Biosynthetic Gene Cluster of a D-Tryptophan-Containing Lasso Peptide, MS-271

Zhi Feng, Dr. Yasushi Ogasawara*, Satoshi Nomura, and Prof. Dr. Tohru Dairi*

Graduate School of Engineering, Hokkaido University, Sapporo, Hokkaido 060-8628, Japan

*Corresponding authors.

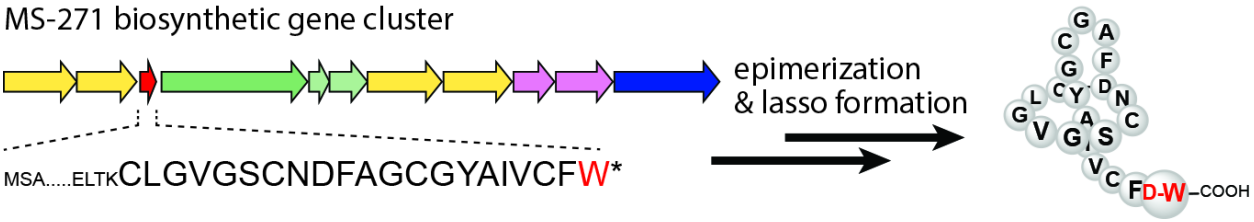
E-mail: yogasawa@eng.hokudai.ac.jp and dairi@eng.hokudai.ac.jp

Tel: +81-11-706-7815

Keywords

antibiotics, biosynthesis, D-amino acid, epimerase, lasso peptide

Graphical Abstract:



The biosynthetic gene cluster of the D-tryptophan-containing lasso peptide MS-271 of *Streptomyces* sp. M-271 was identified. Participation of a novel peptide epimerase in MS-271 biosynthesis was postulated by sequence analysis and heterologous expression of the gene cluster.

Abstract

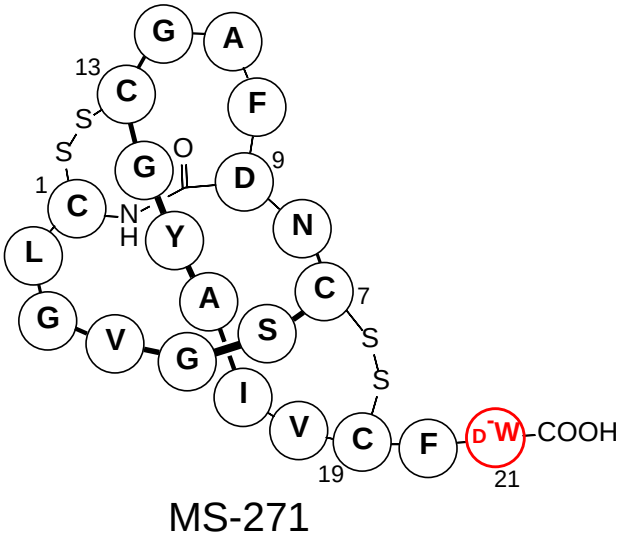
MS-271, produced by *Streptomyces* sp. M-271, is a lasso peptide natural product comprising 21 amino acid residues with a D-tryptophan (D-Trp) at its C-terminus. Because lasso peptides are ribosomal peptides, the biosynthesis of MS-271, especially the mechanism of D-Trp introduction, is of great interest. The MS-271 biosynthetic gene cluster was identified by draft genome sequencing of the MS-271 producer and it was revealed that the precursor peptide contains all 21 amino acid residues including the C-terminal tryptophan. This suggested that the D-Trp residue is introduced via epimerization. Modification enzyme genes such as a macrolactam synthetase (MslC), precursor peptide recognition element (MslB1), cysteine protease (MslB2), disulfide oxidoreductases (MslE, MslF), and a function-unknown protein (MslH) were found in the flanking region of the precursor peptide gene. Although obvious epimerase genes were absent in the cluster, heterologous expression of the putative MS-271 cluster in *Streptomyces lividans* showed that it contains all the necessary genes for MS-271 production including a gene for a novel peptide epimerase. Furthermore, a gene deletion experiment indicated that MslB1, B2, C and H were indispensable for MS-271 production and that some interactions of the biosynthetic enzymes were essential for the biosynthesis of MS-271.

Lasso peptides are a class of peptide natural products that have an intriguing interlocked lasso topology.[1] The lasso scaffold features an N-terminal macrocyclic ring composed of 7–9 amino acid residues and a C-terminal tail that threads through the ring, and this structural feature confers high stability against heat treatment and proteases. Since the first discovery of a lasso peptide, anantin[2], in 1991, many lasso peptides with a wide range of biological activities such as antimicrobial, antitumor, and receptor antagonistic activity have been isolated. Because of their promising bioactivity and chemical stability, lasso peptides have attracted attention from pharmacologists and natural product chemists in recent decades.

Concerning their biosynthesis, it has been shown that lasso peptides are derived from gene-encoded precursor peptides. A minimum of two modification enzymes are required to generate the lasso topology.[3] One is a cysteine protease, called the B enzyme, which cleaves the amide bond between the leader peptide and the core peptide in the precursor to provide the core peptide with a free N-terminal amino group.[4] The B enzyme consists of an N-terminal precursor peptide recognition element domain[5] and a C-terminal ATP-dependent cysteine protease domain. The other is a macrolactam synthetase that catalyzes ATP-dependent isopeptide bond formation between the N-terminal amino group and the carboxylic acid side chain of aspartic acid or glutamic acid to generate a macrocyclic ring in the lasso peptide.[4] The lasso peptide biosynthetic gene clusters from actinobacteria generally encode B enzymes in two open reading frames (ORFs), which correspond to the N- and C-terminal domains of the full-length B protein, and are called split B enzymes (B1 and B2).

MS-271 is a lasso peptide natural product isolated from *Streptomyces* sp. M-271 as a potent inhibitor of calmodulin-activated myosin light chain kinase.[6] It consists of 21 amino acid residues and contains a D-isomer of tryptophan (D-Trp) at its C-terminus (Figure 1) although all

known lasso peptides are ribosomal peptides.[7] To gain insight into this novel enzyme chemistry, we studied the biosynthesis of MS-271. In this study, we identified the biosynthetic gene cluster of MS-271 (*msl*) and showed that C-terminal Trp epimerization is catalyzed by an unprecedented epimerase encoded in the *msl* cluster.



MS-271

Figure 1. Structure of MS-271.

To investigate the biosynthesis of MS-271, the genome of *Streptomyces* sp. M-271 was sequenced using an Illumina HiSeq platform. Because all known lasso peptides are ribosomal peptides, we employed a tBlastn search (a search of translated nucleotide databases using a protein query) against the sequenced genome using the amino acid sequence of MS-271 as the query and identified a precursor peptide gene, *mslA*. The *mslA* gene encoded 42-residue peptide with a leader peptide at its N-terminus, and most importantly, the C-terminal core region contained all 21 amino acid residues of MS-271 including the C-terminal tryptophan (Figure 2). This suggested that the D-Trp residue is introduced via epimerization into a ribosomal peptide as a post-translational modification. Direct downstream of *mslA* gene were *mslC* (a macrolactam synthetase gene), *mslB1*, and *mslB2* (genes for the split B enzymes) genes, all of which are

required for the biosynthesis of lasso peptides. Furthermore, the following ORFs were also identified (Figure 2, Table 1). MslE had similarity to DoxX family enzymes[8], which were suggested to catalyze thiosulfate oxidation to generate tetrathionate. MslF showed homology to thiol–disulfide oxidoreductases (TDORs) and contained a conserved CXXC motif.[9] Thus, the two disulfide bonds in MS-271 may be formed by MslE and MslF. MslD1 to D4 exhibited homology to ABC transporters, which are a large family of membrane-associated proteins involved in natural product export and resistance. MslH was homologous to CapA family proteins, which are proteins of unknown function involved in the biosynthesis of poly-gamma-glutamate (PGA).[10] Although the role of MslH in MS-271 biosynthesis is not clear, all the genes between *mslC* and *mslH* apparently constitute an operon as indicated by translational coupling of these genes, and thus MslH likely participates in the biosynthesis of MS-271.

The *msl* cluster contained three homologues of known regulators of secondary metabolism. MslG and MslR1 exhibited sequence similarity to a large number of histidine protein kinases and LuxR-type response regulator proteins, respectively, and are likely regulators for two-component signal transduction.[11] In addition, *mslR2*, present in the far downstream region of the cluster, showed significant similarity to LuxR. Both MslR1 and MslR2 contained a C-terminal helix-turn-helix (HTH) DNA-binding domain typical for the NarL/FixJ family of regulators.

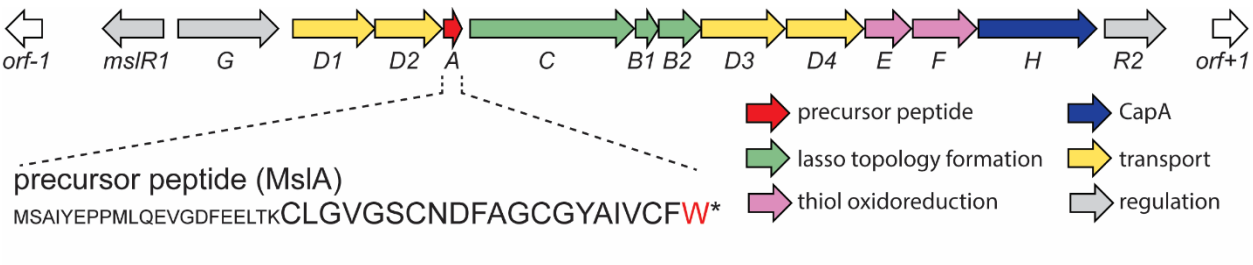


Figure 2. Gene organization of the *msl* cluster and the deduced amino acid sequence of the precursor peptide (MslA).

Table 1. Proposed functions of ORFs in the *msl* cluster.

ORFs	Length (AA)	Proposed functions	Typical homology, protein accession (% identity/% similarity) [organism]
MslR1	230	LuxR type regulator	DNA-binding response regulator, WP_067372316 (94/97), [*Streptomyces olivochromogenes*]
MslG	377	sensor histidine kinase	sensor histidine kinase, WP_062022673 (84/90) [*Streptomyces phaeopurpureus*]
MslD1	307	ABC transporter	ABC transporter ATP-binding protein, WP_059005878 (74/84) [*Streptomyces specialis*]
MslD2	250	ABC transporter	ABC transporter ATP-binding protein, WP_059005879 (70/82) [*Streptomyces specialis*]
MslA	42	precursor peptide	
MslC	613	lasso peptide isopeptide bond-forming cyclase	lasso peptide isopeptide bond-forming cyclase. WP_086700785 (77/84) [*Streptomyces vinaceus*]
MslB1	87	lasso peptide protease B1 protein	lasso peptide biosynthesis PqqD family chaperone, WP_059005881 (69/82) [*Streptomyces specialis*]
MslB2	156	lasso peptide protease B2 protein	lasso peptide biosynthesis B2 protein, WP_059005882 (87/90) [*Streptomyces specialis*]
MslD3	321	ABC transporter	ABC transporter ATP-binding protein, WP_059005883 (79/85) [*Streptomyces specialis*]
MslD4	289	ABC transporter	ABC transporter ATP-binding protein, WP_059005884 (82/88) [*Streptomyces specialis*]
MslE	181	Disulfide bond formation	DoxX family membrane protein, WP_075027174 (90/91) [*Streptomyces mirabilis*]
MslF	246	Disulfide bond formation	disulfide bond formation protein DsbA, WP_099941913 (94/96) [*Streptomyces* sp. 93]
MslH	440	Unknown	CapA family protein WP_051857567 (52/64) [*Streptomyces cellulosae*]
MslR2	224	LuxR type regulator	DNA-binding response regulator, WP_073930679 (90/93) [*Streptomyces* sp. CB02400]

Although the *msl* cluster lacked obvious candidate genes for the epimerization of the tryptophan residue, we examined whether it contained all necessary genes for MS-271 biosynthesis by heterologous expression, useful tool for characterization of lasso peptide genes.[12] The *msl* cluster spanning a ca. 11-kbp region from *mslR1* to *mslH* was cloned into pWHM3[13] to generate the plasmid pWHM3-msl and expressed in *Streptomyces lividans* TK23. The constructed strain (*S. lividans*/pWHM3-msl) was cultured along with wild-type *Streptomyces* sp. M-271 on agar plates containing the production medium. LC-MS analysis of the metabolites revealed the production of MS-271 (at retention time 23.8 min) in both

heterologous expression and wild type cultures (Figure 3, traces i and ii and Figures S1, S2). In addition, chiral analysis confirmed the tryptophan in the metabolites had D configuration, indicating that the product was indeed MS-271 (Figure S3). These results strongly indicated that the gene cluster contained all the genes responsible for the biosynthesis of MS-271, including a peptide epimerase gene. To date, only a small number of enzymes that catalyze peptide epimerization have been identified. These include a radical SAM epimerase (PoyD)[14] in polytheonamide biosynthesis and a glycopeptidyl-glutamate epimerase (MurL)[15] in the alternative peptidoglycan biosynthesis discovered in *Xanthomonas oryzae*. Apparently, the enzyme catalyzing the epimerization in MS-271 biosynthesis is distinct from the known epimerases.

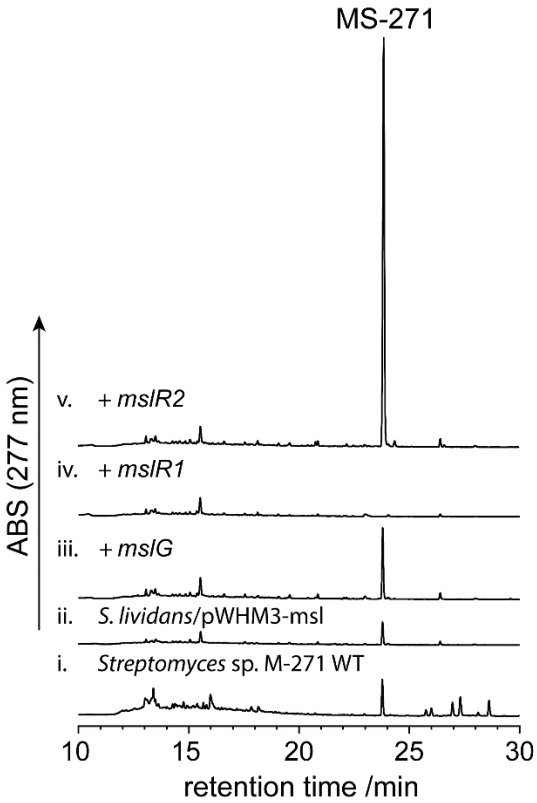


Figure 3. LC-MS analysis of products accumulated in the cultures. i) *Streptomyces* sp. M-271 wild type, ii) *S. lividans*/pWHM3-msl, iii) *S. lividans*:*mslG*/pWHM3-msl, iv) *S.*

lividans:*mslR1*/pWHM3-msl, v) *S. lividans*:*mslR2*/pWHM3-msl. Mass and MS/MS spectra of MS-271 produced by *S. lividans*/pWHM3-msl are shown in Figures S1 and S2.

Because the *msl* cluster contained three regulators including two putative pathway-specific LuxR type activators, we next investigated whether these regulatory genes were involved in the biosynthesis of MS-271. The *mslR1*, *mslG*, and *mslR2* genes were independently cloned downstream of an *ermE*P* promoter of pTYM18ep[16] and integrated into the *attB* site[17] on the genome of *S. lividans* TK23 to generate *S. lividans*:*mslR1*, *S. lividans*:*mslG*, and *S. lividans*:*mslR2*, respectively. The resulting strains were transformed with pWHM3-msl and their MS-271 productivities were evaluated by HPLC. The product yield was significantly increased by 18-fold (2.1 mg per plate) and 4-fold when pWHM3-msl was co-expressed with *mslR2* and *mslG*, respectively (Figure 3, traces iii–v). In contrast, the production of MS-271 decreased with *mslR1* overexpression. These results demonstrated that *mslR2* was a pivotal positive regulator for MS-271 biosynthesis. We thus used *S. lividans*:*mslR2* as a heterologous host for the following experiments.

To analyze the biosynthetic mechanism of peptide modification, we constructed pWHM3-msl-derived in-frame deletion plasmids lacking *mslB1*, *mslB2*, *mslC*, *mslE*, *mslF*, or *mslH* genes using Red/ET recombination. Each deletion construct was introduced into *S. lividans*:*mslR2*, and the resulting metabolite profiles were analyzed by LC-MS. The mutants of *mslE* and *mslF* showed production of MS-271, although their production levels were reduced (Figure 4, traces v and vi). Both MslE and F were predicted to be disulfide bond formation enzymes. We then searched for possible intermediates such as reduced forms of the lasso peptide; however, no intermediates were identified by LC-MS analysis. Considering that some gene clusters of disulfide-containing lasso peptides do not contain genes for disulfide bond

formation, disulfide bonds in lasso peptides are likely formed spontaneously, and MslE and MslF would assist the formation of correct disulfide bonds in the biosynthesis of MS-271.These observations are consistent with the previous report that the formation of disulfide bonds in sviceucin does not require SviF, MslF homolog in the sviceucin cluster.[18] In contrast, deletion of *mslB1*, *mslB2*, *mslC*, or *mslH* completely abolished MS-271 production (Figure 4, traces ii–iv and vii), which clearly demonstrated that these genes are indispensable for MS-271 biosynthesis. Furthermore, no intermediate was accumulated in any of the deletion mutants. These results suggested that interactions of all modification enzymes, including MslH, are required to produce MS-271. Similarly, post-translational modification enzymes were essential for the biosynthesis of an acetylated lasso peptide, albusnodin, in *Streptomyces albus*. Elimination of acetyltransferase gene (*albT*) from the albusnodin cluster completely diminished the formation of lasso peptide.[12a]

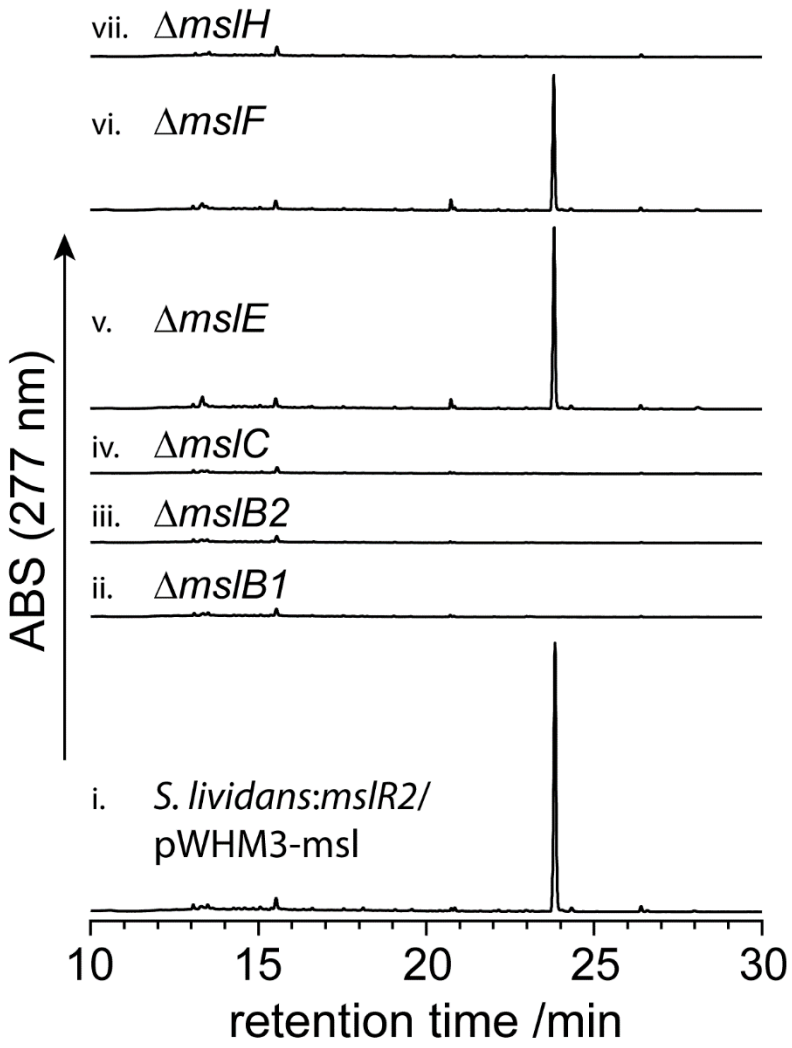


Figure 4. LC-MS analysis of metabolites produced by the mutants i) *S. lividans*:mslR/pWHM3-msl, ii) /pWHM3-msl-Δ*mslB1*, iii) /pWHM3-msl-Δ*mslB2*, iv) /pWHM3-msl-Δ*mslC*, v) /pWHM3-msl-Δ*mslE*, vi) /pWHM3-msl-Δ*mslF*, and vii) /pWHM3-msl-Δ*mslH*.

Among the proteins required for the biosynthesis of MS-271, MslH was the only protein of unknown function. We speculated MslH is responsible for the epimerization of the C-terminal tryptophan and carried out an *in vitro* experiment of recombinant MslH (Figure S4). Because all attempts to obtain full-length MslA were unsuccessful, we used a chemically synthesized core peptide (21 aa) as the substrate. However, the enzyme showed no epimerase activity even with reaction mixtures containing various co-factors, suggesting that an intrinsic substrate of MslH is either the full-length MslA or the cyclic peptide generated by MslB1, B2 and C.

To obtain a clue to the function of MslH, we then searched for *msl*-like gene clusters using the core amino acid sequences of MS-271 as the probe and identified several clusters in genome databases. All the clusters identified contained *mslH* orthologs like MS-271 and the core peptides in these clusters were almost identical to that of MslA (Figure S5). We also searched for gene clusters of lasso peptides possessing two disulfide bonds like MS-271 and found the sviceucin biosynthetic gene cluster as the sole example. Sviceucin consists of 20 amino acid residues and its sequence is different from that of MS-271 including its C-terminal amino acid (valine).[18] However, the genes in the biosynthetic gene cluster and their organization are the same as in MS-271 except for the absence of a *mslH* ortholog (Figure S5). These results suggested that MslH may be a novel epimerase responsible for the epimerization of the C-terminal tryptophan in MS-271 biosynthesis.

In conclusion, we identified the biosynthetic gene cluster of the D-amino acid-containing lasso peptide MS-271 from *Streptomyces* sp. M-271. Besides the gene for the precursor peptide (MslA), the cluster contained genes encoding enzymes for post-translational modifications such as a macrolactam synthetase (MslC), precursor peptide recognition element (MslB1), cysteine

protease (MslB2), disulfide oxidoreductases (MslE, MslF), and a protein of unknown function (MslH). An important finding was that the epimerization of D-Trp is catalyzed by a novel epimerase encoded in the *msl* cluster, as shown by heterologous expression of the cluster. Although it remains to confirm MslH unambiguously as responsible for the epimerization, this work has laid the foundation for further biochemical and mechanistic studies of MS-271 biosynthesis. Because the introduction of D-amino acids into peptides increases their stability, an understanding of the mechanism of the novel peptide epimerization would be important for biosynthetic pathway redesign to prepare unnatural peptide secondary metabolites.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgments

This study was supported in part by Grants-in-Aid for Research on Innovative Areas from MEXT, Japan (JSPS KAKENHI Grant Number 16H06452 to T.D.) and Grants-in-Aid for Scientific Research from JSPS (15H03110 and 18H03937 to T.D. and 18K05449 to Y.O.). We are grateful to Kyowa Hakko Bio Co. Ltd., Tokyo, Japan, for providing *Streptomyces* sp. MS-271. We also thank Prof. Hiroyasu Onaka at The University of Tokyo for providing pTYM19ep and Prof. Shinya Kodani at Shizuoka University for helpful discussion. We thank Robbie Lewis, MSc, from Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

Short Text for the Table of Contents

The biosynthetic gene cluster of the D-tryptophan-containing lasso peptide MS-271 of *Streptomyces* sp. M-271 was identified. Participation of a novel peptide epimerase in MS-271 biosynthesis was postulated by sequence analysis and heterologous expression of the gene cluster.

References

[1] a) J. D. Hegemann, M. Zimmermann, X. Xie, M. A. Marahiel, *Acc Chem Res* **2015**, *48*, 1909; b) M. O. Maksimov, S. J. Pan, A. James Link, *Nat Prod Rep* **2012**, *29*, 996.

[2] a) W. Weber, W. Fischli, E. Hochuli, E. Kupfer, E. K. Weibel, *J Antibiot* **1991**, *44*, 164; b) D. F. Wyss, H. W. Lahm, M. Manneberg, A. M. Labhardt, *J Antibiot* **1991**, *44*, 172.

[3] S. Duquesne, D. Destoumieux-Garzon, S. Zirah, C. Goulard, J. Peduzzi, S. Rebuffat, *Chem Biol* **2007**, *14*, 793.

[4] K. P. Yan, Y. Li, S. Zirah, C. Goulard, T. A. Knappe, M. A. Marahiel, S. Rebuffat, *Chembiochem* **2012**, *13*, 1046.

[5] B. J. Burkhart, G. A. Hudson, K. L. Dunbar, D. A. Mitchell, *Nat Chem Biol* **2015**, *11*, 564.

[6] a) R. Katahira, M. Yamasaki, Y. Matsuda, M. Yoshida, *Bioorg Med Chem* **1996**, *4*, 121; b) K. Yano, S. Toki, S. Nakanishi, K. Ochiai, K. Ando, M. Yoshida, Y. Matsuda, M. Yamasaki, *Bioorg Med Chem* **1996**, *4*, 115.

[7] P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, P. D. Cotter, D. J. Craik, M. Dawson, E. Dittmann, S. Donadio, P. C. Dorrestein, K. D. Entian, M. A. Fischbach, J. S. Garavelli, U. Goransson, C. W. Gruber, D. H. Haft, T. K. Hemscheidt, C. Hertweck, C. Hill, A. R. Horswill, M. Jaspars, W. L. Kelly, J. P. Klinman, O. P. Kuipers, A. J. Link, W. Liu, M. A. Marahiel, D. A. Mitchell, G. N. Moll, B. S. Moore, R. Muller, S. K. Nair, I. F. Nes, G. E. Norris, B. M. Olivera, H. Onaka, M. L. Patchett, J. Piel, M. J. Reaney, S. Rebuffat, R. P. Ross, H. G. Sahl, E. W. Schmidt, M. E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Stein, R. D. Sussmuth, J. R. Tagg, G. L. Tang, A. W. Truman, J. C. Vederas, C. T. Walsh, J. D. Walton, S. C. Wenzel, J. M. Willey, W. A. van der Donk, *Nat Prod Rep* **2013**, *30*, 108.

[8] S. Nambi, J. E. Long, B. B. Mishra, R. Baker, K. C. Murphy, A. J. Olive, H. P. Nguyen, S. A. Shaffer, C. M. Sassetti, *Cell Host Microbe* **2015**, *17*, 829.

[9] H. Kadokura, F. Katzen, J. Beckwith, *Annu Rev Biochem* **2003**, *72*, 111.

[10] M. Ashiuchi, *Microb Biotechnol* **2013**, *6*, 664.

[11] a) M. Y. Galperin, *Curr Opin Microbiol* **2010**, *13*, 150; b) S. Horinouchi, T. Beppu, *Gene* **1992**, *115*, 167; c) H. Rodriguez, S. Rico, M. Diaz, R. I. Santamaria, *Microb Cell Fact* **2013**, *12*, 127.

[12] a) C. Zong, W. L. Cheung-Lee, H. E. Elashal, M. Raj, A. J. Link, *Chem Commun (Camb)* **2018**, *54*, 1339; b) J. Mevaere, C. Goulard, O. Schneider, O. N. Sekurova, H. Ma, S. Zirah, C. Afonso, S. Rebuffat, S. B. Zotchev, Y. Li, *Sci Rep* **2018**, *8*, 8232.

[13] J. Vara, M. Lewandowska-Skarbek, Y. G. Wang, S. Donadio, C. R. Hutchinson, *J Bacteriol* **1989**, *171*, 5872.

[14] M. F. Freeman, C. Gurgui, M. J. Helf, B. I. Morinaka, A. R. Uria, N. J. Oldham, H. G. Sahl, S. Matsunaga, J. Piel, *Science* **2012**, *338*, 387.

[15] R. Feng, Y. Satoh, Y. Ogasawara, T. Yoshimura, T. Dairi, *J Am Chem Soc* **2017**, *139*, 4243.

[16] H. Onaka, S. Taniguchi, H. Ikeda, Y. Igarashi, T. Furumai, *J Antibiot* **2003**, *56*, 950.

[17] P. Combes, R. Till, S. Bee, M. C. Smith, *J Bacteriol* **2002**, *184*, 5746.

[18] Y. Li, R. Ducasse, S. Zirah, A. Blond, C. Goulard, E. Lescop, C. Giraud, A. Hartke, E. Guittet, J. L. Pernodet, S. Rebuffat, *ACS Chem Biol* **2015**, *10*, 2641.