Bioorganic & Medicinal Chemistry Letters 21 (2011) 4041–4044

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

The geranyl-modified tryptophan residue is crucial for $ComX_{RO\text{-}E\text{-}2}$ pheromone biological activity

Fumitada Tsuji [a], Ko Kobayashi [a], Masahiro Okada [b], Hisao Yamaguchi [a], Makoto Ojika [a], Youji Sakagami [a,*]

[a] Graduate School of Bioagricultural Sciences, Nagoya University, Chikusa, Nagoya, Aichi 464-8601, Japan
[b] Graduate school of Bioscience and Biotechnology, Chubu University, Kasugai, Aichi 487-8501, Japan

ARTICLE INFO

Article history:
Received 15 January 2011
Revised 20 April 2011
Accepted 27 April 2011
Available online 12 May 2011

Keywords:
Bacillus subtilis
ComX pheromone
Posttranslational modification
Quorum sensing
Structure–activity relationships

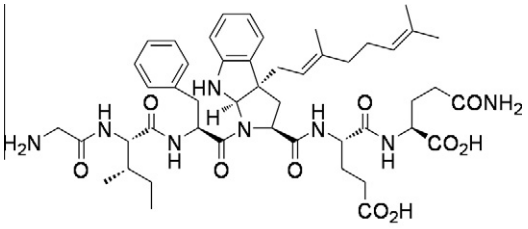
ABSTRACT

The ComX pheromone is an isoprenoidal oligopeptide containing a modified tryptophan residue, which stimulates natural genetic competence in gram-positive bacteria, *Bacillus*. We have reported the structure of the $ComX_{RO\text{-}E\text{-}2}$ pheromone, which is produced by the RO-E-2 strain of *Bacillus subtilis*. $ComX_{RO\text{-}E\text{-}2}$ analogs with substituted amino acids and isoprenoid modified tryptophan residues (e.g., prenyl, geranyl, and farnesyl), were synthesized and examined for biological activity. These results indicate that Phe-Trp*(Ger)-NH_2 is the minimum pharmacophore of the $ComX_{RO\text{-}E\text{-}2}$ pheromone. Furthermore, the length of the isoprenoid moiety (i.e., modification style), and the presence of double bonds, are crucial for biological activity. The modification style of the ComX pheromone is more important than the peptide sequence with respect to biological activity.

Quorum sensing is a well known phenomenon in which bacteria regulate their behavior, dependent on cell density.[1] Bacteria secrete specific extracellular signaling molecules and utilize such molecules to sense their cell density. *Bacilli* secrete the ComX pheromone, which stimulates natural genetic competence controlled by quorum sensing.[2] In the presence of the ComX pheromone, a membrane-located receptor (histidine kinase ComP) autophosphorylates and subsequently donates a phosphate via a two-component system ultimately activating competence gene expression.[3,4]

Recent efforts from our group have demonstrated that the $ComX_{RO\text{-}E\text{-}2}$ pheromone (**1**), from the *B. subtilis* strain RO-E-2, is a hexapeptide containing a geranyl-modified tryptophan residue.[5] This novel functional group modification results in the formation of a tricyclic structure (Fig. 1). Our group has also reported the structure of the farnesylated pheromone, $ComX_{RO\text{-}C\text{-}2}$.[6] We have been interested in defining the structure–activity relationships of $ComX_{RO\text{-}E\text{-}2}$ analogs, with the ultimate goal being the identification of a simple synthetic molecule that could control bacterial growth. These structure–activity relationships provide new insight into the interactions between the endogenous pheromone and its respective receptor. The ComX pheromone bears a unique structure and the ComP receptor belongs to the histidine kinase receptor-protein family. This constitutes a receptor class in which ligand–receptor interactions have not been adequately investigated. Several virulent bacteria are known to be taxonomically similar to *B. subtilis* (e.g., *Bacillus anthracis*). Their respective quorum sensing pheromones most likely share similar structures with those of the ComX pheromone, although studies on their quorum sensing systems have not been well established to date. We hope that this study will provide the foundation necessary for the further development of small molecules to be used in controlling these bacteria by the quorum sensing system.

We have reported that the tricyclic structure and stereochemical configuration of the modified tryptophan residue is necessary for pheromone function. In addition, we have explored the structure–activity relationships using an alanine scan and with truncated peptides.[7–9]



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

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

Figure 1. Chemical structure of $ComX_{RO\text{-}E\text{-}2}$ pheromone. Bold Trp*(Ger) represents the modified tryptophan residue with geranyl group in the $ComX_{RO\text{-}E\text{-}2}$ pheromone (bold lines).

* Corresponding author. Tel: +81 52 789 4116; fax: +81 52 789 4118.
E-mail address: ysaka@agr.nagoya-u.ac.jp (Y. Sakagami).

0960-894X/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2011.04.123

Here, we disclose the active subunit of the $ComX_{RO\text{-}E\text{-}2}$ pheromone and highlight the importance of the geranyl side chain pendant on the modified tryptophan residue. Various $ComX_{RO\text{-}E\text{-}2}$ analogs were synthesized using solid-phase synthetic techniques.[5,7–10] Biological activities of these peptides, as well as the $ComX_{RO\text{-}E\text{-}2}$ pheromone, were investigated using β-galactosidase assays.[3,5,7,8,10,11] These data were used to construct the respective dose–response curves. The biological activities of these peptides were represented by three values normalized to the activity of the $ComX_{RO\text{-}E\text{-}2}$ pheromone (**1**) (Table 1). The first and second values correspond to the concentrations of each peptide required to elicit the same β-galactosidase activity as that observed with the $ComX_{RO\text{-}E\text{-}2}$ pheromone at EC_{25} and EC_{50}, respectively. The third is the maximum activity as a percentage of that obtained with the $ComX_{RO\text{-}E\text{-}2}$ pheromone.

The structure–activity relationships of the tripeptides was investigated in comparison with [3-5]$ComX_{RO\text{-}E\text{-}2}$ (**2**), the smallest peptide known to exhibit biological activity as demonstrated in previous work.[9] [E5A][3-5]$ComX_{RO\text{-}E\text{-}2}$ (**3**) showed slightly higher biological activity than tripeptide **2**, and dose-dependent activity was observed with [E5G][3-5]$ComX_{RO\text{-}E\text{-}2}$ (**4**) (Table 1). These results indicate that the C-terminal amino acid residue of the tripeptide can be replaced by other amino acids. In previous work, alanine scan experiments with the original $ComX_{RO\text{-}E\text{-}2}$ hexapeptide pheromone have demonstrated that the third phenylalanine is exchangeable with alanine.[9] However, the replacement of the N-terminal phenylalanine residue with alanine (tripeptide **5**) in the tripeptide **2**, significantly attenuates biological activity. Among the tripeptides, a fixed C-terminal alanine, tyrosine and isoleucine exchange (tripeptide **7** and **9**) was tolerated, but other amino acid replacements led to very low activities (**6**, **8**, **10**, **11**, **12**; Table 1). A structure–activity relationships as a function of aromaticity and/or steric bulk among the N-terminal amino acid residues of these tripeptides was not observed. These results indicate that the phenylalanine residue is not exchangeable with another amino acid residue on the tripeptide. Next, we synthesized the corresponding dipeptides to determine the minimum bioactive structure-subunit (i.e., the pharmacophore). We attempted to synthesize a C-terminal blocked dipeptide, [4-5]$ComX_{RO\text{-}E\text{-}2}$-OMe, because peptides containing a modified tryptophan residue at the C- or N-terminus led to low biological activity in previous studies.[9] Efforts to synthesize this dipeptide methyl ester were unfruitful, but [4-5]$ComX_{RO\text{-}E\text{-}2}$ (diketopiperazine **13**) was readily obtained. To prevent formation of the diketopiperazine, the N-terminus was functionalized as the acetyl amide **14**. Unfortunately, the dipeptides **13** and **14** did not elicit biological activity. Gratifyingly, bioactivity was observed with [4-5]$ComX_{RO\text{-}E\text{-}2}$-$NH_2$ (**15**), with activity comparable to that of **4**. Moreover, Bz-Trp*(Ger)-NH_2 (**16**) showed no biological activity. Therefore, it can be posited that the minimal bioactive subunit of the $ComX_{RO\text{-}E\text{-}2}$ pheromone is Phe-Trp*(Ger)-NH_2.

The effect of modifying the tryptophan side chain was also studied. [Pre]$ComX_{RO\text{-}E\text{-}2}$ (**17**) and [Far]$ComX_{RO\text{-}E\text{-}2}$ (**18**) were prepared, which contain prenyl and farnesyl side chains, respectively (Table 2). The peptide **17** lost all activity and **18** lost most activity. Previous studies have demonstrated that the tricyclic structure and the stereochemical configuration of the modified tryptophan moiety are important for biological activity.[7,8] Despite this, [Pre]-$ComX_{RO\text{-}E\text{-}2}$ (**17**) and [Far]$ComX_{RO\text{-}E\text{-}2}$ (**18**), retaining the typical tricyclic structure and stereochemical configuration of the parent tryptophan residue, led to reduced biological activities. These results indicate that the geranyl side chain of the tryptophan residue is an important factor in expressing biological activity in $ComX_{RO\text{-}E\text{-}2}$. These results are also in accord with the finding that the RO-E-2 tester strain exhibits cross activities with other geranylated pheromones, but little activity in the presence of $ComX_{RO\text{-}C\text{-}2}$ and $ComX_{168}$ pheromone, both of which have farnesyl modified tryptophan groups.[3]

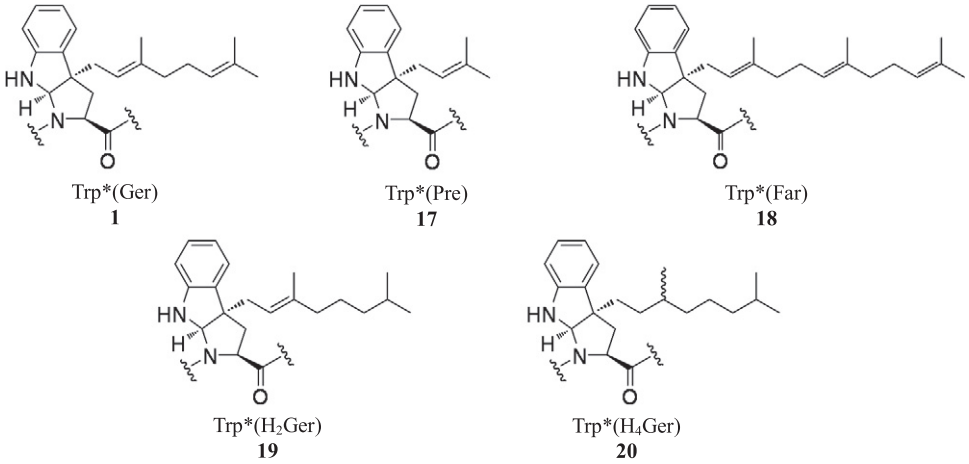
Next, we hydrogenated the Fmoc-geranylated tryptophan with Rhodium on Alumina and obtained two Fmoc-amino acids with saturated/partially saturated side chains. Using these compounds,

Table 1
Biological activities of $ComX_{RO\text{-}E\text{-}2}$ and $ComX_{RO\text{-}E\text{-}2}$ analogs

Peptide	Abbreviation	Amino acid sequence	EC_{25}[a] (nM)	EC_{50}[b] (nM)	ED_{max}[c] (%)
1	$ComX_{RO\text{-}E\text{-}2}$	Gly-Ile-Phe-**Trp*(Ger)**-Glu-Gln	0.3	1	100
2	[3-5]$ComX_{RO\text{-}E\text{-}2}$	Phe-**Trp*(Ger)**-Glu	8	20	60
3	[E5A][3-5]$ComX_{RO\text{-}E\text{-}2}$	Phe-**Trp*(Ger)**-Ala	3	60	70
4	[E5G][3-5]$ComX_{RO\text{-}E\text{-}2}$	Phe-**Trp*(Ger)**-Gly	200	>300	40
5	[F3A][3-5]$ComX_{RO\text{-}E\text{-}2}$	Ala-**Trp*(Ger)**-Glu	>300	>300	20
6	[F3A,E5A][3-5]$ComX_{RO\text{-}E\text{-}2}$	Ala-**Trp*(Ger)**-Ala	>300	>300	15
7	[F3Y,E5A][3-5]$ComX_{RO\text{-}E\text{-}2}$	Tyr-**Trp*(Ger)**-Ala	100	>300	25
8	[F3W,E5A][3-5]$ComX_{RO\text{-}E\text{-}2}$	Trp-**Trp*(Ger)**-Ala	>300	>300	15
9	[F3I,E5A][3-5]$ComX_{RO\text{-}E\text{-}2}$	Ile-**Trp*(Ger)**-Ala	20	>300	35
10	[F3L,E5A][3-5]$ComX_{RO\text{-}E\text{-}2}$	Leu-**Trp*(Ger)**-Ala	>300	>300	20
11	[F3P,E5A][3-5]$ComX_{RO\text{-}E\text{-}2}$	Pro-**Trp*(Ger)**-Ala	>300	>300	5
12	[F3N,E5A][3-5]$ComX_{RO\text{-}E\text{-}2}$	Asn-**Trp*(Ger)**-Ala	>300	>300	5
13	[4-5]$ComX_{RO\text{-}E\text{-}2}$	┌Phe-**Trp*(Ger)**┐	>300	>300	5
14	Ac-[4-5]$ComX_{RO\text{-}E\text{-}2}$-OMe	Ac-Phe-**Trp*(Ger)**-OMe	>300	>300	10
15	[4-5]$ComX_{RO\text{-}E\text{-}2}$-$NH_2$	Phe-**Trp*(Ger)**-NH_2	50	>300	35
16	Bz-Trp*(Ger)-NH_2	Bz-**Trp*(Ger)**-NH_2	>300	>300	0

[a] concentration of each peptide showing the same activity to the observed activity at EC_{25} of $ComX_{RO\text{-}E\text{-}2}$ pheromone.
[b] concentration of each peptide showing the same activity to the observed activity at EC_{50} of $ComX_{RO\text{-}E\text{-}2}$ pheromone.
[c] ratio of maximum activity compared with $ComX_{RO\text{-}E\text{-}2}$ pheromone.

Table 2
Chemical structures and biological activities of $ComX_{RO\text{-}E\text{-}2}$ and $ComX_{RO\text{-}E\text{-}2}$ analogs



Peptide	Abbreviation	Amino acid sequence	EC_{25}[a] (nM)	EC_{50}[b] (nM)	ED_{max}[c] (%)
1	$ComX_{RO\text{-}E\text{-}2}$	Gly-Ile-Phe-**Trp*(Ger)**-Glu-Gln	0.3	1	100
17	[Pre]$ComX_{RO\text{-}E\text{-}2}$	Gly-Ile-Phe-**Trp*(Pre)**-Glu-Gln	>300	>300	0
18	[Far]$ComX_{RO\text{-}E\text{-}2}$	Gly-Ile-Phe-**Trp*(Far)**-Glu-Gln	20	>300	35
19	[H_2Ger]$ComX_{RO\text{-}E\text{-}2}$	Gly-Ile-Phe-**Trp*(H_2Ger)**-Glu-Gln	2	20	75
20	[H_4Ger]$ComX_{RO\text{-}E\text{-}2}$	Gly-Ile-Phe-**Trp*(H_4Ger)**-Glu-Gln	200	>300	40

[a] concentration of each peptide showing the same activity to the observed activity at EC_{25} of $ComX_{RO\text{-}E\text{-}2}$ pheromone.
[b] concentration of each peptide showing the same activity to the observed activity at EC_{50} of $ComX_{RO\text{-}E\text{-}2}$ pheromone.
[c] ratio of maximum activity compared with $ComX_{RO\text{-}E\text{-}2}$ pheromone.

we synthesized two peptides [H_2Ger]$ComX_{RO\text{-}E\text{-}2}$ (**19**) and [H_4Ger]-$ComX_{RO\text{-}E\text{-}2}$ (**20**), incorporating side chains with the same carbon skeleton as $ComX_{RO\text{-}E\text{-}2}$ (Table 2). [H_4Ger]$ComX_{RO\text{-}E\text{-}2}$ (**20**) was obtained as a mixture of epimers that were inseparable by HPLC. Therefore, [H_4Ger]$ComX_{RO\text{-}E\text{-}2}$ (**20**) was tested as a mixture of epimers. [H_2Ger]$ComX_{RO\text{-}E\text{-}2}$ (**19**) retained partial activity, but [H_4Ger]$ComX_{RO\text{-}E\text{-}2}$ (**20**) led to a reduced response. These results indicate that the double bonds of the geranyl group, as well as the length, are important for biological activity.

Tremerogen A-10 is a dodecapeptide mating pheromone of *Tremella mesenterica*, which has an *S*-farnesyl modified cysteine methyl ester at the C-terminus.[12,13] Earlier structure–activity relationship studies of tremerogen A-10 analogs have indicated that the longer isoprenoid chain generates the greatest biological activity, and that the alkyl chain is replaceable with an isoprenoid side chain.[14] These results suggest that the farnesyl moiety is not crucial in eliciting biological activity with tremerogen A-10.

In contrast, the geranyl moiety is critical for the observed biological activity associated with the ComX pheromone as mentioned above. Furthermore, the peptide chain length is crucial in tremerogen A-10, but it is not necessary in inducing biological activity with the ComX pheromone.[13,14] Despite the similar post-translational modifications between these two classes of isoprenoidal peptides, it is interesting that the structural requirements necessary for biological activity is completely different. This could imply that the interactions between the peptide containing the isoprenoidal tryptophan and its corresponding receptor are different from that of the peptide containing the isoprenoidal cysteine. It can be suggested that the ComX receptor (ComP), which is a membrane-located histidine kinase, distinguishes the length of the isoprenoid chain. FTase (farnesyltransferase) and GGTase-I (geranylgeranyltransferase type I), enzymes that catalyze the isoprenylation of cysteine, are known to distinguish the length of the isoprenoidal chain.[15] The mechanism of chain length distinction has been termed the 'ruler hypothesis'. This hypothesis is supported by X-ray crystallographic analysis of FTase and GGTase-I and enables one to account for the substrate specificity of the respective enzymes. Therefore, further analysis of the X-ray crystallographic data of ComP, as well as FTase and GGTase-I, could provide insight into the mechanism of isoprenoidal chain length distinction. Perhaps ComP can be classified as a new type of receptor that distinguishes the isoprenoidal chain length, but not the peptide sequence. Because, receptors that have a signal transduction system (two-component system) are distributed widely throughout bacteria, fungi and plants, it is possible that ComP exists as a novel receptor type in various organisms. According to our results from the structure–activity relationship studies of the $ComX_{RO\text{-}E\text{-}2}$ pheromone, the structure of the labeled ligand has been optimized to facilitate investigation of the molecular interactions between the ComX pheromone and the ComP receptor.

The quorum sensing pheromone of gram-negative bacteria, a so called autoinducer, is an acylhomoserine lactone, and the species specificity of the pheromones is dependent on the acyl group length.[16–18] Gram-positive bacteria, *Bacillus*, may also use the peptide pheromone side chain length with respect to species specificity. *B. subtilis* and related *Bacilli* may be unique gram-positive bacteria, because not only the modification, but also the receptors were evolved for the acquisition of species-specificity. In gram-negative bacteria, an autoinducer (acylhomoserine lactone) is a lead compound in the development of a novel type of drug for use in controlling bacterial growth.[19–21] This work provides the foundation necessary to aid in similar trials of virulent gram-positive bacteria in the near future.

Acknowledgements

This work was supported by a Grant-in-Aid for JSPS Fellows (no. 21 2118) for F.T. and a Grant-in-Aid for Scientific Research (S) (no. 18101009) for Y.S. from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2011.04.123.

References and notes

1. Fuqua, W. C.; Winans, S. C.; Greenberg, E. P. *J. Bacteriol.* **1994**, *176*, 269.
2. Magnuson, R.; Solomon, J.; Grossman, A. D. *Cell* **1994**, *77*, 207.
3. Ansaldi, M.; Marolt, D.; Stebe, T.; Mandic-Mulec, I.; Dubnau, D. *Mol. Microbiol.* **2002**, *44*, 1561.
4. Hamoen, L. W.; Venema, G.; Kuipers, O. P. *Microbiology* **2003**, *149*, 9.
5. Okada, M.; Sato, I.; Cho, S. J.; Iwata, H.; Nishio, T.; Dubnau, D.; Sakagami, Y. *Nat. Chem. Biol.* **2005**, *1*, 23.
6. Okada, M.; Yamaguchi, H.; Sato, I.; Tsuji, F.; Dubnau, D.; Sakagami, Y. *Biosci. Biotechnol. Biochem.* **2008**, *72*, 914.
7. Okada, M.; Sato, I.; Cho, S. J.; Dubnau, D.; Sakagami, Y. *Tetrahedron* **2006**, *62*, 8907.
8. Okada, M.; Sato, I.; Cho, S. J.; Suzuki, Y.; Ojika, M.; Dubnau, D.; Sakagami, Y. *Biosci. Biotechnol. Biochem.* **2004**, *68*, 2374.
9. Okada, M.; Yamaguchi, H.; Sato, I.; Cho, S. J.; Dubnau, D.; Sakagami, Y. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 1705.
10. The Fmoc-protected modified tryptophan residue, with a geranyl group and farnesyl group, were synthesized as previously reported.[5–9] The Fmoc-protected modified tryptophan residue functionalized with a prenyl group was synthesized using prenylbromide as the electrophile.[5,7,8] The ammonolysis of the tryptophan residue was effected using 10% aqueous ammonia. Trp*(Ger)-OMe, the synthetic intermediate of the Fmoc-protected modified tryptophan,[5,7,8] was used as the starting material in this transformation. Benzoylation was accomplished by condensing H-Trp*(Ger)-NH_2 and benzoic acid with *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HATU) and 1-hydroxy-7-azabenzotriazole (HOAt) in CH_2Cl_2, *N*,*N*-diisopropylethylamine (DIPEA) under nitrogen. The Fmoc-protected modified tryptophan residues containing the tetrahydrogeranyl group and the dihydrogeranyl group were prepared by hydrogenation using Rhodium–Alumina as the catalyst under an atmosphere of hydrogen gas. The Fmoc-protected tryptophan residue bearing a geranyl group was used as a starting material in this transformation. The structures of all modified tryptophan residues were confirmed by ^{1}H NMR spectroscopy. Other residues were purchased from commercial sources (Watanabe chemical, Nova biochem). Peptide bond formation was accomplished with a peptide synthesizer (Applied Biosystems 433A) except for the modified tryptophan coupling. After which, each crude peptide was purified by HPLC and the resulting product structures were confirmed by MS/MS analysis.
11. Biological activity was investigated as previously reported using the *B. subtilis* tester strain in the expression of a *srfA-lacZ* fusion, which responds to added $ComX_{RO\text{-}E\text{-}2}$ pheromone.[3,5,7,8] Briefly, the strain was cultured overnight and subsequently diluted 100-fold. The culture was added to a sample solution and incubated at 37 °C for 5 h at 150 rpm, then β-galactosidase activity was measured at 420 nm by a standard method using *o*-nitrophenyl-β-D-galactopyranoside at 30 °C. Each analog was tested at concentrations up to 300 nM due to the fact that the $ComX_{RO\text{-}E\text{-}2}$ pheromone response, and the analogs possessing the modified tryptophan residue, achieved saturation at this concentration.
12. Sakagami, Y.; Isogai, A.; Suzuki, A.; Tamura, S.; Kitada, C.; Fujino, M. *Agric. Biol. Chem.* **1979**, *43*, 2643.
13. Sakagami, Y.; Yoshida, M.; Isogai, A.; Suzuki, A. *Science* **1981**, *212*, 1525.
14. Fujino, M.; Kitada, C.; Sakagami, Y.; Isogai, A.; Tamura, S.; Suzuki, A. *Naturwissenschaften* **1980**, *67*, 406.
15. Lane, K. T.; Beese, L. S. *J. Lipid Res.* **2006**, *47*, 681.
16. Bassler, B. L.; Losick, R. *Cell* **2006**, *125*, 237.
17. Camilli, A.; Bassler, B. L. *Science* **2006**, *311*, 1113.
18. Williams, P.; Winzer, K.; Chan, W. C.; Camara, M. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **2007**, *362*, 1119.
19. Castang, S.; Chantegrel, B.; Deshayes, C.; Dolmazon, R.; Gouet, P.; Haser, R.; Reverchon, S.; Nasser, W.; Hugouvieux-Cotte-Pattat, N.; Doutheau, A. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5145.
20. Frezza, M.; Castang, S.; Estephane, J.; Soulere, L.; Deshayes, C.; Chantegrel, B.; Nasser, W.; Queneau, Y.; Reverchon, S.; Doutheau, A. *Bioorg. Med. Chem.* **2006**, *14*, 4781.
21. Morohoshi, T.; Shiono, T.; Takidouchi, K.; Kato, M.; Kato, N.; Kato, J.; Ikeda, T. *Appl. Environ. Microbiol.* **2007**, *73*, 6339.