Isolation and Structure Determination of New Antibacterial Peptide Curacomycin Based on Genome Mining

Isolation and structure determination of new antibacterial peptide curacomycin based on genome mining

Issara Kaweewan, [a] Hisayuki Komaki, [b] Hikaru Hemmi, [c] Shinya Kodani,*[a,d]

Abstract: Based on genome mining, a new antibacterial cyclic peptide named curacomycin (**1**) and its analogue dechlorocuracomycin (**2**) were isolated from *Streptomyces curacoi* NBRC 12761^T and *Streptomyces noursei* NBRC 15452^T, respectively. The chemical structures of the two peptides were determined by the combination of ESI-MS and NMR analyses. The structure of **1** was determined to be a cyclic peptide which consisted of 6 amino acids including Val, Leu, Ile, Orn, β-hydroxyasparagine, and 5-chlorotryptophan. The NMR spectral data of **2** was very similar to that of **1**, which indicated the structure of **2** to be a dechlorinated analogue of **1**. The comparison of antimicrobial activities of the two peptides indicated that the presence of chlorine in the molecule of **1** was critical for antimicrobial activity. The proposed biosynthetic gene clusters for **1** and **2** were found in the genome data of *S. curacoi* and *S. noursei*, respectively. The functions of biosynthetic genes were discussed by comparison between the two gene clusters.

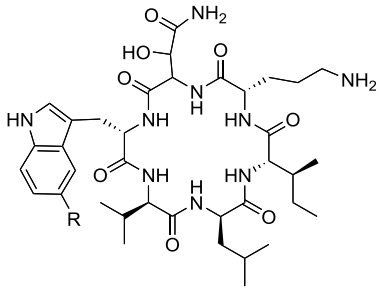
Introduction

Nonribosomal peptides (NRP) are a class of peptide secondary metabolites, usually produced by microorganisms like bacteria and fungi.[1] Nonribosomal peptides are synthesized by nonribosomal peptide synthetases (NRPS), which build a large protein complex. The NRPS proteins are mostly large molecular multi-enzymes which function as the assembly-line to produce the resulting peptide through a series of amino acid building modules lined up in parallel.[2-6] Each module in NRPS uptake one specific amino acid and causes successive elongation of the peptide with enzymatic function of condensation (C) domain. Each module normally contains a basic set of three domains following; 1) an adenylation (A) domain for the recognition of specific amino acid substrate; 2) a peptidyl carrier domain (PCP) that binds the growing peptide chain and the incoming aminoacyl unit; 3) a condensation (C) domain that catalyses peptide bond formation between amino acids on neighboring PCPs. In addition, a thioesterase (TE) domain normally exists at the end of assembly-line of NRPS to release the full-length of the peptide. Genome mining has become a powerful method to search for new bioactive secondary metabolites according to accumulation of genome data of bacteria.[7-17] High percentages of halogenase gene-possessing actinobacteria were reported to contain polyketide synthase genes and/or NRPS genes, indicating the genetic and physiological potentials of halogenases for producing secondary metabolites, possibly halometabolites.[18] So far, more than 4,000 halometabolites have been discovered,[19] including commercially important antibiotics such as chloramphenicol,[20] vancomycin,[21] and teicoplanin.[22]

The presence of tryptophan halogenase encoding gene (ADW94630.1) within NRPS gene cluster was reported in the genome sequence of *Streptomyces toxytricini* NRRL 15443.[23] The heterologous expression experiment indicated that the function of the gene was chlorination of 6th position in the indole of Trp.[23] However, it was not clear whether the tryptophan halogenase (ADW94630.1) could function to provide chlorinated Trp to the neighboring NRPS genes (stnrps-1 and stnrps-2) in vivo. We performed blastp similarity search using the tryptophan halogenase encoding gene (ADW94630.1), and the result indicated the wide distribution of the similar genes in actinomycetes. Among those genes, two genes (AQI70_05465 of *Streptomyces curacoi* and SNOUR_12610 of *Streptomyces noursei*) were found to be located close to NRPS as same as the query gene (ADW94630.1). The strains were already known as producers of antibiotics (*Streptomyces curacoi*: curamycin [24], *Streptomyces noursei*: nystatin[25]). Although the strains are genetically different considering that identity of 16S DNA genes (AB184841.1: *S. curacoi* and CP011533.1: *S. noursei*) was 97%, they have a similar set of NRPS genes. We proposed that cooperative function of the NRPS and the tryptophan halogenase in those strains may give new halogenated peptides. This hypothesis prompted us to perform chemical investigation on two strains including *S. curacoi* and *S. noursei* using HPLC and ESI-MS. As a result, new peptides including curacomycin (**1** in Fig. 1) and dechlorocuracomycin (**2**) were isolated from *S. curacoi* and *S. noursei*, respectively (Fig. 1). Here we describe isolation and structure determination of curacomycin (**1**) and dechlorocuracomycin (**2**). In addition, the proposed biosynthetic pathway of both peptides via NRPS was discussed based on the genome data of the producing strains.

[a] Graduate School of Integrated Science and Technology, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka 422-8529 Japan
[b] Biological Resource Center, National Institute of Technology and Evaluation (NBRC), 2-5-8 Kazusakamatari, Kisarazu, Chiba 292–0818, Japan
[c] Food Research Institute, National Agriculture and Food Research Organization (NARO), 2–1-12 Kannondai, Tsukuba, Ibaraki 305-8642 Japan
[d] Academic Institute, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka 422-8529 Japan
E-mail: kodani.shinya@shizuoka.ac.jp

Supporting information for this article is given via a link at the end of the document.



curacomycin (**1**) R=Cl
dechlorocuracomycin (**2**) R=H

Figure 1. Chemical structures of curacomycin (**1**) and dechlorocuracomycin (**2**)

Results

The cultivation of *S. curacoi* was performed with 2 L of ISP2 agar media. After 10 days of cultivation, cells of spore and aerial hyphae were harvested by scraping off from the surface of agar medium with a steel spatula. The cells of *S. curacoi* were extracted with double volume of MeOH. The MeOH extract was subjected to open column chromatography using hydrophobic polymer resin and water/MeOH system for elution. The MeOH fraction was repeatedly subjected to preparative HPLC to yield the new peptide curacomycin (**1** in Fig. 1, 4.0 mg). The cultivation of *S. noursei* was performed with 2 L of nutrient agar medium complemented with leucine. In the same manner as curacomycin, dechlorocuracomycin (**2**, 3.0 mg) was purified from the extract of MeOH. Interestingly, *S. noursei* did not produce **2** when it was cultured using ISP2 medium.

The molecular formula of **1** was established to be $C_{37}H_{56}N_9O_8Cl$ by HR ESI-MS analysis, that is, the ion corresponding to $[M+H]^+$ (the calculated *m/z* value, 790.4018) was observed at m/z 790.4013. To determine the structure, the analyses of NMR spectra including ^1H, ^{13}C, DEPT-135, DQF-COSY, TOCSY, NOESY, HMBC, HSQC, and ^1H-^{15}N HSQC were performed on **1** dissolved in 0.5 mL of mixed solvent: MeCN-d_3/DMSO-d_6 (4:1). In the ^1H NMR spectrum, α-protons of amino acid residues (3.72, 3.90, 4.18, 4.38, 4.66, and 4.69 ppm) and amide protons (7.24, 7.29, 7.58, 7.69, 7.77, and 8.33 ppm) were observed. The assignments of the constituent 6 amino acids including Val, Leu, Ile, Orn, 5-chlorotryptophan (ClTrp), and β-hydroxyasparagine (OHAsn) were completed using spin system identification (Fig. 2a and Table 1). For the assignment of OHAsn, the existence of hydroxylated methine was indicated by characteristic chemical shifts (δH:4.54 ppm and δC:72.1ppm). The structure of OHAsn was established by the HMBC correlations (α-H and β-H to γ-CO:174.8 ppm, δ-NH:6.71 ppm to β-C:72.1 ppm) and TOCSY correlations indicated in bold line. The assignment of ClTrp was accomplished mainly by HMBC correlations. Four protons (7.14, 7.59, 7.05, 7.33 ppm) in aromatic system and one indole imino proton (10.37 ppm) were observed for assignment of indole moiety, which indicated the substituted indole. The key HMBC correlations from ε3-H to γ-C, η2-C and ε2-C and HMBC correlations from η2-H to ε3-C and ε2-C confirmed the structure of 5-chlorinated tryptophan. As shown Fig. 2a by arrow, the amino acid sequence and its cyclic structure was determined by key HMBC correlations from α-proton and amide proton to the same α-carbonyl carbon (α-H-ClTrp1 and NH-Val to CO:173.3, α-H-Val2 and NH-Leu3 to CO:172.2 ppm, α-H-Leu3 to and NH-Ile4 to CO:174.0, α-H-Ile4 and NH-Orn5 to CO:173.5 ppm, α-H-Orn5 and NH-OHAsn6 to CO:171.9 ppm, α-H-OHAsn6 and NH-Cl-Trp1 to CO:171.1 ppm).

The molecular formula of **2** was established to be $C_{37}H_{57}N_9O_8$ by accurate ESI-MS analysis, that is, the ion corresponding to $[M+H]^+$ (the calculated *m/z* value, 756.4408) was observed at *m/z* 756.4427. The analyses of NMR spectra including ^1H, ^{13}C, DEPT-135, DQF-COSY, TOCSY, NOESY, HMBC, and HSQC were performed on **2** dissolved in 0.5 mL of mixed solvent: MeCN-d_3/DMSO-d_6 (4:1). The chemical shift values of **2** (Table 1) were very similar to those of **1** except that the proton (ζ3-H) was observed in the indole ring of **2**. The assignments of the other constituent 6 amino acids including Val, Leu, Ile, Orn, Trp, β-OHAsn in **2** were completed using spin system identification. The sequence of amino acids was established by HMBC correlations from α-proton and amide proton to the same α-carbonyl carbon in the same manner with **1** (Fig. 2b). Above all, the structure of dechlorocuracomycin was determined to be **2**.

To elucidate the absolute stereochemistries of Orn, Ile, Leu, Val, Trp and ClTrp in **1** and **2**, the modified Marfey's method [26] was applied. The hydrolysate of **1** and **2** were derivatized with N-α-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide (L-FDLA), respectively. The derivatives were subjected to HPLC analysis to compare with standard amino acid derivatives. As a result, the stereochemistries of Orn, Ile, Trp and ClTrp in **1** and **2** were determined to be L, and those of Val and Leu in **1** and **2** were determined to be D.

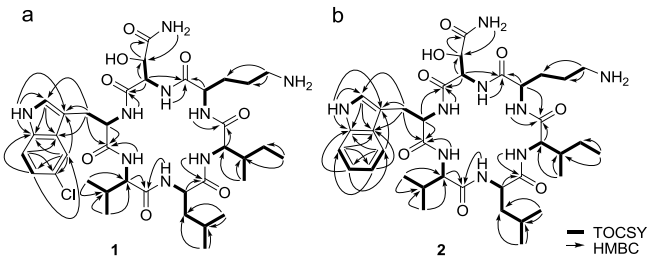


Figure 2. Key 2D NMR correlations of a) curacomycin (**1**) and b) dechlorocuracomycin (**2**)

Table1. NMR chemical shift values of **1** and **2** in MeCN-d_3/DMSO-d_6 (4:1)

curacomycin (**1**)					dechlorocuracomycin (**2**)				
AA	POS	δH (*J* = Hz)	δC	δN	AA	POS	δH (*J* = Hz)	δC	δN
Cl-Trp1	CO		173.3		Trp1	CO		173.0	
	NH	7.77 (d, 7.3)		121.1		NH	7.83 (d, 8.0)		121.2
	α	4.69 (m)	54.8			α	4.68 (m)	54.7	
	β	2.98 (m)	27.8			β	2.93(m)	27.9	
		3.14 (m)					3.18 (m)		
	ε1	10.37 (s)		130.7		ε1	10.43 (s)		130.0
	δ1	7.14 (d, 2.3)	126.8			δ1	7.07 (d, 2.2)	124.7	
	γ		110.9			γ		110.9	
	ε3	7.59 (d, 2.0)	118.8			ε3	7.55 (d, 8.0)	119.3	
	ζ3		124.7			ζ3	6.98 (ddd, 8.0, 6.8, 1.0)	119.4	
	η2	7.05 (dd, 8.6, 2.0)	122.1			η2	7.05 (ddd, 8.0, 6.8, 1.0)	122.0	
	ζ2	7.33 (d, 8.6)	113.8			ζ2	7.31 (dd, 8.0, 1.0)	112.3	
	ε2		135.8			ε2		137.2	
	δ2		129.6			δ2		128.4	
Val2	CO		172.2		Val2	CO		172.0	
	NH	7.69 (d, 6.6)		120.5		NH	7.68 (d, 6.8)		119.6
	α	3.72 (t, 6.9)	62.2			α	3.75 (t, 7.1)	61.9	
	β	1.90 (m)	30.3			β	1.87 (m)	30.3	
	γ1	0.72 (d, 6.9)	19.5			γ1	0.72 (d, 7.0)	19.6	
	γ2	0.75 (d, 6.9)	19.0			γ2	0.73 (d, 7.0)	19.2	
Leu3	CO		174.0		Leu3	CO		173.7	
	NH	7.29 (d, 7.7)		114.1		NH	7.32 (ov*)		114.2
	α	4.38 (m)	52.4			α	4.33 (m)	52.4	
	β	1.52 (m)	40.8			β	1.50 (m)	40.8	
		1.58 (m)					1.56 (m)		
	γ	1.52 (m)	25.5			γ	1.50 (m)	25.4	
	δ1	0.83 (d, 6.5)	21.9			δ1	0.82 (d, 6.5)	22.2	
	δ2	0.87 (ov*)	23.5			δ2	0.86 (d, ov*)	23.5	
Ile4	CO		173.5		Ile4	CO		173.4	
	NH	7.58 (ov*)		115.2		NH	7.71 (d, 7.6)		115.1
	α	4.18 (m)	58.7			α	4.18 (m)	58.4	
	β	1.93 (m)	37.1			β	1.90 (m)	37.2	
	γ1	1.18 (m)	26.8			γ1	1.16 (m)	26.7	
		1.40 (m)					1.38 (m)		
	δ	0.89 (ov*)	11.9			δ	0.86 (ov*)	11.9	
	γ2	0.88 (ov*)	15.3			γ2	0.85 (ov*)	15.3	
Orn5	CO		171.9		Orn5	CO		171.8	
	NH	8.33 (d, 6.2)		119.3		NH	8.50 (d, 6.4)		119.5
	α	3.90 (m)	54.7			α	3.91 (m)	54.6	
	β	1.74 (m)	27.5			β	1.71 (m)	27.7	
		1.99 (m)					1.95 (m)		
	γ	1.60(m)	25.0			γ	1.55(m)	25.3	
	δ	2.83 (m)	40.0			δ	2.77 (m)	40.0	
OHAsn6	CO		171.1		OHAsn6	CO		170.9	
	NH	7.24 (d, 9.0)		105.6		NH	7.33 (ov*)		106.6
	α	4.66 (dd, 8.9, 2.6)	56.6			α	4.63 (m)	56.8	
	β	4.54 (m)	72.1			β	4.50 (m)	72.0	
	OH	5.72 (brs)				OH	5.87 (brs)		
	γ		174.8			γ		174.6	
	δ	6.71 (brs)		102.6		δ	6.98 (ov*)		103.2
		7.11 (brs)					7.17 (brs)		

ov*: overlapped

To clearify antimicrobial spectrum, the antimicrobial assay on **1** and **2** was performed by using a paper disk diffusion assay. In the paper disk diffusion assay, the sensitivity of testing microorganisms (bacterial strains including *Escherichia coli*, *Pseudomonas aeruginosa*, *Serratia marcescens*, *Bacillus subtilis*, *Staphylococcus aureus, Micrococcus luteus*; yeast strains including *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kloeckera apiculata*; fungal strains including *Aspergillus niger*, *Aspergillus oryzae*, *Mucor hiemalis*) was tested against **1** and **2** following previously reported procedure.[27] At the dosage of 10 μg/disk, compound **1** showed antimicrobial activities against Gram-positive bacteria including *B. subtilis*, *S. aureus*, and *M. luteus*. Weak antibacterial activity of **1** was observed against *P. aeruginosa*. The compound **1** did not show antibacterial activity against the other testing microorganisms. The inhibitory zones of **1** were following; 9 mm against *B. subtilis*, 9 mm against *S. aureus,* 10.5 mm against *M. luteus,*15 mm against *P. aeruginosa*,. On the other hand, compound **2** did not show any antimicrobial activity. This result indicated that chlorination of Trp was important for the antimicrobial activity. Tetracycline was used as a positive control reagent against the bacterial strains. At the dosage of 50 μg per disk, tetracycline showed inhibitory zones with diameter of 15, 10, 17, 17.5, 22.5, and 30 mm against *E. coli, P. aeruginosa, S. marcescens, B. subtilis, S. aureus,* and *M. luteus*, respectively. Cycloheximide was used as a positive control reagent against yeast and fungal strains. At the dosage of 50 μg per disk, cycloheximide showed inhibitory zones with diameter of 33, 22, and 9 mm against *S. cerevisiae, S. pombe,* and *K. apiculata*, respectively. At the dosage of 50 μg per disk, cycloheximide showed inhibitory zones with diameter of 13.5 and 9 mm against *A. niger*, and *M. hiemalis*, respectively. At the dosage of 200 μg per disk, cycloheximide showed inhibitory zones with diameter of 12 mm against *A. oryzae.*

Discussion

In this study, we found tryptophan halogenase genes locating close to NRPS genes in the genome of *S. curacoi* and *S. noursei*, and then isolated **1** and **2** from these strains. To clarify the relationship between the biosynthetic genes and the peptide compounds, bioinformatics analysis was conducted for the gene clusters of the two strains. The gene organizations of the clusters are shown in Fig. 3. The cluster of *S. noursei* is completely sequenced and encodes two NRPS genes (SNOUR_12645 and SNOUR_12640, pink in Fig. 3b). Genes involved in transport of peptide (SNOUR_12690, SNOUR_12665, and SNOUR_12660, blue in Fig 3b) and regulation (SNOUR_12675 and SNOUR_12670, yellow in Fig 3b) are also present near the NRPS genes. On the other hand, the sequence of the cluster of *S. curacoi* is incomplete and divided into two scaffolds in the draft genome sequence, (Fig. 3a). As same as the gene cluster of *S. nousei* (Fig 3b), the gene cluster of *S. curacoi* includes two NRPS genes (one incomplete NRPS “AQI70_32580 – gap - AQI70_05500”, and one complete NRPS AQI70_05495, pink in Fig 3a). The gene organizations including transport and regulation genes (blue and yellow in Fig. 3a and 3b) are same between the two strains.

The peptide backbones of **1** and **2** are composed of six amino acids: L-ClTrp or L-Trp, D-Val, D-Leu, L-Ile, L-Orn and OHAsn. Consistent with the number of amino acids, the consecutive NRPSs encode a total of six modules (Figure 3). The first NRPS comprises four modules composed of 13 domains (A/PCP-C/A/PCP/E-C/A/PCP/E-C/A/PCP) and the second comprises two modules with seven domains (C/A/PCP-C/A/PCP/E). Analysis using antiSMASH [28, 29] suggested that substrates of A domains in the module-1 to -6 are Trp, Val, Leu, Ile, Orn and Asn, respectively. Presence of E domains in module-2, -3 and -6 accounts for D-configurations of Val, Leu, and Asn. Previously, a dioxygenase SyrP in the biosynthetic gene cluster of syringomycin E was found being involved in hydroxylation of Asp.[30] Interestingly, SyrP was suggested to hydroxylate Asp residue in the premature syringomycin E attached on the NRPS.[30] In the biosynthetic gene clusters of **1** and **2**, there are SyrP-like dioxygenase genes (indicated in purple, AQI70_05465 in Fig 3a and SNOUR_12630 in Fig 3b) which are suggested to be involved in the hydroxylation of Asn to yield OHAsn in the same manner as syringomycin E. Chlorination of Trp can be catalyzed by tryptophan halogenase (AQI70_05465) encoded downstream of the dioxygenase gene. Previously, tryptophan halogenase (ADW94630.1) was reported to chlorinate 6th position in the indole of Trp by accomplishing in vitro experiment.[23] The tryptophan halogenases (AQI70_05465 and SNOUR_12610, green in Fig 3) showed high identities (AQI70_05465, 67%; SNOUR_12610, 65%) in the amino acid sequences to the previously reported tryptophan halogenase (ADW94630.1),[23] respectively. We assumed that the tryptophan halogenase (AQI70_05465) also could accept Trp as a substrate and supply it to NRPS. Interestingly Trp was chlorinated at 5th position of indole in **1,** which was possibly caused by the tryptophan halogenase (AQI70_05465). The tryptophan halogenases (AQI70_05465 and ADW94630.1) belong to flavin-dependent halogenase.[31] The regioselectivities of this class of tryptophan halogenases have been studied very well.[31] X-ray crystallography was performed on tryptophan 7-halogenase (PrnA)[32] and tryptophan 5-halogenase (PyrH).[33] The comparative study on the tryptophan binding site of PyrH and PrnA indicated the loop structures holding the substrate tryptophan in these proteins are different and important for the regioselectivity.[33] The crystal structure of tryptophan 6-halogenase SttH (ADW94630.1) was reported.[34] Interestingly, Structure-guided mutagenesis on SttH resulted in a triple mutant of Trp-recognition region (L460F/P461E/P462T) that exhibited a complete switch in regioselectivity to 5-halogenase.[34] The amino acid sequences are very similar between 6-halogenase SttH (ADW94630.1) and the tryptophan halogenase (AQI70_05465) with high identity (67%). The Trp-recognition region (“LPP”, 460-462th amino acids in SttH) was turned out to be “FES” in the tryptophan halogenase (AQI70_05465) by the alignment analysis. This data explains the difference of regioselectivities of the tryptophan halogenases (ADW94630.1and AQI70_05465), although the heterologous

expression analysis of the tryptophan halogenase (AQI70_05465) is needed.

Above all, the amino acids expected from the biosynthetic pathway (Fig. 4) completely correspond to the compositions of **1** and **2**, strongly suggesting the gene clusters to be involved in syntheses of these compounds. We here propose the biosynthetic pathway as shown in Fig. 4. Each amino acid building block (L-ClTrp or L-Trp, L-Val, L-Leu, L-Ile, L-Orn and L-OHAsn) is converted to aminoacyl adenylate by each A domain and transferred to the adjacent PCP domain within each module to form the corresponding aminoacyl thioesters. The L-Val, L-Leu and L-OHAsn residues loaded onto module-2, -3 and -6 undergo α-carbon epimerization by E domain within each module, and then C domains of module-2 to -6 catalyze five successive N-acylation to yield a L, D, D, L, L, D-hexapeptidyl thioester attached to the PCP domain of module 6. The absolute stereochemistries of Val, Leu and OHAsn residues in **1** and **2** are expected to be D. For the stereochemistries of Val and Leu in **1** and **2**, they were determined to be D which seemed reasonable, although the stereochemistry of OHAsn could not be determined in this study. In association with the release of the linear hexapeptide chain from the PCP domain, cyclization would take place to yield **1** and **2**. Unexpectedly, *S. noursei* did not produce **1** but **2**, although its gene cluster also encodes a tryptophan halogenase gene as that of *S. curacoi*. The reason is not clear, but the halogenase gene in *S. noursei* might be inactive or weakly expressed. Above all, we concluded that the NRPS gene clusters shown in Fig. 3 are responsible for syntheses of **1** and **2**, respectively.

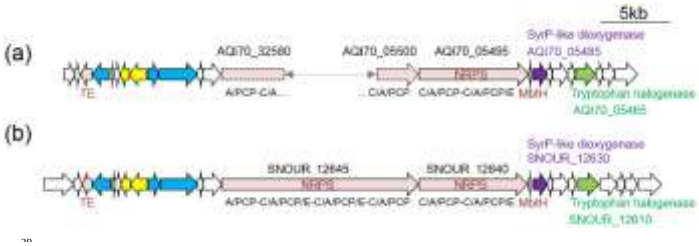


Figure 3. Gene organization of **1**- and **2**-biosynthetic gene clusters found in genome sequences of *Streptomyces curacoi* (a) and *Streptomyces noursei* (b) type strains, respectively. Pink, NRPS and related; green, ClTrp biosynthesis; Purple, OHAsn biosynthesis; yellow, regulation; blue, transport; gray ellipse, sequence gap between AQI70_32580 (LMWJ01000028) and AQI70_05500 (LMWJ01000003); TE, thioesterase. Domain organizations are shown under ORFs encoding NRPS genes. (a) AQI70_32520 to AQI70_32580 in LMWJ01000028 and AQI70_05500 to AQI70_05450 in LMWJ01000003 of the draft genome sequence of *S. curacoi* DSM 40107^T(=NBRC 12761^T). (b) SNOUR_12705 to SNOUR_12590 in CP011533 of the complete genome sequence of *S. noursei* ATCC 11455^T (=NBRC 15452^T).

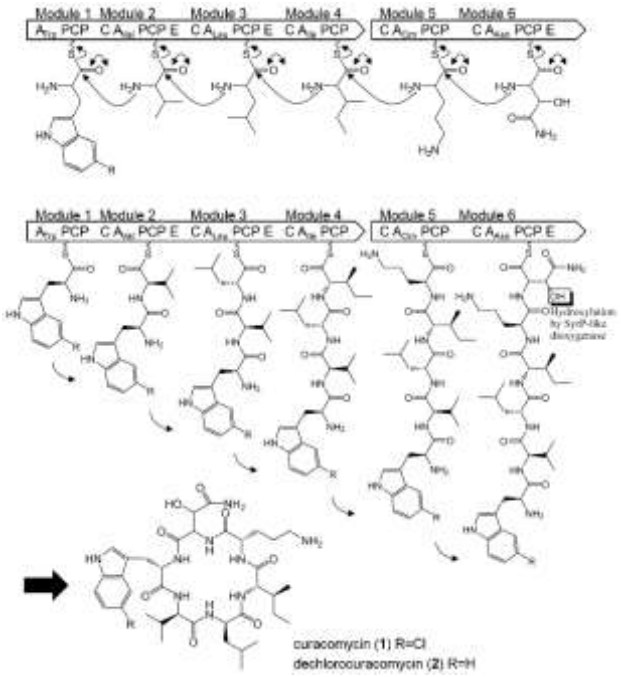


Figure 4. Proposed biosynthetic pathway for **1** and **2**. Capital letters A, C, E PCP, represent adenylation, condensation, epimerization and a peptidyl carrier protein domains, respectively.

Conclusions

We isolated two new peptides curacomycin and dechlorocuracomycin based on genome mining. The structures were determined by NMR and ESI-MS experiments. Curacomycin showed antibacterial activities against several bacteria,while dechlorocuracomycin did not, which indicated the importance of chlorine for antibacterial activity. The biosynthetic pathway of peptides was also proposed from the genome data of producing streptomycetes. The genome mining which focuses on modification enzyme like tryptophane halogenase was indicated to be useful to find new bioactive secondary metabolites.

Experimental Section

Microbial strains: The microorganisms (Bacterial strains including *Streptomyces curacoi* NBRC 12761^T, *Streptomyces noursei* NBRC 15452^T, *Escherichia coli* NBRC 102203^T, *Pseudomonas aeruginosa* NBRC 12689^T, *Serratia marcescens* NBRC102204^T, *Bacillus subtilis* NBRC 13719^T, *Staphylococcus aureus* NBRC 100910^T, *Micrococcus luteus* NBRC 3333^T; Yeast strains including *Saccharomyces cerevisiae* NBRC 2376, *Schizosaccharomyces pombe* NBRC0340, *Kloeckera apiculata* NBRC 0154; fungal strains including *Aspergillus niger* NBRC 33023^T, *Aspergillus oryzae* NBRC 4290, *Mucor hiemalis* NBRC

9405^T) were obtained from the NBRC culture collection (NITE Biological Resource Center, Japan).

Isolation of curacomycin (1): *Streptomyces curacoi* was cultured in 2 L of ISP2 [35] agar medium. After incubation at 30˚C for 10 days, the aerial hyphae and spore cells on the agar surface were harvested. For extraction, double volume of MeOH was added to the harvested cells, followed by filtration using filter paper (Whatman No.1, GE Healthcare Life Sciences, Little Chalfont, UK). The MeOH extract was concentrated by rotary evaporator and was subjected to open column chromatography (styrene-divinylbenzene resin, CHP-20P, Mitsubishi Chemical Corp., Tokyo, Japan) eluted with 10%MeOH, 60%MeOH and 100%MeOH. The 100%MeOH fraction was concentrated by rotary evaporator and subjected to HPLC separation using an ODS column (4.6 × 250 mm, Wakopak Handy ODS, Wako Pure Chemical Industries, Tokyo, Japan) with isocratic elution of 35% MeCN containing 0.05% TFA (trifluoroacetic acid) at flow rate of 1 mL /min. The UV detector of HPLC was set at the wavelength of 220 nm to yield 4.0 mg of **1** (Retention time; 17 min)

Isolation of dechlorocuracomycin (2): Nutrient agar medium supplemented with leucine (peptone 5 g, beef extract 3 g, NaCl 5g, leucine 1g, agar 15 g in 1 L of distilled water, pH 7.3) was used for the cultivation of *Streptomyces noursei*. After incubation at 30 ˚C for 7 days, aerial hyphae and spore cells were harvested and extracted with double volume of MeOH. For HPLC analysis, compound **2** was isolated using an ODS column (4.6 × 250 mm, Wakopak Handy ODS, Wako) with isocratic elution of 32% MeCN containing 0.05% TFA at flow rate of 1 mL /min. The UV detector of HPLC was set at the wavelength of 220 nm to yield 3.0 mg of **2** (Retention time; 15.6 min).

NMR analysis: A NMR sample was prepared by dissolving curacomycin (**1**) or dechlorocuracomycin (**2**), respectively, in 500 µl of MeCN-d_3:DMSO-d_6 = 4:1. One-dimensional (1D) ^{1}H, ^{13}C, and DEPT-135 spectra were obtained on Bruker Avance III HD 800 spectrometer. The 1D ^{1}H, ^{13}C, DEPT-135 spectra were recorded at 25 °C with 15 ppm for proton and 239 ppm or 222 ppm for carbon. All two-dimensional (2D) ^{1}H-NMR spectra were obtained on Bruker Avance 600 and Avance III HD 800 spectrometers with quadrature detection in the phase-sensitive mode by States-TPPI (time proportional phase incrementation) or in the echo-antiecho mode. The following spectra were recorded at 25 °C with 15 ppm spectral widths in t1 and t2 dimensions: 2D double quantum filtered correlated spectroscopy (DQF-COSY), recorded with 512 and 1024 complex points in t1 and t2 dimensions; 2D homonuclear total correlated spectroscopy (TOCSY) with DIPSI2 mixing sequence, recorded with mixing time of 80 ms, 512 and 1024 complex points in t1 and t2 dimensions; 2D nuclear Overhauser effect spectroscopy (NOESY), recorded with mixing times of 200 and 400 ms, 512 and 1024 complex points in t1 and t2 dimensions. 2D ^{1}H-^{13}C heteronuclear single quantum correlation (HSQC) and heteronuclear multiple bond connectivity (HMBC) spectra were acquired at 25 °C in the echo-antiecho mode. The ^{1}H-^{13}C HSQC and HMBC spectra were recorded with 1024 × 512 complex points for 15 ppm in the ^{1}H dimension and 160 ppm in the ^{13}C dimension or for 15 ppm in the ^{1}H dimension and 222 ppm in the ^{13}C dimension, respectively, at a natural isotope abundance. 2D ^{1}H-^{15}N HSQC spectrum was recorded with 1024 × 128 complex points for 15 ppm in the ^{1}H dimension and 40 or 60 ppm in the ^{15}N dimension at a natural isotope abundance. All NMR spectra were processed using TOPSPIN 3.5 (Bruker). Before Fourier transformation, the shifted sinebell window function was applied to t1 and t2 dimensions. All ^{1}H and ^{13}C dimensions were referenced to MeCN-d_3 at 25°C.

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

Modified Marfey's method: Each peptide was subjected to acid hydrolysis with 6N HCl containing 3% Phenol at 166 °C for 25 min to recover Trp following previous report [36]. The hydrolysate was completely evaporated by rotary evaporator and resuspended in H_2O (200µL). The 10 µL of solution of Nα-(5-fluoro-2,4-dinitrophenyl)-L-leucinamide (L-FDLA, Tokyo Chemical Industry Co., LTD, Tokyo, Japan) in acetone (10 µg/µL) and 100 µL of 1M $NaHCO_3$ solution were added to the hydrolysate and the mixture was incubated at 80 °C for 3 min. The reaction mixture was cooled down at room temperature before neutralized with 50 µL of 2N HCl and diluted with 1 mL of 50% MeCN. For standard amino acid, each amino acid was derivatized with L-FDLA and D-FDLA in the same method. The standard amino acids including Orn, Ile, Leu, Val, Trp were purchased from Wako chemical. The unusual amino acid L-chloro-tryptophan was purchased from Carbosynth limited (Berkshire, UK). Approximately 30 µL of each FDLA derivatives was subjected to HPLC analysis with C18 column (4.6 × 250 mm, Wakopak Handy ODS, Wako). The DAD detector (MD-2018, JASCO, Tokyo, Japan) was used for detection of the amino acid derivatives accumulating the data of the absorbance from 220nm to 420 nm. The HPLC analysis was performed at a flow rate of 1 mL/min using solvent A (distilled water containing 0.05% TFA) and solvent B (MeCN containing 0.05%TFA) with a linear gradient mode from 0 min to 70 min, increasing percentage of solvent B from 25% to 60%. The retention times (min) of L- or D-FDLA derivatized amino acids in this HPLC condition were following; L-Orn-D-FDLA (13.97 min), L-Orn-L-FDLA (18.47 min), L-Val-L-FDLA (37.35 min), L-allo-Ile-L-FDLA (42.55 min), L-Ile-L-FDLA (42.92 min), L-Trp-L-FDLA (44.25 min), L-Leu-L-FDLA (43.35 min), L-ClTrp-L-FDLA (49.61 mm), L-Trp-D-FDLA (51.20 min), L-Val-D-FDLA (51.55 min), L-ClTrp-D-FDLA (56.44 mm), L-allo-Ile-D-FDLA (58.27 min), L-Ile-D-FDLA (58.41 min), and L-Leu-D-FDLA (58.88 min).

Antimicrobial assay: The antimicrobial activity of the compound was tested by using a paper disk diffusion assay (6 mm diameter, thick type) against testing microorganisms including; *E. coli*, *P. aeruginosa*, *S. marcescens*, *B. subtilis*, *S. aureus*, *M. luteus*, *S. cerevisiae*, *S. pombe*, *K. apiculata*, *A. niger*, *A. oryzae* and *M. hiemalis*. The compound **1** or **2** was dissolved in DMSO at the concentration 10 mg/mL and 1 mg/mL. The testing microorganisms were cultured using nutrient agar medium (peptone 5 g, beef extract 3 g, NaCl 5g, leucine 1g, agar 15 g in 1 L of distilled water, pH 7.3). After all testing microorganisms were inoculated onto nutrient agar medium, paper disks with 10 µL of each compound (1 µg/µL) in DMSO were placed onto the surface of the agar medium, and paper

disk with 10 µL DMSO was used as a negative control. After incubation at 30 ˚C for 2 days, the formation of inhibition zone was measured for evaluation of antimicrobial activity.

Acknowledgements

This study was supported by the Japan Society for the Promotion of Science by Grants-in-aids (grant number 16K01913).

Keywords: genome mining • Tryptophan halogenase • Nonribosomal peptide synthetase• cyclic peptide• *Streptomyces curacoi*• *Streptomyces noursei*

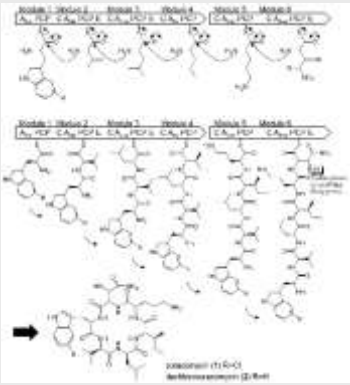
[1] D. Schwarzer, R. Finking, M. A. Marahiel, *Nat Prod Rep* **2003**, *20*, 275-287.
[2] S. A. Sieber, M. A. Marahiel, *Chem. Rev.* **2005**, *105*, 715-738.
[3] S. Zhu, J. D. Hegemann, C. D. Fage, M. Zimmermann, X. Xie, U. Linne, M. A. Marahiel, *J Biol Chem* **2016**, *291*, 13662-13678.
[4] C. T. Walsh, *Nat Prod Rep* **2016**, *33*, 127-135.
[5] S. Sundaram, C. Hertweck, *Curr Opin Chem Biol* **2016**, *31*, 82-94.
[6] R. D. Sussmuth, A. Mainz, *Angew Chem Int Ed Engl* **2017**, *56*, 3770-3821.
[7] G. L. Challis, *Microbiology* **2008**, *154*, 1555-1569.
[8] A. Hornung, M. Bertazzo, A. Dziarnowski, K. Schneider, K. Welzel, S. E. Wohlert, M. Holzenkampfer, G. J. Nicholson, A. Bechthold, R. D. Sussmuth, A. Vente, S. Pelzer, *Chembiochem* **2007**, *8*, 757-766.
[9] E. Zazopoulos, K. Huang, A. Staffa, W. Liu, B. O. Bachmann, K. Nonaka, J. Ahlert, J. S. Thorson, B. Shen, C. M. Farnet, *Nat Biotechnol* **2003**, *21*, 187-190.
[10] C. Corre, G. L. Challis, *Nat Prod Rep* **2009**, *26*, 977-986.
[11] K. J. Hetrick, W. A. van der Donk, *Curr Opin Chem Biol* **2017**, *38*, 36-44.
[12] B. Aigle, S. Lautru, D. Spiteller, J. S. Dickschat, G. L. Challis, P. Leblond, J. L. Pernodet, *J Ind Microbiol Biotechnol* **2014**, *41*, 251-263.
[13] H. Mohimani, W. T. Liu, R. D. Kersten, B. S. Moore, P. C. Dorrestein, P. A. Pevzner, *J Nat Prod* **2014**, *77*, 1902-1909.
[14] K. Kleigrewe, J. Almaliti, I. Y. Tian, R. B. Kinnel, A. Korobeynikov, E. A. Monroe, B. M. Duggan, V. Di Marzo, D. H. Sherman, P. C. Dorrestein, L. Gerwick, W. H. Gerwick, *J Nat Prod* **2015**, *78*, 1671-1682.
[15] C. Kurth, S. Schieferdecker, K. Athanasopoulou, I. Seccareccia, M. Nett, *J Nat Prod* **2016**, *79*, 865-872.
[16] Hindra, T. Huang, D. Yang, J. D. Rudolf, P. Xie, G. Xie, Q. Teng, J. R. Lohman, X. Zhu, Y. Huang, L. X. Zhao, Y. Jiang, Y. Duan, B. Shen, *J Nat Prod* **2014**, *77*, 2296-2303.
[17] C. Wu, C. Du, K. Ichinose, Y. H. Choi, G. P. van Wezel, *J Nat Prod* **2017**, *80*, 269-277.
[18] P. Gao, Y. Huang, *Appl Environ Microbiol* **2009**, *75*, 4813-4820.
[19] G. W. Gribble, *J. Chem. Edu.* **2004**, *81*, 1441-1449.
[20] G. P. Dinos, C. M. Athanassopoulos, D. A. Missiri, P. C. Giannopoulou, I. A. Vlachogiannis, G. E. Papadopoulos, D. Papaioannou, D. L. Kalpaxis, *Antibiotics (Basel)* **2016**, *5*.
[21] D. P. Levine, *Clin Infect Dis* **2006**, *42 Suppl 1*, S5-12.
[22] K. E. Henson, M. T. Levine, E. A. Wong, D. P. Levine, *Expert Rev Anti Infect Ther* **2015**, *13*, 1265-1278.
[23] J. Zeng, J. Zhan, *Biotechnol Lett* **2011**, *33*, 1607-1613.
[24] O. L. D. V. Galmarini, *Tetrahedron* **1961**, *15*, 76-86.
[25] P. C. Gotzsche, H. K. Johansen, *Cochrane Database Syst Rev* **2014**, CD002033.
[26] K. I. Harada, K. Fujii, K. Hayashi, M. Suzuki, Y. Ikai, H. Oka, *Tetrahedron Lett.* **1996**, *37*, 3001-3004.
[27] S. Kodani, Y. Inoue, M. Suzuki, H. Dohra, T. Suzuki, H. Hemmi, M. Ohnishi-Kameyama, *Eur J Org Chem* **2017**, 1177-1183.
[28] T. Weber, K. Blin, S. Duddela, D. Krug, H. U. Kim, R. Bruccoleri, S. Y. Lee, M. A. Fischbach, R. Muller, W. Wohlleben, R. Breitling, E. Takano, M. H. Medema, *Nucleic Acids Res* **2015**, *43*, W237-243.
[29] K. Blin, M. H. Medema, R. Kottmann, S. Y. Lee, T. Weber, *Nucleic Acids Res* **2017**, *45*, D555-D559.
[30] G. M. Singh, P. D. Fortin, A. Koglin, C. T. Walsh, *Biochemistry* **2008**, *47*, 11310-11320.
[31] V. Weichold, D. Milbredt, K. H. van Pee, *Angew Chem Int Ed Engl* **2016**, *55*, 6374-6389.
[32] C. Dong, S. Flecks, S. Unversucht, C. Haupt, K. H. van Pee, J. H. Naismith, *Science* **2005**, *309*, 2216-2219.
[33] X. Zhu, W. De Laurentis, K. Leang, J. Herrmann, K. Ihlefeld, K. H. van Pee, J. H. Naismith, *J Mol Biol* **2009**, *391*, 74-85.
[34] S. A. Shepherd, B. R. Menon, H. Fisk, A. W. Struck, C. Levy, D. Leys, J. Micklefield, *Chembiochem* **2016**, *17*, 821-824.
[35] E. B. Shirling, D. Gottlieb, *Int. J. Syst. Bacteriol.* **1966**, *16*, 313-340.
[36] K. Muramoto, H. Kamiya, *Anal Biochem* **1990**, *189*, 223-230.

Entry for the Table of Contents (Please choose one layout)

Layout 1:

FULL PAPER

A new antibacterial cyclic peptide named curacomycin and its analogue dechlorocuracomycin were isolated based on genome mining. The biosynthetic pathways of the peptides were proposed by analysis on biosynthetic gene clusters found in the genome sequences.



Key Topic*: genome mining, NRPS

Issara Kaweewan, Hisayuki Komaki, Hikaru Hemmi, Shinya Kodani*

Page No. – Page No.

Title

*one or two words that highlight the emphasis of the paper or the field of the study

Layout 2:

FULL PAPER

((Insert TOC Graphic here; max. width: 11.5 cm; max. height: 2.5 cm; NOTE: the final letter height should not be less than 2 mm.))

Key Topic*

*Author(s), Corresponding Author(s)**

Page No. – Page No.

Title

Text for Table of Contents

*one or two words that highlight the emphasis of the paper or the field of the study