Supporting Information

Phormidepistatin from the Cyanobacterium UIC 10484 – Assessing the Phylogenetic Distribution of the Statine Pharmacophore

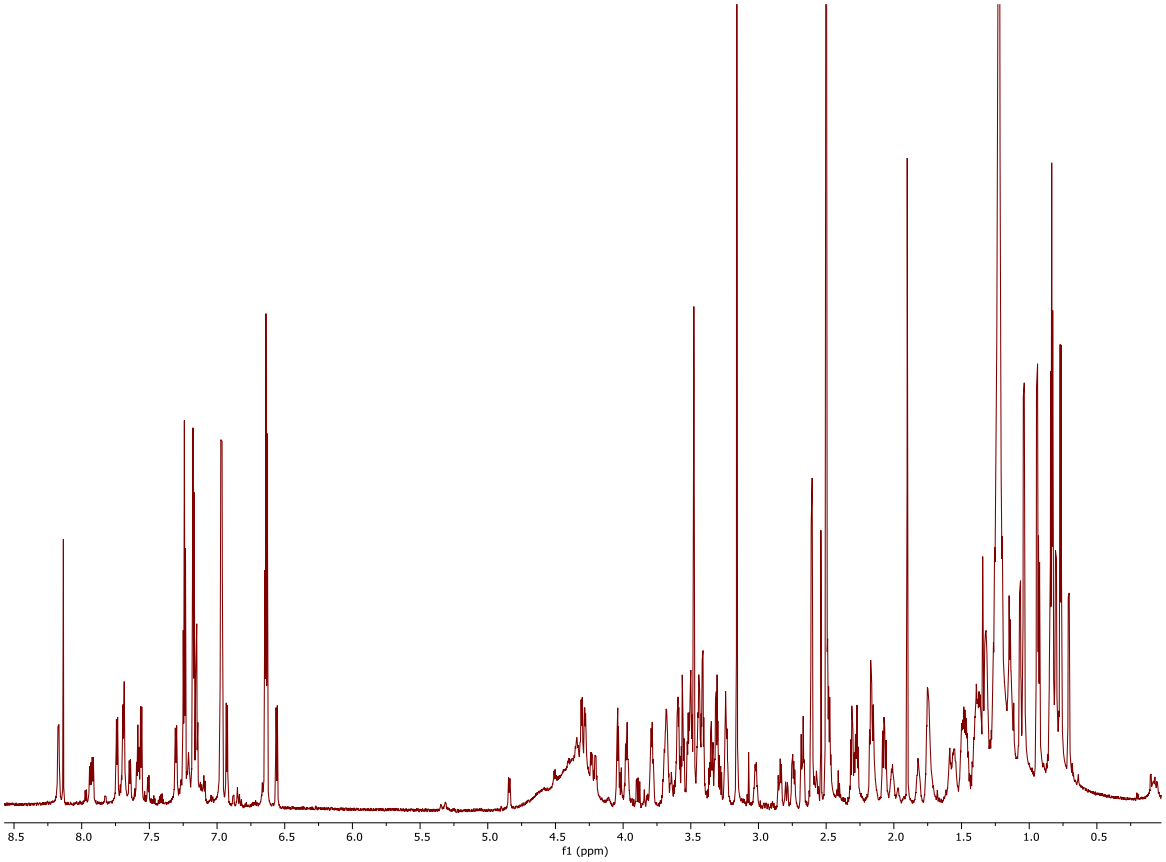
*Peter Sullivan, Aleksej Krunic, Lydia J. Davis, Hwan Seung Kim, Joanna E. Burdette, and Jimmy Orjala**

Department of Pharmaceutical Sciences, College of Pharmacy, University of Illinois at Chicago, Chicago, IL, 60612, United States

*correspondence to orjala@uic.edu

Table of Contents

Figure S1. ^{1}H NMR spectrum of phormidepistatin (900 MHz, T=300 K, DMSO-d_6 + TFA)



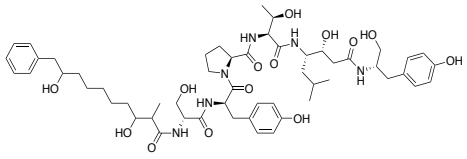
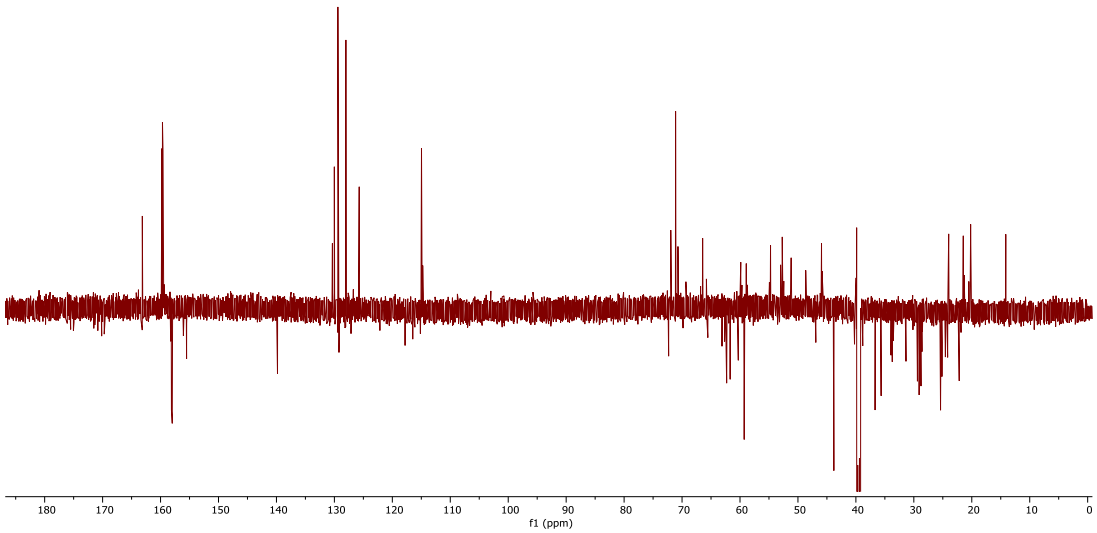


Figure S2. DEPTQ spectrum of phormidepistatin (226 MHz, T=300 K, DMSO-d_6 + TFA)



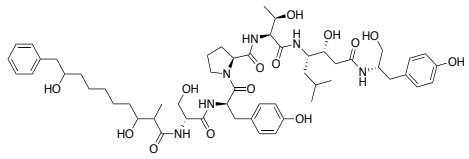


Figure S3. HSQC spectrum of phormidepistatin (900 MHz, T=300 K, DMSO-d_6 + TFA)

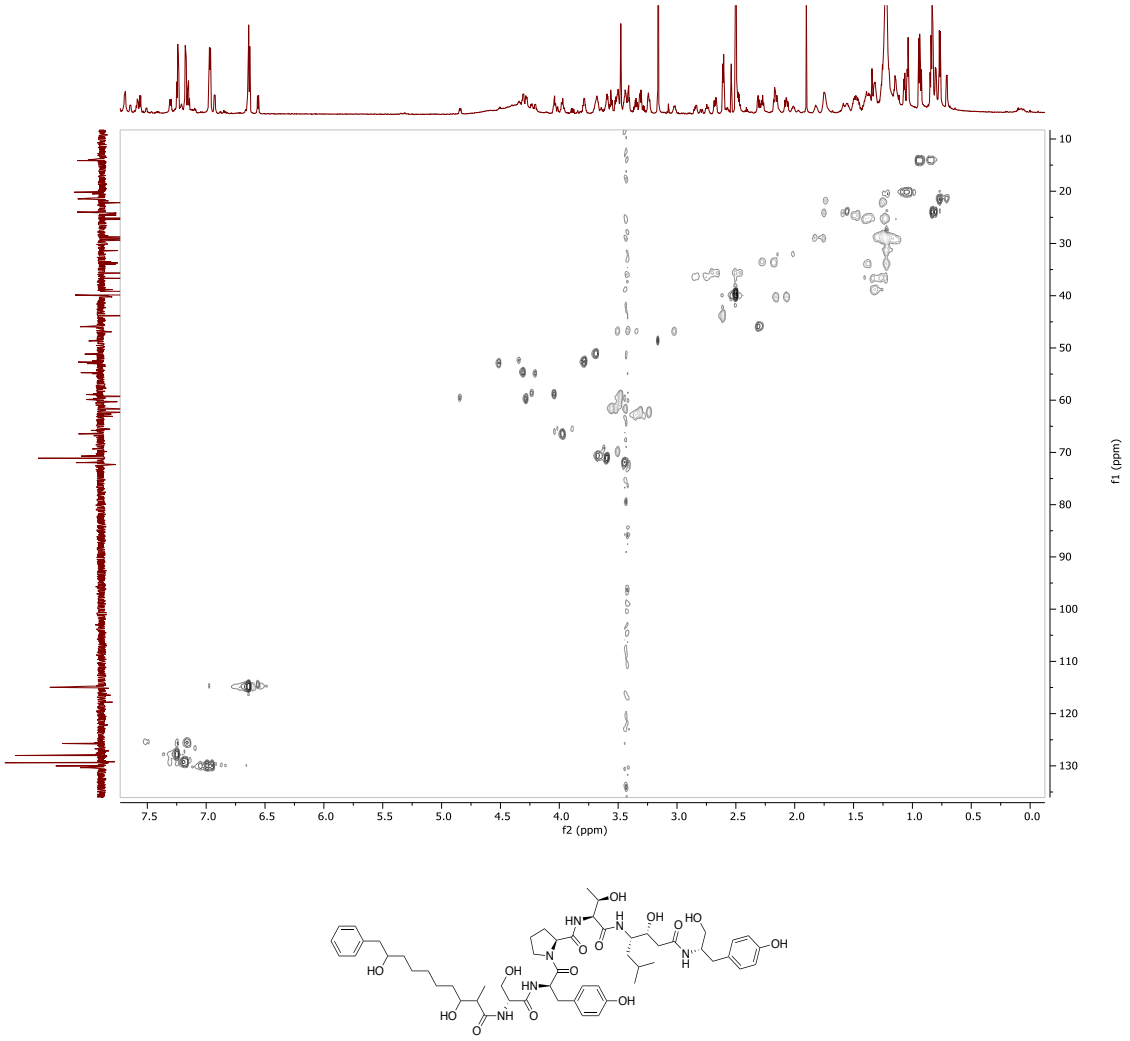


Figure S4. COSY spectrum of phormidepistatin (900 MHz, T=300 K, DMSO-d_6 + TFA)

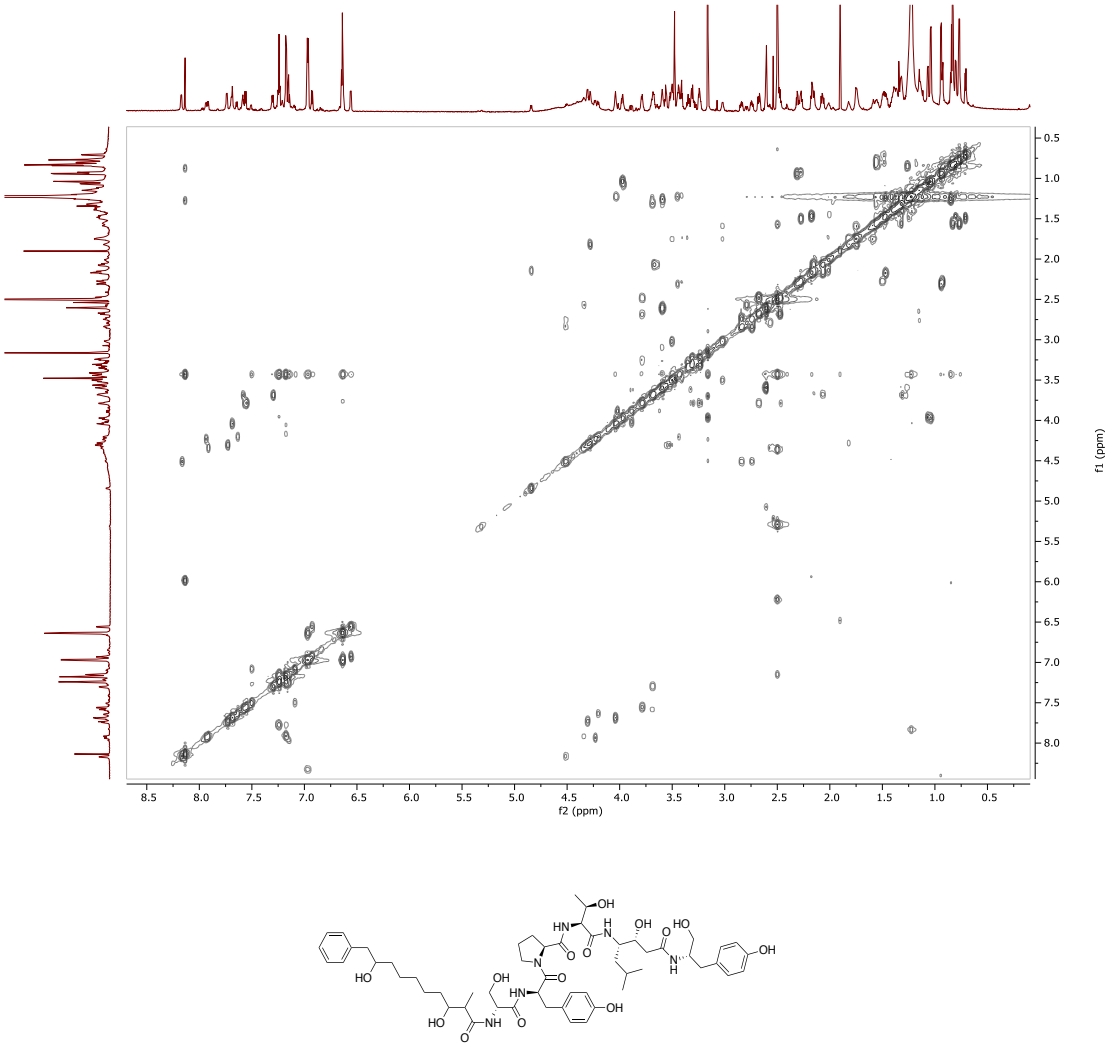


Figure S5. TOCSY spectrum of phormidepistatin (900 MHz, T=300 K, DMSO-d_6 + TFA)

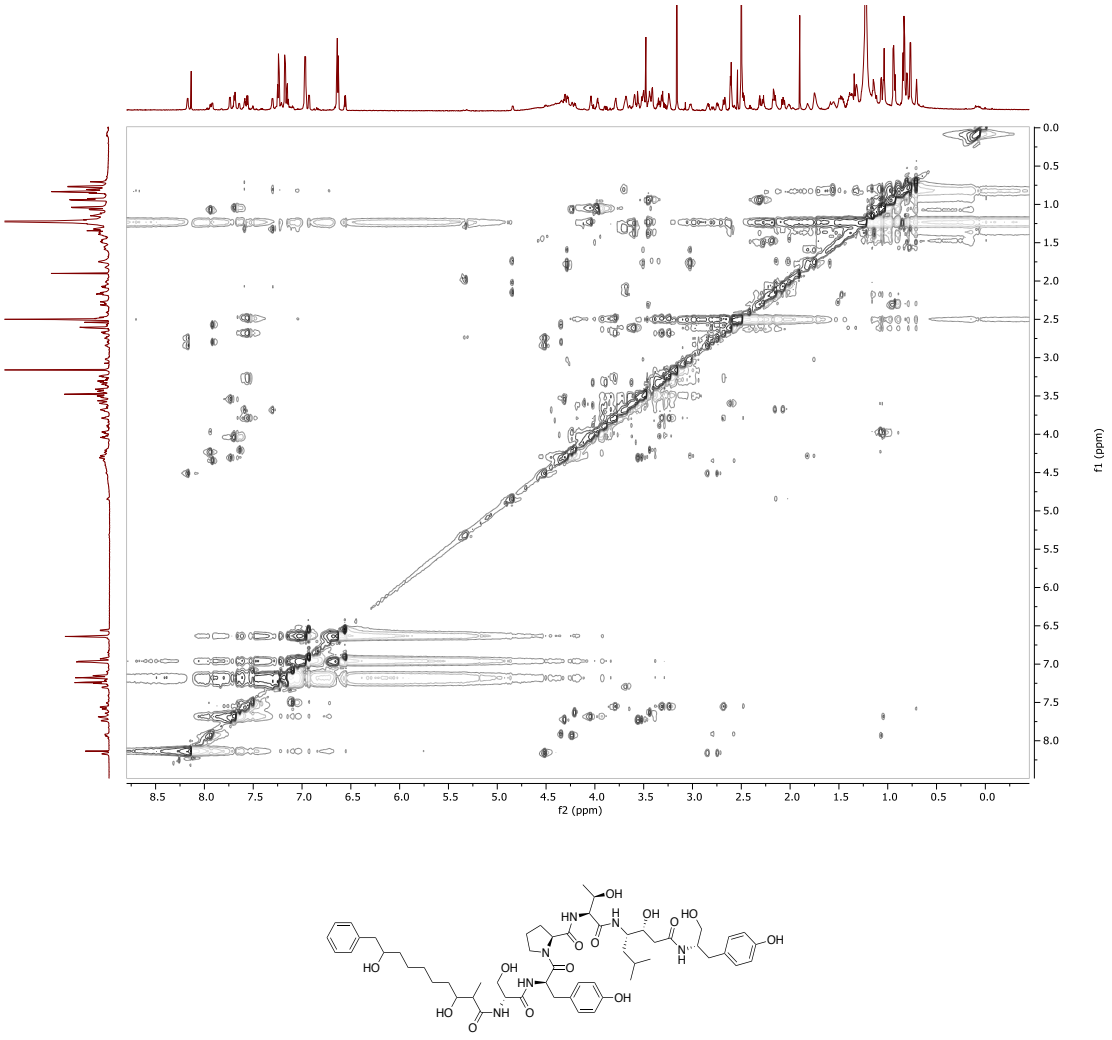


Figure S6. HMBC spectrum of phormidepistatin (900 MHz, T=300 K, DMSO-d_6 + TFA)

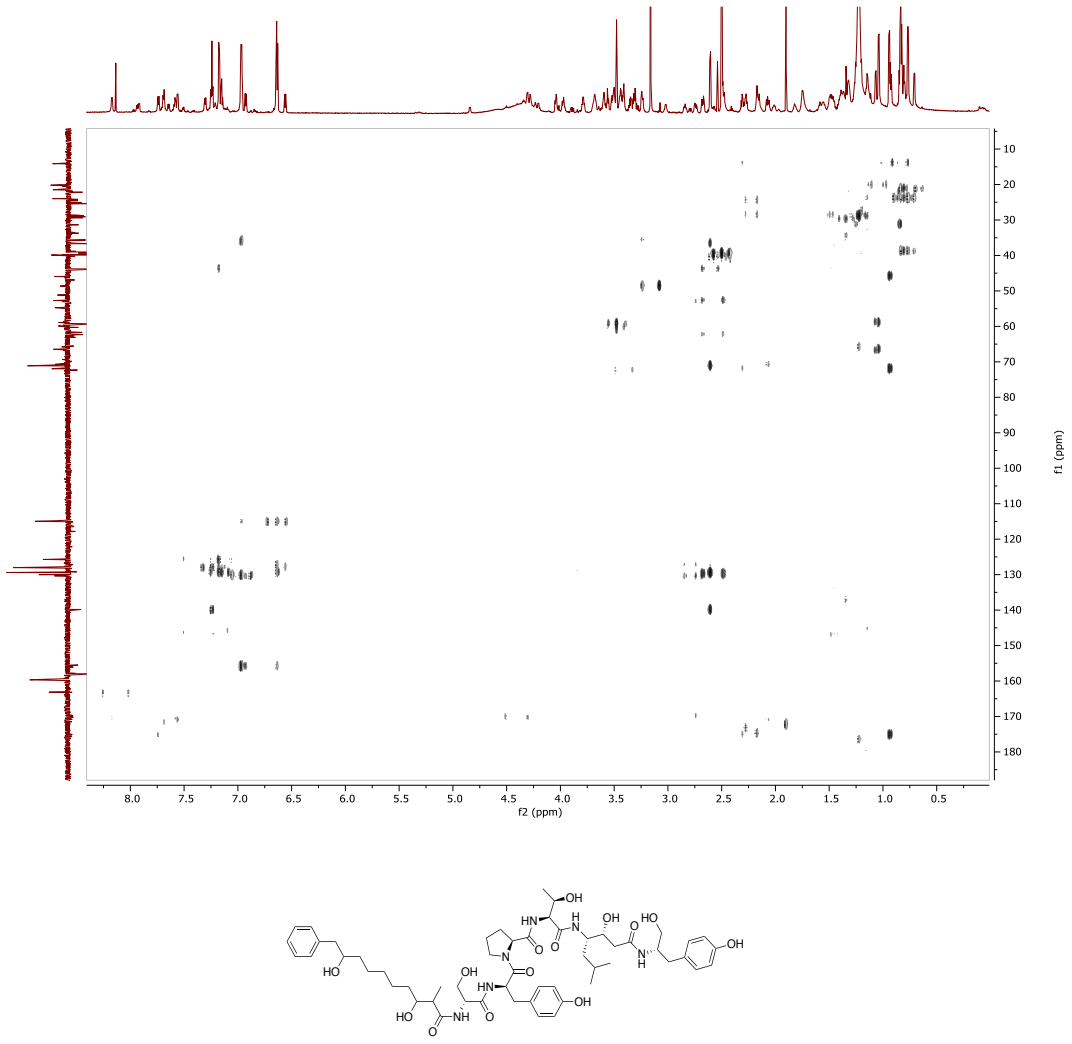


Figure S7. Band-selective HMBC spectrum of phormidepistatin (900 MHz, T=300 K, DMSO-d_6 + TFA)

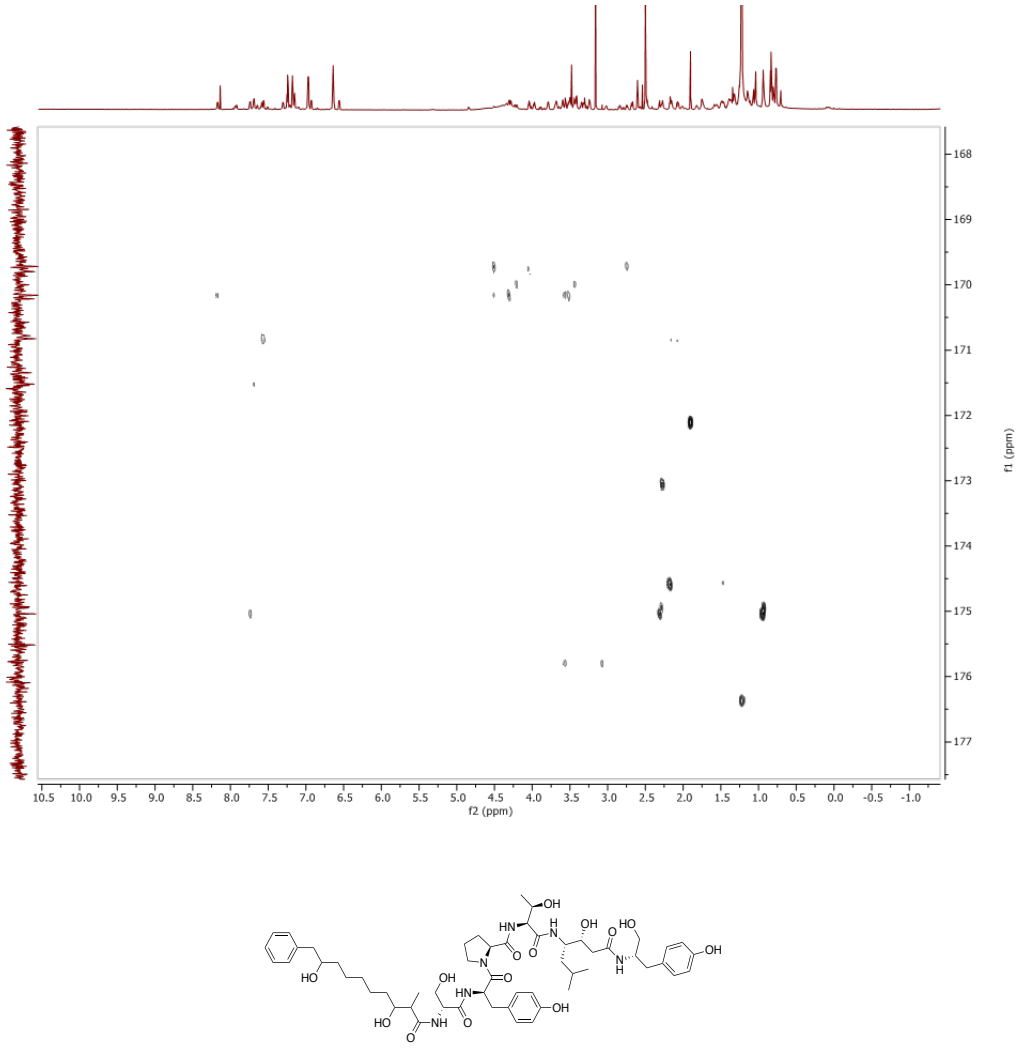


Figure S8. HSQC-TOCSY spectrum of phormidepistatin (900 MHz, T=300 K, DMSO-d_6 + TFA)

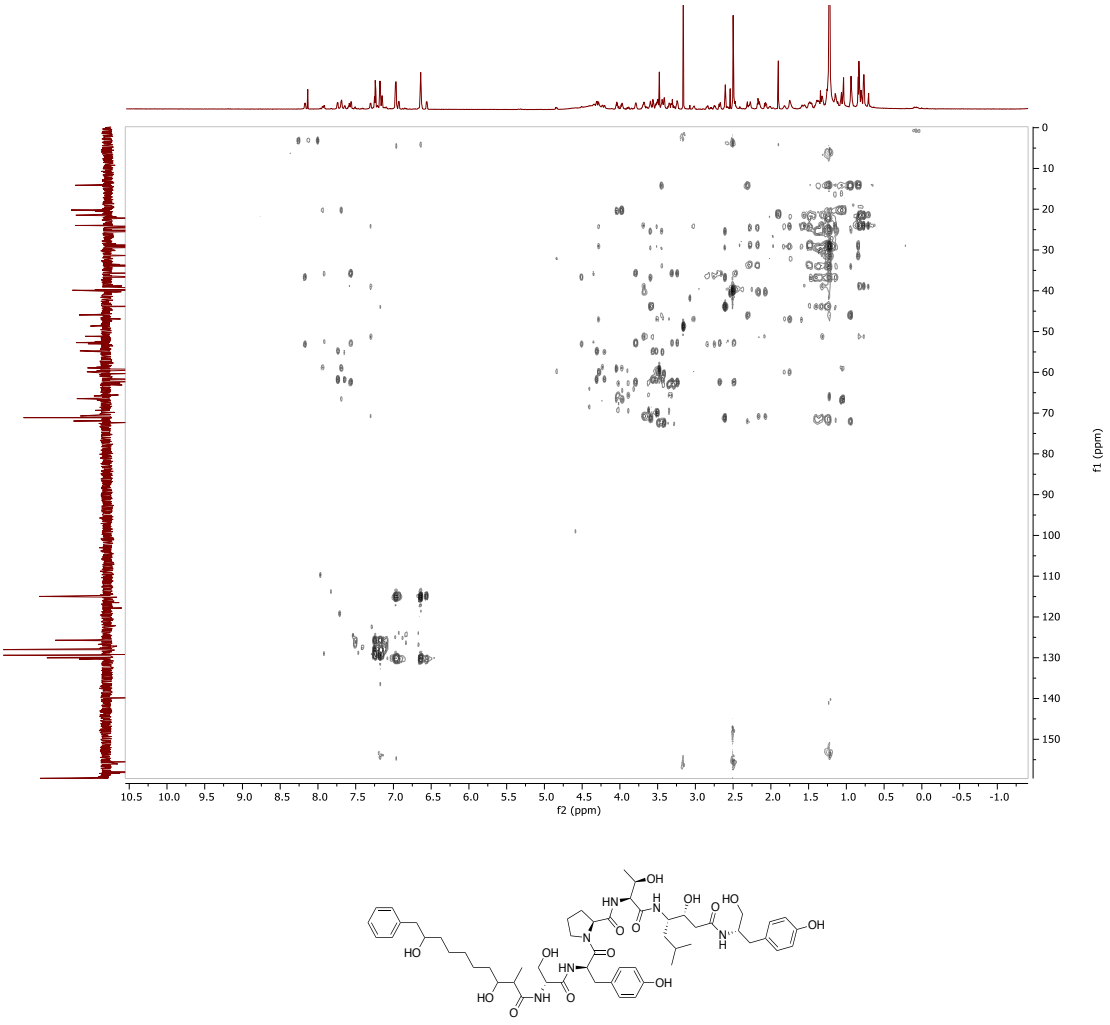


Figure S9. MS analysis of hydrolysis

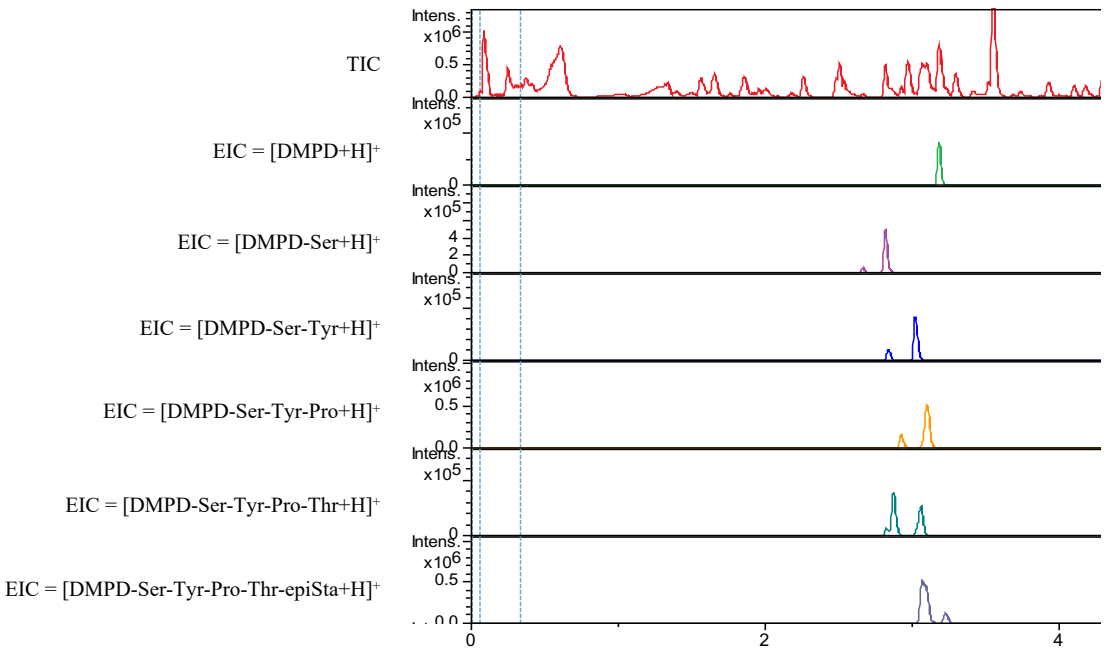


Figure S9: EIC values confirming the sequence of phormidepistatin. EICs with two or more peaks indicate racemization as a result of the hydrolysis.

Figure S10. DQF-COSY spectrum of phormidepistatin (900 MHz, T=300 K, DMSO-d_6 + TFA)

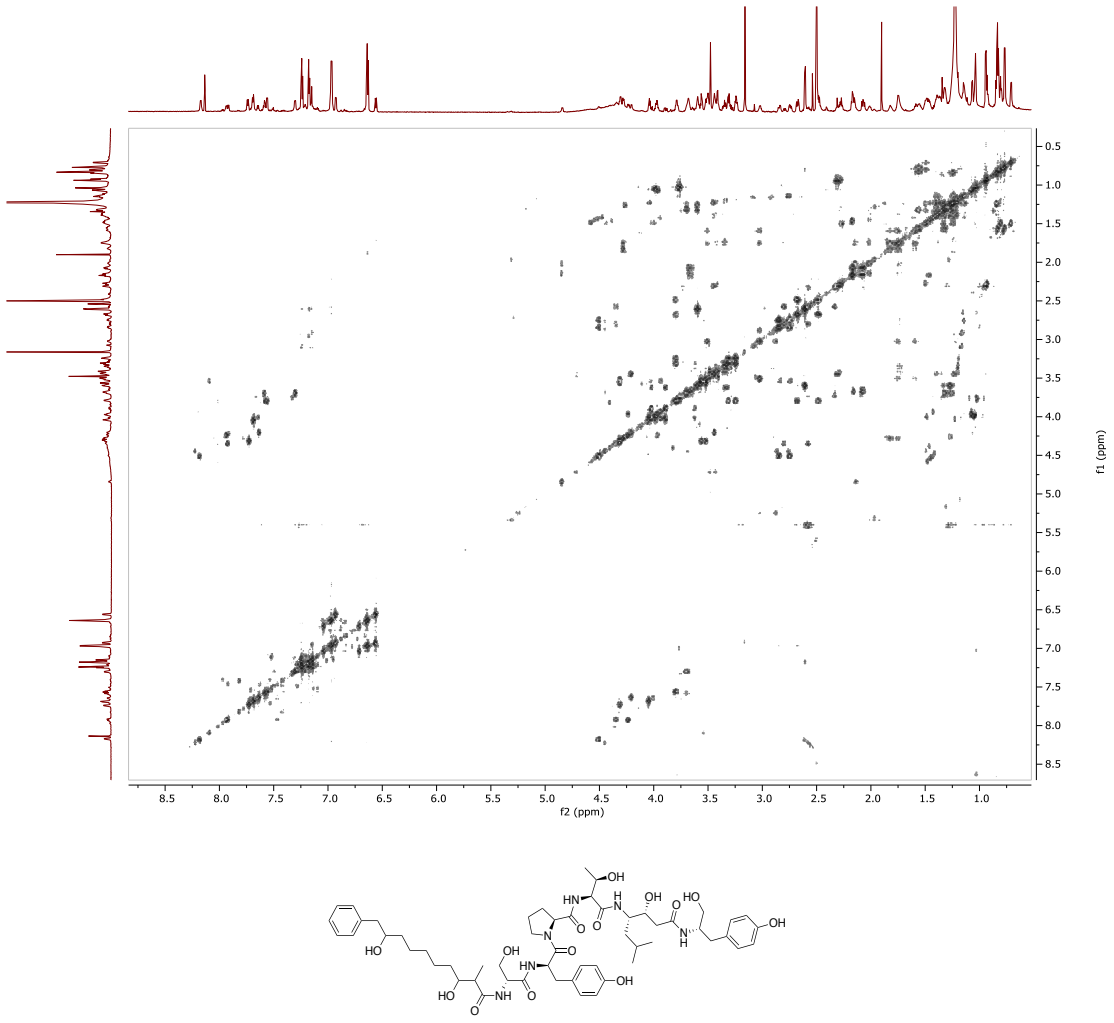


Figure S11. HSQMBC spectrum of phormidepistatin (900 MHz, T=300 K, DMSO-d_6 + TFA)

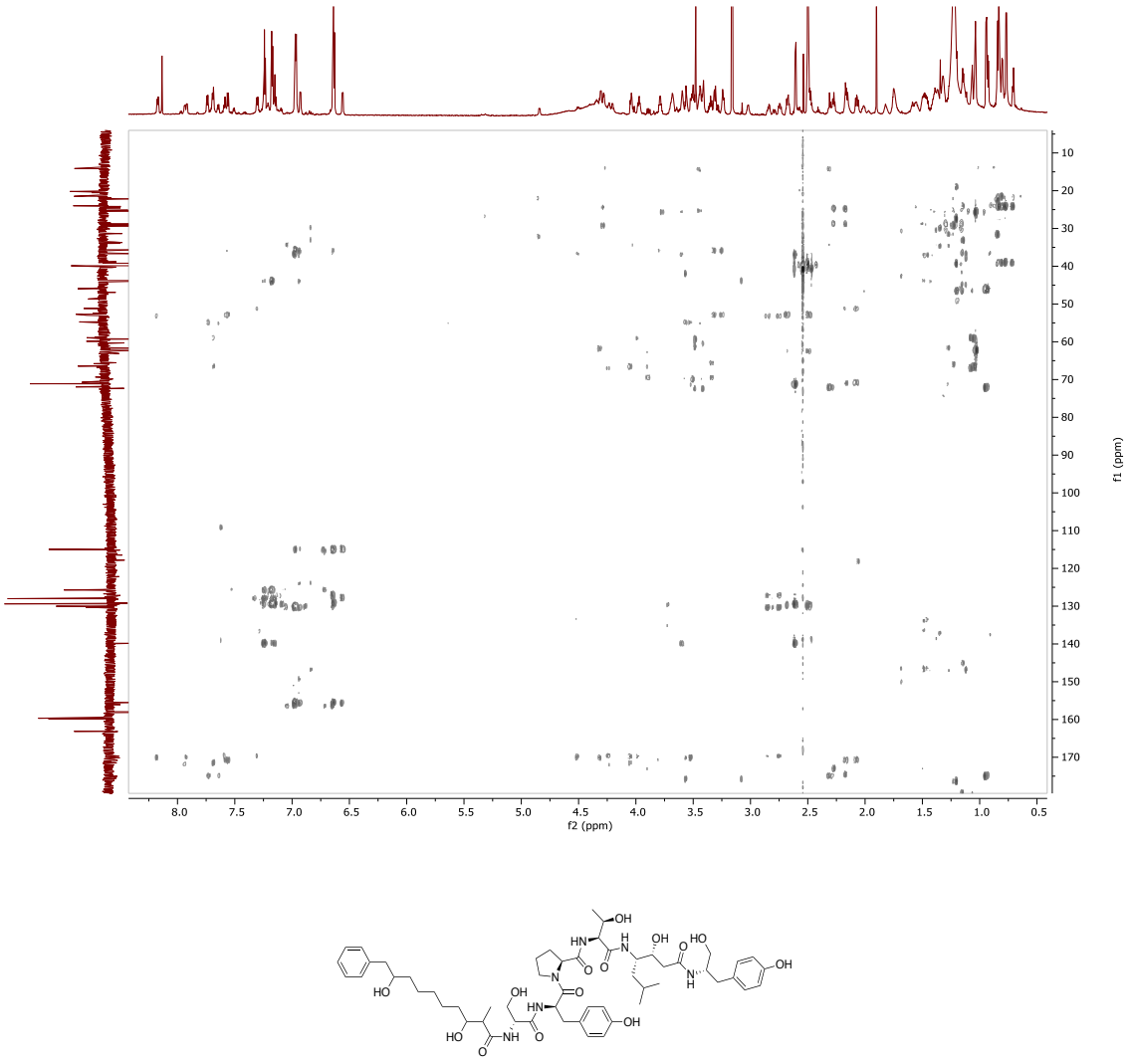


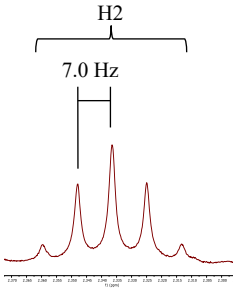
Figure S12. Evaluation of relative configuration based on J-coupling NMR data

No.	Conform. analysis	Coupling (Hz)	DMPD magnitude	Anti/gauche$^-$ magnitude
1	$^3J_{H2\text{-}H3}$	7.0	medium	medium
2	$^3J_{H2\text{-}C4}$	4.8	medium	medium
3	$^3J_{H3\text{-}C1}$	small*	small	small
4	$^3J_{H3\text{-}C17}$	5.1	medium	medium
5	$^2J_{H2\text{-}C3}$	7.3	large	large

*Exact Hz value not determined as it was below the experiment's digital resolution

1)

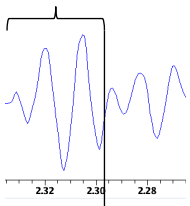
1D ROESY spectrum



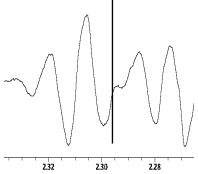
quintet – 7.0 Hz

2)

psHMBC slice

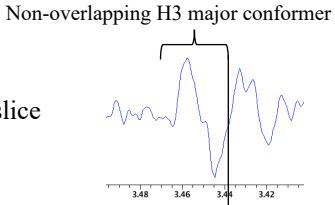


^{1}H modified reference spectra sum – 4.8 Hz

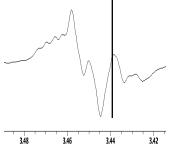


4)

HSQMBC slice

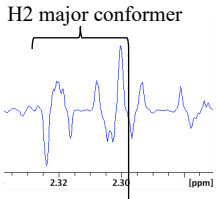


^{1}H reference spectra sum – 5.1 Hz



5)

HSQMBC slice



1H reference spectra sum – 7.3 Hz

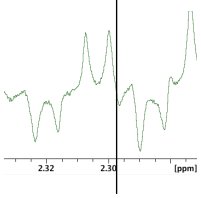
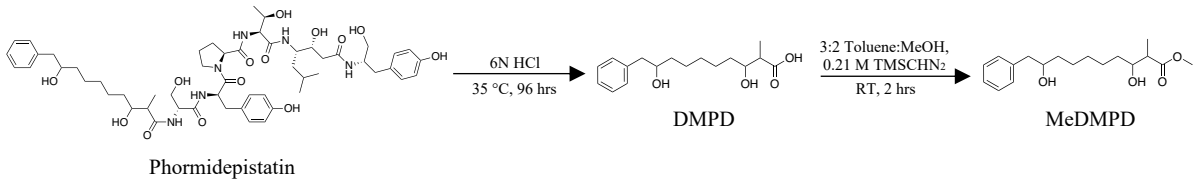


Figure S13. DMPD - Mosher ester analysis preparation and experimentation



Phormidepistatin (5.0 mg) was hydrolyzed in 6N HCl (1.0 mg per 1.0 mL) at 35 °C for 96 hours. The sample vials were then dried under an ambient air stream and washed with 500 µL H_2O twice and evaporated off. The 3,9-dihydroxy-2-methyl-10-phenyldecanoic acid (DMPD) substructure was then isolated on an Agilent 1100 HPLC system using a Phenomenex Kinetex reversed-phase C_{18} column (250x4.6 mm). An initial 25% CH_3CN (+0.1% formic acid) isocratic step for five minutes was followed by a gradient from 25% to 33.6% CH_3CN over 30 minutes to isolate DMPD. The compound eluted at 18.3 minutes. Total yield was 0.4 mg of DMPD (30%).

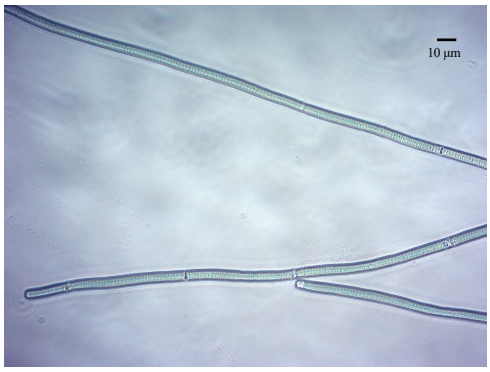
Pure DMPD was esterified in a solution of 3:2 toluene:methanol with 0.21 M Trimethylsilyldiazomethane ($TMSCHN_2$) resulting in MeDMPD. This was done to reduce the number of reactive groups susceptible to derivatization in the reaction required for Mosher ester analysis. The reaction was performed to completion. MeDMPD was purified using the same instrumentation, column, and method as the DMPD isolation. MeDMPD eluted at 32.7 minutes.

Sufficient derivatization of the two MeDMPD secondary alcohols with (*R*)-(−)-α-Methoxy-α-(trifluoromethyl)phenylacetyl chloride (*R*-MTPA-Cl) and (*S*)-(+)-α-Methoxy-α-(trifluoromethyl)phenylacetyl chloride (S-MTPA-Cl) was not achieved. Various derivatization conditions using pyridine (Pyr) and 4-dimethylaminopyridine (DMAP) were attempted as listed in the table below.

Attempt	Rxn time	Catalyst	MTPA	Relative amount of MTPA*
1	18 hrs	Pyr	*R, S*	15
2	18 hrs	DMAP (in $CHCl_3$)	*R, S*	4
3	18 hrs	DMAP + Pyr	*R, S*	4

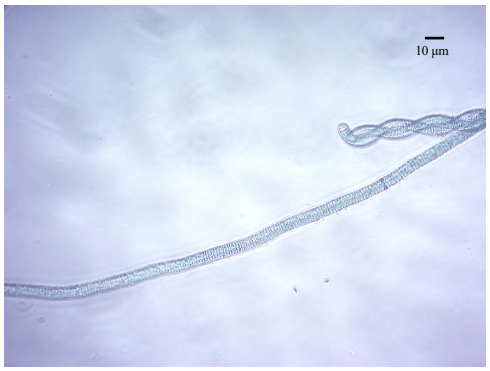
*Molar equivalence compared to MeDMPD

Figure S14. Photomicrographs of cf. *Phormidium* sp. UIC 10045 and 10484 and cf. *Trichormus* sp. UIC 10339



UIC 10045; 400x, brightfield

UIC 10339; 400x, brightfield



UIC 10484; 400x, brightfield

Figure S15. MS evaluation of phormidepistatin in the 10484, 10339, and 10045 extracts

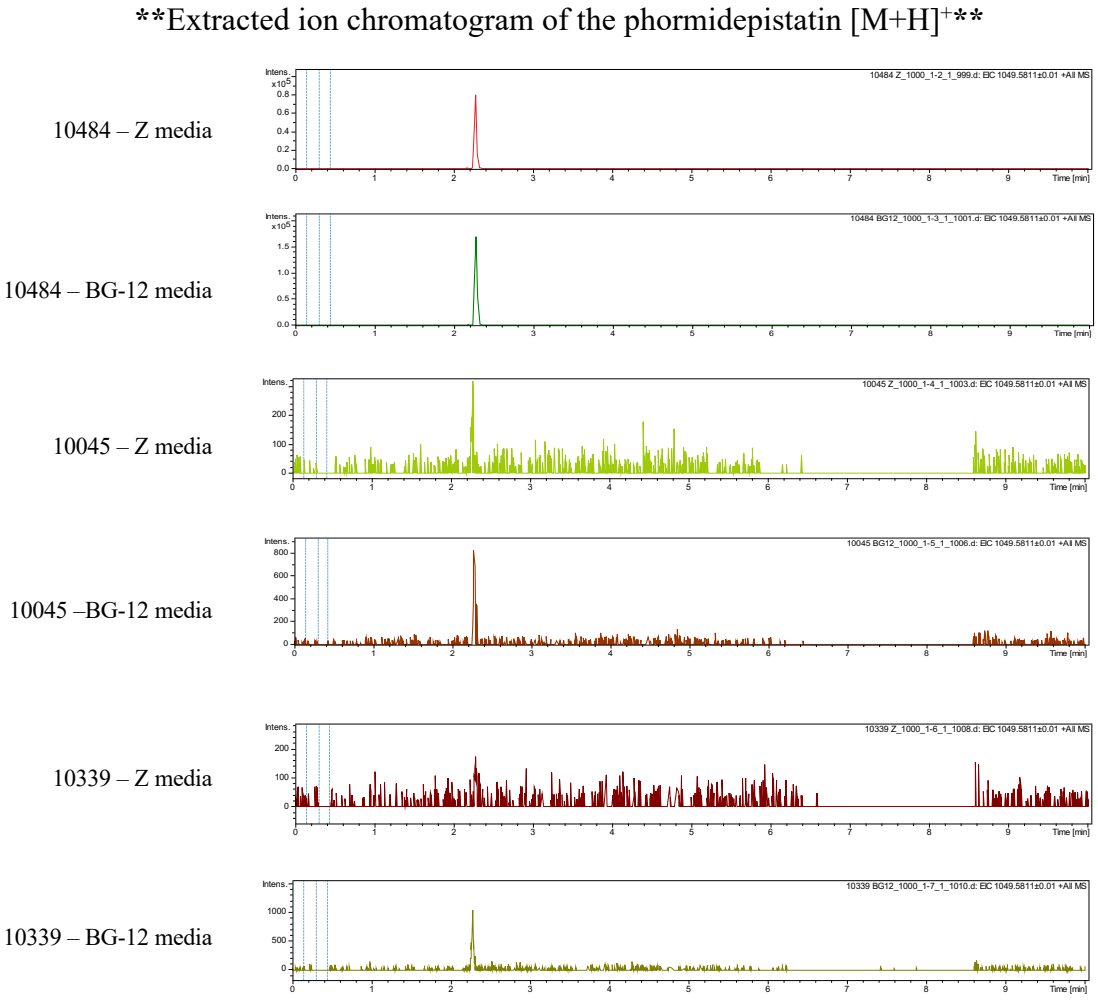


Figure S14: Extracted ion chromatogram of the phormidepistatin $[M+H]^+$ ion in all six samples evaluated (three strains grown in two media).

Figure S16. MS/MS spectra of phormidepistatin analogues from UIC 10484

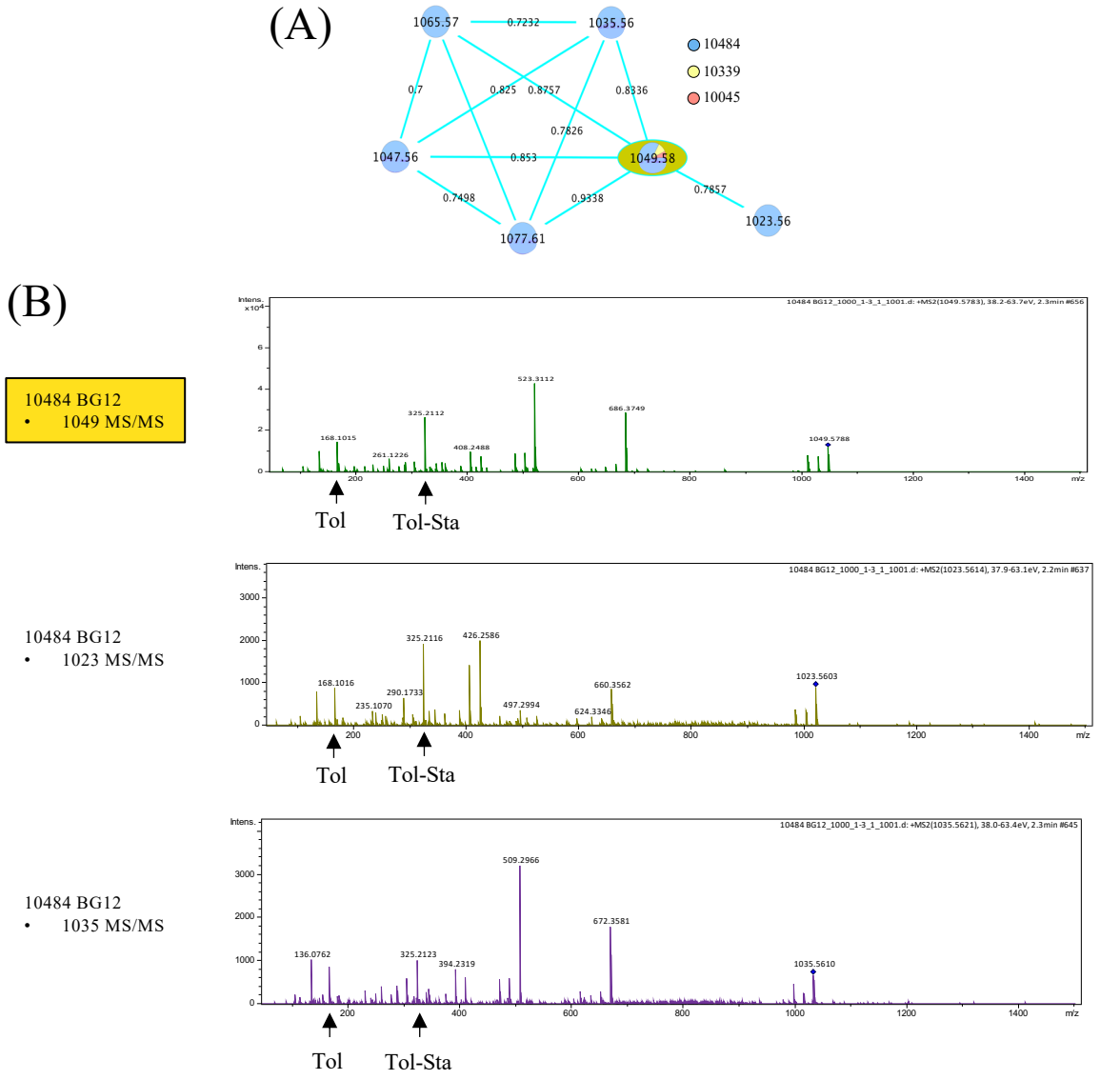


Figure S15: Survey of phormidepistatin and analogues produced by UIC 10484, 10339, and 10045. Analogues were only identified from UIC 10484. (A) The molecular networking analysis was run through the Global Natural Products Social Molecular Networking platform. The gold-highlighted node is phormidepistatin. Node color indicates producing strain. A cosine score was set to 0.7 with minimum matched peaks = 6. Parent mass and MS/MS fragment mass ion tolerances were set to 0.02 and 0.5 Da, respectively. The network included doubly charged nodes of phormidepistatin and analogues, however, these nodes were removed from the visual to improve clarity. (B) Three of the 10484 MS/MS spectra that correspond to half of the nodes in the network. The phormidepistatin MS/MS spectrum is at the top with the label highlighted gold. The two spectra below are analogues, both of which contain a Sta/epi-Sta moiety. Additional analogue spectra are in Figure S16. (Refer to the Experimental Section description for MS/MS parameters.)

Figure S17. MS/MS spectra of phormidepistatin analogues from UIC 10484 (cont.)

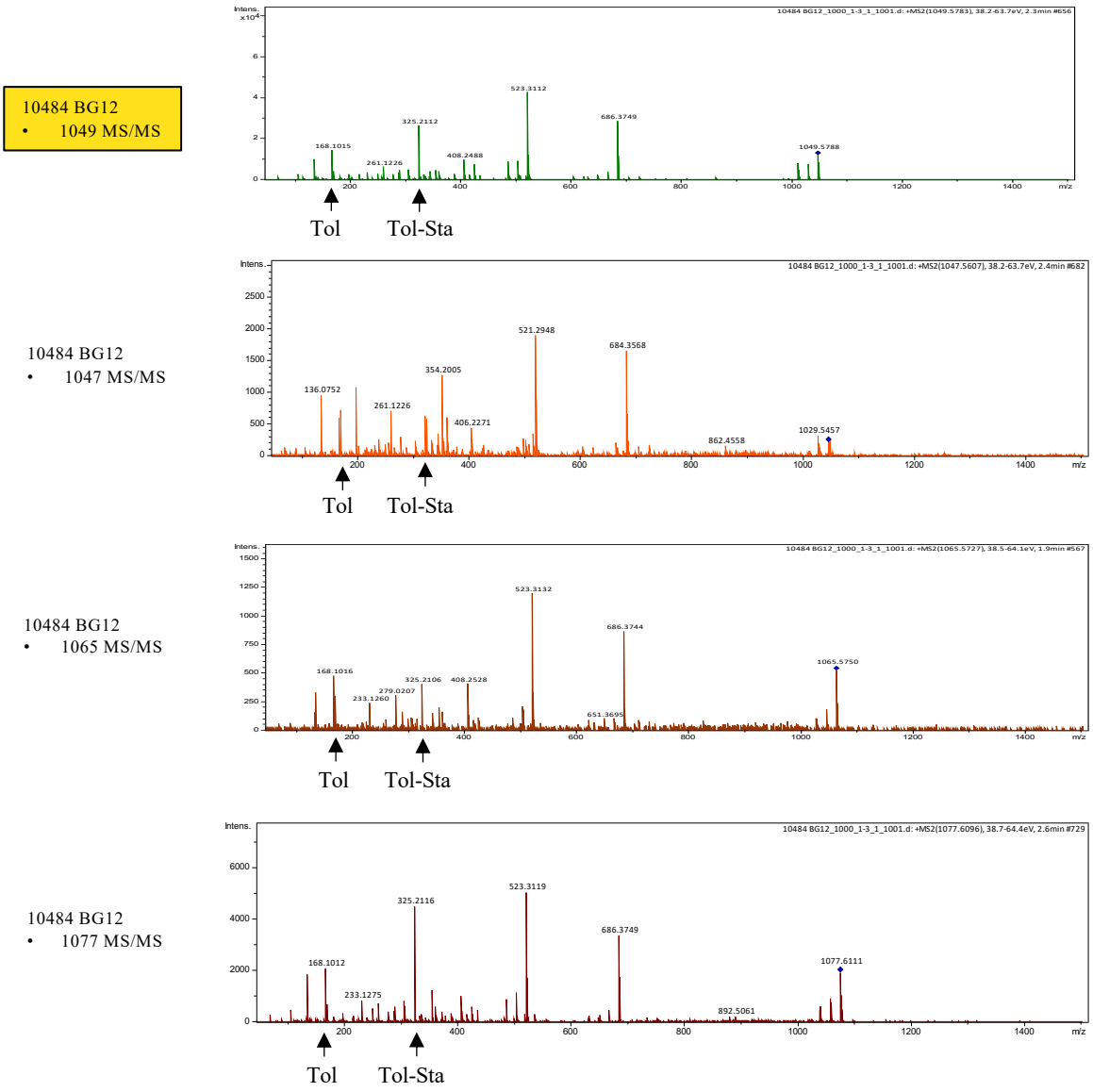
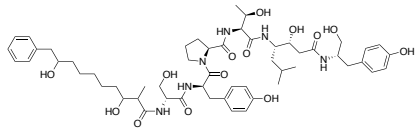


Figure S16: The phormidepistatin MS/MS spectrum is at the top with the label highlighted gold. The three spectra below are analogues, all of which contain a Sta/epi-Sta moiety. These metabolites correspond to nodes in the molecular network in Figure S15.

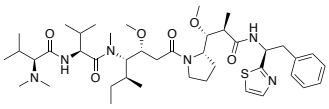
Figure S18. Structures of cyanobacterial metabolites with a Sta/Sta-like moiety

Compound Class A

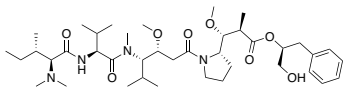


Phormidepistatin

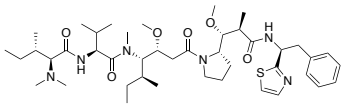
Compound Class B



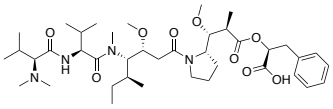
Dolastatin 10



Malevamide D

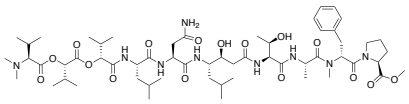


Symplostatin 1

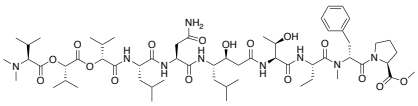


Symplostatin 3

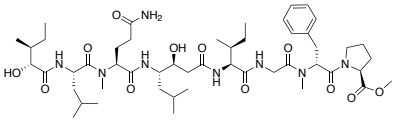
Compound Class C



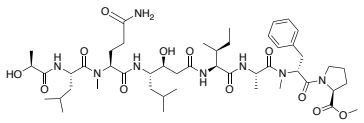
Grassystatin A



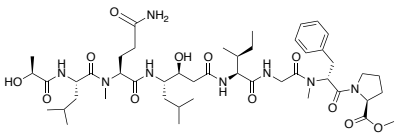
Grassystatin B



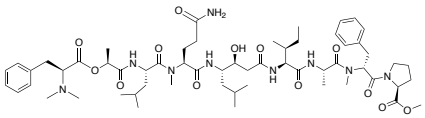
Grassystatin C



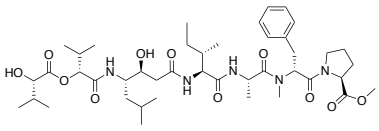
Grassystatin D



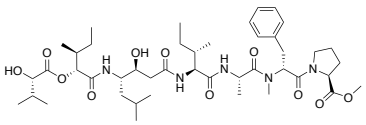
Grassystatin E



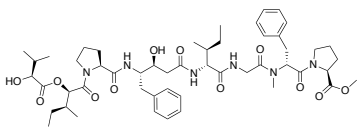
Grassystatin F



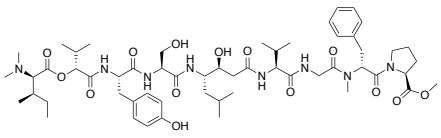
Izenamide A



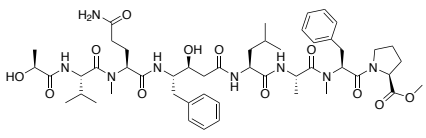
Izenamide B



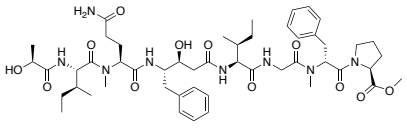
Maedamide



Symplocin A



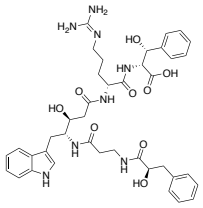
Tasiamide B



Tasiamide F

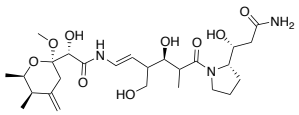
Figure S19. Structures of cyanobacterial metabolites with a Sta/Sta-like moiety (cont.)

Compound Class D



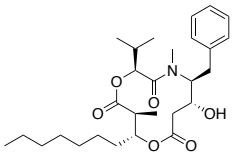
Kasumigamide

Compound Class E



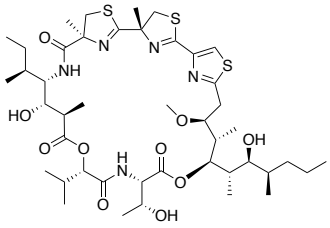
Nosperin

Compound Class F

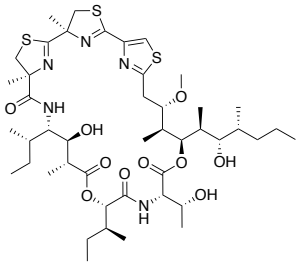


Hapalosin

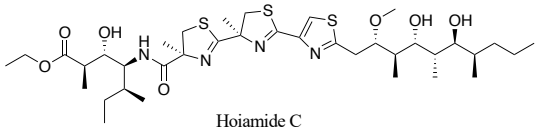
Compound Class G



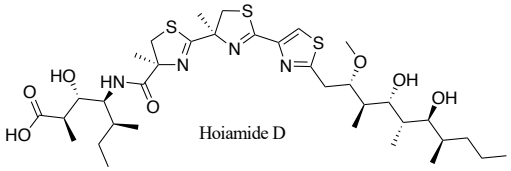
Hoiamide A



Hoiamide B



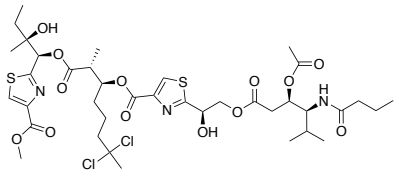
Hoiamide C



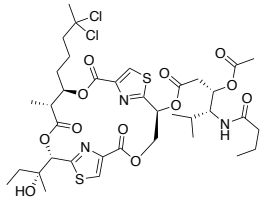
Hoiamide D

Figure S20. Structures of cyanobacterial metabolites with a Sta/Sta-like moiety (cont.)

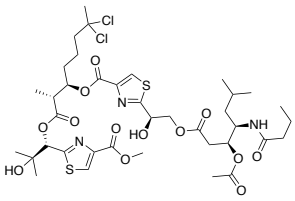
Compound Class H



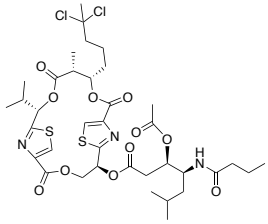
Lyngbyabellin D



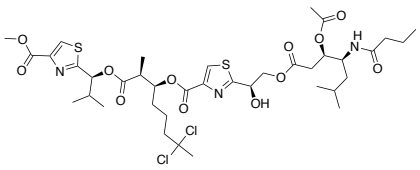
Lyngbyabellin E



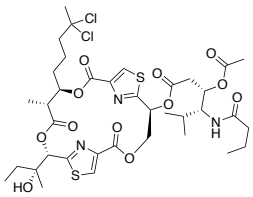
Lyngbyabellin F



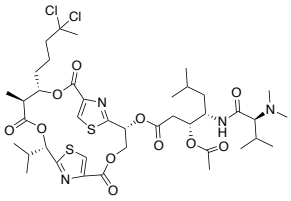
Lyngbyabellin H



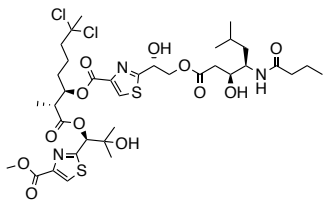
Lyngbyabellin I



Lyngbyabellin J

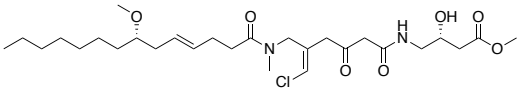


Lyngbyabellin N



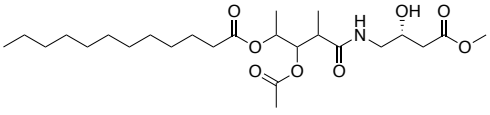
Lyngbyabellin P

Compound Class I



Malyngamide 3

Compound Class J



Guamamide

Table S1. Phormidepistatin advanced Marfey's data

Amino acid	Retention time (min)			Assignment
	L	D	Measured	
Tol[a]	22.4	22.5	22.4	L
Pro[b]	2.6	3.2	2.6	L
Tyr[b]	3.0	3.5	3.5	D
Ser[a]	32.0	33.1	33.1	D

Amino acid	Retention time (min)					Assignment
	L	D	L-allo	D-allo	Measured	
Thr[b]	1.8	2.0	2.3	2.7	1.8	L

Amino acid	Retention time (min)					Assignment
	3S,4S	3R,4R	3S,4R	3R,4S	Measured	
Sta[b]	3.4	5.3	5.3	3.6	3.6	*3R,4S*

[a]Absolute configuration determined by co-injection with commercial standards by HPLC. (See Experimental Section)
[b]Absolute configuration determined by LC-MS with retention time compared to commercial standard retention times. (See Experimental Section)

Table S2. Structure similarity analysis – compound classes

Software package – MayaChemTools
Python Script – RDKitClusterMolecules

Parameters
Molecular fingerprint – Topological torsion
Clustering method – Butina
Tanimoto coefficient threshold – 0.6
fpsize – 16,384

Name	ClusterNumber
Phormidepistatin	1
Symplostatin 3	2
Dolastatin 10	2
Malevamide D	2
Symplostatin 1	2
Izenamide B	3
Grassystatin A	3
Grassystatin B	3
Grassystatin C	3
Grassystatin D	3
Grassystatin E	3
Grassystatin F	3
Izenamide A	3
Maedamide	3
Symplocin A	3
Tasiamide B	3
Tasiamide F	3
Kasumigamide	4
Nosperin	5
Hapalosin	6
Hoiamide D	7
Hoiamide A	7
Hoiamide B	7
Hoiamide C	7
Lyngbyabellin J	8
Lyngbyabellin D	8
Lyngbyabellin E	8
Lyngbyabellin F	8
Lyngbyabellin H	8
Lyngbyabellin I	8
Lyngbyabellin N	8
Lyngbyabellin P	8
Malyngamide 3	9
Guamamide	10

Table S3. Strains known to produce a compound class containing a Sta/Sta-like moiety

Class	Compound	Strain	Acc. No.	Source
A	Phormidepistatin (**1**)	UIC 10484	MN453282	freshwater
		UIC 10339	KF444210	freshwater
		UIC 10045	KF444211	freshwater
B	Dolastatin 10	FK15-1	KC207935	marine
		FK09-8	KP164815	marine
		FK09-1	KP164816	marine
		FK08-3	KP164817	marine
		FK09-6	KP164818	marine
		FK09-4	KP164820	marine
		FK09-3	KP164821	marine
		DRTO-57	KP164823	marine
		SJVI08-1	KP164819	marine
		DRTO-59	KP164822	marine
		BCBC11-25	KF746590	marine
		BCBC11-38	KF746592	marine
		BCBC12-2	KF746593	marine
		NAB11-8	KF746594	marine
		NAB11-21	KF746595	marine
		NAB11-28	KF746596	marine
		NAB11-29	KF746597	marine
		NAB11-32	KF746598	marine
		NAC11-67	KF746599	marine
		NAC11-68	KF746600	marine
		FK12-2	KF746601	marine
		FK12-20	KF746602	marine
		FK12-26	KF746603	marine
		FK13-1	KF746605	marine
		HMC13-6	KF746606	marine
		HMC13-9	KF746607	marine
	Malevamide D	x	x	marine
	Symplostatin 1	FK09-8	KP164815	marine
		FK09-1	KP164816	marine
		FK08-3	KP164817	marine
		FK09-6	KP164818	marine
		FK09-4	KP164820	marine
		FK09-3	KP164821	marine
		DRTO-57	KP164823	marine
		SJVI08-1	KP164819	marine
		DRTO-59	KP164822	marine
		BCBC11-25	KF746590	marine
		BCBC11-38	KF746592	marine
		NAB11-8	KF746594	marine
		NAB11-21	KF746595	marine
		NAB11-28	KF746596	marine
		NAB11-29	KF746597	marine
		NAC11-67	KF746599	marine
		NAC11-68	KF746600	marine
		FK12-2	KF746601	marine
		FK12-20	KF746602	marine
		FK12-26	KF746603	marine
		FK13-1	KF746605	marine
		HMC13-6	KF746606	marine
		HMC13-9	KF746607	marine
	Symplostatin 3	VP452	x	marine
C	Grassystatin A	FK12-17	KC196268	marine
		FK12-27	KC986935	marine
		BCBC12-1	KC986932	marine
		L. cf. confervoides	x	marine
	Grassystatin B	FK12-17	KC196268	marine
		FK12-27	KC986935	marine
		BCBC12-1	KC986932	marine
		L. cf. confervoides	x	marine
	Grassystatin C	L. cf. confervoides	x	marine
	Grassystatin D	VPG 14-61	MG098886	marine
	Grassystatin E	VPG 14-61	MG098886	marine
	Grassystatin F	VPG 14-61	MG098886	marine
	Izenamide A	1605-5	LC315181	marine
	Izenamide B	1605-5	LC315181	marine
	Maedamide	MA2	AB857842	marine
	Symplocin A	10-10-039	x	marine
	Tasiamide B	NIH304	x	marine
	Tasiamide F	NIH399	x	marine
D	Kasumigamide	NIES-87	D89031	freshwater
E	Nosperin	N6	CP026681	freshwater
F	Hapalosin	UTEX B1830	x	freshwater
		IC-52-3	KJ767019	freshwater
		SAG 46.79	x	freshwater
G	Hoiamide A	PNG5-28-02-12	x	marine
		PNG06-65.1	HM072001	marine
	Hoiamide B	PNG06-65.1	HM072001	marine
	Hoiamide C	PNG05-8	HM072003	marine
	Hoiamide D	x	x	marine
H	Lyngbyabellin D	VP417	x	marine
	Lyngbyabellin E	PNG5-27- 02-1	x	marine
	Lyngbyabelln F	PNG5-27- 02-1	x	marine
		S1501	KY889150	marine
	Lyngbyabellin H	PNG5-27-02-1	x	marine
	Lyngbyabellin I	PNG5-27-02-1	x	marine
	Lyngbyabellin J	VP417	x	marine
	Lyngbyabellin N	PAL 8/16/08-3	x	marine
	Lyngbyabellin P	S1501	KY889150	marine
I	Malyngamide 3	ECO 27	x	marine
J	Guamamide	VP727	x	marine

x = no data

Table S4. The ten cyanobacterial metabolite classes containing a Sta/Sta-like moiety

Class	Collection source	Secondary metabolite	Standard statine?		Aspartic protease inhibition (IC_{50}, nM)
			Planar structure	Configuration	
A	freshwater	Phormidepistatin (**1**)	Y	N	21300[a]
B	marine	Dolastatin 10	N	N	x
		Malevamide D	N	N	x
		Symplostatin 1	N	N	x
		Symplostatin 3	N	N	x
C	marine	Grassystatin A	Y	Y	26.5[a], 0.886[b]
		Grassystatin B	Y	Y	7.24[a], 0.354[b]
		Grassystatin C	Y	Y	1620[a], 42.9[b]
		Grassystatin D	Y	Y	2000[a], 30[b]
		Grassystatin E	Y	Y	900[a], 5[b]
		Grassystatin F	Y	Y	50[a], 0.5[b]
		Izenamide A	Y	Y	380[a]
		Izenamide B	Y	Y	270[a]
		Maedamide	N	Y	x
		Symplocin A	N	Y	0.3[b]
		Tasiamide B	N	Y	50/182[a], 9.0/66[b], 189[c]
		Tasiamide F	N	Y	57[a], 23[b], 690[c]
D	freshwater	Kasumigamide	N	N	x
E	freshwater	Nosperin	N	N	x
F	freshwater	Hapalosin	N	N	x
G	marine	Hoiamide A	N	Y	x
		Hoiamide B	N	Y	x
		Hoiamide C	N	Y	x
		Hoiamide D	N	Y	x
H	marine	Lyngbyabellin D	N	Y	x
		Lyngbyabellin E	N	N	x
		Lyngbyabellin F	N	x	x
		Lyngbyabellin H	N	N	x
		Lyngbyabellin I	N	x	x
		Lyngbyabellin J	N	N	x
		Lyngbyabellin N	N	N	x
		Lyngbyabellin P	Y	N	x
I	marine	Malyngamide 3	N	N	x
J	marine	Guamamide	N	N	x

Standard *3S,4S* statine; Y = contains, N = does not contain; x = no data; [a]Cathepsin D, [b]Cathepin E, [c]BACE1