and L-Leu-D-FDLA (62.57 min). HPLC analysis for Val was performed using solvent A (distilled water containing 0.05% TFA) and solvent B (MeCN containing 0.05% TFA) at a flow rate of 1 mL/min in isocratic mode 40% of solvent B for 55 min (HPLC condition 2). The retention times (min) for L- and D-FDLA derivatized amino acids (HPLC condition 2) were as follows; L-Val-L-FDLA (21.05 min) and L-Val-D-FDLA (40.23 min). HPLC analysis for Trp was performed using solvent A (distilled water containing 0.05% TFA) and solvent B (MeCN containing 0.05% TFA) at a flow rate of 1 mL/min in isocratic mode 45% of solvent B for 55 min (HPLC condition 3). The retention times (min) for L- and D-FDLA derivatized amino acids (HPLC condition 3) were as follows; L-Trp-L-FDLA (18.57 min) and L-Trp-D-FDLA (28.04 min).

Antibacterial activity test

The antibacterial activities of peptides were assessed using the minimum inhibitory concentrations (MICs) test in 96-well microplates according to the previous report.[26] The peptides were tested against bacterial strains including *E. coli*, *P. aeruginosa*, *S. aureus*, *B. subtilis and M. luteu*s. Tetracycline was used as a positive control. Using the broth dilution technique with Mueller-Hinton Broth (MHB), serial twofold dilution of

the compounds (**1** and **5**) and tetracycline were prepared (50 μL/well), ranging from 64 to 0.0625 μg/mL. Test bacteria were diluted with MHB to give a final inoculum of 10^5 CFU/mL. The bacterial solution (50 μL) was transferred to each well which contain a various concentrations of testing compound. The MICs was determined as the lowest concentration that gave no visible growth after incubate the microplates at 30 °C for 24h.

Acknowledgments

This study was supported by the Japan Society for the Promotion of Science by Grants-in-aids (grant number 16K01913). The NMR spectra were recorded on Bruker Avance 600 and Avance III HD 800 spectrometers at Advanced Analysis Center, NARO.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Hancock, R. E. & Chapple, D. S. Peptide antibiotics. *Antimicrob Agents Chemother.* **43**, 1317-1323 (1999)

2. Levine, D. P. Vancomycin: a history. *Clin Infect Dis.* **42 Suppl 1**, S5-12 (2006)

3. Hubbard, B. K. & Walsh, C. T. Vancomycin assembly: nature's way. *Angew Chem Int Ed.* **42**, 730-765 (2003)

4. Dinos, G. P., et al. chloramphenicol derivatives as antibacterial and anticancer agents: historic problems and current solutions. *Antibiotics (Basel).* **5**, E20 (2016)

5. Steenbergen, J. N., Alder, J., Thorne, G. M. & Tally, F. P. Daptomycin: a lipopeptide antibiotic for the treatment of serious gram-positive infections. *J Antimicrob Chemother.* **55**, 283-288 (2005)

6. Henson, K. E., Levine, M. T., Wong, E. A. & Levine, D. P. Glycopeptide antibiotics: evolving resistance, pharmacology and adverse event profile. *Expert Rev Anti Infect Ther.* **13**, 1265-1278 (2015)

7. Marks, T. A. & Venditti, J. M. Potentiation of actinomycin D or adriamycin antitumor activity with DNA. *Cancer Res.* **36**, 496-504 (1976)

8. Dorr, R. T., Meyers, R., Snead, K. & Liddil, J. D. Analytical and biological inequivalence of two commercial formulations of the antitumor agent bleomycin. *Cancer Chemother Pharmacol.* **42**, 149-154 (1998)

9. Tedesco, D. & Haragsim, L. Cyclosporine: a review. *J Transplant.* **2012**, 230386

(2012)

10. Süssmuth, R. D. & Mainz, A. Nonribosomal peptide synthesis - principles and prospects. *Angew Chem Int Ed.* **56**, 3770-3821 (2017)

11. Curtis, R. W. Curvatures and malformations in bean plants caused by culture filtrate of *Aspergillus niger*. *Plant Physiol.* **33**, 17-22 (1958)

12. Curtis, R. W. Root curvatures induced by culture filtrates of *Aspergillus niger*. *Science.* **128**, 661-662 (1958)

13. Wang, J., et al. Study of malformin C, a fungal source cyclic pentapeptide, as an anti-cancer drug. *PLoS One.* **10**, e0140069 (2015)

14. Liu, Y., et al. Malformin A1 promotes cell death through induction of apoptosis, necrosis and autophagy in prostate cancer cells. *Cancer Chemother Pharmacol.* **77**, 63-75 (2016)

15. Koizumi, Y., Nagai, K., Hasumi, K., Kuba, K. & Sugiyama, T. Structure–activity relationship of cyclic pentapeptide malformins as fibrinolysis enhancers. *Bioorg Med Chem Lett.* **26**, 5267-5271 (2016)

16. Ma, Y.-M., Liang, X.-A., Zhang, H.-C. & Liu, R. Cytotoxic and antibiotic cyclic pentapeptide from an endophytic *Aspergillus tamarii* of *Ficus carica*. *J Agric Food Chem.* **64**, 3789-3793 (2016)

17. Nakamura, F., et al. Kakeromamide A, a new cyclic pentapeptide inducing astrocyte differentiation isolated from the marine cyanobacterium *Moorea bouillonii*. *Bioorg Med Chem Lett.* **28**, 2206-2209 (2018)

18. Kojiri, K., et al. Endothelin-binding inhibitors, BE-18257A and BE-18257B. I. Taxonomy, fermentation, isolation and characterization. *J Antibiot (Tokyo).* **44**, 1342-1347 (1991)

19. Nakajima, S., Niiyama, K., Ihara, M., Kojiri, K. & Suda, H. Endothelin-binding inhibitors, BE-18257A and BE-18257B II. Structure determination. *J Antibiot (Tokyo).* **44**, 1348-1356 (1991)

20. Jang, J.-P., et al. Pentaminomycins A and B, hydroxyarginine-containing cyclic pentapeptides from *Streptomyces* sp. RK88-1441. *J Nat Prod.* **81**, 806-810 (2018)

21. Labeda, D. P. Taxonomic evaluation of putative *Streptomyces scabiei* strains held in the ARS Culture Collection (NRRL) using multi-locus sequence analysis. *Antonie Van Leeuwenhoek.* **109**, 349-356 (2016)

22. Shirling, E. B. & Gottlieb, D. Methods for characterizing *Streptomyces* species. *Int J Syst Bacteriol.* **16**, 313-340 (1966)

23. Harada, K. I., Fujii, K., Hayashi, K., Suzuki, M., Ikai, Y. & Oka, H. Application

of D,L-FDLA derivatization to determination of absolute configuration of constituent amino acids in peptide by advanced Marfey's method. *Tetrahedron Lett.* **37**, 3001-3004 (1996)

24. Inglis, A. S., Nicholls, P. W. & Roxburgh, C. M. Hydrolysis of the peptide bond and amino acid modification with hydriodic acid. *Aust J Biol Sci.* **24**, 1235-1240 (1971)

25. Muramoto, K. & Kamiya, H. Recovery of tryptophan in peptides and proteins by high-temperature and short-term acid hydrolysis in the presence of phenol. *Anal Biochem.* **189**, 223-230 (1990)

26. Di Modugno, E., Erbetti, I., Ferrari, L., Galassi, G., Hammond, S. M. & Xerri, L. In vitro activity of the tribactam GV104326 against gram-positive, gram-negative, and anaerobic bacteria. *Antimicrob Agents Chemother.* **38**, 2362-2368 (1994)

27. Prieto, C., Garcia-Estrada, C., Lorenzana, D. & Martin, J. F. NRPSsp: non-ribosomal peptide synthase substrate predictor. *Bioinformatics.* **28**, 426-427 (2012)

28. Blin, K., et al. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **47**, W81-W87 (2019)

29. Kuranaga, T., et al. Total synthesis of the nonribosomal peptide surugamide B and Identification of a new offloading cyclase family. *Angew Chem Int Ed Engl.* **57**, 9447-9451 (2018)

30. Matsuda, K., et al. SurE is a trans-acting thioesterase cyclizing two distinct non-ribosomal peptides. *Org Biomol Chem.* **17**, 1058-1061 (2019)

31. Kodani, S., et al. Sphaericin, a lasso peptide from the rare actinomycete *Planomonospora sphaerica*. *Eur J Org Chem.* 1177-1183 (2017)

340 Table1. NMR chemical shift values of **5** in DMSO-d_6

AA	Position	δH (*J* = Hz)	δC	δN
Leu	CO		172.4	
	NH	7.52 (d, 9.0)		122.0
	α	4.37 (m)	50.6	
	β	1.32 (m)	41.1	
	γ	1.40 (m)	24.4	
	δ1	0.79 (d, 6.5)	22.1	
	δ2	0.82 (d, 6.6)	22.9	
Phe	CO		170.8	
	NH	8.80 (d, 8.2)		127.4
	α	4.45 (m)	53.9	
	β	2.78 (d, 14.0, 9.4)	34.5	
		2.93 (d, 14.0, 5.9)		
	γ		138.0	
	δ	7.21 (m)	129.2	
	ε	7.21 (m)	128.2	
	ζ	7.16 (m)	126.4	
OHArg	CO		170.6	
	NH	7.31 (d, 7.2)		115.3
	α	4.14 (m)	52.9	
	β	1.51 (m)	28.2	
	γ	1.16 (m), 1.33 (m)	22.2	
	δ	3.40 (m)	50.6	
	N-OH	10.52 (s)		
	C=NH	ND	157.4	
	NH_2	ND		
Trp	CO		171.9	
	NH	8.55 (d, 7.8)		125.8
	α	4.27 (m)	55.5	
	β	2.89 (dd, 14.6, 11.3)	27.1	
		3.17 (dd, 14.6, 3.6)		
	ε1	10.75 (s)		130.8
	δ1	7.15 (m)	124.0	
	γ		110.2	
	ε3	7.50 (d, 7.9)	118.1	
	ζ3	6.96 (m)	118.5	
	η2	7.03 (m)	121.1	
	ζ2	7.29 (d, 8.1)	111.5	
	ε2		136.3	
	δ2		127.0	
Val	CO		171.6	
	NH	8.37 (d, 7.5)		123.3
	α	3.67 (dd, 9.9, 7.5)	60.1	
	β	1.62 (m)	28.7	
	γ1	0.31 (d, 6.7)	18.6	
	γ2	0.73 (d, 6.6)	19.2	

ND:not detected

341

342

343 Figure legends

344 Fig.1. Chemical structures of BE-18257A (**1**) and B (**2**), and pentaminomycins A (**3**), B

345 (**4**), and C (**5)**

346 Fig. 2. Key 2D NMR correlations of pentaminomycin C (**5**)

347 Fig. 3. Biosynthetic gene cluster for pentaminomycin C (**5)** and BE-18257A (**1**). ORFs

348 of BZY53_RS11015 to BZY53_RS10895, BZY53_R10880 plus BZY53_RS01305, and

349 BZY53_RS35155 are indicated. BZY53_RS11015 to BZY53_R10880 are encoded in

350 MUBL01000147, while BZY53_RS01305 and BZY53_RS35155 are in

351 MUBL01000010. The two contig sequences were assembled via 78 bp overlapping

352 between 3'-terminal of MUBL01000010 and 5'-terminal of MUBL01000147. As the

353 sequence of the right NRPS is divided into BZY53_R10880 and BZY53_RS01305, the

354 ORF was indicated as BZY53_R10880 plus BZY53_RS01305. Domain organizations

355 are shown under NRPS genes and each module is underlined. A, adenylation domain; C,

356 condensation domain, E epimerase domain, T, peptidyl carrier protein, which is also

357 called thiolation domain; m, module. TE, thioesterase. Regulatory and transport-related

358 genes are dotted and gray-colored, respectively. The same gene cluster is also present in

359 the genome of strain NBRC 12748^{T} (BJMM010000001－BJMM010000197) but its

360 sequence is fragmentated in to so many contigs (BJMM010000017, BJMM010000182,

361 BJMM010000163, BJMM010000158, BJMM01000085 and BJMM01000063 *etc*) that
362 whole the sequence could not be determined.
363 Fig. 4. Proposed biosynthetic pathways for a) pentaminomycin C (**5**) and b) BE-18257A
364 (**1**). Letters in the figure represent as following: A, adenylation domain; C, condensation
365 domain; E, epimerase domain; m, module; PCP, peptidyl carrier protein, which is also
366 called thiolation domain.
367

Fig. 1

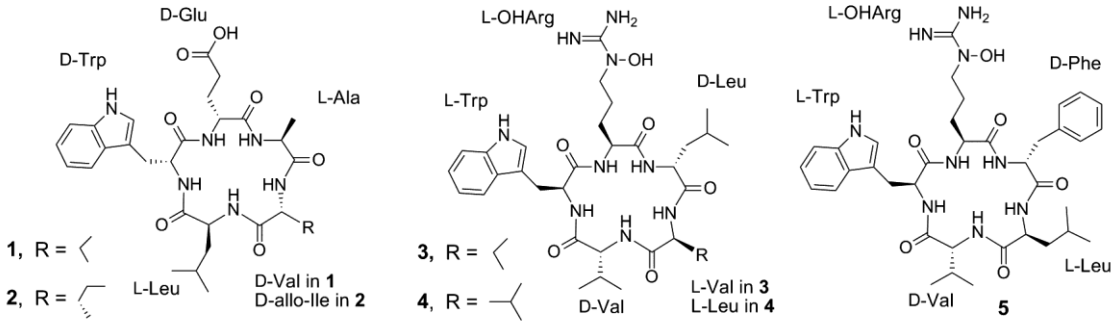

Fig. 1 Kaweewan et al.

371 **Fig. 2**

372

373

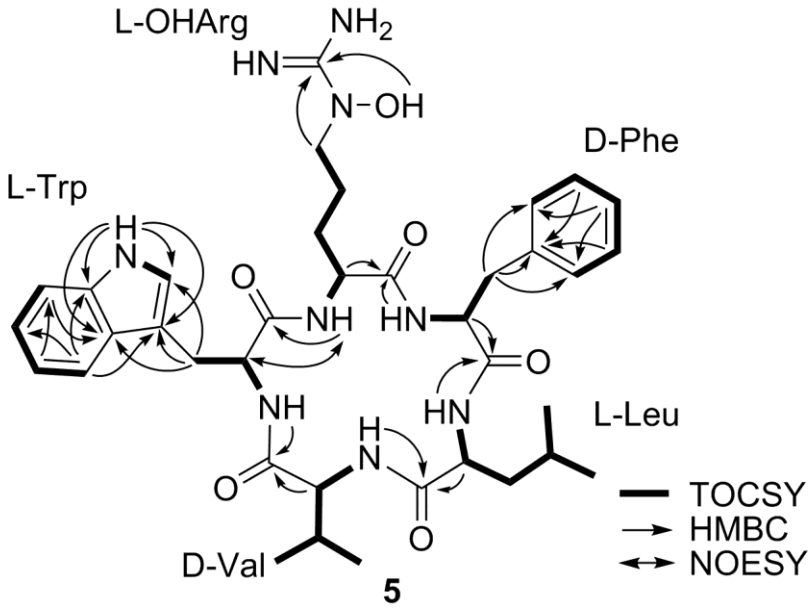

Fig. 2. Kaweewan et al.

374 **Fig. 3**

375

376

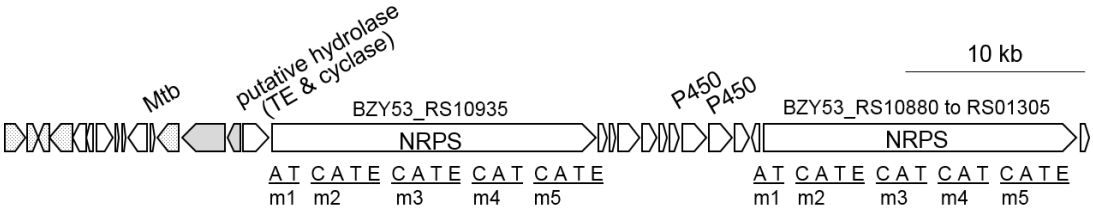

Fig. 3. Kaweewan et al.

377 **Fig. 4**

378

379

380

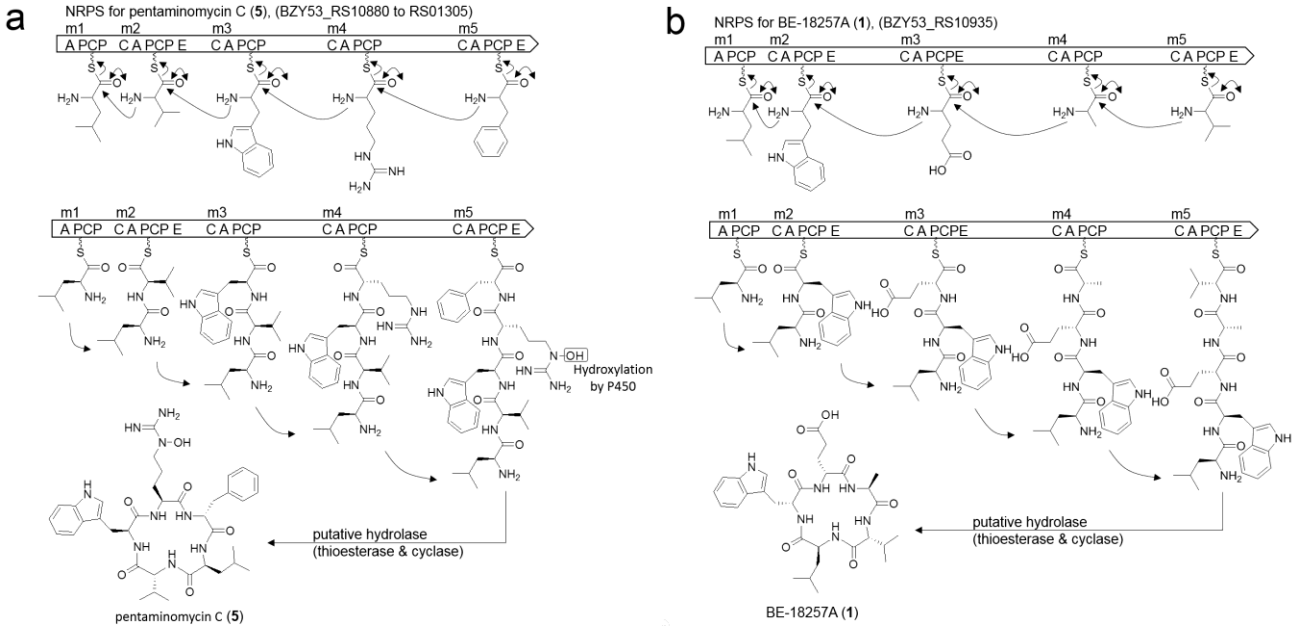

Fig.4. Kaweewan et al.