Communication

# A methyltransferase initiates terpene cyclization in teleocidin B biosynthesis

Takayoshi Awakawa, Lihan Zhang, Toshiyuki Wakimoto, Shotaro Hoshino, Takahiro Mori, Takuya Ito, Jun Ishikawa, Martin E. Tanner, and Ikuro Abe

*J. Am. Chem. Soc.*, **Just Accepted Manuscript** • DOI: 10.1021/ja505224r • Publication Date (Web): 03 Jul 2014

# A methyltransferase initiates terpene cyclization in teleocidin B biosynthesis

Takayoshi Awakawa,[†] Lihan Zhang,[†] Toshiyuki Wakimoto,[†] Shotaro Hoshino,[†] Takahiro Mori,[†] Takuya Ito,[‡] Jun Ishikawa,[¶] Martin E. Tanner,[§] and Ikuro Abe*[,†]

[†] Graduate School of Pharmaceutical Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan.

[‡] Faculty of Pharmaceutical Sciences, Tokushima Bunri University, Yamashiro-cho, Tokushima 770-8514, Japan

[¶] National Institute of Infectious Diseases, Shinjuku-ku, Tokyo 162-8640, Japan

[§] Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada

*Supporting Information Placeholder*

**ABSTRACT:** Teleocidin B is an indole terpenoid isolated from *Streptomyces*. Due to its unique chemical structure and ability to activate protein kinase C, it has attracted interest in the areas of organic chemistry and cell biology. Here, we report the identification of genes encoding enzymes for teleocidin B biosynthesis, including non-ribosomal peptide synthetase (*tleA*), P-450 monooxygenase (*tleB*), prenyltransferase (*tleC*), and methyltransferase (*tleD*). The *tleD* gene, which is located outside of the *tleABC* cluster on the chromosome, was identified by transcriptional analysis and heterologous expression. Remarkably, TleD not only installs a methyl group on the geranyl moiety of the precursor, but also facilitates the nucleophilic attack from the electron-rich indole to the resultant cation, to form the indole-fused six-membered ring. This is the first demonstration of a cation, generated from methylation, triggering successive terpenoid-ring closure.

Teleocidin B (**1**) is a unique indole alkaloid, with an indolactam and a monoterpenoid moiety fused with C-6 and C-7 of the indole (Figure 1). It was first isolated from *Streptomyces mediocidicus*,[1] and consists of a mixture of four stereoisomers (B-1, B-2, B-3, and B-4).[2] **1** is also known as a potent protein kinase C activator,[3] which makes it a potentially valuable reagent in pharmaceutical and biochemical research. Thus, the unique structures and specific bioactivities have attracted the attention of organic chemists, leading to many isolation studies[2, 4] and total syntheses[5] of its derivatives. The biosynthetic routes to these molecules are also interesting, and have been investigated by chemical conversion or isotope-feeding experiments.[6] These studies revealed that the (−)-indolactam V (**2**) moiety of **1** (Figure 1) is biosynthesized from L-valine and L-tryptophan, and the monoterpenoid moiety is derived from the non-mevalonate pathway. However, the reaction mechanism for the formation of its indole-fused six-membered ring remained unclear. Therefore, we set out to identify the genes encoding the biosynthetic enzymes, and investigated their reactions *in vitro*, to shed light on the later stage of the teleocidin biosynthesis.

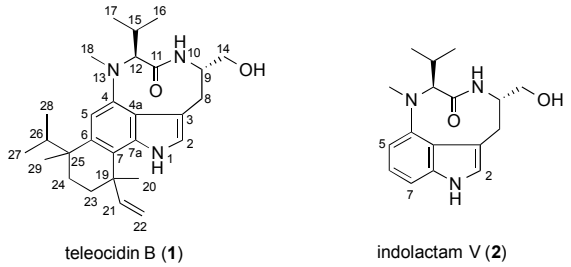



Figure 1 The chemical structures of teleocidin B and related compounds. (19*R*, 25*S*), (19*S*, 25*R*), (19*S*, 25*S*), and (19*R*, 25*R*)-teleocidin B are named teleocidin B-1 (**1a**), B-2 (**1b**), B-3 (**1c**), and B-4 (**1d**), respectively.

Lyngbyatoxin A (**3**, Scheme 1), which is structurally related to **1**, was isolated from the cyanobacterium *Moorea producens* (formerly *Lyngbya majuscula*).[7] In 2004, Edwards and Gerwick identified the lyngbyatoxin biosynthetic gene cluster, and characterized the reaction of an aromatic prenyltransferase (LtxC) to attach a geranyl group to the C-7 position of **2**.[8] A non-ribosomal peptide synthetase (LtxA) and a P-450 monooxygenase (LtxB) encoded in the *ltx* gene cluster were also identified as biosynthetic enzymes. Subsequently, LtxB was shown to oxidize and cyclize methyl-L-valyl-L-tryptophanol to yield **2** *in vitro*.[9] Given the structural similarity between **1** and **3**, we hypothesized that enzymatic homologs of LtxABC would be involved in teleocidin B biosynthesis. A Blast search of the draft genome sequence of the teleocidin B producing strain *Streptomyces blastmyceticus* NBRC 12747, using the amino acid sequence of LtxC as a query, revealed a gene encoding a prenyltransferase with

40% amino acid identity to LtxC, which was designated as TleC. The upstream region of *tleC* contained two genes, encoding a non-ribosomal peptide synthetase and a P-450 monooxygenase, which were named *tleA* and *tleB,* respectively. TleA has the same catalytic domain structure as LtxA with 49% identity, and TleB shares 47% identity with LtxB. Thus, we identified a 23.2 kb contig containing *tleABC,* as a candidate for the teleocidin biosynthetic gene cluster (Figure S1). On the other hand, no genes related to a *C*-methyltransferase (C-MT) or a terpene cyclase (TPC), which are thought to be involved in teleocidin B biosynthesis, were found in the region adjacent to the *tle* cluster.

To test whether the *tle* cluster is responsible for teleocidin B production, the fosmid bearing the *tle* cluster (pCC-*tle*) was introduced into *Streptomyces lividans* TK21. We engineered pCC-*tle* by λ-RED recombination for integration into the chromosome,[10] and introduced the resultant plasmid, pCCST-*tle,* into the host. The compounds that accumulated within the *S.lividans*/pCCST-*tle* cells, cultured in TP2 medium for 3 days, were analyzed by LC-ESIMS. As a result, compound **3**, which was not detected in the *S. lividans* harboring pCCST (void vector), was newly detected in the *S. lividans*/pCCST-*tle* (Figure 2A). In accordance with the HR-MS and $^{1}$H NMR spectra, and the rotation value in the literature,[7] **3** was identified as lyngbyatoxin A (Supplementary data). This result suggested that the enzymes encoded in the *tle* cluster are responsible for the production of **3**, but not for the methylation or the cyclization of the geranyl moiety required for the biosynthesis of teleocidin B. Therefore, we anticipated that one or more genes encoding the enzyme(s) that catalyze the *C*-methylation and/or cyclization of the geranyl moiety of lyngbyatoxin A are located outside of the *tle* cluster.

We searched for the genes encoding a putative C-MT responsible for the methylation of **3** in the draft genome sequence of *S. blastmyceticus,* and found nine candidate C-MT genes. Since an RT-PCR analysis revealed that three of them were not co-transcribed with *tleC* (Figure S3), they were excluded from the candidates. Each of the six remaining C-MT genes was introduced into *S. lividans* harboring the *tle* cluster. The LC-ESIMS analyses revealed that three compounds (**1a, 1d,** and **4**) (Scheme 1) were newly detected only in the cell-extract of the transformant that harbors *tleD*, which shares moderate homology with other C-MTs that catalyze the *C*-methylation of the prenyl group, such as the sterol 24-C-MT from *Magnaporthe oryzae*[11] and the geranyl diphosphate (GPP) 2-C-MT from *Streptomyces coelicolor* A3(2)[12] (Figure 2A).

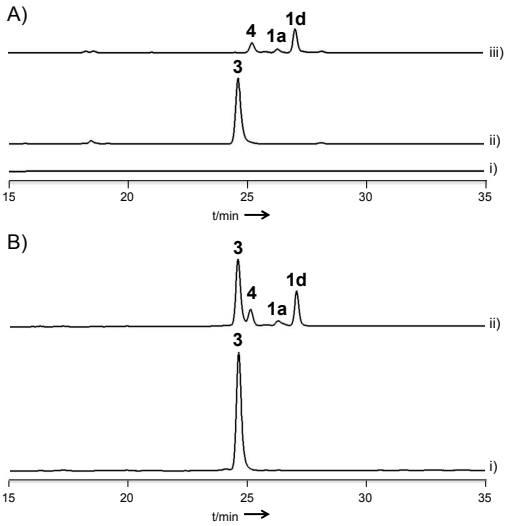



Figure 2 (A) LC-ESIMS profiles of cell extracts from transformants harboring i) the void vector, ii) the *tle* cluster, and iii) the *tle* cluster and *tleD.* (B) LC-ESIMS profiles of the products of the *in vitro* enzyme assays, in the i) absence or ii) presence of TleD, with **3** as a substrate. All of the chromatograms were monitored at 280 nm.

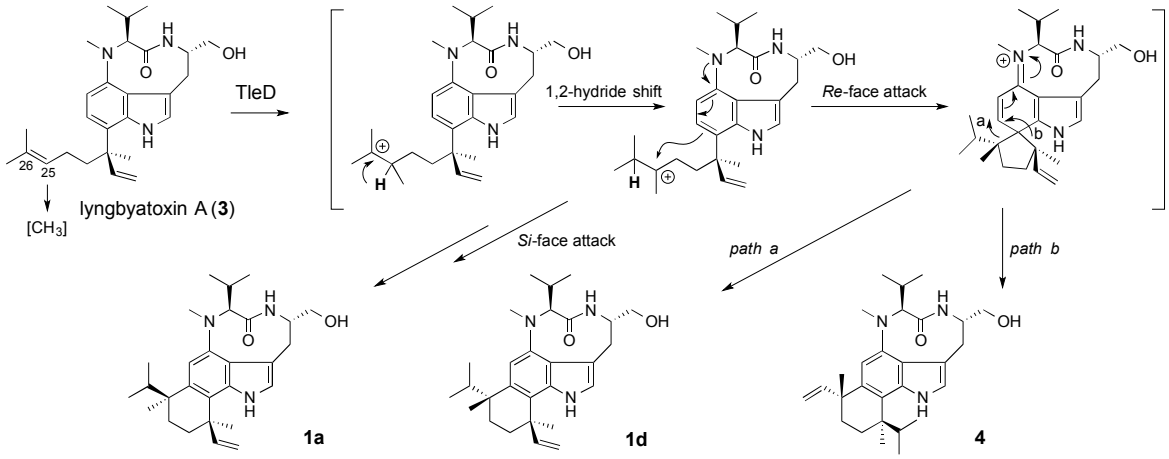
All three compounds showed psuedomolecular ions at *m/z* = 451 consistent with **1**. After a large scale fermentation and intensive spectroscopic analyses, **1d** was identified as teleocidin B-4, by comparing its $^{1}$H- and $^{13}$C- NMR, HR-MS spectra, and optical rotation value (supplementary data) with data from the literature.[13,14] The $^{1}$H and $^{13}$C NMR spectra of **1a** were similar to those of **1d**, except for a characteristic difference observed in the $^{13}$C chemical shift at C-20: 26.9 ppm for **1a,** and 21.5 ppm for **1d**. This difference is probably attributable to a γ-gauche effect between C-20 and the substituent at C-25.[2] Thus, **1a** was determined to be teleocidin B-1, which is the C-25 epimer of **1d**. Compound **4** was identified as des-*O*-methyl-olivoretin C, by comparing its $^{1}$H- and $^{13}$C- NMR, HR-MS spectra, and optical rotation value (supplementary data) with the published data of olivoretin C.[15] These data demonstrated that the TleD-mediated methylation of the geranyl moiety of **3** resulted in the formation of three cyclized products: **1a**, **1d**, and **4**. In the process of the assay, we also identified a MT that catalyzes the *O*-methylation of the C-14 hydroxyl group of **3** (Supplementary information). Interestingly, like *tleD,* this gene is also located outside of the *tle* cluster.

To understand the catalytic function of TleD, we performed *in vitro* enzyme reactions, using purified His-tagged recombinant TleD with **3** as a substrate. Our results confirmed that TleD alone successfully afforded the cyclized products **1a**, **1d** and **4** (Figure 2B). This finding clearly ruled out the involvement of

other enzymes in the cyclization reaction, and suggested that a cation generated by methylation initiated the cyclization of the geranyl moiety (Scheme 1). Next, to further characterize the enzyme reaction, we measured the steady-state kinetic parameters of TleD. The $k_{cat}$ and $K_m$ values for the formation of **4** were 0.37 $min^{-1}$ and 2.7 μM, while those for the formation of **1a** were 0.17 $min^{-1}$ and 3.4 μM, and for the formation of **1d** were 0.84 $min^{-1}$ and 3.2 μM. These values are comparable to those of other MTs,[16] thus indicating that TleD accepts **3** as a genuine substrate, and is responsible for the biosynthesis of teleocidin B. To confirm the native role of TleD in the biosynthesis, gene deletion study is required. Otherwise, we cannot rule out the possibility of having other genes to perform the same reaction, or the possibility for TleD to be involved in other pathways. Interestingly, TleD also accepted 14-*O*-methyllyngbyatoxin A, but with less efficiency (Figure S15).

Considering that both the TleD-catalyzed methylation and subsequent nucleophilic attack take place on the same carbon, we supposed that the C-25 cationic intermediate should be formed prior to cyclization, presumably *via* a 1,2-hydride shift (Scheme 1). To confirm the 1,2-hydride shift in the proposed pathway, we performed a deuterium labeling experiment. The TleC enzyme reaction with [6-$^2$H]-GPP and **2** afforded [25-$^2$H]-**3**, which was then used as a substrate for the TleD reaction. Mass spectrometry of the products from the reaction using [25-$^2$H]-**3** gave an *m/z* value of 452, which revealed that all three products, **1a**, **1d** and **4**, retained the deuterium atom (Figure S17**)**. In the $^1$H NMR spectrum of D-labeled **1d**, the multiplet signal corresponding to H-26 was no longer detected, and the two doublet methyl signals corresponding to H-27 and H-28 were observed as a singlet (Figure S19). These data provided convincing evidence that the deuterium atom at C-25 is shifted to C-26, and confirmed the 1,2-hydride shift in the cationic intermediate. The $k_{cat}$ and $K_m$ values for the formation of D-labeled **1a**, **1d** and **4** were comparable to those in the reaction using unlabeled **3**: the $k_{cat}$ and $K_m$ values for the formation of D-labeled **4** were 0.41 $min^{-1}$ and 2.4 μM, while those for the formation of D-labeled **1a** were 0.14 $min^{-1}$ and 3.7 μM, and for the formation of D-labeled **1d** were 0.88 $min^{-1}$ and 2.4 μM. These results excluded the deuterium isotope effect during the methylation/cyclization reaction, and confirmed that the 1,2-hydride shift in the cationic intermediate is not the rate-limiting step.

**Scheme 1** Proposed mechanism of TleD-catalyzed reaction.




Although the previous feeding experiments suggested that **3** was a precursor of **1**,[6b] the details of the reaction mechanism had not been reported. Our *in vitro* study clearly illustrated that **3** is the direct precursor of **1a**, **1d** and **4**, and the reaction is catalyzed by one enzyme TleD. Moreover, the deuterium labeling experiment corroborated the biosynthetic pathway originally proposed by Irie *et al* (Scheme 1).[6b] First, TleD installs a methyl group at C-25, and the hydride at C-25 migrates to C-26. The resultant cation at C-25 is then attacked by the nucleophilic C-7, to afford a spiro-fused intermediate. Finally, the C-C bond, which consists of the spiro-ring, migrates *via* path a, to afford **1d**. A similar spiro-ring compound was reported as the product in the Pictet-Spengler reaction of $N_b$-hydroxytryptamines and cysteinals,[17] thus supporting this proposed reaction mechanism. During the spiro-ring formation in the pathway to **1**, the facial selectivity of the nucleophilic attack to the C-25 cationic intermediate determines the C-25 stereochemistry of **1**. Since the two bulky substituents, the isopropyl and vinyl groups, should be on the same face of the five-membered ring

intermediate to afford **1a**, the facile and preferred production of **1d** over **1a** is presumably due to steric factors. Furthermore, **4** could be generated from the same spiro-ring intermediate, via path b. The production ratio of **1a** : **1d** : **4** is 0.5 : 2.3 : 1.0 (both in the *in vitro* and *in vivo* assay), indicating that path a is favored. Finally, it should be noted that no byproducts derived from deprotonation or hydroxylation of the C-25 cationic intermediate were detected, which suggested that the 1,2-hydride shift and the cyclization reaction take place in a concerted manner.

TleD shares 16.5%, 28.3%, and 33.9% amino acid sequence identities with the GPP 2-*C*-MT from *S. coelicolor* A3(2),[12] the sterol 24-*C*-MT from *M. oryzae*,[11] and the rebeccamycin sugar 4'-*O*-MT RebM from *Lechevalieria aerocolonigenes*,[16] respectively. Sequence comparisons revealed that TleD has the consensus GXGXG motif (Figure S20), which is a conserved SAM-binding site in MTs. Considering that TleD possesses Y200, corresponding to F222 of GPPMT, which is proposed to stabilize the cationic intermediate through a cation-□ interaction in the deprotonation of the intermediate,[12b] TleD may facilitate the methylation/cyclization reaction by a similar cation stabilization mechanism. However, we cannot totally exclude a possibility that TleD simply methylates the substrate while the remaining reactions are spontaneous. To further understand the intimate structural details of the enzyme reaction, an X-ray crystallographic study of TleD is now in progress.

In conclusion, we identified a suite of genes involved in teleocidin B biosynthesis. Notably, TleD is the first MT capable of triggering terpene cyclization. Similar to the Class I TPCs,[18] that generate a carbocation by ionizing the allylic diphosphate, TleD also generates a carbocation that electrophilically adds to the aromatic ring and then undergoes a subsequent alkyl migration. However, the TleD-catalyzed reaction differs from those of the Class I TPCs, in that methylation initiates cyclization. Our study illustrates the first representative case of a new mode of terpene cyclization, thus broadening the scope of TPC chemistry.

## ASSOCIATED CONTENT

**Supporting Information**.
Experimental details and spectral data. This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

**Corresponding Author**

abei@mol.f.u-tokyo.ac.jp

**Notes**

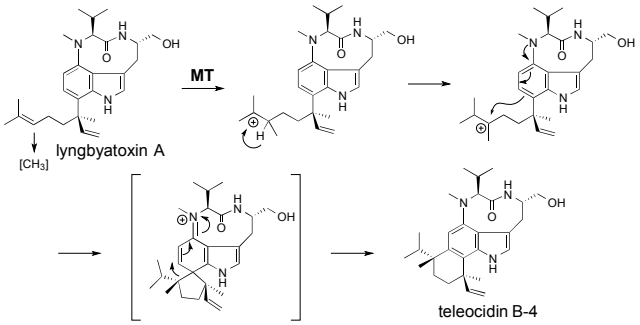
The authors declare no competing interest.

## ACKNOWLEDGMENTS

We thank Prof. H. Onaka (Univ. Tokyo), and Prof. E. Takano (Univ. Manchester) for kindly providing the expression vectors. We also thank Prof. H. Ikeda (Kitasato Univ.) for useful advice to this work. This work was supported by Grants-in-Aid for Scientific Research from the MEXT, Japan.

## REFERENCES

(1) Takashima, M.; Sakai, H. *Bull. Agric. Chem. Soc. Jpn.* **1960**, *24*, 647.
(2) Hitotsuyanagi, Y.; Fujiki, H.; Suganuma, S.; Aimi, N.; Sakai, S.; Endo, Y.; Shudo, K.; Sugimura, T. *Chem. Pharm. Bull*. **1984**, *32*, 4233.
(3) Fujiki, H.; Tanaka, Y.; Miyake, R.; Kikkawa, U.; Nishizuka, Y.; Sugimura, T. *Biochem. Biophys. Res. Commun.* **1984**, *120*, 339.
(4) a) Sakai, S.; Aimi, N.; Yamaguchi, K.; Hitotsuyanagi, Y.; Watanabe, C.; Yokose, K.; Koyama, Y.; Shudo, K.; Itai, A. *Chem. Pharm. Bull*. **1984**, *32*, 354. b) Sakai, S., Hitotsuyanagi, Y.; Aimi, N.; Fujiki, H.; Suganuma, M.; Sugimura, T.; Endo, Y.; Shudo, K. *Tetrahedron Lett.* **1986**, *27*, 5219. c) Sakai, S.; Hitotsuyanagi, Y.; Yamaguchi, K.; Aimi, N.; Ogata, K.; Kuramochi, T.; Seki, H.; Hard, R.; Fujiki, H.; Suganuma, M.; Sugimura, T.; Endo, Y.; Shudo, K.; Koyama, Y. *Chem. Pharm. Bull*. **1986**, *2*, 4883.
(5) a) Muratake, H.; Natsume, M. *Tetrahedron Lett*. **1987**, *28*, 2265. b) Nakatsuka, S.; Masuda, T.; Goto, T. *Tetrahedron Lett.* **1987**, *28*, 3671. c) Dangel B.D.; Godula, K.; Youn, S. W.; Sezen, B.; Sames. D. *J. Am. Chem. Soc*. **2002**, *124*, 11856.
(6) a) Irie, K.; Kajiyama, S.; Funaki, A.; Koshimizu, K.; Hayashi, H.; Arai, M. *Tetrahedron Lett.* **1990**, *31*, 101. b) Irie, K.; Kajiyama, S.; Funaki, A.; Koshimizu, K.; Hayashi, H.; Arai, M. *Tetrahedron* **1990**, *46*, 2773. c) Irie, K.; Nakagawa, Y.; Tomimatsu, S.; Ohigashi, H. *Tetrahedron Lett.* **1998**, *39*, 7929.
(7) Cardellina II, J. H.; Marner, F.-J.; Moore, R. E. *Science* **1979**, *204*, 193.
(8) Edwards, D. J.; Gerwick, W. H. *J. Am. Chem. Soc.* **2004**, *126*, 11432.
(9) Huynh, M. U.; Elston, M. C.; Hernandez, N. M.; Ball, D. B.; Kajiyama, S.; Irie, K.; Gerwick, W. H.; Edwards, D. J. *J. Nat. Prod.* **2009**, *73*, 71.
(10) Zhang, Y.; Buchholz, F.; Muyrers, J. P. P.; Stewart, A. F. *Nat. Genet.* **1998**, *20*, 123.
(11) Nes, W. D.; Jayasimha, P.; Song, Z. *Arch. Biochem. Biophys*. **2008**, *477*, 313.
(12) a) Komatsu, M.; Tsuda, M.; Omura, S.; Oikawa, H.; Ikeda, H. Proc. Natl. Acad. Sci. USA **2008**, 105, 7422-7427. b) Köksal, M.; Chou, W. K. W.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2012**, *51*, 3003.
(13) Kawamura, T.; Matsubara, K.; Otaka, H.; Tashiro, E.; Shindo, K.; Yanagita, R. C.; Irie, K.; Imoto, M. *Bioorg. Med. Chem*. **2011**, *19*, 4377.
(14) Takashima, M.; Sakai, K.; Arima, K. *Agr. Biol. Chem*. **1962**, *26*, 669.
(15) Hitotsuyanagi, Y.; Yamaguchi, K.; Ogata, K.; Aimi, N.; Sakai, S.; Koyama, Y.; Endo, Y.; Shudo, K.; Itai, A.; Iitaka, Y. *Chem. Pharm. Bull*. **1984**, *32*, 3774.
(16) Singh, S.; McCoy, J. G.; Zhang, C.; Bingman, C. A.; Phillips, G. N.; Thorson, J. S. *J. Biol. Chem*. **2008**, *283*, 22628.
(17) Liu, J.; Nakagawa, M.; Hino, T. *Tetrahedron* **1989**, *45*, 7729.
(18) Christianson, D. W. *Chem. Rev*. **2006**, *106*, 3412.


MT
lyngbyatoxin A
[CH3]
teleocidin B-4