



## OPEN ACCESS

**Edited by:**

Prasat Kittakoop,  
Chulabhorn Graduate  
Institute, Thailand

**Reviewed by:**

Fernando Reyes,  
Fundación MEDINA, Spain  
Jianhua Ju,  
South China Sea Institute of  
Oceanology, China

**\*Correspondence:**

Bin Hong  
hongbin@imb.pumc.edu.cn  
Yuning Xie  
xieyy@imb.pumc.edu.cn

†These authors have contributed  
equally to this work

**‡Present address:**

Li Liu,  
Institute of Chinese Materia Medica,  
China Academy of Chinese Medical  
Sciences, Beijing, China  
Wenqiang Guo,  
Department of Medicine,  
Jinggangshan University, Ji'an, China

**Specialty section:**

This article was submitted to  
Medicinal and Pharmaceutical  
Chemistry,  
a section of the journal  
Frontiers in Chemistry

**Received:** 13 January 2020

**Accepted:** 03 March 2020

**Published:** 20 March 2020

**Citation:**

Li Y, Liu L, Zhang G, He N, Guo W,  
Hong B and Xie Y (2020)  
Potashchelins, a Suite of Lipid  
Siderophores Bearing Both L-threo  
and L-erythro Beta-Hydroxyaspartic  
Acids, Acquired From the  
Potash-Salt-Ore-Derived Extremophile  
*Halomonas* sp. MG34.  
*Front. Chem.* 8:197.  
doi: 10.3389/fchem.2020.00197

# Potashchelins, a Suite of Lipid Siderophores Bearing Both L-threo and L-erythro Beta-Hydroxyaspartic Acids, Acquired From the Potash-Salt-Ore-Derived Extremophile *Halomonas* sp. MG34

Yihong Li<sup>1†</sup>, Li Liu<sup>1‡</sup>, Gengxin Zhang<sup>2</sup>, Ning He<sup>1</sup>, Wenqiang Guo<sup>1‡</sup>, Bin Hong<sup>1,3\*</sup> and Yuning Xie<sup>1\*</sup>

<sup>1</sup> CAMS Key Laboratory of Synthetic Biology for Drug Innovation, Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China, <sup>2</sup> Key Laboratory of Alpine Ecology, Institute of Tibetan Plateau Research, China Academy of Sciences, Beijing, China, <sup>3</sup> NHC Key Laboratory of Biotechnology of Antibiotics, Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

Four new lipid siderophores bearing both L-threo- and L-erythro- $\beta$ -hydroxyaspartic acids, potashchelins A-D (**1-4**), were isolated from the potash-salt-ore-derived extremophile *Halomonas* sp. MG34. The planar structures of **1-4** were elucidated on the basis of extensive 1D and 2D NMR studies and MS/MS data. Potashchelins **1-4** contain a hydrophilic nonapeptide headgroup sequentially consisting of  $\beta$ -hydroxyaspartic acid, serine, glycine, serine, serine,  $\beta$ -hydroxyaspartic acid, threonine, serine, and cyclic N( $\delta$ )-hydroxy-ornithine, which is appended by one of a series of fatty acids ranging from dodecanoic acid to tetradecanoic acid. The absolute configurations of the amino acids of potashchelins **1-4** were determined by C<sub>3</sub> and advanced Marfey's reaction, partial hydrolysis, and bioinformatics analysis, which revealed that potashchelins **1-4** bear both L-threo- and L-erythro- $\beta$ -hydroxyaspartic acid. Phylogenetic analysis showed that the stand-alone  $\beta$ -hydroxylase, PtcA, and the fused domain with  $\beta$ -hydroxylase activity in PtcB are expected to be responsible for the formation of L-erythro and L-threo diastereomers, respectively. Additionally, utilizing a comparative genomics approach, we revealed an evolutionary mechanism for lipid siderophores in *Halomonas* involving horizontal transfer. Bioassays showed that potashchelin A and D had weak antibacterial activity against *B. subtilis* CPCC 100029 with an MIC value of 64  $\mu$ g/mL.

**Keywords:** potashchelins, lipid siderophore, L-threo- $\beta$ -hydroxyaspartic acid, L-erythro- $\beta$ -hydroxyaspartic acid, *Halomonas*,  $\beta$ -hydroxylases

## INTRODUCTION

Extremophiles are microbes that inhabit extreme environments and represent a fruitful source of natural products (Wilson and Brimble, 2009; Zhang et al., 2018). Extremophiles can be classified as halophiles, alkaliphiles, acidophiles, piezophiles, psychrophiles, thermophiles, etc. based on their optimum growth conditions (Rothschild and Mancinelli, 2001). For example, halophiles thrive in

environments with a high salt concentration. In our research, we are dedicated to discovering novel natural products produced by halophiles collected from the Qinghai-Tibet Plateau.

Nonribosomal peptides are assembled by nonribosomal ribosomal peptide synthetases (NRPSs) and play a pivotal role in the treatment of diseases (Sussmuth and Mainz, 2017) and the survival of producer strains (Boiteau et al., 2016). With the deciphering of the biosynthetic logic of nonribosomal peptides and advances in genome sequencing technologies and bioinformatics, many genome mining methods targeting NRPSs have been developed to exploit novel nonribosomal peptides (Chen et al., 2019). Among these methods, PCR-based genome mining using degenerate primers is valued for its ability to rapidly identify strains potentially producing the desired structural class of compounds before the whole genomes of the microbes are sequenced. In recent years, PCR-based genome mining has been successfully utilized to discover novel natural products (Liu et al., 2018; Zhou et al., 2018).

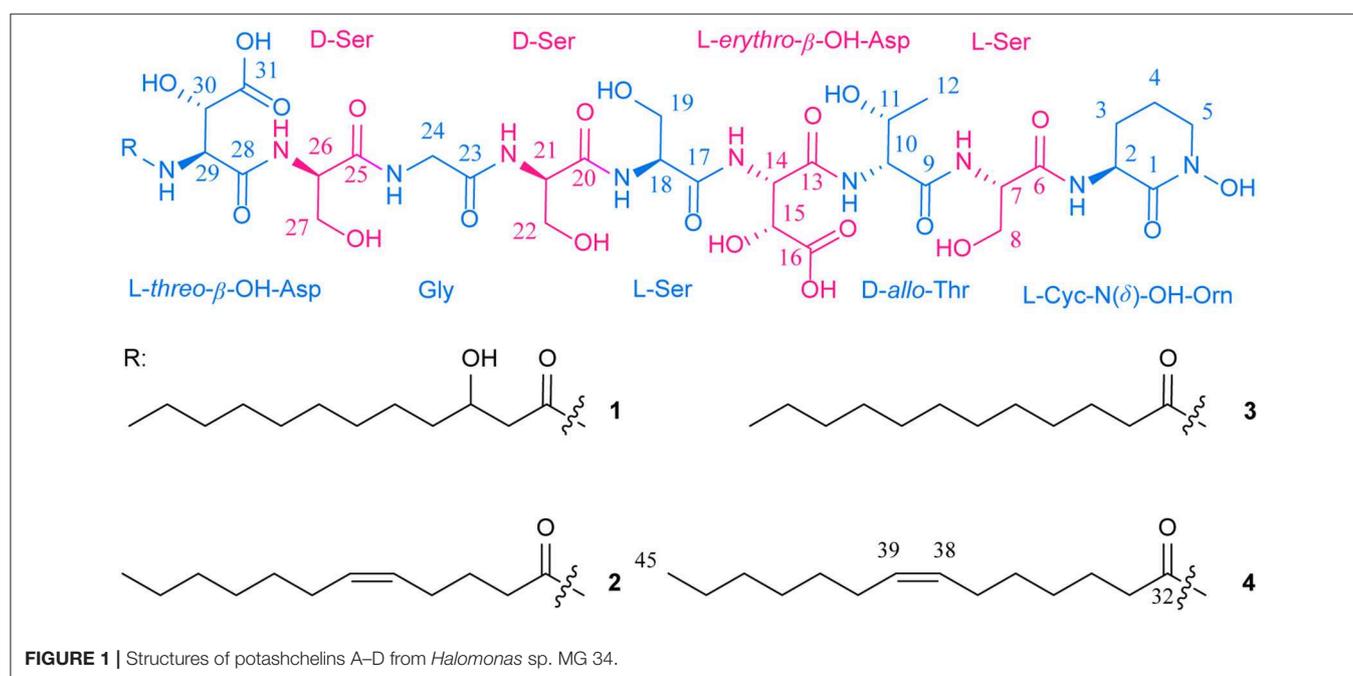
Here, we report the discovery of four novel lipid siderophores, potashchelins A-D (1-4, **Figure 1**), from *Halomonas* sp. MG34 based on bioassays and PCR-guided screening. The planar structures of 1-4 were elucidated by NMR and MS/MS spectral data. The absolute configurations of the amino acids forming these peptides were determined by a combination of Marfey's analysis, partial hydrolysis, and bioinformatics analysis of their biosynthetic gene cluster. Unexpectedly, both *L*-threo- and *L*-erythro- $\beta$ -hydroxyaspartic acids are present in potashchelins 1-4. Further phylogenetic analysis revealed that the stereospecificity of the  $\beta$ -hydroxylases in the lipid siderophore biosynthesis gene clusters can be predicted, which in turn indicated that the *L*-threo and *L*-erythro diastereomers in potashchelins 1-4 are synthesized

by the stand-alone  $\beta$ -hydroxylase PtcA and the fused domain with  $\beta$ -hydroxylase activity in PtcB, respectively. Additionally, using a comparative genomics approach, we disclosed an evolutionary mechanism of lipid siderophores in *Halomonas* involving horizontal transfer. Potashchelins A-D (1-4) were assayed for their antibacterial activity and potashchelin A and D showed weak antibacterial activity against *B. subtilis* CPC 100029 (64  $\mu$ g/mL MIC).

## MATERIALS AND METHODS

### General Experimental Details

UV data were recorded on a Shimadzu UV-2550 spectrophotometer. NMR data were acquired with Varian Mercury 600 spectrometers using DMSO- $d_6$  as solvent. HRESIMS and ESIMS/MS data were recorded on a Thermal LTQ Orbitrap XT mass spectrometer. HPLC analyses were performed on an Agilent 1200 or Shimadzu DGU-20A instrument using an XBridge C<sub>18</sub> column (3.5  $\mu$ m, 4.6  $\times$  150 mm) on a binary LC system [solvent A: 0.1% (v/v) TFA aqueous], solvent B: acetonitrile containing 0.1% (v/v) TFA as modifier; flow rate, 1 mL/min; 0–40 min, 25–45% B (linear gradient); UV detection at 215 nm and oven temperature at 25°C. HPLC purifications were carried out using a XBridge<sup>TM</sup> Prep C<sub>18</sub> (5  $\mu$ m, 10  $\times$  150 mm); eluted with 0.1% (v/v) TFA aqueous (solvent A)–0.1% (v/v) TFA in MeCN (solvent B) from 37% to 45% B during 50 min; flow rate, 2.5 mL/min; UV detection at 215 nm and oven temperature at 25°C. The optimized eluents 37, 40, 42, and 45% acetonitrile aqueous inclusive of 0.1% TFA modifier were used to repurify potashchelins A-D, respectively.



## Isolation, Screening, and Identification of *Halomonas* sp. MG34

Fifteen extremophile strains were isolated from potash salt ore, collected in Qinghai-Tibet Plateau, China and cultured in the medium (S5): 30 g/L NaCl, 100 g/L MgCl<sub>2</sub>, 5 g/L MgSO<sub>4</sub>·7H<sub>2</sub>O, 5 g/L soya peptone, 3 g/L yeast extract, 18 g/L agar, adjusted to pH 7.2, supplemented with 3% NaCl.

For PCR screening of potential producers of nonribosomal peptides, the genomic DNAs from all 15 strains were extracted using standard protocols (Kieser et al., 2000). The degenerate primers of A3F (5'-GCSTACSYSATSTACACSTCSGG-3') and A7R (5'-SASGTCVCCSGTSCGGTAS-3') were used in the process of screening (Ayuso-Sacido and Genilloud, 2005). A 50 μL PCR system containing 2 μL of forward primer (10 μM), 2 μL of reverse primer (10 μM), 2 μL genomic DNA, 25 μL of Easy Taq Polymerase (Beijing TransGen Biotech, Beijing, China), and 19 μL of sterilized water, was used. The PCR program was 95°C/5 min [95°C/30 s, 59°C/2 min, 72°C/4 min] × 35 cycles, 72°C/10 min. The PCR products were analyzed by agarose gel electrophoresis and the expected size was 700–800 bp. In all of the 15 halophile strains screened, NRPS sequences were detected (Figure S1).

For bioassay screening of potential producers with antibacterial activity, the spores of each of the 15 strains were inoculated into thirteen different media (see Table S1) and cultured at 28°C, 180 rpm for 5 days. The fermentation broth was centrifuged and the supernatant was tested for their activity against *B. subtilis* CMCC 100027, *M. phlei* CMCC 160023, *S. aureus* ATCC 29213, *E. coli* ATCC 25922, *P. aeruginosa* ATCC 27853, and *C. albicans* ATCC 10231 using a modified cylinder plate method. The cylinders were put on the surface of agar plate with various tested bacteria and the fermentation broth supernatant was added into the cylinders. The antibacterial activity can be detected by the inhibition zone. The strain MG34 was picked out because it exhibited antibacterial activity only in high salt media of DEF-15 (+) containing 3% NaCl (Table S1).

The *Halomonas* sp. MG34 was taxonomically identified based on the housekeeping 16S rRNA gene. Briefly, the housekeeping 16S rRNA gene was amplified by PCR and sequenced using the primers 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3'). Blast on the EzTaxon-e server (<http://www.ezbiocloud.net/>) disclosed that its 16S rRNA gene sequence (GenBank accession no. MN636765) shows high identity with those of *Halomonas* strains. The phylogenetic trees, based on the 16S rRNA gene sequences of the strain MG34 and the related *Halomonas* homologs, identified MG34 as a *Halomonas* strain (Figure S2).

## Scale-Up Fermentation and Isolation

The spores of *Halomonas* sp. MG34 were inoculated into high salt medium of DEF-15 (+) (Table S1) to be precultured at 28°C and 220 rpm for 48 h. Then, 50 mL of preculture was transferred into 5 L of Erlenmeyer flasks containing 1 L of DEF-15 (+) medium and incubated on a rotary shaker at 220 rpm, and 28°C. After 5 days, the culture broth was harvested. The fermentation broth was filtered to remove the mycelia and 10 L of filtrate was

obtained. Then the target compounds were enriched from the filtrate using a column of macroporous absorbent resin 4006 (1 L, 7.2 × 27 cm). After washing with 5 L of water, the active absorbed materials were eluted with 2 L of 20% and 2 L of 50% aqueous acetone, which were combined and lyophilized to afford 500 mg of crude extract. Antibiotic activity was determined by a paper-disk agar diffusion assay against *S. aureus* ATCC 29213 on Mueller-Hinton medium. The crude extract was further purified utilizing semi-preparative reversed phase HPLC chromatography (Figure S3), running with H<sub>2</sub>O/MeCN containing 0.1% TFA, by repeated preparation to afford potashchelins A (1, 3 mg), B (2, 5 mg), C (3, 3 mg), D (4, 4 mg).

Potashchelin A (1): white powder;  $[\alpha]_D^{25} +77.8$  (c 0.05, MeOH); UV (MeOH)  $\lambda_{\max}$  (log  $\epsilon$ ) 212 (5.13); 1D and 2D NMR (600 MHz, DMSO-*d*<sub>6</sub>) (see Table 1 and Supplementary Material); HRESI(+)-MS  $[M + H]^+$  *m/z* 1097.4900 (calcd for C<sub>43</sub>H<sub>73</sub>N<sub>10</sub>O<sub>23</sub>, 1097.4850).

Potashchelin B (2): white powder;  $[\alpha]_D^{25} 0$  (c 0.05, MeOH); UV (MeOH)  $\lambda_{\max}$  (log  $\epsilon$ ) 202 (3.75); 1D and 2D NMR (600 MHz, DMSO-*d*<sub>6</sub>) (see Table 1 and Supplementary Material); HRESI(+)-MS  $[M + H]^+$  *m/z* 1079.46915 (calcd for C<sub>43</sub>H<sub>71</sub>N<sub>10</sub>O<sub>22</sub>, 1079.4744).

Potashchelin C (3): white powder;  $[\alpha]_D^{25} -158.4$  (c 0.05, MeOH); UV (MeOH)  $\lambda_{\max}$  (log  $\epsilon$ ) 206 (3.42); 1D and 2D NMR (600 MHz, DMSO-*d*<sub>6</sub>) (see Table 1 and Supplementary Material); HRESI(+)-MS  $[M + H]^+$  *m/z* 1081.48462 (calcd for C<sub>43</sub>H<sub>73</sub>N<sub>10</sub>O<sub>22</sub>, 1081.4901).

Potashchelin D (4): white powder;  $[\alpha]_D^{25} -7.92$  (c 0.05, MeOH); UV (MeOH)  $\lambda_{\max}$  (log  $\epsilon$ ) 206 (3.47); 1D and 2D NMR (600 MHz, DMSO-*d*<sub>6</sub>) (see Table 1 and Supplementary Material); HRESI(+)-MS  $[M + H]^+$  *m/z* 1107.50574 (calcd for C<sub>45</sub>H<sub>75</sub>N<sub>10</sub>O<sub>22</sub>, 1107.5057).

## Marfey's Analyses

### C<sub>3</sub> Marfey's Analysis

C<sub>3</sub> Marfey's analysis was carried out following the reported method (Kreutzer et al., 2012; Vijayasathy et al., 2016). Briefly, compounds 1–4 (50 μg each) were hydrolyzed in 6 M HI (100 μL) at 115°C for 5 h. Then, the hydrolysates were concentrated to dryness at 115°C for 1 h under a stream of dry N<sub>2</sub>. Subsequently, the hydrolysates were treated with 1 M NaHCO<sub>3</sub> (30 μL), and then with L-FDAA (1% solution in acetone, 40 μL) at 40°C for 12 h, after which the reaction was neutralized with 1 M HCl (30 μL) and diluted with 500 μL 10% acetonitrile in water prior to HPLC-ESIMS analysis. Authentic standards of L-Ser, D-Ser, L-Orn, D-Orn, L-Thr, D-Thr, DL-threo-β-OH-Asp, and DL-allo-Thr were derivatized with L-FDAA according to the above method. 2 μL of each derivative was analyzed using HPLC-ESIMS on an Agilent Zorbax SB-C<sub>3</sub> column (5 μm, 150 × 4.6 mm, 50°C, 1 mL/min) with a gradient elution using H<sub>2</sub>O (mobile phase A) and MeCN (mobile phase B) containing formic acid (0.1%). The gradient elution program was 17% B from 0 to 20 min, 17–45% B from 20 to 30 min, 45–17% B from 30 to 32 min, and 17% B from 32 to 40 min. The presence of each amino acid was assessed by UV (340 nm) and MS, and then their retention times were compared with those from the authentic standard derivatives. The measured retention times (*t*<sub>R</sub>, min) of

**TABLE 1** | <sup>1</sup>H (600 MHz) and <sup>13</sup>C (150 MHz) NMR data for potashchelin A (**1**), B (**2**), C (**3**), and D (**4**) in DMSO-*d*<sub>6</sub>.

Pos	Potashchelin A ( <b>1</b> )		Potashchelin B ( <b>2</b> )		Potashchelin C ( <b>3</b> )		Potashchelin D ( <b>4</b> )	
	$\delta_c$	$\delta_H$ (J in Hz)						
<b>Cyclic, N-OH Ornithine</b>								
C1	164.9		164.9		164.9		164.9	
C2	49.8	4.31, m	49.8	4.31, m	49.8	4.33, m	49.8	4.33, m
C3	27.1	1.91, m	27.1	1.86, m	27.1	1.87, m	27.1	1.87, m
		1.69, m		1.69, m		1.69, m		1.69, m
C4	20.3	1.92, m	20.3	1.90, m	20.3	1.89, m	20.3	1.89, m
C5	51.2	3.49, m						
		3.46, m		3.46, m		3.46, m		3.46, m
N1		8.11, d (8.9)		8.11, d (8.4)		8.11, d (8.4)		8.11, d (8.4)
<b>Serine</b>								
C6	169.7		169.7		169.7		169.7	
C7	54.9	4.37, dd (13.3, 5.6)	54.9	4.36, dd (13.3, 5.6)	54.9	4.36, dd (13.3, 5.6)	54.9	4.36, dd (13.2, 5.6)
C8	61.8	3.59, dd (10.9, 5.5)	61.8	3.59, dd (10.9, 5.5)	61.8	3.59, dd (10.8, 5.4)	61.8	3.58, m
N2		7.97, d (7.7)		7.96, d (7.7)		7.96, d (7.7)		7.96, d (7.6)
<b>Threonine</b>								
C9	169.7		169.7		169.7		169.7	
C10	58.5	4.17, dd (7.7, 7.0)	58.6	4.17, t (7.2)	58.6	4.17, t (6.6)	58.6	4.17, t (7.3)
C11	66.8	3.83, m	66.8	3.82, m	66.8	3.83, m	66.8	3.83, m
C12	19.7	1.04, d (6.3)						
N3		7.62, d (8.1)		7.61, d (8.1)		7.61, d (8.1)		7.61, d (8.1)
<b><math>\beta</math>-OH aspartic acid</b>								
C13	168.2		168.2		168.2		168.2	
C14	55.7	4.80, dd (8.6, 3.4)	55.7	4.79, dd (8.6, 3.4)	55.7	4.79, dd (8.6, 3.4)	55.7	4.79, dd (8.6, 3.3)
C15	71.0	4.12, d (3.4)	71.0	4.12, d (3.3)	71.0	4.12, d (3.3)	71.0	
C16	172.5		172.5		172.5		172.5	
N4		8.41, d (8.6)		8.41, d (8.6)		8.39, d (8.6)		8.40, d (8.5)
<b>Serine</b>								
C17	170.2		170.2		170.2		170.2	
C18	55.4	4.29, m	55.4	4.29, m	55.4	4.30, m	55.4	4.30, m
C19	61.8	3.61, d (5.7)	61.8	3.61, d (5.7)	61.8	3.61, d (5.4)	61.8	3.61, d (5.0)
N5		7.76, d (8.0)		7.76, d (7.7)		7.74, d (7.6)		7.74, d (7.5)
<b>Serine</b>								
C20	170.0		170.0		170.0		170.0	
C21	55.3	4.28, m						
C22	61.6	3.54, m	61.6	3.54, m	61.5	3.54, m	61.6	3.54, m
N6		7.80, d (7.8)		7.76, d (7.7)		7.76, d (8.2)		7.76, d (8.0)
<b>Glycine</b>								
C23	169.3		169.4		169.4		169.4	
C24	42.2	3.79, d (6.0)	42.2	3.79, d (5.7)	42.2	3.79, t (5.6)	42.2	3.79, t (5.6)
N7		8.12, t (5.7)		8.13, t (5.7)		8.13, t (5.7)		8.13, t (5.7)
<b>Serine</b>								
C25	170.2		170.2		170.2		170.2	
C26	54.6	4.44, m	54.6	4.44, dd (13.3, 6.5)	54.6	4.44, dd (13.3, 6.5)	54.6	4.44, dd (13.2, 6.5)
C27	61.9	3.54, m						
		3.66, dd (10.6, 5.7)		3.66, dd (10.5, 5.6)		3.66, dd (10.5, 5.7)		3.66, m
N8		7.97, d (7.7)		7.96, d (7.7)		7.96, d (7.7)		7.96, d (7.6)
<b><math>\beta</math>-OH aspartic acid</b>								
C28	168.8		168.8		168.7		168.7	
C29	55.4	4.73, dd (9.0, 2.6)	55.4	4.73, dd (9.2, 2.6)	55.3	4.75, dd (9.2, 2.6)	55.2	4.75, dd (9.2, 2.5)

(Continued)

TABLE 1 | Continued

Pos	Potashchelin A (1)		Potashchelin B (2)		Potashchelin C (3)		Potashchelin D (4)	
	$\delta_c$	$\delta_H$ (J in Hz)	$\delta_c$	$\delta_H$ (J in Hz)	$\delta_c$	$\delta_H$ (J in Hz)	$\delta_c$	$\delta_H$ (J in Hz)
C30	70.1	4.50, d (2.7)	70.1	4.49, d (2.5)	70.4	4.49, d (1.8)	70.4	4.49, br s
C31	173.0		172.9		172.9		172.9	
N9		7.91, d (9.0)		7.88, d (9.2)		7.86, d (9.2)		7.87, d (9.2)
<b>Fatty acid tail</b>								
C32	171.1		172.4		172.5		172.5	
C33	43.3	2.31, dd (14.1, 5.1) 2.23, dd (14.1, 7.0)	34.9	2.16, dd (8.2, 6.5)	35.3	2.15, t (7.5)	35.2	2.15, t (7.4)
C34	67.4	3.79, m	25.5	1.51, m	25.3	1.46, br s	25.1	1.47, m
C35	36.5	1.35, m	26.3	1.99, m	28.6	1.24, br s	28.3	1.24, m
C36	25.2	1.26, m	129.1	5.34, m	28.9	1.24, br s	29.0	1.26, m
C37	29.1	1.24, m	130.1	5.34, m	28.9	1.24, br s	26.5	1.98, m
C38	29.1	1.24, m	26.6	1.97, m	29.1	1.24, br s	129.6	5.33, m
C39	29.0	1.24, m	29.1	1.24, m	29.0	1.24, br s	129.6	5.33, m
C40	28.7	1.24, m	28.3	1.24, m	28.7	1.24, br s	26.6	1.98, m
C41	31.3	1.24, m	31.2	1.24, m	31.3	1.24, br s	29.1	1.26, m
C42	22.1	1.26, m	22.1	1.26, m	22.1	1.27, m	28.3	1.24, m
C43	14.0	0.86, t (6.9)	14.0	0.86, t (6.9)	14.0	0.85, t (7.0)	31.1	1.24, m
C44	–	–	–	–	–	–	22.1	1.24, m
C45	–	–	–	–	–	–	13.9	0.85, t (6.8)

authentic L-FDAA derivatives were as follows (min): L-FDAA-L-Ser (9.4), L-FDAA-D-Ser (10.6), L-FDAA-L-Orn (3.7), L-FDA-D-Orn (3.3), L-FDAA-L-*threo*- $\beta$ -OH-Asp (6.0), L-FDAA-D-*threo*- $\beta$ -OH-Asp (5.5), L-FDAA-L-*allo*-Thr (11.7) and L-FDAA-D-*allo*-Thr (15.4) (**Figure S4**). L-FDAA derivatized hydrolysates of **1-4** gave retention times of ( $t_R$ , min): L-FDAA-L-Ser (9.4), L-FDAA-D-Ser (10.6), L-FDAA-L-Orn (3.7), L-FDAA-L-*threo*- $\beta$ -OH-Asp (6.0), L-FDAA-D-*allo*-Thr (15.4) and D or L-*erythro*- $\beta$ -OH-Asp (10.1) (**Figure S4**).

### Advanced Marfey's Method

To determine the stereochemistry at C-2 in *erythro*- $\beta$ -OH-Asp residues in potashchelins, advanced Marfey's method was used. Briefly, **1-4** (50  $\mu$ g each) were hydrolyzed completely and derivatized with L- and D- FDLA using the same method as above. Authentic standard DL-*threo*- $\beta$ -OH-Asp was derivatized with L-FDLA according to the above method. An aliquot (2  $\mu$ L) of each derivative was analyzed using HPLC-ESIMS on an Agilent Zorbax SB-C<sub>3</sub> column (5  $\mu$ m, 150  $\times$  4.6 mm, 50°C, 1 mL/min) with a gradient elution using H<sub>2</sub>O (mobile phase A) and MeCN (mobile phase B) containing formic acid (0.1%). The gradient elution program was 27% B from 0 to 20 min, 27–45% B from 20 to 30 min, 45–27% B from 30 to 32 min and 27% B from 32 to 40 min. Authentic L-FDLA derivatized DL-*threo*- $\beta$ -OH-Asp gave retention times ( $t_R$ , min): L-FDLA-D-*threo*- $\beta$ -OH-Asp (7.2), and L-FDLA-L-*threo*- $\beta$ -OH-Asp (7.3). L-FDLA derivatized hydrolysates of **1-4** gave retention times of ( $t_R$ , min): D-FDLA-L-*threo*- $\beta$ -OH-Asp (7.2), L-FDLA-L-*threo*- $\beta$ -OH-Asp (7.3), D-FDLA-L-*erythro*- $\beta$ -OH-Asp

(10.0), and L-FDLA-L-*erythro*- $\beta$ -OH-Asp (10.5). Therefore, the *erythro*- $\beta$ -OH-Asp residues in potashchelins were all determined as L configuration.

### Partial Hydrolysis

To determine the position of L-*erythro*- and L-*threo*- $\beta$ -OH-Asp residues in potashchelins, **2** (300  $\mu$ g) was partially hydrolyzed in 0.5 M HCl (100  $\mu$ L) at 100°C for 40 min and then the hydrolysates were concentrated to dryness at 100°C for 40 min under a stream of dry N<sub>2</sub>. The hydrolysates were treated with 1 M NaHCO<sub>3</sub> (30  $\mu$ L), diluted with 500  $\mu$ L 10% acetonitrile in water, and then subjected to HPLC (Agilent Zorbax SB-C<sub>3</sub> column, 5  $\mu$ m, 150  $\times$  4.6 mm, 30°C, 1 mL/min) with a gradient elution using H<sub>2</sub>O (mobile phase A) and MeCN (mobile phase B) containing formic acid (0.1%). The gradient elution program was 35% B from 0 to 10 min, 35–95% B from 10 to 11 min, 95% B from 11 to 15 min, 95–35% B from 15 to 16 min and 35% B from 16 to 20 min. The main hydrolysis product giving an ion at  $m/z$  496 corresponding to fatty-acid-L- $\beta$ -OH-Asp-Ser-Gly was picked out and purified (**Figure S5**). Then, it was treated and analyzed using the above mentioned advanced Marfey's method. The retention time of the L-FDLA derivatized L- $\beta$ -OH-Asp based on SIE ( $m/z$  466 [M+Na]<sup>+</sup>) was 7.5 min ( $m/z$  466), which is identical to that of L-FDLA derivatized L-*threo*- $\beta$ -OH-Asp.

### Genomic DNA Sequencing, Assembly and Bioinformatics Analysis

The strain *Halomonas* sp. MG34 was grown in tryptic soy broth liquid medium and genomic DNA was extracted using

a standard chloroform protocol (Nikodinovic et al., 2003). The genome was sequenced using a next generation sequencing platform in the paired-end ( $2 \times 300$ ) format, resulting in 824 Mb clean data (2,764,928 reads with 298 bp average insert size and 110-fold average coverage). The genome was assembled with the SPAdes algorithm (V3.13.1) (Bankevich et al., 2012) to yield the first version of the draft genome. Then, to obtain the complete biosynthetic gene cluster (BGC) of potashchelins, antiSMASH and end-extending method (Huang et al., 2013) were employed. The genome was submitted to the National Center for Biotechnology Information (NCBI) database with the accession No. WJPH00000000. The biosynthetic gene cluster of potashchelins and its domain were identified in the assembled genome using standalone antiSMASH 5. The homologs of PtcA were collected from MIBiG (<https://mibig.secondarymetabolites.org/>, as of October, 2019). The maximum-likelihood phylogenetic tree of PtcA and its homologs were reconstructed in PhyML 3.2.0, using the LG amino acid substitution model (Guindon et al., 2010).

## Antibacterial Bioassay Method

The MICs for *B. subtilis* CMCC 100027, *M. phlei* CMCC 160023, *S. aureus* ATCC 29213, *E. coli* ATCC 25922, *P. aeruginosa* ATCC 27853, and *C. albicans* ATCC 10231 were determined by a microdilution method (Cockerill et al., 2012). Briefly, the bacterial strain was grown on Mueller-Hinton broth (MHB), and the final suspension of bacteria (in MHB medium) was  $10^6$  cells/mL. The fungus, *C. albicans* ATCC 10231, was cultured on SDB, and the final suspension concentration was the same as the bacteria. Tested samples were dissolved in DMSO and diluted serially. Then 1  $\mu$ L of each diluted sample was added into a 96-well plate in triplicate containing 100  $\mu$ L of the bacterial suspension in each well. After incubation at 37°C for 18 h, the growth of the tested organism was detected by eye, and the MIC was identified as the lowest concentration that completely inhibited growth of the organism. The positive controls were as the following (MIC,  $\mu$ g/mL): streptomycin for *B. subtilis* (8), *M. phlei* (16) and *M. smegmatis* (16), gentamycin for *E. coli* (2) and *P. aeruginosa* (2), vancomycin for *S. aureus* (2), amphotericin B for *C. albicans* (8). The MIC values of potashchelins A–D are shown in **Table 2**.

## RESULTS

### Strain Prioritization for Nonribosomal Peptide Discovery From Halophiles Based on the Combination of PCR and Bioassay Screening

The combination of PCR and bioassay screening was used to discover active nonribosomal peptides from halophiles isolated from samples collected from the Qinghai-Tibet Plateau. The degenerate primers A3F and A7R, deduced from the conserved sequences of the adenylation domains (A) of NRPSs, were designed to amplify NRPS gene sequences from Actinomycetes and proved to be very useful for strain prioritization (Ayuso-Sacido and Genilloud, 2005; Lemetre et al., 2017; Masand et al., 2018; Zhou et al., 2018). First, 15 halophile strains were screened using PCR primers A3F and A7R. To our surprise, NRPS sequences were detected in all 15 halophile strains screened by PCR (**Figure S1**). The positive rate of 100% in the PCR assay indicates that all of the tested strains have the potential to produce nonribosomal peptides. Then, each of the 15 strains was fermented on a small scale (100 mL) using 13 different types of media and prioritized by antibacterial activity. Strain MG34 was selected because it exhibited antibacterial activity only in DEF-15 (+) high-salt medium, which contained 3% NaCl (**Table S1**). To identify the genus of strain MG34, the housekeeping 16S rRNA gene was amplified by PCR and sequenced. Blast on the EzTaxon-e server (<http://www.ezbiocloud.net/>) revealed that the 16S rRNA gene sequence of MG34 showed high identity with those of *Halomonas* strains. Phylogenetic trees based on the 16S rRNA gene sequences of MG34 and related *Halomonas* homologs identified MG34 as a *Halomonas* strain (**Figure S2**).

### Production and Isolation of Potashchelins A–D

*Halomonas* sp. MG34 was fermented in DEF-15 (+) high-salt medium, and potashchelins A–D were isolated and purified using a bioassay-guided method from a 5-day broth culture. Briefly, the fermentation broth of *Halomonas* sp. MG34 was filtered, and potashchelins were enriched from the filtrate by a macroporous absorbent resin 4006 column. The yielded active

**TABLE 2** | Antimicrobial bioassay results (MIC,  $\mu$ g/mL) for potashchelins A–D (1–4).

Compounds	<i>B. subtilis</i> CPCC 100029	<i>M. phlei</i> CPCC 160023	<i>M. smegmatis</i>	<i>E. coli</i> ATCC 25922	<i>P. aeruginosa</i> 11	<i>S. aureus</i> ATCC 29213	<i>C. albicans</i> ATCC 10231
Potashchelin A (1)	64	>64	>64	>64	>64	>64	>64
Potashchelin B (2)	>64	>64	>64	>64	>64	>64	>64
Potashchelin C (3)	>64	>64	>64	>64	>64	>64	>64
Potashchelin D (4)	64	>64	>64	>64	>64	>64	>64
Streptomycin	8	16	16	ND	ND	ND	ND
Gentamycin	ND	ND	ND	2	2	ND	ND
Vancomycin	ND	ND	ND	ND	ND	2	ND
Amphotericin B	ND	ND	ND	ND	ND	ND	8

ND, Not detected.

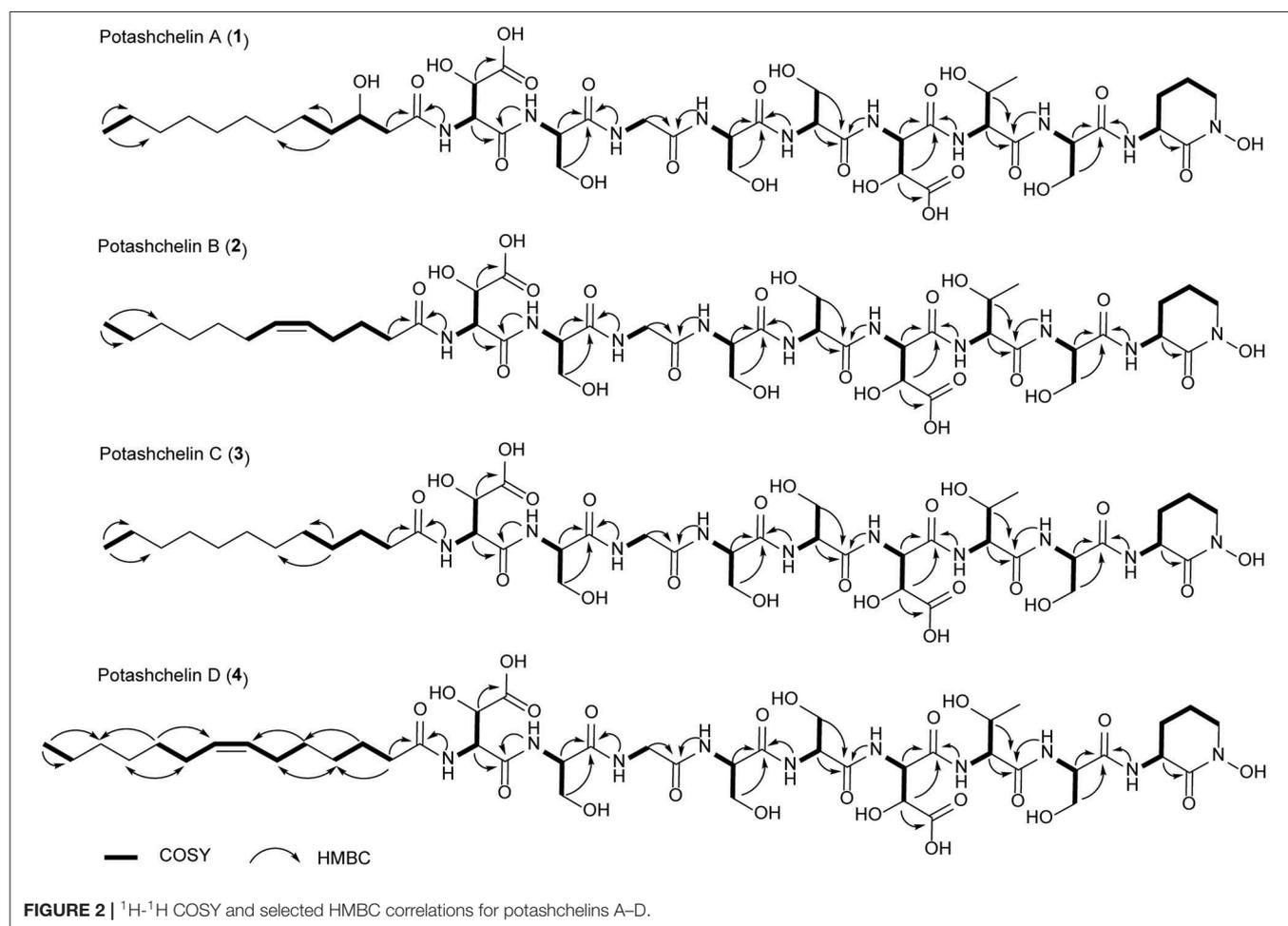
crude extract was analyzed by HPLC, but no apparent peaks appeared when an aqueous solution of MeOH or MeCN was used as an eluent. When TFA was added to the eluent as a modifier, five peaks appeared on the HPLC spectrum with excellent resolution (**Figure S3**). Therefore, the target peaks were further purified utilizing semi-preparative reversed phase HPLC chromatography (run with H<sub>2</sub>O/MeCN containing 0.1% TFA) to yield potashchelins A (**1**, 3 mg), B (**2**, 5 mg), C (**3**, 3 mg), and D (**4**, 4 mg).

## Planar Structural Elucidation by Spectroscopic Analysis

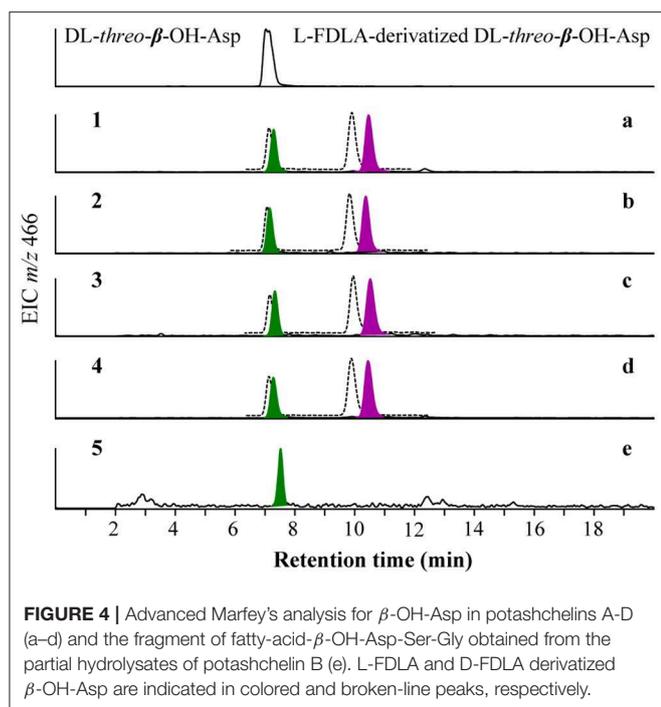
Potashchelin A (**1**) was obtained as a white powder and proved to have a molecular formula of C<sub>43</sub>H<sub>73</sub>N<sub>10</sub>O<sub>23</sub>, based on high resolution electrospray ionization mass spectrometry (HR-ESIMS) [M + H]<sup>+</sup> *m/z* 1097.4900 (calcd for C<sub>43</sub>H<sub>73</sub>N<sub>10</sub>O<sub>23</sub>, 1097.4850). Analysis of the <sup>1</sup>H and <sup>13</sup>C NMR data for **1** (**Table 1**) revealed the presence of twelve ester/amide carbonyls ( $\delta_C$  164.9–172.4), accounting for twelve double bond equivalents and requiring that **1** incorporate one ring. Further interpretation of <sup>1</sup>H-<sup>1</sup>H COSY, HSQC, and HMBC spectroscopic data (**Supplementary Material**) disclosed correlations, indicative of a

nonapeptide consisting of a cyclic N( $\delta$ )-hydroxyornithine (C1-C5), four serines (C6-C8, C17-C19, C20-C22, and C25-C27), two  $\beta$ -hydroxyaspartic acid (C13-C16 and C28-C31), one threonine (C9-C12), and one glycine (C23 and C24), as well as a fatty acid tail (C32-C43). The presence of these amino acid residues was also confirmed by C<sub>3</sub> Marfey analysis (see below). The fatty acid tail was evidenced to be 3-hydroxydodecanoic acid by comprehensive analysis of 1D and 2D NMR data, as well as comparison with the reported data in the literature (Frost and Gunstone, 1975; Gunstone et al., 1977).

The two-bond HMBC correlations (**Figure 2**) from amide NH protons of cyclic N( $\delta$ )-hydroxyornithine ( $\delta_H$  8.11), serine ( $\delta_H$  7.97), threonine ( $\delta_H$  7.62),  $\beta$ -hydroxyaspartic acid ( $\delta_H$  8.41), serine ( $\delta_H$  7.76), serine ( $\delta_H$  7.80), glycine ( $\delta_H$  8.12), serine ( $\delta_H$  7.97), and  $\beta$ -hydroxyaspartic acid ( $\delta_H$  7.91) to the carbonyl carbons of serine ( $\delta_C$  169.7), threonine ( $\delta_C$  169.7),  $\beta$ -hydroxyaspartic acid ( $\delta_C$  168.2), serine ( $\delta_C$  170.2), serine ( $\delta_C$  170.0), glycine ( $\delta_C$  169.3), serine ( $\delta_C$  170.2),  $\beta$ -hydroxyaspartic acid ( $\delta_C$  168.8), and the fatty acid tail ( $\delta_C$  171.1), respectively, preliminarily identified the sequence of nine amino acids and the fatty acid as fatty-acid- $\beta$ -OH-Asp-Ser-Gly-Ser-Ser- $\beta$ -OH-Asp-Thr-Ser-cyclic-N( $\delta$ )-OH-Orn. This sequence was further confirmed by the tandem mass spectrometry fragmentation







the similar “y” MS/MS fragments as 1–3, while different “b” fragments, each of which was 24 mass units greater than those of 2, hinting that the extra  $C_2H_2$  was located in the fatty acid tail. These results suggested that 4 could have a tetradecenoic acid tail. This hypothesis was further identified by the  $^1H$  and  $^{13}C$  NMR spectra of 4, which exhibited extra signals [ $\delta_C$  28.3, 29.0 and  $\delta_H$  1.24 (2H, m), 1.26 (2H, m)]. The extra signals were assigned as  $CH_2$ -35 ( $\delta_C$  28.3,  $\delta_H$  1.24) and  $CH_2$ -36 ( $\delta_C$  29.0,  $\delta_H$  1.26), based on the  $^1H$ - $^1H$  COSY connections between  $\delta_{H-34}$  1.47 and  $\delta_{H-35}$  1.24, as well as the HMBC connections from  $\delta_{H-33}$  2.15 to  $\delta_{C-35}$  28.3, and  $\delta_{H-34}$  1.47 to  $\delta_{C-36}$  29.0 (Figure 2). Furthermore, the  $^1H$ - $^1H$  COSY spectrum disclosed that H-36 ( $\delta_H$  1.26) correlated to allylic methylene protons ( $\delta_H$  1.98), which in turn connected with the olefinic proton ( $\delta_H$ -38 5.33), indicative of the position of the double bond between C-38 and C-39. The  $\Delta^7$  position and *cis* configuration of the double bond in the fatty acid of 4 were further confirmed by comprehensive interpretation of 1D and 2D NMR data of 4 (Table 1 and Figure 3) and comparison with those of 2. Therefore, 4 contained (*Z*)-tetradec-7-enoic acid.

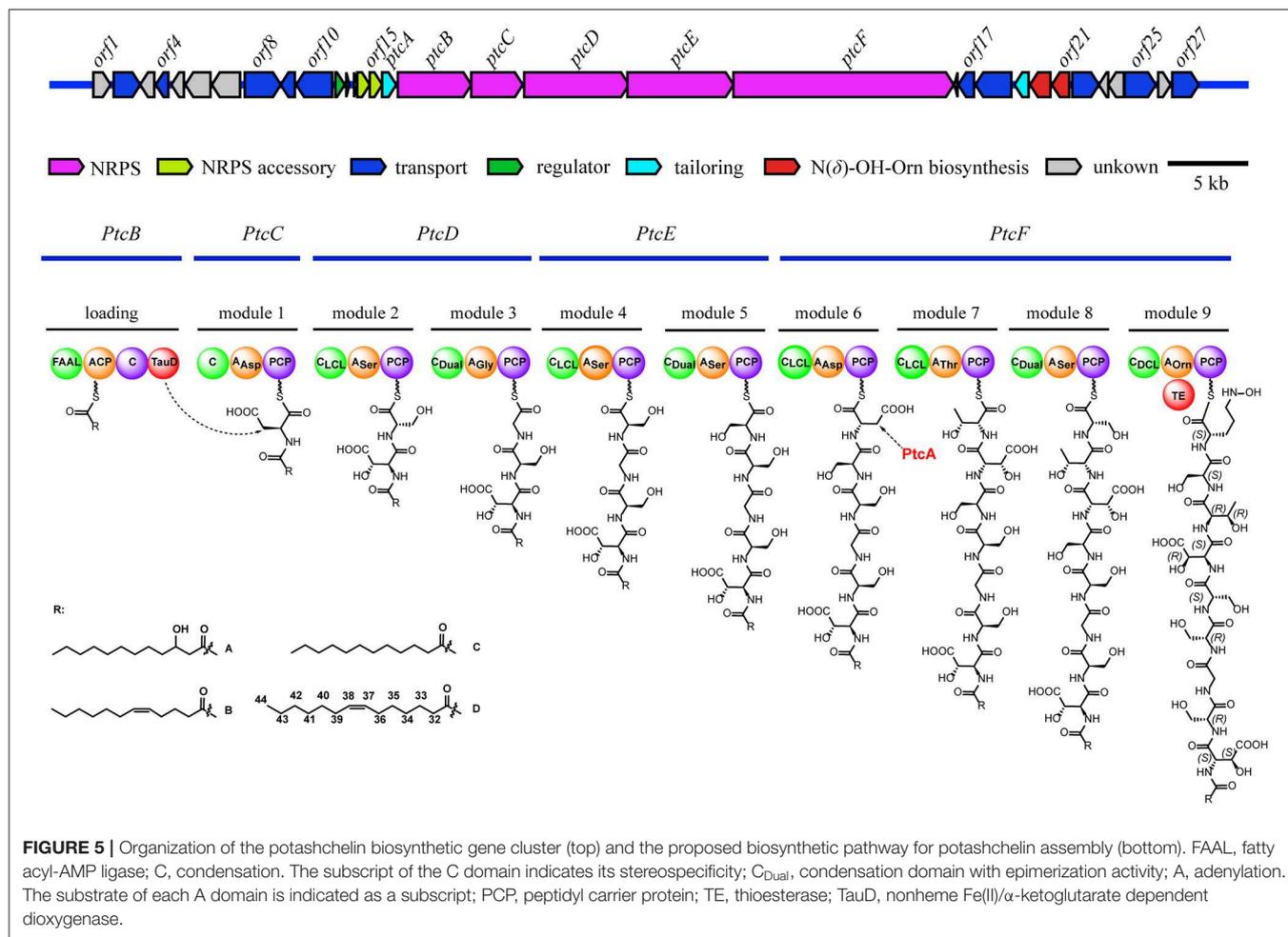
### Absolute Configuration Determination by the Combination of Marfey's Method, Partial Hydrolysis and Bioinformatics Analysis

To resolve the absolute configurations of the amino acid residues, we first applied  $C_3$  Marfey's method (Vijayasathy et al., 2016) based on acidic hydrolysis of potashchelins. Reductive HI cleavage was carried out to release two  $\beta$ -hydroxyaspartic acid moieties and an ornithine residue (Kreutzer et al., 2012). After derivatization with Marfey's reagent (L-FDAA), the hydrolysates of 1–4 were analyzed by HPLC-DAD-MS and compared with

authentic amino acid standards. The analyses revealed the presence of L-Orn, two L-Ser, two D-Ser, D-*allo*-Thr, L-*threo*- $\beta$ -OH-Asp, and D- or L-*erythro*- $\beta$ -OH-Asp residues in 1–4 (Figure S4). The advanced Marfey's method, in which the hydrolysates of 1–4 were derivatized with L-FDLA or D-FDLA, was employed because the standard of *erythro*- $\beta$ -OH-Asp was unavailable (Fujii et al., 1997a,b). According to the elution order of the diastereomeric pairs of L- and D-FDLA-derivatized *erythro*- $\beta$ -OH-Asp (Fujii et al., 1997b), the *erythro*- $\beta$ -OH-Asp residues in 1–4 were all determined to exist in the L configuration (Figure 4). Partial hydrolysis was performed to address the regiochemistry of the enantiomeric and epimeric amino acid residues in 1–4. The main fragment, determined to be fatty-acid- $\beta$ -OH-Asp-Ser-Gly based on mass analysis (Figure S5), was purified, hydrolyzed, derivatized with L-FDLA, and subjected to  $C_3$  Marfey's analysis. The regiochemistry of the L-*threo*- $\beta$ -OH-Asp adjacent to the fatty acid tail was unambiguously established (Figure 4), but, unfortunately, that of D-Ser and L-Ser were not determined despite significant effort because only trace amounts of the corresponding partial hydrolysates were present.

Bioinformatics analyses were used to completely clarify the regiochemistry of the enantiomeric serine residues in 1–4. A draft genome sequence of *Halomonas* sp. MG 34 was obtained using next generation sequencing platform and evaluated using antiSMASH 5.0.0 for the presence of a biosynthetic gene cluster (BGC) identical to the NRPS-derived backbone of 1–4. The analysis showed that the putative potashchelin BGC was divided into four subclusters, which were arranged according to the structures of 1–4 and the substrate specificities of the adenylation domains of NRPSs. The end-extending method (Huang et al., 2013) was employed to fill the gaps between contigs. All three gaps between these four contigs were filled, and a connected contig containing the complete potashchelin BGC was obtained, which made it possible to determine the regiochemistry of the enantiomeric serine residues in 1–4.

The architecture and features of the potashchelin BGC were interrogated, and a model for potashchelin biosynthesis was therefore deduced (Figure 5). Based on the antiSMASH and BLAST analyses, ten genes within the BGC (*orf2*, 4, 8, 9, 10, 17, 18, 22, 25, and 27) were determined to be involved in siderophore transport, while one gene (*orf11*) was associated with regulation of this cluster. *orf20* and *orf21* encode a lysine/ornithine N-monooxygenase and an acyltransferase, respectively, which are proposed to act concertedly to supply the amino acid precursor N( $\delta$ )-OH-Orn (Kreutzer et al., 2012), whereas a lipase encoded by *orf19* may be responsible for providing the fatty acid starter unit. In addition, an MbtH-like protein (Zhang et al., 2010), type-II thioesterase (Schwarzer et al., 2002), and 4'-phosphopantethenyl transferase (Flugel et al., 2000) encoded by *orf13*, *orf14* and *orf15*, respectively, are expected to activate and improve the functionality of the NRPS biosynthesis machinery. The five NRPSs derived from the successive genes *ptcB*, *ptcC*, *ptcD*, *ptcE*, and *ptcF* are responsible for the assembly of the nonapeptide potashchelin backbone. The resulting proteins form ten modules: one for initiation and nine for extension and termination of the NRPS assembly line. The loading module, which has the

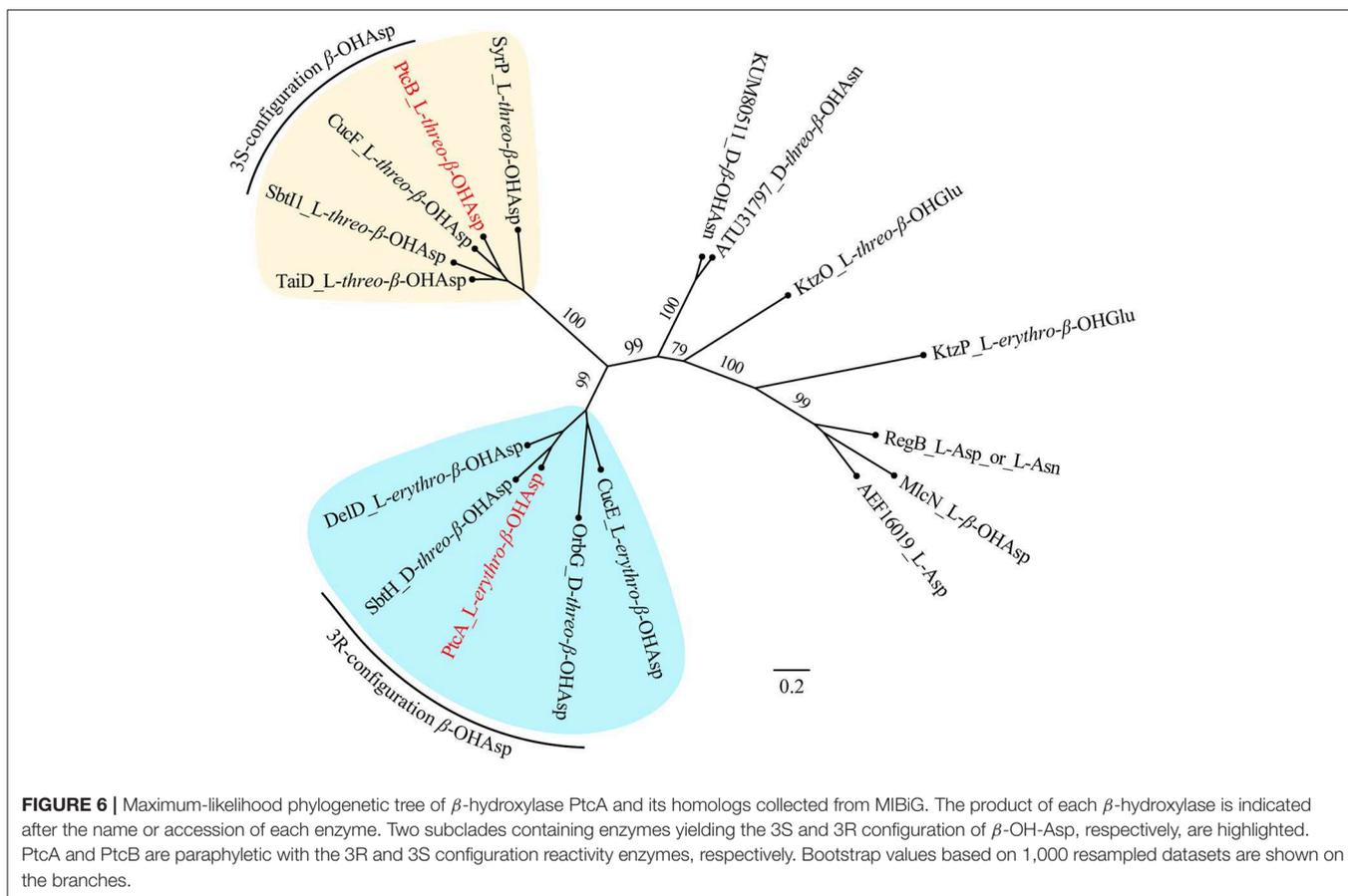


same domain arrangement as the taiwachelin BGC (Kreutzer and Nett, 2012), primes the fatty acid tail onto the ACP domain and initiates biosynthesis. Next, the nine amino acids constituting a potashchelin are assembled consecutively by the nine extension modules harbored by PtcC, PtcD, PtcE, and PtcF. The last extension module has a TE domain, which may be responsible for cyclization of N( $\delta$ )-OH-Orn and offloading the final product.

The adenylation (A) domains, from the specificity-conferring code of which the building blocks of peptides can be predicted, and the condensation (C) domains, from which the configuration of the peptides in these modules can be deduced, are of particular interest, as they account for the selection and stereochemistry of the monomer amino acids incorporated in potashchelins. The sequence of amino acid residues deduced by the biosynthesis analysis was the same as that determined by NMR and tandem MS interpretation. There are no *E* domains in the BGC, but the  $C_{Dual}$  domains (Balibar et al., 2005) within the BGC function both as an epimerase and in condensation, so the configurations of amino acids inserted in the peptide backbone were determined by these C domains (Figures S31, S32). Based on the arrangement of the domains in modules 1-9,

FAAL-ACP-C-TauD-C-A<sub>Asp</sub>-PCP- $C_{LCL}$ -A<sub>Ser</sub>-PCP- $C_{Dual}$ -A<sub>Gly</sub>-PCP- $C_{LCL}$ -A<sub>Ser</sub>-PCP- $C_{Dual}$ -A<sub>Ser</sub>-PCP- $C_{LCL}$ -A<sub>Asp</sub>-PCP- $C_{LCL}$ -A<sub>Thr</sub>-PCP- $C_{Dual}$ -A<sub>Ser</sub>-PCP- $C_{DCL}$ -A<sub>Orn</sub>-PCP-TE, the complete siderophore should feature L- $\beta$ -OH-Asp-D-Ser-Gly-D-Ser-L-Ser-L- $\beta$ -OH-Asp-D-*allo*-Thr-L-Ser-L-Cyc-N( $\delta$ )-OH-Orn as the peptide backbone. The presence of two L-Ser and two D-Ser amino acid residues is consistent with the results of Marfey's analysis.

NMR and Marfey's methods revealed the presence of L-*threo*- $\beta$ -OH-Asp and L-*erythro*- $\beta$ -OH-Asp at the first and sixth amino acid residues, respectively, of 1-4. PtcA, which has TauD activity, and the TauD domain at the C-terminus of PtcB are expected to be responsible for hydroxylation at the beta carbon of aspartic acid after it bonds to ACP, as in cupriachelin biosynthesis (Kreutzer et al., 2012). It was reported that cupriachelin contains two L-*threo*- $\beta$ -OH-Asp residues (Kreutzer et al., 2012), but a more recent study identified one L-*threo*-OH-Asp residue and one L-*erythro*- $\beta$ -OH-Asp residue (Reitz et al., 2019). To assess the stereospecificity of PtcA and the TauD domain at the C-terminus of PtcB, MiBIG-provided homologs of PtcA involved in the beta-hydroxylation of Asp, Glu and Asn were aligned (as of November, 2019). Interestingly, all stand-alone

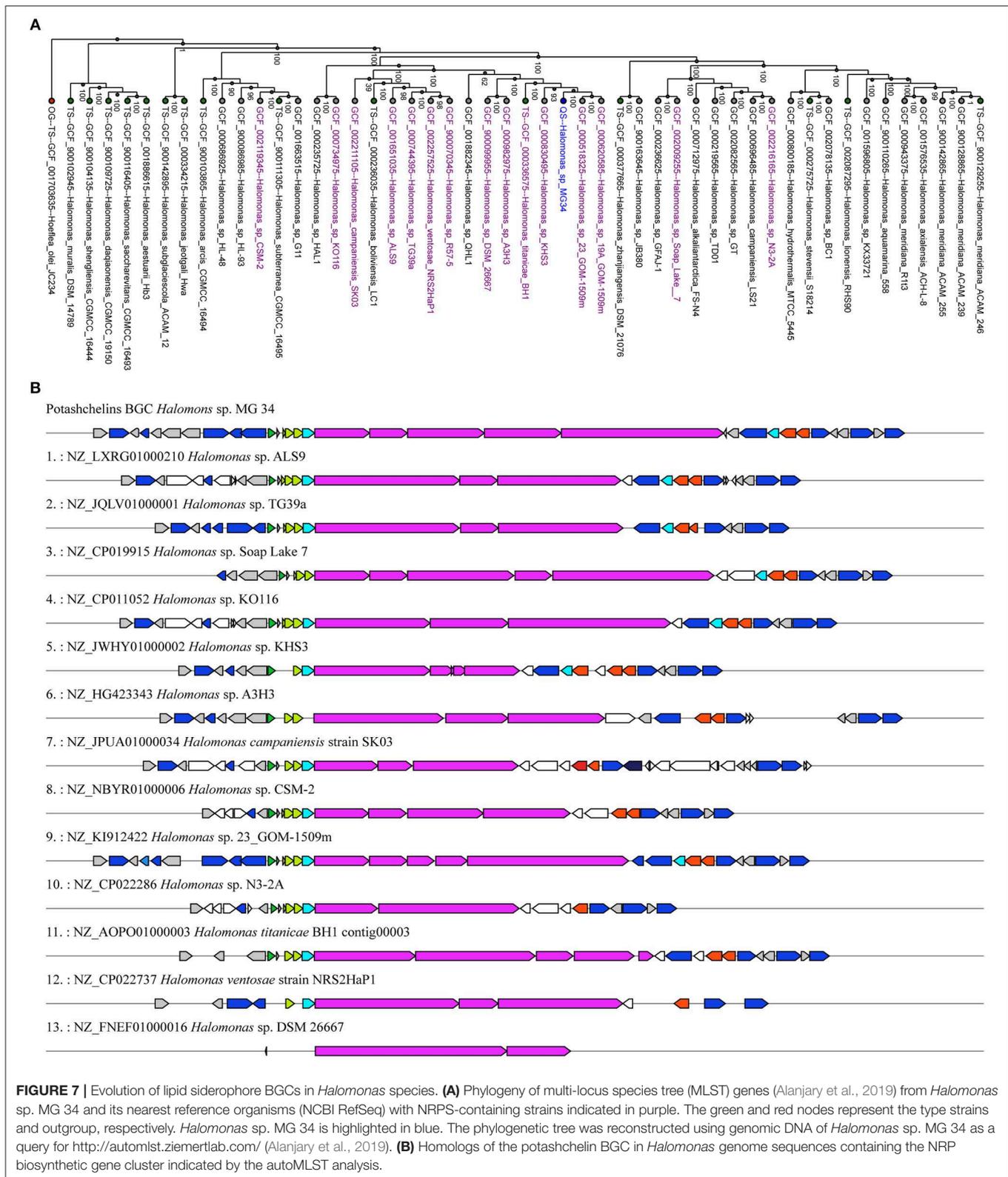


enzymes and fused domains catalyzing the beta carbon of Asp to produce the *R* configuration were clustered together, while those producing the *S* configuration were clustered separately (Figure 6). When PtcA and the PtcB domain with TauD activity were considered together, the latter was clustered with the *S* configuration subclade, while the former was located in the *R* configuration subclade. Therefore, the TauD domain in PtcB can be envisioned to catalyze ACP-Asp in its adjacent module to form *L*-threo- $\beta$ -OH-Asp, as in taiwachelin biosynthesis (Kreutzer and Nett, 2012), while PtcA is proposed to be involved in the production of *L*-erythro- $\beta$ -OH-Asp. As this manuscript was nearing completion, a similar analysis about stereospecificity of  $\beta$ -hydroxylases was reported by Reitz et al. (2019), in which a method to predict  $\beta$ -hydroxylase stereochemistry *in silico* was developed through mapping stereochemically characterized  $\beta$ -OH-Asp residues in siderophores to the phylogenetic tree of  $\beta$ -hydroxylases. Our result confirms that the stereospecific reactivity of  $\beta$ -hydroxylases in siderophore biosynthesis can be predicted by aligning their amino acid sequences.

## Evolution of the Potashchelin Gene Cluster

Siderophores play a particularly important role in the bioavailability of iron in iron-scarce areas (Boiteau et al., 2016). Lipid siderophores biosynthesized by NRPSs have been found in several *Halomonas* species (Martinez et al.,

2000; Homann et al., 2009; Figueroa et al., 2015), which often inhabit iron-poor environments. However, it remains unknown whether *Halomonas* sp. acquired the biosynthetic gene clusters of lipid siderophores by vertical or horizontal gene transfer. To explore this question, we first reconstructed a phylogenetic tree (Figure 7A) using the genomic DNA of *Halomonas* sp. MG34 as a query for the autoMLST server (<http://automlst.ziemertlab.com/>) (Alanjary et al., 2019), which can quickly provide related strains and annotate their secondary metabolite types. NRPSs were detected in only a few *Halomonas* species (Figure 7A), suggesting that lipid siderophore BGCs were likely acquired by *Halomonas* species via horizontal gene transfer. Then, to better understand the distribution of the potashchelin BGC in *Halomonas*, which is, to the best of our knowledge, the first reported lipid siderophore BGC from this genus, we used this BGC as a multigene BLAST query against all *Halomonas* genome sequences containing a NRP biosynthetic gene cluster indicated by autoMLST analysis (Figure 7B). The potashchelin biosynthesis gene cluster was not detected in any of the tested genomes, suggesting that it is relatively rare in *Halomonas* species. However, this analysis revealed some homologs of potashchelin BGC with highly conserved boundary regions, which might help to define the edges of the potashchelin BGC.



## Antibacterial Activities of Potashchelins

The antibacterial activity of potashchelins A–D (1–4) was assayed by the micro broth dilution method. Streptomycin, gentamycin,

vancomycin, and amphotericin B were selected as positive control treatments. As shown in **Table 2**, only potashchelins A (1) and D (4) exhibited weak antibacterial activity against *B. subtilis*

CPCC 100029 (MIC = 64  $\mu\text{g}/\text{mL}$ ), but they had no activity against *E. coli*, *P. aeruginosa*, *M. phlei*, *M. smegmatis*, *S. aureus*, or *C. albicans* (MIC > 64  $\mu\text{g}/\text{mL}$ ).

## DISCUSSION

PCR screening based on the conserved sequences of adenylation domain (A) in nonribosomal peptides (NRPs) has been successfully used for the discovery of NRPs. Degenerate primers A3F and A7R were designed to amplify NRPS gene sequences from actinomycetes and proved to be very useful for strain prioritization of actinomycetes, with positive rates ranging from 0% to 100% (Ayuso-Sacido and Genilloud, 2005; Lemetre et al., 2017; Masand et al., 2018; Zhou et al., 2018). In this study, using this degenerate primer pair, we successfully amplified NRPS genes from extremophiles and prioritized extremophile strains. The high positive rate of 100% demonstrated the efficacy of the degenerate primers in the amplification of NRPS genes from extremophiles and revealed the richness of the NRPSs in the extremophiles tested in this study.

Bioassay-guided prioritization and isolation are traditional methods for the discovery of new compounds with activity. It showed efficiency during the prioritization of tested strains in this study. However, bioassay-guided isolation did not seem to be successful, because the bioactivity of potashchelins A-D was not in line with that detected in the crude extract; the latter showed activity against *S. aureus*, while none of the former compounds exhibited activity against *S. aureus* at the concentration of less than 64  $\mu\text{g}/\text{mL}$ . We speculate that this divergence of bioactivity may be aroused by the weak activity of the metabolites against *S. aureus*, or the minor components in the crude extract were not detected, which is often encountered during the process of bioassay-guided isolation.

Lipid siderophores are often secreted by microbes to obtain iron from iron-scarce environments, and these molecules play an important role in the recycling of iron in the environment (Boiteau et al., 2016).  $\beta\text{-OH-Asp}$  is often found in lipid siderophores and takes part in covalent bonding with iron ions (Hardy and Butler, 2018).  $\beta\text{-OH-Asp}$  bears two stereocenters at C2 and C3, and it exists as four diastereomers: *L-threo*-, *D-threo*-, *L-erythro*-, and *D-erythro*- $\beta\text{-OH-Asp}$ . Most lipid siderophores contain one  $\beta\text{-OH-Asp}$  with either *L-erythro*-, *L-threo*-, or *D-threo* stereochemistry (Stephan et al., 1993; Agnoli et al., 2006; Kreutzer and Nett, 2012; Johnston et al., 2013; Reitz et al., 2019). Only serobactin (Kreutzer et al., 2012; Rosconi et al., 2013) and pacifibactin (Hardy and Butler, 2019) were reported to hold two  $\beta\text{-OH-Asp}$  with *LD-threo* diastereomers, while cupriachelin has been reported to bear *L-threo*- and *L-erythro*- $\beta\text{-OH-Asp}$  residues (Kreutzer et al., 2012; Reitz et al., 2019). However, the locations of the  $\beta\text{-OH-Asp}$  diastereomers of serobactin, pacifibactin, and cupriachelin have not been chemically determined. Here, we report that potashchelins contain both *L-threo*- and *L-erythro*- $\beta\text{-OH-Asp}$ , and we report the locations of these residues for the first time based on the results of partial hydrolysis and Marfey's analysis.

To determine the configuration of the four serine residues of the potashchelins, the potashchelin biosynthesis gene cluster was deduced by bioinformatics analysis. To the best of our

knowledge, this is the first reported lipid siderophore BGC from the genus *Halomonas*. Five NRPSs, encoded by the consecutive genes *ptcB*, *ptcC*, *ptcD*, *ptcE*, and *ptcF*, containing one loading and nine extending modules, are expected to be responsible for assembly of the potashchelin backbone. Importantly, we were able to deduce the stereochemistry of Ser based on our analysis of the C domains in modules 3, 5, 6, and 9 (Figure 1). Additionally, we further explored the evolutionary history of lipid siderophore BGCs using the genomic sequence of *Halomonas* sp. MG34 and the potashchelin BGC as a query for the autoMLST server and Multigeneblast analysis. This analysis suggested that the evolutionary mechanism of lipid siderophores in *Halomonas* involved horizontal transfer, despite the presence of lipid siderophores in several *Halomonas* species (Martinez et al., 2000; Homann et al., 2009; Figueroa et al., 2015).

Two  $\beta$ -hydroxylases (one stand-alone PtcA and one fused with the C-domain in PtcB) belonging to the TauD/TdfA family of nonheme Fe(II)/ $\alpha$ -ketoglutarate dependent dioxygenases (Singh et al., 2008) were discovered in the potashchelin BGC and proposed to be involved in hydroxylation of Asp at C-3. When homologs of the potashchelin BGC are considered together, it is interesting that the stereospecificity of PtcA and PtcB can be predicted from the phylogenetic analysis, which coincides with the results reported recently by Reitz et al. (2019). In contrast to the analysis of Reitz et al., we chose to analyze hydroxylases producing  $\beta\text{-OH-Asn}$ , as in the synthesis of curacomycin (Kaweewan et al., 2017) and ulleungmycin (Son et al., 2017), and  $\beta\text{-OH-Glu}$ , as in the synthesis of kutzneride (Strieker et al., 2009), as well as  $\beta\text{-OH-Asp}$ , as found in non-siderophore compounds such as malacidin A (Hover et al., 2018) and syringomycin (Gross and deVay, 1977). The phylogenetic analysis revealed that the prediction was meaningful only when the hydroxylases in the BGCs of siderophores and phytotoxins were taken into account. To better understand this phenomenon, we reconstructed the phylogenetic tree based on the 16S rRNA of the producing strains listed in Table S3 and Figure S30. It seems that the phylogenetic analysis based on  $\beta$ -hydroxylases protein sequences (Figure 6) are likely to emphasize genus-specific mutational evolution rather than functional evolution. Although the phylogeny reported by Reitz et al. and in this study can be used to illuminate the stereochemistry of  $\beta$ -hydroxylases in related genera, a more detailed phylogenetic analysis is needed to clarify the stereospecificity of  $\beta$ -hydroxylases derived from wider evolutionary origins.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in NCBI: genomic data, accession no. WJPH00000000 (<https://www.ncbi.nlm.nih.gov/nuccore/WJPH00000000>) and 16S rRNA sequence, accession no. MN636765 (<https://www.ncbi.nlm.nih.gov/nuccore/MN636765>).

## AUTHOR CONTRIBUTIONS

YX and BH initiated and oversaw all research. GZ provided the isolated extremophiles. NH cultured all strains. LL performed

the screening and potashchelin isolation. YL performed Marfey's analysis and partial hydrolysis. NH and WG extracted genomic DNA and carried out bioassays. YX performed data analysis, structure elucidation, genome assembly, and bioinformatics analysis. YX and YL co-drafted the manuscript. BH improved the manuscript.

## FUNDING

This research was co-funded by The Drug Innovation Major Project (2018ZX09711001-007-001), The CAMS Innovation Fund for Medical Sciences (2016-I2M-2-002),

## REFERENCES

- Agnoli, K., Lowe, C. A., Farmer, K. L., Husnain, S. I., and Thomas, M. S. (2006). The ornibactin biosynthesis and transport genes of *Burkholderia cenocepacia* are regulated by an extracytoplasmic function sigma factor which is a part of the fur regulon. *J. Bacteriol.* 188, 3631–3644. doi: 10.1128/JB.188.10.3631-3644.2006
- Alanjary, M., Steinke, K., and Ziemert, N. (2019). AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Res.* 47, W276–W282. doi: 10.1093/nar/gkz282
- Ayuso-Sacido, A., and Genilloud, O. (2005). New PCR primers for the screening of NRPS and PKS-I systems in actinomycetes: detection and distribution of these biosynthetic gene sequences in major taxonomic groups. *Microb. Ecol.* 49, 10–24. doi: 10.1007/s00248-004-0249-6
- Balibar, C. J., Vaillancourt, F. H., and Walsh, C. T. (2005). Generation of D amino acid residues in assembly of arthrofactin by dual condensation/epimerization domains. *Chem. Biol.* 12, 1189–1200. doi: 10.1016/j.chembiol.2005.08.010
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021
- Boiteau, R. M., Mende, D. R., Hawco, N. J., McIlvin, M. R., Fitzsimmons, J. N., Saito, M. A., et al. (2016). Siderophore-based microbial adaptations to iron scarcity across the eastern Pacific Ocean. *Proc. Natl. Acad. Sci. U.S.A.* 113, 14237–14242. doi: 10.1073/pnas.1608594113
- Chen, R., Wong, H. L., and Burns, B. P. (2019). New approaches to detect biosynthetic gene clusters in the environment. *Med. (Basel)* 6:32. doi: 10.3390/medicines610032
- Cockerill, F., Wikler, M., Alder, J., Dudley, M., Eliopoulos, G., Ferraro, M., et al. (2012). *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria that Grow Aerobically: Approved Standard*. Wayne, PA: Clinical and Laboratory Standards Institute.
- Figuerola, L. O., Schwarz, B., and Richards, A. M. (2015). Structural characterization of amphiphilic siderophores produced by a soda lake isolate, *Halomonas* sp. SL01, reveals cysteine-, phenylalanine- and proline-containing head groups. *Extremophiles* 19, 1183–1192. doi: 10.1007/s00792-015-0790-x
- Flugel, R. S., Hwangbo, Y., Lambalot, R. H., Cronan, J. E. Jr., and Walsh, C. T. (2000). Holo-(acyl carrier protein) synthase and phosphopantetheinyl transfer in *Escherichia coli*. *J. Biol. Chem.* 275, 959–968. doi: 10.1074/jbc.275.2.959
- Frost, D. J., and Gunstone, F. D. (1975). The PMR analysis of non-conjugated alkenoic and alkenoic acids and esters. *Chem. Phys. Lipids* 15, 53–85. doi: 10.1016/0009-3084(75)90032-8
- Fujii, K., Ikai, Y., Mayumi, T., Oka, H., Suzuki, M., and Harada, K.-I. (1997a). A nonempirical method using lc/ms for determination of the absolute configuration of constituent amino acids in a peptide: elucidation of limitations of Marfey's method and of its separation mechanism. *Anal. Chem.* 69, 3346–3352. doi: 10.1021/ac9701795
- Fujii, K., Ikai, Y., Oka, H., Suzuki, M., and Harada, K.-I. (1997b). A nonempirical method using lc/ms for determination of the absolute configuration of constituent amino acids in a peptide: combination of Marfey's method with mass spectrometry and its practical application. *Anal. Chem.* 69, 5146–5151. doi: 10.1021/ac970289b
- Gross, D., and deVay, J. (1977). Production and purification of syringomycin, a phytotoxin produced by *Pseudomonas syringae*. *Physiol. Plant Pathol.* 11, 13–28. doi: 10.1016/0048-4059(77)90083-2
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010
- Gunstone, F. D., Polard, M. R., Scrimgeour, C. M., and Vedanayagam, H. S. (1977). Fatty acids. Part 50. <sup>13</sup>C nuclear magnetic resonance studies of olefinic fatty acids and esters. *Chem. Phys. Lipids* 18, 115–129. doi: 10.1016/0009-3084(77)90031-7
- Hardy, C. D., and Butler, A. (2018).  $\beta$ -Hydroxyaspartic acid in siderophores: biosynthesis and reactivity. *J. Biol. Inorg. Chem.* 23, 957–967. doi: 10.1007/s00775-018-1584-2
- Hardy, C. D., and Butler, A. (2019). Ambiguity of NRPS structure predictions: four bidentate chelating groups in the siderophore pacifibactin. *J. Nat. Prod.* 82, 990–997. doi: 10.1021/acs.jnatprod.8b01073
- Homann, V. V., Sandy, M., Tincu, J. A., Templeton, A. S., Tebo, B. M., and Butler, A. (2009). Lohlichelins A-F, a suite of amphiphilic siderophores produced by the marine bacterium *Halomonas* LOB-5. *J. Nat. Prod.* 72, 884–888. doi: 10.1021/np800640h
- Hover, B. M., Kim, S. H., Katz, M., Charlop-Powers, Z., Owen, J. G., Ternei, M. A., et al. (2018). Culture-independent discovery of the malacidins as calcium-dependent antibiotics with activity against multidrug-resistant Gram-positive pathogens. *Nat. Microbiol.* 3, 415–422. doi: 10.1038/s41564-018-0110-1
- Huang, Y., Fan, H., Zhang, Z., Tong, Y., and Zhou, Y. (2013). Method of gap filling in microbial sequencing whole genome assembly. *Biotechnol. Lett.* 24, 819–821. doi: 10.3969/j.issn.1009-0002.2013.06.015
- Johnston, C. W., Wyatt, M. A., Li, X., Ibrahim, A., Shuster, J., Southam, G., et al. (2013). Gold biomineralization by a metallophore from a gold-associated microbe. *Nat. Chem. Biol.* 9, 241–243. doi: 10.1038/nchembio.1179
- Kawewuan, I., Komaki, H., Hemmi, H., and Kodani, S. (2017). Isolation and structure determination of new antibacterial peptide curacomycin based on genome mining. *Asian J. Org. Chem.* 6, 1838–1844. doi: 10.1002/ajoc.201700433
- Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000). *Practical Streptomyces Genetics: A Laboratory Manual*. Norwich: John Innes Foundation.
- Kreutzer, M. F., Kage, H., and Nett, M. (2012). Structure and biosynthetic assembly of cupriachelin, a photoreactive siderophore from the bioplastic producer *Cupriavidus necator* H16. *J. Am. Chem. Soc.* 134, 5415–5422. doi: 10.1021/ja300620z
- Kreutzer, M. F., and Nett, M. (2012). Genomics-driven discovery of taiwachelin, a lipopeptide siderophore from *Cupriavidus taiwanensis*. *Org. Biomol. Chem.* 10, 9338–9343. doi: 10.1039/c2ob26296g
- Lemetre, C., Maniko, J., Charlop-Powers, Z., Sparrow, B., Lowe, A. J., and Brady, S. F. (2017). Bacterial natural product biosynthetic domain composition in soil correlates with changes in latitude on a continent-wide scale. *Proc. Natl. Acad. Sci. U.S.A.* 114, 11615–11620. doi: 10.1073/pnas.1710262114
- and The National Natural Science Foundation of China (81603007).

## ACKNOWLEDGMENTS

We thank Biomarker Technologies for sequencing the genome DNA sequence.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fchem.2020.00197/full#supplementary-material>

- Liu, B., Li, J., Chen, M., Hao, X., Cao, F., Tan, Y., et al. (2018). Seco-Tetracenomycins from the marine-derived actinomycete *saccharothrix* sp. 10-10. *Mar Drugs* 16:345. doi: 10.3390/md16100345
- Martinez, J. S., Zhang, G. P., Holt, P. D., Jung, H. T., Carrano, C. J., Haygood, M. G., et al. (2000). Self-assembling amphiphilic siderophores from marine bacteria. *Science* 287, 1245–1247. doi: 10.1126/science.287.5456.1245
- Masand, M., Sivakala, K. K., Menghani, E., Thinesh, T., Anandham, R., Sharma, G., et al. (2018). Biosynthetic potential of bioactive streptomycetes isolated from arid region of the Thar Desert, Rajasthan (India). *Front. Microbiol.* 9:687. doi: 10.3389/fmicb.2018.00687
- Nikodinovic, J., Barrow, K. D., and Chuck, J. A. (2003). High yield preparation of genomic DNA from *Streptomyces*. *Biotechniques* 35, 932–934, 936. doi: 10.2144/03355bm05
- Reitz, Z. L., Hardy, C. D., Suk, J., Bouvet, J., and Butler, A. (2019). Genomic analysis of siderophore  $\beta$ -hydroxylases reveals divergent stereocontrol and expands the condensation domain family. *Proc. Natl. Acad. Sci. U.S.A.* 116, 19805–19814. doi: 10.1073/pnas.1903161116
- Rosconi, F., Davy, D., Martinez, V., Martinez, M., Abin-Carriquiry, J. A., Zane, H., et al. (2013). Identification and structural characterization of serobactins, a suite of lipopeptide siderophores produced by the grass endophyte *Herbaspirillum seropedicae*. *Environ. Microbiol.* 15, 916–927. doi: 10.1111/1462-2920.12075
- Rothschild, L. J., and Mancinelli, R. L. (2001). Life in extreme environments. *Nature* 409, 1092–1101. doi: 10.1038/35059215
- Schwarzer, D., Mootz, H. D., Linne, U., and Marahiel, M. A. (2002). Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. *Proc. Natl. Acad. Sci. U.S.A.* 99, 14083–14088. doi: 10.1073/pnas.2123.82199
- Singh, G. M., Fortin, P. D., Koglin, A., and Walsh, C. T. (2008).  $\beta$ -Hydroxylation of the aspartyl residue in the phytotoxin syringomycin E: characterization of two candidate hydroxylases AspH and SyrP in *Pseudomonas syringae*. *Biochemistry* 47, 11310–11320. doi: 10.1021/bi801322z
- Son, S., Hong, Y. S., Jang, M., Heo, K. T., Lee, B., Jang, J. P., et al. (2017). Genomics-driven discovery of chlorinated cyclic hexapeptides ulleungmycins A and B from a *Streptomyces* species. *J. Nat. Prod.* 80, 3025–3031. doi: 10.1021/acs.jnatprod.7b00660
- Stephan, H., Freund, S., Beck, W., Jung, G., Meyer, J. M., and Winkelmann, G. (1993). Ornibactins—a new family of siderophores from *Pseudomonas*. *Biometals* 6, 93–100. doi: 10.1007/BF00140109
- Strieker, M., Nolan, E. M., Walsh, C. T., and Marahiel, M. A. (2009). Stereospecific synthesis of threo- and erythro-beta-hydroxyglutamic acid during kutzneride biosynthesis. *J. Am. Chem. Soc.* 131, 13523–13530. doi: 10.1021/ja9054417
- Sussmuth, R. D., and Mainz, A. (2017). Nonribosomal peptide synthesis—principles and prospects. *Angew. Chem. Int. Ed Engl.* 56, 3770–3821. doi: 10.1002/anie.201609079
- Vijayasarathy, S., Prasad, P., Fremlin, L. J., Ratnayake, R., Salim, A. A., Khalil, Z., et al. (2016). C3 and 2D C3 Marfey's methods for amino acid analysis in natural products. *J. Nat. Prod.* 79, 421–427. doi: 10.1021/acs.jnatprod.5b01125
- Wilson, Z. E., and Brimble, M. A. (2009). Molecules derived from the extremes of life. *Nat. Prod. Rep.* 26, 44–71. doi: 10.1039/B800164M
- Zhang, W., Heemstra, J. R. Jr., Walsh, C. T., and Imker, H. J. (2010). Activation of the pacidamycin PacL adenylation domain by MbtH-like proteins. *Biochemistry* 49, 9946–9947. doi: 10.1021/bi101539b
- Zhang, X., Li, S. J., Li, J. J., Liang, Z. Z., and Zhao, C. Q. (2018). Novel natural products from extremophilic fungi. *Mar. Drugs* 16:194. doi: 10.3390/md16060194
- Zhou, M., Liu, F., Yang, X., Jin, J., Dong, X., Zeng, K. W., et al. (2018). Bacillibactin and bacillomycin analogues with cytotoxicities against human cancer cell lines from marine *Bacillus* sp. PKU-MA00093 and PKU-MA00092. *Mar. Drugs* 16:22. doi: 10.3390/md16010022

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Li, Liu, Zhang, He, Guo, Hong and Xie. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.