



## Heterologous production of new lasso peptide koreensin based on genome mining

著者	Fuwa Hiroki, Hemmi Hikaru, Kaweewan Issara, Kozaki Ikko, Honda Hiroyuki, Kodani Shinya
journal or publication title	The Journal of Antibiotics
year	2020-08-27
出版者	Springer Nature
URL	<a href="http://hdl.handle.net/10297/00027718">http://hdl.handle.net/10297/00027718</a>

doi: 10.1038/s41429-020-00363-5

1 Title: Heterologous production of new lasso peptide koreensin based on genome mining

2 Running headline: Heterologous production of koreensin

3 Authors: Hiroki Fuwa,<sup>1</sup> Hikaru Hemmi,<sup>2</sup> Issara Kaweewan,<sup>3</sup> Ikko Kozaki,<sup>4</sup> Hiroyuki

4 Honda,<sup>4</sup> Shinya Kodani<sup>1,3,5\*</sup>

5 Affiliations: <sup>1</sup>Graduate School of Integrated Science and Technology, Shizuoka

6 University, Shizuoka, Japan, <sup>2</sup>Food Research Institute, National Agriculture and Food

7 Research Organization (NARO), Ibaraki, Japan, <sup>3</sup>Graduate School of Science and

8 Technology, Shizuoka University, Shizuoka, Japan, <sup>4</sup>Department of Biomolecular

9 Engineering, Graduate School of Engineering, Nagoya University, <sup>5</sup>Academic

10 Institute, Shizuoka University, Shizuoka, Japan

11

12 \*To whom correspondence should be addressed: Shinya Kodani, College of Agriculture,

13 Academic Institute, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka 422-8529

14 Japan, Tel/Fax; +81(54)238-5008, E-mail; kodani.shinya@shizuoka.ac.jp

15 Keywords: lasso peptide, heterologous production, *Sphingomonas koreensis*

16

17 **Abstract**

18 Lasso peptides are a class of ribosomally biosynthesized and posttranslationally  
19 modified peptides with a knot structure as a common motif. Based on a genome search,  
20 a new biosynthetic gene cluster of lasso peptides was found in the genome of the  
21 proteobacterium *Sphingomonas koreensis*. Interestingly, the amino acid sequence of the  
22 precursor peptide gene includes two cell adhesion motif sequences (KGD and DGR).  
23 Heterologous production of the new lasso peptide was performed using the cryptic  
24 biosynthetic gene cluster of *S. koreensis*. As a result, a new lasso peptide named  
25 koreensin was produced by the gene expression system in the host strain *S. subtarrenea*.  
26 The structure of koreensin was determined by NMR and ESI-MS analysis. The three-  
27 dimensional structure of koreensin was obtained based on an NOE experiment and the  
28 coupling constants. A variant peptide (koreensin-RGD) which had RGD instead of KGD  
29 was produced by heterologous production with site-directed mutagenesis experiment.  
30 Koreensin and koreensin-RGD did not show cell adhesion inhibitory activity, although  
31 the molecules possessed cell adhesion motifs. The possible presence of a salt bridge  
32 between the motifs in koreensin was indicated, and it may prevent the cell adhesion  
33 motif from functioning.  
34

## 35 **Introduction**

36 Lasso peptides are a class of ribosomally biosynthesized and posttranslationally  
37 modified peptides<sup>1-4</sup> (RiPPs) with a knot structure as a common motif.<sup>5-8</sup> The amino  
38 group of the *N*-terminal amino acid forms a peptide bond with the  $\beta$ - or  $\gamma$ -carboxyl  
39 group of Asp or Glu in the 7th-9th position from the *N*-terminus, resulting in the  
40 formation of a macrolactam. The macrolactam looks like a loop of a “lasso” with the C-  
41 terminal linear peptide, which normally passes through the ring, as the tail. The lasso  
42 peptide microcin J25 was isolated from *Escherichia coli*, and it is regarded as the  
43 archetype of lasso peptides.<sup>9</sup> Its biosynthetic gene cluster consists of four genes coding  
44 a precursor peptide (gene A: *mcjA*), two maturation enzymes (gene B: *mcjB* and gene C:  
45 *mcjC*) and an ATP-binding cassette transporter (gene D: *mcjD*) in one gene cluster  
46 (approximately 4.8 kbp).<sup>10</sup> Normally, lasso peptide biosynthetic genes in proteobacteria  
47 have the same set of genes, although the transporter gene is optional.

48 Through genome mining, Maksimov et al. investigated the genome data of more than  
49 3000 prokaryotic species in search of BGCs of lasso peptides and identified 76 bacteria  
50 that had BGCs of lasso peptides.<sup>11</sup> A prediction system of RiPPs named RODEO (Rapid  
51 ORF Description and Evaluation Online) was developed, and six lasso peptides were  
52 found based on the RODEO system.<sup>12</sup> A genome-mining approach and heterologous

53 expression using *Escherichia coli* as the host cell was performed to produce new lasso  
54 peptides.<sup>8, 13-22</sup> Attempts have also been made to utilize lasso peptides as fine chemical  
55 materials. Two cysteines were introduced into the amino acid sequence of microcin J25,  
56 followed by enzymatic cleavage.<sup>23</sup> Catenane-like molecules were generated by self-  
57 association.<sup>23</sup> The lasso peptide benenodin-I was reported to form two thermally  
58 interchangeable conformers.<sup>24</sup> The benenodin-I variant was prepared and digested with  
59 trypsin to afford a rotaxane-like molecule.<sup>24</sup> It is difficult to synthesize lasso peptides in  
60 vitro due to the instability of the modification enzymes. Recently, the in vitro synthesis  
61 of lasso peptides were reported by two different groups (fuscanodin/fusilassin).<sup>25, 26</sup>  
62 However, heterologous production is still the main method for exploring new lasso  
63 peptides due to the instability of the modification enzymes. For the heterologous  
64 production of lasso peptides, we established a gene expression system using the host  
65 strain *Sphingomonas subteranea*, a natural producer of the lasso peptide subterisin.<sup>27</sup> As  
66 a result, a new lasso peptide named brevunsin was produced by this system with a high  
67 yield.<sup>28</sup> Previously, the genetic engineering of lasso peptides afforded a microcin J25  
68 (RGDF) variant that had the cell adhesion motif RGD in the loop structure.<sup>29</sup> This  
69 engineering turned lasso peptide microcin J25 into a highly potent and selective  $\alpha v \beta 3$   
70 integrin inhibitor.<sup>29</sup> A lasso peptide precursor gene that possesses the cell adhesion

71 motifs KGD<sup>30, 31</sup> and DGR<sup>32, 33</sup> in the amino acid sequence was found in the genome of  
72 the proteobacterium *Sphingomonas koreensis* through genome mining. The KGD motif  
73 was reported to exist in the snake venom protein barbourin, which functions as a platelet  
74 aggregation inhibitor by binding specifically to integrin IIb-IIIa receptors on the platelet  
75 surface.<sup>30, 31</sup> To clarify characteristic of KGD motif, the mimicking cyclic peptide  
76 containing RGD (cyclo-(S, S)-GCGRGDWPCA-NH<sub>2</sub>) or KGD (cyclo-(S, S)-  
77 GCGKGDWPCA-NH<sub>2</sub>) was subjected to inhibition assays of fibrinogen binding to  
78 integrin IIb-IIIa, and vitronectin binding to the receptor  $\alpha_v\beta_3$ .<sup>34</sup> As a result, both  
79 peptides indicated more than 90% inhibition at the concentration of 10  $\mu$ M in fibrinogen  
80 binding to integrin IIb-IIIa.<sup>34</sup> Interestingly, inhibition of vitronectin binding to the  
81 receptor was observed in the assay of the cyclic RGD peptide, but not in the assay of the  
82 cyclic KGD peptide. This result indicated that the cyclic KGD peptide possessed higher  
83 inhibition selectivity to integrin IIb-IIIa.<sup>34</sup> The DGR motif is known to be the cell  
84 adhesion motif of the bone-cell secreted-signal peptide osteopontin.<sup>32, 33</sup> The DGR  
85 containing peptide was reported to inhibit the attachment of cells to laminin and type I  
86 collagen.<sup>35</sup> However, DGR containing peptides were indicated to be less than 1000  
87 times as active as the RGD containing peptides in inhibiting the attachment of cells to  
88 fibronectin.<sup>36</sup> These reports prompted us to accomplish the heterologous production of

89 a new lasso peptide using the cryptic gene cluster of *S. koreensis*. As a result, we  
90 succeed in the heterologous production of a new lasso peptide named korensin (**1** in  
91 Fig. 1A). In addition, korensin variant peptide containing RGD (korensin-RGD, **2** in  
92 Fig. 1A) was obtained by the site-directed mutagenesis experiment to investigate the  
93 effect on cell adhesion.

94

## 95 **Results and Discussion**

96 During genome mining, lasso peptide precursor coding genes were found in the  
97 genome data of bacteria belonging to the *Sphingomonadaceae* family by a BLASTP  
98 search using the amino acid sequence of the lasso peptide benenodin-1<sup>24</sup> precursor  
99 peptide coding gene (*benAI*). The alignment revealed the consensus motif -Leu-Ile/Val-  
100 Asp-Leu-Gly- in the leader peptide sequence (Fig. 2). However, the core peptide amino  
101 acid sequences of the core peptides had no conserved motifs except for 1<sup>st</sup> Gly and 8<sup>th</sup>  
102 Asp/Glu. The precursor-encoding genes had a common motif, -Thr-X- as the last two  
103 amino acids in the leader peptide. The core sequences started with Gly and had Asp/Glu  
104 in the 8<sup>th</sup> position counting from the 1<sup>st</sup> Gly of the core peptide (Fig. 2). Normally, the  
105 Lasso peptide forms a macrolactam between the 1<sup>st</sup> Gly and the 7<sup>th</sup>-9<sup>th</sup> Asp or Glu.  
106 Therefore, the macrolactam was expected to be formed between the 1<sup>st</sup> Gly and 8<sup>th</sup>  
107 Asp/Glu in the peptides (Fig. 2). Among the precursors, a precursor peptide (*korA*,  
108 coding gene accession number: WP\_107525062.1) had the natural cell adhesion motifs  
109 KGD and DGR in ring and tail sequences (gray background letters in Fig. 2). Since the  
110 amino acid sequence of *korA* contains two motifs (KGD and DGR), we expect that the  
111 mature lasso peptide of *S. koreensis* may have cell adhesion inhibitory activity.  
112 Therefore, we performed heterologous production using a gene cluster containing *korA*.



113 As shown in Fig. 3, the lasso peptide biosynthetic gene cluster consisted of 8 genes  
114 coding GntR family transcriptional regulator (WP\_066573229.1), precursor peptide  
115 (korA: WP\_107525062.1), modification enzymes (korB: WP\_157926355.1 and korC:  
116 WP\_066573225.1), isopeptidase (WP\_157926354.1), TonB dependent receptor  
117 (WP\_075152903.1), FecR-like protein (WP\_066573215.1), and FecI-like protein  
118 (WP\_082737884.1). The organization in the gene cluster is the same as that of the lasso  
119 peptide astexin.<sup>21</sup> The minimum gene set required to produce a mature lasso peptide is  
120 three genes (*korA*, *korB*, and *korC*). The gene *korB* includes a peptidase domain that  
121 may cleave the leader peptide, and *korC* codes cyclase to form an isopeptide bond. As  
122 shown in Fig. 3, the new lasso peptide korensin was expected to be produced from  
123 precursor KorA by modification by enzymes KorB and KorC.

124 To perform the heterologous production of the expected peptide, the minimum set of  
125 biosynthetic genes (*korA*, *korB*, and *korC*) was cloned. Previously, the expression vector  
126 pHSG396Sp was established to produce a new lasso peptide in the proteobacterium  
127 *Sphingomonas subterranea*.<sup>28</sup> The DNA fragment including the genes *korA*, *korB*, and  
128 *korC* was amplified by PCR and integrated into the expression shuttle vector  
129 pHSG396Sp to afford pHSG396Sp-16723L (Fig. S1). The plasmid was cloned using *E.*  
130 *coli* DH5 $\alpha$  and transformed into the host bacterial strain *S. subterranea* by

131 electroporation following a previous report.<sup>28</sup> The transformant (*S. subterranea*  
132 possessing the plasmid pHSG396Sp-16723L) was cultured in liquid modified basal  
133 medium at 30 °C for 9 days. The culture was evaporated and extracted with MeOH. The  
134 MeOH extract was analyzed by HPLC (Fig. S2) and ESI-MS. As a result, the  
135 transformant was indicated to produce the expected peptide korensin (**1** in Fig. 1A). To  
136 obtain enough **1** for structure determination and biological testing, the transformant was  
137 cultured using 1 L of modified basal medium. After open column chromatography, the  
138 60% MeOH fraction was repeatedly subjected to HPLC purification to obtain 2.8 mg of  
139 **1** from 1 L of culture.

140 The molecular formula of **1** was established as C<sub>73</sub>H<sub>113</sub>N<sub>21</sub>O<sub>23</sub> by accurate ESI-MS  
141 analysis (Fig. S3), since the ion corresponding to [M+2H]<sup>2+</sup> (the calculated m/z value,  
142 826.9237) was observed at m/z 826.9239. To determine the planar chemical structure,  
143 NMR spectroscopy, including <sup>1</sup>H, <sup>13</sup>C, DEPT135, DQF-COSY, TOCSY, NOESY,  
144 HMBC, and HSQC experiments (Fig. S4-S11), was performed on **1** (2.0 mg) in 0.5 mL  
145 of DMSO-*d*<sub>6</sub>. All 17 amino acids in the peptide were assigned using spin system  
146 identification from the 2D NMR spectral data (Table 1). Three linear peptide sequences  
147 (Gly1-Pro2-Lys3-Gly4, Asp5-Phe6-Pro7-Asp8-Val9-Gly10-Asp11-Gly12-Arg13-Ile14,  
148 and Leu15-Ala16-Gly17) were established by NOESY and HMBC data (Fig. 1B). The

149 *N*-terminus of Gly1 was indicated to be attached to the side chain of Asp8 to form an  
150 isopeptide bond based on HMBC cross-peaks from the amide proton of Gly1 to the  $\beta$ -  
151 carbonyl carbon of Asp8 and the NOESY correlation between the amide proton of Gly1  
152 and the  $\beta$ -proton of Asp8. Above all, the assignment of all the constituent amino acids  
153 of **1** was accomplished. Interestingly, a C-terminal peptide with four amino acids  
154 (LTDE) was truncated during heterologous production, possibly by endogenous  
155 proteases of *S. subterranea*.

156 The absolute stereochemistries of the constituent amino acids in **1** were analyzed by  
157 modified Marfey's method (Fig. S12-S25).<sup>37</sup> All the constituent amino acids were  
158 determined to be in the L-form, as shown in Fig. 1A. Three-dimensional analysis was  
159 performed on **1** based on NOE correlations and the coupling constants obtained in the  
160 NMR experiments. As shown in Fig. 4, three-dimensional (3D) structure of **1** was  
161 obtained. The amino acid sequence of Gly1 to Asp8 formed a macrolactam (yellow in  
162 Fig. 4B). The amino acids from Val9 to Asp11 formed a loop structure, and the amino  
163 acids of the C-terminus sequence from Gly12 to Gly17 formed the tail. The 3D structure  
164 of **1** was indicated to be a typical "lasso" structure (the atomic coordinate data were  
165 deposited in the Protein Data Bank, PDB ID: 7BW5). The bulky amino acid Arg13  
166 seemed to be a steric lock preventing the tail from becoming unthreaded from the

167 macrolactam. The cell adhesion motif KGD was in the macrolactam exposed to the  
168 outside, and the other cell adhesion motif (DGR) was located inside the macrolactam.  
169 The thermostability of **1** was tested at 50, 70, and 95 °C for 3 h. Interestingly,  
170 compound **1** was stable during incubation even at 95 °C for 3 h (Fig. S26). The lasso  
171 peptide astexin-2 gave unthreaded form due to thermal stress, while the related lasso  
172 peptide astexin-3 was thermostable.<sup>38</sup> As same as astexin-3, compound **1** was resistant  
173 to the thermal stress.

174 To obtain korensin variant peptide (korensin-RGD, **2** in Fig. 1A) which have RGD  
175 sequence instead of KGD sequence, the site-directed mutagenesis experiment was  
176 performed on the korensin expression vector pHSG396Sp-16723L. As a result, the  
177 mutated vector pHSG396Sp-16723L-RGD was obtained. In the same manner with  
178 korensin, isolation of korensin variant (korensin-RGD, **2** in Fig. 1A) which had Arg3  
179 instead of Lys3 (Fig. 1A) was performed. As a result, **2** was obtained by HPLC  
180 purification with the yield of 5.2 mg from 1 L of culture. The structure of **2** was  
181 determined by ESI-MS and NMR (Fig. S27 and S28, 2D NMR data of **2** is not shown)  
182 in the same manner with **1**.

183 To determine whether **1** and **2** had inhibitory activity on cell adhesion, a cell adhesion  
184 assay using human umbilical vein endothelial cells (HUVECs) was performed in the

185 presence of **1** or **2**. As a result, compounds **1** and **2** did not show inhibitory activity  
186 against cell adhesion at doses from 5 to 5000 nM (Fig. S29). This result was unexpected  
187 since compounds **1** and **2** had cell adhesion motif of KGD or RGD (Fig. S29). Further  
188 investigation of the 3D structure of **1** indicated the possibility of the formation of a salt  
189 bridge between the amino residue of Lys3 and the carboxyl residue of Asp11. The  
190 distance between the nitrogen in the amino residue of the side chain of Lys3 and the  
191 oxygen in the carboxyl residue of Asp11 was calculated to be approximately 4.4  
192 angstroms in the lowest energy structure (Fig. 4C). This seems a little bit long for a salt  
193 bridge since a salt bridge is normally within 4 angstroms. The peptide could be  
194 conformational flexible in solution, so we could not exclude the possibility of salt  
195 bridge formation. On the other hand, the distance between the nitrogen in the guanidine  
196 group of Arg13 and the oxygen in the carboxyl residue of Asp5 was calculated to be  
197 more than 7 angstroms, which seemed to be too far to form a salt bridge. The formation  
198 of a salt bridge between Lys3 and Asp11 may be one possible reason why compound **1**  
199 did not show cell adhesion inhibition activity. Although the 3D structure of peptide  
200 variant korensin-RGD (**2**) was not obtained, korensin-RGD may possess similar 3D  
201 structure with **1**. The reason of the inactivity of **2** may be explained by the  
202 intramolecular salt bridge between Arg3 and Asp11. The existence of a salt bridge

203 explains the structural stability of **1** against high thermal stress. Further mutation  
204 experiments on the expression plasmid are required to obtain recombinant korensin, in  
205 which Asp11 is replaced with a neutral amino acid such as Ala to preclude salt bridge  
206 formation and achieve inhibitory activity on cell adhesion.

207

## 208 **Materials and methods**

### 209 **Bacterial strains**

210 The microorganisms (Bacterial strains including *Sphingomonas subterrenea*  
211 NBRC16086<sup>T</sup> and *Sphingomonas koreensis* NBRC16723<sup>T</sup>) were obtained from the  
212 NBRC culture collection (NITE Biological Resource Center, Japan).

### 213 **Construction of expression vector for korensin**

214 For the template for PCR amplification, genome DNA was extracted from the cells of  
215 *S. koreensis* following previous report.<sup>28</sup> The DNA fragment including a lasso peptide  
216 korensin biosynthetic gene cluster (about 2.7 kbp, korA, korB, and korC) was  
217 amplified by PCR with template and the primer pair of 16723L-F  
218 (TGTCTCTAGATGCATCATGCGCAGCGCTCA) and 16723L-R  
219 (ATTACGGTACCGCCAAGCACGCGCAATATACG), using EmeraldAmp PCR Master

220 Mix (Takara Bio Inc., Shiga, Japan) by following the manufacturer's instructions. The  
221 insert DNA fragment including korensin biosynthetic gene cluster and the shuttle  
222 vector pHSG396Sp<sup>28</sup> were digested with XbaI (New England Biolabs Inc., Ipswich,  
223 United States) and KpnI (New England Biolabs Inc.) according to the manufacturer's  
224 instructions. The DNA products were ligated using T4 DNA ligation mix (Takara Bio  
225 Inc.) to afford the shuttle vector pHSG396Sp-16723L (Fig. S2). *E. coli* DH5 $\alpha$  cells  
226 were transformed with 2  $\mu$ L of the ligation mixture by chemical competence  
227 transformation and cells were placed onto LB agar plates containing chloramphenicol  
228 (final concentration of 20  $\mu$ g/mL).

#### 229 **Transformation of expression vector to *Sphingomonas subterranea***

230 The expression shuttle vector pHSG396Sp-16723L was transformed into *S.*  
231 *subterranea* using electroporation. The bacterium *S. subterranea* was cultured in 10 mL  
232 of NBRC medium number 802 liquid medium (peptone 10 g, yeast extract 2 g,  
233 MgSO<sub>4</sub>·7H<sub>2</sub>O 1 g in distilled water 1 L, pH 7.0) at 30 °C for 24h with shaking at 50  
234 rpm. The cells were cooled on ice for 30 min, followed by centrifugation (4000 rpm,  
235 4 °C, 10 min). The harvested cells were suspended in 10 ml of cold 10% glycerol  
236 solution. After centrifugation (4000 rpm, 4 °C, 10 min), The cells were harvested and  
237 re-suspended in 10 ml of cold 10% glycerol solution. After centrifugation (4000 rpm,

238 4 °C, 10 min), the harvested cells were re-suspended in 0.1 mL of cold 10% glycerol  
239 solution for electroporation. The electroporation experiment to 0.1 mL of suspension of  
240 the cells was performed with the program “Ec3” (Voltage: 3.0 kV, number of pulse: 1)  
241 using MicroPulser (Bio-Rad Laboratories, California, USA). After electroporation, 0.5  
242 mL of SOC medium was immediately added to the suspension of cells. For recovery of  
243 damage, and the cells were incubated at 30 °C for 2 h. Then, the cells of *S. subterranea*  
244 were spread onto NBRC medium number 802 agar medium (peptone 10 g, yeast extract  
245 2 g, MgSO<sub>4</sub>·7H<sub>2</sub>O 1 g, agar 15 g in distilled water 1 L, pH 7.0) containing  
246 chloramphenicol (final concentration of 10 µg/mL). After incubation at 30 °C for 6  
247 days, colonies were picked and checked by colony PCR method to obtain *S.*  
248 *subterranea* harboring pHSG396Sp-16723L.

#### 249 **Isolation of koreensin**

250 *Sphingomonas subterranea* harboring pHSG396Sp-16723L was cultured using 1 L of  
251 modified basal medium<sup>39</sup> containing chloramphenicol (20 µg/mL, final concentration)  
252 with shaking of 120 rpm at 30 °C for 9 days. The modified basal medium was prepared  
253 by adding the inorganic compounds (K<sub>2</sub>SO<sub>4</sub>, 2 g; K<sub>2</sub>HPO<sub>4</sub>, 3 g; NaCl, 1 g; NH<sub>4</sub>Cl, 5 g;  
254 MgSO<sub>4</sub>·7H<sub>2</sub>O, 80 mg; CuCl<sub>2</sub>, 0.5 mg; MnSO<sub>4</sub>·H<sub>2</sub>O, 0.35 mg; FeCl<sub>3</sub>, 0.5 mg;  
255 CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.5 mg) in 1L of distilled water with adjusting pH 7.0. After autoclaving,  
256 the medium was supplemented with separately sterilized glucose and yeast extract at  
257 final concentrations of 0.25%, 0.005%, respectively. The culture including bacterial



258 cells was evaporated using rotary evaporator to aqueous residue (about 20 mL). The  
259 aqueous residue was extracted with MeOH (300 mL). After filtration with paper filter  
260 (Wattman No. 1 filter, GE Healthcare Life Sciences, Illinois, USA), the MeOH extract  
261 was concentrated to aqueous residue using rotary evaporator. The aqueous residue was  
262 subjected to open column chromatography using hydrophobic resin CHP-20P  
263 (Mitsubishi Chemical Co., Tokyo, Japan), eluted with 10% MeOH, 60% MeOH, and  
264 100% MeOH. The 60% MeOH fraction was subjected to HPLC analysis using ODS  
265 column (4.6 × 250 mm, 5 μm, Wakopak Handy-ODS, Wako Pure Chemical Industries  
266 Ltd., Osaka, Japan) with gradient elution from 20% to 50% MeCN containing 0.05%  
267 trifluoroacetic acid for 20 min. The UV detector in HPLC system was set at 220 nm to  
268 detect **1** (retention time: 11.2min, Fig. S2), along with subterisin (retention time:  
269 14.3min, Fig. S2). For isolation of koreensin, 60% MeOH fraction was repeatedly  
270 subjected to HPLC purification using ODS column (4.6 × 250 mm, 5 μm, Wakopak  
271 Handy-ODS, Wako Pure Chemical Industries, Ltd., Osaka, Japan) with isocratic elution  
272 of 22% MeCN containing 0.05% trifluoroacetic acid with UV detector set at 220 nm to  
273 isolate 2.8 mg of **1** (retention time, 15.8min).

#### 274 **ESI-MS experiment**

275 Koreensin (**1**) was dissolved in 30% MeCN (0.1 mg/mL) and accurate ESI-MS of **1**  
276 was measured with JMS-T100LP (JEOL Co. ltd, Tokyo, Japan) using polyethylene  
277 glycol 1000 as internal standard.

#### 278 **NMR experiments and Structure calculations**

279 A NMR sample was prepared by dissolving **1** (2.0 mg) or **2** (2.0 mg) in 500  $\mu\text{L}$  of  
280 DMSO- $d_6$ . All NMR spectra were obtained on Bruker Avance 600 and Avance III HD  
281 800 spectrometers (Bruker, Massachusetts, United States) with quadrature detection in  
282 the phase-sensitive mode by States-TPPI (time proportional phase incrementation) and  
283 in the echo-antiecho mode. NMR spectra were acquired according to previous report.<sup>28</sup>  
284 Structure calculations of the three-dimensional structure of koreensin in solution state  
285 were also performed by using distant restraints and backbone  $\Phi$  dihedral angle  
286 restraints obtained from the NMR data according to the previous report.<sup>23</sup>

### 287 **Modified Marfey's analysis**

288 The modified Marfey's analyses of **1** were carried out in sealed vacuum hydrolysis tube.  
289 Compound **1** (0.5 mg) was hydrolyzed in 500  $\mu\text{L}$  of 6 N HCl at 110  $^\circ\text{C}$  for 16 h. After  
290 cooling to room temperature, the hydrolysate was completely evaporated using rotary  
291 evaporator and freeze-dried under vacuum. The hydrolysate was resuspended in 200  $\mu\text{L}$   
292 of water, followed by adding 10  $\mu\text{L}$  of the solution of *N* $\alpha$ -(5-fluoro-2,4-dinitrophenyl)-  
293 L-leucinamide (L-FDLA; Tokyo Chemical Industry Co., LTD, Tokyo, Japan) in acetone  
294 (10  $\mu\text{g}/\mu\text{L}$ ). The 100  $\mu\text{L}$  of 1M NaHCO<sub>3</sub> solution was added to the hydrolysate,  
295 followed by incubation at 80  $^\circ\text{C}$  for 3 min. The reaction mixture was cooled down to  
296 room temperature and neutralized with 50  $\mu\text{L}$  of 2N HCl. The 50% MeCN/water (1 mL)

297 was added to the mixture and then subjected to HPLC analysis. Each standard amino  
298 acid (1 mg) was derivatized with L-FDLA and D-FDLA in the same manner.

299 Approximately 20  $\mu$ L of each FDLA derivative was analyzed by HPLC (C18 column,  
300  $4.6 \times 250$  mm, Wakopak Handy ODS, Wako Pure Chemical Industries, Tokyo, Japan).

301 DAD detector (MD-2018, JASCO, Tokyo, Japan) was used to detect amino acid  
302 derivatives, accumulating data of absorbance from 220 nm to 420 nm. HPLC analysis  
303 for standard amino acids, including Lys, Arg, Asp, Pro, Ala, Val, allo-Ile, Ile, Leu and  
304 Phe, was performed using solvent A (distilled water containing 0.05% TFA) and solvent  
305 B (MeCN containing 0.05% TFA) in linear gradient mode from 0 min to 70 min,  
306 increasing percentage of solvent B from 25% to 60% at a flow rate of 1 mL/min (HPLC  
307 condition 1). The retention times (min) for L- and D-FDLA derivatized amino acids  
308 (HPLC condition 1) were as follows; L-Lys-L-FDLA (21.03 min), L-Lys-D-FDLA  
309 (16.60 min), L-Arg-L-FDLA (21.36 min), L-Arg-D-FDLA (18.31 min), L-Asp-L-FDLA  
310 (26.97 min), L-Asp-D-FDLA (28.74 min), L-Pro-L-FDLA (33.28 min), L-Pro-D-FDLA  
311 (38.92 min), L-Ala-L-FDLA (33.77 min), L-Ala-D-FDLA (40.09 min), L-Val-L-FDLA  
312 (40.72 min), L-Val-D-FDLA (54.24 min), L-allo-Ile-L-FDLA (46.02 min), L-allo-Ile-D-  
313 FDLA (61.71 min), L-Ile-L-FDLA (46.32 min), L-Ile-D-FDLA (61.84 min), L-Leu-L-  
314 FDLA (47.28 min), L-Leu-D-FDLA (62.65 min), L-Phe-L-FDLA (48.66 min) and L-Phe-

315 D-FDLA (59.40 min). HPLC analysis for Val was performed using solvent A (distilled  
316 water containing 0.05% TFA) and solvent B (MeCN containing 0.05% TFA) in isocratic  
317 mode 40% of solvent B for 55 min at a flow rate of 1 mL/min (HPLC condition 2). The  
318 retention times (min) for L- and D-FDLA derivatized amino acids (HPLC condition 2)  
319 were as follows; L-Val-L-FDLA (20.77 min) and L-Val-D-FDLA (39.14 min).

### 320 **Thermostability test of korensin**

321 Concentration of **1** was adjusted to 0.5 mg/mL in DMSO. Aliquot sample (100  $\mu$ L  
322 each) was heated at 50  $^{\circ}$ C, 70  $^{\circ}$ C, and 95  $^{\circ}$ C for 3 h, followed by immediate cooling to  
323 4  $^{\circ}$ C. Each sample (50  $\mu$ L) was subjected to HPLC analysis using ODS column (4.6  $\times$   
324 250 mm, 5  $\mu$ m, Wakopak Handy-ODS, Wako Pure Chemical Industries Ltd., Osaka,  
325 Japan) with gradient elution from 20 to 50% MeCN containing 0.05% trifluoroacetic  
326 acid for 20 min with UV detector set at 220 nm.

### 327 **Site-directed mutagenesis experiment on expression vector pHSG396Sp-16723L to** 328 **obtain korensin variant peptide (korensin-RGD)**

329 The DNA fragment (about 2.7 kbp) was obtained by PCR method using primer set  
330 (16723L-KtoR-F: ATTAAGGGCCCCAGGGGGGACTTCCCCGAT, 16723L-KtoR-R:  
331 TTGAGTGAGCTGATACCGCTCGCCGCAGCC) and template DNA (pHSG396Sp-  
332 16723L). The insert DNA fragment and the vector pHSG396Sp-16723L<sup>28</sup> were

333 digested with ApaI (New England Biolabs Inc.) and KpnI (New England Biolabs Inc.)  
334 according to the manufacturer's instructions. The DNA fragments were ligated using T4  
335 DNA ligation mix (Takara Bio Inc.) to afford the vector pHSG396Sp-16723L-RGD.  
336 The mutation in the vector was confirmed by DNA sequencing. Transformation and  
337 expression experiments using pHSG396Sp-16723L-RGD were performed in the same  
338 manner with the vector pHSG396Sp-16723L.

### 339 **Cell culture**

340 Human Umbilical Vein Endothelial Cells (HUVEC) cells were provided as frozen  
341 cells after primary culture by the supplier (Kurabo, Osaka, Japan). HUVEC were  
342 maintained in 75 cm<sup>2</sup> flasks (658170; Greiner Bio-One, Frickenhausen, Germany)  
343 cultured in HuMedia-EG2 (KE-2150S; Kurabo) supplemented with attached additive.

### 344 **Cell adhesion assay**

345 A cell adhesion assay was conducted on SPOT arrays according to a previously  
346 described method with slight modifications.<sup>40</sup> The spots which synthesized RGDS  
347 peptide were punched out as a disk and embedded in a 96-well plate. Cells were stained  
348 with calcein-AM (Life Technologies Corporation, Carlsbad, CA, USA) for 30 min, and  
349  $1.0 \times 10^4$  cells/well were directly seeded on the disks with serum-free Dulbecco's  
350 modified Eagle medium (DMEM, 08458-16; Nacalai Tesque) containing RGDS or

351 Lasso peptide. Cells, soluble peptides and peptide disks were incubated for 5 h for cell  
352 adhesion on the peptide disks. After two washes using phosphate-buffered saline (PBS)  
353 to remove unattached cells, fluorescence intensity was measured on a Fluoroskan  
354 Ascent (type 374; Labsystems, Helsinki, Finland) at 485/505 nm (excitation/emission)  
355 wavelengths. For reproducibility, data from triplicate spots were averaged.

356

### 357 **Acknowledgments**

358 This study was supported by the Japan Society for the Promotion of Science by  
359 Grants-in-aids (grant number 20K05848), Koyanagi-Foundation, and Takahashi  
360 Industrial and Economic Research Foundation. The NMR spectra were recorded on  
361 Bruker Avance 600 and Avance III HD 800 spectrometers at Advanced Analysis  
362 Center, NARO.

### 363 **Conflict of interest**

364 The authors declare that they have no conflict of interest.

365

366 **References**

- 367 1. Budisa, N. Expanded genetic code for the engineering of ribosomally synthesized  
368 and post-translationally modified peptide natural products (RiPPs). *Curr Opin*  
369 *Biotechnol.* **24**, 591-598 (2013)
- 370 2. Letzel, A. C., Pidot, S. J. & Hertweck, C. Genome mining for ribosomally  
371 synthesized and post-translationally modified peptides (RiPPs) in anaerobic  
372 bacteria. *BMC Genomics.* **15**, 983 (2014)
- 373 3. Link, A. J. Biosynthesis: Leading the way to RiPPs. *Nat Chem Biol.* **11**, 551-552  
374 (2015)
- 375 4. Sardar, D. & Schmidt, E. W. Combinatorial biosynthesis of RiPPs: docking with  
376 marine life. *Curr Opin Chem Biol.* **31**, 15-21 (2016)
- 377 5. Maksimov, M. O., Pan, S. J. & James Link, A. Lasso peptides: structure, function,  
378 biosynthesis, and engineering. *Nat Prod Rep.* **29**, 996-1006 (2012)
- 379 6. Cheung-Lee, W. L. & Link, A. J. Genome mining for lasso peptides: past, present,  
380 and future. *J Ind Microbiol Biotechnol.* **46**, 1371-1379 (2019)
- 381 7. Maksimov, M. O. & Link, A. J. Prospecting genomes for lasso peptides. *J Ind*  
382 *Microbiol Biotechnol.* **41**, 333-344 (2014)
- 383 8. Hegemann, J. D., Zimmermann, M., Xie, X. & Marahiel, M. A. Lasso peptides:

- 384 an intriguing class of bacterial natural products. *Acc Chem Res.* **48**, 1909-1919  
385 (2015)
- 386 9. Salomon, R. A. & Farias, R. N. Microcin 25, a novel antimicrobial peptide  
387 produced by *Escherichia coli*. *J Bacteriol.* **174**, 7428-7435 (1992)
- 388 10. Solbiati, J. O., Ciaccio, M., Farias, R. N. & Salomon, R. A. Genetic analysis of  
389 plasmid determinants for microcin J25 production and immunity. *J Bacteriol.* **178**,  
390 3661-3663 (1996)
- 391 11. Maksimov, M. O., Pelczer, I. & Link, A. J. Precursor-centric genome-mining  
392 approach for lasso peptide discovery. *Proc Natl Acad Sci U S A.* **109**, 15223-15228  
393 (2012)
- 394 12. Tietz, J. I., et al. A new genome-mining tool redefines the lasso peptide  
395 biosynthetic landscape. *Nat Chem Biol.* (2017)
- 396 13. Jeanne Dit Fouque, K., Bisram, V., Hegemann, J. D., Zirah, S., Rebuffat, S. &  
397 Fernandez-Lima, F. Structural signatures of the class III lasso peptide BI-32169  
398 and the branched-cyclic topoisomers using trapped ion mobility spectrometry-  
399 mass spectrometry and tandem mass spectrometry. *Anal Bioanal Chem.* **411**,  
400 6287-6296 (2019)
- 401 14. Hegemann, J. D., et al. Xanthomonins I-III: a new class of lasso peptides with a



- 402 seven-residue macrolactam ring. *Angew Chem Int Ed Engl.* **53**, 2230-2234 (2014)
- 403 15. Hegemann, J. D., Zimmermann, M., Zhu, S., Klug, D. & Marahiel, M. A. Lasso  
404 peptides from proteobacteria: Genome mining employing heterologous  
405 expression and mass spectrometry. *Biopolymers.* **100**, 527-542 (2013)
- 406 16. Hegemann, J. D., Zimmermann, M., Xie, X. & Marahiel, M. A. Caulosegnins I-  
407 III: a highly diverse group of lasso peptides derived from a single biosynthetic  
408 gene cluster. *J Am Chem Soc.* **135**, 210-222 (2013)
- 409 17. Hegemann, J. D., Schwalen, C. J., Mitchell, D. A. & van der Donk, W. A.  
410 Elucidation of the roles of conserved residues in the biosynthesis of the lasso  
411 peptide paeninodin. *Chem Commun (Camb).* **54**, 9007-9010 (2018)
- 412 18. Cheung-Lee, W. L., et al. Discovery of ubonodin, an antimicrobial lasso peptide  
413 active against members of the *Burkholderia cepacia* complex. *Chembiochem.* **21**,  
414 1335-1340 (2019)
- 415 19. Cheung-Lee, W. L., Parry, M. E., Jaramillo Cartagena, A., Darst, S. A. & Link, A.  
416 J. Discovery and structure of the antimicrobial lasso peptide citrocin. *J Biol Chem.*  
417 **294**, 6822-6830 (2019)
- 418 20. Cheung-Lee, W. L., Cao, L. & Link, A. J. Pandonodin: a proteobacterial lasso  
419 peptide with an exceptionally long C-terminal tail. *ACS Chem Biol.* **14**, 2783-2792

- 420 (2019)
- 421 21. Zimmermann, M., Hegemann, J. D., Xie, X. & Marahiel, M. A. The astexin-1  
422 lasso peptides: biosynthesis, stability, and structural studies. *Chem Biol.* **20**, 558-  
423 569 (2013)
- 424 22. Martin-Gomez, H., Linne, U., Albericio, F., Tulla-Puche, J. & Hegemann, J. D.  
425 Investigation of the biosynthesis of the lasso peptide chaxapeptin using an *E. coli*-  
426 based production system. *J Nat Prod.* **81**, 2050-2056 (2018)
- 427 23. Allen, C. D. & Link, A. J. Self-assembly of catenanes from lasso peptides. *J Am*  
428 *Chem Soc.* **138**, 14214-14217 (2016)
- 429 24. Zong, C., Wu, M. J., Qin, J. Z. & Link, A. J. Lasso peptide benenodin-1 is a  
430 thermally actuated [1] rotaxane switch. *J Am Chem Soc.* **139**, 10403-10409 (2017)
- 431 25. Koos, J. D. & Link, A. J. Heterologous and *in vitro* reconstitution of fuscanodin,  
432 a lasso peptide from *Thermobifida fusca*. *J Am Chem Soc.* **141**, 928-935 (2019)
- 433 26. DiCaprio, A. J., Firouzbakht, A., Hudson, G. A. & Mitchell, D. A. Enzymatic  
434 reconstitution and biosynthetic investigation of the lasso peptide fusilassin. *J Am*  
435 *Chem Soc.* **141**, 290-297 (2019)
- 436 27. Kuroha, M., Hemmi, H., Ohnishi-Kameyama, M. & Kodani, S. Isolation and  
437 structure determination of a new lasso peptide subterisin from *Sphingomonas*

- 438 subterranea. *Tetrahedron Letters*. **58**, 3429-3432 (2017)
- 439 28. Kodani, S., Hemmi, H., Miyake, Y., Kaweewan, I. & Nakagawa, H. Heterologous  
440 production of a new lasso peptide brevunsin in *Sphingomonas subterranea*. *J Ind*  
441 *Microbiol Biotechnol*. **45**, 983-992 (2018)
- 442 29. Hegemann, J. D., et al. Rational improvement of the affinity and selectivity of  
443 integrin binding of grafted lasso peptides. *J Med Chem*. **57**, 5829-5834 (2014)
- 444 30. Minoux, H., Chipot, C., Brown, D. & Maignet, B. Structural analysis of the KGD  
445 sequence loop of barbourin, an  $\alpha$ IIb $\beta$ 3-specific disintegrin. *J Comput Aided Mol*  
446 *Des*. **14**, 317-327 (2000)
- 447 31. Scarborough, R. M., et al. Barbourin. A GPIIb-IIIa-specific integrin antagonist  
448 from the venom of *Sistrurus m. barbouri*. *J Biol Chem*. **266**, 9359-9362 (1991)
- 449 32. Horii, A., Wang, X., Gelain, F. & Zhang, S. Biological designer self-assembling  
450 peptide nanofiber scaffolds significantly enhance osteoblast proliferation,  
451 differentiation and 3-D migration. *PLoS One*. **2**, e190 (2007)
- 452 33. Tsutsumi, H., Kawamura, M. & Mihara, H. Osteoblastic differentiation on  
453 hydrogels fabricated from Ca<sup>2+</sup>-responsive self-assembling peptides  
454 functionalized with bioactive peptides. *Bioorg Med Chem*. **26**, 3126-3132 (2018)
- 455 34. Scarborough, R. M., et al. Design of potent and specific integrin antagonists.

- 456 Peptide antagonists with high specificity for glycoprotein IIb-IIIa. *J Biol Chem.*  
457 **268**, 1066-1073 (1993)
- 458 35. Yamada, K. M. & Kennedy, D. W. Peptide inhibitors of fibronectin, laminin, and  
459 other adhesion molecules: unique and shared features. *J Cell Physiol.* **130**, 21-28  
460 (1987)
- 461 36. Hautanen, A., Gailit, J., Mann, D. M. & Ruoslahti, E. Effects of modifications of  
462 the RGD sequence and its context on recognition by the fibronectin receptor. *J*  
463 *Biol Chem.* **264**, 1437-1442 (1989)
- 464 37. Harada, K. I., Fujii, K., Hayashi, K., Suzuki, M., Ikai, Y. & Oka, H. Application  
465 of D,L-FDLA derivatization to determination of absolute configuration of  
466 constituent amino acids in peptide by advanced Marfey's method. *Tetrahedron*  
467 *Lett.* **37**, 3001-3004 (1996)
- 468 38. Allen, C. D., Chen, M. Y., Trick, A. Y., Le, D. T., Ferguson, A. L. & Link, A. J.  
469 Thermal unthreading of the lasso peptides astexin-2 and astexin-3. *ACS Chem Biol.*  
470 **11**, 3043-3051 (2016)
- 471 39. Imbert, M. & Blondeau, R. Effect of light on germinating spores of *Streptomyces*  
472 *viridosporus*. *FEMS Microbiol Lett.* **181**, 159-163 (1999)
- 473 40. Kanie, K., et al. Collagen type IV-specific tripeptides for selective adhesion of

474 endothelial and smooth muscle cells. *Biotechnol Bioeng.* **109**, 1808-1816 (2012)

475

476

477 Figure legends

478 Figure 1. A) Chemical structure of **1** and **2**, B) Key correlations of 2D NMR of **1**

479 Figure 2. Alignment of amino acid sequences of lasso peptide precursor coding genes

480 similar to *benAI*. Bold letters: conserved amino acids. Underlined letters: core peptide.

481 Gray back ground letters: cell adhesion motif.

482 Figure 3. Biosynthetic gene cluster for production of **1**

483 Figure 4. NMR-derived structures of **1**: (A) superposition of the 15 lowest-energy

484 structures and (B) the lowest-energy structure of **1**. The isopeptide bond between Gly1

485 and Asp8 is shown in red. The ring-forming residues are shown in yellow, and the loop

486 and the tail in blue. The side chain of “steric lock” Arg13 is shown. (C) The distance

487 between nitrogen (orange) in amino residue of Lys3 and oxygen (cyan) in carboxyl

488 residue of Asp11

489

490 Figure 1

491

492

493

494

495

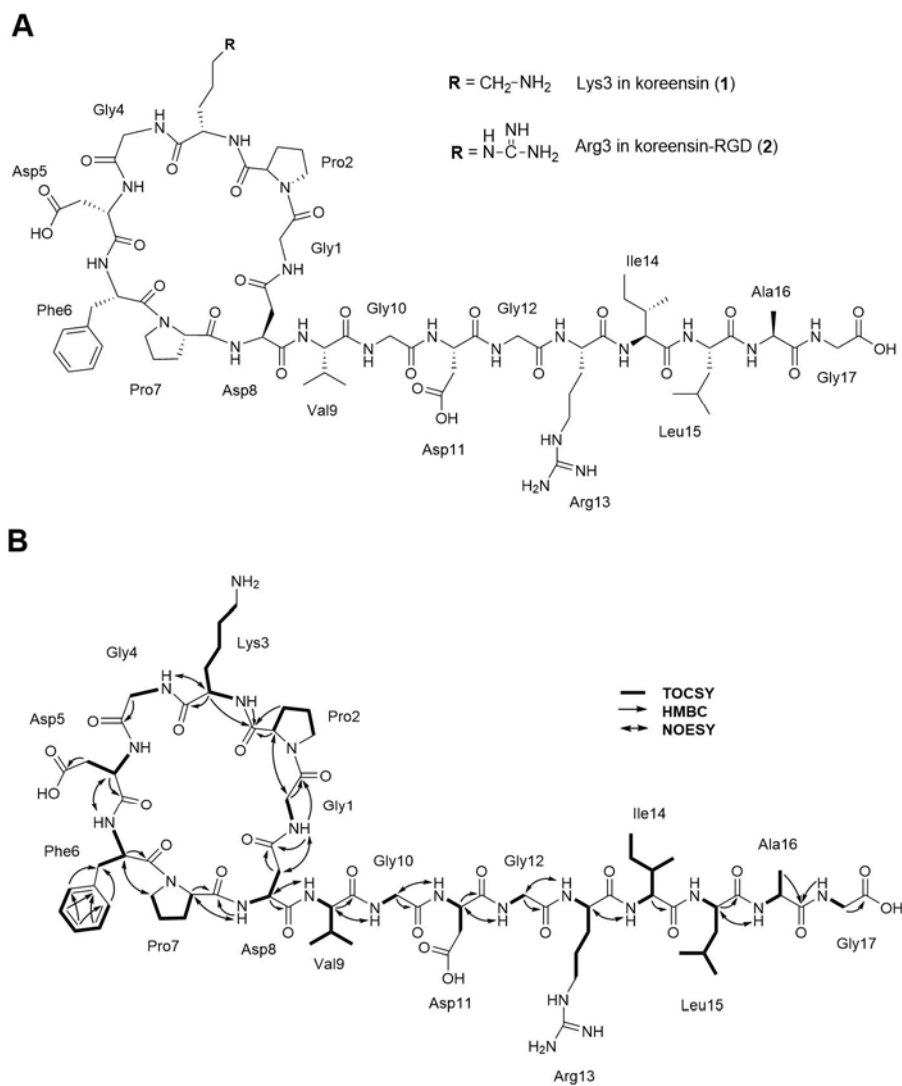


Figure 1. Fuwa et al.

496 Figure 2

497

498

499

<i>benA1</i> [ <i>Asticcacaulis benevestitus</i> ]	MEK--IETHED <u>LIDLGAASSETKGVGFGRPDSILTQEQA</u> KPMGLDRD
WP_107525062.1 [ <i>Sphingomonas koreensis</i> ]	MERN-HETPSD <u>LIDLGAASVETKGP</u> KGDFPDVGDGRILAGLTDE
WP_116091775.1 [ <i>Sphingomonas crusticola</i> ]	MER----INEEL <u>LIDLGAASVETKGP</u> GGKPGDVQLGRFELGLVED
WP_150127136.1 [ <i>Sphingomonas panacis</i> ]	MQRNDRLEGD <u>LIDLGDARVETKGN</u> DGNDIDMAHQRLQGGLSDD
WP_144033815.1 [ <i>Sphingomonas laterariae</i> ]	MER--IHEHDE <u>LIDLGTASVETRGP</u> WGPRFDPGAGEVLIPIGISDD
WP_128830833.1 [ <i>Sphingobium barthaii</i> ]	MQRE-METESD <u>LIDLGPVTVETKGI</u> AGLSRDQDQSPKGGAGILDD
WP_153002958.1 [ <i>Sphingomonas sanguinis</i> ]	MERNNDKVP-TLVDLGEARSL <u>TQGP</u> PAGTPIDEVQGFLLAAGLSDD
KEQ53303.1 [ <i>Sphingobium chlorophenolicum</i> ]	MERNDRHD-D <u>LIDLGAAS</u> TETQGRPGLQLEFGVIAQPIGIEAE

Figure 2. Fuwa et al.



500 Figure 3

501

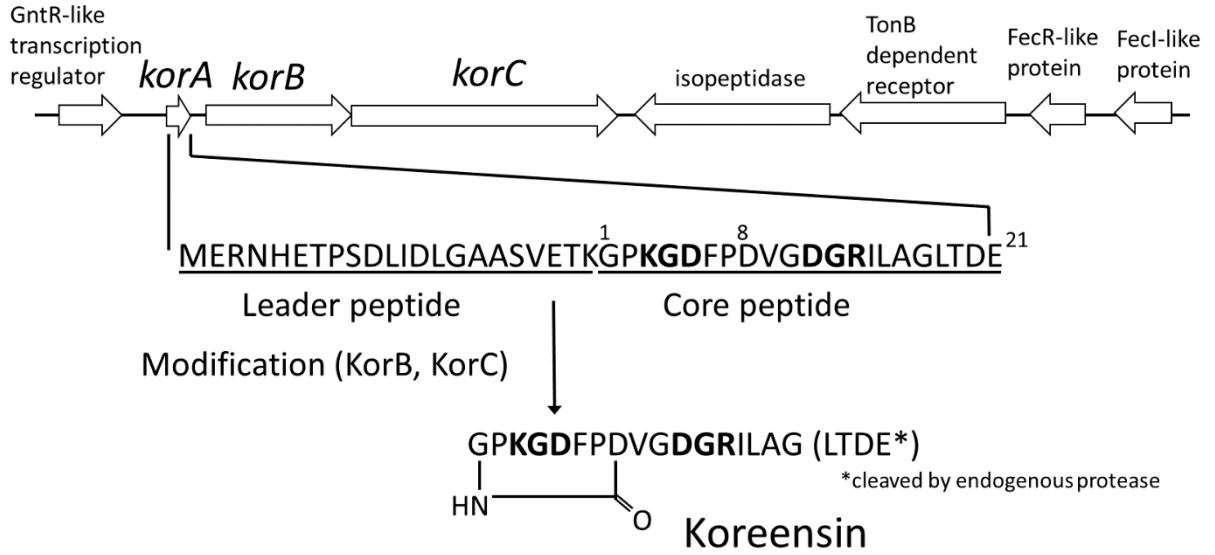


Figure 3. Fuwa et al.

502 Figure 4

503

504

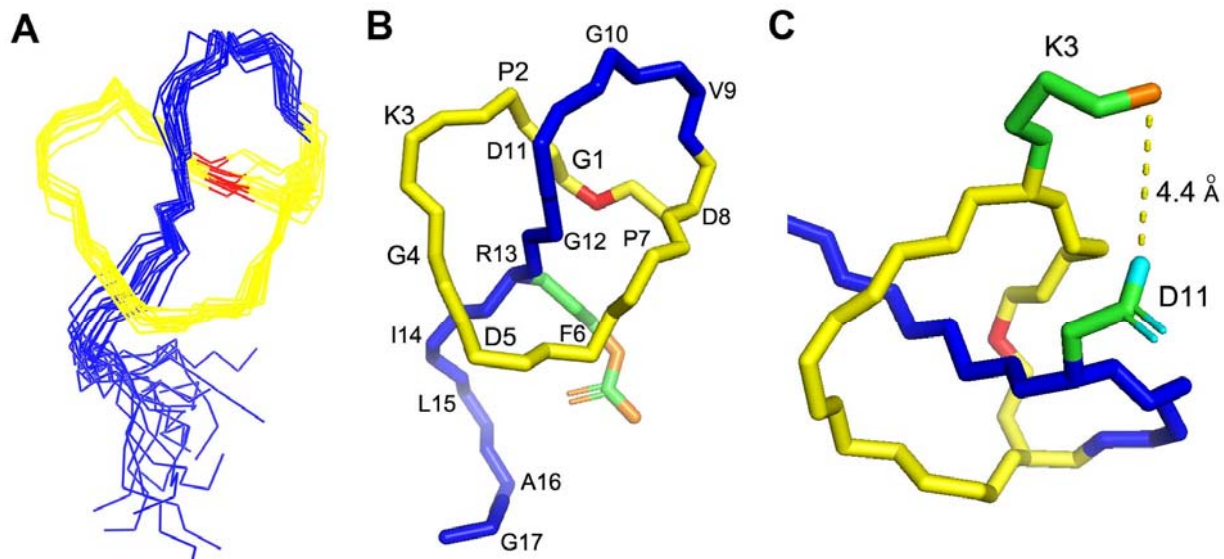


Figure 4. Fuwa et al.