The Isolation of Pyrroloformamide Congeners and Characterization of Its Biosynthetic Gene Cluster from *Streptomyces* sp. CB02980 Revealed a Unified Mechanism for Dithiolopyrrolone Biosynthesis

Wenqing Zhou[†,#], Haoyu Liang[†,#], Xiangjing Qin[‡], Danfeng Cao[†],

Xiangcheng Zhu[†,§,¶], Jianhua Ju[‡], Ben Shen [Δ,∥,&], Yanwen Duan[†, §,⊥,¶ *],

Yong Huang[†,⊥,¶ *]

[†]Xiangya International Academy of Translational Medicine at Central South University, Changsha, Hunan 410013, China

[‡]CAS Key Laboratory of Tropical Marine Bio-resources and Ecology, Guangdong Key Laboratory of Marine Materia, RNAM Center for Marine Microbiology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China

[§]Hunan Engineering Research Center of Combinatorial Biosynthesis and Natural Product Drug Discover, Changsha, Hunan 410011, China

[⊥]National Engineering Research Center of Combinatorial Biosynthesis for Drug Discovery, Changsha, Hunan 410011, China

[Δ] Departments of Chemistry, [∥] Molecular Medicine, [&]Natural Products Library Initiative, Jupiter, Florida 33458, United States

[#] These authors contributed equally

Abstract

Dithiolopyrrolones are microbial natural products containing a disulfide or thiosulfonate bridge embedded in a unique bicyclic structure. By interfering with zinc ion homeostasis in living cells, they showed strong antibacterial activity against a variety of bacterial pathogens, as well as potent cytotoxicity against human cancer cells. In the current study, two new dithiolopyrrolones, pyrroloformamide C (**3**) and pyrroloformamide D (**4**), were isolated from *Streptomyces* sp. CB02980, together with the known pyrroloformamides **1** and **2**. The biosynthetic gene cluster for pyrroloformamides was identified from *S.* sp. CB02980, which shared high sequence similarity with those of dithiolopyrrolones, including holomycin and thiolutin. Gene replacement of *pyfE*, which encodes a non-ribosomal peptide synthetase, abolished the production of **1**－**4**. Overexpression of *pyfN*, a type II thioesterase gene, increased the production of **1** and **2**. Genome neighborhood network analysis of the characterized and orphan gene clusters of dithiolopyrrolones revealed a unified mechanism for their biosynthesis. The structure elucidation and biosynthetic characterization of pyrroloformamides **1-4** may inspire future efforts to discover new dithiolopyrrolones, which are promising drug leads for the treatment of infectious diseases or cancer.

Keywords: dithiolopyrrolones; pyrroloformamide; *Streptomyces*

Introduction

Sulfur is the 10th most abundant element in the Universe, which is essential to all life forms found on Earth, and has been proposed to be metabolized by ancient microbes on Mars.[1] Dithiolopyrrolones are sulfur-containing natural products, such as holomycin,[2] thiolutin,[3] xenorhabdins[4] and thiomarinols,[5-7] characterized by the presence of a disulfide or thiosulfonate bridge embedded in a unique bicyclic scaffold (Figure 1A). The dithiolopyrrolones have been identified from diverse bacteria species, such as soil-dwelling *Streptomyces*, or Gram-negative bacteria, including *Xenohabdus*, *Alteromonas*, *Yersinia* and *Photobacterium*, which suggests their important chemical ecological roles for the producing organisms.[4,5,8,9]
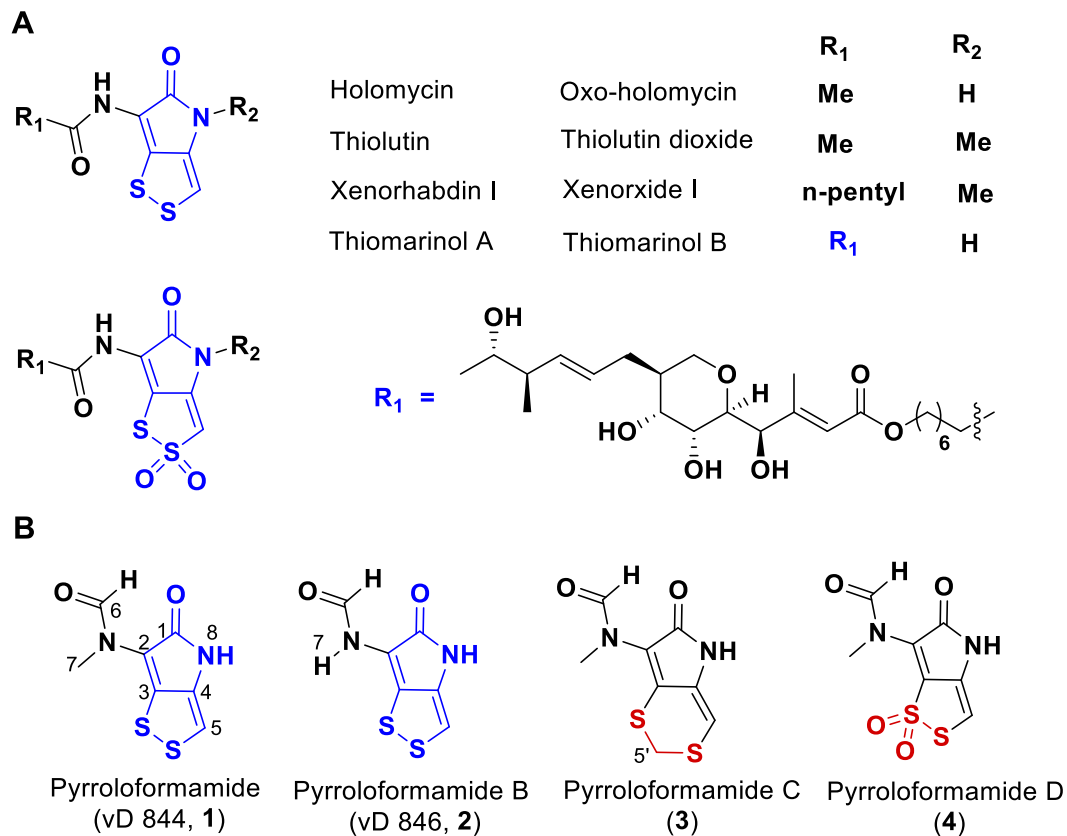
A

		R_1	R_2
Holomycin	Oxo-holomycin	Me	H
Thiolutin	Thiolutin dioxide	Me	Me
Xenorhabdin I	Xenorxide I	n-pentyl	Me
Thiomarinol A	Thiomarinol B	R_1	H

B

Pyrroloformamide (vD 844, **1**) Pyrroloformamide B (vD 846, **2**) Pyrroloformamide C (**3**) Pyrroloformamide D (**4**)

Figure 1. Structures of selected dithiolopyrrolone natural products. (A) The

structures of N-acylpyrrothine or N-methyl-N-acylpyrrothine dithiolopyrrolones. (B) The structures of pyrroloformamides, the N-formylated dithiolopyrrolones from *Streptomyces sp.* CB02980, in which **1** and **2** were previously reported.

Both holomycin and thiolutin are broad-spectrum antibiotics against many Gram-positive and Gram-negative pathogens.[2-7, 10] They were also shown to be cytotoxic against the tested human cancer cell lines. Due to their promising biological activities and unprecedented molecular structures, the mode of action of holomycin and thiolutin was recently revealed to be metal-chelating agents in *Escherichia coli*, fungi or human cells. The reduced holomycin was able to chelate Zn^{2+} ion of the essential *E. coli* class II fructose bisphosphate aldolase, and the metallo-β-lactamases responsible for clinical carbapenem resistance,[11] while reduced thiolutin was also a Zn^{2+} ion chelator for bacterial metalloproteinase thermolysin in vitro, as well as AB1/MPN/Mov34 (JAMM) domain-containing metalloprotease Rpn11 and other JAMM metalloproteases in fungi and mammals.[12] The inhibition of Rpn11, a de-ubiquinating enzyme of the 19S proteasome, or other JAMM proteases, may disrupt protein or genome homeostasis of cancer cells, which are critically dependent on the ubiquitin-proteasome systems.[13] The elucidation of these fundamental mechanisms for the observed inhibitory activity of holomycin and thiolutin is significant and would be instrumental for the future drug development of this family of metal-chelating natural products. In addition, thiomarinols showed

superior antibacterial activity, which consists of the clinically-used topical antibiotic mupirocin and a dithiolopyrrolone moiety, connected by an aliphatic linker. It suggested a natural hybridity for antibiotic development, which was further elaborated by Thomas, Li and their co-workers in the mechanistic characterization of a unique TmlU/HolE enzyme pair and the mutasynthetic preparation of dozens of thiomarinol analogues.[14,15]

Pyrroloformamides **1** and **2** (previously named vD 844 and vD 846) were unique among the isolated dithiolopyrrolones, due to the modification of their amino terminal by N-formylation (Figure 1B).[16,17] The additional N-methyl group on **1** results the presence of two rotamers in ambient temperature. Both **1** and **2** were potent antibiotics against methicillin-resistant *Staphylococcus aureus* and tested Gram-negative pathogens. Pyrroloformamide (**1**) acted as a cytokinesis modulator to interrupt cell cycle, and was cytotoxic against metastatic prostate cancer cells with an IC_{50} of 1.67 μM. Interestingly, unlike other dithiolopyrrolones, no new analogue of pyrroloformamides was reported since their initial discovery in 1969.[16-18] In addition, to our knowledge, the biosynthetic gene cluster of pyrroloformamides has not been reported, despite the recent strides for the biosynthesis, heterologous expression or mutasynthesis of holomycin, thiolutin or thiomarinols.[19-24]

We have been interested in discovering new natural product drug leads from un- or under-explored environmental niches, and also recently identified disulfide-containing guangnanmycins through a genome mining approach.[25]

During our continuing search for new natural products from actinomycete, we came across *S.* sp. CB02980 which isolated from Yandang Mountain in Eastern China. This strain produced pyrroloformamides **1** and **2**, along with two new congeners pyrroloformamide C (**3**) and pyrroloformamide D (**4**) (Figure 1B). In this study, we report the isolation and structure characterization of these two new pyrroloformamides. In addition, the identification and preliminary characterization of the biosynthetic gene cluster for pyrroloformamides in *S.* sp. CB02980 revealed a unified mechanism for dithiolopyrrolone biosynthesis. Our study should inspire the discovery of new dithiolopyrrolones, especially pyrroloformamides from bacteria, which hold great promise against infectious diseases or cancer.

Results and Discussion

Fermentation，Isolation and Structure Elucidation of Pyrroloformamide C (3) and D (4).

During the activity-based metabolite profiling of the collected actinomycete strains, the fermentation broth of *S.* sp. CB02980 showed strong inhibitory activity against the tested *S. aureus* ATCC 29213 and *E. coli*. Further fermentation of *S.* sp. CB02980 led to the isolation of pyrroloformamide (**1**) (Table S1 and Figure S1-S8). Since **1** shares the common disulfide-containing bicyclic core with other dithiolopyrrolones, the addition of L-cysteine to the production medium of *S.* sp. CB02980 may increase the yield of **1**. Various concentrations of L-cysteine of 5, 10, 20 and 40 mM were added into its production medium in 250-mL shaking flasks. The yield of **1** reached to 32.6 ± 6.2 mg/L when 10 mM of L-cysteine was used, which was about 3-fold improvement over the original medium (Figures S9 and S10). It was consistent with the previous report that addition of L-cysteine may increase the titers of thiolutin and the related dithiolopyrrolone antibiotics.[26] A large-scale fermentation (12 L) using this optimized medium led to the isolation of compounds **1**–**3** with the addition of macroporous resins, while **4** was isolated from the fermentation culture (4 L) without resins. Compounds **2**–**4** were close analogues of pyrroloformamide (**1**), in which **2** was the previously isolated vD 846 (renamed as pyrroloformamide B in this study) (Table S2 and Figures S11-S16), while pyrroloformamide C (**3**) and pyrroloformamide D (**4**)
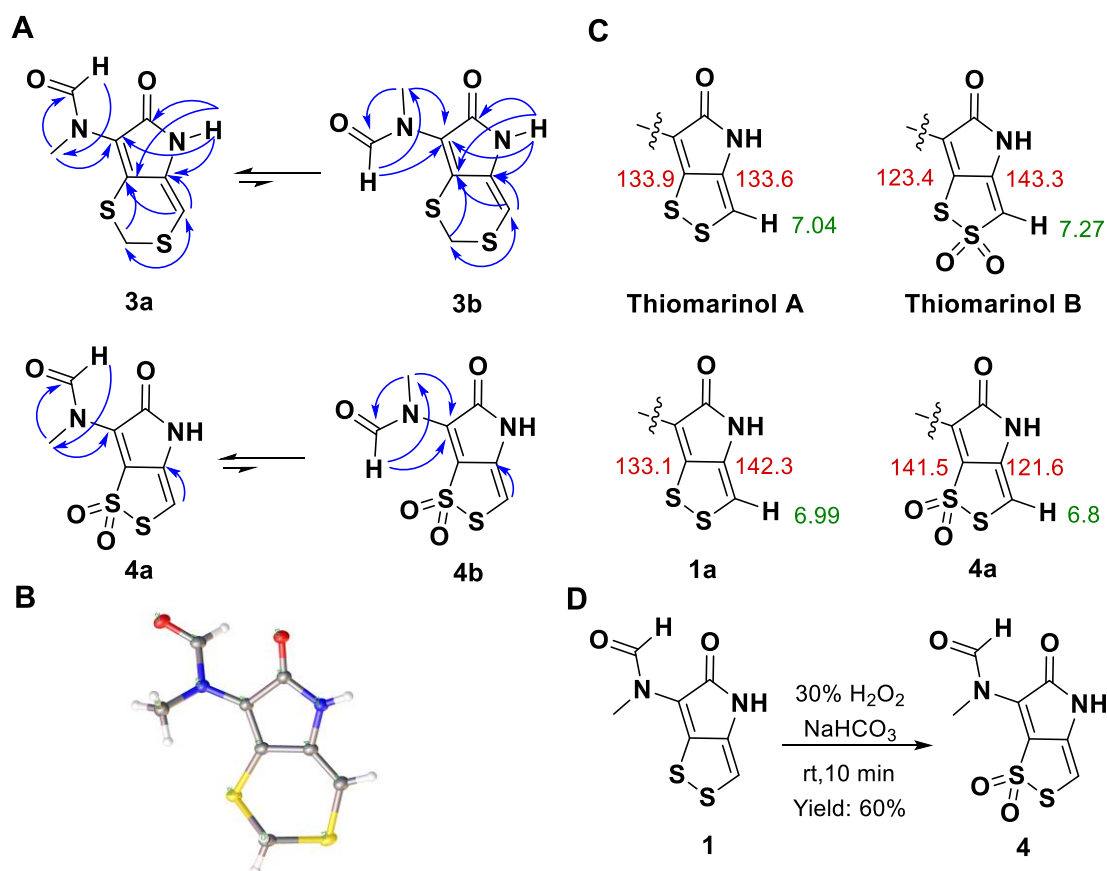
were new congeners (Figure 2).

Figure 2. Structure elucidation of **3** and **4**. (A) HMBC (⌒→) correlations of **3** and **4**. (B) The X-ray crystallographic study of **3**. (C) The chemical shifts of holothin part of thiomarinol A, thiomarinol B, **1** and **4**. (D) Semisynthesis of **4** from **1**.

Pyrroloformamide C (**3**) was obtained as light-yellow needle crystals. Its molecular formula $C_8H_8N_2O_2S_2$, was established upon analysis of the HR-ESI-MS peak at m/z 229.0095 $[M + H]^+$ (calculated for $[M + H]^+$ ion at m/z 229.0105), differing from the molecular formula of **1** by one methylene unit (Figure S18). The ^1H and ^{13}C-NMR spectra indicated that **3** existed as a

mixture of two rotamers **3a** and **3b**, similar to compound **1** (Table S3). The difference was the presence of a methylene singlet in pyrroloformamide C (**3**) [δ_H 4.40 (2H, s), δ_C 30.4 (CH$_2$, C-5′) for **3a** and δ_H 4.28 (2H, s), δ_C 30.3 (CH$_2$, C-5′) for **3b**]. It indicated that the methylene group was embedded in sulfur-containing groups, by comparing with other 1,3-dithiolane containing natural products, including bis (methylthiomethyl) disulfide and another dithiolopyrrolone derivative, holomycin A.[27] The assignment was supported by the HMBC correlations from H-5 to C-5′ and H-5′ to C-3, C-5 for **3a**, and H-5 to C-5′ and H-5′ to C-3, C-5 for **3b** (Figure 2A). The complete structure of **3a** and **3b** was assigned based on their ^{1}H and ^{13}C-NMR, HSQC and HMBC spectra (Figures S20-S23). A single-crystal X-ray crystallographic analysis was applied to further confirm the structure of **3**, in which only rotamer **3a** was observed (Figures 2B and S24，Table S4). This was consistent with the presence of a major conformer **3a** (appr. 8:2) in the ^{1}H and the other NMR spectra, suggesting that some steric effects imposed by 1, 3-dithiine moiety in **3** might restrict the dynamic equilibrium between **3a** and **3b**. In contrast, both rotamers (appr. 6:4) were present in the crystal of compound **1**.[17]

Compound **4** was obtained as yellow oil. The molecular formula C$_7$H$_6$N$_2$O$_4$S$_2$, was established upon analysis of the HR-ESI-MS peak at m/z 244.9694 [M - H]$^-$ (calculated for [M - H]$^-$ ion at m/z 244.9691), differing from the molecular formula of **1** by two more oxygen atoms (Figure S26). The ^{1}H, ^{13}C NMR, HSQC and HMBC spectra of **4** were similar to those of **1**, with the presence of

two set of signals (appr. 5:3), suggesting the formation of two rotamers **4a** and **4b** (Figures S28-S31). The only difference was the replacement of the disulfide moiety in **1** by a thiosulfonate unit in **4**, supported by the comparison of chemical shifts among **1a**, **4a**, as well as thiomarinol A and B. Comparing to compound **1a**, the downfield-shifted C-3 (δ_C 141.5), upfield-shifted C-4 (δ_C 121.6) and H-5 (δ_H 6.8) in **4a** suggested that the position of sulfone in **4a** was different from those of thiomarinol B,[28] as well as other thiosulfonate analogues oxo-holomycin, thiolutin dioxide and xenorxides (Figures 1 and 2C).[6,29-31] This assignment was also confirmed by the preparation of **4** through the treatment of **1** with 30% H_2O_2 in aqueous acetone solution with an isolated yield of 60% (Figure 2D).[28] The semisynthetic compound **4** was identical to the natural product **4** in ^1H NMR spectra (Figure S34).

Antibacterial Activities and Cytotoxicity of Compounds 3 and 4

The newly isolated pyrroloformamide congeners **3** and **4,** were first evaluated for their in vitro antibacterial activities against *S. aureus* ATCC 29213 and methicillin-resistant *S. aureus* (MRSA), as well as Gram-negative *E. coli* and *Klebsiella pneumoniae*, using **1, 2** and linezolid as controls (Table 1). Comparing to compounds **1** and **2**, compound **4** had attenuated antibacterial activities against the tested pathogens with its minimal inhibitory concentration (MIC) ranging from 8－64 μg/mL, while compound **3** showed no antibacterial activity against all four tested strains. Compounds **3** and **4** were also not

cytotoxic against non-small cell lung cancer cell line A549 and human epithelial colorectal adenocarcinoma (Caco-2) in the MTT assay.[32] In contrast, pyrroloformamides **1** and **2** had MICs of 0.5 µg/mL or 1−4 µg/mL against the tested pathogens, respectively. Compound **1** also showed strong cytotoxicity against A549 and Caco-2 cells. Since only the reduced holomycin or thiolutin could chelate the essential Zn^{2+} ion from Zn^{2+}-dependent enzymes, the loss of antibacterial activity and cytotoxicity of **3** was likely due to its inability to interfere zinc ion homeostasis in living cells. In contrast, compound **4** still had moderate antibacterial activity against the tested *S. aureus* and *E. coli*, which was consistent with the reported antibacterial activity of many thiosulfonate- or thiosulfinate-containing natural products, such as pseudoallicin and allicin found in garlic,[33] as well as leinamycin from soil bacteria *Streptomyces*, whose unique redox-active and catalytic properties have been previously documented.[34]

Table 1. The antibacterial activity (MICs, µg/mL) and cytotoxicity (IC$_{50}$, µM) of pyrroloformamides.

Bioactivity	Bacteria/Cancer cells	Compounds				
		Linezolid	1	2	3	4
Antibacterial activities	*S. aureus* ATCC 29213	1	0.5	4	> 64	8
	MRSA	1	0.5	2	> 64	8
	E. coli	> 64	0.5	1	> 64	32
	K. pneumoniae	> 64	0.5	1	> 64	64
Cytotoxicity	A549	-	1.34±0.01	-	> 64	> 64
	Caco-2	-	0.57±0.01	-	> 64	> 64

-, not tested.

Pyrroloformamide Biosynthesis

The isolation of pyrroloformamides **1**－**4** in *S.* sp. CB02980 provided an excellent opportunity to study their biosynthesis, and the evolutionary relationships with other dithiolopyrrolone natural products. Shotgun genome sequencing of *S.* sp. CB02980 and the following bioinformatic analysis led to an assumption of candidate pyrroloformamide biosynthetic gene cluster with high sequence similarity with those of holomycin and thiolutin (Figure 3 and Table S9).[35] The pyrroloformamide gene cluster containing essential genes for dithiolopyrrolone core biosynthesis, including the multi-domain containing non-ribosomal peptide synthetase PyfE for the biosynthesis of the core Cys-Cys dipeptide.
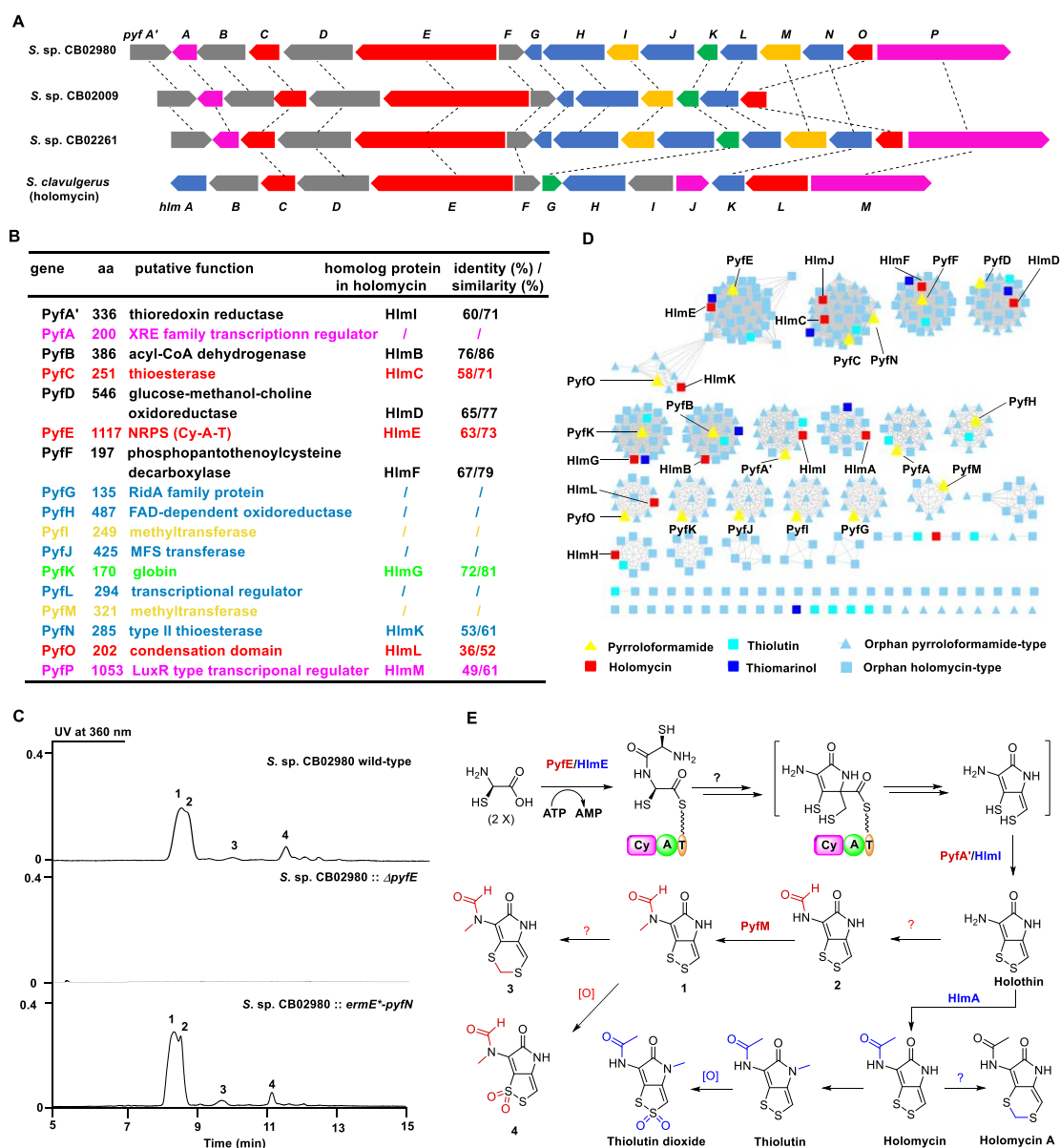
B

gene	aa	putative function	homolog protein in holomycin	identity (%) / similarity (%)
PyfA'	336	thioredoxin reductase	HlmI	60/71
PyfA	200	XRE family transcriptionn regulator	/	/
PyfB	386	acyl-CoA dehydrogenase	HlmB	76/86
PyfC	251	thioesterase	HlmC	58/71
PyfD	546	glucose-methanol-choline oxidoreductase	HlmD	65/77
PyfE	1117	NRPS (Cy-A-T)	HlmE	63/73
PyfF	197	phosphopantothenoylcysteine decarboxylase	HlmF	67/79
PyfG	135	RidA family protein	/	/
PyfH	487	FAD-dependent oxidoreductase	/	/
PyfI	249	methyltransferase	/	/
PyfJ	425	MFS transferase	/	/
PyfK	170	globin	HlmG	72/81
PyfL	294	transcriptional regulator	/	/
PyfM	321	methyltransferase	/	/
PyfN	285	type II thioesterase	HlmK	53/61
PyfO	202	condensation domain	HlmL	36/52
PyfP	1053	LuxR type transcriponal regulator	HlmM	49/61

Figure 3. (**A**) Comparison of the pyrroloformamide gene clusters in S. sp. CB02980, with other similar gene clusters. (**B**) Predicted functions of ORFs in the pyrroloformamide-type gene cluster from *S.* sp. CB02980. (**C**) HPLC analysis of pyrroloformamides production in *S.* sp. CB02980 wild-type strain and mutants**. (D)** GNN analysis of 28 orphan dithiolopyrrolone gene clusters and 4 characterized dithiolopyrrolone gene clusters (pyrroloformamide, holomycin, thiolutin and thiomarinol). ■: orphan holomycin-type; ▲: orphan

pyrroloformamide-type; ■: holomycin; ■: thiolutin; ■: thiomarinols; ▲: pyrroloformamide in *S.* sp. CB02980. (**E**) The unified dithiolopyrrolone biosynthetic pathways.

To study if the *pyfE* was responsible for the biosynthesis of pyrroloformamides in *S.* sp. CB02980, we inactivated *pyfE* by replacing it with a mutant copy in which *pyfE* was disrupted by the thiostrepton-resistance gene with a *k*asOp* promoter.[36] The gene replacement of *pyfE* completely abolished the production of compounds **1**－**4** in *S.* sp. CB02980, in comparison to the wild-type strain (Figure 3C). There are two standalone thioesterases (TE) PyfC and PyfN in the *pyf* gene cluster, which shared 58% and 53% sequence identity with HlmC and HlmK from holomycin biosynthetic gene cluster, respectively. Since co-expression of *hlmK* with thiolutin and aureothricin biosynthetic gene cluster in *S. albus* led to the increasement of their production, we hypothesized that overexpression of PyfN, which shared the same S92G mutation with HlmK, may also act as a proofreading TE for pyrroloformamide biosynthesis, can also improve the production titer of pyrroloformamides (Figure S39).[37] A *pyfN* overexpression plasmid containing the constitutive promoter *ermE**, was thus constructed in the integrative vector pSET152 and introduced into *S.* sp. CB02980 by conjugation to obtain a *pyfN* overexpression mutant. Comparing to the *S.* sp. CB02980 wild-type strain, the production of pyrroloformamides **1** and **2** increased more than 2-fold in the mutant strain with

one additional copy of *pyfN*. These data confirmed that the *pyf* gene cluster is responsible for pyrroloformamides biosynthesis in *S.* sp. CB02980.

To study the evolutionary differences among the *pyf* gene cluster and other dithiolopyrrolone biosynthetic gene clusters, a genome neighborhood network (GNN) analysis was conducted (Figure 3D). Using PyfE as a query sequence, the survey of the NCBI GenBank resulted in a total of 115 proteins with sequence identity > 43%. Using a 95% cutoff to de-duplicate gene clusters containing PyfE homologs of extremely high similarity, and removal of 37 gene clusters for lacking of essential genes for dithiolopyrrolone biosynthesis, a total of 28 putative dithiolopyrrolone biosynthetic gene clusters were identified (Tables S9-S36). Although a set of highly conserved dithiolopyrrolone core biosynthetic genes are present in almost all of those gene clusters, it seemed that two distinct types of dithiolopyrrolone gene clusters were present, which were tentatively named as holomycin-type or pyrroloformamide-type gene cluster. For example, the Cys-specific adenylation domain in the NRPSs, such as PyfE and HolE, shared highly conserved sequence motifs with the Cys-specific adenylation domain of LnmI, which loads an L-Cys to its cognate PCP in leinamycin biosynthesis (Figure S39).[38,39] A standalone thioesterase presumably responsible for proofreading the peptide intermediate, as well as a thioredoxin-disulfide reductase and acyl-CoA dehydrogenase for the formation of the ene-thiol core motif were also present. In contrast, comparing to the holomycin-type gene clusters, the pyrroloformamide-type gene clusters

contained a set of unique biosynthetic genes, such as RidA family proteins and methyltransferases, while they lacked the representative acetyl- or acyl-transferase, such as HlmA, in holomycin-type gene cluster. Interestingly, the homolog protein of HlmI, a thioredoxin reductase catalyzing the formation of disulfide bond, was only present in 14 out of 28 orphan dithiolopyrrolone gene clusters, indicating an alternative mechanism for the disulfide bond formation and the resistance mechanism for these putative dithiolopyrrolones. Based on the above analysis, the dithiolopyrrolone core biosynthesis for pyrroloformamides may share with the well-studied holomycin (Figure 3E).[19] However, they diverged significantly after holothin or its thiol-containing precursor was off-loaded from the NRPS assembly line. Since either holothin or its biosynthetic intermediates were also heavily modified in pyrroloformamide biosynthesis, the detailed enzymatic picture of on-line tailoring of the Cys-Cys dipeptide intermediate remains obscure. In addition, the mechanism responsible for N-formylation in pyrroloformamides could not be deduced based on the current analysis of its biosynthetic pathway. Interestingly, two orphan pyrroloformamide-type gene clusters from the in-house strains, *S.* sp. CB02009 and *S.* sp. CB02261, were also identified, which shared high sequence similarity and similar gene organization with *pyf* in *S.* sp. CB002980 (Figure 3A). Therefore, our identification and preliminary characterization of these pyrroloformamide biosynthetic gene clusters might potentially provide us a great opportunity to study their biosynthesis, and to

generate "unnatural" dithiolopyrrolones by manipulating unique enzymes from both holomycin-type and pyrroloformamide-type biosynthetic machinery.

Conclusions

Dithiolopyrrolone natural products are promising drug leads, due to their ability to chelate Zn^{2+} ion in essential enzymes in pathogenic bacteria or cancer cells. In this study, we have isolated two new pyrroloformamides **3** and **4** from *S.* sp. CB02980, and established their structures by HRESIMS, 1D and 2D NMR analysis. Their structures were further confirmed by the single crystal X-ray crystallographic analysis or semisynthesis. The identification and preliminary characterization of the pyrroloformamide biosynthetic gene cluster revealed a unified mechanism for the biosynthesis of dithiolopyrrolones. Our study may inspire future efforts to study their biosynthesis and discover new pyrroloformamides.

Experimental Section

General experimental procedures.

IR spectra were recorded on Nicolet iS50 FT-IR (Thermo Scientific). HRMS spectra were recorded on a LTQ-ORBITRAP-ETD instrument. NMR spectra were acquired using a Brucker 500/400 MHz spectrometer. Chemical shifts were reported in ppm relative to CD_3CN (δ = 1.94 ppm), CD_3OD (δ = 3.31 ppm) or DMSO-d_6 (δ = 2.50 ppm) for ^{1}H-NMR and CD_3CN (δ = 1.32 and 118.26 ppm), CD_3OD (δ = 49.00 ppm) or DMSO-d_6 (δ = 39.60 ppm) for ^{13}C NMR spectroscopy. Column chromatography (CC) was carried out on silica gel (100–200 mesh and 300–400 mesh, Yantai Jiangyou Silica Gel Development Co., Ltd., Yantai, China). Semi-preparative reversed phase-high-performance liquid chromatography (RP-HPLC) was performed using a Waters 1525 Binary HPLC Pump equipped with a Waters 2489 UV/Visible Detector and using a Welch Ultimate AQ-C18 column (250 × 10 mm, 5 μm). Crystal data were acquired on a Rigaku APEX-II XtaLAB PRO MM007HF diffractometer using Cu Kα radiation at 100 K. The structure was analyzed by a SHELXS-97 software and refined by means of full-matrix least-squares.

Bacterial strains.

The *S.* sp. CB02980 strain was grown G1 agar plates (2% soluble starch, 0.1% KNO_3, 0.05% K_2HPO_4, 0.05% $MgSO_4·7H_2O$, 0.05% NaCl, 0.001% $FeSO_4·7H_2O$, 2% agar) and incubated at 28 °C to obtain spores.

Fermentation.

The *S*. sp. CB02980 spores were inoculated into 250 mL Erlenmeyer flasks containing 50 mL tryptic soy broth medium (1.7% tryptone, 0.3% soya peptone, 0.5% NaCl, 0.25% K_2HPO4, 0.25% dextrose, pH 7.3 ± 0.2) at 30 °C on a rotary shaker at 220 rpm for 24 h. Then 10% (v/v) seed cultures were transferred into 50 mL production medium (2% soluble starch, 2% corn flour, 0.05% KH_2PO_4, 0.025% $MgSO_4$, trace element solution (V/V=1/1000) (0.1% $ZnSO_4·7H_2O$, 0.1% $FeSO_4·7H_2O$, 0.1% $MnCl_2·4H_2O$, 0.1% $CaCl_2$), adjust pH to 7.1, then added 0.5% $CaCO_3$), with 6.0% (v/v) wet macroporous resins DA201-H (Jiangsu Su Qing Water Treatment Engineering Group Co., Ltd., Jiangyin, China) in a 250 mL flask. Before inoculation, the production medium was supplemented with different concentration sterile cysteine, and then cultured for 8 days at 30 °C on a rotary shaker at 230 rpm. For large-scale fermentation, the seed cultures (50 mL) were aseptically transferred to 2 L Erlenmeyer flasks containing 500 mL of production medium.

Extraction and isolation.

After large-scale fermentation (12 L), the resins were filtered and dried in air. They were extracted with MeOH and dried in vacuum, further extracted by EtOAc to afford the crude extracts (10 g). The crude extracts were subjected to silica gel chromatography (CC) and eluted with a gradient of petroleum ether (PE)/EtOAc (100:0, 9:1, 8:2, 7:3, 6:4, 1:1, 4:6, 3:7, 2:8, 1:9, 0:100, v/v) to give 7 fractions (Fr. A1–Fr. A7). Fr. A7 was purified by silica gel CC, using

dicholoromethane/MeOH (100:0, 99:1, 98:2, 97:3, 96:4, 95:5, 94:6, 93:7, 92:8, 91:9, 9:1, 1:1, 0:100, v/v) to obtain compound **1** (250 mg) and compound **2** (284 mg). Fr. A6 was further purified by Sephadex LH-20 chromatography and then silica gel CC to obtain compound **3** (24 mg).

Compound **4** was obtained from the culture without the addition of resins in the medium. After fermentation, the culture broth (4 L) was centrifuged (4000 rpm, 15 min) to obtain the supernatant, which was extracted with EtOAc (3 × 4 L) and dried over anhydrous Na_2SO_4. The crude extracts (188 mg) were obtained after concentration under reduced pressure, which were subjected to semipreparative RP-C18 HPLC (Welch Ultimate AQ 5 μm, 250 × 10 mm) with a flow rate of 2 mL/min and a gradient elution of CH_3CN /H_2O in 16 min (10% to 95% for 8 min, followed by 95% for 2 min, and 95% to 10% for 2 min, followed by 10% for 4 min), to afford compounds **4** (t_R = 11.5 min, 69 mg).

X-ray Single-Crystal Data of 3.

The crystals of **3** were obtained by crystallization of **3** from a solution of CH_3CN/H_2O (v/v, 1:1). Monoclinic crystals, a = 3.9579 (15) Å, b = 9.5320 (3) Å, c = 13.2614 (5) Å, α= 70.820(3)°, β = 85.006(3)°, γ =79.423(3)°, Z = 2, μ = 5.006 mm^{-1}, F (000) = 238, and T = 100 K; Crystal dimensions: 0.3 × 0.01 × 0.01 mm^3, Volume = 464.32 (3) $Å^3$, 1898 reflections measured, 1809 independent reflections (R_{int} = 0.0420), the final R indices [I > 2σ(I)] R_1 = 0.0388, wR_2 = 0.1107, R indices (all data) R_1 = 0.0400. The goodness of fit on F^2 was 1.0377. Crystallographic data for the structure reported in this paper

have been deposited with the Cambridge Crystallographic Data Centre deposition with the deposition no. CCDC 1900736 (**3**). Copies of the data can be obtained, free of charge, on application to the Director, CCDC, 12 Union Road, Cambridge CB2 1EZ, UK (fax: +44-(0)1223-336033 or e-mail: deposit@ccdc.cam.ac.uk).

Preparation of semisynthetic 4 from pyrroloformamide (1).

Compound **1** (20 mg) was dissolved in acetone/water solution (4 mL of acetone and 2 mL of water), and then 2 drops of sodium hydrogen carbonate dilute aqueous solution and 30% aqueous H_2O_2 (14 µL) was added. The whole reaction mixture was stirred at room temperature for 5 -10 min and concentrated under reduced pressure. The resulting residue was subjected to semi-preparative RP-C18 HPLC (Welch Ultimate AQ 5 µm, 250 × 10 mm) separation to afford semisynthetic **4** (t_R = 11.5 min, 12 mg) with a yield of 60%.

Pyrroloformamide A (1): yellow needle crystal. UV (CH_3CN) λmax 231.3, 364.4 nm. ^1H NMR (500 MHz, CD_3CN) and ^{13}C NMR (126 MHz, CD_3CN) see Table S1. HRESIMS m/z 214.9947 $[M + H]^+$ (calcd for $C_7H_7O_2N_2S_2$, 214.9949).

Pyrroloformamide B (2): yellow powder. UV (CH_3CN) λmax 245.5, 297.7, 386.1 nm. ^1H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6) see Table S2. ESIMS m/z 200.96 $[M + H]^+$ (calcd for $C_6H_5N_2O_2S_2$, 200.98).

Pyrroloformamide C (3): light yellow needle crystal. UV (CH_3CN): λmax 357 nm. IR v_{max} (film) 3136, 3037, 2999, 2937, 2814, 1686, 1667, 1587, 1342, 1303, 1125, 799, 775, 715 cm^{-1}. HRESIMS m/z 229.0095 $[M + H]^+$ (calcd for

$C_8H_9O_2N_2S_2$, 229.0105).

Rotamer **3a**: ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.55 (1H, s, H-5), 4.40 (2H, s, H-5'), 8.23 (1H, s, H-6), 3.07 (3H, s, H-7), 10.53 (1H, s, H-8); ^{13}C NMR (126 MHz, DMSO-d_6) δ 163.6 (s, C-1), 127.4 (s, C-2), 125.1 (s, C-3), 130.1 (s, C-4), 108.2 (s, C-5), 30.4 (s, C-5'), 163.0 (s, C-6), 30.4 (s, C-7). Rotamer **3b**: ^{1}H NMR (500 MHz, DMSO-d_6) δ 6.51 (1H, s, H-5), 4.28 (2H, s, H-5'), 8.17 (H, s, H-6), 3.25 (3H, s, H-7), 10.37 (H, s, H-8); ^{13}C NMR (126 MHz, DMSO-d_6) δ 163.8 (s, C-1), 125.1 (s, C-2), 130.7 (s, C-3), 130.2 (s, C-4), 108.1 (s, C-5), 30.3 (s, C-5'), 162.7 (s, C-6), 34.3 (s, C-7).

Pyrroloformamide D (4): yellow oil. UV (CH$_3$CN): λmax 296.5, 364.4 nm. IR v_{max} (film) 3435, 3074, 2922, 2850, 1739, 1667, 1620, 1311, 1173, 1099, 549 cm[-1]. HRESIMS m/z 244.9694 [M - H][-] (calcd for $C_7H_6O_4N_2S_2$, 244.9691).

Rotamer **4a**: ^{1}H NMR (500 MHz, CD$_3$CN) δ 6.80 (1H, s, H-5), 8.84 (1H, s, H-6), 3.15 (3H, s, H-7); ^{13}C NMR (126 MHz, CD$_3$CN) δ 165.0 (s, C-1), 127.2 (s, C-2), 141.5 (s, C-3), 121.6 (s, C-4), 107.7 (s, C-5), 161.9 (s, C-6), 29.3 (s, C-7).

Rotamer **4b**: ^{1}H NMR (500 MHz, CD$_3$CN) δ 6.63 (1H, s, H-5), 8.10 (1H, s, H-6), 3.46 (3H, s, H-7); ^{13}C NMR (126 MHz, CD$_3$CN) δ 165.3 (s, C-1), 124.1 (s, C-2), 141.7 (s, C-3), 126.5 (s, C-4), 108.0 (s, C-5), 163.6 (s, C-6), 33.7 (s, C-7).

Gene replacement of *pyfE* in *S.* sp. CB02980

A pOJ260-based plasmid pXY2001 was constructed to generate the Δ*pyfE* gene replacement mutant in *S.* sp. CB02980 via a double crossover homologous recombination. To inactivate *pyfE*, a 2,052-bp fragment of *pyfE*

gene was replaced with thiostrepton resistance gene assembled with *k*asOp*

using the In-Fusion cloning kit (Tsingke Biological Technology Co., Ltd., Beijing,

China), and the mutated *pyfE* gene was cloned into pOJ260 between the

*Hind*III and *Xba*I restriction sites. This plasmid was introduced into *S.* sp.

CB02980 by conjugation and selected for thiostrepton resistance and

apramycin-sensitive phenotype to isolate the desired double-crossover mutant

strains *S.* sp. YX2001. The genotypes of the mutants were confirmed by PCR

and DNA sequencing. The PCR primers were shown in Table S6.

Overexpression of *pyfN* gene in *S.* sp. CB02980

A pSET152-based plasmid pXY2002, in which the expression of *pfyN* is under

the control of the constitutive promoter *ermE**, was constructed. A 900-bp *pyfN*

gene was PCR-amplified by high-fidelity DNA polymerase from *S.* sp.

CB02980 using the primers listed in Table S6 and the resultant PCR fragment

was subcloned into the pSET152 between *BamH*I and *Xba*I restriction sites

under the control of a constitutive promoter *ermE*p*. This plasmid pXY2002

was introduced into *S.* sp. CB02980 by conjugation and selected for apramycin

resistant conjugants to afford *S.* sp. YX2001.

Genome neighborhood network analysis of dithiolopyrrolone gene

clusters in GenBank

Using PyfE as the query, we searched for its homologous proteins in the

GenBank database on March 6, 2019. A total of 115 proteins with sequence

identity over 43% were collected, and their lengths range between 1,000 to

1,300 amino acids. Protein homologs below 43% were unlikely involved in dithiolopyrrolone biosynthesis due to the lack of key dithiolopyrrolone biosynthetic homologue genes, such as *hlmB*, *hlmC*, *hlmD* and *hlmF*. The 115 PyfE homologs were winnowed into 65 representative sequences by filtering them to a maximum identity of 95% using CD-HIT.[40] A total of 37 proteins were further removed from the 65 representative sequences, whose gene clusters also lack the essential *hlmB*, *hlmC*, *hlmD* and *hlmF* homologue genes for dithiolopyrrolone biosynthesis. All the proteins in the remaining 28 putative dithiolopyrrolone gene clusters were used in an all versus all BLAST using BLAST+ and an *E* value limit of 10.[41] Cytoscape v3.4 was used for GNN visualization and analysis.[42]

Phylogenetic analysis of the adenylation domain in NRPS

The amino acid sequences of adenylation domains in NRPSs of dithiolopyrrolone gene clusters were aligned using ClustalW in MEGA 7.[43] A maximum likelihood phylogenetic tree was generated using the Jones–Taylor–Thornton model of amino acid substitution and 100 bootstrap replications. The adenylation domain of LnmI in leinamycin gene cluster was used as an outer group.

Biological Assays

The cytotoxicity of compounds **1**, **3**, **4** against the human cancer cell lines A549 and Caco-2 were evaluated with the MTT assay.[32] The antibacterial

activity of these compounds was tested against *S. aureus* ATCC 29213 and MRSA, as well as Gram-negative pathogen *E. coli* and *K. pneumoniae*, in accordance with the previously reported method.[44] All the tested compounds were dissolved in DMSO.

ASSOCIATED CONTENT

AUTHOR INFORMATION

ORCID

Jianhua Ju: 0000-0001-7712-8027

Ben Shen: 0000-0002-9750-5982

Yong Huang: 0000-0002-3163-1716

Supporting Information

The supporting information is available free of charge on the Publication website.

Availability of data and materials

The shotgun genome sequence of *S.* sp. CB02980 (PRJNA524931) was submitted to GenBank and may be accessed.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

(1)   Mandeville, C. W. *Elements* **2010**, *6*, 75– 80.

(2)   Okamura, K.; Soga, K.; Shimauchi, Y.; Ishikura, T.; Lein, J. *J. Antibiot.* **1977**, *30*, 334– 336.

(3)   Celmer, W. D.; Solomons, I. A. *J. Am. Chem. Soc.* **1955**, *77*, 2861– 2865.

(4)   Mcinerney, B. V.; Gregson, R. P.; Lacey, M. J.; Akhurst, R. J.; Lyons, G. R.; Rhodes, S. H.; Smith, D. R.; Engelhaedt, L. M.; White, A. H. *J. Nat. Prod.* **1991**, *54*, 774– 784.

(5)   Shiozawa, H.; Kagasaki, T.; Kinoshita, T.; Haruyama, H.; Domon, H.; Utsui, Y.; Kodama, K.; Takahashi, S. *J. Antibiot.* **1993**, *46*, 1834– 1842.

(6)   Shiozawa, H.; Kagasaki, T.; Torikata, A.; Tanaka, N.; Fujimoto, K.; Hata, T.; Furukawa, Y.; Takahashi, S. *J. Antibiot.* **1995**, *48*, 907– 909.

(7)   Shiozawa, H.; Shimada, A.; Takahashi, S. *J. Antibiot.* **1997**, *50*, 449– 452.

(8)   Qin, Z.; Baker, A. T.; Raab, A.; Huang, S.; Wang, T.; Yu, Y.; Jaspars, M.; Secombes, C. J.; Deng, H. *J. Biol. Chem.* **2013**, *288*, 14688– 14697.

(9)   Wietz, M.; Mansson, M.; Gotfredsen, C. H.; Larsen, T. O.; Gram, L. *Mar. drugs* **2010**, *8*, 2946– 2960.

(10) Li, B.; Wever, W. J.; Walsh, C. T.; Bowers, A. A. *Nat. Prod. Rep.* **2014**, *31*, 905– 923.

(11)   Chan, A.; Shiver, A. L.; Wever, W. J.; Razvi, S. Z.; Traxler, M. F.; Li, B.

Proc. Natl. Acad. Sci. U S A. **2017**, *114*, 2717– 2722.

(12) Lauinger, L.; Li, J.; Shostak, A.; Cemel, I. A.; Ha, N.; Zhang, Y.; Merkl, P. E.; Obermeyer, S.; Stankovic-valentin, N.; Schafmeier, T. Wever, W. J.; Bowers, A. A.; Carter, K. P.; Palmer, A. E.; Tschochner, H.; Melchior, F.; Deshaies, R. J.; Brunner, M.; Diernfellner, A. *Nat. Chem. Biol.* **2017**, *13*, 709– 714.

(13) Deshaies, R. J. *BMC Biol.* **2014**, *12*, 94– 107.

(14) Murphy, A. C.; Fukuda, D.; Song, Z.; Hothersall, J.; Cox, R. J.; Willis, C. L.; Thomas, C. M.; Simpson, T. J. *Angew. Chem. Int. Ed. Engl.* **2011**, *50* 3271– 3274.

(15) Dunn, Z. D.; Wever, W. J.; Economou, N. J.; Bowers, A. A.; Li, B. *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 5137– 5141.

(16) Daehne, W.; Godtfredsen, O. W.; Tybring, L.; Schaumburg, K. *J Antibiot.* ***1969****, 22, 233– 236.*

(17) Jensen, B. *J. Antibiot.* **1969**, *22*, 231– 232.

(18) Abreu, P. A.; Sousa, T. S.; Jimenez, P. C.; Wilke, D. V; Rocha, D. D.; Freitas, H. P.; Pessoa, O. D.; La Clair, J. J.; Costa-Lotufo, L. V. *Chembiochem* **2014**, *15*, 501– 506.

(19) Li, B.; Walsh, C. T. *Proc. Natl. Acad. Sci. U S A.* **2010**, *107*, 19731– 19735.

(20) Li, B.; Walsh, C. T. *Biochemistry* **2011**, *50*, 4615– 4622.

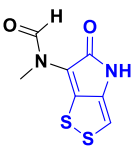(21) Fukuda, D.; Haines, A. S.; Song, Z.; Murphy, A. C.; Hothersall, J.;

Stephens, E. R.; Gurney, R.; Cox, R. J.; Crosby, J.; Willis, C. L.; Simpson, T. J., Thomas, C. M. *PLoS One* **2011**, *6*, e18031.

(22) Huang, S.; Zhao, Y.; Qin, Z.; Wang, X.; Onega, M.; Chen, L.; He, J.; Yu, Y.; Deng, H. *Process Biochemistry* **2011**, *46*, 811– 816.

(23) Saker, S.; Chacar, S.; Mathieu, F. *Enzyme Microb. Technol.* **2015**, *72*, 35– 41.

(24) Huang, S.; Tong, M. H.; Qin, Z.; Deng, Z.; Deng, H.; Yu, Y. *Anticancer Agents Med. Chem.* **2015**, *15*, 277– 284.

(25) Pan, G.; Xu, Z.; Guo, Z.; Hindra; Ma, M.; Yang, D.; Zhou, H.; Gansemans, Y.; Zhu, X.; Huang, Y.; Zhao, L. X.; Jiang, Y.; Cheng, J.; Van Nieuwerburgh, F.; Suh, J. W.; Duan, Y.; Shen, B. *Proc. Natl. Acad. Sci. U S A.* **2017**, *114*, E11131– E11140.

(26) Bouras, N.; Mathieu, F.; Sabaou, N.; Lebrihi, A. *J. Appl. Microbiol.* **2006**, *100*, 390– 397.

(27) Ding, H.; Wang, J. N.; Zhang, D. S.; Ma, Z. *Chem. Biodivers.* **2017**, *14*, e1700140.

(28) Takahashi, S.; Shiozawa, H.; Fujimoto, K.; Iwano, Y.; Hirai, K.; Torikata, A.; Kagasaki, T.; Ogawa, K.; Sakaida, Y.; Kodama, K.; Ishii, A. Thiomarinol compounds. U.S. Patent 5399711, **1995**.

(29) Eder, C.; Kurz, M.; Wink, J. E.U. Patent 20010105959, **2004**.

(30) Webster, J. M.; Li, J.; Chen, G. U.S. Patent 19960685211, **1996**.

(31) Chan, A. N.; Wever, W. J.; Massolo, E.; Allen, S. E.; Li, B. *Chem. Res.*
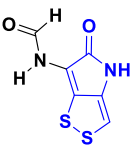
Toxicol. **2019**, *32*, 400– 404.

(32) Korzeniewski, C.; Callewaert, D. M. *J. Immunol. Methods* **1983**, *64*, 313–320.

(33) Belous, M. A.; P. I. A., *Zh. Obshch. Khim.* **1950**, *20*, 1701– 1710.

(34) Behroozi, S. J.; Kim, W.; Dannaldson, J.; Gates, K. S. *Biochemistry* **1996**, *35*, 1768– 1774.

(35) Blin, K.; Wolf, T.; Chevrette, M. G.; Lu, X.; Schwalen, C. J.; Kautsar, S. A.; Suarez Duran, H. G.; de Los Santos, E. L. C.; Kim, H. U.; Nave, M.; Dickschat, J. S.; Mitchell, D. A.; Shelest, E.; Breitling, R.; Takano, E.; Lee, S. Y.; Weber, T.; Medema, M. H. *Nucleic Acids Res.* **2017**, *45*, W36– W41.

(36) Wang, W.; Li, X.; Wang, J.; Xiang, S.; Feng, X.; Yang, K. An Engineered Strong Promoter for Streptomycetes. *Appl. Environ. Microbiol.* **2013,** *79*, 4484– 4492.

(37) Zhai, Y.; Bai, S.; Liu, J.; Yang, L.; Han, L.; Huang, X.; He, J. *Biochem. Biophys. Res. Commun.* **2016**, *473*, 329– 335.

(38) Conti, E.; Stachelhaus, T.; Marahiel, M. A.; Brick, P. *EMBO J.* **1997**, *16*, 4174– 4183.

(39) Tang, G. L.; Cheng, Y. Q.; Shen, B. *Chem. Biol.* **2004**, *11*, 33– 45.

(40) Huang, Y.; Niu, B.; Gao, Y.; Fu, L.; Li, W.; *Bioinformatics* **2010**, *26*, 680–682.

(41)   Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.;
       Bealer, K.; Madden, T. L. *BMC Bioinformatics* **2009**, *10*, 421– 429.

(42)   Kohl, M.; Wiese, S.; Warscheid, B. *Methods Mol. Biol.* **2011**, *696*, 291–
       303.

(43)   Kumar, S.; Stecher, G.; Tamura, K. *Mol. Biol. Evol.* **2016**, *33*, 1870–
       1874.

(44)   Elshikh, M.; Ahmed, S.; Funston, S.; Dunlop, P.; McGaw, M.; Marchant,
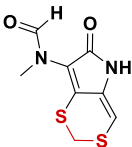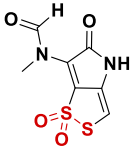   R.; Banat, I. M. *Biotechnol. Lett.* **2016**, *38*, 1015– 1019.

TOC

Pyrroloformamide
(vD 844, **1**)

Pyrroloformamide B
(vD 846, **2**)

Pyrroloformamide C
(**3**)

Pyrroloformamide D
(**4**)

S. sp. CB02980

S. sp. CB02009

S. sp. CB02261

S. clavulgerus
(holomycin)