OL | Organic Letters

Letter

Cite This: Org. Lett. 2019, 21, 6825−6829

pubs.acs.org/OrgLett

Genome Mining and Enzymatic Total Biosynthesis of Purincyclamide

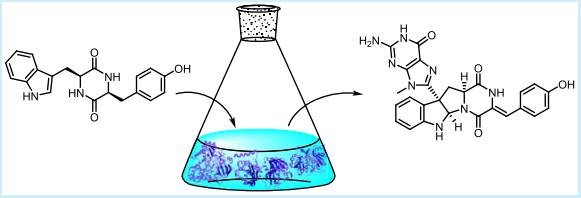
Jing Shi,[†,§] Xiang Xu,[†,§] Er Juan Zhao,[†] Bo Zhang,[†] Wei Li,[†] Yang Zhao,[†] Rui Hua Jiao,*[,†] Ren Xiang Tan,[†,‡] and Hui Ming Ge*[,†]

[†]State Key Laboratory of Pharmaceutical Biotechnology, Institute of Functional Biomolecules, School of Life Sciences, Nanjing University, Nanjing 210023, China

[‡]State Key Laboratory Cultivation Base for TCM Quality and Efficacy, Nanjing University of Chinese Medicine, Nanjing 210023, China

Supporting Information

ABSTRACT: A cyclodipeptide synthase-containing gene cluster (*pcm*) was identified in *Streptomyces chrestomyceticus* NA4264 through genome mining. Heterologous expression of the cryptic *pcm* gene cluster in *Streptomyces albus* led to the production of a new highly modified cyclodipeptide named purincyclamide (**1**). Bioinformatic analysis and gene knockout experiments revealed the biosynthetic pathway. The production of **1** was achieved through a one-pot enzymatic reaction.



Natural cyclodipeptides (CDPs) make up a large class of complex secondary metabolites produced mainly by microorganisms.[1−3] Recently, there is growing interest in CDPs due to their broad bioactivities ranging from antibacterial,[4,5] cytotoxic,[6,7] and antiviral[8] to anti-inflammatory activities.[9] The physiological roles of CDPs are poorly understood, but some of them are thought to be quorum-sensing signals in microorganisms.[10] CDPs contain a common 2,5-diketopiperazine structure core that is generated by either nonribosomal peptide synthases (NRPSs) or aminoacyl tRNA-dependent cyclodipeptide synthases (CDPSs). Unlike the well-documented NRPSs,[11] which utilize free amino acids as direct substrates through adenylation, CDPSs catalyze the formation of two tandem amide bonds in an ATP-independent manner by hijacking aminoacyl tRNAs in the ribosomal biosynthetic machinery.[12,13]

To date, the biosynthetic pathways for several highly modified CDPs have been identified in different Actinobacteria strains, including albonoursin,[14] nocazines,[15] bicyclomycin,[16] 1-(8-guaninyl)-cyclo-L-Trp-L-Trp,[17] pulcherriminic acid,[18] and drimentines[19] (Figure 1). Genes encoding CDPSs are often clustered with other tailoring enzymes, including cytochrome P450s, cyclodipeptide oxidases (CDOs), prenyltransferases, and methyltransferases,[1] which further expand the structural diversity and complexity of CDP-derived compounds. Recent advances in microbial genome sequencing indicate that the biosynthetic function for CDPs is underexplored. In this study, we report the genome mining of a CDP gene cluster and activation of this cryptic gene cluster through heterologous expression of the intact cluster in a heterologous host, as well as the enzymatic total biosynthesis of the final product *in vitro*.

During our natural product discovery project using a genome mining strategy,[20−22] we identified a CDPS-encoding gene locus (named *pcmA*) from the genome of an insect-derived *Streptomyces chrestomyceticus* (strain NA4264). Adjacent to the CDPS gene are a P450, a methyltransferase, and a pair of CDOs (Figure 2). As the organization of these gene loci has not been reported previously, we therefore speculated that this gene cluster may encode a new modified CDP.

Our initial attempts to isolate the corresponding CDP from *S. chrestomyceticus* NA4264 were not successful, probably because the gene cluster is inactive under our fermentation conditions. To activate the cryptic gene cluster, we constructed a genomic library of *Streptomyces* sp. NA4264 using a *Streptomyces* integrative cosmid vector pJTU2554.[22] Specific primers were designed as probes to screen for the *pcm* gene cluster in the library, which resulted in the identification of three positive clones bearing the targeting sequence. One of the isolated cosmids, pHG7001, covering the intact *pcm* gene cluster was identified by end sequencing.

Cosmid pHG7001 was then transferred and integrated into the chromosome of *S. albus* J1074 by conjugation, yielding the *S. albus*/pHG7001 strain. The empty vector was also transferred into *S. albus* as a negative control. These two recombinant strains were then fermented, and the metabolic extracts were then analyzed by LC−MS. A new peak with a molecular ion at *m*/*z* 511.2 ([M + H]$^+$) was detected in *S. albus*/pHG7001, in contrast to the control strain with the empty vector (Figure 3). After large-scale fermentation, compound **1** was isolated, purified, and subjected to spectroscopic analyses.

The molecular formula of **1** was determined to be $C_{26}H_{22}N_8O_4$ according to its HRSEIMS ion. The ^{1}H NMR

Received: July 15, 2019
Published: August 13, 2019

DOI: 10.1021/acs.orglett.9b02461
Org. Lett. 2019, 21, 6825−6829

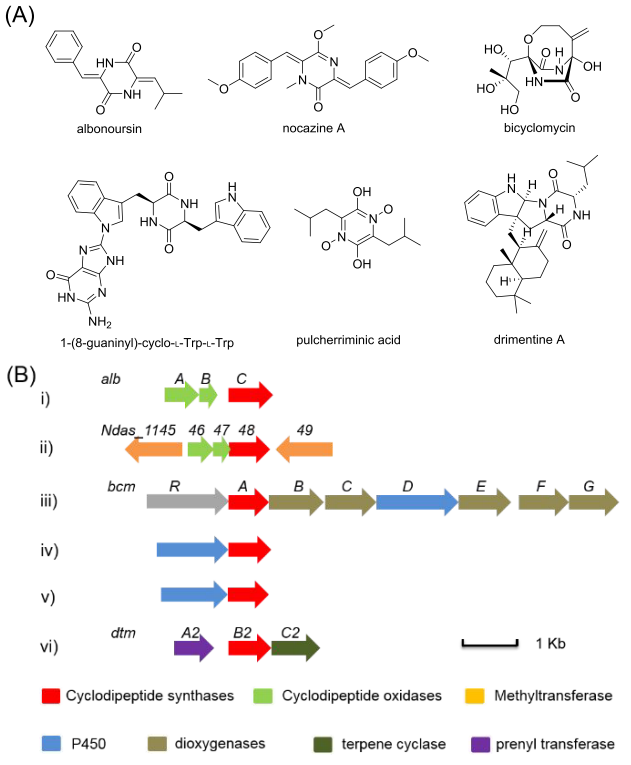


Figure 1. Representative examples of CDPs and their corresponding gene clusters. (A) Structures of six CDPs. (B) Biosynthetic gene clusters for (i) albonoursin, (ii) nocazine A, (iii) bicyclomycin, (iv) 1-(8-guaninyl)-cyclo-L-Trp-L-Trp, (v) pulcherriminic acid, and (vi) drimentine A.

and $^{1}H-^{1}H$ COSY data displayed two sets of aromatic proton spin systems corresponding to the presence of a 1,2-disubstituted phenyl ring (δ_H 6.83, 6.63, 7.10, and 6.66) and a 1,4-disubstituted phenyl ring (δ_H 7.43 and 6.79) (Figure 4 and Table S4). The HMBC correlations of an olefinic proton

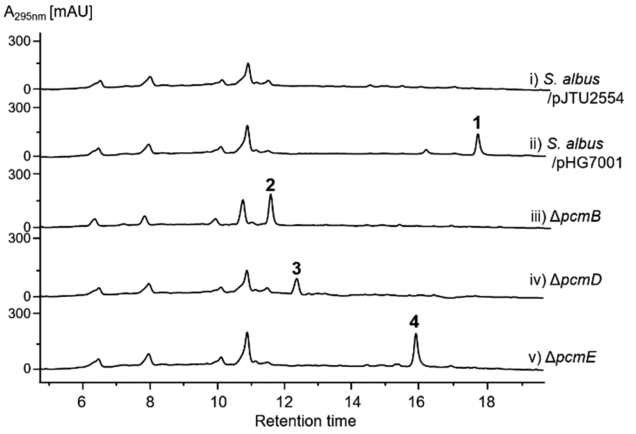


Figure 3. LC−MS analysis of metabolic extracts from *Streptomyces albus* carrying plasmids pJTU2554 and pHG7001 as well as gene mutant cosmids.

(δ_H 6.68) with C-14 (δ_C 161.2) and C-19/C-23 (δ_C 131.8), together with a correlation of phenolic OH (δ_H 9.80) with C-22/C-20 (δ_C 115.9), indicated the presence of a desaturated tyrosyl group. The HMBC correlations of H-3 (δ_H 3.80 and 2.01) with C-1 (δ_C 167.2) and C-5 (δ_C 130.1) and of H-12 (δ_H 5.91) with C-3 (δ_C 40.5) and C-10 (δ_C 149.3) established the diketopiperazine skeleton. In addition, five quaternary carbons (δ_C 157.0, 154.2, 154.0, 146.6, and 115.0) were assigned to those of a guanine group.[17] The HMBC correlation of H-12 with C-1′ indicated the connection of the guanine group with diketopiperazine via C-1′−C-4. Finally, the HMBC correlations of the remaining methyl group (δ_H 3.02) with C-1′ and C-9′ completed the planar structure of **1**. The clear NOE correlations of H-12/H-11′ and H-12/H-2 as well as NH-16/H-23(19) established the sterochemistry of **1**. As both amino acid building blocks are of the L type,[1] we further assigned the absolute configurations of **1** as shown in Figure 4.

With the structure of **1** in hand, we intended to clarify the functions of the associated proteins encoded by the *pcm* gene

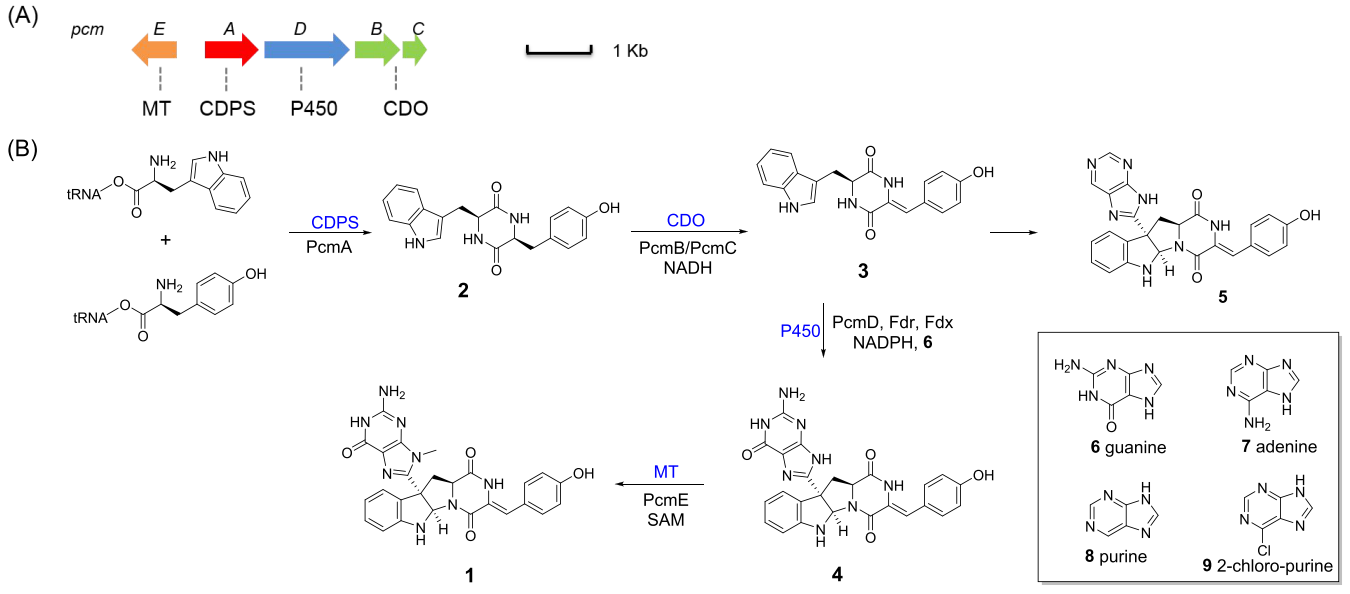


Figure 2. Biosynthesis of purincyclamide (**1**). (A) *pcm* gene cluster from *Streptomyces chrestomyceticus* NA4264 (GenBank accession number MN190176). (B) Proposed biosynthetic pathway for **1**.

DOI: 10.1021/acs.orglett.9b02461
Org. Lett. 2019, 21, 6825−6829

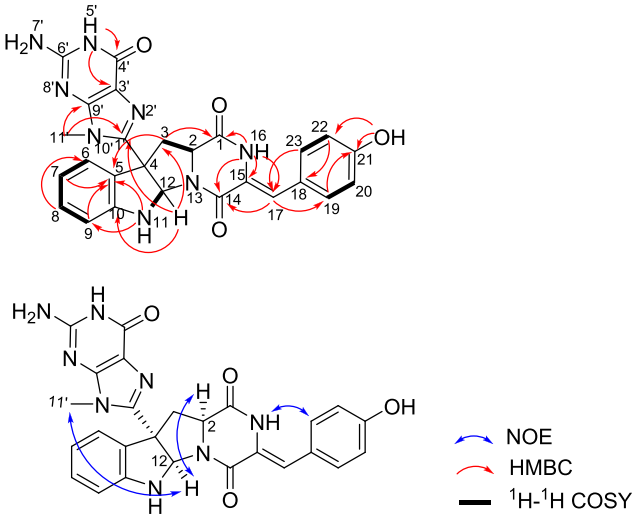


Figure 4. Key two-dimensional nuclear magnetic resonance correlations of **1**.

cluster. Thus, we systematically inactivated *pcmB*, *pcmD*, and *pcmE* genes through λ-Red-mediated PCR-targeting mutagenesis in cosmid pHG7001.[23] The resulting cosmid mutants were then reintroduced into *S. albus* J1074, yielding mutant strains Δ*pcmB*, Δ*pcmD*, and Δ*pcmE*, respectively. Each of these mutant strains, together with *S. albus*/pHG7001 as a control, was then fermented to investigate the effects of these genes on **1** biosynthesis.

As shown in Figure 3, relative to *S. albus*/pHG7001, all mutant strains completely abolished the production of **1**, confirming these genes, *pcmB*,[14] *pcmD*, and *pcmE*, are involved in the biosynthesis of **1**. Instead, compound **2** clearly accumulated in the Δ*pcmB* mutant strain. Complete MS, ^{1}H NMR, and ^{13}C NMR analyses established the structure of **2** as cyclo-L-Trp-L-Tyr,[24] indicating that the CDPS enzyme, PcmA, accepts L-Trp-tRNA and L-Tyr-tRNA as substrates to form **2**. In addition, compound **3**, which was obtained from the Δ*pcmD* mutant strain, has a molecular formula of $C_{20}H_{17}N_3O_3$, 2 Da less than that of **2**. From the ^{1}H and ^{13}C NMR data of **3**, the signals related to protons at α- and β-carbons of the tyrosine part have disappeared. Instead, a new olefinic proton signal (δ_H 6.14) was observed, indicating that the tyrosine became desaturated. The geometry of the double bond was determined to be *Z* according to the clear NOE correlation of NH-16 (δ_H 9.21) with H-19(23) (δ_H 6.49) (Table S5), consistent with that in **1**. Further explanation of its HSQC, HMBC, and ^{1}H−^{1}H COSY confirmed the structure of **3** as shown in Figure 2.

Compound **4**, which had accumulated in the Δ*pcmE* mutant strain (Figure 3),[14] has a molecular formula of $C_{25}H_{20}N_8O_4$ on the basis of the HRESIMS data. Therefore, **4** is 14 Da less than that of **1**. The ^{1}H and ^{13}C NMR data of **4** are almost identical to those of **1** but revealed the disappearance of the signal from an *N*-methyl group (δ_C 29.7 and δ_H 3.02) (Tables S4 and S6). This assignment is consistent with the predicted function of PcmE as a methyltransferase.

The experimental results presented above led us to propose the biosynthetic pathway for **1** as shown in Figure 2. The CDPS (PcmA) generates two successive amide bonds by using L-Trp-tRNA and L-Tyr-tRNA as substrates, resulting in the formation of **2** as the sole product. Two enzymes encoded by *pcmB* and *pcmC* that were 51% and 49% identical to AlbA and AlbB (Table S1),[14] respectively, were annotated as two subunits of the CDO complex. PcmB and PcmC are thus responsible for introducing the double bond only into the tyrosyl part in **2**, although AlbA and AlbB or other known CDO homologues can successively generate two double bonds in two amino acid residues.[25] The P450 enzyme (PcmD) is intriguing and may catalyze intermolecular and intramolecular C−C bond formation through a radical cascade manner to afford **4**.[26,27] Finally, the methyltransferase (PcmE) transfers a methyl group to the guaninyl residue in **4** to afford **1**.

To gain a better understanding of the roles of individual enzymes in **1** biosynthesis, we intended to reconstitute the whole biosynthetic pathway *in vitro*. First, we placed the *pcmA* gene under the control of the inducible T7 promotor in the pET22b(+) vector and then introduced it into *Escherichia coli* BL21(DE3). After induction with 0.1 mM IPTG at 16 °C for 18 h, the culture was centrifuged and the supernatant was extracted with ethyl acetate. LC−MS analysis of the organic extract showed a sole product with an expected molecular weight and an expected retention time like those for **2**. Two CDO genes, *pcmB* and *pcmC*, were inserted into the multiple cloning sites of the pACYC-Duet-1 vector. Co-expression of these two genes in *E. coli* BL21(DE3) was achieved and led to the production of bright yellow proteins. SDS−PAGE analysis showed two bands with molecular weights of 23 and 12 kDa, which are in agreement with the expected molecular weights for PcmB and PcmC, respectively (Figure S2). Incubation of the CDO complex with **2** in the presence of 2 mM NADH led to the production of a singly dehydrogenated compound **3**. A longer incubation time did not yield the doubly dehydrogenated product, indicating the PcmB−PcmC complex desaturated only the tyrosyl part, consistent with the proposed pathway.

The P450-encoding gene (*pcmD*) was cloned into pET28a-(+) and overexpressed in *E. coli*. The P450s from *Streptomyces* usually belong to a class I redox system, which require ferredoxin reductase (Fdr) and ferredoxin (Fdx) for their activities. A new peak identical to **4** was observed via LC−MS upon incubation of PcmD, NADPH, guanine,[17] and **3**, as well as Fdr and Fdx from *Kitasatospora* sp. CGMCC 16924.[21] In contrast, no product formation was observed in the negative control in which active PcmD was replaced with the boiled enzyme. To evaluate the substrate specificity of PcmD, **3** was incubated with different purine derivatives in the aforementioned enzymatic reaction condition. On the basis of the LC−MS analysis (Figure 5), purine (**8**) can be also accepted by PcmD (Figure 5 and Figure S3), while PcmD was inactive toward adenine (**7**) or 2-chloro-purine (**9**). It is worth noting that compound **5** or its methylated derivative cannot be detected in the *S. albus*/pHG7001 strain via LC−MS.

To complete the biosynthetic pathway, we expressed the remaining methyltransferase (PcmE) as a soluble His_6-tagged protein in *E. coli*. PcmE was incubated with compound **4** and *S*-adenosylmethionine (SAM). Gratifyingly, the final product **1** of the biosynthetic pathway was observed during LC−MS analysis. The control reaction without SAM did not produce any new peaks (Figure 5). Thus, our stepwise *in vitro* biochemical experiments supported the proposed biosynthetic pathway of **1** and confirmed the functions of individual enzymes. Finally, to examine whether we can achieve the production of **1** through a one-pot enzymatic reaction, six enzymes, PcmB, PcmC, PcmD, PcmE, and a pair of Fdr and

DOI: 10.1021/acs.orglett.9b02461
Org. Lett. 2019, 21, 6825−6829

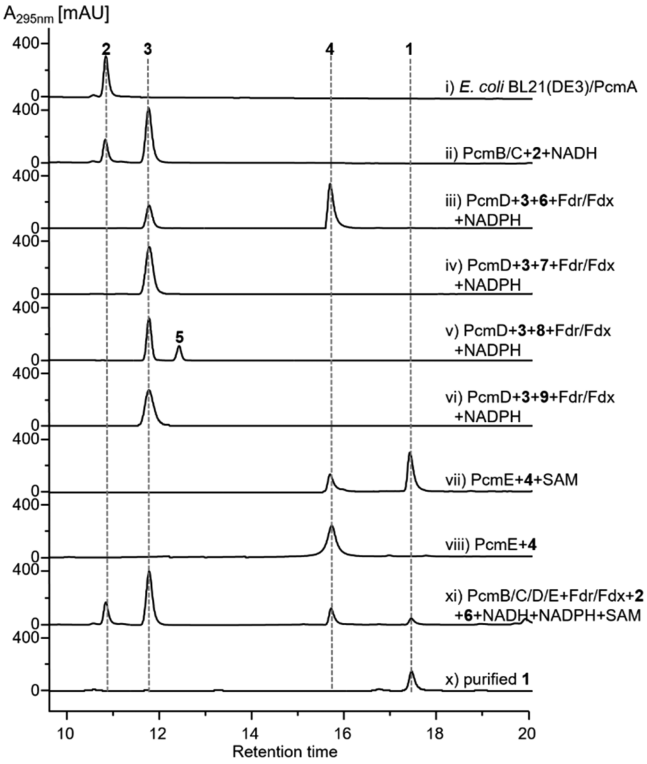


Figure 5. *In vitro* enzymatic reactions in **1** biosynthesis.

Fdx, were incubated with NADPH, NADH, SAM, and substrate **2** in HEPES buffer (pH 8.2) for 2 h. LC−MS analysis of enzymatic reaction mixtures indicated the presence of **1** together with biosynthetic intermediates **2**−**4**. The identity of **1** together with other biosynthetic intermediates was verified with those of authentic compounds. The time course of the experiment (Figure S4) indicated that the P450 enzyme (PcmD) could be the bottleneck of this multiple-enzyme reaction.

In conclusion, a highly modified cyclodipeptide, purincyclamide (**1**), was isolated and characterized by genome mining in conjunction with the heterologous expression of the intact gene cluster. Systematic inactivation of *pcmB*, *pcmD*, and *pcmE* genes enabled us to isolate **2**−**4** as biosynthetic intermediates, establishing a biosynthetic pathway for **1**. The stepwise *in vitro* experiments further supported the proposed pathway. An atypical P450 enzyme (PcmD) was identified in **1** biosynthesis that catalyzes an intramolecular C−C bond formation as well as an unusual intermolecular coupling of guanine with cyclo-L-Trp-L-ΔTyr (**3**) (Figure S5).[26] Finally, the production of **1** was successfully achieved through a one-pot enzymatic reaction using cyclo-L-Trp-L-Tyr (**2**) as a starting material. This work sets the stage for the production of new cyclodipeptide analogues by combinatorial biosynthesis or enzymatic total synthesis.[28]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.orglett.9b02461.

Detailed experimental procedures and NMR spectra of **1**−**4** (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: hmge@nju.edu.cn.
*E-mail: rhjiao@nju.edu.cn.

ORCID

Ren Xiang Tan: 0000-0001-6532-6261
Hui Ming Ge: 0000-0002-0468-808X

Author Contributions

§J.S. and X.X. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was financially supported by NSFC (21572100, 81773591, 21861142005, 81530089, 81673333, 81803380, 21761142001, and 21661140001), MOST (2018YFC1706200 and 2018YFA0902000), and the Fundamental Research Funds for the Central Universities (020814380092 and 020814380103).

■ REFERENCES

(1) Belin, P.; Moutiez, M.; Lautru, S.; Seguin, J.; Pernodet, J. L.; Gondry, M. *Nat. Prod. Rep.* **2012**, *29*, 961−979.
(2) Gondry, M.; Sauguet, L.; Belin, P.; Thai, R.; Amouroux, R.; Tellier, C.; Tuphile, K.; Jacquet, M.; Braud, S.; Courcon, M.; Masson, C.; Dubois, S.; Lautru, S.; Lecoq, A.; Hashimoto, S.-i.; Genet, R.; Pernodet, J.-L. *Nat. Chem. Biol.* **2009**, *5*, 414−420.
(3) Borthwick, A. D. *Chem. Rev.* **2012**, *112*, 3641−3716.
(4) Clark, B.; Capon, R. J.; Lacey, E.; Tennant, S.; Gill, J. H. *J. Nat. Prod.* **2005**, *68*, 1661−1664.
(5) Cain, C. C.; Lee, D. H.; Waldo, R. H.; Henry, A. T.; Casida, E. J.; Wani, M. C.; Wall, M. E.; Oberlies, N. H.; Falkinham, J. O. *Antimicrob. Agents Chemother.* **2003**, *47*, 2113−2117.
(6) Park, H. B.; Kim, Y. J.; Park, J. S.; Yang, H. O.; Lee, K. R.; Kwon, H. C. *J. Nat. Prod.* **2011**, *74*, 2309−2312.
(7) Meng, L. H.; Wang, C. Y.; Mandi, A.; Li, X. M.; Hu, X. Y.; Kassack, M. U.; Kurtan, T.; Wang, B. G. *Org. Lett.* **2016**, *18*, 5304−5307.
(8) Rodriguez, P. L.; Carrasco, L. *J. Virol.* **1992**, *66*, 1971−1976.
(9) Shimazaki, N.; Shima, I.; Hemmi, K.; Hashimoto, M. *J. Med. Chem.* **1987**, *30*, 1706−1709.
(10) Park, D. K.; Lee, K. E.; Baek, C. H.; Kim, I. H.; Kwon, J. H.; Lee, W. K.; Lee, K. H.; Kim, B. S.; Choi, S. H.; Kim, K. S. *J. Bacteriol.* **2006**, *188*, 2214−2221.
(11) Walsh, C. T. *Nat. Prod. Rep.* **2016**, *33*, 127−135.
(12) Moutiez, M.; Schmitt, E.; Seguin, J.; Thai, R.; Favry, E.; Belin, P.; Mechulam, Y.; Gondry, M. *Nat. Commun.* **2014**, *5*, 5141.
(13) Jacques, I. B.; Moutiez, M.; Witwinowski, J.; Darbon, E.; Martel, C.; Seguin, J.; Favry, E.; Thai, R.; Lecoq, A.; Dubois, S.; Pernodet, J. L.; Gondry, M.; Belin, P. *Nat. Chem. Biol.* **2015**, *11*, 721−727.
(14) Lautru, S.; Gondry, M.; Genet, R.; Pernodet, J. L. *Chem. Biol.* **2002**, *9*, 1355−1364.
(15) Giessen, T. W.; von Tesmar, A. M.; Marahiel, M. A. *Chem. Biol.* **2013**, *20*, 828−38.
(16) Meng, S.; Han, W.; Zhao, J.; Jian, X. H.; Pan, H. X.; Tang, G. L. *Angew. Chem., Int. Ed.* **2018**, *57*, 719−723.
(17) Yu, H.; Xie, X.; Li, S. M. *Org. Lett.* **2018**, *20*, 4921−4925.
(18) Cryle, M. J.; Bell, S. G.; Schlichting, I. *Biochemistry* **2010**, *49*, 7282−96.
(19) Yao, T.; Liu, J.; Liu, Z.; Li, T.; Li, H.; Che, Q.; Zhu, T.; Li, D.; Gu, Q.; Li, W. *Nat. Commun.* **2018**, *9*, 4091.

DOI: 10.1021/acs.orglett.9b02461
Org. Lett. 2019, 21, 6825−6829

(20) Zhang, X.; Wang, T. T.; Xu, Q. L.; Xiong, Y.; Zhang, L.; Han, H.; Xu, K.; Guo, W. J.; Xu, Q.; Tan, R. X.; Ge, H. M. *Angew. Chem., Int. Ed.* **2018**, *57*, 8184−8188.
(21) Shi, J.; Liu, C. L.; Zhang, B.; Guo, W. J.; Zhu, J. P.; Chang, C. Y.; Zhao, E. J.; Jiao, R. H.; Tan, R. X.; Ge, H. M. *Chem. Sci.* **2019**, *10*, 4839−4846.
(22) Shi, J.; Zeng, Y. J.; Zhang, B.; Shao, F. L.; Chen, Y. C.; Xu, X.; Sun, Y.; Xu, Q.; Tan, R. X.; Ge, H. M. *Chem. Sci.* **2019**, *10*, 3042−3048.
(23) Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 1541−1546.
(24) Zhang, Y.; Li, X. M.; Feng, Y.; Wang, B. G. *Nat. Prod. Res.* **2010**, *24*, 1036−1043.
(25) Gondry, M.; Lautru, S.; Fusai, G.; Meunier, G.; Menez, A.; Genet, R. *Eur. J. Biochem.* **2001**, *268*, 1712−1721.
(26) Tian, W. Y.; Sun, C. H.; Zheng, M.; Harmer, J. R.; Yu, M. J.; Zhang, Y. A.; Peng, H. D.; Zhu, D. Q.; Deng, Z. X.; Chen, S. L.; Mobli, M.; Jia, X. Y.; Qu, X. D. *Nat. Commun.* **2018**, *9*, 4428.
(27) Rudolf, J. D.; Chang, C. Y.; Ma, M.; Shen, B. *Nat. Prod. Rep.* **2017**, *34*, 1141−1172.
(28) During the submission of this manuscript, Li and co-workers reported a complementary way of discovery of guanitrypmycin from Streptomyces monomycini NRRL B-24309, see:Liu, J.; Xie, X.; Li, S.-M. *Angew. Chem. Int. Ed.* **2019**, DOI: 10.1002/anie.201906891.

DOI: 10.1021/acs.orglett.9b02461
Org. Lett. 2019, 21, 6825−6829