

Supporting information for NRPquest: Coupling Mass-Spectrometry and Genome Mining for Non Ribosomal Peptide Discovery

Hosein Mohimani,[†] Wei-Ting Liu,[‡] Roland D. Kersten,[§] Bradley S. Moore,^{§,⊥} Pieter C. Dorrestein,^{‡,⊥} Pavel A. Pevzner,^{||,*}

[†]Department of Electrical and Computer Engineering, UC San Diego

[‡]Department of Chemistry and Biochemistry, UC San Diego

[§]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, UC San Diego

[⊥]Skaggs School of Pharmacy and Pharmaceutical Sciences, UC San Diego

^{||} Department of Computer Science and Engineering, UC San Diego

Table S1. (A) Identifying individual spectra from *S. roseosporous* (SR), *S. pristinaespiralis* (SP), *B. subtilis* (BS), and *B. brevis* (BB) extracts. For the spectra that can be dereplicated from the Norine dataset with p-value less than 10^{-10} , the corresponding peptide is shown. (B) Rescoring results of part (A) by spectral networks using multitag algorithm.¹³ The gene clusters correspond to the known NRPs arylomycin, daptomycin, pristinamycin, surfactin, plipastatin, and tyrocidine, respectively. SpecNetScore is shown for each peptide. Brackets stand for cyclic/branched-cyclic structure, and superscripts stand for modification masses.(C) Correct structure of peptides produced by these organisms (only most abundant species is shown for each peptide family)

(A)

	mz	p-value	structure												
SR	1634 ⁺¹	1*10 ⁻¹²	[val ⁺⁶⁹	trp	asn	asp	thr]	gly	orn	asp	ala	asp	gly	ser	me-g
SR	1663 ⁺¹	8*10 ⁻¹¹	[val	trp	asn	asp	thr ⁺⁷⁹	[gly	orn	asp	ala	asp	gly	ser	me-g
SR	818 ⁺²	9*10 ⁻¹¹	[val ⁺⁹²	tyr	asn	asp	thr	gly	gln ⁻³⁸	asp	val	asp	[gly	ser	me-g
SR	976 ⁺¹	8*10 ⁻¹³	[Me-Ser	lyserg ⁺¹⁹⁶	gly ⁺⁶²	Me-hpg	gly]		bht						
SR	617 ⁺¹	2*10 ⁻¹²	[Me-Ser	gly	gly ⁺⁹⁴	Me-hpg ⁻⁹¹	gly	tyr]							
SR	899 ⁺¹	2*10 ⁻¹²	Me-Ser	[gly	gly	Me-hpg ⁺⁹⁴	gly]		tyr ⁺¹⁷²]						
SP	853 ⁺¹	2*10 ⁻¹²	[ala	thr ⁺³²	abu	pro	Me-phe ⁻⁸²	pip	phg]	dab					
SP	897 ⁺¹	1*10 ⁻¹⁰	ile	thr	[val	pro	Me-phe	pip	phg ⁻⁶¹	dab ⁻²]					
SP	867 ⁺¹	5*10 ⁻¹⁰	[abu ⁻¹⁰	thr	ile	pro	Me-phe ⁻⁶⁸	pip	phg]	dab]					
BS	1037 ⁺¹	3*10 ⁻¹⁶	[glu ⁺²²²	leu	leu	val	asp	leu	leu ⁺¹⁸]						
BS	1023 ⁺¹	7*10 ⁻¹⁵	[glu ⁺²⁰⁸	leu	leu	val	asp	leu	leu ⁺¹⁸]						
BS	760 ⁺²	1*10 ⁻¹⁶	[glu	orn	tyr	thr	glu	val	pro ⁻¹	glu	tyr	ile	phe ⁺¹³⁵]		
BS	738 ⁺²	4*10 ⁻¹⁵	[glu	orn	tyr	thr	glu	val	pro	glu ⁻¹	tyr	ile	phe ⁺⁹¹]		
BS	753 ⁺²	1*10 ⁻¹⁴	[glu	orn	tyr	thr	glu	val	pro ⁻¹	glu	tyr	ile	phe ⁺¹²¹]		
BB	663 ⁺²	8*10 ⁻²⁰	[phe	pro	phe ⁺³⁹	phe ⁺³⁹	asn	gln	tyr	val	orn ⁺¹⁵	leu]			
BB	1324 ⁺¹	2*10 ⁻¹⁸	phe	[pro	phe ⁺³⁹	phe ⁺³⁹	asn	gln	tyr	val	orn ⁺¹⁴	leu]			
BB	656 ⁺²	3*10 ⁻¹⁸	[phe	pro	phe ⁺³⁹	phe ⁺³⁹	asn	gln	tyr	val	orn	leu]			
BB	1285 ⁺¹	5*10 ⁻¹⁷	[phe ⁻¹⁷	pro	phe	phe	asn	gln	tyr	val	orn ⁺³¹	leu]			
BB	656 ⁺²	1*10 ⁻¹⁵	[phe	pro	phe ⁺³⁹	phe ⁺³⁹	asn	gln	tyr	val	orn ⁺¹	leu]			

(B)

species	mz	score	sequence											
SR	1634 ⁺¹	108	[val ⁺⁹⁷	trp	asn	asp	[thr	gly	orn	asp	ala	asp	gly	ser
SR	856 ⁺¹	142	[Me-Ser	lyserg ⁺¹⁹⁶	gly ⁺⁶²	Me-hpg	gly]	bht						
SP	867 ⁺¹	314	ala	thr ⁺³²	[abu	pro	Me-phe ⁻⁶⁸	pip	phg]	dab]				
BS	1037 ⁺¹	inf	[glu ⁺²	leu ⁺²³⁸	leu	val	asp	leu]	leu					
BS	753 ⁺²	75	glu	orn	[tyr	thr	glu	val	pro ⁻¹	glu	tyr	ile	phe ⁺¹²¹]	
BB	1324 ⁺¹	349	[phe	pro	phe ⁺³⁹	phe	asn	gln	tyr	val	orn ⁺¹⁴	leu]		

(C)

SR	daptomycin	1664	aC13:0	trp	asn	asp	[thr	gly	orn	asp	ala	asp	gly	ser
SR	arylomycin	853	iC14:0	Nme-ser	ala	gly	[NMe-hpg	ala	tyr]					
SP	pristinamycin	867	hpa	[thr	abu	pro	Nme-Me2A-phe	4oxo-hpr	Ph-gly					
BS	surfactin	1023	[aC15:0-OH3	glu	leu	leu	val	asp	leu	leu				
BS	plipastatin	1507	C16:0-OH3	glu	orn	[tyr	athr	glu	val	pro	glu	tyr	ile]	
BB	tyrocidine	1310	[phe	pro	trp	phe	asn	gln	tyr	val	orn	leu]		

Table 2. (A) Spectral network dereplication results for top scoring networks of Table S1. These networks correspond to daptomycin (1), arylomycin (2), pristinamycin (3), surfactin (4), plipastatin (5) and tyrocidine (6) gene clusters respectively. For each network in Figure 3, all of the nodes are dereplicated from the most abundant known variant. Only variants with statistical significance smaller than 0.0001 are shown.

Organism	mz	p-value	reconstruction															
SR	1664 ⁺¹	2*10 ⁹	ac13:0	trp	asn	asp	[thr	gly	orn	asp	ala	asp	gly	ser	me-glu	kyn]		
SR	819 ⁻²	1*10 ⁸	ac13:0 ²⁸	trp	asn	asp	[thr	gly	orn	asp	ala	asp	gly	ser	me-glu	kyn]		
SR	1635 ⁺¹	8*10 ⁷	ac13:0 ²⁸	trp	asn	asp	[thr	gly	orn	asp	ala	asp	gly	ser	me-glu	kyn]		
SR	1650 ⁻¹	2*10 ⁷	ac13:0 ¹⁴	trp	asn	asp	[thr	gly	orn	asp	ala	asp	gly	ser	me-glu	kyn]		
SR	832 ⁻²	1*10 ⁵	ac13:0	trp	asn	asp	[thr	gly	orn	asp	ala	asp	gly	ser	me-glu	kyn]		
SR	826 ⁻²	1*10 ⁵	ac13:0 ¹⁴	trp	asn	asp	[thr	gly	orn	asp	ala	asp	gly	ser	me-glu	kyn]		
SR	841 ⁻²	1*10 ⁵	ac13:0	trp	asn ¹⁹	asp	[thr	gly	orn	asp	ala	asp	gly	ser	me-glu	kyn]		
SR	1281 ⁺¹	1*10 ⁵				asp	[thr	gly	orn	asp	ala	asp	gly	ser	me-glu	kyn]		
SR	853 ⁺¹	9*10 ¹⁰	ic14:0	N-Me-Ser	ala	gly	[NMe-hpg	ala	tyr]									
SR	825 ⁺¹	5*10 ¹⁰	ic14:0 ⁴²	N-Me-Ser ¹⁴	ala	gly	[NMe-hpg	ala	tyr]									
SR	839 ⁻¹	2*10 ⁸	ic14:0	N-Me-Ser	ala	gly	[NMe-hpg	ala	tyr]									
SR	845 ⁻¹	5*10 ⁷	ic14:0 ⁶	N-Me-Ser	ala	gly	[NMe-hpg	ala	tyr]									
SR	867 ⁻¹	4*10 ⁵	ic14:0 ¹⁴	N-Me-Ser	ala	gly	[NMe-hpg	ala	tyr]									
SP	853 ⁺¹	7*10 ¹⁴	hpa	[thr	abu	pro	NMe-Me2A-phe ¹⁴	4oxo-hpr	Ph-gly]									
SP	867 ⁻¹	3*10 ¹¹	hpa	[thr	abu	pro	NMe-Me2A-phe	4oxo-hpr	Ph-gly]									
SP	810 ⁻¹	3*10 ⁹	hpa	[thr ¹⁴	abu	pro	NMe-Me2A-phe ⁴³	4oxo-hpr	Ph-gly]									
SP	857 ⁻¹	8*10 ⁸	hpa	[thr	abu	pro	NMe-Me2A-phe	4oxo-hpr ¹⁰	Ph-gly]									
SP	905 ⁻¹	3*10 ⁷	hpa	[thr	abu	pro ²⁸	NMe-Me2A-phe	4oxo-hpr	Ph-gly]									
SP	883 ⁻¹	5*10 ⁷	hpa	[thr	abu	pro	NMe-Me2A-phe ¹⁶	4oxo-hpr	Ph-gly]									
SP	825 ⁺¹	2*10 ⁶	hpa	[thr	abu	pro	NMe-Me2A-phe ⁴³	4oxo-hpr	Ph-gly]									
SP	891 ⁺¹	2*10 ⁶	hpa	[thr	abu	pro ²²	NMe-Me2A-phe ¹	4oxo-hpr	Ph-gly]									
SP	889 ⁻¹	4*10 ⁶	hpa	[thr	abu	pro ²²	NMe-Me2A-phe	4oxo-hpr	Ph-gly]									
SP	881 ⁻¹	7*10 ⁶	hpa	[thr	abu ¹⁸	pro	NMe-Me2A-phe	4oxo-hpr ³³	Ph-gly]									
SP	899 ⁻¹	1*10 ⁵	hpa	[thr	abu	pro	NMe-Me2A-phe ¹⁶	4oxo-hpr	Ph-gly ^{16]}									
SP	849 ⁻¹	1*10 ⁵	hpa	[thr	abu ¹⁸	pro	NMe-Me2A-phe	4oxo-hpr	Ph-gly]									
SP	875 ⁺¹	2*10 ⁵	hpa	[thr	abu	pro ⁸	NMe-Me2A-phe	4oxo-hpr	Ph-gly]									
SP	847 ⁻¹	3*10 ⁵	hpa	[thr	abu ¹⁸	pro	NMe-Me2A-phe	4oxo-hpr	Ph-gly ¹									
SP	843 ⁻¹	3*10 ⁵	hpa	[thr	abu	pro	NMe-Me2A-phe	4oxo-hpr ²⁴	Ph-gly]									
SP	873 ⁻¹	4*10 ⁵	hpa	[thr	abu	pro	NMe-Me2A-phe ¹⁶	4oxo-hpr ¹⁰	Ph-gly]									
BS	1041 ⁺¹	2*10 ¹¹	[aC15:0-OH3	glu	leu	leu	val	asp	leu ⁻²	leu ^{16]}								
BS	1023 ⁺¹	2*10 ¹⁰	[aC15:0-OH3	glu	leu	leu	val	asp	leu	leu]								
BS	1051 ⁺¹	4*10 ¹⁰	[aC15:0-OH3	glu ¹⁴	leu	leu	val	asp	leu ⁻¹⁵	leu]								
BS	1009 ⁻¹	2*10 ⁹	[aC15:0-OH3	glu ¹⁰	leu	leu	val	asp	leu	leu ^{3]}								
BS	955 ⁻¹	6*10 ⁸	[aC15:0-OH3	glu ¹⁴	leu	leu	val	asp	leu ⁻¹³	leu]								
BS	1019 ⁺¹	7*10 ⁸	[aC15:0-OH3	glu	leu ¹⁸	leu	val	asp	leu ⁻¹⁵	leu]								
BS	1037 ⁺¹	2*10 ⁷	[aC15:0-OH3	glu	leu	leu	val	asp	leu	leu ^{14]}								
BS	1065 ⁻¹	2*10 ⁷	[aC15:0-OH3	glu ²⁸	leu	leu	val	asp	leu ⁻¹⁵	leu]								
BS	1059 ⁺¹	7*10 ⁵	[aC15:0-OH3	glu	leu	leu	val ²³	asp	leu	leu ^{14]}								
BS	1071 ⁺¹	7*10 ⁵	[aC15:0-OH3	glu	leu	leu ³⁵	val	asp	leu	leu ^{14]}								
BS	1031 ⁺¹	9*10 ⁵	[aC15:0-OH3 ²²	glu	leu	leu	val	asp	leu ⁻¹³	leu]								
BS	753 ⁻²	4*10 ¹¹	C16:0-OH3 ¹⁴	glu	orn	[tyr	athr	glu	val	pro	glu	tyr	ile					
BS	739 ⁻²	2*10 ¹⁰	C16:0-OH3 ¹²	glu	orn	[tyr	athr	glu	val	pro ²⁹	glu	tyr	ile					
BS	732 ⁻²	3*10 ¹⁰	C16:0-OH3	glu	orn	[tyr	athr	glu	val ²⁹	pro	glu	tyr	ile					
BS	1507 ⁻¹	4*10 ⁹	C16:0-OH3 ¹⁴	glu	orn	[tyr	athr	glu	val	pro	glu	tyr	ile					
BS	760 ⁻²	1*10 ⁸	C16:0-OH3 ²⁸	glu	orn	[tyr	athr	glu	val	pro	glu	tyr	ile					
BS	746 ⁻²	2*10 ⁸	C16:0-OH3	glu	orn	[tyr	athr	glu	val	pro	glu	tyr	ile					
BS	1535 ⁻¹	1*10 ⁵	C16:0-OH3 ⁴¹	glu	orn	[tyr	athr	glu	val	pro	glu	tyr	ile					
BS	724 ⁻²	1*10 ⁵	C16:0-OH3 ¹⁶	glu	orn	[tyr	athr ²⁸	glu	val	pro	glu	tyr	ile					
BS	1525 ⁺¹	2*10 ⁵	C16:0-OH3 ¹⁴	glu	orn	[tyr	athr	glu	val	pro	glu	tyr	ile ¹⁸					
BB	675 ⁻²	3*10 ²²	[phe	pro	trp	phe ⁴⁰	asn	gln	tyr	val	orn	leu]						
BB	656 ⁻²	3*10 ²⁰	[phe	pro	trp	phe	asn	gln	tyr	val	orn	leu]						
BB	663 ⁻²	2*10 ¹⁶	[phe	pro	trp	phe	asn	gln	tyr	val	orn ¹⁵	leu]						
BB	1324 ⁺¹	5*10 ¹⁶	[phe	pro	trp	phe	asn	gln	tyr	val	orn ¹⁵	leu]						
BB	1363 ⁺¹	6*10 ¹⁵	[phe	pro	trp	phe ³⁹	asn	gln	tyr	val	orn ¹⁵	leu]						
BB	682 ⁻²	2*10 ¹⁴	[phe	pro	trp	phe ⁴⁰	asn	gln	tyr	val	orn ¹⁴	leu]						
BB	667 ⁻²	2*10 ¹²	[phe	pro	trp	phe ²⁴	asn	gln	tyr	val	orn	leu]						
BB	1285 ⁺¹	3*10 ¹²	[phe	pro	trp	phe ³⁸	asn	gln	tyr	val	orn ¹⁴	leu]						
BB	1349 ⁺¹	3*10 ¹²	[phe	pro	trp	phe ⁴⁰	asn	gln	tyr	val	orn	leu]						
BB	1372 ⁺¹	1*10 ¹⁰	[phe	pro	trp	phe ⁴⁰	asn	gln	tyr ²³	val	orn	leu]						
BB	1310 ⁺¹	1*10 ¹⁰	[phe	pro	trp	phe	asn	gln	tyr	val	orn	leu]						
BB	1295 ⁺¹	7*10 ¹⁰	[phe	pro	trp	phe	asn ¹⁵	gln	tyr	val	orn	leu]						
BB	691 ⁻²	1*10 ⁹	[phe	pro	trp	phe ⁴⁰	asn	gln	tyr	val	orn ³³	leu]						
BB	1271 ⁺¹	8*10 ⁹	[phe	pro	trp	phe	asn ³⁸	gln	tyr	val	orn	leu]						
BB	1339 ⁺¹	2*10 ⁸	[phe	pro	trp	phe	asn	gln	tyr	val ¹⁵	orn ¹⁵	leu]						
BB	669 ⁻²	2*10 ⁸	[phe	pro	trp	phe ⁴⁰	asn ¹³	gln	tyr	val	orn	leu]						
BB	672 ⁻²	1*10 ⁶	[phe	pro	trp	phe	asn	gln	tyr	val ³³	orn	leu]						

NRPquest webservice user manual guide

NRPquest webservice is available at cyclo.ucsd.edu. The user can upload their genomic and mass spectral datasets, and the webservice reports all the significant hits between the spectral dataset and NRPs predicted from the genomic datasets. Samples of datasets used in this paper from bacterial strains *Streptomyces roseosporus* NRRL 15998, *Streptomyces pristinaespiralis* ATCC 25486, *Bacillus subtilis* subsp. *subtilis* str. NCIB 3610, and *Bacillus brevis* ATCC 8185 are provided. The runtime is about 1 min/spectrum. **Fig. S1** shows a snapshot of how the webservice works for *Bacillus subtilis* str. NCIB 3610 dataset.

NRPquest

mzXML File:	<input type="button" value="Choose File"/> PY79.mzXML	
Genome File:	<input type="button" value="Choose File"/> 11379.fasta	<input checked="" type="radio"/> protein <input type="radio"/> DNA
Email:	<input type="text" value="g.hosein.mohimani@g"/>	
Charge(default 1-3):	<input type="text"/>	
Precursor Tolerance (Th):	<input type="text"/>	
Product Ion Tolerance (Th):	<input type="text"/>	
Spectral network	<input type="checkbox"/>	
<input type="button" value="Submit"/>		

[Sample spectral datasets and documentation](#)

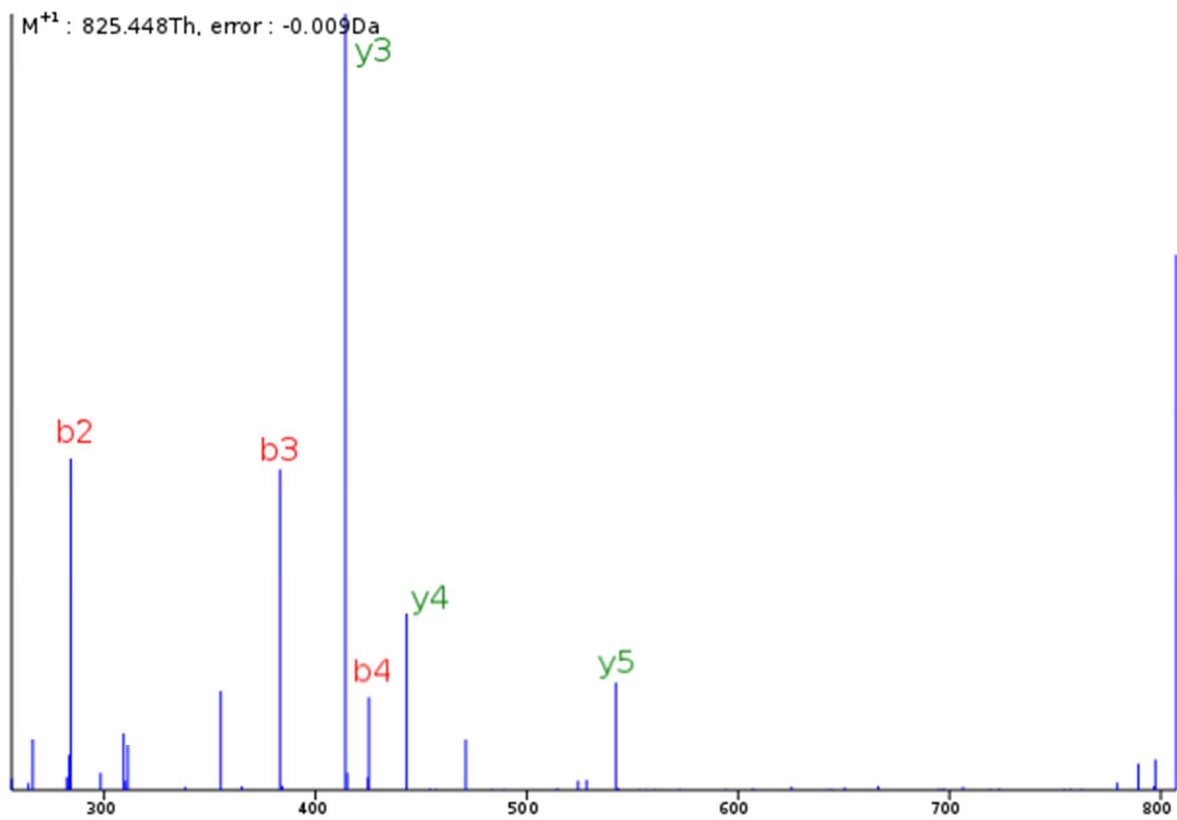
(a)

Scan Index	Spectrum Mass	Charge	Genome cluster	Score	Hmm Pvalue	Structure	Peptide
2622	1008.69	1	378344 ... 400366	16.00	1.55e-10	branch-cyclic	D_glu(+242.121);D_leu(-29.9146);[D_leu;D_val;asp;D_leu;leu]
2823	1036.71	1	378344 ... 400366	12.00	3.28e-10	branch-cyclic	D_glu(+240.12);D_leu;D_leu;D_val;asp;[D_leu;leu]
2901	1022.70	1	378344 ... 400366	19.00	1.57e-09	branch-cyclic	D_glu(+222.111);[D_leu;D_leu;D_val;asp;D_leu;leu(+4.10442)]
707	1074.84	1	378344 ... 400366	12.00	2.22e-09	branch-cyclic	[D_gln;D_leu(+220.11);D_leu;D_val(+59.2339);asp;D_leu;leu]
2669	1036.70	1	378344 ... 400366	13.00	3.15e-09	branch-cyclic	[D_glu(+258.129);D_leu;D_leu;D_val];asp(-17.9058);D_leu;leu]
2795	1008.68	1	378344 ... 400366	11.00	1.09e-08	linear	D_gln(+227.114);D_leu;D_leu;D_val;asp(-31.9387);D_leu;leu]
2979	1036.72	1	378344 ... 400366	19.00	1.68e-08	cyclic	[D_glu(+222.111);D_leu;D_leu;D_val;asp;D_leu;leu(+18.125)]
2851	1036.71	1	378344 ... 400366	9.00	5.53e-08	branch-cyclic	D_gln(+286.143);D_leu(-44.9331);D_leu;D_val;asp;[D_leu;leu]
3591	1077.93	1	378344 ... 400366	12.00	1.19e-07	branch-cyclic	D_gln;[D_leu;D_leu(+199.1);D_val;asp;D_leu(+83.3377);leu]
3116	882.64	1	378344 ... 400366	11.00	1.50e-07	branch-cyclic	[D_gln(-45.9216);D_leu;D_leu(+133.067);D_val;asp;]D_leu;leu]
3747	769.55	2	1994855 ... 1954732	10.00	4.84e-07	branch-cyclic	glu;D_D_orn;tyr;D_thr(+34.3715);D_glu;D_val;D_pro;[D_glu;D_D_tyr;ile(+118.059);phe]
3751	893.63	1	378344 ... 400366	11.00	6.03e-07	branch-cyclic	D_gln;D_leu(+140.07);[D_leu;D_val;asp;D_leu;leu(-41.9335)]
937	815.73	1	378344 ... 400366	11.00	6.50e-07	branch-cyclic	D_glu;D_leu;[D_leu;D_val(-32.7805);asp;D_leu(+52.026);leu]
319	621.26	2	1994855 ... 1954732	23.00	6.92e-07	branch-cyclic	[glu;D_D_orn(-53.0265);tyr;D_thr;D_glu;D_val;D_pro;D_glu;D_D_tyr(-91.1162)];ile;phe]
1638	757.49	1	378344 ... 400366	11.00	1.00e-06	branch-cyclic	[D_gln;D_leu(-13.0065);D_leu(-25.0027);D_val;asp;]D_leu;leu]
1674	879.49	1	378344 ... 400366	11.00	1.13e-06	branch-cyclic	[D_glu;D_leu;D_leu(+49.0245);D_val;asp;]D_leu;leu(+33.9864)
2751	849.50	1	378344 ... 400366	9.00	1.16e-06	branch-cyclic	[D_glu;D_leu(+56.028);D_leu;D_val];asp;D_leu;leu(-3.01043)
3830	931.60	1	378344 ... 400366	10.00	1.26e-06	branch-cyclic	D_glu;D_leu;[D_leu;D_val(+15.0608);asp(+120.06);D_leu;leu]
329	407.23	2	378344 ... 400366	15.00	1.40e-06	branch-cyclic	[D_glu;D_leu(-31.0578);D_leu;D_val(+48.024);asp;D_leu;leu]
412	472.31	3	1994855 ... 1954732	26.00	1.48e-06	branch-cyclic	[glu(+96.048);D_D_orn;tyr(-66.7867);D_thr;D_glu;D_val;D_pro;D_glu;D_D_tyr;ile];phe]
520	805.75	1	378344 ... 400366	10.00	1.53e-06	branch-cyclic	[D_glu(+26.013);D_leu;D_leu;D_val;asp;]D_leu(-16.744);leu]
226	424.15	2	378344 ... 400366	14.00	1.78e-06	branch-cyclic	D_gln;D_leu(-11.2287);[D_leu;D_val;asp(+63.0315);D_leu;leu]
1751	951.64	1	378344 ... 400366	12.00	1.80e-06	branch-cyclic	[D_glu;D_leu;D_leu;D_val(+67.1184);asp;D_leu(+88.044)];leu]
3382	868.61	1	378344 ... 400366	7.00	1.83e-06	branch-cyclic	[D_gln;D_leu(+20.0877)];D_leu;D_val;asp;D_leu(+53.0265);leu]
3049	936.78	1	378344 ... 400366	9.00	2.14e-06	branch-cyclic	D_glu(+96.048);D_leu;D_leu;D_val;[asp;D_leu;leu(+44.2528)]
1683	480.32	2	378344 ... 400366	16.00	2.17e-06	branch-cyclic	[D_gln(+204.102);D_leu(-39.9648);D_leu;D_val;asp;D_leu;leu]

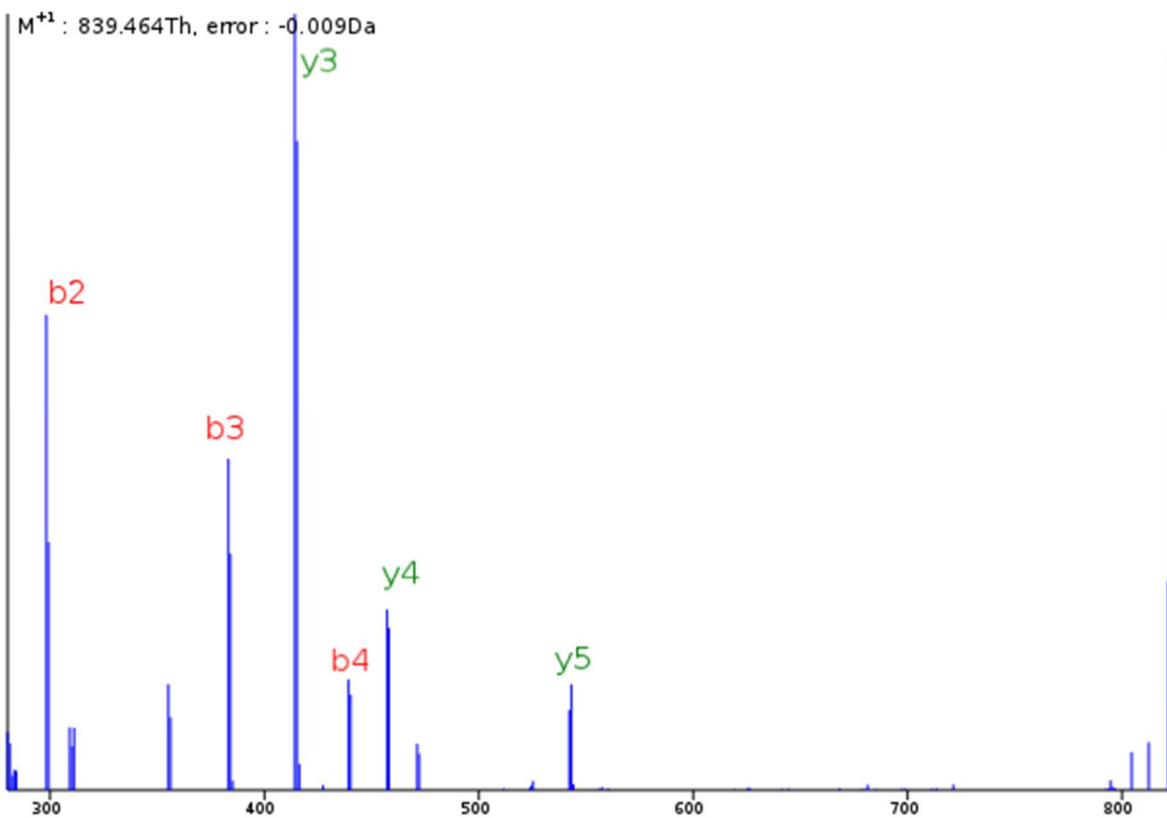
(b)

Fig. S1 NRPquest webserver (a) and results for *Bacillus subtilis* (b). The top hits with masses 1008, 1022, 1036 and 1074 represent surfactins.

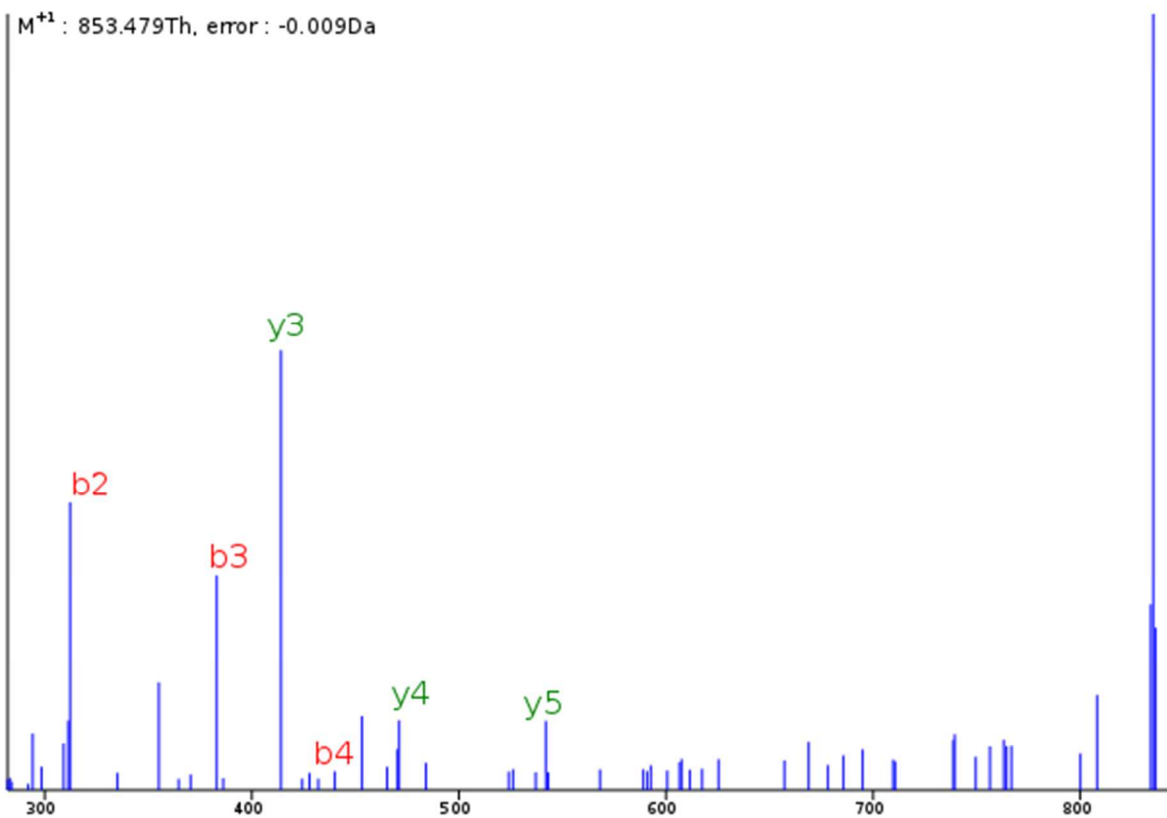
The whole analysis takes about an hour. The webserver and datasets are available at cyclo.ucsd.edu.



(a)



(b)



(c)

Fig. S2 Annotation of the spectra of (a) arylomycin A2, (b) arylomycin A4 and (c) arylomycin A5