Available online at www.sciencedirect.com
ScienceDirect
Tetrahedron
ELSEVIER
Tetrahedron 62 (2006) 8907–8918

Chemical synthesis of ComX pheromone and related peptides containing isoprenoidal tryptophan residues

Masahiro Okada,[a,*,†] Isao Sato,[a] Soo Jeong Cho,[b] David Dubnau[b] and Youji Sakagami[a,*]

[a]*Graduate School of Bioagricultural Sciences, Nagoya University, Chikusa, Nagoya 464-8601, Japan*
[b]*Public Health Research Institute, 225 Warren Street, Newark, NJ 07103, USA*

Received 18 May 2006; revised 17 June 2006; accepted 21 June 2006
Available online 28 July 2006

Abstract—The ComX pheromone is a post-translationally modified oligopeptide that stimulates natural genetic competence controlled by quorum sensing in *Bacillus subtilis*. Recently, the structure of the $ComX_{RO\text{-}E\text{-}2}$ pheromone produced by strain RO-E-2 was determined. Based on the NMR analysis, a geranyl group is bound to the tryptophan residue, which results in the formation of a tricyclic ring structure. It was proposed that one of the four possible stereochemical isomers was based on a conformational search for model compounds and the assumption that amino acid residues in the natural pheromone have the L-configuration. All possible modified tryptophan residues and the corresponding $ComX_{RO\text{-}E\text{-}2}$ peptides were synthesized to confirm the precise stereochemistry. Here, the synthesis of the modified tryptophan derivatives was reported in detail. It was succeeded in synthesizing four optically active modified tryptophan methyl esters from which the four diastereomeric $ComX_{RO\text{-}E\text{-}2}$ peptides were prepared. Since only one of the four diastereomers was spectroscopically identical to the natural pheromone and exhibited biological activity, the absolute structure of the $ComX_{RO\text{-}E\text{-}2}$ pheromone was able to be established unambiguously. Furthermore, it was noticed that two other bioactive pheromones were present in the culture broth that were co-purified with $ComX_{RO\text{-}E\text{-}2}$ pheromone. These pheromones were presumed to be the N-terminal truncated peptides of $ComX_{RO\text{-}E\text{-}2}$ pheromone, i.e., [2-6]$ComX_{RO\text{-}E\text{-}2}$ and [3-6]$ComX_{RO\text{-}E\text{-}2}$, by LC-MS and NMR analyses. Using Fmoc solid-phase peptide synthesis, $ComX_{RO\text{-}E\text{-}2}$ pheromone and the [2-6]$ComX_{RO\text{-}E\text{-}2}$ and [3-6]$ComX_{RO\text{-}E\text{-}2}$ peptides were prepared. The synthetic peptides were identical to the natural pheromones and also showed significant biological activity.

1. Introduction

Bacteria constitutively secrete specific pheromones with the pheromone concentration increasing with cell density. When these concentrations reach threshold levels, bacteria respond to the pheromone by altering their gene expression. The regulation that depends on cell density is called 'quorum sensing'.[1–4] The quorum sensing pheromones are oligopeptides in Gram-positive bacteria, while Gram-negative pheromones are mainly *N*-acylhomoserine lactones. *Bacillus subtilis* and related bacilli produce a unique pheromone that stimulates natural genetic competence controlled by the quorum sensing.[5] This *B. subtilis* competent factor is a post-translationally modified oligopeptide known as ComX pheromone. The ComX pheromone is thought to activate the signal transduction cascade for natural genetic competence by binding to the membrane-embedded histidine autokinase ComP.[6,7]

The amino acid sequence of the ComX pheromone varies according to the *B. subtilis* strain, but each possesses a modified tryptophan residue.[8–11] This modification increases the hydrophobicity and molecular weight of the peptide encoded by *comX*.[5] ComX is biosynthesized as a pre-protein, which is then processed and modified by ComQ to become the active ComX pheromone.[12] ComQ contains an isoprenyl transferase domain, and this putative isoprenoid-binding site in ComQ is required for the expression of pheromone activity.[13] Furthermore, [^{3}H]-5-mevalonate, a precursor of isoprenoid, is incorporated into ComX pheromone.[10] These results

Keywords: *Bacillus subtilis*; ComX; Post-translational modification; Quorum sensing; Tryptophan.
Abbreviations: Clt, 2-chlorotrityl; [Dα]$ComX_{RO\text{-}E\text{-}2}$, $ComX_{RO\text{-}E\text{-}2}$ peptide containing the modified D-tryptophan residue with an α-geranyl group; [Dβ]$ComX_{RO\text{-}E\text{-}2}$, $ComX_{RO\text{-}E\text{-}2}$ peptide containing the modified D-tryptophan residue with a β-geranyl group; DIPEA, *N*,*N*′-diisopropylethylamine; HATU, *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HBTU, *O*-benzotriazol-1-yl-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate; [Lα]$ComX_{RO\text{-}E\text{-}2}$, $ComX_{RO\text{-}E\text{-}2}$ peptide containing the modified L-tryptophan residue with an α-geranyl group; [Lβ]$ComX_{RO\text{-}E\text{-}2}$, $ComX_{RO\text{-}E\text{-}2}$ peptide containing the modified L-tryptophan residue with a β-geranyl group; Pp, 2-phenyl-2-propyl; Su, succinimidyl; TAS-F, tris(dimethylamino)sulfonium difluorotrimethylsilicate; Teoc, 2-(trimethylsilyl)-ethoxycarbonyl.

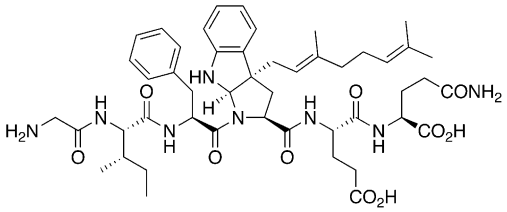
* Corresponding authors. Tel.: +81 22 795 6555; fax: +81 22 795 6557; e-mail addresses: okada-org@mail.tains.tohoku.ac.jp; ysaka@agr.nagoya-u.ac.jp
† Present address: Graduate School of Science, Tohoku University, Aoba, Sendai 980-8578, Japan.

0040–4020/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tet.2006.06.074

indicate that ComX pheromone is activated by isoprenylation from an inactive precursor. Isoprenylation of a tryptophan residue during post-translational modification has not been previously reported, although the modification of cysteine residues in proteins by farnesyl or geranylgeranyl groups is well known.[14]

We recently isolated the $ComX_{RO\text{-}E\text{-}2}$ pheromone using an *Escherichia coli* expression system. This required only *comQXP′*$_{RO\text{-}E\text{-}2}$ from *B. subtilis* strain RO-E-2, which was driven by a T7 phage promoter to produce biologically active $ComX_{RO\text{-}E\text{-}2}$ pheromone.[15] We determined the planar structure of the tryptophan residue by NMR analysis, which showed a tricyclic ring structure with a geranyl group attached to the 3-position of tryptophan. A conformational search was then carried out on the L-tryptophan derivative model compound using a Monte Carlo protocol to elucidate the stereochemistry of $ComX_{RO\text{-}E\text{-}2}$ pheromone. This search was based on the assumption that the pheromone was composed only of L-amino acids. We found that only one stereostructure could explain the NMR data. We thus proposed the absolute structure of the $ComX_{RO\text{-}E\text{-}2}$ pheromone **1** and its modified tryptophan residue to be [2*S*,3a*R*,8a*S*]-3α-geranyl-1,2,3,3α,8,8α-hexahydropyrrolo[2,3-*b*]indole-2-carboxylic acid [L-Trp*(α-Ger)] (Fig. 1). To confirm the structure of the $ComX_{RO\text{-}E\text{-}2}$ pheromone, synthesis of all possible $ComX_{RO\text{-}E\text{-}2}$ peptides and investigation of their biological activities and physiological data were necessary. There are four possible structures **2**–**5** for the modified tryptophan residue (Fig. 2), i.e., [Lα]$ComX_{RO\text{-}E\text{-}2}$ **2** has the modified L-tryptophan residue with an α-geranyl group [L-Trp*(α-Ger)], [Lβ]$ComX_{RO\text{-}E\text{-}2}$ **3** has the modified L-tryptophan residue with a β-geranyl group [L-Trp*(β-Ger)], [Dα]$ComX_{RO\text{-}E\text{-}2}$ **4** has the modified D-tryptophan residue with an α-geranyl group [D-Trp*(α-Ger)], and [Dβ]$ComX_{RO\text{-}E\text{-}2}$ **5** has the modified D-tryptophan residue with a β-geranyl group [D-Trp*(β-Ger)]. We therefore synthesized four optically active geranyl tryptophan methyl esters, and prepared four diastereomeric $ComX_{RO\text{-}E\text{-}2}$ peptides to confirm our proposed structure for the $ComX_{RO\text{-}E\text{-}2}$ pheromone.

Two additional peptides were co-isolated with $ComX_{RO\text{-}E\text{-}2}$ pheromone from the culture broth of *E. coli* ED413. These were determined to be N-terminal truncated peptides of $ComX_{RO\text{-}E\text{-}2}$ pheromone. Using solid-phase peptide synthesis as described previously,[16] we synthesized these two peptides in addition to $ComX_{RO\text{-}E\text{-}2}$ pheromone and investigated their biological activity.



H-Gly-Ile-Phe-Trp*(Ger)-Glu-Gln-OH

$ComX_{RO\text{-}E\text{-}2}$ pheromone **1**

Figure 1. The proposed structure of $ComX_{RO\text{-}E\text{-}2}$ pheromone.

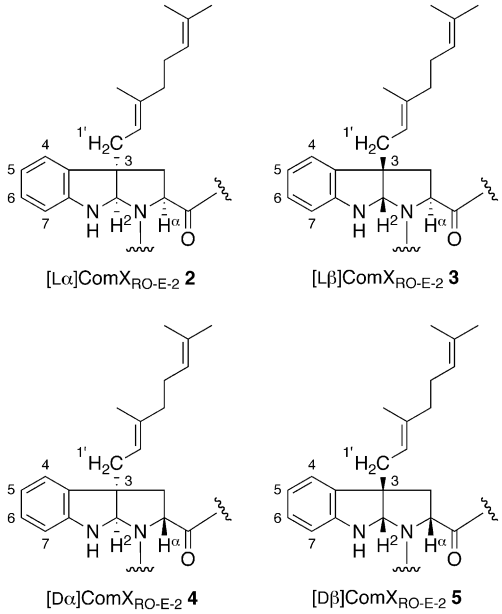


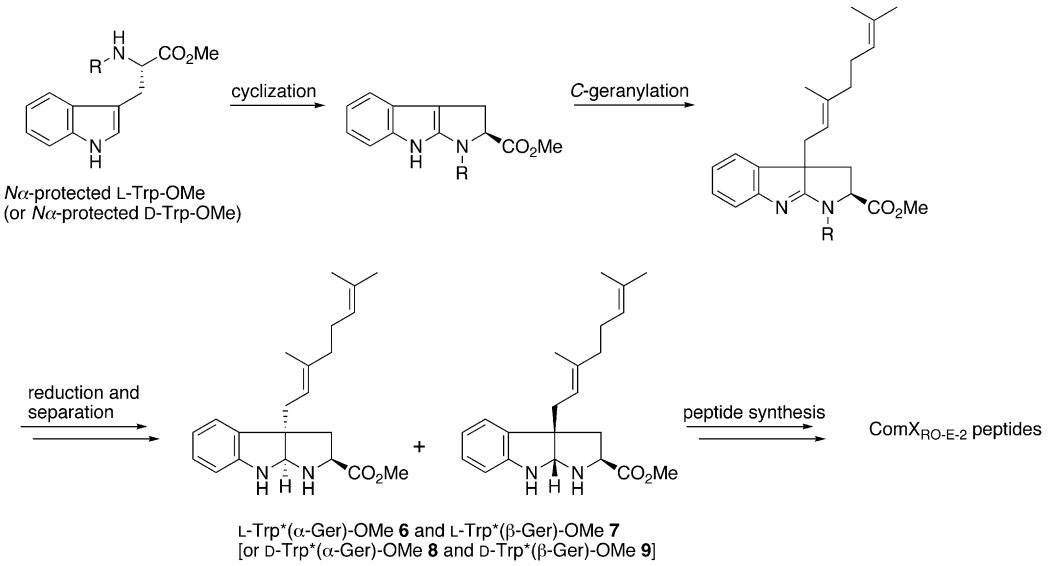
Figure 2. Four possible structures of the modified tryptophan residue in the $ComX_{RO\text{-}E\text{-}2}$ pheromone based on the NMR analysis.

2. Results

2.1. Synthesis of modified tryptophan and stereoisomers

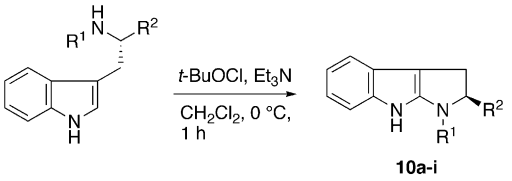
For efficient synthesis of the four stereoisomeric geranyl-modified tryptophan residues and the corresponding peptides composed of the $ComX_{RO\text{-}E\text{-}2}$ sequence, we selected L- or D-tryptophan methyl esters as chiral starting materials and geranyl bromide as the geranylating reagent. The optically active tricyclic tryptophan derivatives were synthesized by cyclization of tryptophan methyl ester derivatives, followed by geranylation at the C3 position. The resulting C3 diastereomers were separated, and each isomer was converted to the protected tryptophan derivatives **6**–**9** prior to peptide synthesis (Scheme 1).

The cyclization of *N*-acetyl L-tryptophan methyl ester was carried out with *t*-BuOCl (Table 1, entry 1).[17,18] The geranyl group was then introduced at the C3 position of the tricyclic compound **10a** with NaH and geranyl bromide in DMF (Table 2, entry 1).[18] Unfortunately, the acetyl deprotection of the C3-geranylated diastereomers did not proceed under the various conditions tried (data not shown). Thus, we screened for removable protecting groups and conditions for C3-geranylation. Using N-protected tryptophan derivatives, we synthesized nine tricyclic compounds **10a**–**i** that were purified by recrystallization (Table 1).[19] Using **10a**–**i**, we optimized the conditions for C3-geranylation (Table 2).[20] The usual quenching method with methanol or water resulted in product decomposition, and quenching with 5% aqueous $KHSO_4$ gave the C3-β-geranylated compound in poor yield in addition to the N1-geranylated compound. We found that immediate neutralization with a phosphate buffer prevented the decomposition of the C3-geranylated products. When carbamates were selected as the N-protecting group, C-geranylation of **10b**–**e** did not proceed at all (Table 2). We also found that the ester was important for the reaction because the geranylation of 2-chlorotrityl (Clt)



Scheme 1. Synthetic strategy toward four $ComX_{RO\text{-}E\text{-}2}$ peptides.

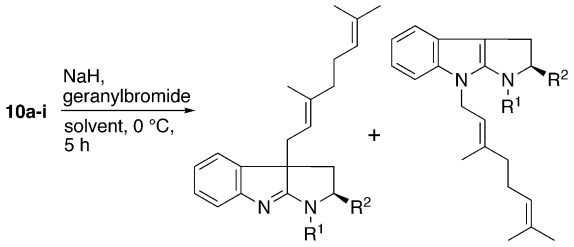
Table 1. Cyclization of various tryptophan derivatives



Entry	R^1	R^2	Yield, %[a] (product)
1	Ac	CO_2Me	92 (**10a**)
2	Boc	CO_2Me	42 (**10b**)
3	CO_2Bn	CO_2Me	49 (**10c**)
4	CO_2Me	CO_2Me	78 (**10d**)
5	CO_2Et	CO_2Me	69 (**10e**)
6	$COCCl_3$	CO_2Me	85 (**10f**)
7	$COCF_3$	CO_2Me	79 (**10g**)
8	Bz	CO_2Me	95 (**10h**)
9	Bz	CH_2OClt	62 (**10i**)

[a] Isolated yield with recrystallization.

Table 2. *C*- or *N*-selective geranylation under various conditions

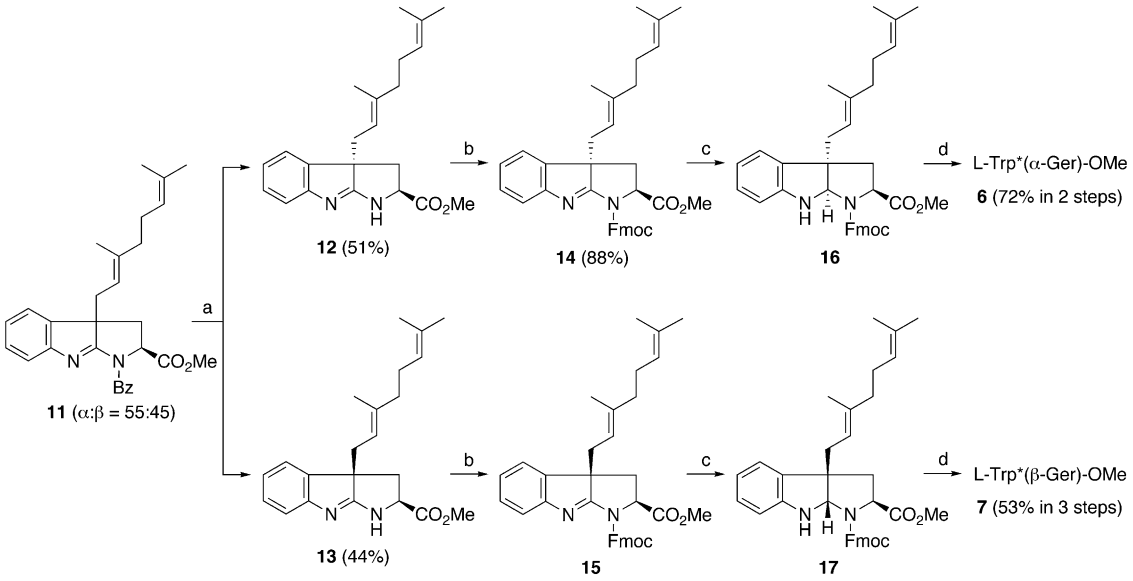


Substrate (R^1, R^2)	Solvent	Yield, %[a]	
		C-Ger (α:β)[b]	*N*-Ger
10a (Ac, CO_2Me)	DMF	63 (33:30)	9
10b (Boc, CO_2Me)	DMF	0	0
10c (CO_2Bn, CO_2Me)	DMF	0	0
10d (CO_2Me, CO_2Me)	DMF	0	14
10e (CO_2Et, CO_2Me)	DMF	0	8
10f ($COCCl_3$, CO_2Me)	DMF	0	0
10g ($COCF_3$, CO_2Me)	DMF	0	0
10h (Bz, CO_2Me)	DMF	81 (44:37)	15
10i (Bz, CH_2OClt)	DMF	0	72
10h (Bz, CO_2Me)	NMP	77 (32:45)	18
10h (Bz, CO_2Me)	THF	83 (47:36)	11

[a] Isolated yield.
[b] α/β ratios were determined by ^1H NMR, and the stereochemistries were determined by NOE.

ether **10i** gave only the N-geranylated compound. The Bz-protected substrate **10h** gave a better yield of C3-geranylated diastereomers when compared with **10a**. The solvent for geranylation also affected the yield and diastereomeric ratio of the geranylated products. THF increased the yield slightly and the ratio of the C3-α-geranyl product improved, but the reverse was true for NMP. The key intermediate **11** was obtained in high yield using the optimized conditions. Mixtures of the C3-geranyl isomers were acceptable for the purpose of synthesizing the four diastereomers. Since the C3-geranylated diastereomers could not be separated by column chromatography, the diastereomixture **11** was used directly as the starting material in the next step.

Next, deprotection, separation of diastereomers, and chemoselective reduction of the imine were attempted (Scheme 2). Treatment of **11** with DIBAL-H selectively cleaved the Bz group. The reductive diastereomeric products **12** and **13** were easily separated by column chromatography using a solvent system of chloroform and methanol to afford the optically active tryptophan derivatives **12** and **13**. We determined the α-geranyl configuration of **12** based on NOE analysis, however purified **13** was not sufficiently stable for direct determination of its stereochemistry. Because compound **13** was readily prepared from the β-geranyl isomer of **11** by treatment with DIBAL-H in nearly quantitative yield, and compound **13** was also determined to have the β-geranyl configuration. It was necessary to protect the unstable secondary amines **12** and **13** with an Fmoc group before the reduction of the imine. After Fmoc protection, the reduction of the imines **14** and **15** was carried out with catecholborane to give amines **16** and **17**, respectively. Fmoc was found to be the best protecting group for the reduction. Bz-protected **11** required prolonged reduction with catecholborane and was difficult to deprotect,[20] and the free amine **12** was

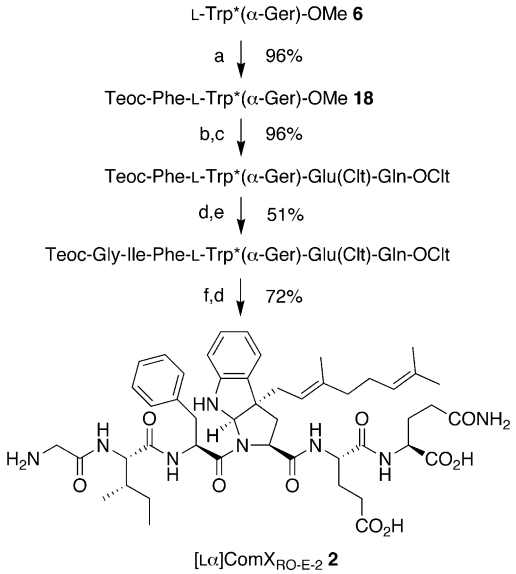


Scheme 2. Synthesis of L-Trp*(α-Ger)-OMe **6** and L-Trp*(β-Ger)-OMe **7**. Reaction conditions: (a) DIBAL-H, THF, –78 °C, 2 h; (b) Fmoc-OSu, 1.0 M aqueous $NaHCO_3$, dioxane, rt, 1.5 h; (c) catecholborane, THF, 0 °C, 2 h; (d) piperidine, CH_3CN, rt, 1.5 h.

decomposed by reduction. Cleavage of the Fmoc group gave L-Trp*(α-Ger)-OMe **6** for the synthesis of [Lα]$ComX_{RO-E-2}$ **2**, which has our proposed configuration for natural $ComX_{RO-E-2}$ pheromone. We confirmed that L-Trp*(α-Ger)-OMe **6** had a cis configuration based on the NOE between the H-2 proton and the geranyl group. We also synthesized L-Trp*(β-Ger)-OMe **7** from intermediate **13** using the same procedure (Scheme 2), and D-Trp*(α-Ger)-OMe **8** and D-Trp*(β-Ger)-OMe **9** from the corresponding *N*-benzoyl-D-tryptophan methyl esters.

2.2. Synthesis of $ComX_{RO-E-2}$ peptides

We synthesized peptides containing the same amino acid sequence as the $ComX_{RO-E-2}$ pheromone using each modified tryptophan methyl ester (Scheme 3). The $ComX_{RO-E-2}$ pheromone is very acid labile.[15,16] Therefore, we used Clt as a protecting group for the carboxy groups because it can be cleaved under mild acidic conditions. Methyl ester **6** was treated with 2-(trimethylsilyl)ethoxycarbonylphenylalanine (Teoc-Phe) to obtain dipeptide **18**. From the analysis of the HMBC spectrum, we found that L-Trp*(α-Ger)-OMe **6** reacts with amino acids only at the *N*α-position under common amino acid coupling conditions (data not shown). Dipeptide methyl ester **18** was hydrolyzed under alkaline conditions, followed by coupling with Glu(Clt)-Gln-OClt. Purification of the resulting tetrapeptide was carried out by silica gel column chromatography with 1% triethylamine so as not to cleave the Clt esters. After cleavage of the Teoc group of the purified tetrapeptide, the peptide was coupled to Teoc-Gly-Ile. Finally, the Clt esters were cleaved with 50% aqueous AcOH, followed by removal of the Teoc group with tris(dimethylamino)sulfonium difluorotrimethylsilicate (TAS-F). The resultant compound was purified by HPLC to give the desired [Lα]$ComX_{RO-E-2}$ peptide **2**. The other three diastereomers, [Lβ]$ComX_{RO-E-2}$ peptide **3**, [Dα]$ComX_{RO-E-2}$ peptide **4**, and [Dβ]$ComX_{RO-E-2}$ peptide **5**, were synthesized using L-Trp*(β-Ger)-OMe **7**, D-Trp*(α-Ger)-OMe **8**, and D-Trp*(β-Ger)-OMe **9**, respectively, as shown in Figure 3.



Scheme 3. Synthesis of [Lα]$ComX_{RO-E-2}$ **2**. Reaction conditions: (a) Teoc-Phe, HATU, HOAt, Et_3N, CH_2Cl_2, 0 °C, 2 h; (b) LiOH, THF, MeOH, H_2O, rt, 1 h; (c) Glu(Clt)-Gln-OClt, HATU, HOAt, Et_3N, CH_2Cl_2, 0 °C, 2 h; (d) TAS-F, CH_2Cl_2, rt, 24 h; (e) Teoc-Gly-Ile, HATU, HOAt, Et_3N, CH_2Cl_2, 0 °C, 2 h; (f) 50% aqueous AcOH, 4 °C, 24 h.

2.3. Comparison of synthetic $ComX_{RO-E-2}$ peptides and natural pheromone

The ^1H NMR spectra of the four synthetic peptides were compared (Fig. 4). [Lα]$ComX_{RO-E-2}$ peptide **2** possessed the proposed stereostructure and exhibited the same ^1H NMR spectrum as the natural peptide pheromone. The spectra of the three other diastereomers **3**–**5** showed considerable differences, particularly in the region of the modified tryptophan residues.

We then investigated the biological activities of these peptides by β-galactosidase assay with *o*-nitrophenyl-β-D-

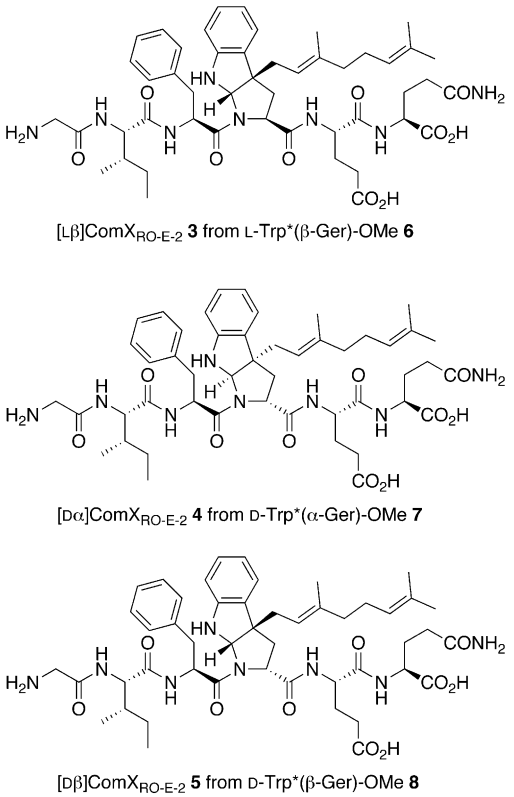


Figure 3. Other three ComX$_{RO-E-2}$ peptides.

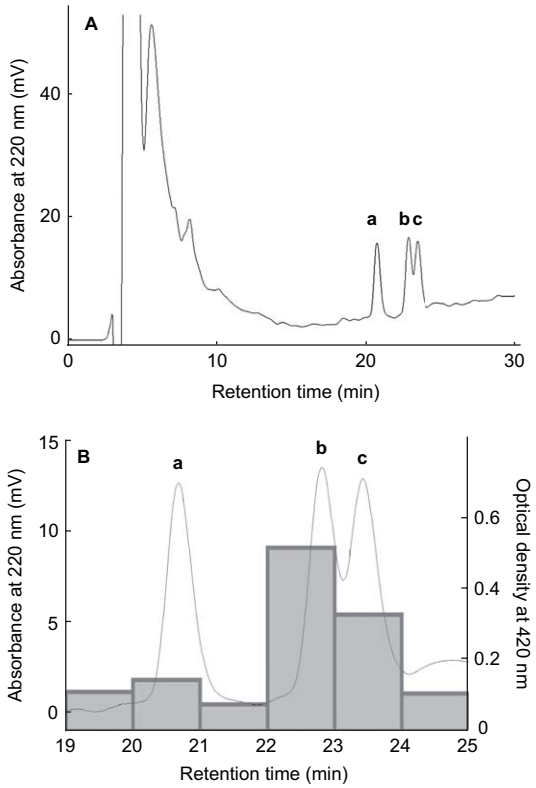


Figure 5. (a) HPLC of a ComX$_{RO-E-2}$ solution partially purified. (b) The effect of each fraction in bioassay and the corresponding HPLC. Gray bars represent optical density at 420 nm.

galactopyranoside using a *B. subtilis* tester strain (BD3020) as described previously.[9,10] In addition to the natural ComX$_{RO-E-2}$ pheromone, only one of the four synthetic peptides, [Lα]ComX$_{RO-E-2}$ **2**, showed obvious biological activity at 1.0 nM, while diastereomers **3–5** showed no biological activity up to 300 nM (Fig. 5).

2.4. Identification and solid-phase synthesis of N-terminal truncated natural pheromones

We analyzed the natural ComX$_{RO-E-2}$ pheromone solution from the culture broth using an *E. coli* expression system,

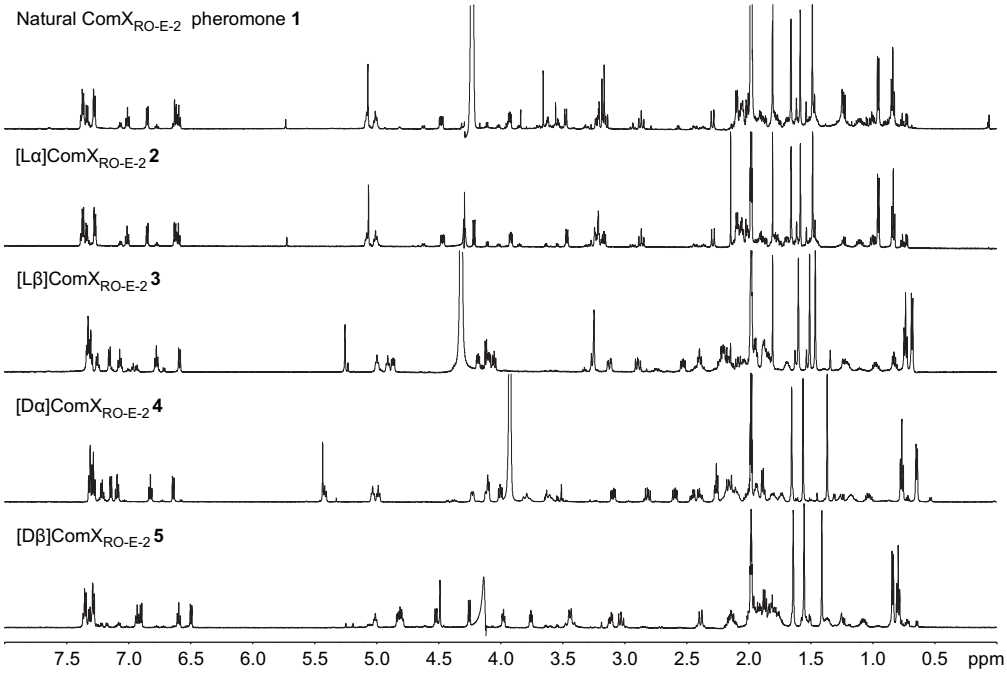


Figure 4. ^{1}H NMR spectra of the natural ComX$_{RO-E-2}$ pheromone and the synthetic peptides. Each DHO signal of [Lα]ComX$_{RO-E-2}$ **2** or [Dβ]ComX$_{RO-E-2}$ **5** was irradiated for suppression.

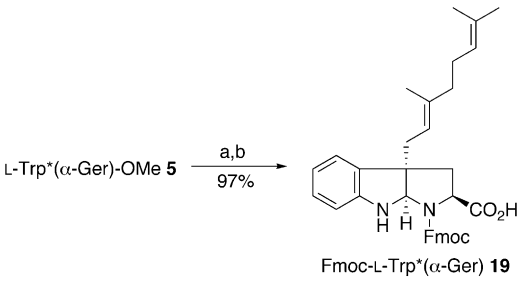
Table 3. HRMS of the $ComX_{RO\text{-}E\text{-}2}$ pheromone and its truncated peptides

Peak	HRMS $[M+H]^+$	Molecular composition (Δ ppm)	Amino acid sequence	Pheromone
a	745.39152	$C_{40}H_{53}N_6O_8$ (−0.56)	FW*EQ	[3-6]$ComX_{RO\text{-}E\text{-}2}$
b	915.49830	$C_{48}H_{67}N_8O_{10}$ (+0.91)	GIFW*EQ	$ComX_{RO\text{-}E\text{-}2}$
c	858.47591	$C_{46}H_{64}N_7O_9$ (−0.11)	IFW*EQ	[2-6]$ComX_{RO\text{-}E\text{-}2}$

W* represents the modified tryptophan residue with a geranyl group.

as reported previously.[9,10,15] Following the final purification step, we identified three biologically active fractions, including the $ComX_{RO\text{-}E\text{-}2}$ pheromone, as shown in Figure 5. Because the other two peptides in the active fractions showed $[M+H]^+$ molecular ions of 858.47591 and 745.39152 on HRMS analysis, and these peptides were assigned molecular formulas of $C_{46}H_{64}N_7O_9$ and $C_{40}H_{53}N_6O_8$, respectively (Table 3). These results suggest that the two peptides were N-terminal truncated peptides of $ComX_{RO\text{-}E\text{-}2}$ pheromone, i.e., [2-6]$ComX_{RO\text{-}E\text{-}2}$ and [3-6]$ComX_{RO\text{-}E\text{-}2}$.

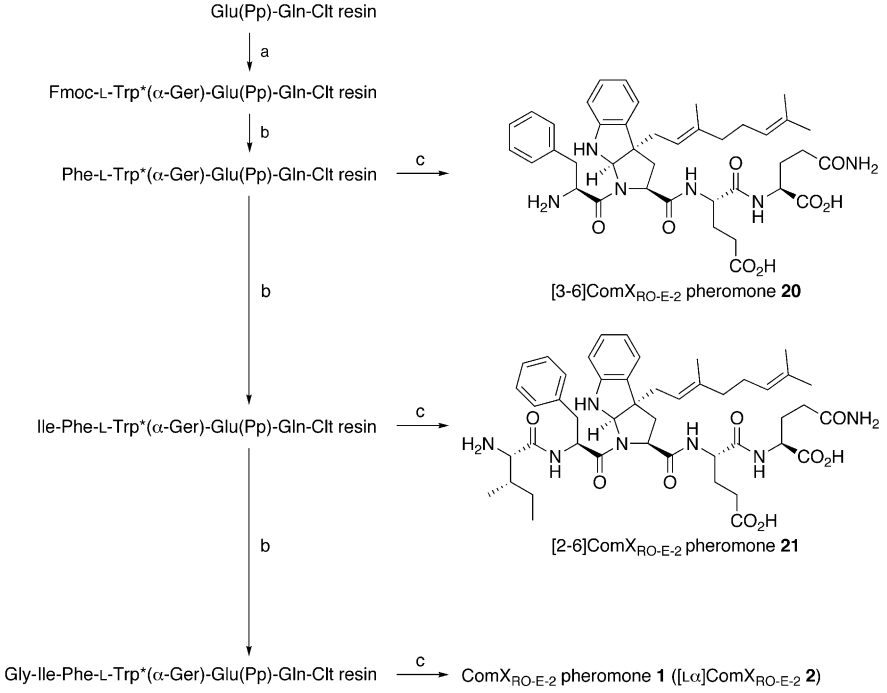
In order to confirm that these active compounds are N-terminal truncated pheromones, we attempted to synthesize



Scheme 4. Synthesis of Fmoc-L-Trp*(α-Ger) **19**. Reaction conditions: (a) LiOH, THF, CH_3OH, H_2O, rt, 1.5 h; (b) Fmoc-OSu, dioxane, 1.0 M aqueous $NaHCO_3$, rt, 1.5 h.

$ComX_{RO\text{-}E\text{-}2}$ pheromone and two N-terminal truncated $ComX_{RO\text{-}E\text{-}2}$ peptides, [3-6]$ComX_{RO\text{-}E\text{-}2}$ peptide **20** and [2-6]$ComX_{RO\text{-}E\text{-}2}$ peptide **21**, using solid-phase synthesis. We first synthesized the desired Fmoc-protected modified tryptophan for Fmoc solid-phase peptide synthesis. L-Trp*(α-Ger)-OMe **6** was hydrolyzed under alkaline conditions, followed by Fmoc protection to give Fmoc-L-Trp*(α-Ger) **19** (Scheme 4). Peptide bond formation was accomplished with a peptide synthesizer, although a previously reported manual synthesis procedure was also used (Scheme 5).[16] Both resin and protecting groups were smoothly cleaved under mild acidic conditions by treatment with 5% TFA at 4 °C for 20 h. The desired peptides were obtained after HPLC purification and analyzed by ^1H NMR spectroscopy and HRMS. The two synthetic N-terminal truncated peptides, [3-6]$ComX_{RO\text{-}E\text{-}2}$ peptides **20** and [2-6]$ComX_{RO\text{-}E\text{-}2}$ peptides **21**, were found to be identical to the natural truncated pheromones.

The biological activity of these peptides was investigated by β-galactosidase assay (Fig. 6). The two peptides showed strong activities, although these were approximately 10-fold weaker than that of the $ComX_{RO\text{-}E\text{-}2}$ (Fig. 7). The dose–response curve of [2-6]$ComX_{RO\text{-}E\text{-}2}$ (peptide **21**) was saturated at 100 nM with an EC_{50} value of 7 nM. [3-6]$ComX_{RO\text{-}E\text{-}2}$ (peptide **20**) showed an EC_{50} value of 8 nM, but its dose–response curve did not reach 100%, even at 1 μM.



Scheme 5. Solid-phase peptide synthesis of $ComX_{RO\text{-}E\text{-}2}$ peptides. Reaction conditions: (a) Fmoc-L-Trp*(α-Ger) **19**, HATU, HOAt, Et_3N, CH_2Cl_2, 0 °C, 2 h; (b) Stepwise solid-phase peptide synthesis with peptide synthesizer using HBTU/HOBt/DIPEA chemistry; (c) 5% TFA, 5% trifluoroethanol, CH_3CN, 4 °C, 20 h.

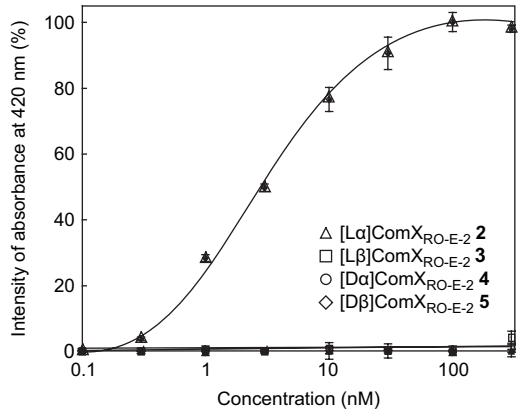


Figure 6. Dose–response curves obtained using the ComX$_{RO-E-2}$ peptide and its diastereomers. Error bars are SD of the means of triplicate samples.

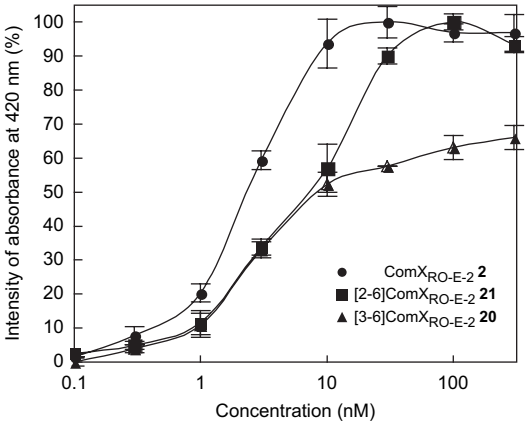


Figure 7. Dose–response curves obtained using the N-terminal truncated ComX$_{RO-E-2}$ peptides. Error bars are SD of the means of triplicate samples.

3. Discussion

The first key step in the synthesis of the modified tryptophan residue was C-geranylation. We screened a variety of conditions for C3-geranylation of the tricyclic N-protected tryptophan methyl esters. We found that Bz-Trp-OMe was the most suitable reactant for C-geranylation. Immediate neutralization with buffer was particularly important to prevent degradation of the C-geranylated products in the subsequent workup. The conditions established are also useful for synthesizing related compounds using other functional groups, such as prenyl, farnesyl, and geranylgeranyl.[21] Based on the hard–soft acid–base theory, we expected C3-geranylation to occur predominantly in THF, a relatively nonpolar solvent. However, THF gave little enhancement of C-geranylation when compared with relatively polar solvents such as DMF or NMP. C/N selectivity in the dimethylallylation of **10a** using alkylating methods involving quaternary ammonium salts was reported to be ineffective by Cardoso et al.[18] On the other hand, geranylation of ether **10i** gave only the N1-geranylated compound in good yield and we were thus able to preferentially synthesize either the N1- or C3-geranylated compound. The synthetic conditions were optimized to yield diastereomeric mixtures of C3-geranyl-modified tryptophans. Because we required both diastereomers for this synthesis, we did not investigate the α/β selectivity in the C3-geranylation in this work, although we observed that selectivity was affected by the choice of solvent. At present, we require only the α-diastereomer to synthesize the natural peptides. We are currently optimizing α-selectivity at the C3 position by investigating alternative solvents and various ester protecting groups for the effective synthesis of ComX pheromones and related peptides.

The second key step was chemoselective reduction to produce modified tryptophan methyl esters **6** and **7** (or **8** and **9**). Compounds **11**, **14**, and **15** contain various reactive functional groups such as olefin, methyl ester, imine, and benzamide (or carbamate). Reduction of **11** using an equivalent amount of DIBAL-H mediated the chemoselective cleavage of benzamide, rather than the reduction of other functional groups. Employment of $NaBH_3CN$, $NaBH_4$, excess DIBAL-H or $BH_3 \cdot SMe_2$ gave various reductive compounds non-chemoselectively. Employment of pinacholborane gave only L-Trp*(β-Ger)-OMe **7** in 51% yield (data not shown). Reduction of **14** (or **15**) with Lindlar's reagent[22] or Et_3SiH under acidic conditions did not facilitate imine reduction. By contrast, catecholborane was found to be the best reagent for the chemoselective reduction of the imine.[23]

Additionally, it was very important to obtain the modified tryptophans in optically pure form. We were able to separate diastereomers **12** and **13** by column chromatography easily, although we were not able to separate the diastereomers in other steps at preparative scale. However, a single isomer is afforded after cleavage of the Bz group, which allowed us to synthesize all four tryptophan residues. The ^1H NMR spectrum of Fmoc-Trp*(Ger)-OMe **16** indicated two different compounds in a 1:1 ratio, although both have the cis configuration as indicated by an NOE between the H-2 proton and the geranyl group. We believe that the two compounds are conformational isomers, since NMR and other analyses showed that cleavage of the Fmoc group gives only optically active esters **6**. We also observed similar phenomena on compounds **17** and **18**, although the ratio of conformers is different on **18** (10:1). Furthermore, the conformational search previously reported for the model study showed two stable conformers.[15] One of the conformers should have a similar three-dimensional structure to the modified tryptophan residue in ComX$_{RO-E-2}$ pheromone. The coupling constants between α and one of the β protons are nearly zero, indicating that the dihedral angle between α and one of the β protons is about 90°.[24] The conformational search clearly suggested this stereostructure, which is important for the expression of biological function. Subsequent investigation into the structure–activity relationships of ComX$_{RO-E-2}$ analogs will clarify the importance of this stereostructure for pheromone activity.

The four peptides were synthesized using a liquid-phase method instead of a solid-phase method for two reasons. First, we wanted to confirm the amino acid coupling reaction with the modified tryptophans at the Nα-position rather than the N1-position. L-Trp*(α-Ger)-OMe **6** was reacted with Teoc-Phe under typical coupling conditions, and HMBC analysis showed that the reaction occurred only at the Nα-position. Second, we were not able to obtain modified Fmoc-protected D-tryptophan with a free carboxyl group,

which is a requirement for solid-phase synthesis, nor was the compound sufficiently stable for this purpose.

Synthetic peptide **2** and natural $ComX_{RO\text{-}E\text{-}2}$ pheromone exhibited identical ^1H NMR spectra and showed very similar biological activity. Thus, the absolute structure of $ComX_{RO\text{-}E\text{-}2}$ pheromone was confirmed to be the same as our proposed structure [Lα]$ComX_{RO\text{-}E\text{-}2}$ **2**, and the structure of the modified tryptophan residue was determined to be [2*S*,3a*R*,8a*S*]-3α-geranyl-1,2,3,3α,8,8α-hexahydropyrrolo-[2,3-*b*]indole-2-carboxylic acid.

We had previously synthesized putative pheromones with 1-, 2-, 4-, 5-, 6- or 7-geranyl substituted tryptophan residues.[16] All of these peptides, including the three newly synthesized diastereomers reported here, had the same molecular formula and showed similar hydrophobicity to the natural $ComX_{RO\text{-}E\text{-}2}$ pheromone on LC analysis, but none exhibited its biological activity. These results suggest that the structure of the tryptophan residue, including its stereostructure, is essential both for the interaction of $ComX_{RO\text{-}E\text{-}2}$ pheromone with receptor ComP and for the signal transduction that results in the expression of genetic competence.

Furthermore, we identified two additional bioactive peptides in the *E. coli* ED413 culture broth. Their structures were predicted to be [2-6]$ComX_{RO\text{-}E\text{-}2}$ and [3-6]$ComX_{RO\text{-}E\text{-}2}$ based on their molecular weights. The Fmoc-modified L-Trp having a free carboxyl group was sufficiently stable to permit the synthesis of these two peptides and $ComX_{RO\text{-}E\text{-}2}$ pheromone itself by solid-phase methods. This method will be useful for future studies on structure–activity relationships and to synthesize probes for investigating ligand–receptor interactions.

The two synthetic N-terminal truncated peptides were identical to the natural peptides. Even the [3-6]$ComX_{RO\text{-}E\text{-}2}$, which has only four amino acid residues, showed low but apparent biological activity. This suggests that while the modified tryptophan residue is essential for biological activity, the two N-terminal amino acid residues are not absolutely required.

Several other ComX pheromones have been reported, and some of these are thought to possess farnesyl groups. The structure of the modified tryptophan residues in these pheromones may be similar to that of $ComX_{RO\text{-}E\text{-}2}$ pheromone. We are now able to synthesize these pheromones using a method similar to that described here. It will be interesting to determine whether these pheromones have the same stereostructures as $ComX_{RO\text{-}E\text{-}2}$ pheromone.

Numerous secondary metabolites of isoprenylated tryptophan derivatives are known;[25–27] however, only four examples of post-translational modification on tryptophan residues have been reported to date.[28–31] The structure of the geranyl-modified tryptophan residue described here is the fifth example of the post-translational modification of tryptophan. Post-translational isoprenylation, farnesyl or geranylgeranyl modification on the C-terminal cysteine residue is known to be quite common.[14] However, isoprenylation on a tryptophan residue is unprecedented, and geranylation is also unknown. The consensus sequence of cysteine isoprenylation is widely known, but we have not yet been able to confirm the consensus sequence for tryptophan isoprenylation among the ComX pheromones. We have no evidence that this novel modification occurs in any other peptide. However, the present tryptophan isoprenylation in a pheromone that governs genetic exchange in *B. subtilis* reminds us that cysteine isoprenylation was first discovered in the sex pheromones of basidiomycetous yeasts.[32–36]

4. Conclusion

We synthesized four diastereomers of unique tricyclic tryptophan residues modified with a geranyl group and their four corresponding peptides. We established the precise structure of $ComX_{RO\text{-}E\text{-}2}$ pheromone, including the modified tryptophan residue, as [2*S*,3a*R*,8a*S*]-3α-geranyl-1,2,3,3α,8,8α-hexahydropyrrolo[2,3-*b*]indole-2-carboxylic acid by comparing NMR spectra and biological activities of the synthetic peptides with those of the natural pheromone.

We also identified other natural $ComX_{RO\text{-}E\text{-}2}$ pheromones. We found that these pheromones were N-terminal truncated peptides of $ComX_{RO\text{-}E\text{-}2}$ pheromone, i.e., [2-6]$ComX_{RO\text{-}E\text{-}2}$ and [3-6]$ComX_{RO\text{-}E\text{-}2}$. These pheromones were synthesized using the solid-phase method and were identical to the natural pheromones analytically and biologically.

5. Experimental

5.1. General methods

High-performance liquid chromatography was performed on a HPLC system equipped with Jasco LC-980 series. HRMS (ESI-TOF, positive) was recorded on a Mariner system (Applied Biosystems) using either an angiotensin/bradykinin/neurotensin mixture or a polypropylene glycol solution as a calibration standard. Optical densities were measured with an AE-15F photoelectric colorimeter (Erma). NMR spectra were recorded on a Bruker ARX-400 or a Bruker AMX-600 spectrometer. Optical rotations were measured on a Jasco DIP-370 digital polarimeter. Organic solvents were purchased as anhydrous grade except for the following solvents, which were freshly distilled prior to use: THF and diethyl ether were dried by distillation from Na and benzophenone ketyl. Moisture-sensitive reactions were carried out under a dry nitrogen atmosphere in well-dried equipment with a tightly fitted rubber septum. Open column chromatography was performed using silica gel BW-300 (Fuji silysia) or ODS Cosmosil 140C18-OPN (Nacalai Tesque). Solid-phase peptide synthesis was performed with a model 433A peptide synthesizer (Applied Biosystems). Fmoc-protected amino acid derivatives except modified tryptophan residues and Clt resin were purchased from commercial sources (Watanabe chemical, Nova biochem).

5.2. Bacterial strains, pheromone production, and biological activity

The *E. coli* $ComX_{RO\text{-}E\text{-}2}$ producer strain (ED413) and the *B. subtilis* tester strain (BD3020) were grown[4,5] and pheromone production was carried out as described previously.[4]

Biological activity was measured by a β-galactosidase assay using the *B. subtilis* tester strain (BD3020) employed the expression of a *srfA-lacZ* fusion, under control of the natural *srfA* promoter, which responds to add $ComX_{RO\text{-}E\text{-}2}$ pheromone. BD3020 was grown overnight to stationary phase and was then diluted 100-fold. The diluted culture (0.5 ml) was added to a sample solution (3 μl of 50% aqueous CH_3CN solution), and incubated at 37 °C for 5 h. After toluene (5 μl) was added to the mixture, biological activity was measured at 420 nm (%) with a standard method using *o*-nitrophenyl-β-D-galactopyranoside at 30 °C.

5.3. Purification of $ComX_{RO\text{-}E\text{-}2}$ derivatives

The pre-purification was as reported previously. The two active fractions were purified by HPLC on an ODS column (4.6×250 mm ID, Develosil ODS-HG-5, Nomura Chemical) at a flow rate of 1.0 ml/min, with CH_3CN and 0.1% aqueous ammonium acetate to give pure $ComX_{RO\text{-}E\text{-}2}$ derivatives. The solution was concentrated to remove CH_3CN, and freeze-dried several times to give pure $ComX_{RO\text{-}E\text{-}2}$ derivatives.

5.4. Synthesis of $ComX_{RO\text{-}E\text{-}2}$ pheromone

5.4.1. Typical cyclization. To a solution of N-protected tryptophan methyl ester (1.00 mmol) and Et_3N (4.0 equiv) in CH_2Cl_2 (10 ml) at 0 °C under nitrogen was slowly added *t*-BuOCl (1.00 mmol). The reaction mixture was stirred at 0 °C for 1 h, and warmed to room temperature for 12 h. Et_2O (200 ml) and H_2O (200 ml) were added to the mixture. The two layers were separated, and the organic layer was washed with saturated aqueous NaCl, dried over Na_2SO_4, and evaporated. The residue was recrystallized from CH_2Cl_2/Et_2O (or appropriate solvents) to give the tricyclic compound as a powder.

5.4.1.1. Tricyclic compound from Bz-L-Trp-OMe 10h. $[\alpha]_D^{23}$ −104 (*c* 1.0, $CHCl_3$); ^1H NMR (600 MHz, DMSO-d_6) δ 11.50 (br, 1H), 7.60 (d, 2H, *J*=7.1 Hz), 7.53–7.36 (m, 4H), 7.26 (dd, 1H, *J*=4.9, 8.6 Hz), 6.98–6.95 (m, 2H), 5.58 (d, 1H, *J*=8.2 Hz), 3.60 (dd, 1H, *J*=8.2, 9.2 Hz), 3.44 (s, 3H), 3.06 (d, 1H, *J*=9.2 Hz); ^{13}C NMR (150 MHz, DMSO-d_6) δ 171.5, 166.8, 143.1, 138.0, 135.2, 131.0, 128.8, 127.5, 122.9, 120.3, 119.5, 117.2, 112.9, 99.0, 67.1, 52.7, 29.4; Anal. Calcd for $C_{19}H_{16}N_2O_3$: C, 71.24; H, 5.03; N, 8.74. Found: C, 71.24; H, 4.92; N, 8.87.

5.4.1.2. D-Enantiomer of 10h. $[\alpha]_D^{28}$ +106 (*c* 1.0, $CHCl_3$).

5.4.2. Typical geranylation. To a solution of **2** (1.00 mmol) in THF (10 ml) at 0 °C under nitrogen was added NaH (1.3 equiv). After stirring at 0 °C for 1 h, geranyl bromide (1.1 eq) was added to the mixture. After stirring for 5 h at 0 °C, the reaction mixture was quenched and neutralized with 0.1 M phosphate buffer (pH 7). The reaction mixture was extracted with Et_2O (4×20 ml), washed with saturated aqueous NaCl, dried over Na_2SO_4, and evaporated. The residue was purified by silica gel column chromatography using appropriate solvents to give the C-geranylated diastereoisomeric mixtures and N-geranylated compound.

5.4.2.1. β-Geranyl compound of 11. ^1H NMR (600 MHz, $CDCl_3$) δ 1.57 (s, 3H), 1.60 (s, 3H), 1.68 (s, 3H), 1.98–2.05 (m, 4H), 2.37 (dd, 1H, *J*=10.9, 12.9 Hz), 2.56 (dd, 1H, *J*=7.8, 14.2 Hz), 2.57 (d, 1H, *J*=12.9 Hz), 2.77 (dd, 1H, *J*=6.7, 14.2 Hz), 3.86 (s, 3H), 4.86 (m, 1H), 5.05 (m, 1H), 5.60 (d, 1H, *J*=10.9 Hz), 7.03 (dt, 1H, *J*=1.8, 7.2 Hz), 7.18–7.26 (m, 3H), 7.43 (dd, 1H, *J*=7.5, 7.8 Hz), 7.56 (dt, 1H, *J*=1.1, 7.5 Hz), 7.61 (dd, 1H, *J*=1.1, 7.8 Hz); ^{13}C NMR (150 MHz, $CDCl_3$) δ 16.4, 17.6, 25.6, 26.4, 30.8, 34.8, 39.8, 52.8, 60.5, 65.0, 117.2, 119.9, 122.6, 123.7, 123.9, 128.1, 128.5, 129.5, 131.6, 132.5, 133.2, 139.5, 139.7, 158.2, 168.5, 171.4, 181.5; Anal. Calcd for $C_{29}H_{32}N_2O_3$: C, 76.29; H, 7.06; N, 6.14. Found: C, 76.31; H, 7.07; N, 6.24.

5.4.2.2. Secondary amines 12 and 13. To a solution of **11** (2.23 g, 4.89 mmol) in THF (49 ml) was slowly added DIBAL-H (1.01 M solution in hexane, 9.70 ml, 9.80 mmol) at –78 °C. After stirring for 2 h at –78 °C, the reaction mixture was poured into 0.1 M phosphate buffer (pH 7.0, 50 ml). It was extracted with EtOAc (4×50 ml). The organic layer was washed with saturated aqueous NaCl, dried over Na_2SO_4, and evaporated. The residue was purified by silica gel column chromatography ($CHCl_3/CH_3OH$ 100/1 → 80/1) to give α-isomer **12** (871 mg, 2.47 mmol, 51%) as a colorless oil and β-isomer **13** (751 mg, 2.13 mmol, 44%) as a colorless oil.

5.4.2.3. α-Geranyl secondary amine 12. $[\alpha]_D^{26}$= +288 (*c* 1.0, $CHCl_3$); ^1H NMR (400 MHz, $CDCl_3$) δ 1.45 (s, 3H), 1.62 (s, 3H), 1.70 (s, 3H), 2.00–2.09 (m, 4H), 2.09 (dd, 1H, *J*=9.8, 11.9 Hz), 2.18 (dd, 1H, *J*=7.4, 13.9 Hz), 2.44 (dd, 1H, *J*=8.0, 13.9 Hz), 2.67 (dd, 1H, *J*=5.1, 11.9 Hz), 3.77 (s, 3H), 4.95 (dd, 1H, *J*=5.1, 9.8 Hz), 5.09 (m, 1H), 5.22 (m, 1H), 6.83 (ddd, 1H, *J*=1.0, 7.4, 7.5 Hz), 7.01 (dd, 1H, *J*=1.0, 7.7 Hz), 7.09 (dd, 1H, *J*=0.8, 7.4 Hz), 7.15 (ddd, 1H, *J*=0.8, 7.5, 7.7 Hz); ^{13}C NMR (100 MHz, $CDCl_3$) δ 16.0, 17.7, 25.7, 26.5, 34.4, 39.4, 39.9, 52.1, 60.2, 71.0, 111.8, 118.6, 120.4, 123.9, 124.2, 128.1, 131.5, 133.2, 139.3, 149.5, 173.4, 184.6; Anal. Calcd for $C_{22}H_{28}N_2O_2$: C, 74.97; H, 8.01; N, 7.95. Found: C, 74.96; H, 7.87; N, 7.50.

5.4.2.4. D-Enantiomer of 12. $[\alpha]_D^{28}$=−279 (*c* 1.0, $CHCl_3$).

5.4.2.5. β-Geranyl secondary amine 13. ^1H NMR (400 MHz, $CDCl_3$) δ 1.35 (s, 3H), 1.60 (s, 3H), 1.69 (s, 3H), 1.93–2.04 (m, 4H), 2.16 (dd, 1H, *J*=7.1, 13.9 Hz), 2.38–2.49 (m, 2H), 2.64 (d, 1H, *J*=12.8 Hz), 3.81 (s, 3H), 4.93 (d, 1H, *J*=9.6 Hz), 5.04–5.11 (m, 2H), 6.86 (dt, 1H, *J*=0.7, 7.6 Hz), 6.93 (dd, 1H, *J*=0.7, 7.6 Hz), 7.09 (dd, 1H, *J*=0.9, 7.6 Hz), 7.16 (dt, 1H, *J*=0.9, 7.6 Hz); ^{13}C NMR (100 MHz, $CDCl_3$) δ 16.2, 17.6, 25.6, 26.6, 37.4, 38.1, 39.9, 51.9, 53.2, 55.9, 117.2, 117.5, 121.1, 121.7, 124.2, 131.1, 137.5, 138.3, 139.1, 155.8, 172.5, 174.7; HRMS (ESI$^+$) *m/z*: calcd for $C_{22}H_{29}N_2O_2$ ($[M+H]^+$) 353.2224. Found 353.2228.

5.4.2.6. Fmoc-protected compound 14. To a solution of **12** (100 mg, 0.284 mmol) in dioxane (3.0 ml) and 1.0 M aqueous $NaHCO_3$ (3.0 ml) was added Fmoc-OSu (120 mg, 0.356 mmol) at room temperature. After stirring for 2 h, water (10 ml) was added to the reaction mixture and the mixture was extracted with EtOAc (4×10 ml). The organic layer was washed with saturated aqueous NaCl, dried over

Na_2SO_4, and evaporated. The residue was purified by silica gel column chromatography (hexane/EtOAc 150/2→125/2) to give **14** (143 mg, 0.249 mmol, 88%) as a colorless oil.

$[\alpha]_D^{29}$ +317 (*c* 1.0, $CHCl_3$); ^{1}H NMR (600 MHz, CD_3COCD_3) δ 1.42 (s, 3H), 1.56 (s, 3H), 1.64 (s, 3H), 1.90–2.04 (m, 4H), 2.27 (dd, 1H, *J*=9.9, 12.0 Hz), 2.46 (d, 2H, *J*=7.9 Hz), 2.73 (dd, 1H, *J*=5.2, 12.0 Hz), 3.78 (s, 3H), 4.38 (dd, 1H, *J*=6.9, 8.1 Hz), 4.55 (dd, 1H, *J*=8.1, 10.4 Hz), 4.59 (dd, 1H, *J*=6.9, 10.4 Hz), 5.06 (m, 1H), 5.09 (m, 1H), 5.21 (dd, 1H, *J*=5.2, 9.9 Hz), 7.13 (dt, 1H, *J*=0.9, 7.5 Hz), 7.29 (ddd, 1H, *J*=1.2, 7.5, 7.9 Hz), 7.31–7.36 (m, 3H), 7.39–7.44 (m, 2H), 7.82–7.96 (m, 3H), 7.96 (dd, 1H, *J*=0.6, 7.5 Hz), 8.26 (dd, 1H, *J*=0.6, 7.5 Hz); ^{13}C NMR (100 MHz, $CDCl_3$) δ 16.1, 17.8, 25.8, 27.3, 35.8, 39.9, 40.6, 47.4, 52.3, 60.8, 69.4, 74.1, 115.9, 118.3, 118.6, 120.6, 120.7, 124.8, 124.9, 125.0, 126.6, 127.4, 128.0, 128.2, 128.6, 128.7, 129.1, 131.9, 133.1, 140.6, 141.9, 142.2, 143.3, 143.4, 143.8, 151.6, 173.8, 176.4; HRMS (ESI+) *m/z*: calcd for $C_{37}H_{39}N_2O_4$ ([M+H]$^+$) 575.2904. Found 575.2922.

5.4.2.7. D-Enantiomer of 14. $[\alpha]_D^{28}$ –314 (*c* 0.50, $CHCl_3$).

5.4.2.8. L-Trp*(α-Ger)-OMe 6. To a solution of **14** (458 mg, 0.797 mmol) in THF (31 ml) was added catecholborane (1.01 M solution in THF, 1.0 ml, 1.00 mmol) at 0 °C. After stirring for 2 h, water (30 ml) was added to the reaction mixture, followed by extraction with EtOAc (4×30 ml). The organic layer was washed with saturated aqueous NaCl, dried over Na_2SO_4, and evaporated. The residue was purified by silica gel column chromatography (hexane/acetone 10/1→8/1) to give a reductive compound. To a solution of the reductive compound in CH_3CN (2.7 ml) at room temperature was added piperidine (0.3 ml). After stirring for 1 h, the reaction mixture was quenched and neutralized with 0.1 M phosphate buffer (pH 7.0, 3.0 ml). The reaction mixture was extracted with EtOAc (4×5 ml). The organic layer was washed with saturated aqueous NaCl, dried over Na_2SO_4, and evaporated. The residue was purified by silica gel column chromatography (hexane/acetone 6/1→5/1) to give **6** (202 mg, 0.570 mmol, 72%, in two steps) as a colorless oil.

$[\alpha]_D^{26}$=+52.2 (*c* 1.0, $CHCl_3$); ^{1}H NMR (600 MHz, C_6D_6) δ 6.99 (t, 1H, *J*=7.6 Hz), 6.96 (d, 1H, *J*=7.6 Hz), 6.69 (t, 1H, *J*=7.6 Hz), 6.40 (d, 1H, *J*=7.6 Hz), 5.24 (m, 1H), 5.14 (m, 1H), 4.63 (s, 1H), 3.71 (dd, 1H, *J*=3.3, 7.8 Hz), 3.07 (s, 3H), 2.46 (dd, 1H, *J*=3.3, 12.6 Hz), 2.35 (dd, 1H, *J*=7.9, 14.5 Hz), 2.31 (dd, 1H, *J*=7.6, 14.5 Hz), 2.18 (dd, 1H, *J*=7.8, 12.6 Hz), 2.10–2.06 (m, 2H), 2.01–1.97 (m, 2H), 1.67 (s, 3H), 1.53 (s, 3H), 1.46 (s, 3H); ^{13}C NMR (100 MHz, CD_3COCD_3) δ 174.8, 151.2, 137.8, 131.7, 128.4, 125.1, 124.4, 121.4, 120.3, 118.3, 109.6, 83.5, 60.2, 58.1, 51.7, 41.9, 40.6, 37.2, 27.2, 25.8, 17.7, 16.4; Anal. Calcd for $C_{22}H_{30}N_2O_2$: C, 74.54; H, 8.53; N, 7.90. Found: C, 74.52; H, 8.49; N, 7.83.

5.4.2.9. D-Trp*(β-Ger)-OMe 9 (D-enantiomer of 6). $[\alpha]_D^{28}$ –52.1 (*c* 1.0, $CHCl_3$).

5.4.2.10. L-Trp*(β-Ger)-OMe 7. $[\alpha]_D^{29}$ –76.1 (*c* 1.0, $CHCl_3$); ^{1}H NMR (600 MHz, C_6D_6) δ 7.01 (dd, 1H, *J*=7.5, 7.7 Hz), 6.94 (d, 1H, *J*=7.4 Hz), 6.73 (dd, 1H, *J*=7.4, 7.5 Hz), 6.38 (d, 1H, *J*=7.7 Hz), 5.19 (m, 1H), 5.13 (m, 1H), 4.59 (s, 1H), 3.65 (dd, 1H, *J*=5.8, 10.7 Hz), 3.24 (s, 3H), 2.35 (dd, 1H, *J*=7.8, 14.5 Hz), 2.31 (dd, 1H, *J*=7.1, 14.5 Hz), 2.27 (dd, 1H, *J*=5.8, 12.0 Hz), 2.09–2.05 (m, 2H), 2.01–1.96 (m, 2H), 1.89 (dd, 1H, *J*=10.7, 12.0 Hz), 1.68 (s, 3H), 1.53 (s, 3H), 1.44 (s, 3H); ^{13}C NMR (100 MHz, CD_3COCD_3) δ 173.6, 150.9, 136.8, 132.9, 130.6, 127.5, 124.1, 123.1, 120.3, 117.2, 107.8, 82.2, 59.1, 58.1, 50.9, 44.0, 39.5, 36.4, 26.2, 24.8, 16.7, 15.4; Anal. Calcd for $C_{22}H_{30}N_2O_2$: C, 74.54; H, 8.53; N, 7.90. Found: C, 74.52; H, 8.57; N, 7.99.

5.4.2.11. D-Trp*(α-Ger)-OMe 8 (D-enantiomer of 7). $[\alpha]_D^{28}$ +75.4 (*c* 1.0, $CHCl_3$).

5.4.3. Synthesis of the $ComX_{RO\text{-}E\text{-}2}$ peptides. We chose Teoc as the N-terminal protecting group, and Clt as the carboxy-protecting group. The Teoc group was cleaved under neutral condition using TAS-F, and the Clt group was cleaved under mild acidic conditions using 50% aqueous AcOH. All reactions were monitored with LC/MS and TLC. Peptides were purified by silica gel column chromatography containing 1% Et_3N so as not to cleave the Clt group. All Clt esters were unstable, so they were carried on to the next step immediately without further purification.

5.4.3.1. Teoc-Phe-L-Trp*(α-Ger)-OMe 18. To a solution of L-Trp*(α-Ger)-OMe **6** (80.5 mg, 0.227 mmol), Teoc-Phe (150 mg, 0.485 mmol), and Et_3N (94.6 mg, 0.935 mmol) in CH_2Cl_2 (5.0 ml) under nitrogen were added HOAt (66.0 mg, 0.485 mmol) and HATU (185 mg, 0.487 mmol) at 0 °C. After stirring for 2 h, the reaction mixture was quenched and neutralized with 0.1 M phosphate buffer (10 ml). The reaction mixture was extracted with EtOAc (4×10 ml). The organic layer was washed with saturated aqueous NaCl, dried over Na_2SO_4, and evaporated. The residue was purified by silica gel column chromatography (hexane/EtOAc 6/1→4/1) to give **18** (141 mg, 0.217 mmol, 96%) as a colorless oil.

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.34–7.27 (m, 3H), 7.23 (d, 2H, *J*=7.1 Hz), 6.96 (dd, 1H, *J*=7.3, 7.7 Hz), 6.86 (d, 1H, *J*=7.4 Hz), 6.59 (dd, 1H, *J*=7.3, 7.4 Hz), 6.11 (d, 1H, *J*=7.7 Hz), 5.59 (d, 1H, *J*=8.2 Hz), 5.10–5.04 (m, 2H), 5.04 (s, 1H), 4.39 (m, 1H), 4.16–4.05 (m, 2H), 3.57 (d, 1H, *J*=8.3 Hz), 3.14 (dd, 1H, *J*=4.7, 12.7 Hz), 3.10 (s, 3H), 2.85 (dd, 1H, *J*=10.6, 12.7 Hz), 2.34 (d, 1H, *J*=12.8 Hz), 2.24 (dd, 1H, *J*=7.5, 14.4 Hz), 2.19 (dd, 1H, *J*=7.6, 14.4 Hz), 2.08–2.00 (m, 4H), 1.70 (s, 3H), 1.68 (s, 3H), 1.52–1.47 (m, 4H), 0.96 (m, 2H), 0.02 (s, 9H); ^{13}C NMR (150 MHz, $CDCl_3$) δ 171.7, 170.4, 155.4, 149.5, 138.8, 136.2, 131.6, 130.0, 129.4, 128.7, 127.2, 124.1, 123.9, 118.5, 118.4, 109.0, 80.9, 63.1, 59.7, 55.1, 53.7, 52.3, 41.3, 40.0, 37.9, 35.3, 26.6, 25.7, 17.8, 17.7, 16.2, –1.53; HRMS (ESI$^+$) *m/z*: calcd for $C_{37}H_{52}N_3O_5Si$ ([M+H]$^+$) 646.3671. Found 646.3684.

5.4.3.2. [Lα]$ComX_{RO\text{-}E\text{-}2}$ 2 ($ComX_{RO\text{-}E\text{-}2}$ pheromone). ^{1}H NMR (600 MHz, CD_3CN/D_2O 6/4) δ 7.33–7.39 (m, 3H), 7.27 (d, 2H, *J*=7.2 Hz), 7.01 (dd, 1H, *J*=7.4, 7.8 Hz), 6.85 (d, 1H, *J*=7.3 Hz), 6.63 (d, 1H, *J*=7.8 Hz), 6.60 (dd, 1H, *J*=7.3, 7.4 Hz), 5.08 (m, 1H), 5.06 (s, 1H), 5.00 (m, 1H), 4.47 (dd, 1H, *J*=4.9, 11.8 Hz), 4.21 (d, 1H, *J*=8.3 Hz),

3.92 (dd, 1H, *J*=4.8, 8.3 Hz), 3.46 (d, 1H, *J*=8.7 Hz), 3.26 (d, 1H, *J*=16.7 Hz), 3.23–3.20 (m, 2H), 3.17 (dd, 1H, *J*=6.4, 8.5 Hz), 2.87 (dd, 1H, *J*=11.8, 12.2 Hz), 2.28 (d, 1H, *J*=12.7 Hz), 2.14–1.98 (m, 10H), 1.94–1.84 (m, 2H), 1.80–1.75 (m, 2H), 1.68 (m, 1H), 1.66 (s, 3H), 1.58 (s, 3H), 1.49 (m, 1H), 1.48 (s, 3H), 1.10 (m, 1H) 1.00 (dd, 1H, *J*=8.7, 12.8 Hz), 0.96 (d, 3H, *J*=6.8 Hz), 0.83 (dd, 3H, *J*=7.3, 7.4 Hz); HRMS (ESI$^+$) *m*/*z*: calcd for $C_{48}H_{68}N_8O_{10}$ ([M+2H]$^{2+}$) 458.2524. Found 458.2525.

5.4.3.3. [Lβ]ComX$_{RO\text{-}E\text{-}2}$ 3. ^{1}H NMR (600 MHz, CD_3CN/D_2O, 6/4) δ 7.34–7.28 (m, 4H), 7.25 (dd, 1H, *J*=7.2, 7.4 Hz), 7.15 (d, 1H, *J*=7.4 Hz), 7.07 (dd, 1H, *J*=7.3, 7.8 Hz), 6.78 (dd, 1H, *J*=7.3, 7.4 Hz), 6.59 (d, 1H, *J*=7.8 Hz), 5.25 (s, 1H), 5.00 (m, 1H), 4.91 (m, 1H), 4.87 (dd, 1H, *J*=4.8, 9.5 Hz), 4.18 (dd, 1H, *J*=5.3, 8.9 Hz), 4.12 (d, 1H, *J*=7.4 Hz), 4.09 (dd, 1H, *J*=4.9, 8.1 Hz), 4.05 (dd, 1H, *J*=7.5, 8.9 Hz), 3.25 (s, 2H), 3.12 (dd, 1H, *J*=4.8, 14.1 Hz), 2.89 (dd, 1H, *J*=9.5, 14.1 Hz), 2.53 (dd, 1H, *J*=7.5, 12.8 Hz), 2.40 (m, 2H), 2.25–2.11 (m, 5H,), 2.08 (dd, 1H, *J*=8.9, 12.8 Hz), 2.05–1.93 (m, 3H), 1.90–1.81 (m, 4H), 1.69 (m, 1H), 1.60 (s, 3H), 1.51 (s, 3H), 1.46 (s, 3H), 1.20 (m, 1H), 0.98 (m, 1H), 0.74 (dd, 3H, *J*=7.3, 7.4 Hz), 0.68 (d, 3H, *J*=6.8 Hz); HRMS (ESI$^+$) *m*/*z*: calcd for $C_{48}H_{68}N_8O_{10}$ ([M+2H]$^{2+}$) 458.2524. Found 458.2530.

5.4.3.4. [Dα]ComX$_{RO\text{-}E\text{-}2}$ 4. ^{1}H NMR (600 MHz, CD_3CN/D_2O, 6/4) δ 7.26–7.23 (m, 4H) 7.21 (dd, 1H, *J*=7.2, 7.4 Hz), 7.14 (d, 1H, *J*=6.8 Hz), 7.09 (dd, 1H, *J*=7.4, 7.9 Hz), 6.82 (dd, 1H, *J*=6.8, 7.4 Hz), 6.64 (d, 1H, *J*=7.9 Hz), 5.43 (s, 1H), 5.42 (dd, 1H, *J*=6.4, 8.8 Hz), 5.03 (m, 1H), 4.99 (m, 1H), 4.23 (dd, 1H, *J*=3.7, 9.8 Hz), 4.13–4.09 (m, 2H) 4.00 (dd, 1H, *J*=7.3, 9.4 Hz), 3.80 (d, 1H, *J*=15.8 Hz), 3.63 (d, 1H, *J*=15.8 Hz), 3.09 (dd, 2H, *J*=6.4, 13.9 Hz), 2.81 (dd, 1H, *J*=8.8, 13.9 Hz), 2.60 (dd, 1H, *J*=7.3, 13.0 Hz), 2.45 (dd, 1H, *J*=7.6, 14.0 Hz), 2.39 (dd, 1H, *J*=7.9, 14.0 Hz), 2.26 (dd, 2H, *J*=7.2, 8.5 Hz), 2.19–2.06 (m, 5H), 2.04–1.89 (m, 5H), 1.81 (m, 1H), 1.74 (m, 1H), 1.65 (s, 3H), 1.56 (s, 3H), 1.37 (s, 3H), 1.18 (m, 1H), 1.03 (m, 1H), 0.77 (dd, 3H, *J*=7.3, 7.4 Hz), 0.65 (d, 3H, *J*=6.8 Hz); HRMS (ESI$^+$) *m*/*z*: calcd for $C_{48}H_{68}N_8O_{10}$ ([M+2H]$^{2+}$) 458.2524. Found 458.2533.

5.4.3.5. [Dβ]ComX$_{RO\text{-}E\text{-}2}$ 5. ^{1}H NMR (600 MHz, CD_3CN/D_2O, 6/4) δ 7.37–7.27 (m, 5H), 6.94 (dd, 1H, *J*=7.5, 7.7 Hz), 6.90 (d, 1H, *J*=7.4 Hz), 6.60 (dd, 1H, *J*=7.4, 7.5 Hz), 6.49 (d, 1H, *J*=7.7 Hz), 5.01 (m, 1H), 4.84–4.79 (m, 2H), 4.53 (d, 1H, *J*=9.0 Hz), 4.49 (s, 1H), 4.28 (d, 1H, *J*=7.9 Hz), 3.98 (dd, 1H, *J*=6.4, 6.5 Hz), 3.76 (dd, 1H, *J*=5.7, 7.5 Hz), 3.46 (d, 1H, *J*=17.1 Hz), 3.42 (d, 1H, *J*=17.1 Hz), 3.11 (dd, 1H, *J*=5.6, 12.6 Hz), 3.05 (dd, 1H, *J*=12.6, 12.6 Hz), 2.39 (d, 1H, *J*=12.7 Hz), 2.18–2.09 (m, 2H), 1.98–1.71 (m, 13H), 1.64 (s, 3H), 1.55 (s, 3H), 1.54 (m, 1H), 1.41 (s, 3H), 1.37 (m, 1H), 1.08 (m, 1H), 0.84 (d, 3H, *J*=6.8 Hz), 0.79 (d, 3H, *J*=7.3, 7.4 Hz); HRMS (ESI$^+$) *m*/*z*: calcd for $C_{48}H_{68}N_8O_{10}$ ([M+2H]$^{2+}$) 458.2524. Found 458.2521.

5.4.4. Synthesis of Fmoc-protected tryptophan residue for solid-phase peptide synthesis.

5.4.4.1. Fmoc-L-Trp*(α-Ger) 19. To a solution of L-Trp*(α-Ger)-OMe **6** (312 mg, 0.880 mmol) in THF (3 ml) and CH_3OH (3.0 ml), 1 M aqueous LiOH (3.0 ml) was added dropwise at room temperature. After the reaction mixture had been stirred for 30 min, it was neutralized with 0.1 M phosphate buffer (10 ml). The reaction mixture was extracted with EtOAc (5×10 ml). The organic solvent was removed by evaporation. The residue was dissolved in dioxane (8.8 ml) and 1 M aqueous Na_2CO_3 (4.4 ml). To the solution was added Fmoc-OSu (386 mg, 1.16 mmol) at room temperature. After the reaction mixture had been stirred for 2 h, the reaction was neutralized with 0.1 M phosphate buffer (10 ml). The reaction mixture was extracted with EtOAc (4×30 ml). The organic layer was washed with saturated aqueous NaCl, dried over Na_2SO_4, and evaporated. The residue was purified by silica gel column chromatography ($CHCl_3/CH_3OH$ 100/1→50/1) to give Fmoc-L-Trp*(α-Ger) **19** (482 mg, 0.857 mmol, 97% in two steps) as a colorless oil.

Anal. Calcd for $C_{36}H_{38}N_2O_4$: C, 76.84; H, 6.81; N, 4.98. Found: C, 76.95; H, 7.12; N, 5.08.

5.4.5. Syntheses of ComX$_{RO\text{-}E\text{-}2}$ pheromones with solid-phase peptide synthesis. Peptide bond formation was carried out according to the procedure for putative ComX$_{RO\text{-}E\text{-}2}$ peptides[16] except for the final cleavage and the deprotection procedures. Cleavage and deprotection of the resin was carried out as follows. To a suspension of the attached resin in CH_3CN was added 5% TFA and 5% trifluoroethanol at 0 °C. The reaction mixture was shaken in a rotary shaker for 20 h at 4 °C.

5.4.5.1. [3-6]ComX$_{RO\text{-}E\text{-}2}$ 20. ^{1}H NMR (600 MHz, CD_3CN/D_2O, 6/4) δ 7.42–7.34 (m, 3H), 7.28 (d, 2H, *J*=6.9 Hz), 7.02 (dd, 1H, *J*=7.5, 7.7 Hz), 6.88 (d, 1H, *J*=7.2 Hz), 6.63–6.60 (m, 2H), 5.09 (m, 1H), 5.07 (s, 1H), 5.02 (m, 1H), 4.28 (dd, 1H, *J*=4.8, 11.8 Hz), 3.96 (dd, 1H, *J*=4.8, 7.8 Hz), 3.38 (d, 1H, *J*=8.7 Hz), 3.27 (dd, 1H, *J*=4.8, 12.2 Hz), 3.21 (dd, 1H, *J*=5.2, 6.1 Hz), 2.99 (dd, 1H, *J*=11.8, 12.2 Hz), 2.28 (d, 1H, *J*=12.9 Hz), 2.19–2.12 (m, 2H), 2.11 (d, 1H, *J*=6.8 Hz), 2.17–1.90 (m, 12H), 1.76 (m, 1H), 1.65 (s, 3H), 1.58 (s, 3H), 1.52 (m, 1H), 1.48 (s, 3H), 1.17 (dd, 1H, *J*=8.7, 12.9 Hz); HRMS (ESI$^+$) *m*/*z*: calcd for $C_{40}H_{53}N_6O_8$ ([M+H]$^+$) 745.3919. Found 745.3939.

5.4.5.2. [2-6]ComX$_{RO\text{-}E\text{-}2}$ 21. ^{1}H NMR (600 MHz, CD_3CN/D_2O, 6/4) δ 7.39–7.34 (m, 3H), 7.28 (d, 2H, *J*=7.2 Hz), 7.01 (dd, 1H, *J*=7.6, 7.7 Hz), 6.87 (d, 1H, *J*=7.5 Hz), 6.63–6.59 (m, 2H), 5.08 (s, 1H), 5.07 (m, 1H), 5.00 (m, 1H), 4.51 (dd, 1H, *J*=5.1, 11.2 Hz), 3.94 (dd, 1H, *J*=4.8, 7.9 Hz), 3.55 (d, 1H, *J*=8.8 Hz), 3.23–3.17 (m, 3H), 2.87 (dd, 1H, *J*=11.2, 12.2 Hz), 2.29 (d, 1H, *J*=12.8 Hz), 2.09 (d, 2H, *J*=7.2 Hz), 2.08–1.92 (m, 9H), 1.89–1.68 (m, 3H), 1.66 (s, 3H), 1.61 (m, 1H), 1.58 (s, 3H), 1.49 (s, 3H), 1.41 (m, 1H), 1.12–1.05 (m, 2H), 0.91 (d, 3H, *J*=6.8 Hz), 0.82 (dd, 3H, *J*=7.4, 7.5 Hz); HRMS (ESI$^+$) *m*/*z*: calcd for $C_{46}H_{64}N_7O_9$ ([M+H]$^+$) 858.4760. Found 858.4763.

Acknowledgements

The work performed in Nagoya was supported by Grant-in-Aid for COE (14COEA02) and for Scientific Research (No. 18101009). The work in Newark was supported by NIH grant GM57720.

References and notes

1. Williams, P.; Baldwin, T. J.; Downie, J. A. *Microbial Signaling and Communication*; England, R. R., Hobbs, G., Bainton, N. J., Roberts, D. McL., Eds.; Cambridge University Press: Cambridge, 1999; pp 1–32.
2. Lyon, G. J.; Muir, T. W. *Chem. Biol.* **2003**, *10*, 1007.
3. Taga, M. E.; Bassler, B. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4549.
4. Hamoen, L. W.; Venema, G.; Kuipers, O. P. *Microbiology* **2003**, *149*, 9.
5. Magnuson, R.; Solomon, J.; Grossman, A. D. *Cell* **1994**, *77*, 207.
6. Weinrauch, Y.; Penchev, R.; Dubnau, E.; Smith, I.; Dubnau, D. *Genes Dev.* **1990**, *4*, 860.
7. Piazza, F.; Tortosa, P.; Dubnau, D. *J. Bacteriol.* **2001**, *181*, 4540.
8. Tran, L.-S. P.; Nagai, T.; Itoh, Y. *Mol. Microbiol.* **2000**, *37*, 1159.
9. Tortosa, P.; Logsdon, L.; Kraigher, B.; Itoh, Y.; Mandic-Mulec, I.; Dubnau, D. *J. Bacteriol.* **2001**, *183*, 451.
10. Ansaldi, M.; Dubnau, D. *Mol. Microbiol.* **2002**, *44*, 1561.
11. Ansaldi, M.; Marolt, D.; Stebe, T.; Mandic-Mulec, I.; Dubnau, D. *J. Bacteriol.* **2004**, *186*, 15.
12. Weinrauch, Y.; Msadek, T.; Kunst, F.; Dubnau, D. *J. Bacteriol.* **1991**, *173*, 5685.
13. Bacon Schneider, K.; Palmer, T. M.; Grossman, A. D. *J. Bacteriol.* **2002**, *184*, 410.
14. Clarke, S. *Annu. Rev. Biochem.* **1992**, *61*, 355.
15. Okada, M.; Sato, I.; Cho, S. J.; Iwata, H.; Nishio, T.; Dubnau, D.; Sakagami, Y. *Nat. Chem. Biol.* **2005**, *1*, 23.
16. Okada, M.; Sato, I.; Cho, S. J.; Suzuki, Y.; Ojika, M.; Dubnau, D.; Sakagami, Y. *Biosci. Biotechnol. Biochem.* **2004**, *68*, 2374.
17. Ohno, M.; Spande, T. F.; Witkop, B. *J. Am. Chem. Soc.* **1980**, *92*, 343.
18. Cardoso, A. S.; Srinivasan, N.; Lobo, A. M.; Prabhakar, S. *Tetrahedron Lett.* **2001**, *42*, 6663.
19. Other protected compounds (Boc, Teoc, Trt, TBS, and Fmoc) could not be recrystallized or decomposed by the cyclization.
20. Teoc group was mediated to only β-geranylated compound **13** with the treatment with Teoc-*O*-nitrophenyl.
21. We have actually confirmed to react the material with farnesyl bromide in good yield (data not shown).
22. Lindlar, H. *Helv. Chim. Acta* **1952**, *35*, 446.
23. Chemoselective imine reduction with diacid–borane complex has been reported (see: Liu, Z.-H.; Bhongle, N.; Su, X.; Ribe, S.; Senanayake, C. H. *Tetrahedron Lett.* **2002**, *43*, 8617), but the reagent was very low reactivity for **14** (or **15**) (data not shown).
24. Karplus, M. *J. Phys. Chem.* **1959**, *30*, 11.
25. Cui, C. B.; Kakeya, H.; Okada, G.; Onose, R.; Ubukata, M.; Takahashi, I.; Isono, K.; Osada, H. *J. Antibiot.* **1995**, *48*, 1382.
26. Robbers, J. E.; Floss, H. G. *Tetrahedron Lett.* **1969**, *23*, 1857.
27. Ishida, K.; Matsuda, H.; Murakami, M.; Yamaguchi, K. *Tetrahedron* **1996**, *52*, 9025.
28. McIntire, W. S.; Wemmer, D. E.; Chistoserdov, A.; Lindstrom, M. E. *Science* **1991**, *252*, 817.
29. Hofsteenge, J.; Müller, D. R.; de Beer, T.; Löffler, A.; Richter, W. J.; Vliegenthart, J. F. G. *Biochemistry* **1994**, *33*, 13524.
30. Jimenez, E. C.; Craig, A. G.; Watkins, M.; Hillyard, D. R.; Gray, W. R.; Gulyas, J.; Rivier, J. E.; Cruz, L. J.; Olivera, B. M. *Biochemistry* **1997**, *36*, 989.
31. Anderson, L. B.; Maderia, M. M.; Ouellette, A. J. A.; Putnam-Evans, C.; Higgins, L.; Krick, T.; MacCoss, M. J.; Lim, H.; Yates, J. R., III; Barry, B. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14676.
32. Kamiya, Y.; Sakurai, A.; Takahashi, N.; Tsuchiya, E.; Abe, K.; Fukui, S. *Agric. Biol. Chem.* **1979**, *43*, 363.
33. Sakagami, Y.; Isogai, M.; Suzuki, A.; Kitada, C.; Fujino, M. *Agric. Biol. Chem.* **1979**, *43*, 2643.
34. Sakagami, Y.; Yoshida, M.; Isogai, A.; Suzuki, A. *Science* **1981**, *212*, 1525.
35. Rilling, H. C.; Breunger, E.; Epstein, W. W.; Crain, P. F. *Science* **1990**, *247*, 318.
36. Farnsworth, C. C.; Gelb, M. H.; Glomset, J. A. *Science* **1990**, *247*, 320.