Supplementary material

FW0622, a New Siderophore Isolated from Marine *Verrucosispora sp.* by Genome Mining

Wei Zhao [a], Fei Peng [b], Chuan-xi Wang[c], Yang Xie [a], Ru Lin [a], Zhi-kai Fang [a], Fei Sun[a] , Yun-yang Lian[a] and Jiang Hong[a*]

[a]*Fujian Provincial Key Laboratory of Screening for Novel Microbial Products, Fujian Institute of Microbiology, Fuzhou 350007, P.R. China;* [b] *College of Pharmacy, Quanzhou Medical College, Quanzhou 362100, P.R. China;* [c]*College of Pharmacy, Jinan Univercity, Guangzhou 510632, P.R. China;*

Using the draft genome sequence of *Verrucosispora* sp. FIM060022, we have identified a new desferrioxamine-like siderophore, FW0622. This is the first chemically characterized siderophore obtained from *Verrucosispora*. The structure was elucidated by extensive spectral analyses. The biosynthetic pathway of FW0622 was proposed to be via the non-ribosomal peptide synthetase (NRPS)-independent (NIS) synthetase pathway based on the putative biosynthetic siderophore gene clusters in FIM060022. The results from this study demonstrate that marine-derived Verrucosispora species deserve recognition as an important source of new natural products. Furthermore, this study verified that genome mining is an effective way to thoroughly identify compounds that may be overlooked by traditional methods.

Key words: genome mining; siderophore; biosynthesis pathway; *Verrucosispora*.

Experimental Section

Draft genome sequence and analysis

Genome sequencing and assembly

High quality genomic DNA from FIM060022 for sequencing was prepared with an EZ1 DNA Tissue Kit and a Bio Robot EZ1 (Qiagen). Shotgun and paired-end libraries were prepared and subsequently sequenced using next generation sequencing (NGS) technology using the Illumina Miseq (PE400) sequencing platform. The 2,845,130 high-quality reads and 632,396,458 bp were assembled using A5-Miseq (Coil D et al. 2015) to yield 29 scaffolds larger than 1,164 bp. The N50 was 453,524 bp.

Genome scan to identify biosynthetic gene cluster of secondary metabolism products

The underlying biosynthesis gene clusters of secondary metabolism products in the strain FIM0600222 were predicted using antiSMASH and BLASTP searches against the NCBI nr databases for predicting function of proteins encoded in the siderophore biosynthetic gene cluster.

Bioassay of antimicrobial activities

The antimicrobial activities of compound FW0622 were evaluated by an agar dilution method (Zaika et al. 1988). The test strains included bacteria (*Staphylococcus aureus*, *Bacillus subtilis*, *Micrococcus luteus,* and *Escherichia coli*) and fungi (*Candida albicans* and *Aspergillus niger*). The minimum inhibitory concentration (MIC) values were defined as the lowest concentration at which no microbial growth was observed. The assays were run in triplicate.

General experimental procedures

Nuclear magnetic resonance (NMR) experiments were conducted on a Bruker Avance III spectrometer (400 MHz). Optical rotations were measured on a JASCO P1020 digital polarimeter. HR-ESI-MS (High Resolution Electrospray Ionization Mass Spectrometry) were recorded on an Aligent Quattro Premier Triple Quadrupole mass spectrometer and a Bruker MicroToF-QII mass spectrometer, respectively. Thin-layer chromatography (TLC) was carried out on silica gel GF254 (Qingdao Marine Chemical Ltd., China) plates. Column chromatography (CC) was carried out on silica gel (100-200 mesh; Qingdao Marine Chemical Ltd., China), Sephadex LH-20 (50 mm; Amersham Pharmacia Biotech, Sweden) and ODS (Octadecylsilyl) (60–80 mesh; YMC Ltd., Japan). All solvents used in TLC and CC analytical grade (Tianjin Damao Chemical Plant, Tianjin, China).

Fermentation of Verrucosispora sp. FIM060022

A fermentation experiment was conducted to isolate compound FW0622. To prepare seed inoculum, 1.0 ml of a thawed spore suspension was inoculated into a 500 ml Erlenmeyer flask containing 80 ml of seed medium consisting of 4.0% starch, 0.5% glucose, 0.5% peptone, 0.5% yeast extract, 0.05% $(NH_4)_2SO_4$, 0.05%

NaCl, 0.05% K_2HPO_4, 0.05% $MgSO_4$ •$7H_2O$, and 0.1% $CaCO_3$ in distilled water (DI) at pH 7.5. After inoculation for 48 h at 32°C on a rotary shaker at 240 rpm, the mycelia suspension (5%) was transferred into 500 ml Erlenmeyer flasks, each containing 80 ml fermentation medium composed of 2.0% oatmeal, 0.2% peptone, 0.025% (NH4)2SO4, 0.025% NaCl, 0.1% CaCO3, and trace elements (0.1% $FeSO_4$ •$7H_2O$, 0.1% $MnCl_2$ •$4H_2O$, and 0.1% $ZnSO_4$•$7H_2O$ 1mL/L) in DI water at pH 7.5 for 120 h at 32°C on a rotary shaker 240 rpm.

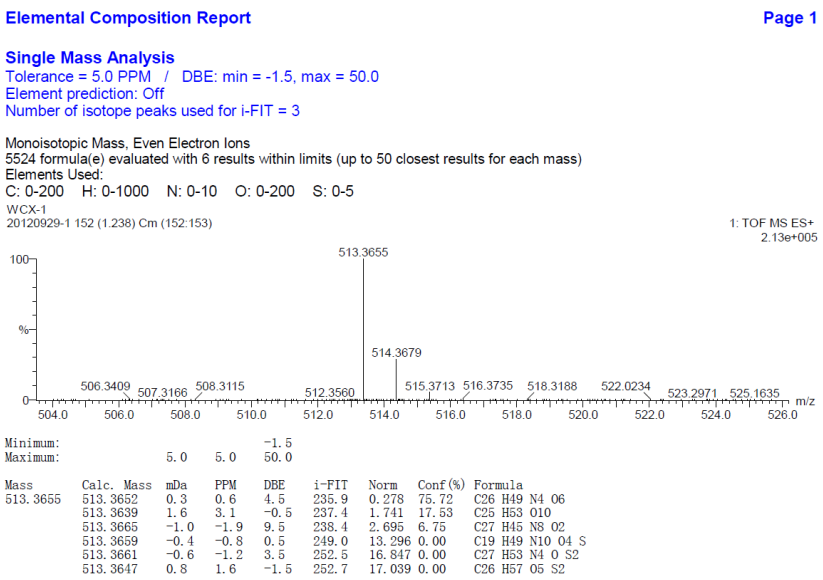
Extraction and isolation

Culture broth was centrifuged to obtain mycelium, which was then soaked with methanol. The methanol solution was concentrated under vacuum to a dry residue, which was dissolved in H_2O, forming a suspension. Subsequently, the mixture of the above suspension and the added supernatant was extracted repeatedly with ethyl acetate (Ea). The Ea layer was evaporated to dryness under vacuum yielding the crude sample (1.9 g), which was fractioned by silica gel CC using a $CHCl_3$-MeOH gradient elution to obtain three fractions (F1-F3). The fraction F3 (0.28g) was subjected to Sephadex LH-20 CC and eluted with $CHCl_3$: MeOH (1:1) in order to yield fractions (F3-1-F3-3). F3-3 was subjected to an ODS medium-pressure liquid chromatography with a gradient of MeOH in H_2O (5-100% for 1 h with a flow rate of 2 mL/min). Fractions were evaporated to dryness (FW0622 12mg).

MS, IR and TLC figures

FW0622: white amorphous powder; [α]20 D 4.0 (c 0.5, MeOH); UV (MeOH) λmax (log ε) No maximum above 210; IR (neat) ν_{max} 3303, 3099, 3009, 2929, 2855, 1620, 1567, 1461, 1426, 1396, 732 cm^{-1}; 1H, 13C NMR, and HMBC (Heteronuclear Multiple Bond Correlation) data see Table S1; HRESIMS m/z 513.3655 $[M + H]^+$ (calculated. for $C_{26}H_{48}N_4O_6$, 513.3652).

Mass spectrum



Elemental Composition Report **Page 1**

Single Mass Analysis
Tolerance = 5.0 PPM / DBE: min = -1.5, max = 50.0
Element prediction: Off
Number of isotope peaks used for i-FIT = 3

Monoisotopic Mass, Even Electron Ions
5524 formula(e) evaluated with 6 results within limits (up to 50 closest results for each mass)
Elements Used:
C: 0-200 H: 0-1000 N: 0-10 O: 0-200 S: 0-5

WCX-1
20120929-1 152 (1.238) Cm (152:153)

1: TOF MS ES+
2.13e+005

Minimum:				-1.5				
Maximum:		5.0	5.0	50.0				
Mass	Calc. Mass	mDa	PPM	DBE	i-FIT	Norm	Conf(%)	Formula
513.3655	513.3652	0.3	0.6	4.5	235.9	0.278	75.72	C26 H49 N4 O6
	513.3639	1.6	3.1	-0.5	237.4	1.741	17.53	C25 H53 O10
	513.3665	-1.0	-1.9	9.5	238.4	2.695	6.75	C27 H45 N8 O2
	513.3659	-0.4	-0.8	0.5	249.0	13.296	0.00	C19 H49 N10 O4 S
	513.3661	-0.6	-1.2	3.5	252.5	16.847	0.00	C27 H53 N4 O S2
	513.3647	0.8	1.6	-1.5	252.7	17.039	0.00	C26 H57 O5 S2

Figure S1 HRESI-MS spectrum of FW0622

IR spectrum

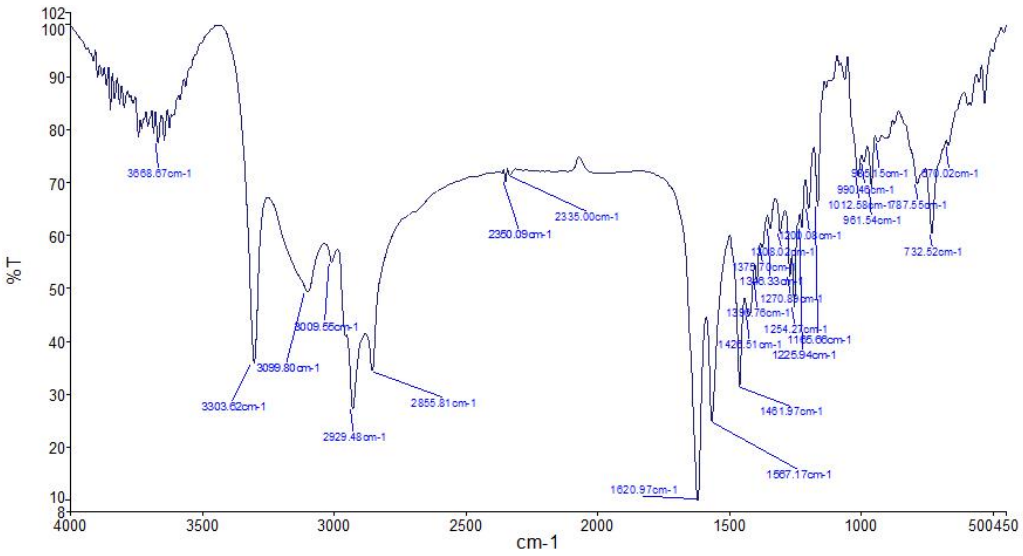


Figure S2 IR spectrum of FW0622

TLC figure

Figure S3 Reaction of hydroxamic acid iron with FW0622

1D, 2D NMR spectra of FW0622 and structural determination

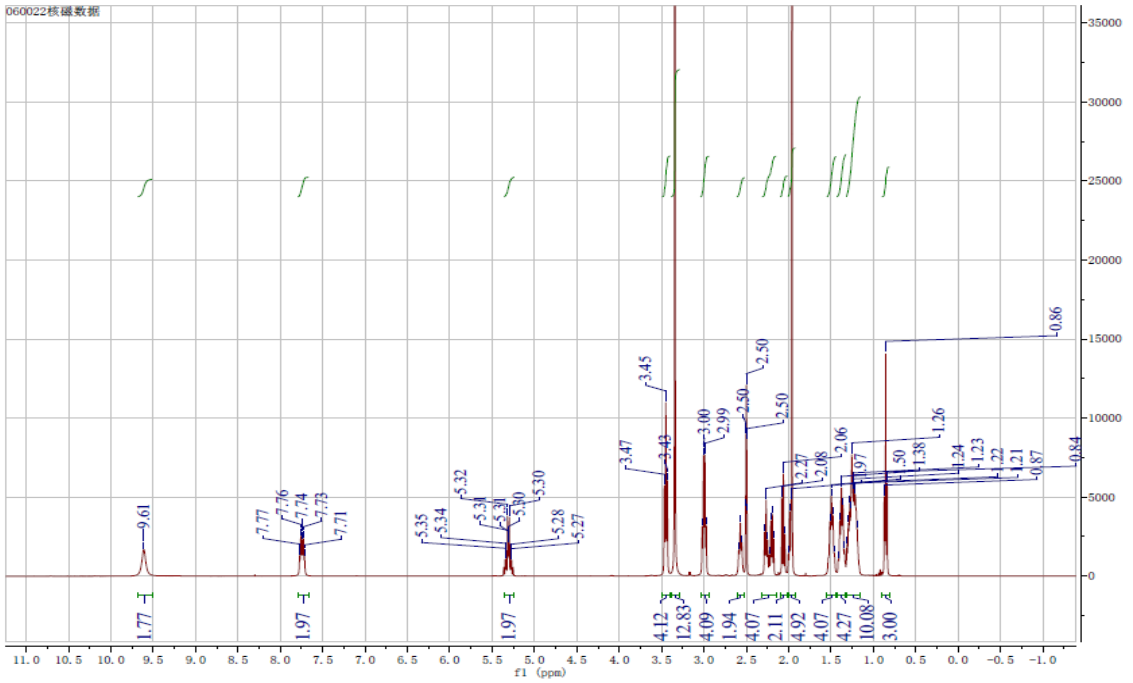


Figure S4 ^{1}H-NMR spectrum of FW0622 (400 MHz, DMSO-d6)

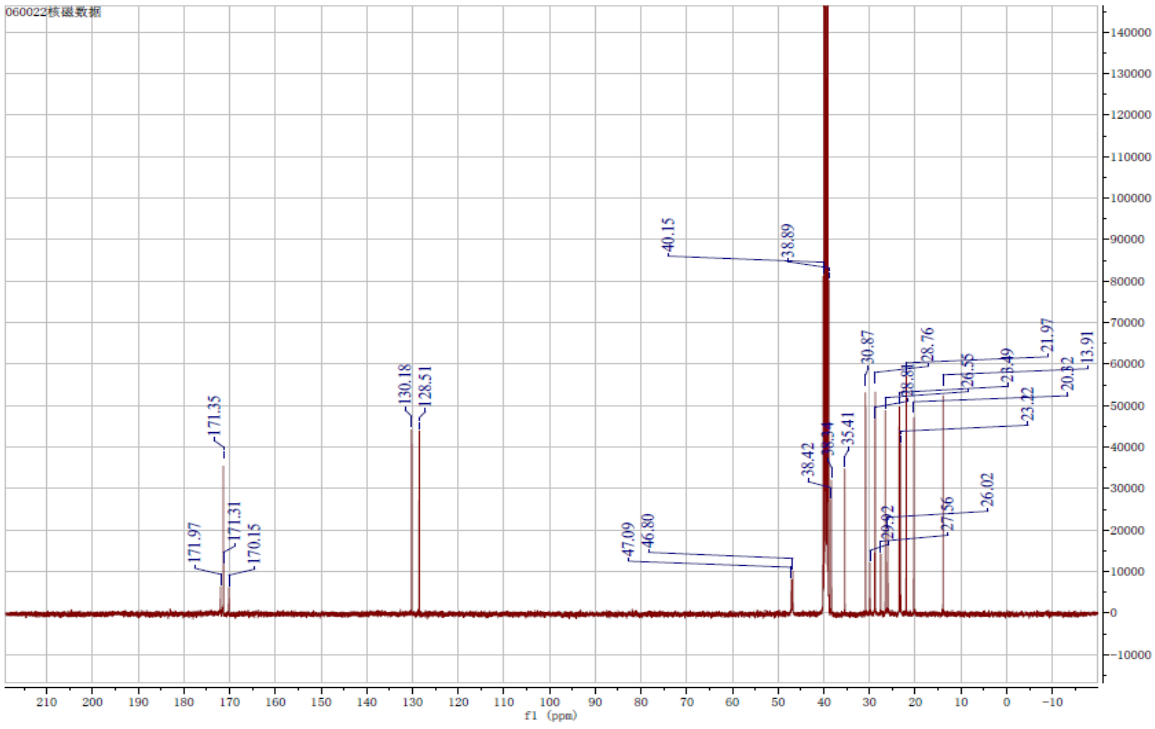


Figure S5 ^{13}C-NMR spectrum of FW0622 (100 MHz, DMSO-d6)

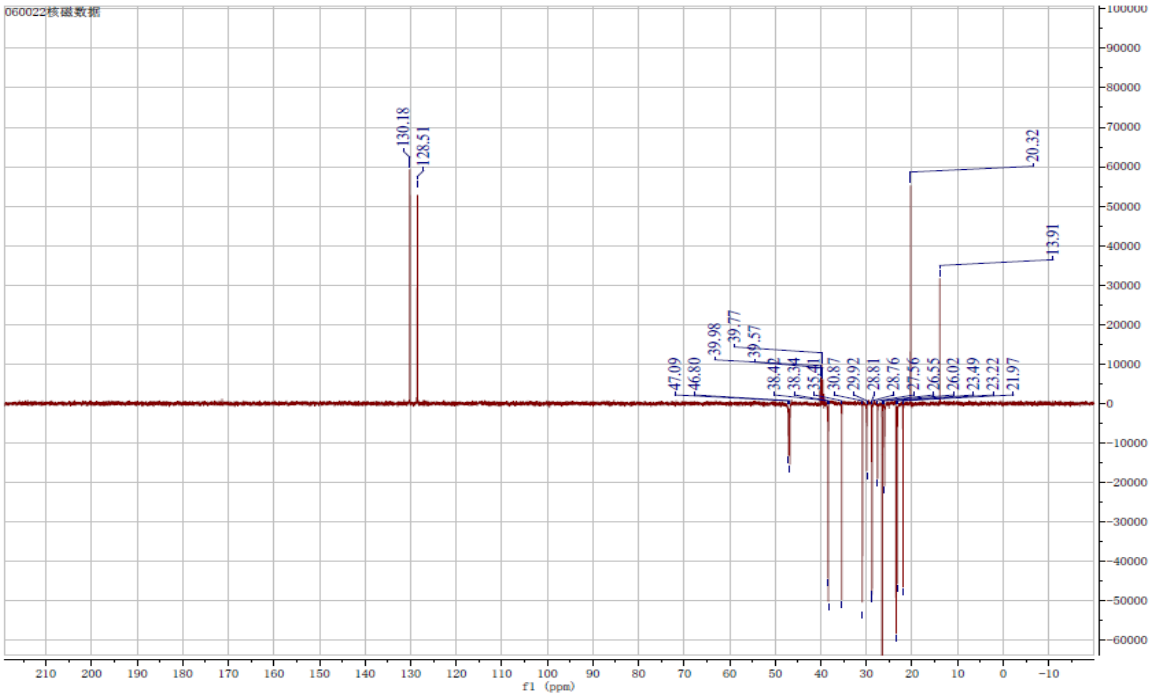


Figure S6 DEPT135 spectrum of FW0622 (100 MHz, DMSO-d6)

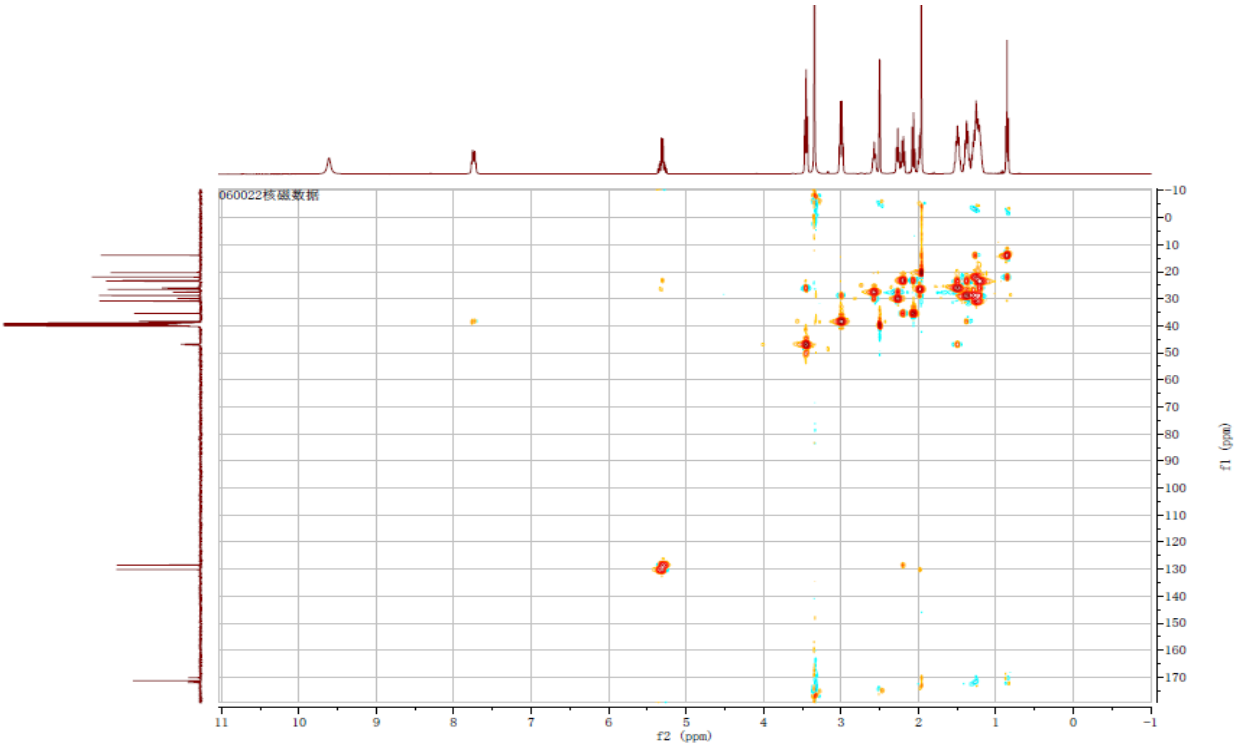


Figure S7 HSQC spectrum of FW0622 (DMSO-d6)

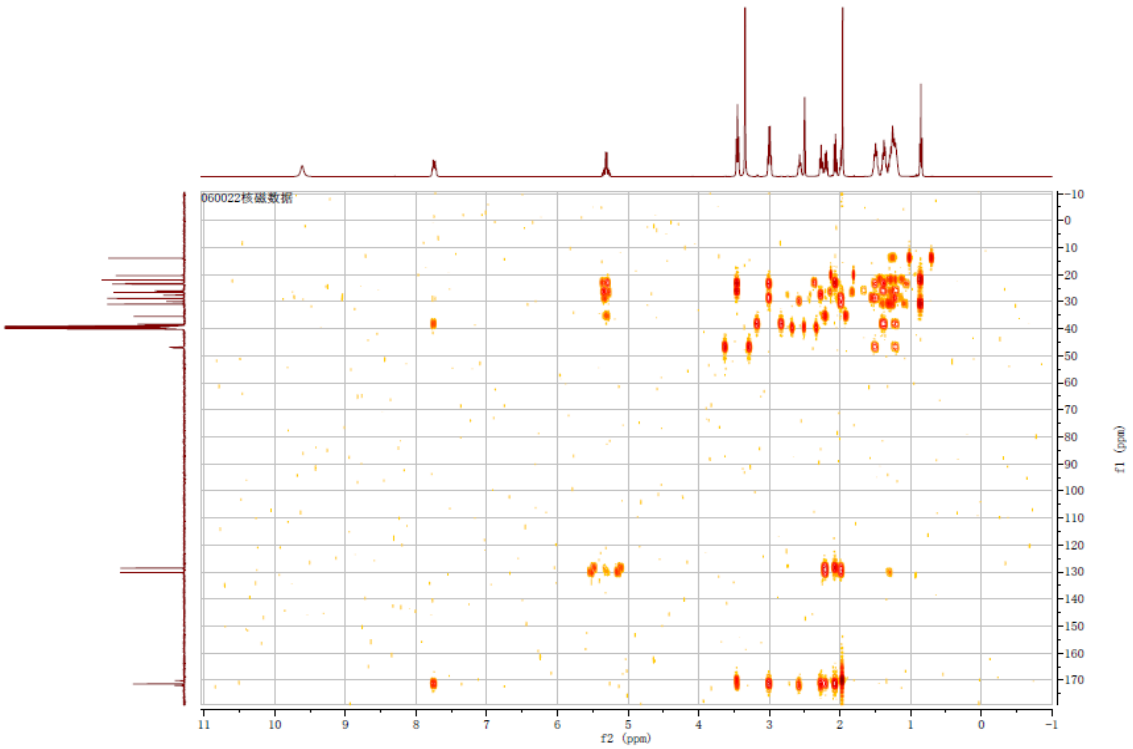


Figure S8 HMBC spectrum of FW0622 (DMSO-d6)

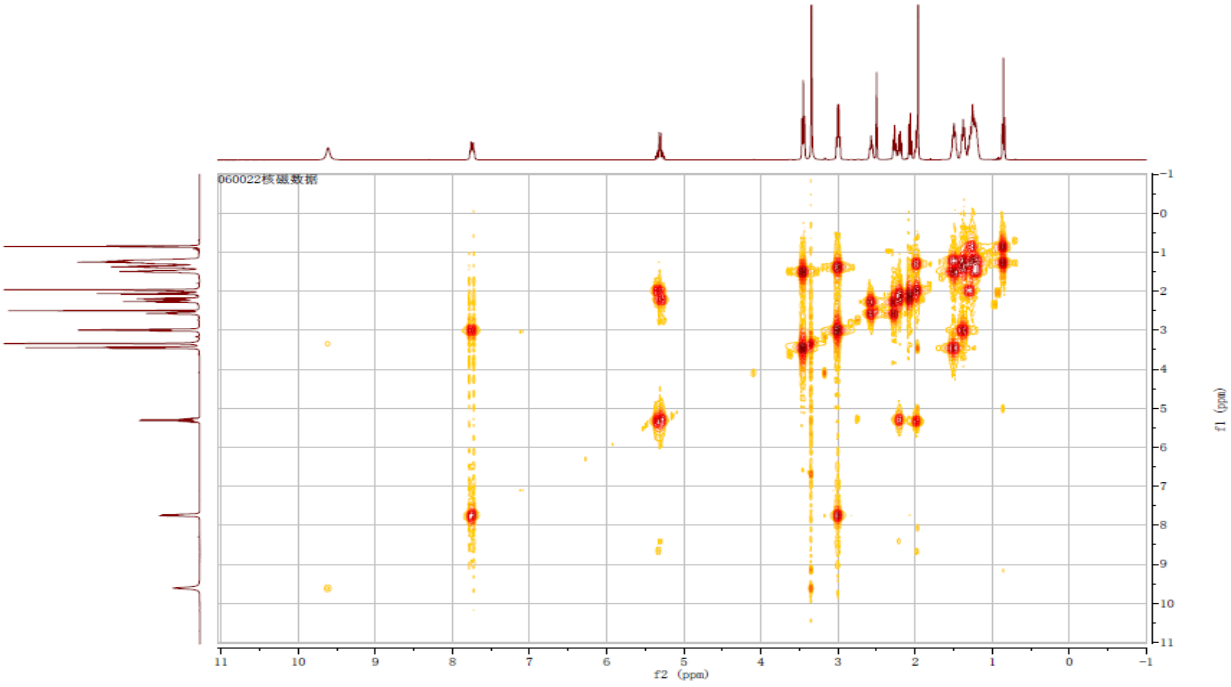


Figure S9 ^{1}H-^{1}H COSY spectrum of FW0622 (DMSO-d6)

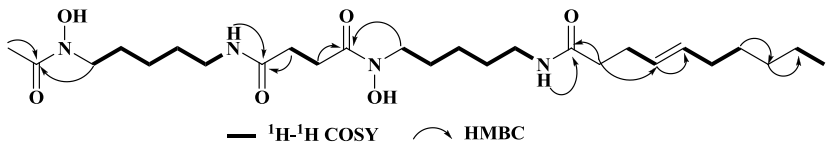


Figure S10 Selected ^{1}H-^{1}H COSY and HMBC correlations of FW0622

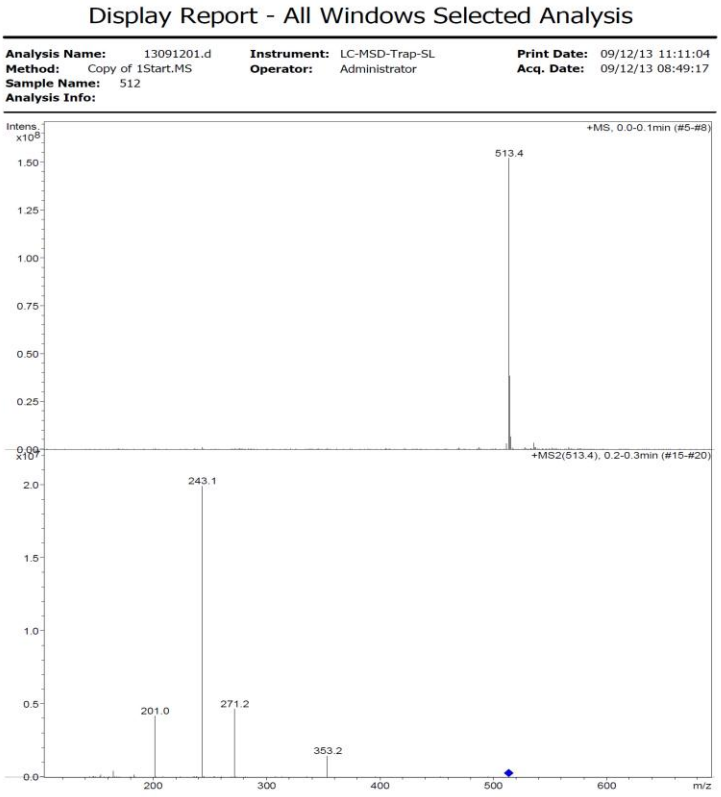


Figure S11 FAB-MS/MS spectrum of FW0622

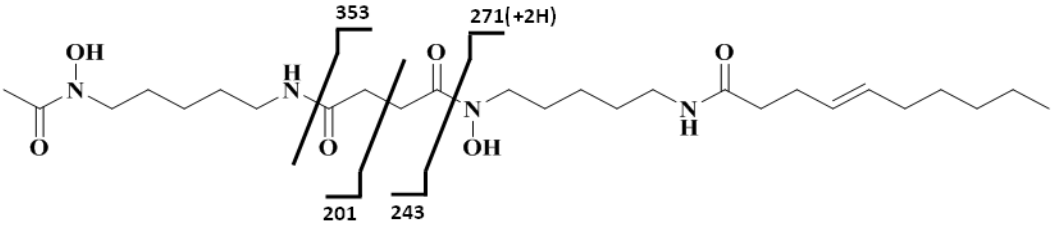


Figure S12 FAB-MS/MS analysis of FW0622

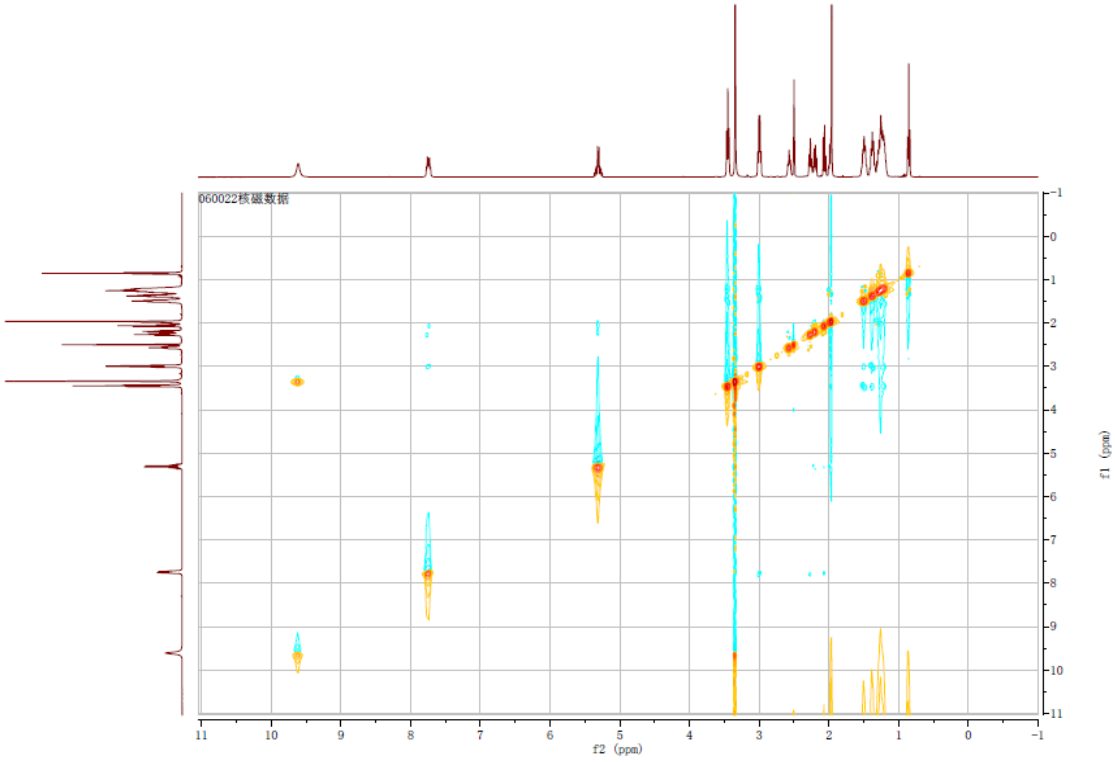


Figure S13 NOESY (Nuclear Overhauser Effect Spectroscopy) spectrum of FW0622 (DMSO-d6)

Relative positions of putative functional genes in the siderophores biosynthetic gene clusters

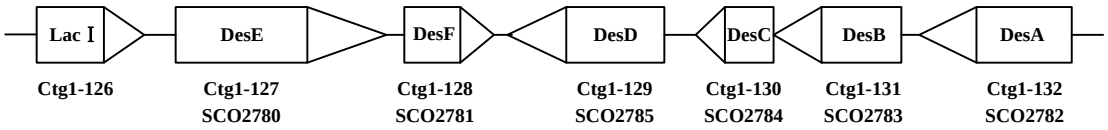


Figure S14. The relative position of enzymes within the desferrioxamine-like biosynthetic gene clusters

Genome sequencing and information

Table S1 Gene clusters involved in biosynthesis of siderophores in *Verrucosispora sp* FIM060022

Cluster name	Scaffold No.	Position (nt)	MIBIG BGC-ID	Gene cluster type	Most similar known cluster
Cluster 5	1	1061600-1084200	BGC0000551_c1	Lantipeptide	SapB biosynthetic gene cluster
Cluster 8	2	647671-688720	BGC0001077_c1	T3pks	Alkyl-O-Dihydrogeranyl-Methoxyhydroquinones biosynthetic gene
Cluster 13	9	19193- 40104	BGC0001087_c4	Terpene	Sioxanthin biosynthetic gene cluster
Cluster 15	11	127607-139415	BGC0000940 c1	Siderophore	Desferrioxamine B biosynthetic gene cluster

Cluster 19	14	74679- 108726	BGC0001296_c1	T1pks	Streptazone E biosynthetic gene cluster

NMR data assignments and physicochemical properties of FW0622

Table S2 NMR data of FW0622 in DMSO-d6 (400 MHz for 1H; 100 MHz for 13C)

No.	Compound FW0622			
	δC (ppm)	δH (ppm)	^{1}H-^{1}H COSY	HMBC (H→C)
1	20.3	1.97 (3H, s)		C-2
2	170.1	-		
3-N	-	9.62 (1H, s)		
4	46.8	3.46 (2H, t, J = 7.0 Hz)	H-5	C-2, C-5, C-6
5	26.0	1.50 (2H, m)	H-4, H-6	C-4, C-6, C-7
6	23.5	1.22 (2H, m)	H-5	C-4, C-5, C-8
7	28.8	1.38 (2H, m)	H-8	C-5, C-6, C-8
8	38.4	3.01 (2H, m)	H-7, H-9	C-6, C-7, C-10
9-N	-	7.73 (1H, m)	H-8	C-8, C-10
10	171.3	-		
11	29.9	2.27 (2H, t, J = 7.2 Hz)	H-12	C-10, C-12, C-13
12	27.6	2.58 (2H, t, J = 7.2 Hz)	H-11	C-10, C-11, C-13
13	172.0	-		
14-N	-	9.62 (1H, s)		
15	47.1	3.46 (2H, t, J = 7.0 Hz)	H-16	C-13, C-17, C-16
16	26.0	1.50 (2H, m)	H-15, H-17	C-15, C-17, C-18
17	23.5	1.22 (2H, m)	H-16	C-15, C-16, C-19
18	28.7	1.38 (2H, m)	H-19	C-16, C-17, C-19
19	38.3	3.01 (2H, m)	H-18, H-20	C-17, C-18, C-21
20-N	-	7.76 (1H, m)	H-19	C-19, C-21
21	171.4	-		
22	35.4	2.07 (2H, t, J = 7.1 Hz)		C-21, C-24
23	23.2	2.20 (2H, m)	H-24	C-21, C-24
24	128.5	5.31 (1H, m)	H-23	C-22, C-23, C-25
25	130.2	5.32 (1H, m)	H-26	C-24, C-26
26	26.5	1.98 (2H, m)	H-25, H-27	C-25, C-27, C-28
27	29.2	1.30 (2H, m)	H-26	C-25, C-28

28	30.8	1.26 (2H, m)		C-29
29	21.9	1.27 (2H, m)	H-30	C-30
30	13.9	0.86 (3H, t, J = 7.1 Hz)	H-29	C-28, C-29

Table S3 Physicochemical properties of compound FW0622

Content	Compound 1
Appearance	White amorphous powder
Melting point	143.3-145.4℃
Molecular formula	$C_{26}H_{48}N_4O_6$
HRFAB-MS(m/z)	Calc. mass: 513.3652
$(M+H)^+$	Found mass: 513.3655
UV λ_{max} nm	No maximum above 210
IR v_{max} (KBr) cm^{-1}	3303, 3099, 3009, 2929, 2855 1620, 1567, 1461, 1426, 1396, 732
Optical rotation	$[\alpha]_{20}$ D 4.0 (c 0.5, MeOH)

Table S4. Open reading frames in the desferrioxamine-like biosynthetic gene cluster of *Verrucosispora* sp. FIM060022

ORF (locus tag)	Size (aa)	Deduced function	Position(nt)	Protein homolog	*Verrucosispora*	
					Identity (%)	Accession number
Ctg1-126	338		127790-128806	Lac Ⅰ family transcription regulator	99	AEB45158.1
Ctg1-127	900	DesE	128851-131553	Secreted Protein (iron compound ABC transporter)	95	WP_013733812.1
Ctg1-128	255	DesF	131590-132600	aminoglycoside phosphotransferase family protein	95	WP_013733811.1
Ctg1-129	602	DesD	132607-134415	IucA/IucC family siderophore biosynthesis protein	95	WP_013733808.1
Ctg1-130	192	DesC	134412-134490	Siderophore biosynthesis N-aceyltransferase protein	93	WP_013733807.1
Ctg1-131	426	DesB	134987-136267	lysine 6-monooxygenase(NADPH)	98	WP_013733806.1
Ctg1-132	507	DesA	136303-137826	pyridoxal-dependent decarboxylase	96	WP_013733805.1

References

Coil D, Jospin G, Darling AE. 2015. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. Bioinformatics. 31(4):587-9.

Zaika LL. 1988. Spices and herbs: their antimicrobial activity and its determination. J. Food Safety. 9: 97–118.

Figure captions:

Table Captions: