

**Large-scale transcriptome mining enables macrocyclic diversification and improved
bioactivity of the stephanotic acid scaffold**

Wang *et al.*

Supplementary Table 1. Plant genomes and plant RNA-seq data in the NCBI sequence read archive.

NCBI Plant genomes (May 2024)	#
Assemblies	1869
Assembly status	
Complete genome	25
Chromosome	868
Scaffold	270
Contig	706
Assemblies with annotation in nr protein database	414
Species	1728
Species with annotation in nr protein database	386
Represented families	225
Rep. families with annotation in nr protein database	112
RNA-seq datasets (SRA) (May 2024)	#
Species	>9000
Transcriptome assemblies (this study)	8897
Species (this study)	7579
Represented families (this study)	498

Supplementary Table 2. Benchmarking of read trimming for *de novo* transcriptome assembly of plant RiPP precursor discovery. 27 RNA-seq datasets were assembled which include 36 precursor genes with 138 core peptides and 65 unique core peptides. Each RNA-seq dataset was assembled with prior read trimming (TrimGalore) or without read trimming in triplicate (n=3) with the standard deviation shown as an error.

<i>De novo</i> assembler	Total time [min]	Total memory usage [GB]	Total detected cores (138 total cores)	Detected unique correct cores (65 total unique cores)	Detected repetitive cores (120 total repetitive cores)	Similarity [sum %] (total similarity: 3600%)	Identity [sum %] (total identity: 3600%)	Gaps [sum %] (total gaps: 3600%)
SPAdes (v3.15.5)	2041 ± 273	269.9 ± 6.6	88	44	69	2948	2915	609
TrimGalore (v0.6.7) + SPAdes (v3.15.5)	(439 + 1628 = 2067) ± 241	(63 + 241 = 304) ± 3	93	48	74	3012	2981	545

Supplementary Table 3. Comparison of scaled transcriptome mining for fused moroidin cyclases and moroidin core peptides in published 1kp SOAPdenovo-trans assemblies and SPAdes-assemblies. Source data are provided as a Source Data file.

Assembly	Species	Detected fused moroidin cyclases	Detected moroidin core peptides
SOAPdenovo-trans (1kp study ⁴)	3 (<i>Lomandra longifolia</i> , <i>Bauhinia tomentosa</i> , <i>Kerria japonica</i>)	3	5 (QLLVWREH, QLLVWRSH, QLLFWFGRH, QLLVWFARH, QLLVWRAH)
SPAdes (v3.15.5, this study)	5 (<i>Kerria japonica</i> , <i>Lomandra longifolia</i> , <i>Orchidantha maxillarioides</i> , <i>Bauhinia tomentosa</i> , <i>Cercis canadensis</i>)	6	14 (QLLFWFGRH, QLLFWTRH, QLLVWFGRH, QLLVWPKH, QLLVWQKH, QLLVWRAH, QLLVWREH, QLLVWRGH, QLLVWRH, QLLVWRNH, QLLVWRSH, QLLVWSRH, QLLVWTDH, QLLVWTRH)

#	SRA#	Organism	Family	Predicted core peptides (stephanotic acid in bold , 1x = single core precursor)	Stephanotic acid Core: QLxxW	Moroidin Core: QLxxWx(xx)H	2.ring-Trp Core: QLxxWx(xx)W	2.ring-Tyr Core: QLxxWx(xx)Y
1	SRR12990041	<i>Aganosma wallichii</i>	Apocynaceae	QLLVWRW (1x)	x		x	
2	SRR6456116	<i>Stellaria aquatica</i>	Caryophyllaceae	QLLVVRNH, QLFVVRNH, QLLVWSKH	x	x		
3	SRR5219180	<i>Mercurialis huetii</i>	Euphorbiaceae	QLFFWRY, QLFFWKY	x			x
4	ERR2597552	<i>Mercurialis annua</i>	Euphorbiaceae	QLFFWKY	x			x
5	SRR8797417	<i>Medicago archiducis-nicolai</i>	Fabaceae	QLFVWGGH	x	x		
6	ERR6688729	<i>Medicago arabica</i>	Fabaceae	QLLIWGGH, QLLVWGGH, QLFVWGGH	x	x		
7	SRR3141655	<i>Vigna mungo</i>	Fabaceae	QLLVWRW (1x)	x		x	
8	SRR3938646	<i>Medicago marina</i>	Fabaceae	QLLWWGGH (1x)	x	x		
9	ERR2021956	<i>Bauhinia blakeana</i>	Fabaceae	QLLVWRQH, QLLVWKIH, QLLIWRQH, QLLVWRGH, QLLVWKGH, QLLVWRSH,	x	x		
10	DRR168799	<i>Bauhinia forficata</i>	Fabaceae	QLRVWTSH, QLSVWTSH, QLRVWQSH, QLLIWMNH, QLLVWRKNH	x	x		
11	ERR2021957	<i>Bauhinia purpurea</i>	Fabaceae	QLSEW, QLLVVRNH, QLLVWRGH, QLVWRRNH, QLFVWRGH, QLRVWTKH, QLLVWREH, QLIVVRNH, QLLVWRRGH, QLLVWRRNH, QLFVWSDH, QLFVWDDGH, QLLVWSSDH, QLFVVGGVH, QLFVWGCH, QLLVWVGH, QLLVWGCGH, QLFVWGCGH, QLLVWGGDH,	x	x		
12	SRR16068779	<i>Medicago sativa</i>	Fabaceae		x	x		
14	SRR18083643	<i>Glechoma longituba</i>	Lamiaceae	QLLVWRW, QLLVWKW	x		x	
15	SRR5150732	<i>Glechoma hederacea</i>	Lamiaceae	QLFVWGW, QLLVWKW, QLLVWRW	x		x	
16	ERR3487362	<i>Salvia sp.</i> IKP	Lamiaceae	QLFVWGW	x			
17	ERR2040758	<i>Orchidantha maxillarioides</i>	Lowiaceae	QLLVKW, QLLVWRSH, QLFNMYTH, QLLVWPKH, QLLVWTDH	x	x	x	
18	SRR16483051	<i>Cinna latifolia</i>	Poaceae	QLLVW, QLLVWRTY, QLLVWRQY, QLLVWRSY, QLFIWRSY, QLFIWRSY, QLLWRSY, QLFIWRSY, QLLVWRSY, QLLIWRTY, QLLVWRTY, QLLVWRQY, QLLWRSY, QLLVWRSY, QLFIWRSY, QLLIWRTY, QLLVWRTY, QLFIWRSY, QLLIWRRTY,	x			x
19	SRR5406165	<i>Digitaria cuyabensis</i>	Poaceae	QLLVVRNH	x	x		
20	SRR16234948	<i>Stipa sibirica</i>	Poaceae	QLLVWRKSW, QLLVWRESW	x		x	
21	SRR13316951	<i>Cinnagrostis crispa</i>	Poaceae	QLLVWRDW	x		x	
22	SRR16483054	<i>Holcus mollis</i>	Poaceae	QLLVWRSY, QLLIWRTY, QLLVWRTY, QLLVWRQY	x			x
23	SRR16243467	<i>Panicum brevifolium</i>	Poaceae	QLLVVRNH	x	x		
24	SRR13316952	<i>Cinnagrostis cabreriae</i>	Poaceae	QLLVWRDW	x		x	
25	SRR23873622	<i>Aegilops uniaristata</i>	Poaceae	QLLIW	x			
26	SRR16299781	<i>Leptothrium senegalense</i>	Poaceae	QLLVWRSH, QLLVVRNH	x	x		
27	SRR16483114	<i>Neomolinia japonica</i>	Poaceae	QLLVWRSW	x		x	
28	ERR3634426	<i>Brachypodium boissieri</i>	Poaceae	QLLVWRSW	x		x	
29	SRR12955278	<i>Bromus catharticus</i>	Poaceae	QLLVWRSW	x		x	
30	SRR6221583, SRR6221613	<i>Sacciolepis striata</i>	Poaceae	QLLVVRNH	x	x		
31	SRR5406195	<i>Zuloagaea bulbosa</i>	Poaceae	QLLVVRNH	x	x		
32	ERR9784826	<i>Dichrosta-chys cinerea</i>	Fabaceae	QLLVWGY	X			X
33	SRR7817914	<i>Hordeum brevisubula-tum</i>	Poaceae	QLRVWRKY (1x)	x			x
34	SRR23631370	<i>Amomum smithiae</i>	Zingiberaceae	QLLVW (1x)	x			
		Known producers						
35	ERR2040423	<i>Kerria japonica</i>	Rosaceae	QLLVWRAH, QLLVWRGH	x	x		
36	ERR706821	<i>Bauhinia tomentosa</i>	Fabaceae	QLLVWRSH, QLLVVRNH, QLLFWGRH, QLLVWSRH, QLLVWTRH, QLLFWTRH,	x	x		
37	ERR706845	<i>Cercis canadensis</i>	Fabaceae	QLLVW, QLKVW, QLLVWOKH, QLLVWRH	x	x		

Supplementary Table 5. Predicted core peptides of stephanotic acid-type burpitides in assembled plant transcriptomes. The number of different sequences from fused burpitide cyclases and split/incomplete transcripts (*italicized*, PHI-BLAST result only) are displayed. Source data are provided as a Source Data file.

#	Moroidin cores	Glechomanin cores	Mercurialin cores	Stephanotic acid core
1	QLLVWRH	QLFVWGW	QLFFWKY	QLFFW
2	QLSFWKH	QLLLWKW	QLFFWRY	QLFIW
3	QLFNWYTH	QLLVWKW	QLLVWGY	QLFNW
4	QLFVWGGH	QLLVWRW	QLFIWRSY	QLFVW
5	QLFVWRGH	QLLVWRDW	QLLIWRTY	QLIVW
6	QLFVWRNH	QLLVWRSW	QLLLWRSY	QLLFW
7	QLFVWSDH	QLLVWRESW	QLLVWRQY	QLLIW
8	QLIVWRNH	QLLVWGW	QLLVWRSY	QLLLW
9	QLLFWTRH	QLLVWCAW	QLLVWRTY	QLLVW
10	QLLIWGGH	QLLVWQSW	QLRVWRKY	QLLWW
11	QLLIWKGH	QLLVWRAW	QLFIWRRSY	QLRVW
12	QLLIWMNH	QLLVWREW	QLLIWRTY	QLSFW
13	QLLIWRQH	Fused: 8	QLLLWRRSY	QLSVW
14	QLLVWGGH	Split/ Incomplete: 5	QLLVWRRSY	QLLMW
15	QLLVWIKH		QLLVWRTY	QLPVW
16	QLLVWPKH		QLVPWEGY	Fused: 13
17	QLLVWQKH		Fused: 15	Split/ Incomplete: 2
18	QLLVWRAH		Split/ Incomplete: 1	
19	QLLVWREH			
20	QLLVWRGH			
21	QLLVWRNH			
22	QLLVWRQH			
23	QLLVWRSH			
24	QLLVWSKH			
25	QLLVWSRH			
26	QLLVWTDH			
27	QLLVWTRH			
28	QLLVWVGH			
29	QLLVWGGH			
30	QLRVWQSH			
31	QLRVWTKH			
32	QLRVWTSH			
33	QLSVWTSH			
34	QLFVWAAGH			
35	QLFVWDDGH			
36	QLFVWGGGH			
37	QLFVWGGVH			
38	QLIVWRRNH			
39	QLLFWFGRH			
40	QLLVWFGRH			
41	QLLVWGGDH			
42	QLLVWGGGH			
43	QLLVWRKNH			
44	QLLVWRRGH			
45	QLLVWRRNH			
46	QLLVWSSDH			
47	QLFVWRGH			
48	QLFIWVRH			
49	QLFVWRSH			
50	QLLIWPRH			
51	QLLIWRNH			
52	QLLMWRGH			
53	QLLVWHNH			
54	QLLVWKTH			
55	QLLVWLNH			
56	QLLVWPEH			
57	QLLVWPRH			
58	QLLVWRTH			
59	QLPVWRGH			
60	QLRVWTQH			
61	QLVPWEGH			
62	QLLVWPPGH			
63	QLLVWPPRH			
64	QLLVWPPSH			
65	QLLVWRRH			
66	QLLVWRRSH			
		Fused: 46		
		Split/ Incomplete: 20		

Supplementary Table 6. ¹H-NMR (800 MHz) and ¹³C-NMR (200 MHz) data for glechomanin in DMSO-d₆.

Residue	C	$\delta(^{13}\text{C})$ [ppm]	H	$\delta(^1\text{H})$ (int, m, J) [ppm]	COSY	TOCSY	HMBC	ROESY
PyroGlu ¹	α	55.2	α	3.98 (1H, dd, 7.9, 4.4)	PyroGlu ¹ : H β 1, H β 2, NH	PyroGlu ¹ : H β 1, H β 2, Hy, NH	PyroGlu ¹ : C β	PyroGlu ¹ : H β 1, H β 2, NH, Leu ² : NH
	β	24.8	β 1	2.10 ^a (1H)	PyroGlu ¹ : Ha, H β 2, Hy	PyroGlu ¹ : Ha, H β 2, Hy, NH	PyroGlu ¹ : Ca, Cy, C δ , C=O	PyroGlu ¹ : Ha, H β 2
	γ	29.3	β 2	1.61 (1H, br s)	PyroGlu ¹ : Ha, H β 1, Hy	PyroGlu ¹ : Ha, H β 1, Hy, NH,	n/a	PyroGlu ¹ : Ha, H β 1, Hy
	δ	177.2	γ	2.10 ^a (2H)	PyroGlu ¹ : Ha, H β 2	PyroGlu ¹ : Ha, H β 2, NH	PyroGlu ¹ : Ca, C β , C δ	PyroGlu ¹ : H β 2
	C=O	171.2	NH	7.74 (1H, s)	PyroGlu ¹ : Ha	PyroGlu ¹ : Ha, H β 1, H β 2, Hy	PyroGlu ¹ : Ca, C β , Cy, C δ	PyroGlu ¹ : Ha, Leu ² : NH
Leu ²	α	55.1	α	4.72 (1H, br t, 10.3)	Leu ² : NH, H β	Leu ² : H β , Hy, H δ 1, H δ 2, NH	n/a	Leu ² : H β , Hy, H δ 2, NH, Phe ³ : NH, Trp ⁵ : H5
	β	52.2	β	2.87 (1H, dd, 11.6, 3.1)	Leu ² : Ha, Hy	Leu ² : Ha, Hy, H δ 1, H δ 2, NH	Leu ² : Ca, Cy, C δ 1, C δ 2, C=O, Trp ⁵ : C5, C6, C7	Leu ² : Ha, Hy, H δ 1, H δ 2, NH, Phe ³ : NH, Trp ⁵ : H7
	γ	26.6	γ	2.05 ^a (1H)	Leu ² : H β , H δ 1, H δ 2	Leu ² : Ha, H β , H δ 1, H δ 2, NH	n/a	Leu ² : Ha, H β , H δ 1, H δ 2, NH, Trp ⁵ : H7
	δ 1	21.8	δ 1	0.66 (3H, d, 6.7)	Leu ² : Hy	Leu ² : Ha, H β , Hy, H δ 2, NH	Leu ² : C β , Cy, C δ 2	Leu ² : H β , Hy, H δ 2, Trp ⁵ : H7
	δ 2	17.3	δ 2	0.82 (3H, br s)	Leu ² : Hy	Leu ² : Ha, H β , Hy, H δ 1, NH	n/a	Leu ² : Ha, H β , Hy, H δ 1, Trp ⁵ : H5
	C=O	169.1	NH	8.09 (1H, d, 9.4)	Leu ² : Ha	Leu ² : Ha, H β , Hy, H δ 1, H δ 2	PyroGlu ¹ : C=O	PyroGlu ¹ : Ha, H β 2, Leu ² : Ha, H β , Hy
Phe ³	α	55.6	α	4.26 (1H, br t, 10.5)	Phe ³ : H β 1, H β 2, NH	Phe ³ : H β 1, H β 2, NH	Phe ³ : C β , C=O	Phe ³ : H β 1, H β 2, H2, H6, NH, Val ⁴ : NH
	β	38.9	β 1	2.78 (1H, br d, 11.6)	Phe ³ : Ha	Phe ³ : Ha, NH	n/a	Phe ³ : Ha, H2, H6, NH, Val ⁴ : Hy1, NH
			β 2	2.85 ^a (1H)	Phe ³ : Ha	Phe ³ : Ha, NH	n/a	Phe ³ : Ha, H2, H6, NH, Val ⁴ : Hy1, NH
	C1	137.5	H2	7.11 (2H, d, 7.2)	Phe ³ : H3	Phe ³ : H3, H4, H5, H6	Phe ³ : Ca, C4, C6	Phe ³ : Ha, H β 1, H β 2, H3
	C2	128.9	H3	7.16 (2H, t, 7.2)	Phe ³ : H2, H4	Phe ³ : H2, H4, H5, H6	Phe ³ : C1, C5	Phe ³ : H2, H4
	C3	127.9	H4	7.13 (1H, t, 7.2)	Phe ³ : H3, H5	Phe ³ : H2, H3, H5, H6	Phe ³ : C2, C6	Phe ³ : H3, H5
	C4	126.0	H5	7.16 (2H, t, 7.2)	Phe ³ : H4, H6	Phe ³ : H2, H3, H4, H6	Phe ³ : C1, C3	Phe ³ : H4, H6
	C5	127.9	H6	7.11 (2H, d, 7.2)	Phe ³ : H5	Phe ³ : H2, H3, H4, H5	Phe ³ : Ca, C2, C4	Phe ³ : Ha, H β 1, H β 2, H5
	C6	128.9	NH	8.35 (1H, s)	Phe ³ : Ha	Phe ³ : Ha, H β 1, H β 2	n/a	Phe ³ : Ha, H β 2, Leu ² : Ha, Val ⁴ : NH, Trp ⁵ : H5
	C=O	169.8						
Val ⁴	α	59.7	α	4.41 (1H, s)	Val ⁴ : NH	Val ⁴ : NH	n/a	Val ⁴ : Hy1, Trp ⁵ : NH, Trp ⁷ : H5
	β	43.2	γ 1	1.05 (3H, s)	n/a	Val ⁴ : Hy2	Val ⁴ : Ca, C β , Cy2, Trp ⁷ : C6	Val ⁴ : Ha, Hy2, Trp ⁷ : H7
	γ 1	29.6	γ 2	1.45 (3H, s)	n/a	Val ⁴ : Hy1	Val ⁴ : Ca	Val ⁴ : Hy1, NH, Trp ⁷ : H5, H7
	γ 2	21.5	NH	7.56 (1H, s)	Val ⁴ : Ha	Val ⁴ : Ha	n/a	Val ⁴ : Ha, Hy2, Phe ³ : H β 2, NH
	C=O	170.3 ^b						
Trp ⁵	α	51.2	α	5.02 (1H, br s)	Trp ⁵ : H β 1, NH	Trp ⁵ : H β 1, H β 2, NH	n/a	Trp ⁵ : H β 1, H4
	β	27.9	β 1	2.71 (1H, dd, 14.0, 11.4)	Trp ⁵ : Ha, H β 2	Trp ⁵ : Ha, H β 2, NH	n/a	Trp ⁵ : Ha, H β 2, H2, NH
	C2	123.8	β 2	3.29 ^a (1H)	Trp ⁵ : Ha	Trp ⁵ : NH, Ha	n/a	Trp ⁵ : H β 1
	C3	108.8	(NH)1	10.52 (1H, s)	Trp ⁵ : H2	Trp ⁵ : H2	Trp ⁵ : C3	Trp ⁵ : H2, H7
	C4	118.3	H2	6.89 (1H, s)	Trp ⁵ : (NH)1	Trp ⁵ : (NH)1	n/a	Trp ⁵ : H β 1, (NH)1
	C5	118.2	H4	7.37 (1H, d, 8.4)	Trp ⁵ : H5	Trp ⁵ : H5	Trp ⁵ : C3, C6, C8	Trp ⁵ : Ha, H β 2, H5
	C6	129.3	H5	6.86 (1H, br d, 8.4)	Trp ⁵ : H4	Trp ⁵ : H4	Trp ⁵ : C7, C9	Leu ² : Ha, H δ 2, Trp ⁵ : H4
	C7	113.8	H7	6.77 (1H, s)	n/a	Trp ⁵ : H4	Leu ² : C β , Trp ⁵ : C5, C9	Leu ² : H β , Hy, H δ 1, Trp ⁵ : (NH)1
	C8	136.2	NH	7.64 (1H, s)	Trp ⁵ : Ha	Trp ⁵ : Ha, H β 1, H β 2	n/a	Val ⁴ : Ha, Trp ⁵ : H β 1
	C9	124.9						
Gly ⁶	α	41.7	α 1	3.96 ^a (1H)	Gly ⁶ : Ha2	Gly ⁶ : Ha2, NH	n/a	Gly ⁶ : Ha2
	C=O	167.3 ^b	α 2	3.30 ^a (1H)	Gly ⁶ : Ha1	Gly ⁶ : NH	n/a	Gly ⁶ : Ha1
Trp ⁷			NH	7.86 ^a (1H)	n/a	Gly ⁶ : Ha1, Ha2	n/a	Gly ⁶ : Ha2
	α	54.8	α	4.52 (1H, br s)	Trp ⁷ : H β 1, H β 2, NH	Trp ⁷ : H β 1, H β 2, NH	n/a	Trp ⁷ : H β 1, H β 2, NH
	β	29.2	β 1	3.22 (1H, br d, 13.2)	Trp ⁷ : Ha, H β 2	Trp ⁷ : Ha, H β 2, NH	Trp ⁷ : C2, C3, C9	Trp ⁷ : Ha, H β 2, H2, H4
	C2	124.1	β 2	2.84 ^a (1H)	Trp ⁷ : Ha, H β 1	Trp ⁷ : Ha, H β 1, NH	n/a	Trp ⁷ : Ha, H β 1, H4
	C3	110.8	(NH)1	10.69 (1H, s)	Trp ⁷ : H2	Trp ⁷ : H2	n/a	Trp ⁷ : H2, H7
	C4	117.8	H2	7.19 (1H, s)	Trp ⁷ : (NH)1	Trp ⁷ : (NH)1	n/a	Trp ⁷ : H β 1, (NH)1
	C5	115.6	H4	7.72 ^a (1H)	Trp ⁷ : H5	Trp ⁷ : H5, H7	n/a	Trp ⁷ : H β 1, H β 2, H5, NH
	C6	137.8	H5	6.79 ^a (1H)	Trp ⁷ : H4	Trp ⁷ : H4, , H7	n/a	Val ⁴ : Ha, Hy1, Hy2, Trp ⁷ : H4,
	C7	110.6	H7	7.42 (1H, s)	n/a	Trp ⁷ : H4, H5	n/a	Val ⁴ : Hy1, Hy2, Trp ⁷ : (NH)1
	C8	136.5	NH	8.49 (1H, s)	Trp ⁷ : Ha	Trp ⁷ : Ha, H β 1, H β 2	n/a	Gly ⁶ : Ha2, Trp ⁷ : Ha, H4
	C9	126.7						
	C=O	172.9 ^b						

^aoverlapping signals, ^bsignals cannot be assigned by 2D-NMR, n/a – not detected, br s – broad singlet, d – doublet, dd – double doublet, t – triplet, m – multiplet.

Supplementary Table 7. ¹H-NMR (800 MHz) and ¹³C-NMR (200 MHz) data for glechomanin in methanol-d₄.

Residue	C	$\delta(^{13}\text{C})$ [ppm]	H	$\delta(^1\text{H})$ (int, m, J) [ppm]	COSY	TOCSY	HMBC	ROESY
PyroGlu ¹	α	56.4	α	4.16 (1H, m)	PyroGlu ¹ : H β 1, H β 2	PyroGlu ¹ : H β 1, H β 2, H γ	n/a	PyroGlu ¹ : H β 1, H β 2, Leu ² : NH
	β	25.3	β 1	1.75 (1H, m)	PyroGlu ¹ : Ha, H β 2, H γ	PyroGlu ¹ : Ha, H β 2, H γ	n/a	PyroGlu ¹ : Ha, H β 2
	γ	29.2	β 2	2.34 ^a (1H)	PyroGlu ¹ : Ha, H β 1, H γ	PyroGlu ¹ : Ha, H β 1, H γ	PyroGlu ¹ : C γ	PyroGlu ¹ : Ha, H β 1
	δ	n/a	γ	2.35 ^a (2H)	PyroGlu ¹ : Ha, H β 1	PyroGlu ¹ : Ha, H β 1	n/a	PyroGlu ¹ : H β 1
	C=O	n/a	NH	n/a	n/a	n/a	n/a	n/a
Leu ²	α	56.6	α	4.78 (1H, t, 10.5)	Leu ² : NH, H β	Leu ² : H β , H γ , H δ 1, H δ 2, NH	Leu ² : C β	Leu ² : H β , H δ 2, NH, Phe ³ : NH, Trp ⁵ : H5
	β	52.7	β	3.00 (1H, br d, 11.7)	Leu ² : Ha, H γ	Leu ² : Ha, H γ , H δ 1, H δ 2, NH	n/a	Leu ² : Ha, H γ , H δ 1, H δ 2, NH, Phe ³ : NH, Trp ⁵ : H7
	γ	27.6	γ	2.18 (1H, m)	Leu ² : H β , H δ 1, H δ 2	Leu ² : Ha, H β , H δ 1, H δ 2, NH	Leu ² : C δ 2	Leu ² : Ha, H β , H δ 1, H δ 2, NH, Trp ⁵ : H7
	δ 1	20.7	δ 1	0.76 (3H, d, 6.4)	Leu ² : H γ	Leu ² : Ha, H β , H γ , H δ 2, NH	n/a	Leu ² : H β , H γ , Trp ⁵ : H7
	δ 2	16.9	δ 2	0.98 (3H, d, 6.8)	Leu ² : H γ	Leu ² : Ha, H β , H γ , H δ 1, NH	Leu ² : C δ 1	Leu ² : Ha, H β , H γ , Trp ⁵ : H5
	C=O	n/a	NH	8.06 ^a	Leu ² : Ha	Leu ² : Ha, H β , H γ , H δ 1, H δ 2	n/a	PyroGlu ¹ : Ha, H β 2, Leu ² : Ha, H β , H γ
Phe ³	α	55.9	α	4.64 (1H, m)	Phe ³ : H β 1, H β 2, NH	Phe ³ : H β 1, H β 2, NH	n/a	Phe ³ : H β 1, H β 2, H2, H6, NH,
	β	39.7	β 1	3.00 (1H, br d, 11.4)	Phe ³ : Ha, H β 2	Phe ³ : Ha, H β 2, NH	n/a	Phe ³ : Ha, H β 2, H2, H6, NH, Val ⁴ : H γ 1
			β 2	2.76 (1H, t, 12.4)	Phe ³ : Ha, H β 1	Phe ³ : Ha, H β 1, NH	n/a	Phe ³ : Ha, H β 1, H2, H6, NH, Val ⁴ : H γ 1, H γ 2
	C1	n/a	H2	7.24 ^a (1H)	Phe ³ : H3	Phe ³ : H3, H4, H5, H6	Phe ³ : C4	Phe ³ : Ha, H β 1, H β 2, H3
	C2	128.9	H3	7.23 ^a (1H)	Phe ³ : H2, H4	Phe ³ : H2, H4, H5, H6	Phe ³ : C5	Phe ³ : H2, H4
	C3	128.0	H4	7.18 ^a (1H)	Phe ³ : H3, H5	Phe ³ : H2, H3, H5, H6	Phe ³ : C2, C6	Phe ³ : H3, H5
	C4	126.3	H5	7.23 ^a (1H)	Phe ³ : H4, H6	Phe ³ : H2, H3, H4, H6	Phe ³ : C3	Phe ³ : H4, H6
	C5	128.0	H6	7.24 ^a (1H)	Phe ³ : H5	Phe ³ : H2, H3, H4, H5	Phe ³ : C4	Phe ³ : Ha, H β 1, H β 2, H5
	C6	128.9	NH	8.08 ^a (1H)	Phe ³ : Ha	Phe ³ : Ha, H β 1, H β 2	n/a	Leu ² : Ha, Phe ³ : Ha, H β 1, H β 2
	C=O	n/a						
Val ⁴	α	60.7	α	4.45 (1H, d, 8.5)	Val ⁴ : NH	Val ⁴ : NH	n/a	Val ⁴ : H γ 1, Trp ⁵ : NH, Trp ⁷ : H5
	β	42.7	γ 1	1.10 (3H, s)	n/a	Val ⁴ : H γ 2	Val ⁴ : C β , C γ 2, Trp ⁷ : C6	Val ⁴ : Ha, Trp ⁷ : H7
	γ 1	28.9	γ 2	1.47 (3H, s)	n/a	Val ⁴ : H γ 1	Val ⁴ : C γ 2	Val ⁴ : H γ 1, NH, Trp ⁷ : H5, H7
	γ 2	20.1	NH	7.25 ^a (1H)	Val ⁴ : Ha	Val ⁴ : Ha	n/a	Val ⁴ : Ha, Phe ³ : H β 1, H β 2, NH
	C=O	n/a						
Trp ⁵	α	51.2	α	5.23 (1H, br s)	Trp ⁵ : H β 1, NH	Trp ⁵ : H β 1, H β 2 NH	n/a	Trp ⁵ : H β 1, H4
	β	27.5	β 1	2.86 (1H, t, 13.5)	Trp ⁵ : Ha, H β 2	Trp ⁵ : Ha, H β 2, NH	n/a	Trp ⁵ : Ha, H β 2, H2, NH
	C2	124.2	β 2	3.48 (1H, dd, 14.3, 4.8)	Trp ⁵ : Ha, H β 1	Trp ⁵ : Ha, H β 1, NH	Trp ⁵ : C α	Trp ⁵ : Ha, H β 1, H4
	C3	107.6	(NH)1	n/a	n/a	n/a	n/a	n/a
	C4	118.8	H2	6.86 (1H, s)	n/a	Trp ⁵ : H β 1, H β 2	Trp ⁵ : C3	Trp ⁵ : H β 1
	C5	117.8	H4	7.46 (1H, d, 8.3)	Trp ⁵ : H5	Trp ⁵ : H5	Trp ⁵ : C3, C6, C8	Trp ⁵ : Ha, H β 2
	C6	127.2	H5	6.97 (1H, d, 8.3)	Trp ⁵ : H4	Trp ⁵ : H4	Trp ⁵ : C7, C9	Leu ² : Ha, H β , H δ 2, Phe ³ : NH, Trp ⁵ : H4
	C7	114.0	H7	6.92 (1H, s)	n/a	Trp ⁵ : H4	Trp ⁵ : C5	Leu ² : H β , H δ 1
	C8	145.8	NH	7.19 ^a (1H)	Trp ⁵ : Ha	Trp ⁵ : Ha, H β 1, H β 2	n/a	Val ⁴ : Ha, Trp ⁵ : H β 1
	C9	125.3						
Gly ⁶	α	41.2	α 1	3.35	Gly ⁶ : Ha2	Gly ⁶ : Ha2	n/a	Gly ⁶ : Ha2, Trp ⁷ : Ha,
	C=O	n/a	α 2	4.31	Gly ⁶ : Ha1	Gly ⁶ : Ha1	n/a	Gly ⁶ : Ha1
			NH	n/a	n/a	n/a	n/a	n/a
Trp ⁷	α	55.5	α	4.73 (1H, br d, 11.2)	Trp ⁷ : H β 1, H β 2	Trp ⁷ : H β 1, H β 2,	n/a	Gly ⁶ : Ha1, Trp ⁷ : H β 2
	β	29.3	β 1	3.00 (1H, m)	Trp ⁷ : Ha, H β 2	Trp ⁷ : Ha, H β 2	n/a	Trp ⁷ : Ha, H β 2, H2, H4
	C2	124.3	β 2	3.41 ^a (1H)	Trp ⁷ : Ha, H β 1	Trp ⁷ : Ha, H β 1,	Trp ⁷ : C α	Trp ⁷ : H β 1, H2
	C3	n/a	(NH)1	n/a	n/a	n/a	n/a	n/a
	C4	117.6	H2	7.21 ^a	n/a	n/a	n/a	Trp ⁷ : H β 1, H β 2
	C5	115.5	H4	7.64 (1H, d, 8.0)	Trp ⁷ : H5	Trp ⁷ : H5, H7	Trp ⁷ : C7	Trp ⁷ : H β 1, H β 2, H5
	C6	138.1	H5	6.97 (1H, d, 8.0)	Trp ⁷ : H4	Trp ⁷ : H4, H7	n/a	Val ⁴ : H γ 1, H γ 2, Trp ⁷ : Ha, H4
	C7	110.6	H7	7.55 (1H, s)	n/a	Trp ⁷ : H4, H5	Trp ⁷ : C5	Val ⁴ : Ha, H γ 1, H γ 2
	C8	n/a	NH	n/a	n/a	n/a	n/a	n/a
	C9	n/a						
	C=O	n/a						

^aoverlapping signals, n/a – not detected, br s – broad singlet, d – doublet, dd – double doublet, t – triplet, m – multiplet.

Supplementary Table 8. Tandem MS fragmentation of linear tryptic core peptide of GheBURP-1xQLFVWGW. Proteomic data: MSV000097169

Fragment (z)	m(obs)	m(calc)	Δm [ppm]
b2 (1+)	228.1342	228.1343	0.4
b3 (1+)	329.1817	329.1820	0.9
y2 (1+)	333.1919	333.1922	0.9
y3 (1+)	390.2133	390.2136	0.8
b4 (1+)	430.2295	430.2297	0.5
y4 (1+)	576.2935	576.2929	1
b5 (1+)	593.2931	593.2930	0.2
y5 (1+)	675.3613	675.3614	0.1
b6 (1+)	707.3354	707.3359	0.7
b7 (1+)	778.3722	778.3730	1
y6 (1+)	822.4287	822.4298	1.3
b8 (1+)	849.4096	849.4102	0.7
y7 (1+)	935.5138	935.5138	0
b9 (1+)	936.4405	936.4422	1.8
b10 (1+)	1051.4697	1051.4691	0.6
y8 (1+)	1063.5707	1063.5724	1.6
b11 (1+)	1165.5113	1165.5120	0.7
y9 (1+)	1177.6152	1177.6153	0.1
y10 (1+)	1234.6355	1234.6368	1.1
b13 (1+)	1336.5736	1336.5764	2.1
y11 (1+)	1348.6780	1348.6797	1.3
y12 (1+)	1463.7054	1463.7067	0.9
b14 (1+)	1464.6520	1464.6350	11.6
y13 (1+)	1550.7360	1550.7387	1.7
b15 (1+)	1577.7193	1577.7191	0.1
y14 (1+)	1621.7744	1621.7758	0.9
y15 (1+)	1692.8098	1692.8129	1.8
b16 (1+)	1724.7881	1724.7875	0.3
y16 (1+)	1806.8535	1806.8559	1.3
b17 (1+)	1823.8510	1823.8559	2.7
y17 (1+)	1969.9164	1969.9192	1.4
b18 (1+)	2009.9360	2009.9352	0.4
y18 (1+)	2070.9632	2070.9669	1.8
y19 (1+)	2172.0123	2172.0145	1
b21 (2+)	2381.1300	2381.1310	0.4
y21 (2+)	2399.1370	2399.1415	1.9

Supplementary Table 9. Tandem MS fragmentation of putative bicyclic tryptic core peptide of GheBURP-1xQLFVWGW. Proteomic data: MSV000097169

Fragment (z)	m(obs)	m(calc)	Δm [ppm]
b2 (1+)	228.1341	228.1343	0.9
b3 (1+)	329.1816	329.1820	1.2
b5 (1+)	593.2915	593.293	2.5
b6 (1+)	707.3357	707.3359	0.3
y7* (1+)	931.4817	931.4826	1.0
b7 (1+)	778.3717	778.3730	1.7
y8* (1+)	1059.5388	1059.5411	2.2
b8 (1+)	849.4096	849.4102	0.7
y9* (1+)	1173.5830	1173.5841	0.9
b9 (1+)	936.4404	936.4422	1.9
y10* (1+)	1230.6054	1230.6055	0.1
b10 (1+)	1051.4704	1051.4691	1.2
y11* (1+)	1344.6453	1344.6484	2.3
b11 (1+)	1165.5071	1165.512	4.2
y12* (1+)	1459.6713	1459.6754	2.8
y13* (1+)	1546.7044	1546.7074	1.9
b13 (1+)	1336.5758	1336.5764	0.4
y14* (1+)	1617.7390	1617.7445	3.4
b14 (1+)	1464.6286	1464.6350	4.4
y15* (1+)	1688.7785	1688.7816	1.8
y16* (1+)	1802.8227	1802.8246	1.1
y17* (1+)	1965.8819	1965.8879	3.1
y18* (1+)	2066.9320	2066.9356	1.7
y19* (1+)	2167.9759	2167.9833	3.4
b21* (2+)	2377.0933	2377.0997	2.7

Supplementary Table 10. ¹H-NMR (800 MHz) and ¹³C-NMR (200 MHz) data for mercurialin in DMSO-d₆

Residue	C	$\delta(^{13}\text{C})$ [ppm]	H	$\delta(^1\text{H})$ (int, m, J) [ppm]	COSY	HMBC	NOESY
PyroGlu ¹	α	55.1	α	3.92 (1H, br d, 7.5)	PyroGlu ¹ : NH, H β 1, H β 2	PyroGlu ¹ : C β , C δ	PyroGlu ¹ : NH, H β 1, H β 2, H γ , Leu ² : NH
	β	24.9	β 1	1.49 (1H, m)	PyroGlu ¹ : Ha, H β 2, H γ	PyroGlu ¹ : Ca, C γ , C δ , C=O	PyroGlu ¹ : Ha, H β 2, H γ , NH
	γ	29.2	β 2	2.04 (1H, m)	PyroGlu ¹ : Ha, H β 1, H γ	PyroGlu ¹ : Ca, C γ , C δ	PyroGlu ¹ : Ha, H β 1, H γ , NH
	δ	177.2	γ	2.04 (2H, m)	PyroGlu ¹ : H β 1, H β 2	PyroGlu ¹ : C β , C δ	PyroGlu ¹ : H β 1, H β 2
	C=O	171.5	NH	7.70 (1H, s)	PyroGlu ¹ : Ha	PyroGlu ¹ : Ca, C β , C γ , C δ	PyroGlu ¹ : Ha, Leu ² : NH, H β 2
Leu ²	α	55.4	α	4.57 (1H, t, 10.5)	Leu ² : NH, H β	PyroGlu ¹ : C=O, Leu ² : C β , C=O	Leu ² : NH, H β , H γ , H δ 2, Phe ³ : NH, Trp ⁵ : H5
	β	52.4	β	2.85 (1H, dd, 2.9, 11.6)	Leu ² : Ha, H γ	Leu ² : Ca, C γ , C δ 1, C δ 2, C=O, Trp ⁵ : C5, C6, C7	Leu ² : NH, Ha, H γ , H δ 1, H δ 2, Trp ⁵ : H7
	γ	26.7	γ	2.04 (1H, m)	Leu ² : H β , H δ 1, H δ 2	n/a	Leu ² : NH, H β , H δ 1, H δ 2
	δ 1	21.8	δ 1	0.66 (3H, d, 6.7)	Leu ² : H γ	Leu ² : C β , C γ , C δ 2	Leu ² : NH, H β , H γ , H δ 2, Trp ⁵ : H7
	δ 2	17.6	δ 2	0.85 (3H, d, 6.7)	Leu ² : H γ	Leu ² : C β , C γ , C δ 1	Leu ² : Ha, H γ , H δ 1, Trp ⁵ : H4, H5
	C=O	169	NH	8.11 (1H, d, 9.1)	Leu ² : Ha	PyroGlu ¹ : C=O, Leu ² : Ca	PyroGlu ¹ : Ha, Leu ² : Ha, H β , H γ
Phe ³	α	54.3	α	3.86 (1H, t, 9.5)	Phe ³ : NH, H β 1, H β 2	Leu ² : C=O, Phe ³ : C β , C=O	Phe ³ : NH, H β 1, H β 2, Trp ⁵ : H2, Arg ⁶ : NH
	β	37.2	β 1	1.66 (1H, t, 13.7)	Phe ³ : Ha, H β 2	Phe ³ : Ca, C1, C2, C6	Phe ³ : NH, Ha, H β 2, H2/6, Phe ⁴ : H2/6, Trp ⁵ : H4
	C1	137.7	β 2	1.18 (1H, d, 13.7)	Phe ³ : H β 1	Phe ³ : C2, C6	Phe ³ : Ha, H β 1, H2/6
	C2	128.3	H2	6.74 (1H, m)	Phe ³ : H3	Phe ³ : C1, C3	Phe ³ : Ha, H β 1/2, H3
	C3	127.8	H3	7.04 (1H, br s)	Phe ³ : H2, H4	Phe ³ : C2, C4	PyroGlu ¹ : H β 1/2, Phe ³ : H2
	C4	125.8	H4	7.04 (1H, br s)	Phe ³ : H3, H5	Phe ³ : C3, C5	Phe ³ : H3/5
	C5	127.8	H5	7.04 (1H, br s)	Phe ³ : H4, H6	Phe ³ : C4, C6	PyroGlu ¹ : H β 1/2, Phe ³ : H2
	C6	128.3	H6	6.74 (1H, m)	Phe ³ : H5	Phe ³ : C1, C5	Phe ³ : Ha, H β 1/2, H5
	C=O	169.2	NH	7.95 (1H, br d, 8.11)	Phe ³ : Ha	Leu ² : C=O	Leu ² : Ha, Phe ³ : Ha, H β 1, Trp ⁵ : H5
Phe ⁴	α	56.4	α	4.44 (1H, dd, 8.8, 7.0)	Phe ⁴ : NH, H β	Phe ³ : C=O, Phe ⁴ : C β , C=O	Phe ⁴ : NH, H β , H2/6, Trp ⁵ : NH, H4
	β	83.2	β	5.35 (1H, d, 7.0)	Phe ⁴ : Ha	Phe ⁴ : C=O, C1, C2, C6, Tyr ⁷ : C4	Phe ⁴ : Ha, H2/6, Tyr ⁷ : H2, H3, H5, H6
	C1	138					
	C2	128.2	H2	7.31 (1H, d, 7.4)	Phe ⁴ : H3	Phe ⁴ : C β , C1, C3, C4	Phe ³ : NH, H β 1/2, Phe ⁴ : Ha, H β , Tyr ⁷ : H3/5
	C3	127.8	H3	7.26 (1H, t, 7.4)	Phe ⁴ : H2, H4	Phe ⁴ : C1, C2, C4	Phe ³ : H β 1/2, Phe ⁴ : H β , H2/6, H4
	C4	127.9	H4	7.16 (1H, t, 7.4)	Phe ⁴ : H3, H5	Phe ⁴ : C2, C3, C5, C6	Phe ⁴ : H3/5
	C5	127.8	H5	7.26 (1H, t, 7.4)	Phe ⁴ : H4, H6	Phe ⁴ : C1, C2, C4	Phe ³ : H β 1/2, Phe ⁴ : H β , H2/6, H4
	C6	128.2	H6	7.31 (1H, d, 7.4)	Phe ⁴ : H5	Phe ⁴ : C β , C1, C4, C5	Phe ³ : NH, H β 1/2, Phe ⁴ : Ha, H β , Tyr ⁷ : H3/5
	C=O	170.2	NH	7.50 (1H, br d, 8.8)	Phe ⁴ : Ha	n/a	Phe ³ : NH, H β 1, Phe ⁴ : Ha
Trp ⁵	α	53.2	α	5.39 (1H, br s)	Trp ⁵ : NH, H β 1, H β 2	n/a	Trp ⁵ : NH, H β 1, H β 2, H4, Arg ⁶ : NH
	β	29.8	β 1	3.03 (1H, t, 14.2)	Trp ⁵ : Ha, H β 2	Trp ⁵ : Ca, C3, C=O	Trp ⁵ : NH, Ha, H β 2, H2, Arg ⁶ : H β
	C2	124.1	β 2	3.19 (1H, br d, 14.2)	Trp ⁵ : Ha, H β 1	n/a	Trp ⁵ : Ha, H β 1, H2, H4, Arg ⁶ : H β
	C3	108.5	(NH)1	10.58 (1H, s)	Trp ⁵ : H2	Trp ⁵ : C2, C3, C8	Trp ⁵ : H2, H7
	C4	118.1	H2	6.98 (1H, m)	Trp ⁵ : (NH)1	Trp ⁵ : C3, C8, C9	Trp ⁵ : (NH)1, H β 1, Arg ⁶ : H β
	C5	118.3	H4	7.46 (1H, d, 8.3)	Trp ⁵ : H5	Trp ⁵ : C3, C5, C6, C8, C9	Leu ² : H δ 2, Trp ⁵ : NH, Ha, H β 1/2, H5
	C6	130.2	H5	6.85 (1H, d, 8.3)	Trp ⁵ : H4	Trp ⁵ : C9	Leu ² : Ha, H β , H δ 2, Phe ³ : NH, Phe ⁴ : H2/6, Trp ⁵ : Ha, H4
	C7	114.1	H7	6.77 (1H, s)	n/a	Trp ⁵ : C5	Leu ² : NH, Ha, H β , Trp ⁵ : (NH)1
	C8	136.4					
	C9	124.8	NH	8.43 (1H, d, 7.0)	Trp ⁵ : Ha	n/a	Phe ⁴ : Ha, Trp ⁵ : Ha, H β 1, Arg ⁶ : NH
	C=O	171.2					
Arg ⁶	α	52.2	α	4.34 (1H, dd, 6.9, 13.6)	Arg ⁶ : NH, H β	Trp ⁵ : C=O	Arg ⁶ : NH, Ha, H β 1, H β 2, H γ , H δ , Tyr ⁷ : NH
	β	30.1	β	1.49 (2H, m)	Arg ⁶ : Ha, H γ 1/2	Arg ⁶ : Ca	Arg ⁶ : NH, Ha, H γ 1/2, H δ
	γ	24.9	γ 1	1.35 (1H, m)	Arg ⁶ : H β , H γ 2, H δ	n/a	Arg ⁶ : NH, Ha, H β , H γ 2, H δ
	δ	40.4	γ 2	1.43 (1H, m)	Arg ⁶ : H β , H γ 1, H δ	Arg ⁶ : C β	Arg ⁶ : NH, Ha, H β , H γ 1, H δ
	ϵ	157.8	δ	3.10 (2H, m)	Arg ⁶ : ϵ (NH), H γ 1/2	n/a	Arg ⁶ : Ha, H β , H γ 1/2, H ϵ
	C=O	169.1	ϵ (NH)	7.54 (1H, br s)	Arg ⁶ : H δ	n/a	Arg ⁶ : H γ , H δ
Tyr ⁷			NH	6.92 (1H, br s)	Arg ⁶ : Ha	n/a	Trp ⁵ : NH, Ha, H β 1, H β 2, Arg ⁶ : Ha, H β , H γ , Tyr ⁷ : Ha
	α	53.2	α	4.41 (1H, dd, 12.3, 11.5)	Tyr ⁷ : H β 1/2, NH	Tyr ⁷ : C=O	Arg ⁶ : NH, Tyr ⁷ : NH, H β 1/2, H2, H6
	β	37.7	β 1	3.13 (1H, d, 12.3)	Tyr ⁷ : Ha, H β 2	Tyr ⁷ : Ca, C1, C2, C6	Tyr ⁷ : NH, Ha, H β 2, H2, H6
	C1	129.9	β 2	2.46 (1H, t, 12.3)	Tyr ⁷ : Ha, H β 1	Tyr ⁷ : Ca, C1, C2, C6	Tyr ⁷ : NH, Ha, H β 1, H2, H6
	C2	130.6	H2	6.98 (1H, br s)	Tyr ⁷ : H3	n/a	Phe ⁴ : H β , Tyr ⁷ : NH, Ha, H β 1/2, H3/5
	C3	130.6	H3	6.65 (1H, d, 7.6)	Tyr ⁷ : H2	Tyr ⁷ : C1, C2, C4	Phe ⁴ : H β , Tyr ⁷ : Ha, H β 1/2, H2, H6
	C4	156.1	H5	6.65 (1H, d, 7.6)	Tyr ⁷ : H6	Tyr ⁷ : C1, C4, C6	Phe ⁴ : H β , Tyr ⁷ : Ha, H β 1/2, H2, H6
	C5	130.6	H6	6.98 (1H, br s)	Tyr ⁷ : H5	n/a	Phe ⁴ : H β , Tyr ⁷ : NH, Ha, H β 1/2, H3/5
	C6	130.6	NH	8.34 (1H, br s)	Tyr ⁷ : Ha	n/a	Arg ⁶ : Ha, Tyr ⁷ : Ha, H β , H2, H6
	C=O	172.5					

n/a – not detected, br s – broad singlet, d – doublet, dd – double doublet, t – triplet, m - multiplet

Supplementary Table 11. MS/MS analysis of linear LysC-proteolytic core peptide of ManBURP-1xQLFFWRY. Proteomic data: MSV000097169

Fragment (z)	m(obs) [Da]	m(calc) [Da]	Δm [ppm]
b3 (1+)	339.2023	339.2027	1.2
b4 (1+)	453.2458	453.2457	0.2
y4 (1+)	462.2379	462.2381	0.4
y5 (1+)	575.3219	575.3222	0.5
b5 (1+)	616.3091	616.3090	0.1
b6 (1+)	730.3513	730.3519	0.8
y7 (1+)	844.5042	844.5074	3.8
b7 (1+)	877.4199	877.4203	0.5
y8 (1+)	957.5915	957.5914	0.1
b8 (1+)	1005.4786	1005.4789	0.3
y9 (1+)	1072.6192	1072.6184	0.7
b9 (2+)	1102.5309	1102.5317	0.7
b10 (1+)	1231.5740	1231.5743	0.2
y11 (1+)	1243.6860	1243.6828	2.6
y12 (2+)	1399.7827	1399.7839	0.8
y13 (2+)	1513.8281	1513.8268	0.9
b13 (2+)	1570.7383	1570.7285	6.3
b14 (2+)	1683.8045	1683.8126	4.8
y16 (3+)	1961.0489	1961.0498	0.4
b17 (2+)	2164.0193	2164.0287	4.3
y18 (3+)	2294.1956	2294.1975	0.8
y19 (3+)	2441.2646	2441.2660	0.6
y20 (3+)	2554.3457	2554.3500	1.7
y21 (3+)	2682.4079	2682.4086	0.2
y22 (3+)	2796.4493	2796.4515	0.8
y23 (3+)	2893.5032	2893.5043	0.4
y24 (3+)	3022.5446	3022.5469	0.7
y25 (3+)	3119.5973	3119.5996	0.7
b25 (3+)	3167.4905	3167.4871	1.1
y26 (3+)	3247.6580	3247.6582	0.1
y27 (3+)	3394.7237	3394.7266	0.9
y28 (3+)	3508.7684	3508.7696	0.3
b28 (4+)	3549.7446	3549.7564	3.3
y29 (3+)	3671.8289	3671.8329	1.1
b31 (4+)	3846.9230	3846.9252	0.6
b32 (4+)	3977.9670	3977.9657	0.3
y32 (5+)	4010.9844	4010.9872	0.7
b33 (5+)	4106.0564	4106.0607	1.1
y33 (5+)	4124.0704	4124.0712	0.2

Supplementary Table 12. MS/MS analysis of putative bicyclic LysC-proteolytic core peptide of ManBURP-1xQLFFWRY. Proteomic data: MSV000097169

Fragment (z)	m(obs) [Da]	m(calc) [Da]	Δm [ppm]
b3 (1+)	339.2027	339.2027	0.0
b4 (1+)	453.2456	453.2457	0.2
y4 (1+)	462.2379	462.2381	0.4
y5 (1+)	575.3225	575.3222	0.5
b5 (1+)	616.3088	616.3090	0.3
b6 (1+)	730.3512	730.3519	1.0
y6 (1+)	731.4170	731.4233	8.6
b7 (1+)	877.4197	877.4203	0.7
y8 (1+)	957.5915	957.5914	0.1
b8 (1+)	1005.4785	1005.4789	0.4
b9 (2+)	1102.5309	1102.5317	0.7
b10 (1+)	1231.5736	1231.5743	0.6
y11 (1+)	1243.6853	1243.6828	2.0
b11 (1+)	1328.6351	1328.6270	6.1
y12 (2+)	1399.7859	1399.7839	1.4
b13 (2+)	1570.7179	1570.7285	6.7
y20* (3+)	2550.3152	2550.3187	1.4
y21* (3+)	2678.3783	2678.3773	0.4
y22* (3+)	2792.4203	2792.4202	0.0
y23* (3+)	2889.4718	2889.4730	0.4
y24* (3+)	3018.5141	3018.5156	0.5
y25* (3+)	3115.5680	3115.5683	0.0
b25* (3+)	3163.4567	3163.4558	0.3
y26* (3+)	3243.6257	3243.6269	0.4
y27* (3+)	3390.6962	3390.6953	0.3
y28* (3+)	3504.7283	3504.7382	2.8
b31* (4+)	3842.8850	3842.8939	2.3
b32* (4+)	3973.9362	3973.9344	0.4
y32* (5+)	4006.9534	4006.9558	0.6
b33* (5+)	4102.0264	4102.0293	0.7

Supplementary Table 13. ¹H-NMR (800 MHz) and ¹³C-NMR (200 MHz) data for moroidin-QLLVWRNH in DMSO-d₆.

Residue		C δ(13C) [ppm]	H	δ(1H) (int, m, J) [ppm]	COSY correlations	HMBC correlations	NOESY correlations
PyroGlu ¹	<i>α</i>	55.1	<i>α</i>	4.10 (1H, dd, 3.0, 8.6)	PyroGlu ¹ : NH, Hβ1, Hβ2	PyroGlu ¹ : Cβ, Cδ	PyroGlu ¹ : NH, Hβ1, Hβ2, Hy, Leu ² : NH
	<i>β</i>	25.4	<i>β1</i>	2.25 (1H, m)	PyroGlu ¹ : Ha, Hβ2, Hy	PyroGlu ¹ : Ca, Cy, C=O	PyroGlu ¹ : Ha, Hβ2, Hy
	<i>γ</i>	28.9	<i>β2</i>	1.70 (1H, m)	PyroGlu ¹ : Ha, Hβ1, Hy	n/a	PyroGlu ¹ : Ha, Hβ1, Hy
	<i>δ</i>	177.3	<i>γ</i>	2.09 (2H, t, 8.7)	PyroGlu ¹ : Hβ1, Hβ2	PyroGlu ¹ : Ca, Cβ, Cδ	PyroGlu ¹ : Hβ1, Hβ2
	C=O	172.2	NH	7.88 (1H, s)	PyroGlu ¹ : Ha	PyroGlu ¹ : Ca, Cβ, Cy, Cδ	PyroGlu ¹ : Ha
Leu ²	<i>α</i>	54.7	<i>α</i>	4.85 (1H, t, 10.3)	Leu ² : NH, Hβ	PyroGlu ¹ : C=O, Leu ² : Cβ, C=O	Leu ² : NH, Hy, Hδ2, Leu ³ : NH, Val ⁴ : NH, Trp ⁵ : H5
	<i>β</i>	51.5	<i>β</i>	3.06 (1H, dd, 3.3, 11.1)	Leu ² : Ha, Hy	Leu ² : Ca, Cy, Cδ1, Cδ2, Trp ⁵ : C5, C6, C7	Leu ² : NH, Ha, Hy, Hδ1, Trp ⁵ : H7
	<i>γ</i>	26.6	<i>γ</i>	2.14 (1H, m)	Leu ² : Hβ, Hδ1, Hδ2	Leu ² : Cδ1	Leu ² : NH, Hβ, Hδ1, Hδ2
	<i>δ1</i>	21.7	<i>δ1</i>	0.75 (3H, m)	Leu ² : Hy	Leu ² : Cβ, Cy, Cδ2	Leu ² : Hβ, Hy, Hδ2, Trp ⁵ : H7
	<i>δ2</i>	17.1	<i>δ2</i>	0.83 (3H, d, 6.4)	Leu ² : Hy	Leu ² : Cβ, Cy, Cδ1	Leu ² : Ha, Hy, Hδ1, Trp ⁵ : H5
	C=O	170.3	NH	8.55 (1H, d, 8.8)	Leu ² : Ha	PyroGlu ¹ : C=O, Leu ² : Ca	PyroGlu ¹ : Ha, Leu ² : Ha, Hβ, Hy
	<i>α</i>	52	<i>α</i>	3.96 (1H, br t, 10.7)	Leu ³ : NH, Hβ	Leu ² : C=O, Leu ³ : Cβ, C=O	Leu ³ : NH, Hβ1, Hβ2, Hδ1, Val ⁴ : NH
Leu ³	<i>β</i>	41.8	<i>β</i>	1.18/1.46 (ea 1H, dd, 5.1, 9.7)	Leu ³ : Ha, Hy	Leu ³ : Ca, Cy, Cδ1, Cδ2	Leu ³ : Ha, Hy, Hδ1, Hδ2, Val ⁴ : NH
	<i>γ</i>	23.8	<i>γ</i>	1.42 (1H, m)	Leu ³ : Hβ, Hδ1, Hδ2	Leu ³ : Cδ1, Cδ2	Leu ³ : Hβ, Hδ1, Hδ2
	<i>δ1</i>	20.8	<i>δ1</i>	0.68 (3H, d, 6.0)	Leu ³ : Hy	Leu ³ : Cβ, Cy, Cδ2	Leu ³ : Ha, Hβ, Hy, Hδ2
	<i>δ2</i>	23	<i>δ2</i>	0.78 (3H, d, 6.4)	Leu ³ : Hy	Leu ³ : Cβ, Cy, Cδ1	Leu ³ : Hβ, Hy, Hδ1
	C=O	171.5	NH	8.38 (1H, d, 9.2)	Leu ³ : Ha	Leu ³ : C=O	Leu ³ : Ha, Leu ³ : Ha, Hβ, Val ⁴ : NH
	<i>α</i>	57.5	<i>α</i>	3.67 (1H, t, 7.7)	Val ⁴ : NH, Hβ	Val ⁴ : Cβ, Cy1, Cy2, C=O	Val ⁴ : NH, Hβ, Hy1, Hy2, Trp ⁵ : NH
	<i>β</i>	30.8	<i>β</i>	1.84 (1H, m)	Val ⁴ : Ha, Hy	Val ⁴ : Ca, Cy1, Cy2	Val ⁴ : NH, Ha, Hy1, Hy2
Val ⁴	<i>γ1</i>	18.2	<i>γ1</i>	0.78 (3H, d, 5.8)	Val ⁴ : Hβ	Val ⁴ : Ca, Cβ, Cy2	Val ⁴ : NH, Ha, Hβ
	<i>γ2</i>	18.3	<i>γ2</i>	0.79 (3H, d, 5.8)	Val ⁴ : Hβ	Val ⁴ : Ca, Cβ, Cy1	Val ⁴ : NH, Ha, Hβ
	C=O	169.2	NH	6.92 ^a (1H)	Val ⁴ : Ha	Leu ⁴ : C=O	Leu ² : Ha, Leu ³ : NH, Hβ, Val ⁴ : Ha, Hy
	<i>α</i>	48.8	<i>α</i>	5.51 (1H, dd, 8.7, 14.8)	Trp ⁵ : NH, Hβ1, Hβ2	n/a	Trp ⁵ : NH, Hβ1, Hβ2, H4, Arg ⁶ : NH
	<i>β</i>	27.7	<i>β1</i>	2.80 (1H, dd, 11.3, 15.3)	Trp ⁵ : Ha, Hβ2	Trp ⁵ : C2	Trp ⁵ : NH, Ha, Hβ2, Arg ⁶ : NH, His ⁸ : H2
Trp ⁵	C2	124.6	<i>β2</i>	3.64 (1H, dd, 4.2, 15.3)	Trp ⁵ : Ha, Hβ1	n/a	Trp ⁵ : Ha, Hβ1, C4 Arg ⁶ : NH, Asn ⁷ : NH, His ⁸ : H2
	C3	103.4	(NH)1	11.80 (1H, s)	n/a	Trp ⁵ : C2, C3	Trp ⁵ : H7, His ⁸ : H4
	C4	119.4	H4	7.55 (1H, d, 8.2)	Trp ⁵ : H5	Trp ⁵ : C2, C3, C6, C8	Leu ³ : Hδ, Trp ⁵ : Ha, Hβ, H5, Arg ⁶ : NH
	C5	119.8	H5	7.01 (1H, d, 8.2)	Trp ⁵ : H4	Trp ⁵ : C2, C7	Leu ³ : Ha, Hδ2, Leu ³ : NH, Trp ⁵ : H4
	C6	131.7	H7	6.93 (1H, s)	n/a	Leu ² : Cβ, Trp ⁵ : C2, C4, C5	Leu ² : Hβ, Hy, Hδ1, Trp ⁵ : H5, (NH)1
	C7	114.2	NH	8.15 (1H, s)	Trp ⁵ : Ha	n/a	Val ⁴ : Ha, Hy, Trp ⁵ : Ha, Hβ
	C8	132.7					
	C9	127					
	C=O	171.6					
Arg ⁶	<i>α</i>	52.8	<i>α</i>	4.35 (1H, dd, 7.5, 13.1)	Arg ⁶ : NH, Hβ1, Hβ2	n/a	Arg ⁶ : NH, Hβ1, Hβ2, Hy, Hδ, Asn ⁷ : NH
	<i>β</i>	28.5	<i>β1</i>	1.65 (1H, m)	Arg ⁶ : Ha, Hβ2, Hy	n/a	Arg ⁶ : NH, Ha, Hβ2, Hy, Hδ
	<i>γ</i>	24.8	<i>β2</i>	1.82 (1H, m)	Arg ⁶ : Ha, Hβ1, Hy	n/a	Arg ⁶ : NH, Ha, Hβ1, Hy, Hδ
	<i>δ</i>	40.3	<i>γ</i>	1.54 ^a (1H)	Arg ⁶ : Hβ1, Hβ2, Hδ	n/a	Arg ⁶ : NH, Ha, Hβ1, Hβ2, Hδ, Hε
	<i>ε</i>	156.8	<i>δ</i>	3.12 (2H, dd, 6.1, 12.2)	Arg ⁶ : ε(NH), Hy	n/a	Arg ⁶ : Ha, Hβ1, Hβ2, Hy, Hε
	C=O	171.3 _b	α(NH)	7.69 (1H, m)	Arg ⁶ : Hδ	n/a	Arg ⁶ : Hy, Hδ
Asn ⁷			NH	8.67 (1H, d, 7.5)	Arg ⁶ : Ha	Trp ⁵ : C=O	Trp ⁵ : Ha, Hβ1, Hβ2, Arg ⁶ : Ha, Hβ1, Hβ2, Hy, Asn ⁷ : NH, His ⁸ : H2
	<i>α</i>	49.5	<i>α</i>	4.54 (1H, m)	Asn ⁷ : NH, Hβ1, Hβ2	Asn ⁷ : Cy	Asn ⁷ : NH, His ⁸ : NH, H2
	<i>β</i>	36.5	<i>β1</i>	2.42 (1H, d, 14.5)	Asn ⁷ : Ha, β2	n/a	Trp ⁵ : H4, Asn ⁷ : Ha, Hβ2, NH
	<i>γ</i>	172.6	<i>β2</i>	2.86 (1H, d, 14.5)	Asn ⁷ : Ha, β1	n/a	Trp ⁵ : H4, Asn ⁷ : Ha, Hβ2, NH
	C=O	169.4 _b	NH	7.79 (1H, br s)	Gly ⁷ : Ha	n/a	Trp ⁵ : Hβ1, Arg ⁶ : Ha, NH Asn ⁷ : Ha
His ⁸	<i>α</i>	51.1	<i>α</i>	4.66 (1H, br s)	His ⁸ : NH, Hβ1, Hβ2	n/a	His ⁸ : NH, Hβ1, Hβ2, H2
	<i>β</i>	27.2	<i>β1</i>	3.12 ^a (1H)	His ⁸ : Ha, Hβ2	His ⁸ : Ca	His ⁸ : Ha, Hβ2
	C1	131.7	<i>β2</i>	3.28 (1H, d, 14.5)	His ⁸ : Ha, Hβ1	n/a	His ⁸ : Ha, Hβ1, H2, NH
	C2	120.8	H2	7.85 (1H, s)	n/a	His ⁸ : Cβ	Trp ⁵ : Hβ1, Hβ2, His ⁸ : NH, Ha, Hβ1, Hβ2
	C4	136.5	H4	9.02 (1H, br s)	n/a	n/a	Trp ⁵ : (NH)1, Arg ⁶ : Ha
	C=O	172.1 _b	NH	7.68 ^a (1H)	His ⁸ : Ha	n/a	Asn ⁷ : NH, Ha, His ⁸ : Ha, Hβ

a – overlapping sig nals, b – signals cannot be assigned by 2D NMR, br s – broad singlet, d – doublet, dd – double doublet, t – triplet, m – multiplet.

Supplementary Table 14. Comparison of burpitide transcriptome mining.

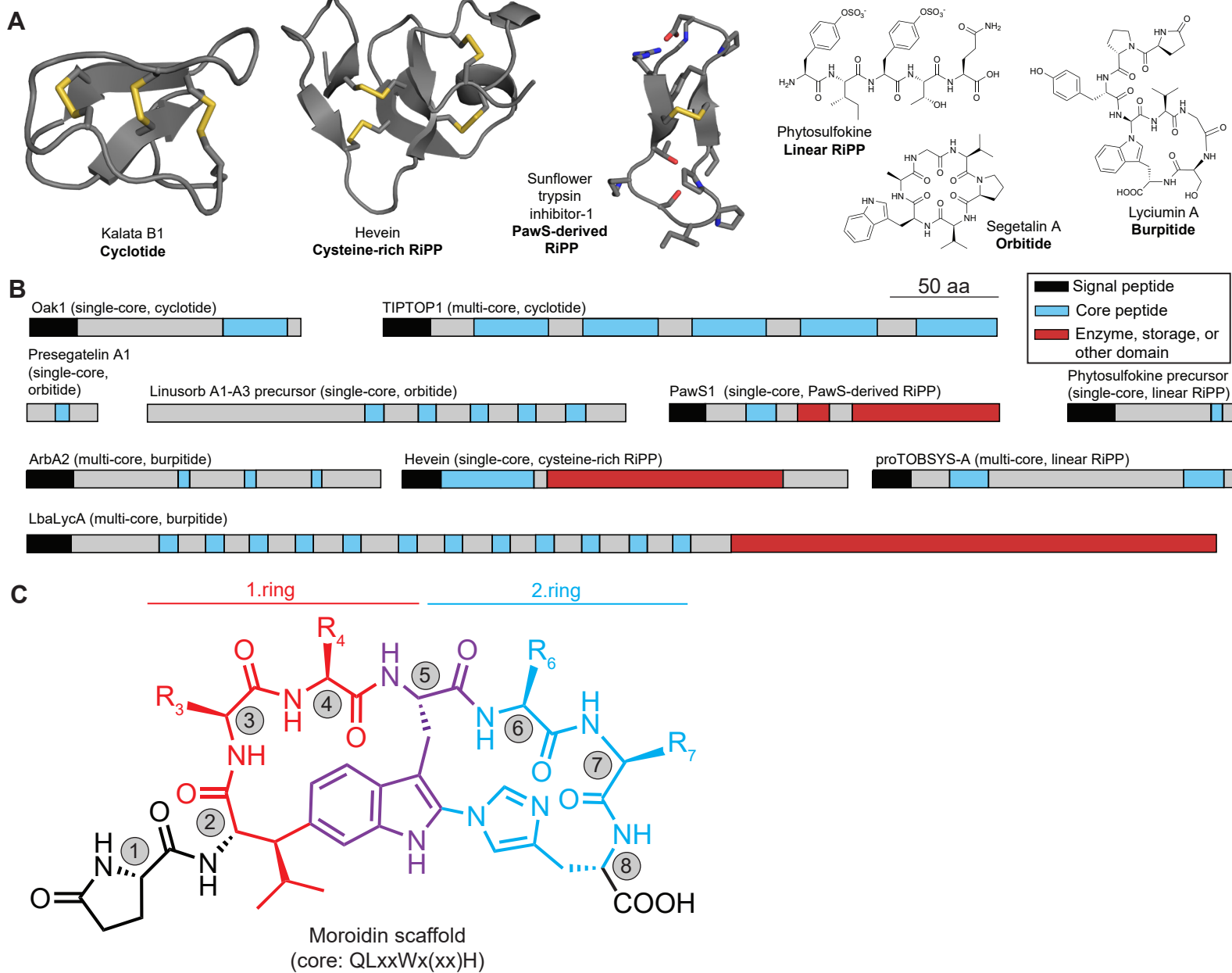
Comparison of RNA-seq studies	Datasets	Detected stephanotic acid core peptides	Characterized new burpitides (NMR, bio-synthesis)	Characterized new PTMs (NMR, bio-synthesis)
Kersten <i>et al.</i> 3	700	3	1	1
Kruger <i>et al.</i> 5	793	5	0	0
This study	8893	55	3	2

Supplementary Table 15. Chemotaxonomic distribution of stephanotic acid-type burpitides.

Species	Family	Order	Stephanotic acid QLxxW	Moroidin QLxxWx(xx)H	Glechomanin QLxxWx(xx)W	Mercurialin QLxxWx(xx)Y
<i>Stellaria aquatica</i>	Caryophyllaceae	Caryophyllales	x*	x*		
<i>Bauhinia tomentosa</i>	Fabaceae	Fabales	x	x		
<i>Stephanotis floribunda</i>	Apocynaceae	Gentianales	x			
<i>Glechoma hederacea</i>	Lamiaceae	Lamiales	x*		x*	
<i>Mercurialis annua</i>	Euphorbiaceae	Malpighiales	x*			x*
<i>Sacciolepis striata</i>	Poaceae	Poales	x*	x*		
<i>Kerria japonica</i>	Rosaceae	Rosales	x	x		
<i>Dendrocnide moroides</i>	Urticaceae	Rosales		x		
<i>Orchidantha maxillarioides</i>	Lowiaceae	Zingiberales	x*	x*	x*	
x - Characterized chemotype & experimentally verified precursor peptide (x* - this study)						
x - Characterized chemotype & no experimentally verified precursor peptide						

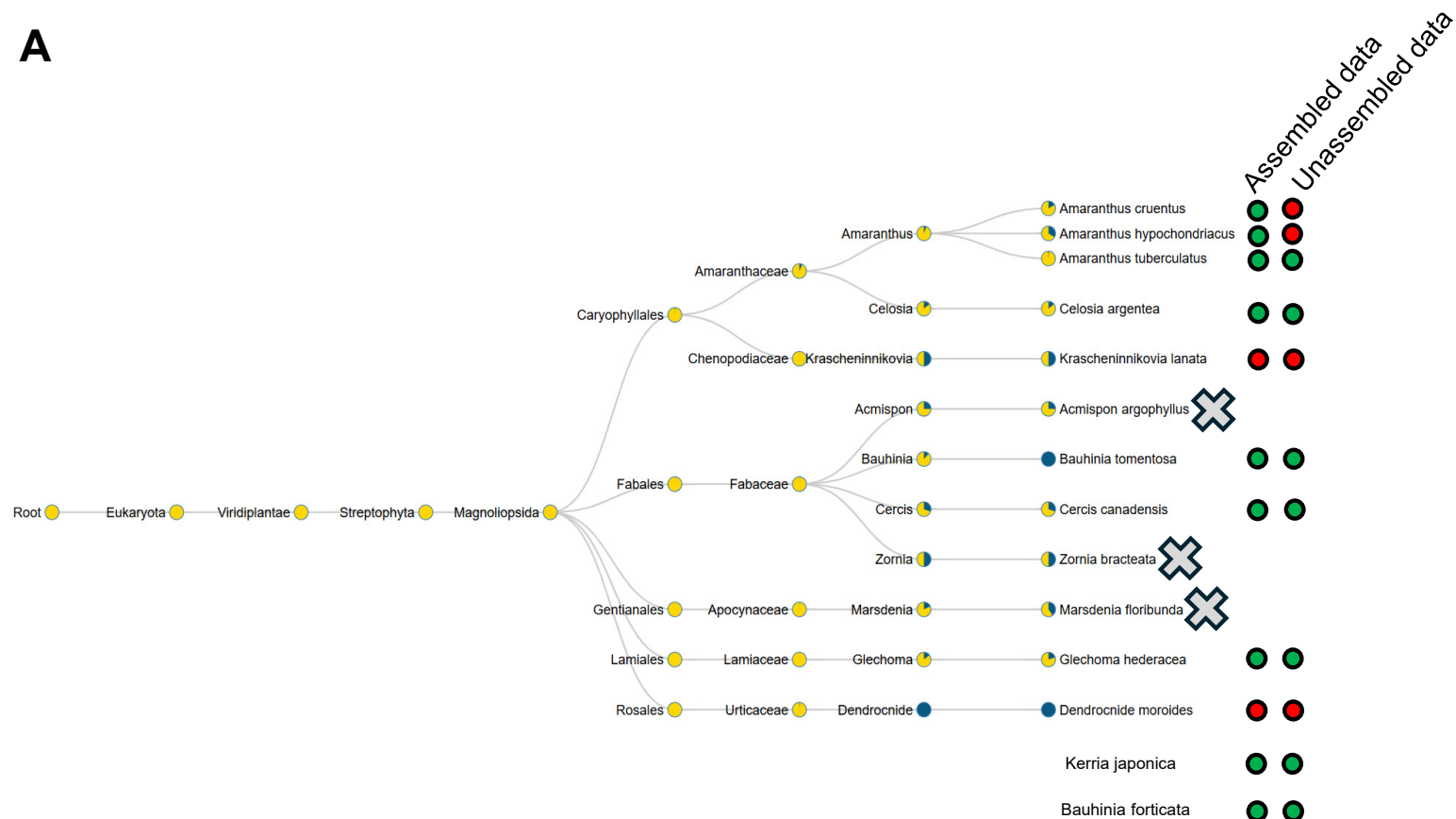
Supplementary Table 16. Cloning primers.

Name	Sequence
GheBURP-pEAQ-HT- fwd	TGCCCAAATTCGCGACCGGTATGGAATTCAAGATTCTCCCC
GheBURP-pEAQ-HT- rev	CCAGAGTTAAAGGCCTCGAGTTAATTTTGGAAACCCACAC
SaqBURP-pH8-fwd	GAAAACCTGTACTTCCAGGCCCATGGCATGAGGTTCAAAATGCTTATTAGTTATATT
SaqBURP-pH8-rev	CTCGAATTCGGATCCGCCATGGTTACATAGCATGAGGAGAAGTATTAGC
OmaBURP1-pEAQ- HT-fwd	TGCCCAAATTCGCGACCGGTATGGATCGCCTCACCTCTG
OmaBURP1-pEAQ- HT-rev	CCAGAGTTAAAGGCCTCGAGTTACACAGTGCTAGTCGGTG
OmaBURP2-pEAQ- HT-fwd	TGCCCAAATTCGCGACCGGTATGGATCGCCTAACCTCTGC
OmaBURP2-pEAQ- HT-rev	CCAGAGTTAAAGGCCTCGAGCTACAGAATGCTACGCGGTGC

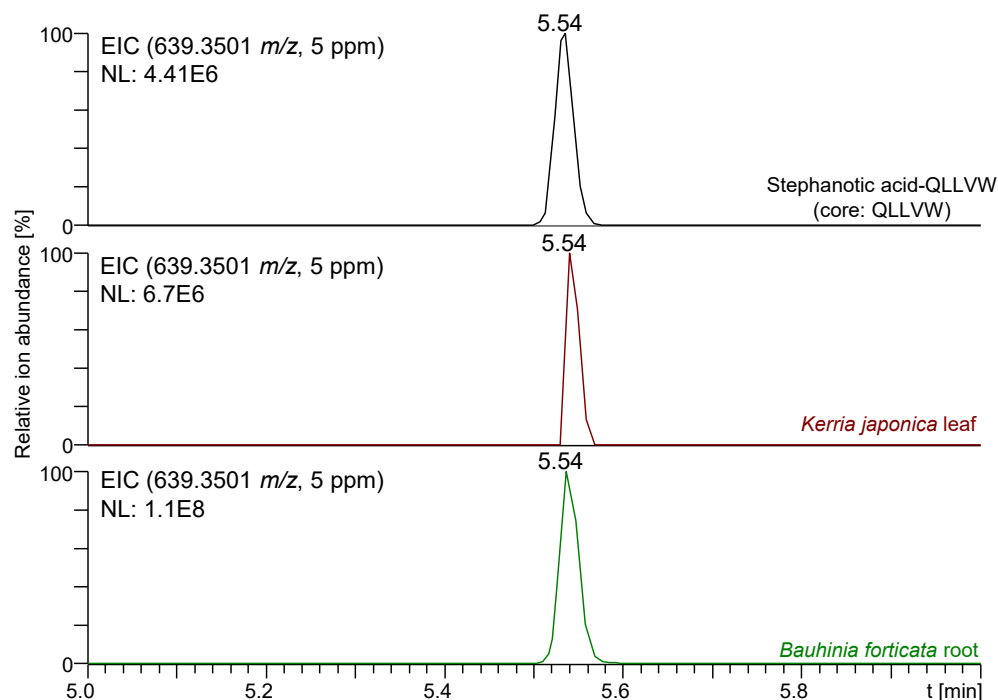


Supplementary Figure 1. Plant RiPP structures and biosynthetic precursor peptides. (A) Chemical structures of representative plant RiPPs. (B) Primary precursor peptide structures of representative plant RiPPs including single- and multi-core precursor peptides. (C) Residue numbering in moroidin scaffold. Residue numbers are in grey circles. The second ring can be contracted or extended by one residue³. x is any proteinogenic amino acid. Abbreviations: aa - amino acid.

A



B

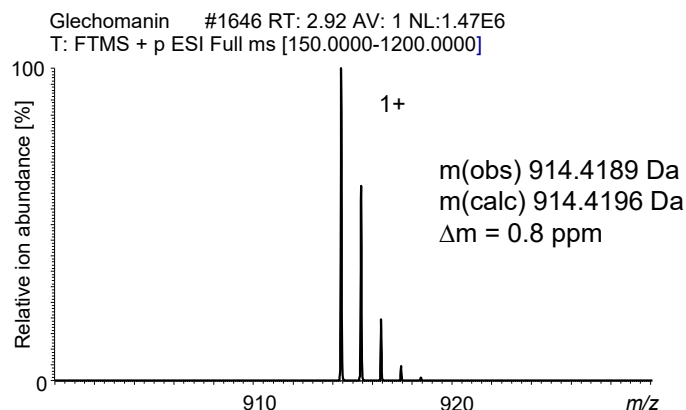


Supplementary Figure 2. Detection of stephanotic acid-QLLVW in plantMASST-LCMS2 data and manually analyzed LCMS1 data. (A) plantMASST (v2025.3.25) ⁶ phylogenetic representation of hits for searching stephanotic acid-QLLVW tandem MS spectrum (CCMSLIB00006709935). A corresponding PHI-BLAST hit (KjaBURP homolog with QLLVW motif) in available transcriptomics data (assembled or unassembled) is indicated by a green circle, no PHI-BLAST hit is indicated by a red circle. plantMASST hits without available RNA-seq data are labeled with a grey cross. (B) LCMS1-based detection of stephanotic acid-QLLVW in methanolic extracts of *Bauhinia forticata* and *Kerria japonica*, which were missed in plantMASST, likely due to the lack of stephanotic acid-QLLVW tandem MS spectra in deposited datasets.

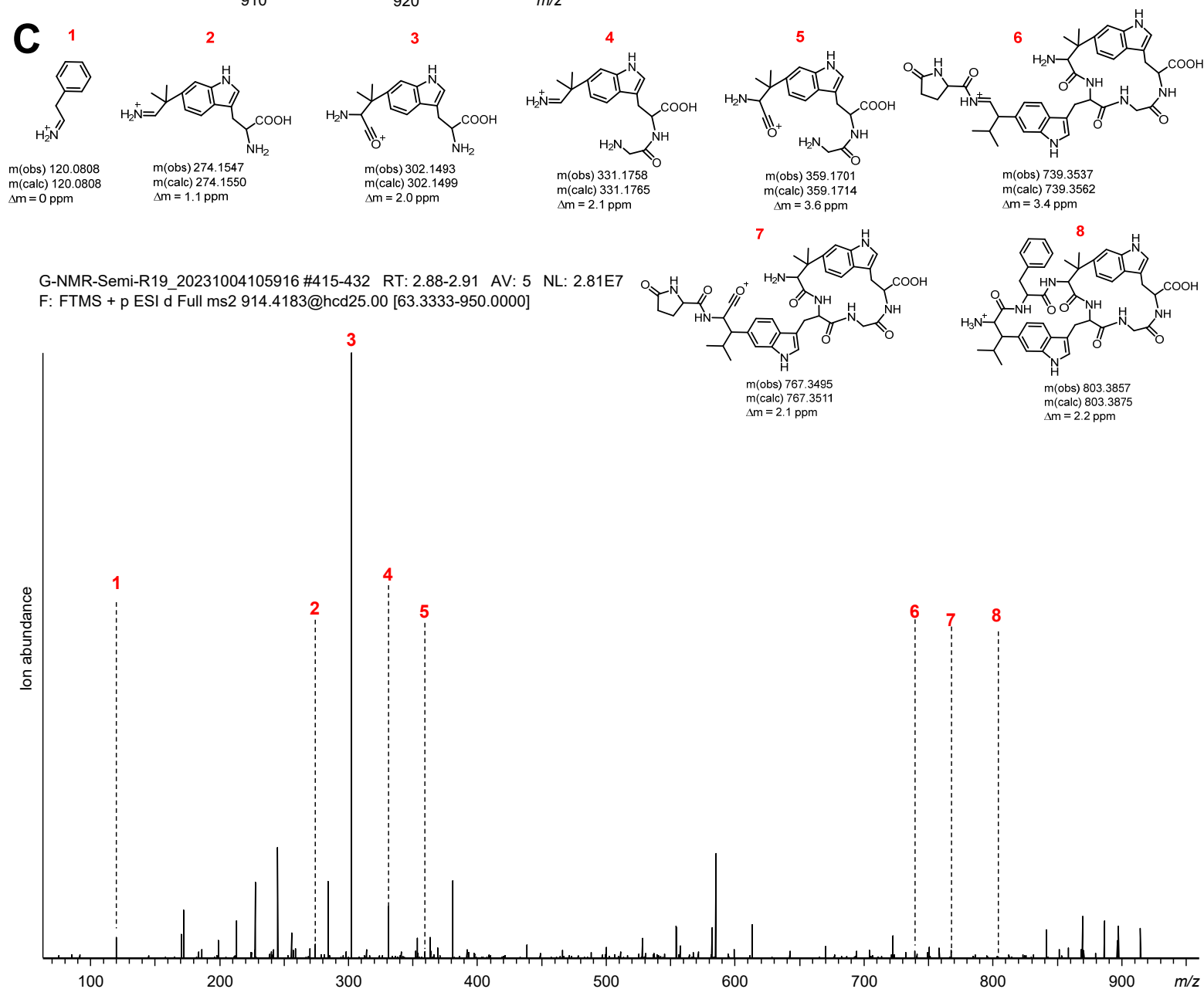
A $m(\text{calc, QLFVWGW}) = 935.4774 \text{ [M+H]}^+$
 $- (1 \times 17.02655) \text{ (pyroglutamate formation from Gln)}$
 $- (4 \times 1.00782) \text{ (bicyclization)}$

 914.4196 Da

B



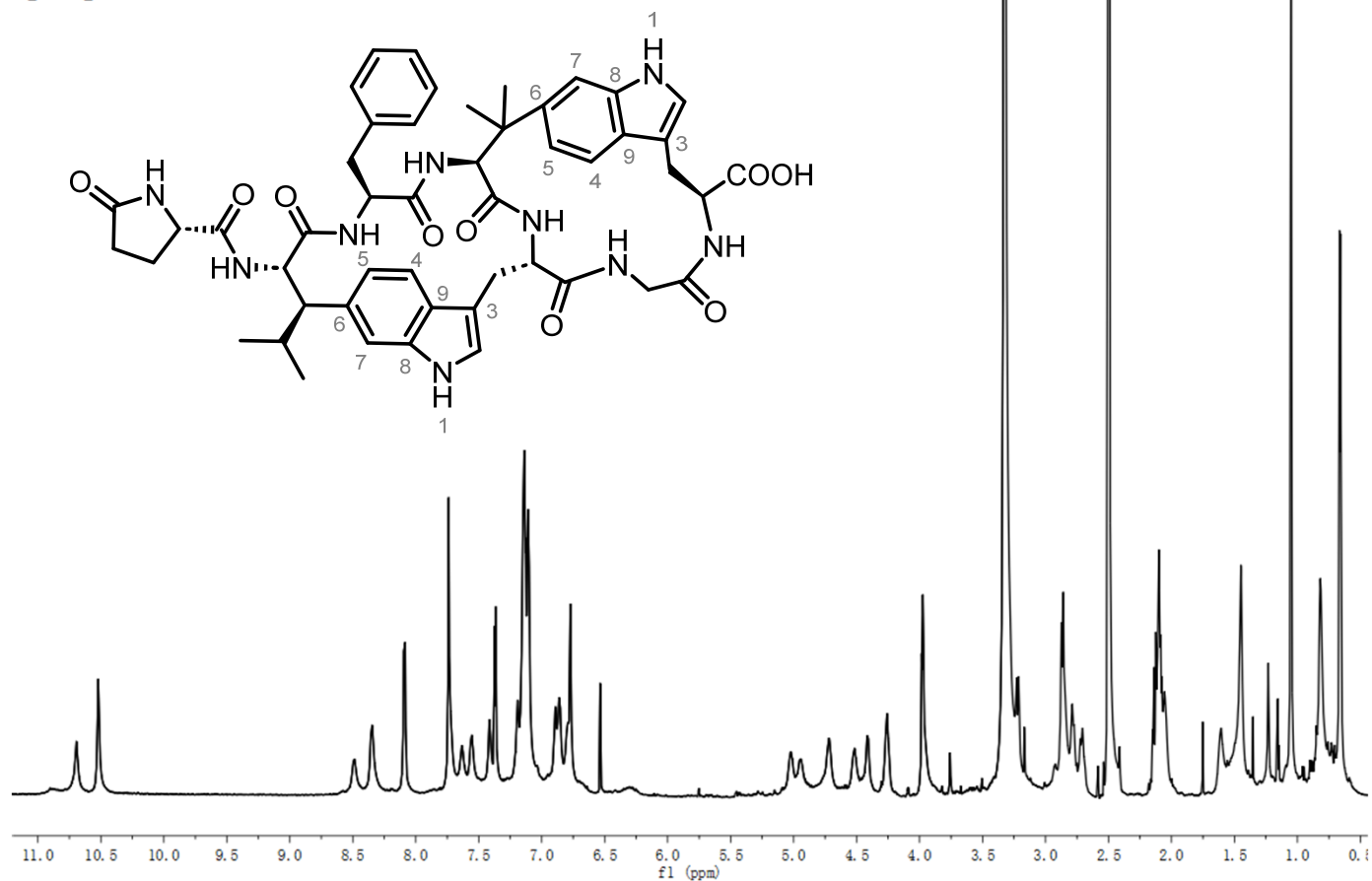
C



Supplementary Figure 3. MS and MS/MS spectrum of glechomanin. (A) Mass calculation of glechomanin from predicted core peptide sequence and posttranslational modifications. (B) MS1 spectrum of glechomanin. (C) MS/MS spectrum of glechomanin and predicted fragments.

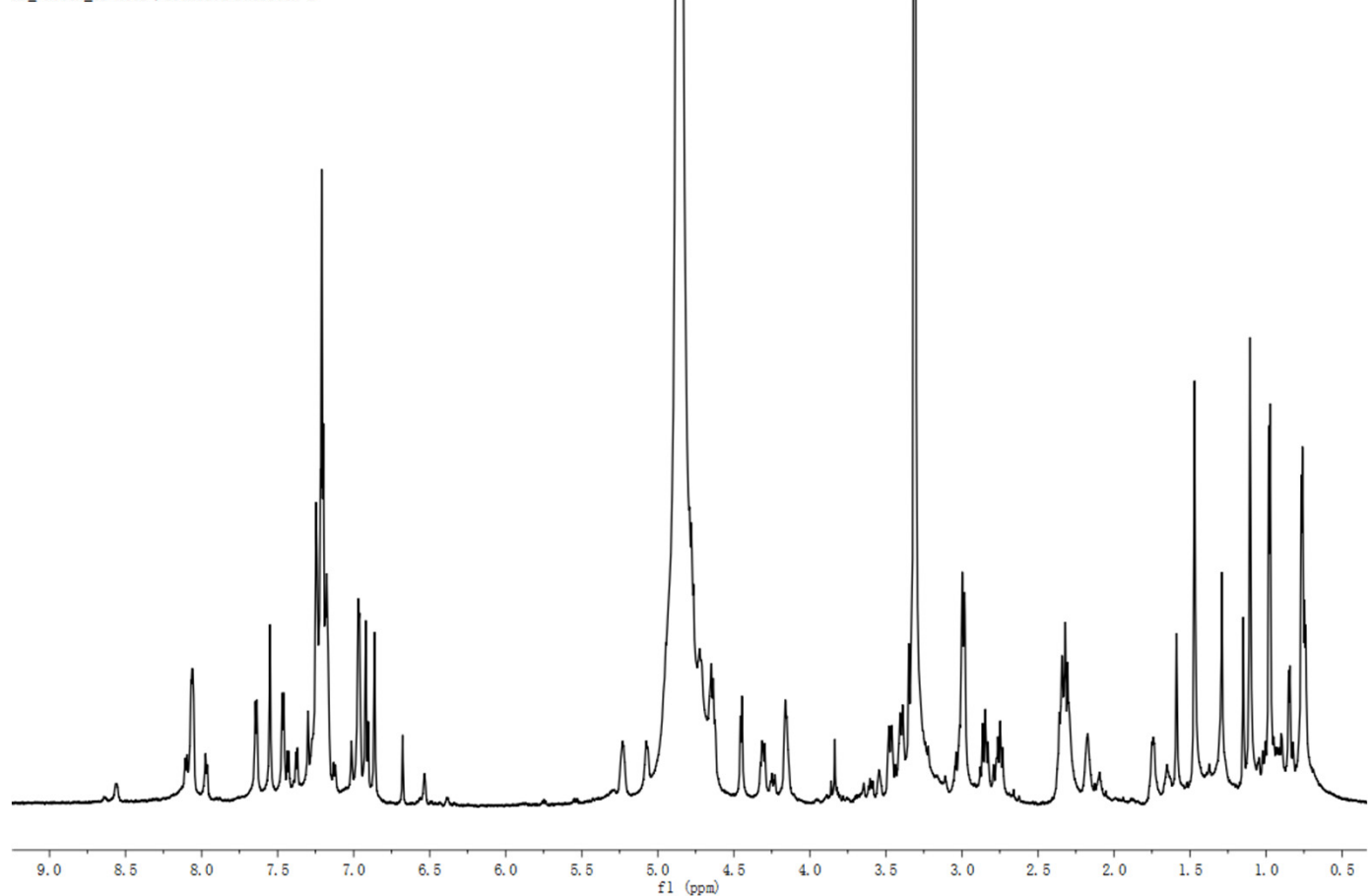
Glechomin-3

RK_PROTON_1D DMSO /icondata rkersten 2

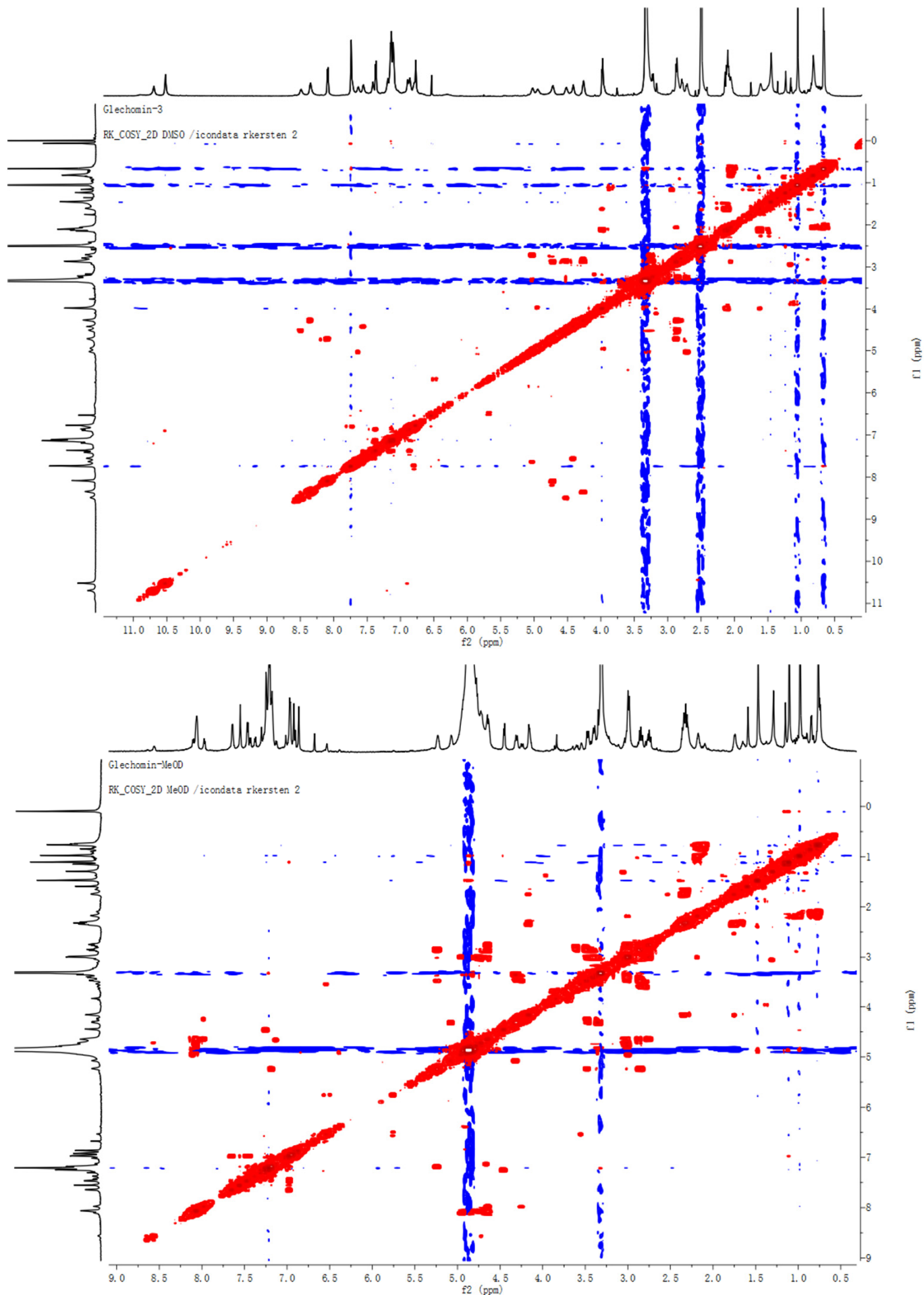


Glechomin-MeOD

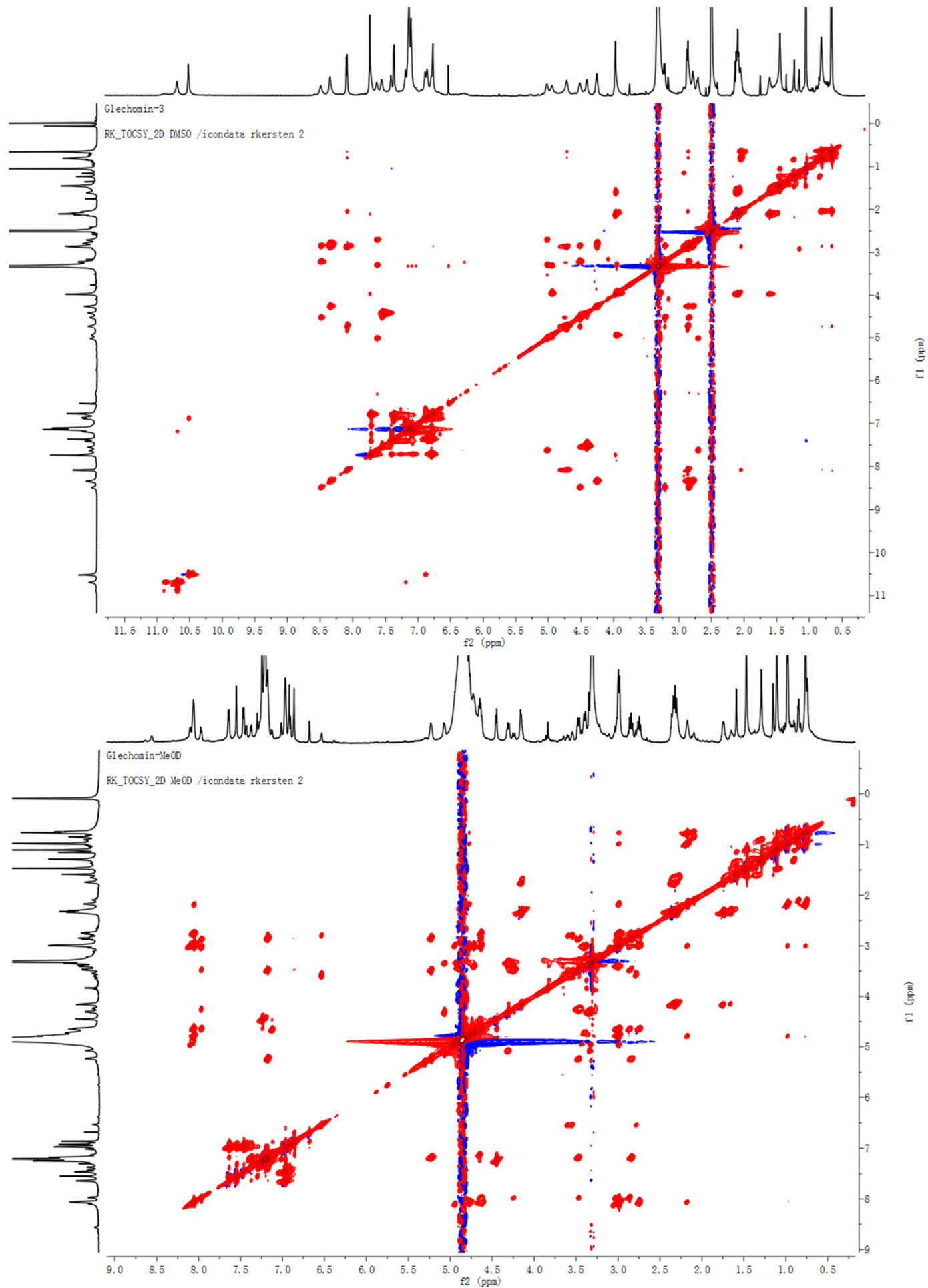
RK_PROTON_1D MeOD /icondata rkersten 2



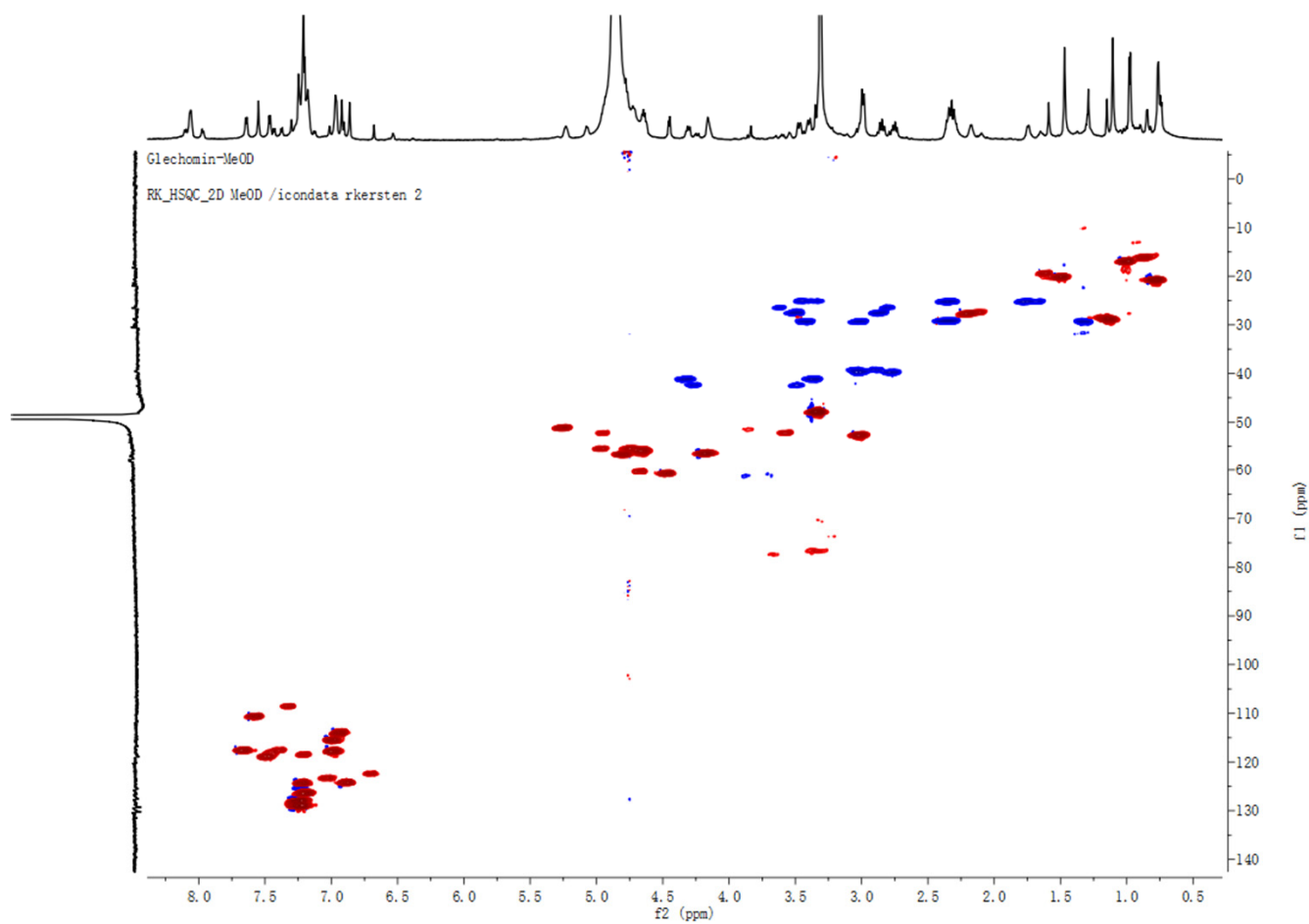
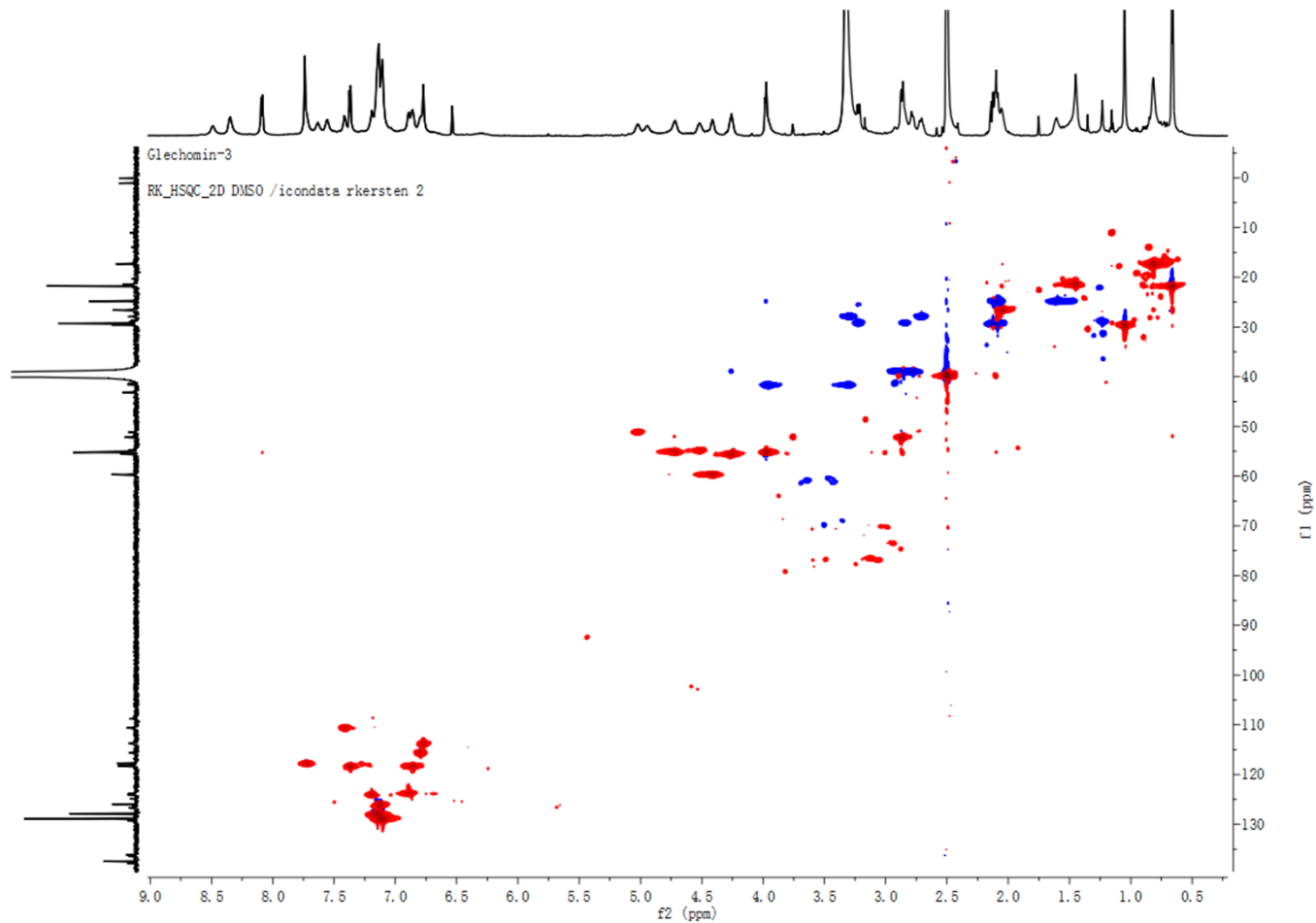
Supplementary Figure 4. ¹H-NMR spectrum of glechomanin in DMSO-d₆ (800 MHz) and in methanol-d₄ (800 MHz).



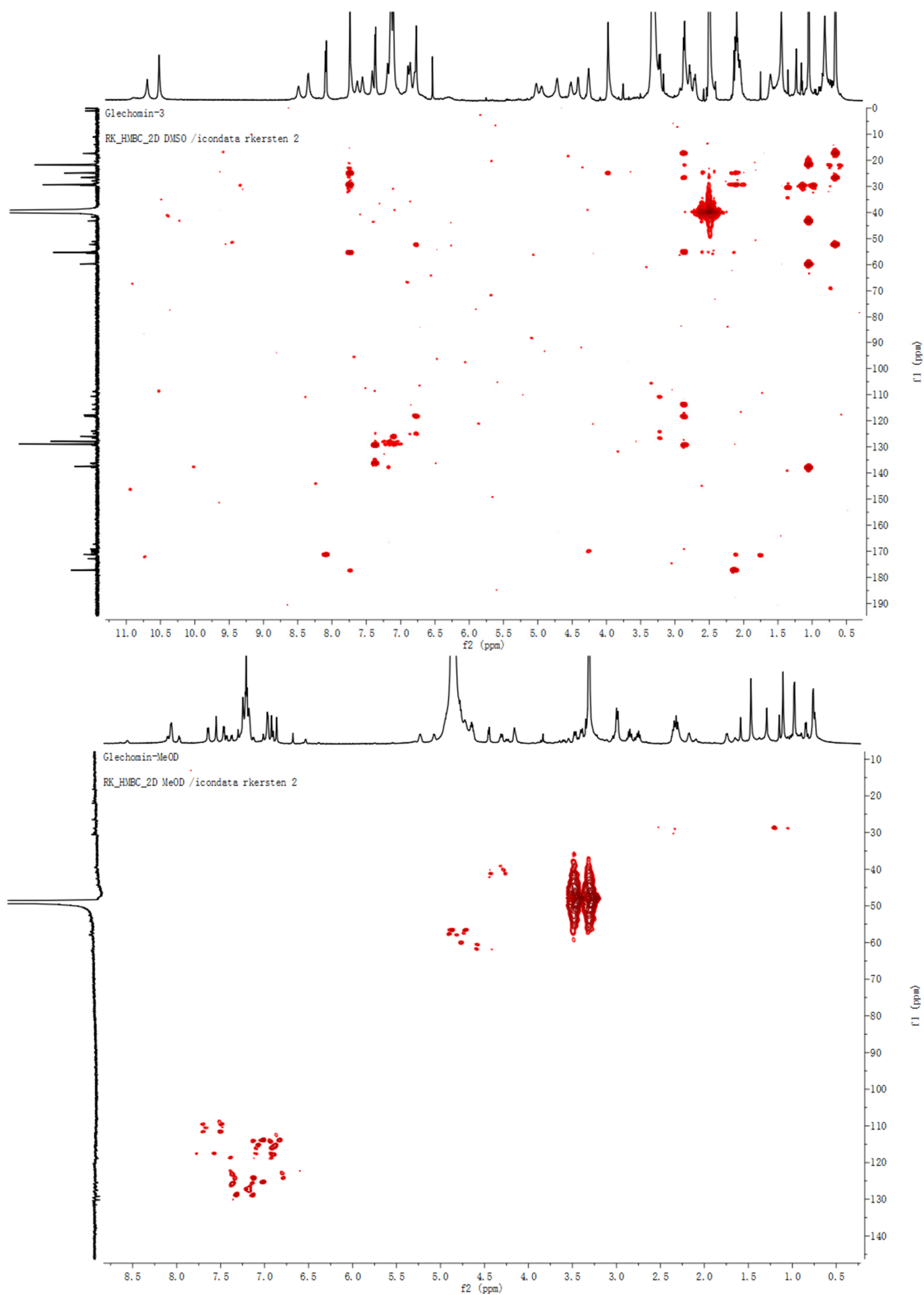
Supplementary Figure 5. ^1H - ^1H COSY-NMR spectrum of glechomanin in $\text{DMSO}-d_6$ and in $\text{methanol}-d_4$.



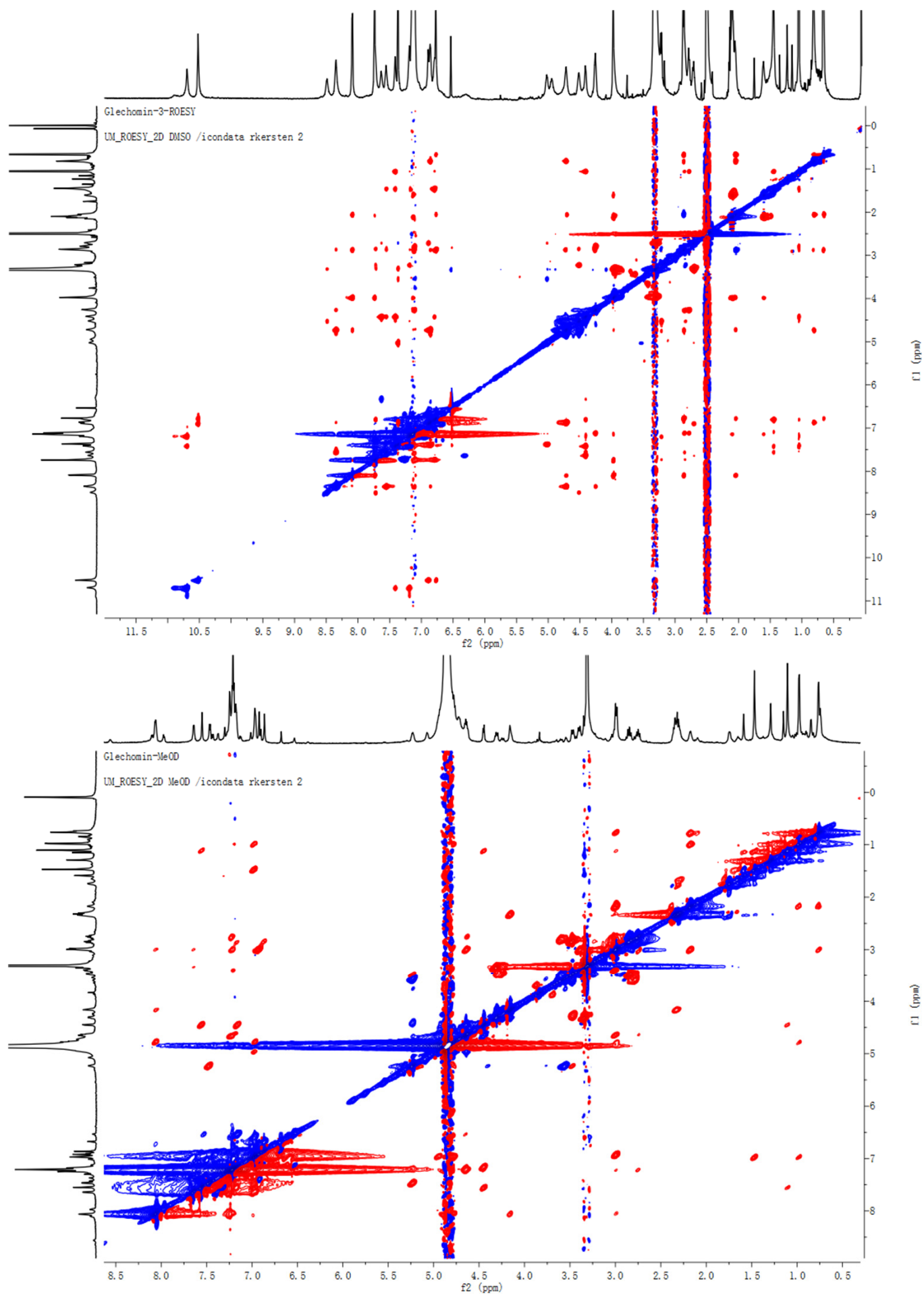
Supplementary Figure 6. ^1H - ^1H -TOCSY-NMR spectrum of glechomatin in DMSO- d_6 and in methanol- d_4 .



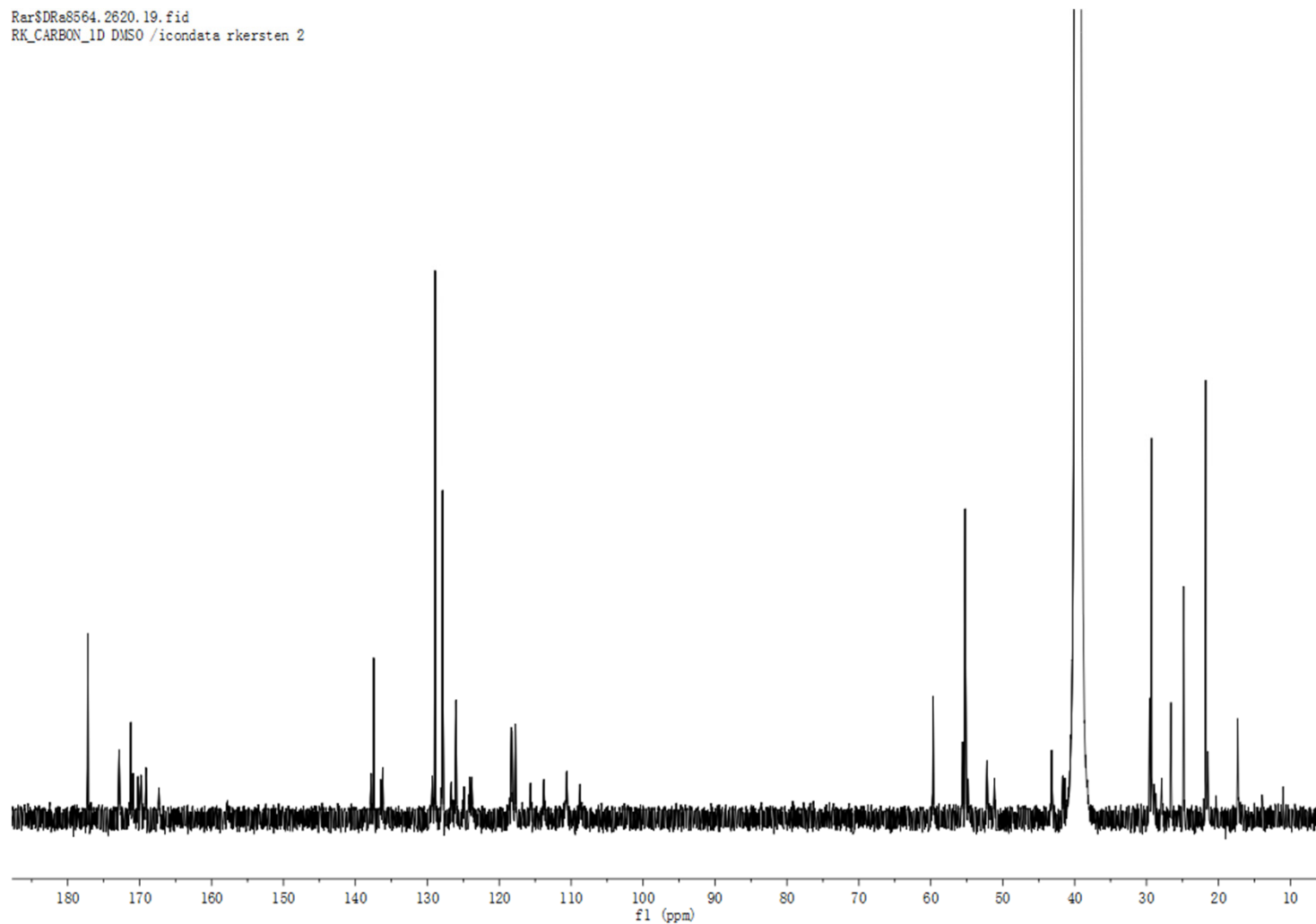
Supplementary Figure 7. HSQC-NMR spectrum of glechomanin in DMSO-d₆ and in methanol-d₄.



Supplementary Figure 8. HMBC-NMR spectrum of glechomanin in DMSO- d_6 and in methanol- d_4 .

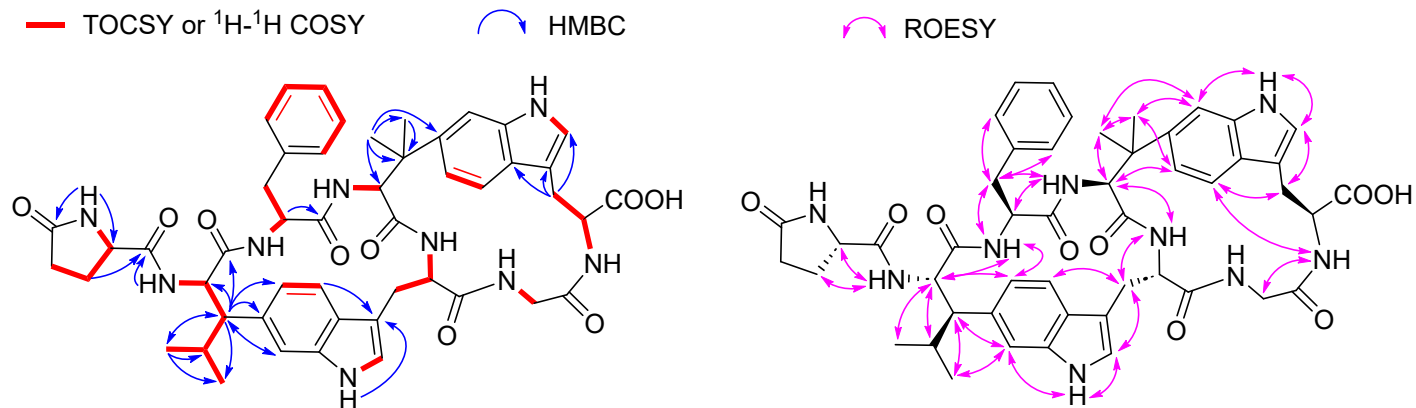


Supplementary Figure 9. ROESY-NMR spectrum of glechomanin in DMSO-d₆ and in methanol-d₄.

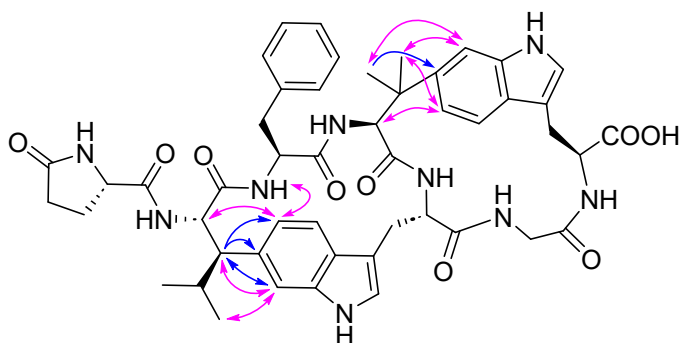


Supplementary Figure 10. ^{13}C -NMR spectrum (200 MHz) of glechomanin in DMSO-d_6 .

2D NMR correlations in DMSO-d₆:

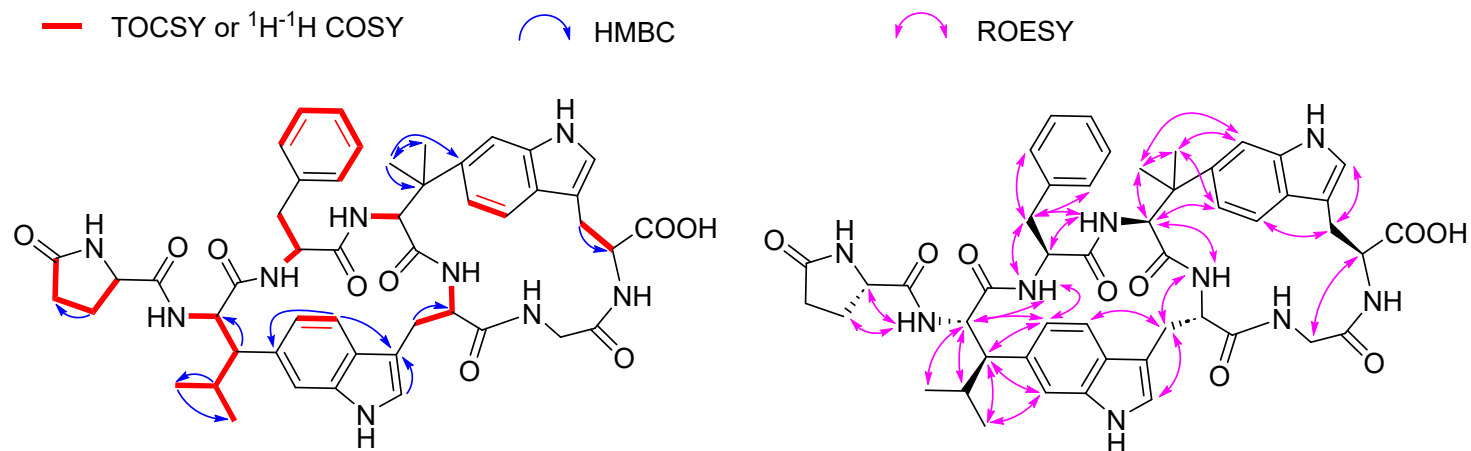


Crosslinks: Key HMBC and ROESY correlations

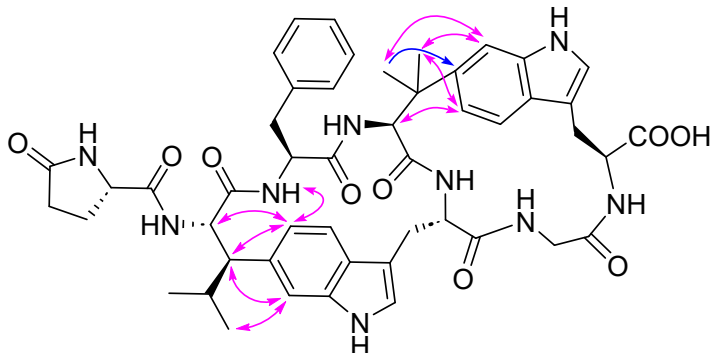


Supplementary Figure 11. Key 2D NMR correlations of glechomanin in DMSO-d₆.

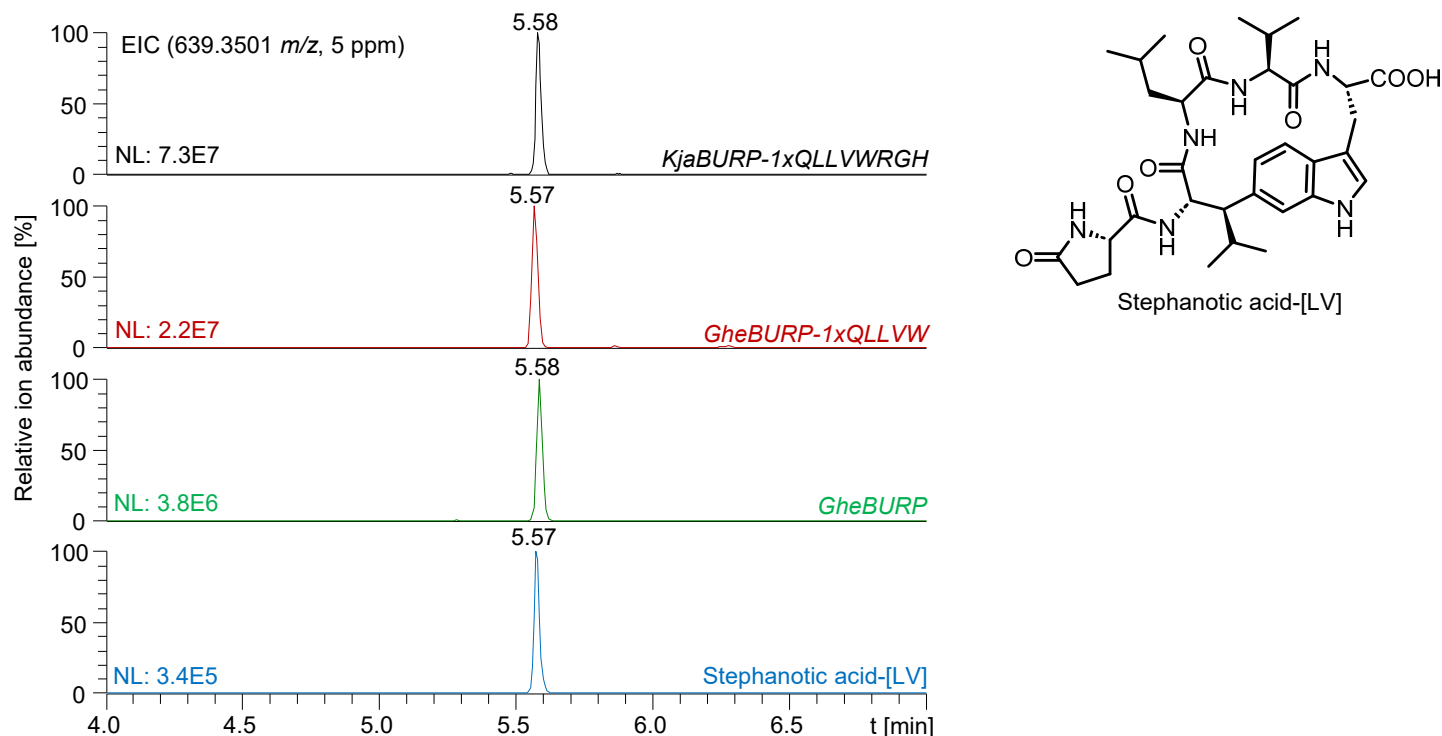
2D NMR correlations in methanol-d₄



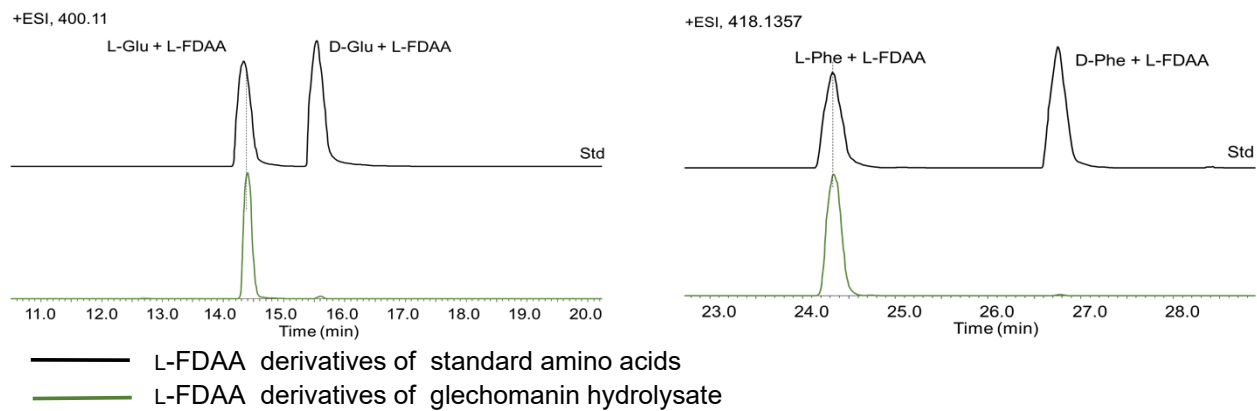
Crosslinks: Key HMBC and ROESY correlations



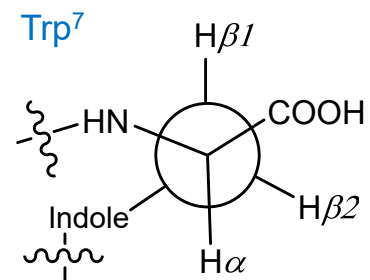
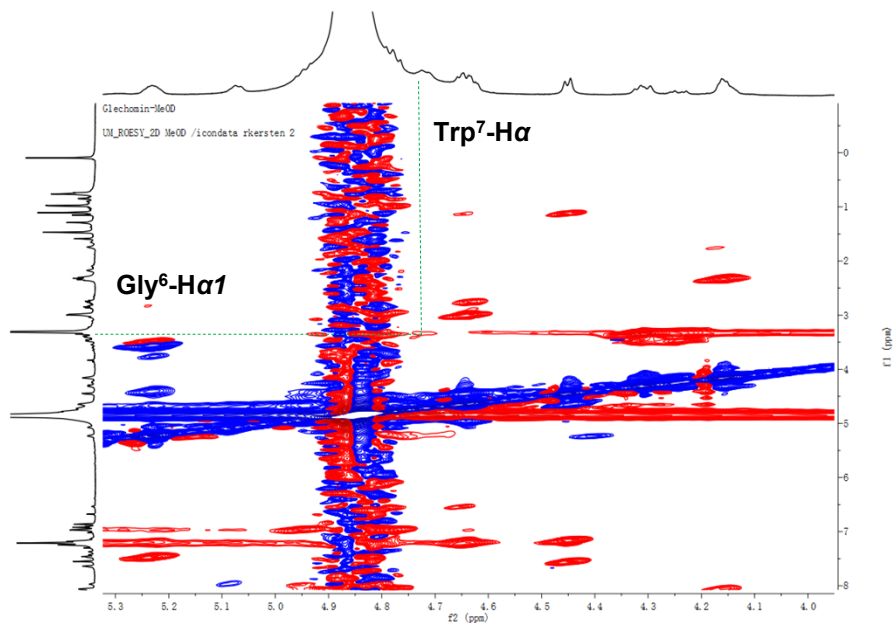
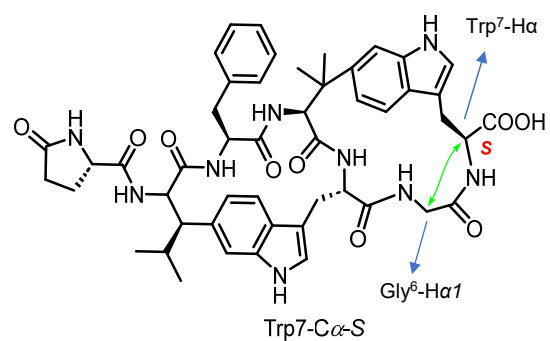
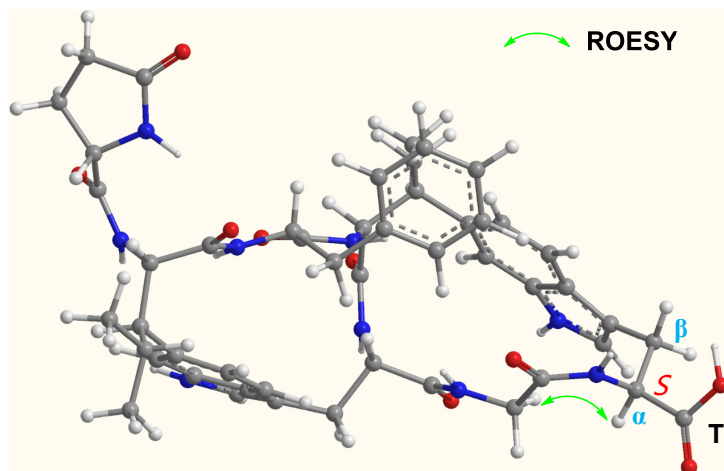
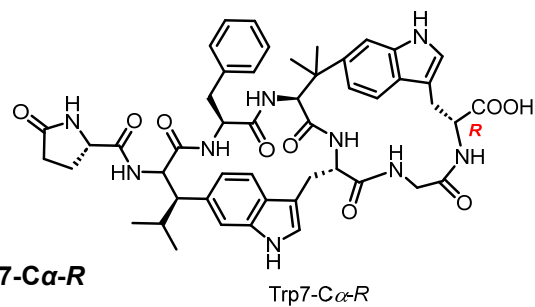
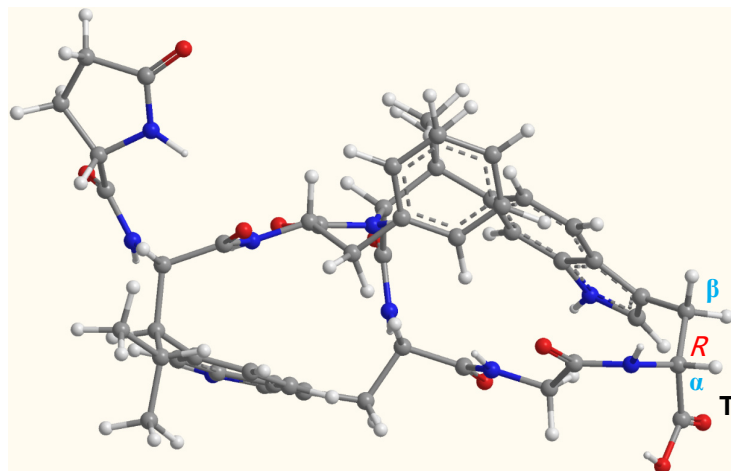
Supplementary Figure 12. Key 2D NMR correlations of glechomanin in methanol-d₄.



Supplementary Figure 13. Stereochemical analysis of glechomanin. LCMS-based comparison of glechomanin first-ring analytes compared to authentic standard stephanotic acid-[LV] after transient expression of *KjaBURP-1xQLLVWRGH* (moroidin cyclase), *GheBURP* (glechomanin cyclase including a QLLVW core peptide) and *GheBURP-1xQLLVW* (glechomin cyclase with truncated substrate domain to include one QLLVW core peptide) for 6 days in *N. benthamiana* via expression vector pEAQ-HT and *Agrobacterium tumefaciens* LBA4404 infiltration. Extracted ion chromatogram (EIC) traces of methanolic *N. benthamiana* leaf extracts are compared to authentic standard of stephanotic acid-[LV] (100 ng/mL).

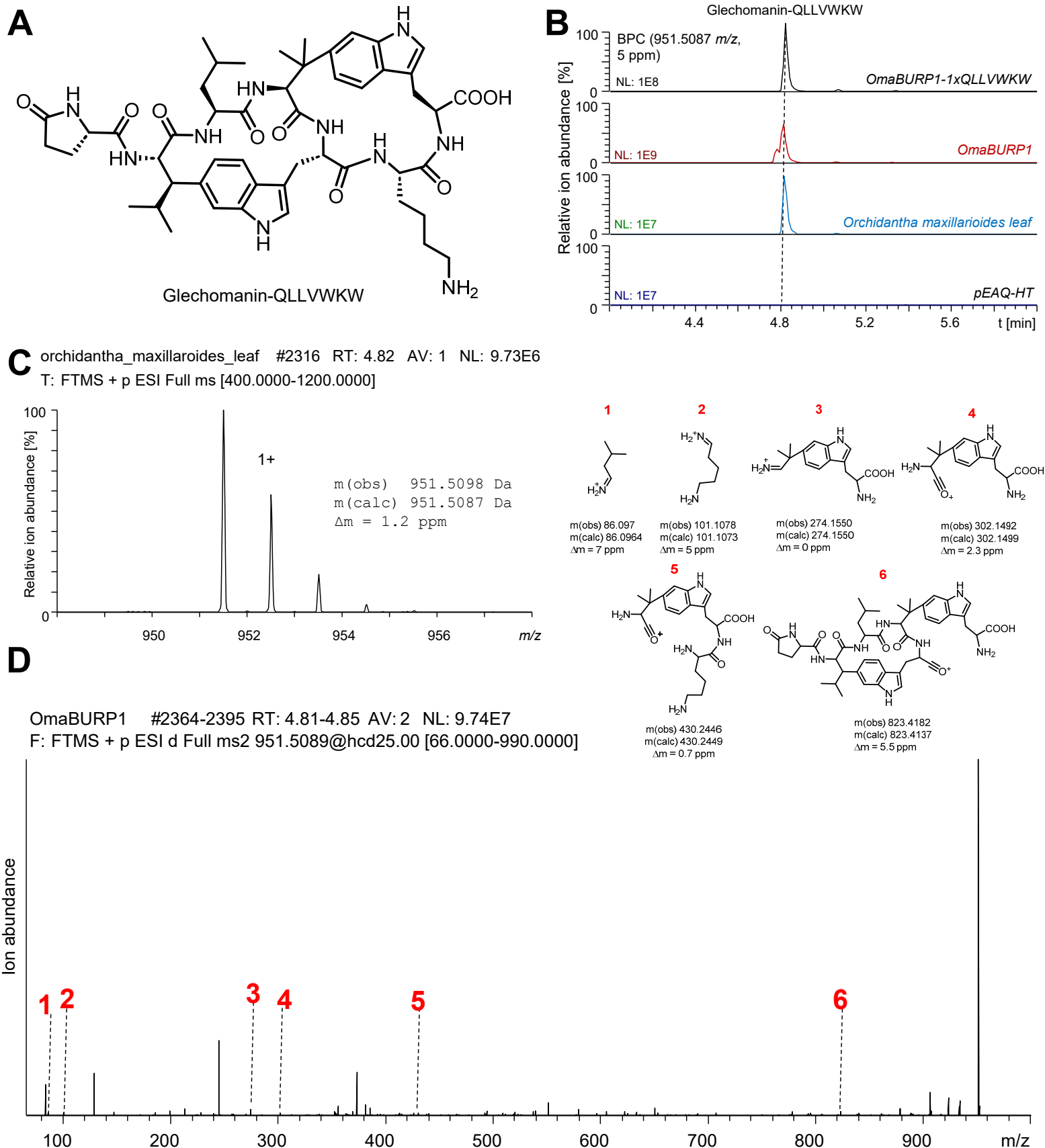


Supplementary Figure 14. L-FDAA derivatization of the hydrolysate of glechomanin and amino acid standards.

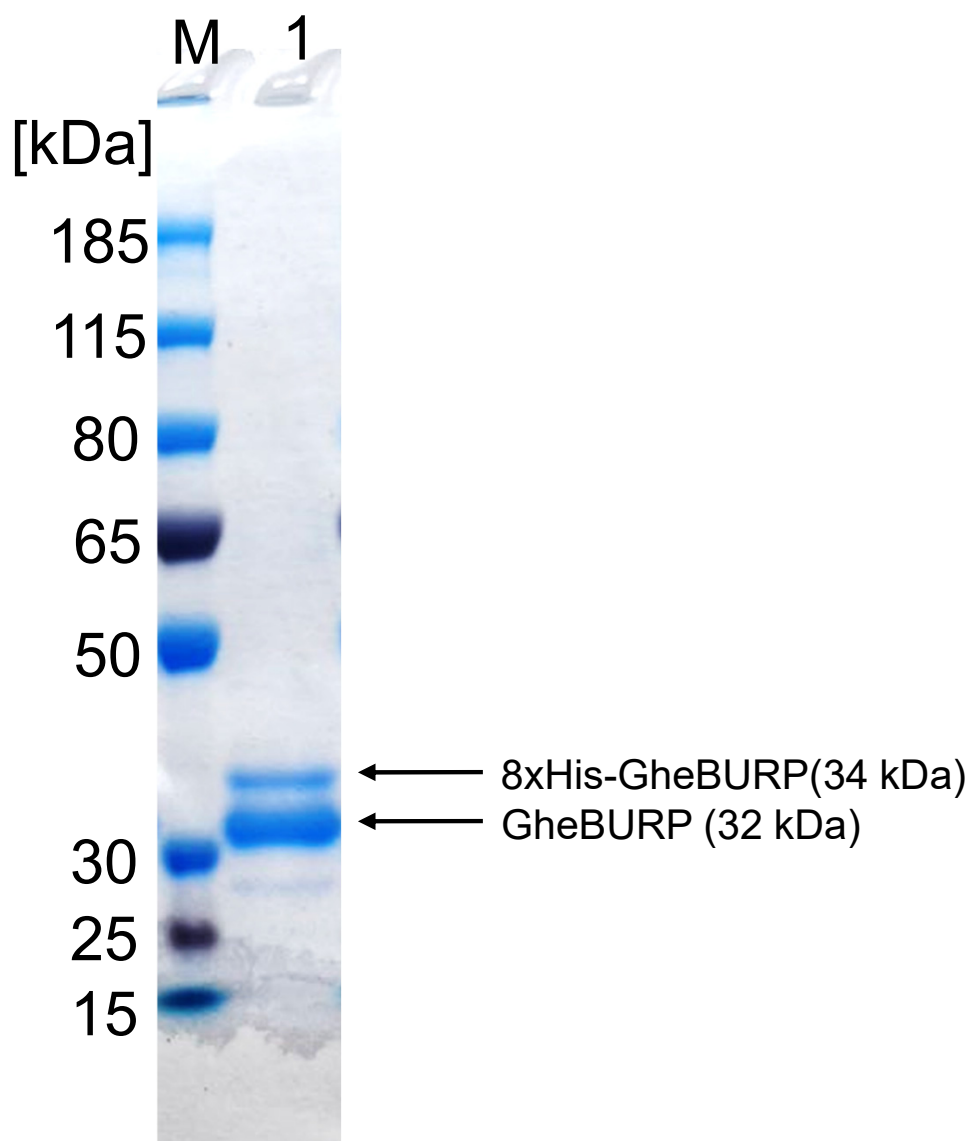


Trp7: $J_{\alpha\beta1} = 11.2$ Hz
S-erythro: $J_{\alpha\beta1} > 6$
R-threo: $J_{\alpha\beta1} < 6$

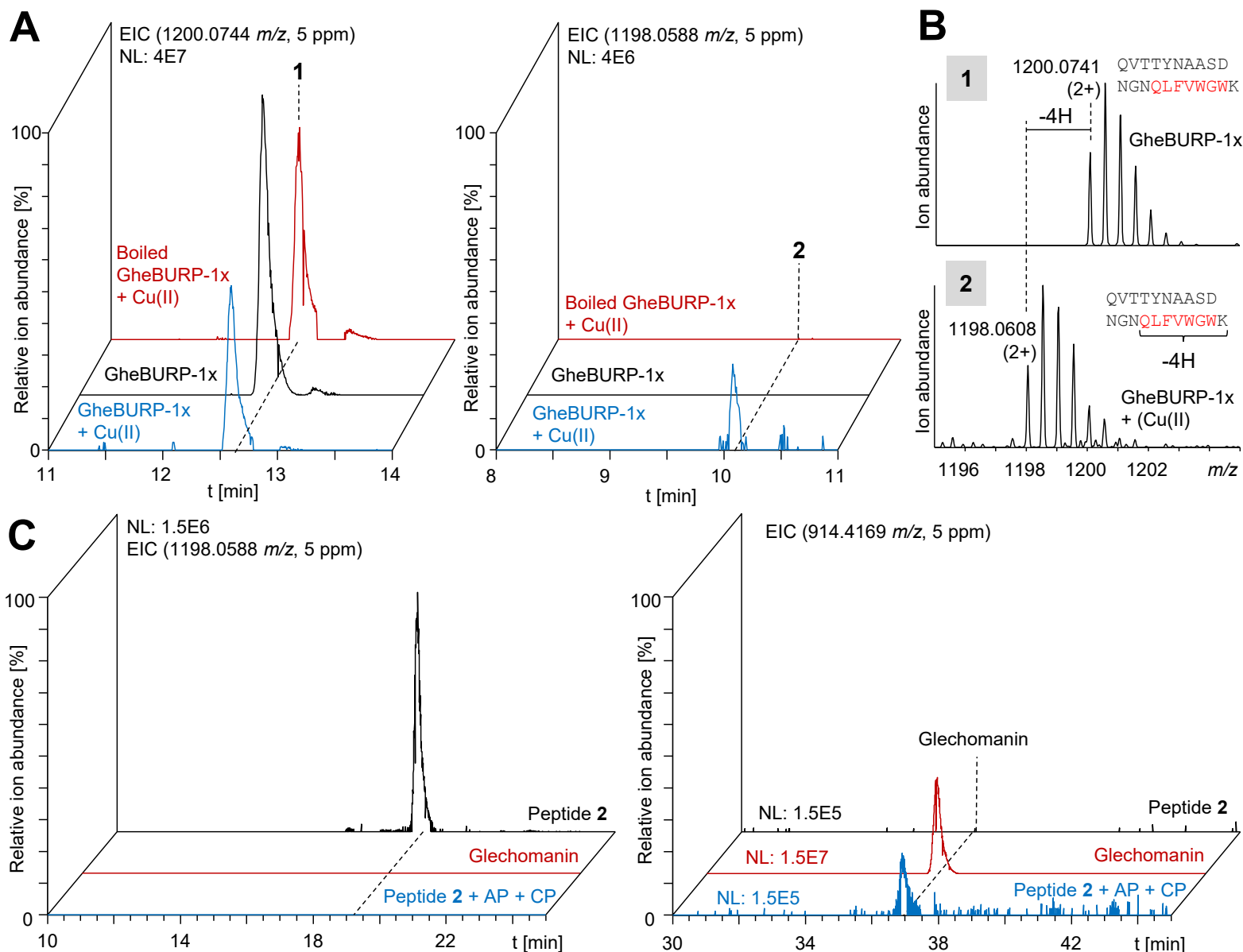
Supplementary Figure 15. Glechomanin geometry optimization and ROESY analysis of Trp7-Cα assignment.



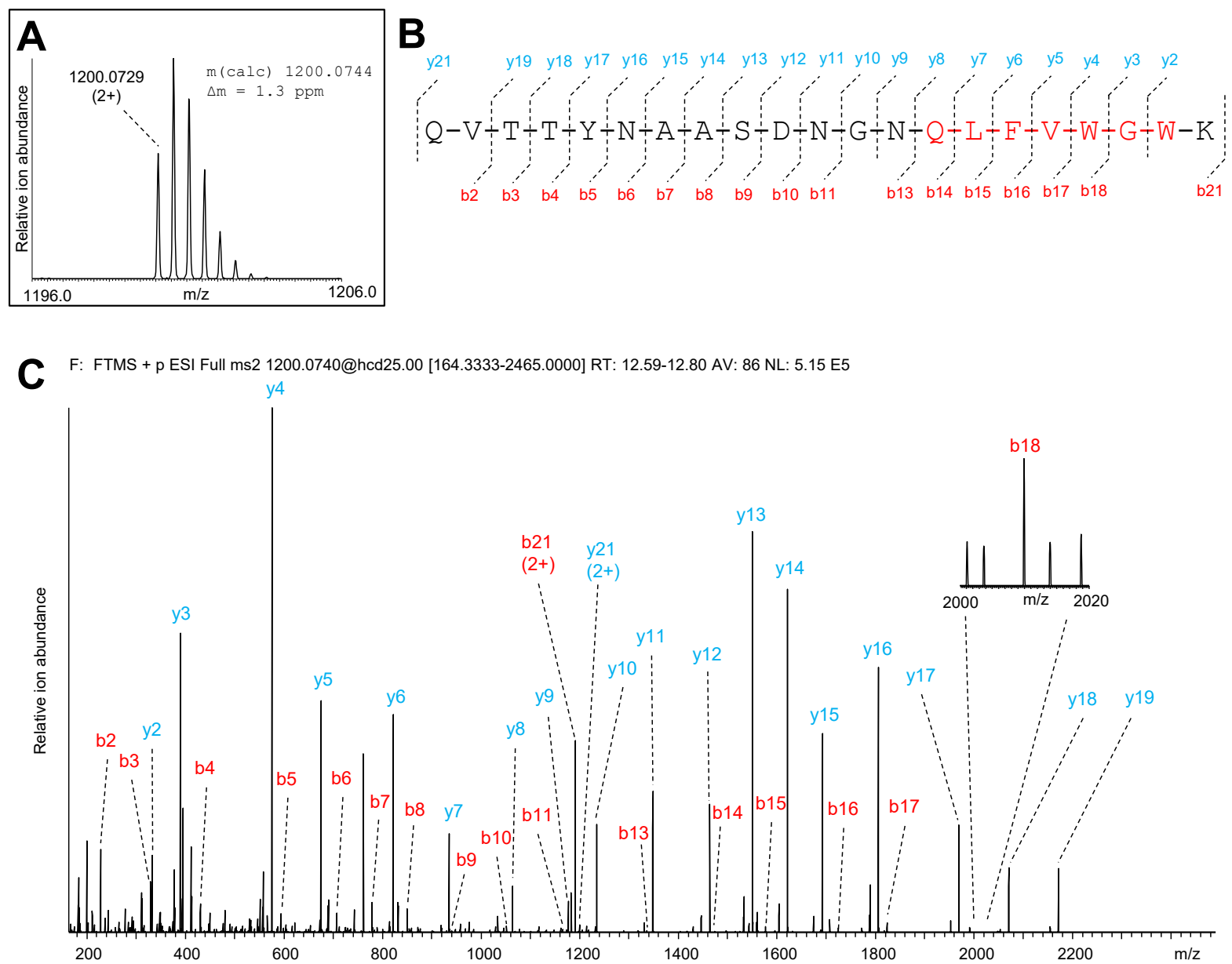
Supplementary Figure 16. Characterization of glechomanin biosynthesis in *Orchidantha maxillarioides*. (A) Predicted glechomanin-QLLVWKW chemotype. (B) LCMS characterization of predicted glechomanin-QLLVWKW in methanolic extracts of *O. maxillarioides* leaf and transgenic *N. benthamiana* after 6 days of expression of *OmaBURP1*. (C) MS analysis of predicted glechomanin-QLLVWKW. (D) MS/MS analysis of predicted glechomanin-QLLVWKW.



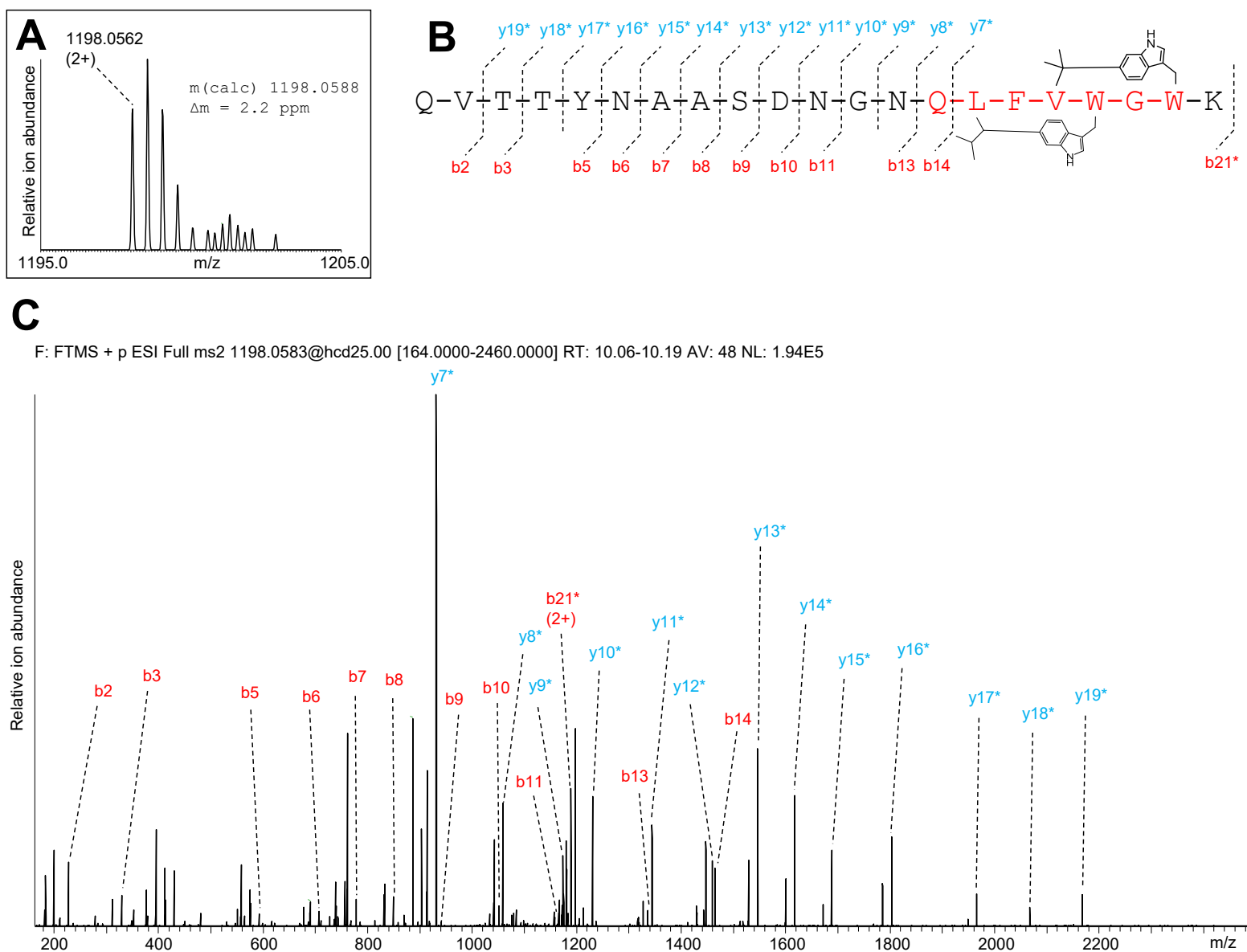
Supplementary Figure 17. Glechomanin cyclase purification. SDS-PAGE of GheBURP-1xQLFVWGW after denature-refold purification. Source data are provided as a Source Data file.



Supplementary Figure 18. LCMS analysis of *in vitro* reconstitution of GheBURP-1xQLFVWG. (A) Detection of bicyclic core peptide modifications in purified GheBURP-1xQLFVWG after incubation with Cu(II) at pH 7. EIC traces of linear and modified tryptic core peptide QVTTYNAASDNGNQ**LFV**WGWK. (B) MS spectra of linear and modified tryptic core peptide. Core peptides are highlighted in red. (C) Detection of glechomanin after incubation of modified bicyclic tryptic core peptide QVTTYNAASDNGNQ**LFV**WGWK (peptide 2) with aminopeptidase (AP) and carboxypeptidase (CP).



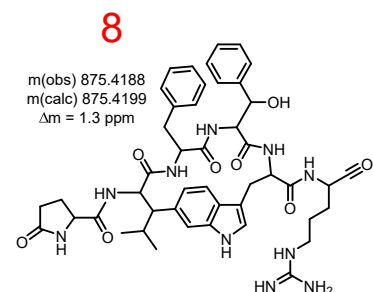
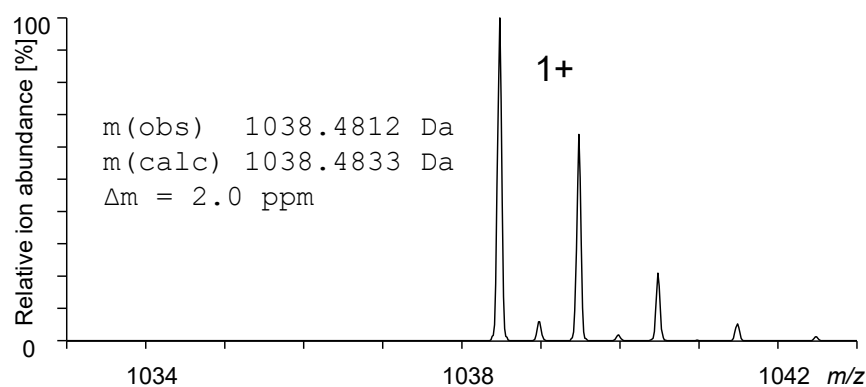
Supplementary Figure 19. MS and MS/MS analysis of unmodified core peptide from *in vitro* reconstitution assay of GheBURP-1xQLFVWGK. (A) MS1 signal of linear tryptic core peptide QVTTYNAASDNGNQLFVWGK. Isotope selected for fragmentation labeled by grey box. (B) Detected tandem MS fragments of linear tryptic core peptide. (C) Tandem MS spectrum of linear tryptic core peptide with labeled fragment peaks. Tandem MS fragmentation table of linear tryptic core peptide is Supplementary Table 8.



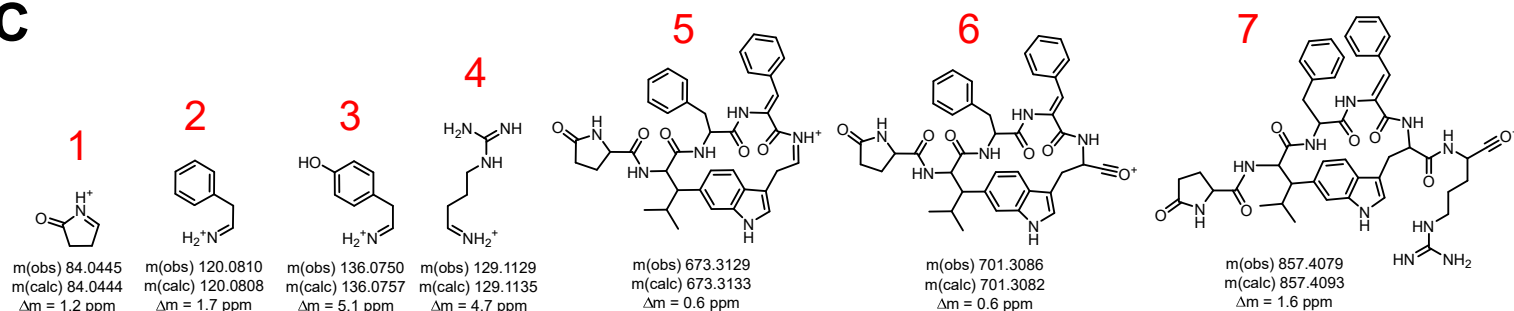
Supplementary Figure 20. MS and MS/MS analysis of modified core peptide from *in vitro* reconstitution of GheBURP-1xQLFVWGW. (A) MS1 signal of putative bicyclic tryptic core peptide QVTTYNAASDNGNQLFVWGWK. Isotope selected for fragmentation labeled by grey box. (B) Detected tandem MS fragments of putative bicyclic tryptic core peptide. (C) Tandem MS spectrum of putative bicyclic tryptic core peptide with labeled fragment peaks. Tandem MS fragmentation table of putative bicyclic tryptic core peptide is Supplementary Table 9.

A $m(\text{calc, QLFFWRY}) = 1059.5411 \text{ [M+H]}^+$
 $- (1 \times 17.02655) \text{ (pyroglutamate formation from Gln)}$
 $- (4 \times 1.00782) \text{ (bicyclization)}$
1038.4833 Da

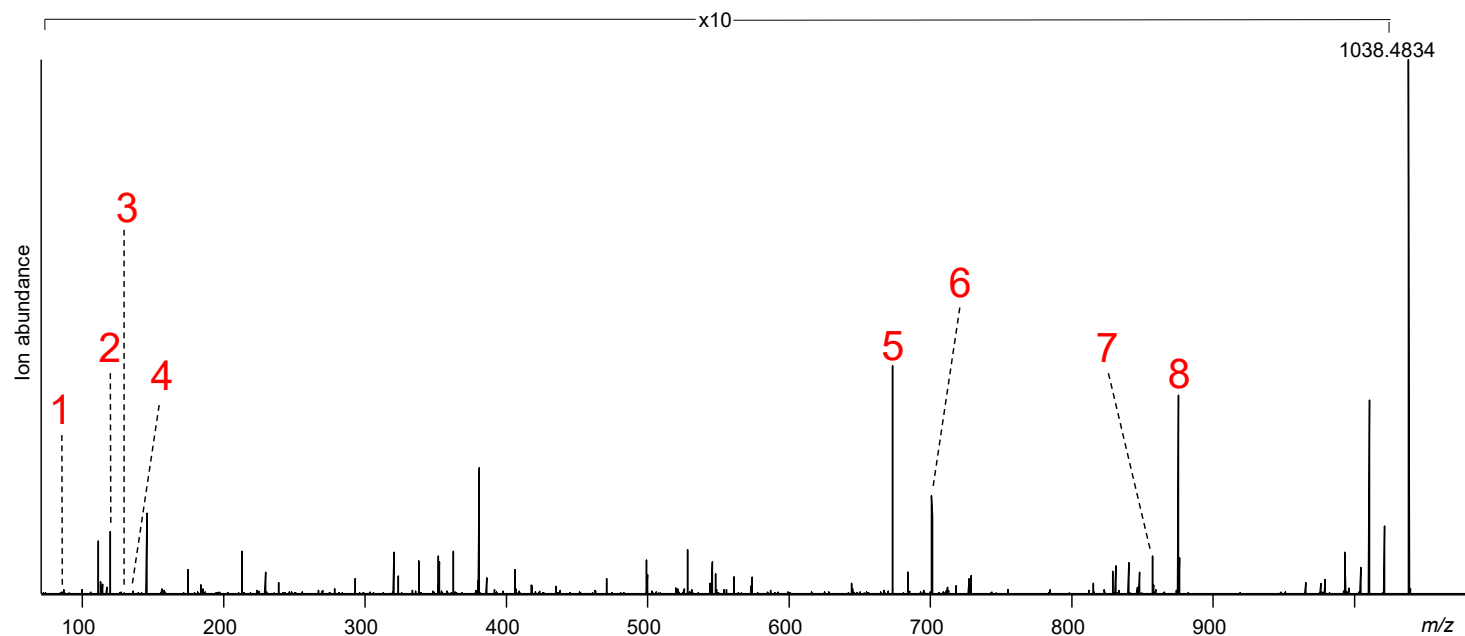
B Mercurialin #2217 RT: 5.20 AV: 1 NL: 1.87E9
T: FTMS + p ESI Full ms [400.0000-1200.0000]



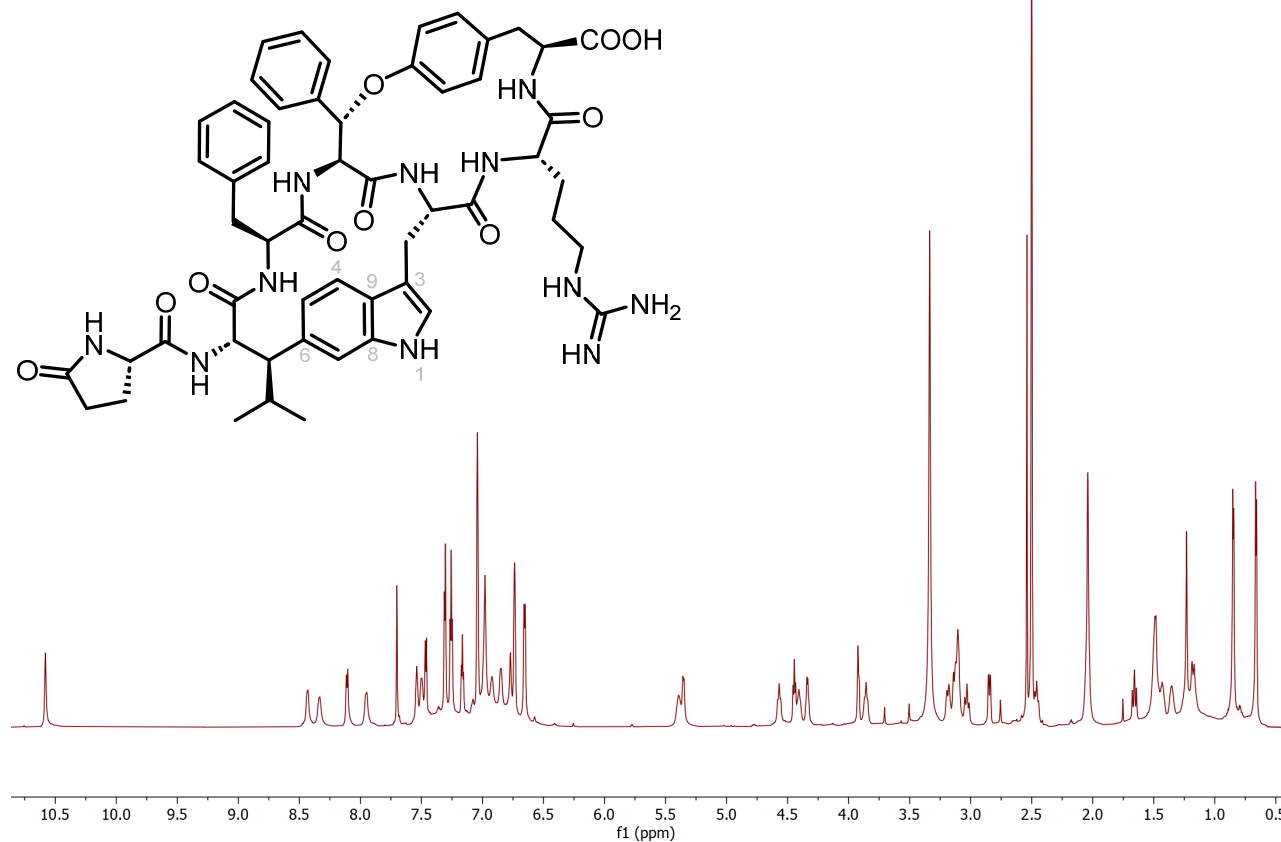
C



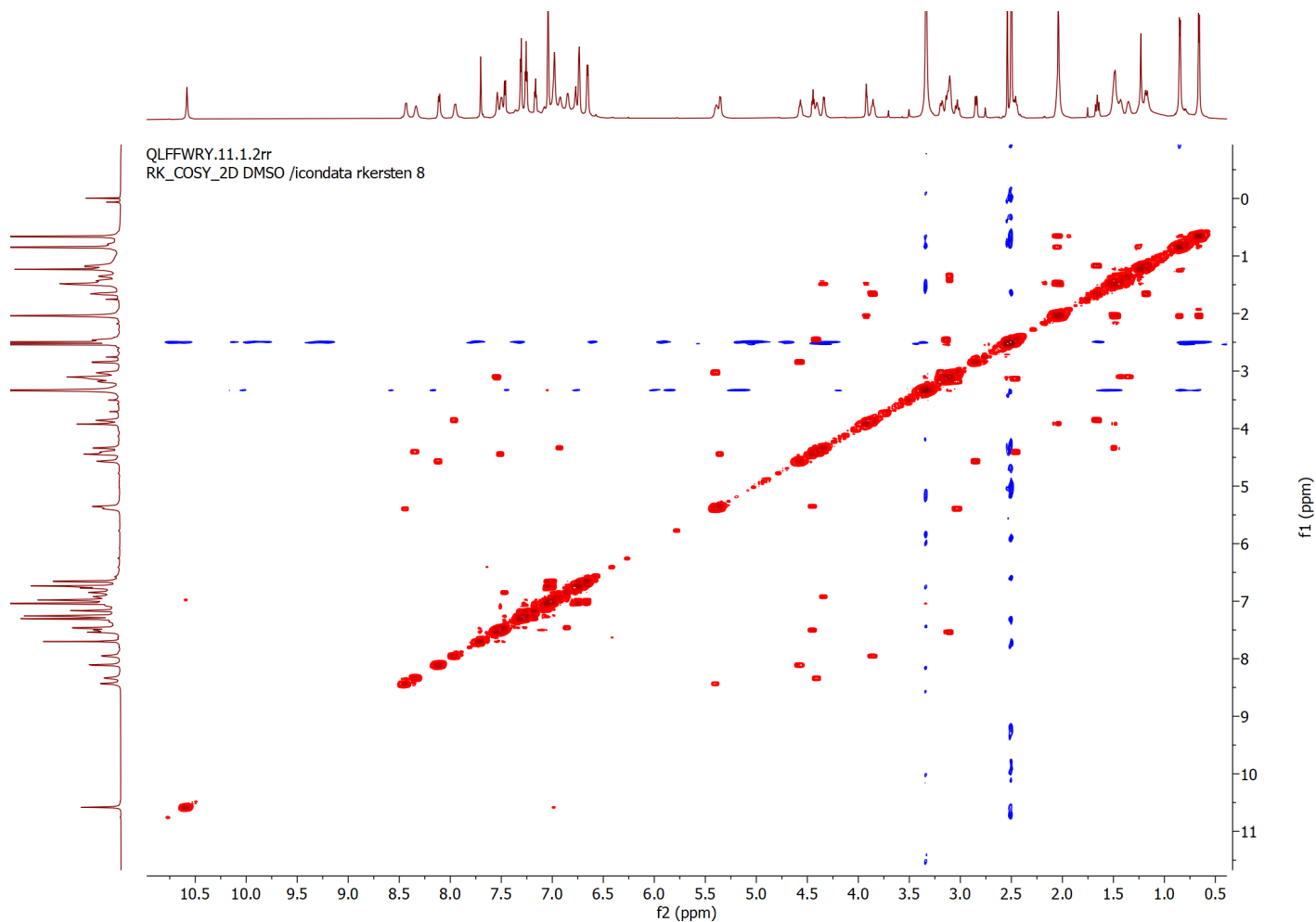
Mercurialin #2150-2286 RT: 5.17-5.31 AV: 8 NL: 2.64E8
F: FTMS + p ESI d Full ms2 1038.4828@hcd25.00 [71.6667-1075.0000]



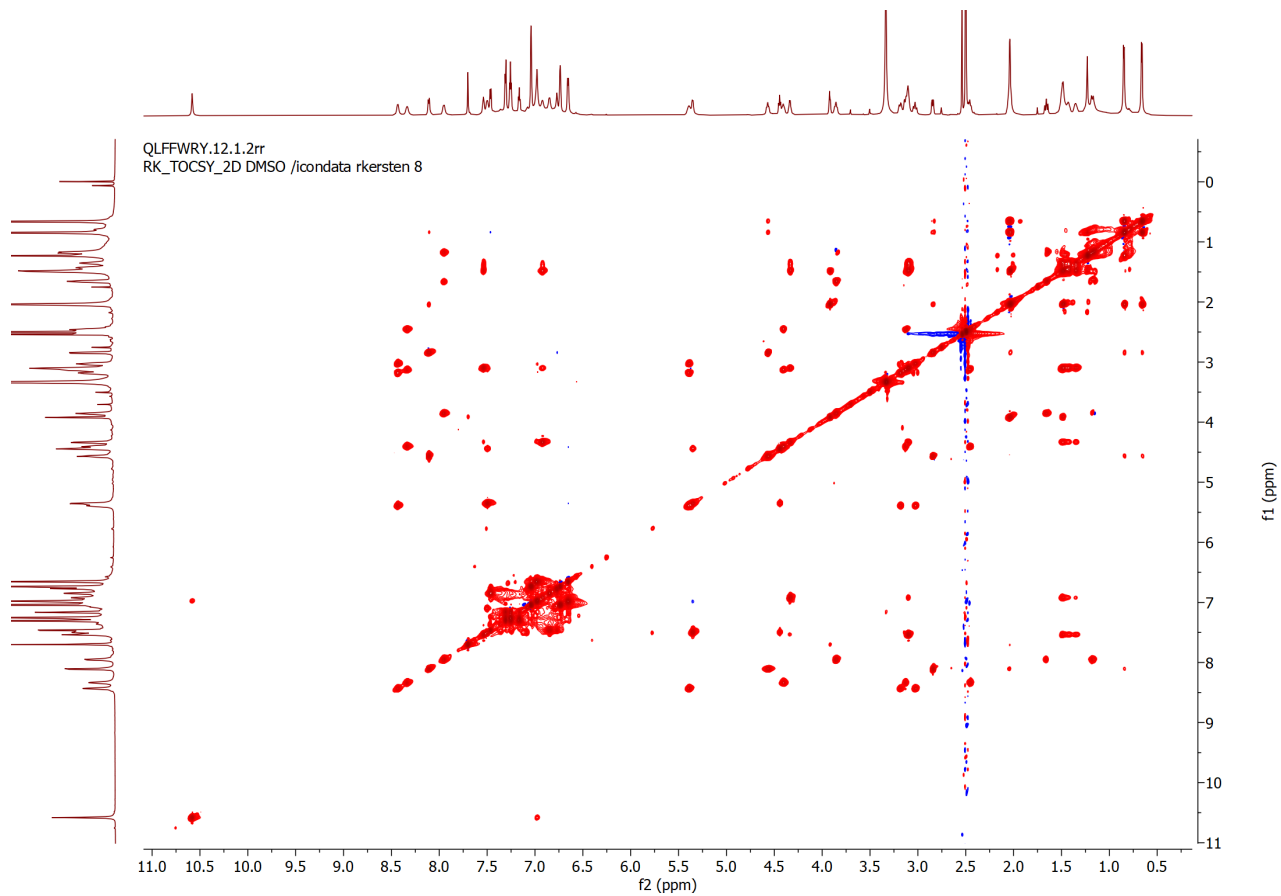
Supplementary Figure 21. MS and MS/MS spectrum of mercurialin. (A) Mass calculation of mercurialin from predicted core peptide sequence and posttranslational modifications. (B) MS1 spectrum of mercurialin. (C) MS/MS spectrum of mercurialin and predicted fragments.



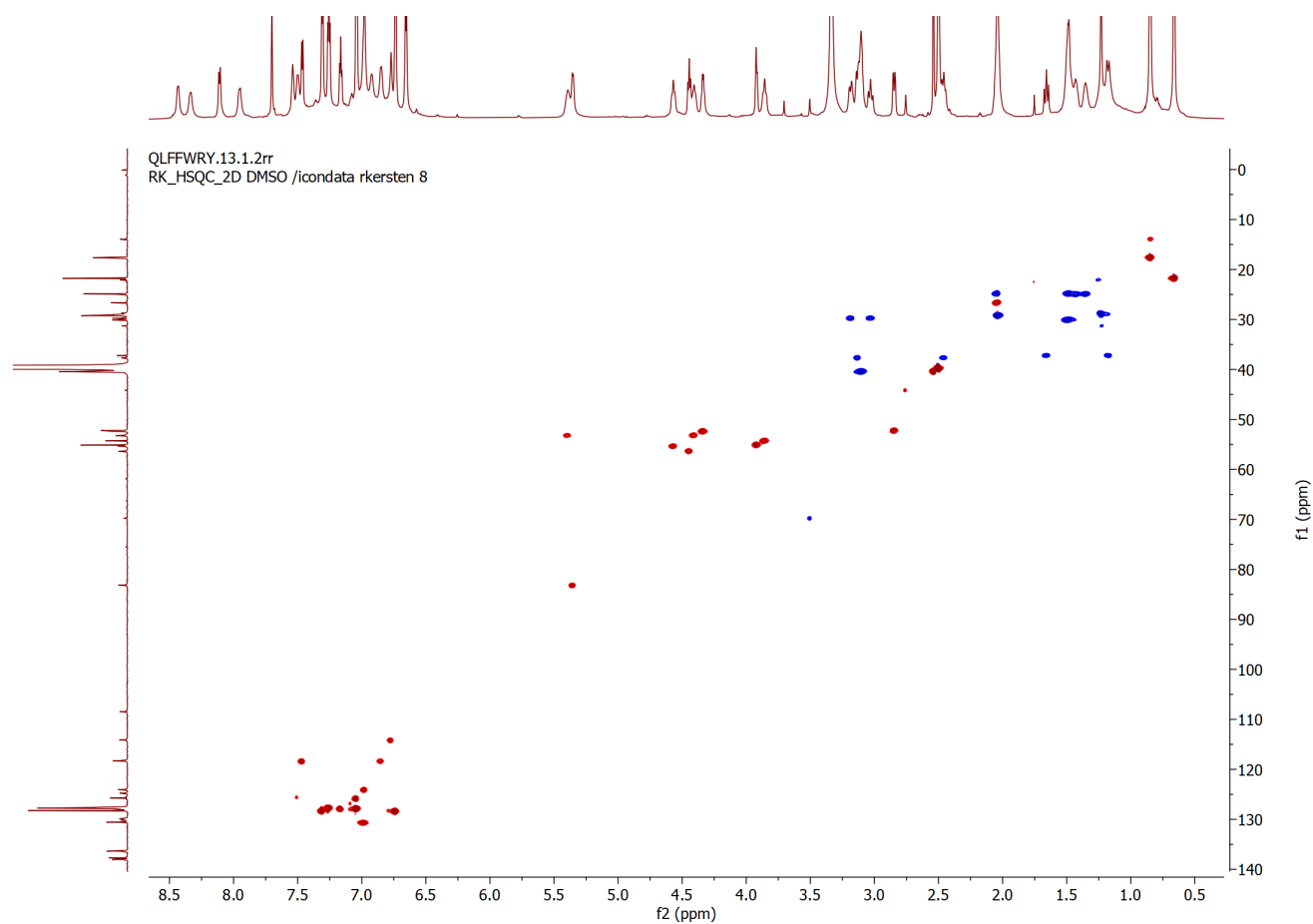
Supplementary Figure 22. ^1H -NMR spectrum of mercurialin in DMSO- d_6 (800 MHz).



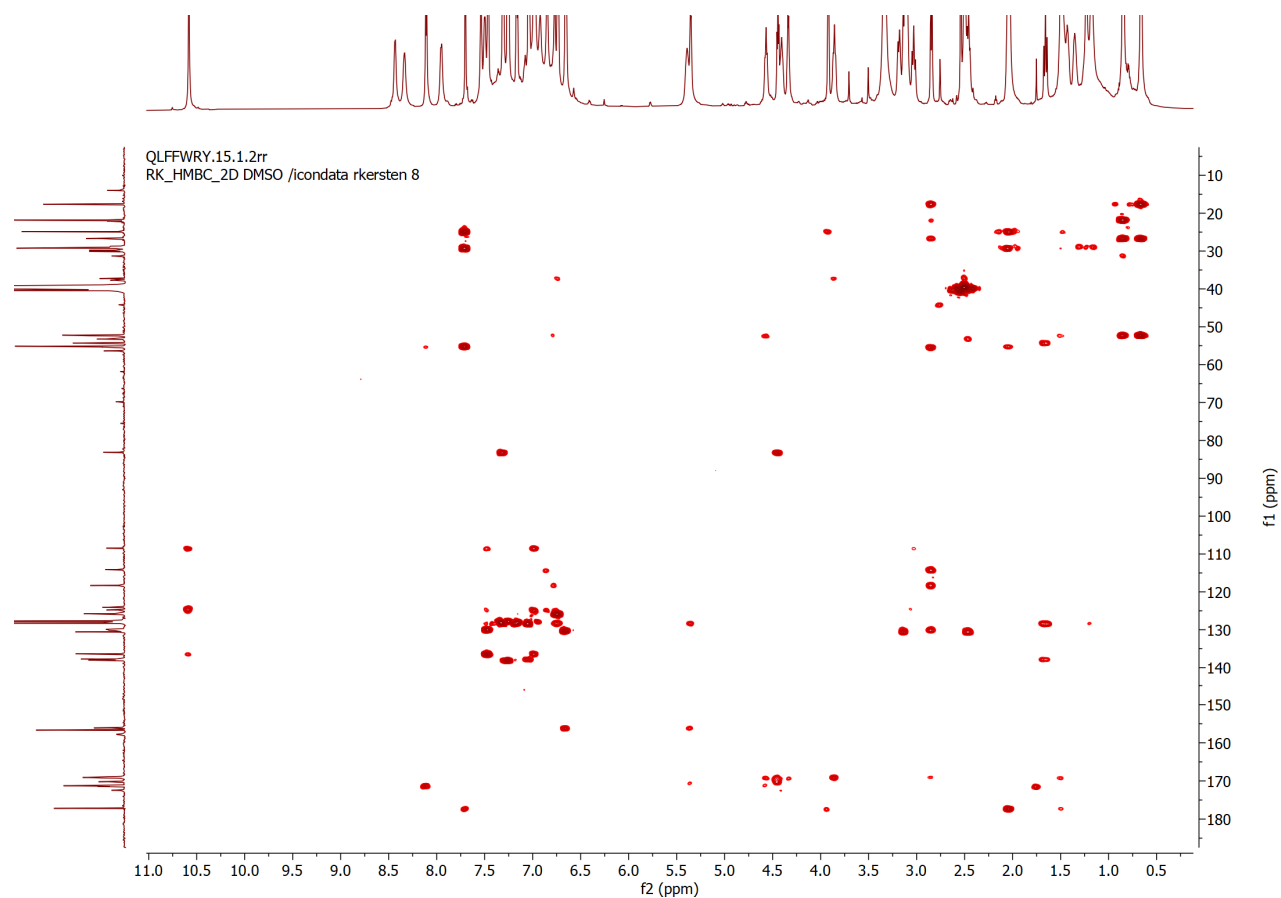
Supplementary Figure 23. ^1H - ^1H -COSY-NMR spectrum of mercurialin in DMSO- d_6 .



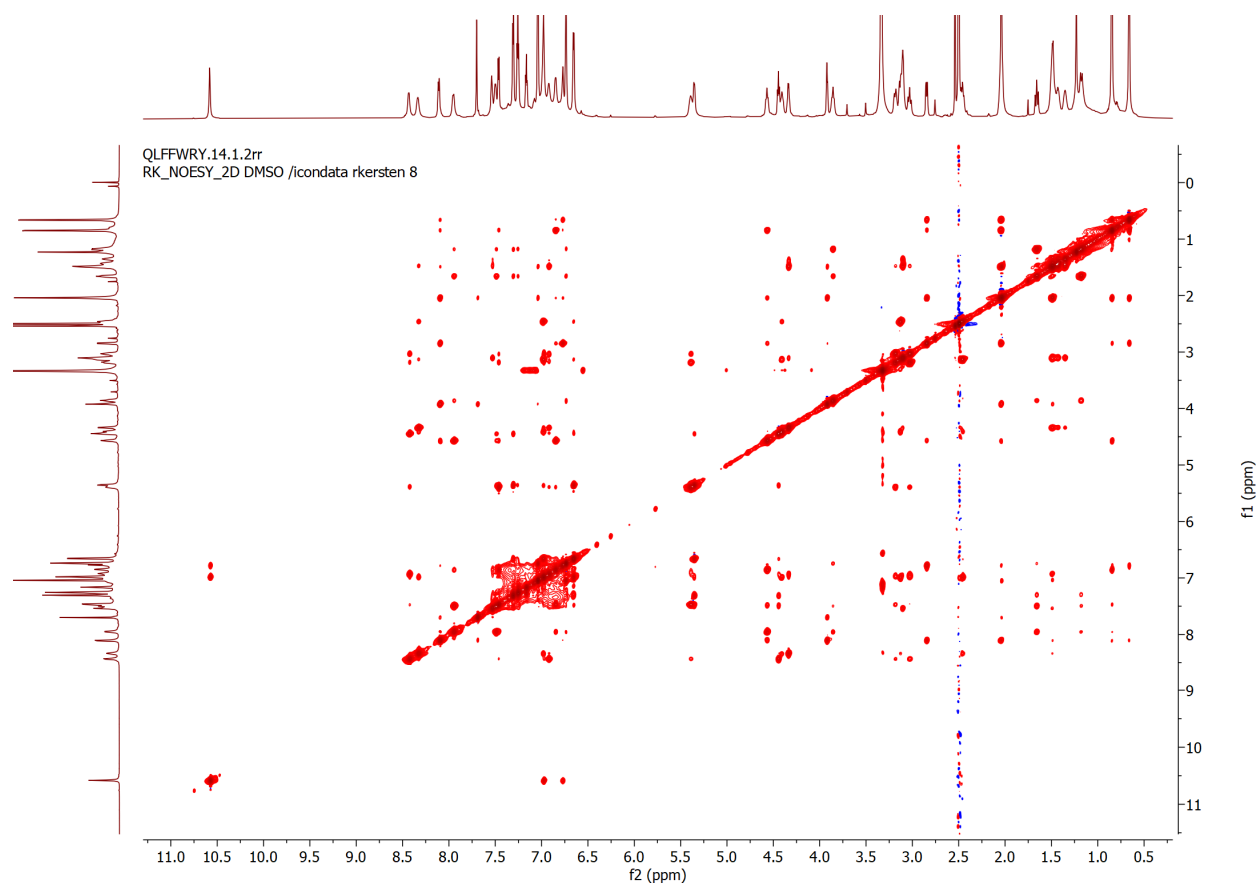
Supplementary Figure 24. ^1H - ^1H -TOCSY-NMR spectrum of mercurialin in DMSO-d_6 .



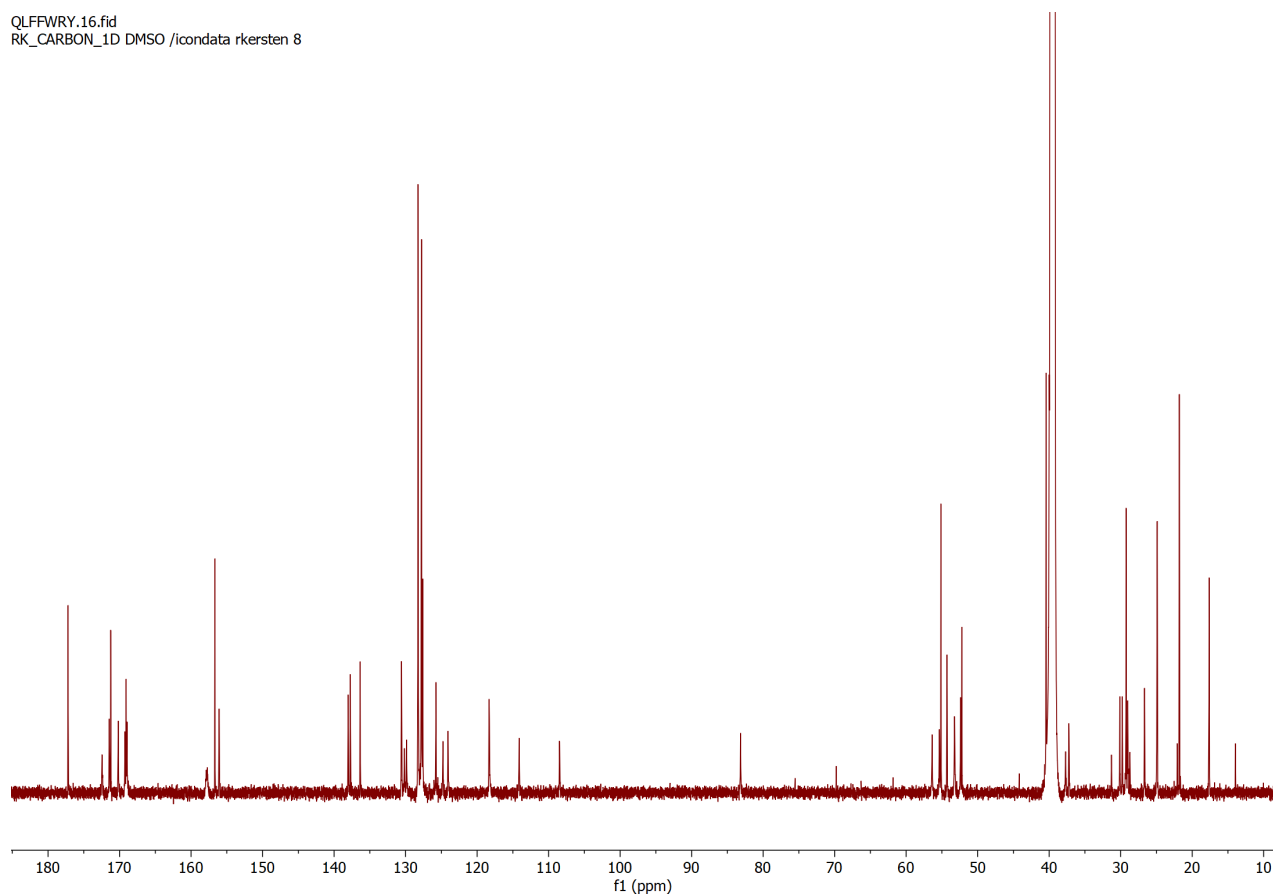
Supplementary Figure 25. HSQC-NMR spectrum of mercurialin in DMSO-d_6 .



Supplementary Figure 26. HMBC-NMR spectrum of mercurialin in DMSO-d₆.

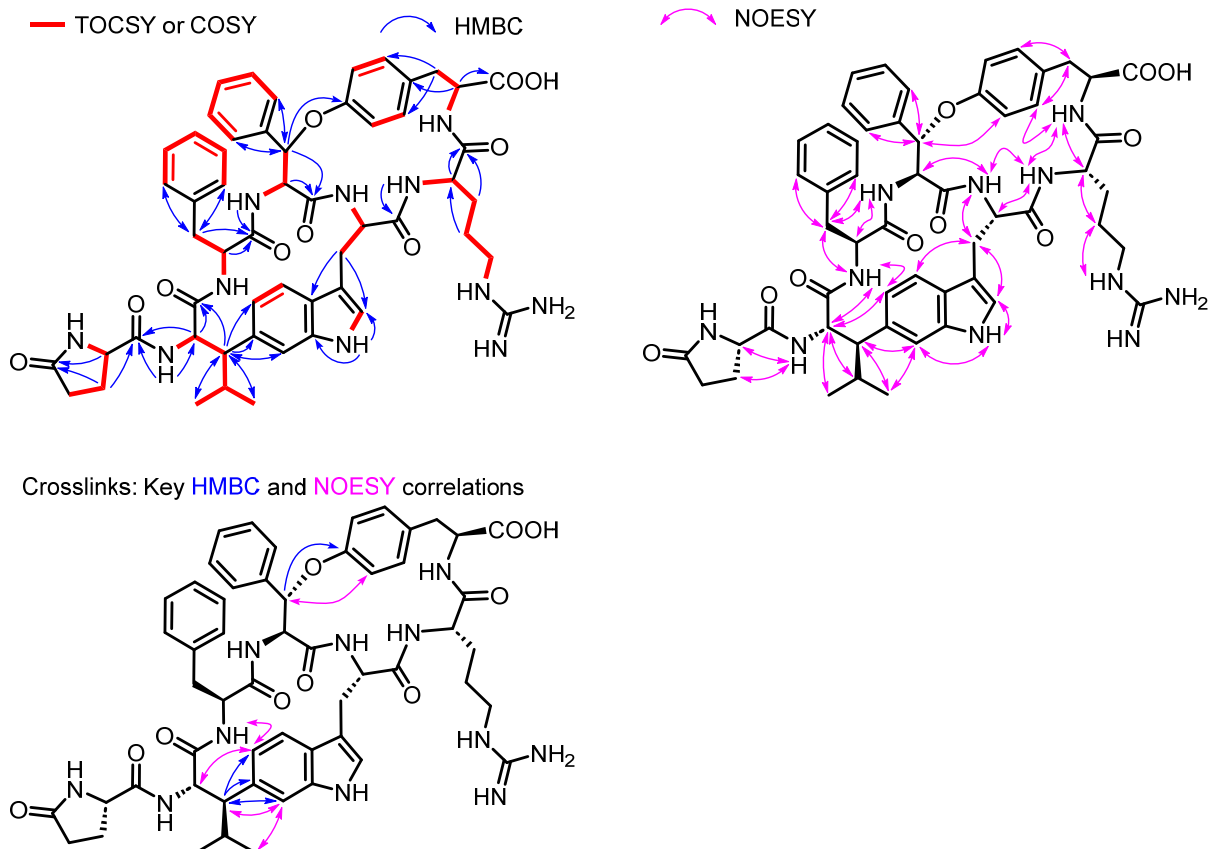


Supplementary Figure 27. NOESY-NMR spectrum of mercurialin in DMSO-d₆.

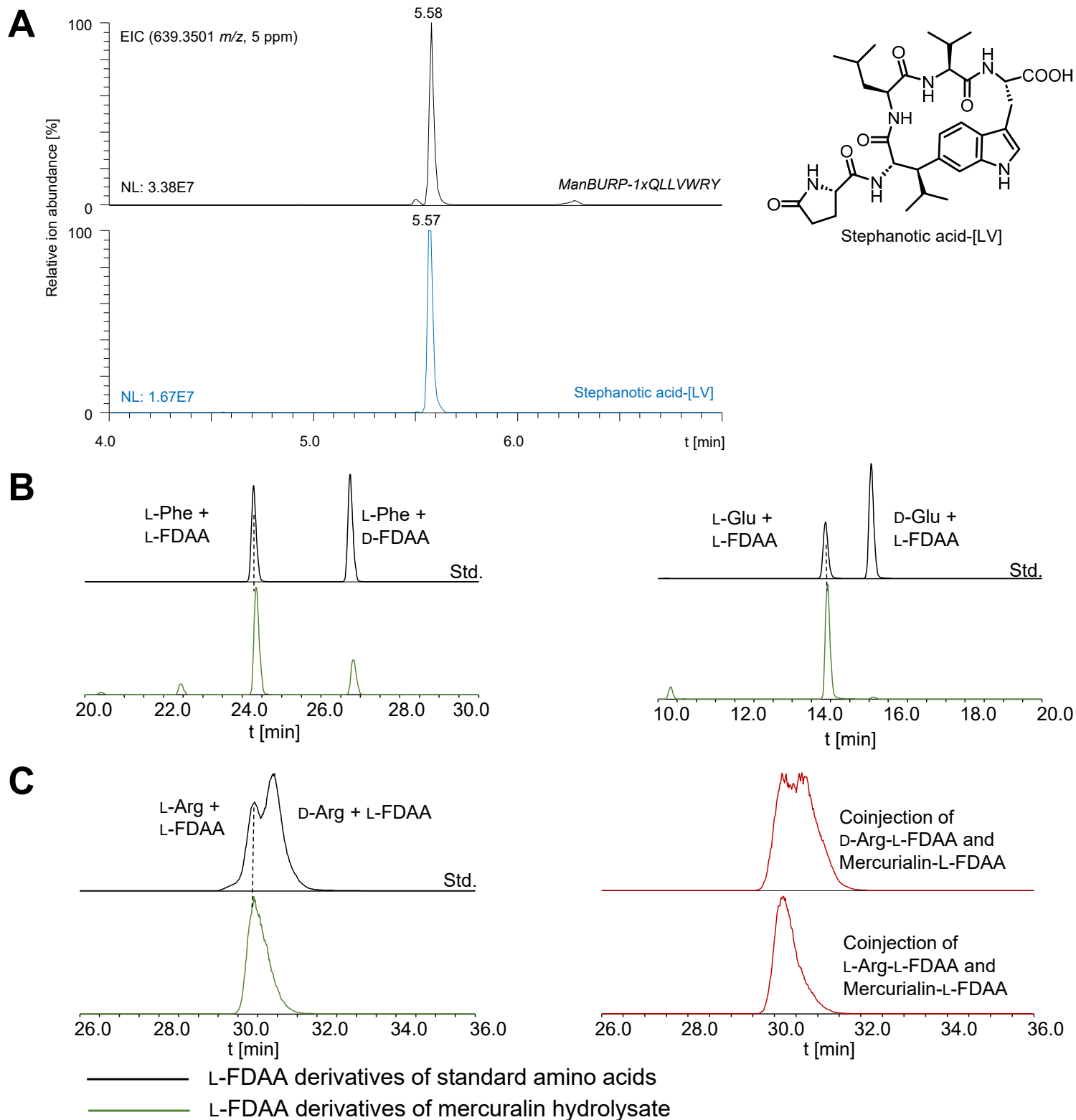


Supplementary Figure 28. ^{13}C -NMR (200 MHz) spectrum of mercurialin in DMSO-d_6 .

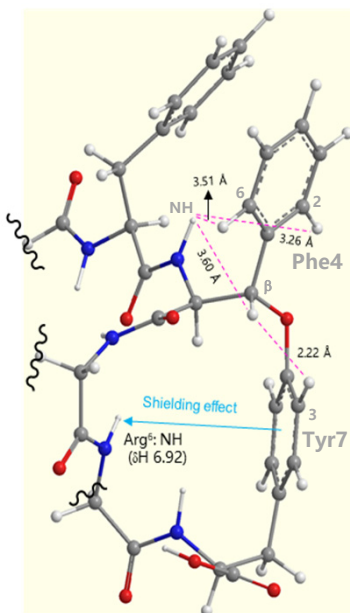
2D NMR correlations in DMSO-d_6 :



Supplementary Figure 29. Key 2D NMR correlations of mercurialin.

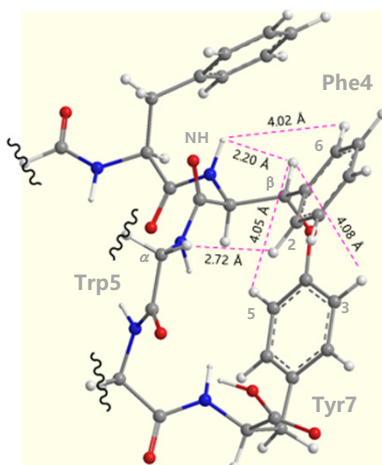
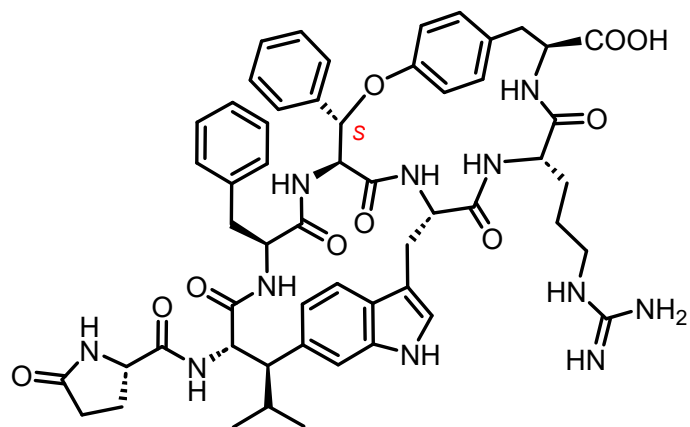


Supplementary Figure 30. Stereochemical analysis of mercurialin. (A) LCMS-based comparison of mercurialin first-ring analyte compared to authentic standard stephanotic acid-[LV] after transient expression of ManBURP-1xQLLVWRY (mercurialin cyclase with truncated substrate domain to include one QLLVWRY core peptide) for 6 days in *N. benthamiana* via expression vector pEAQ-HT and *Agrobacterium tumefaciens* LBA4404 infiltration. Extracted ion chromatogram (EIC) traces of methanolic *N. benthamiana* leaf extract is compared to authentic standard of stephanotic acid-[LV] (100 ng/mL). (B) L-FDAA derivatization of the hydrolysate of mercurialin and amino acid standards. Analyzed by using Phenomenex Kinetex® 2.6 μm C18 reverse phase 100 Å 150 x 3 mm LC column. (C) L-FDAA derivatization of the hydrolysate of mercurialin and amino acid standards. Analyzed by using Higgins Analytical PROTO300 C4 5 μm 250 x 4.6 mm LC column.



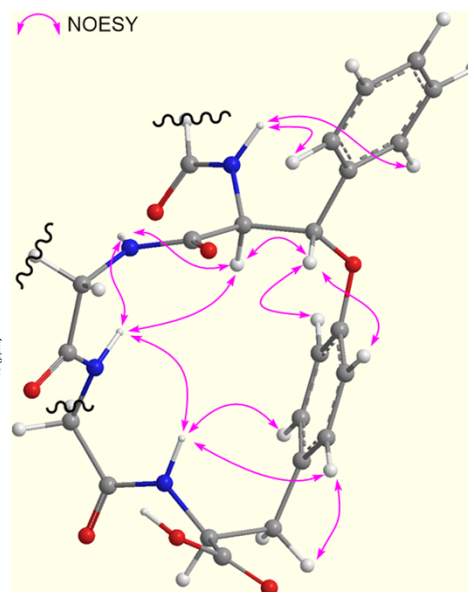
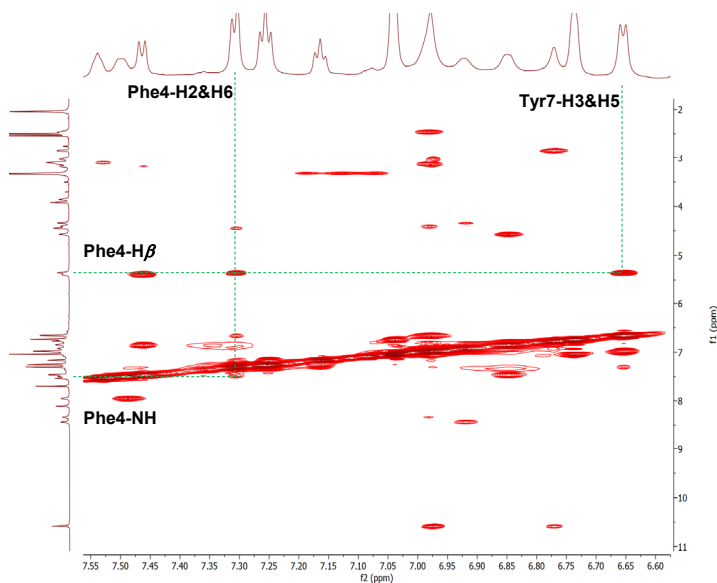
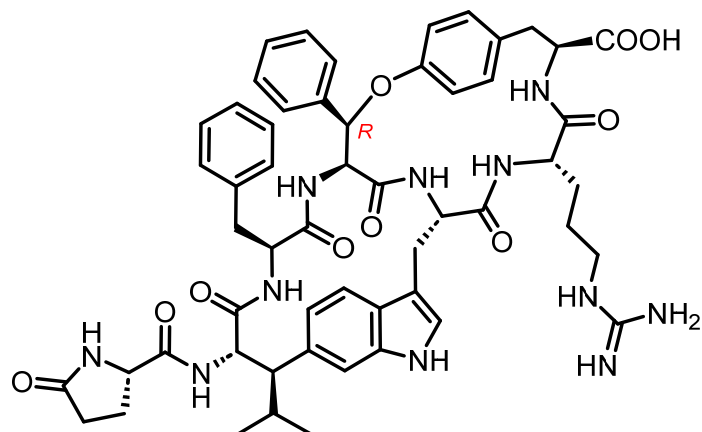
Phe4-C β -S

NOE (Phe⁴-H β -Tyr⁷-H3&H5 and Phe⁴-NH-Phe⁴-H2&H6) detected
NOE (Phe⁴-NH-Phe⁴-H β) not detected

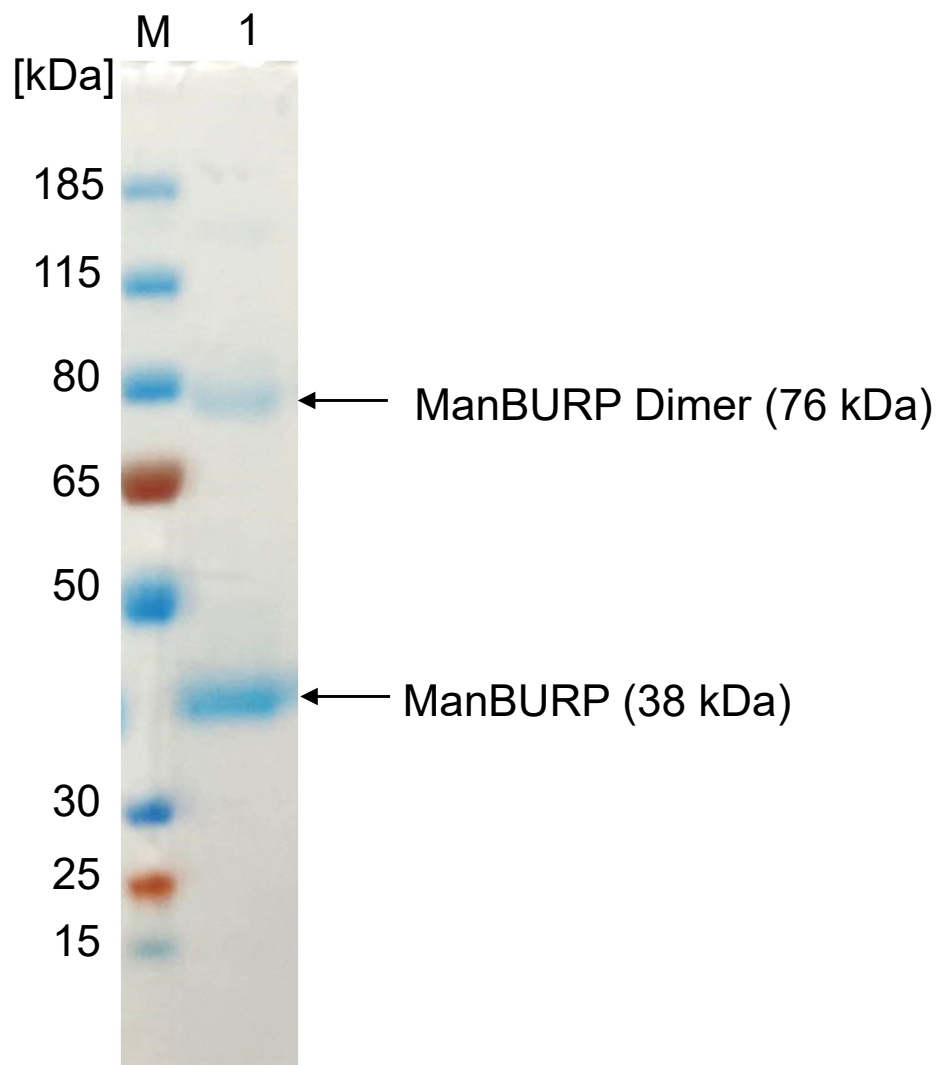


Phe4-C β -R

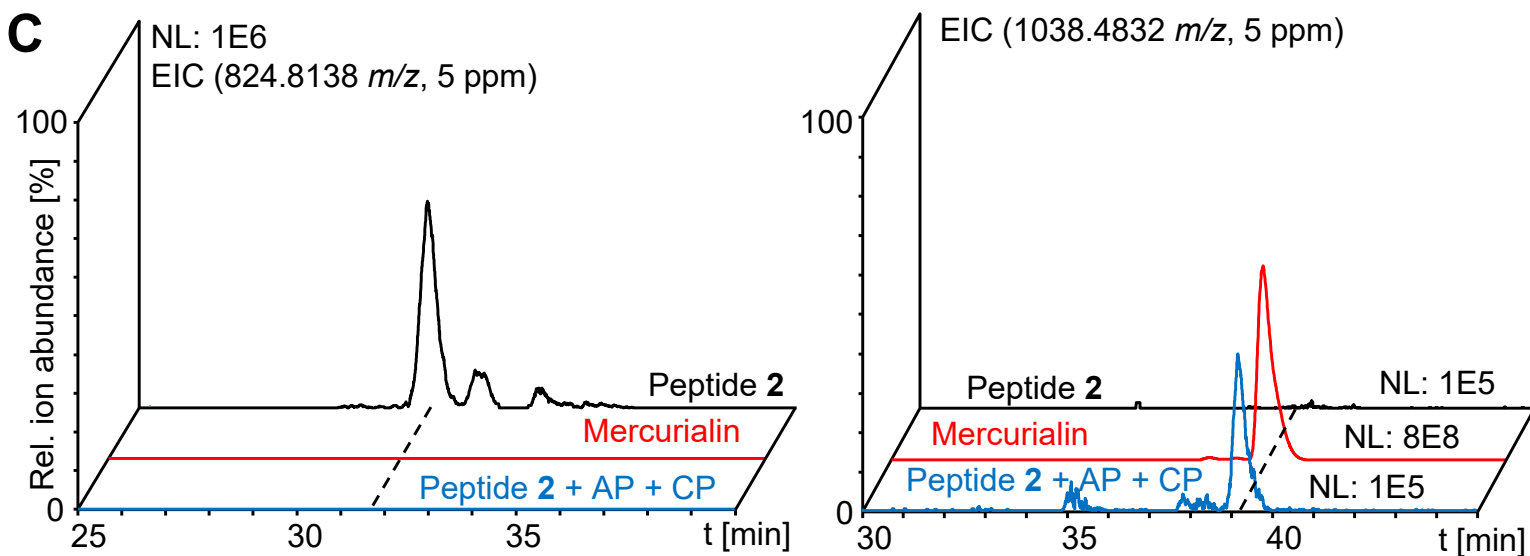
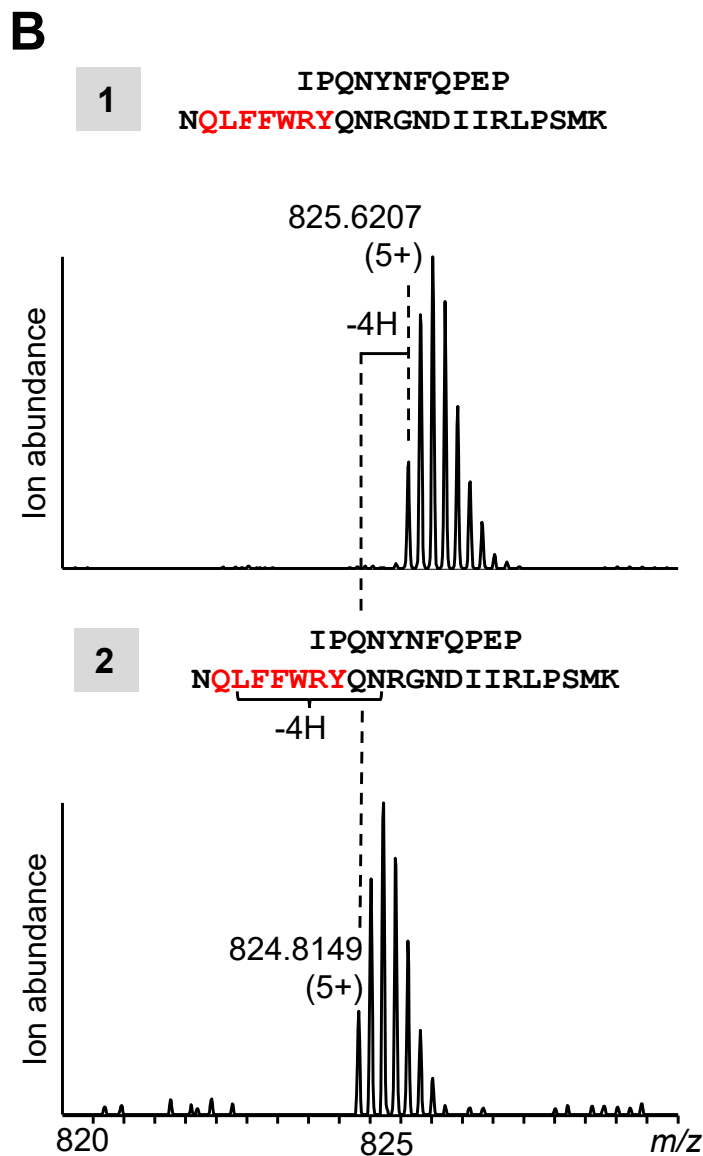
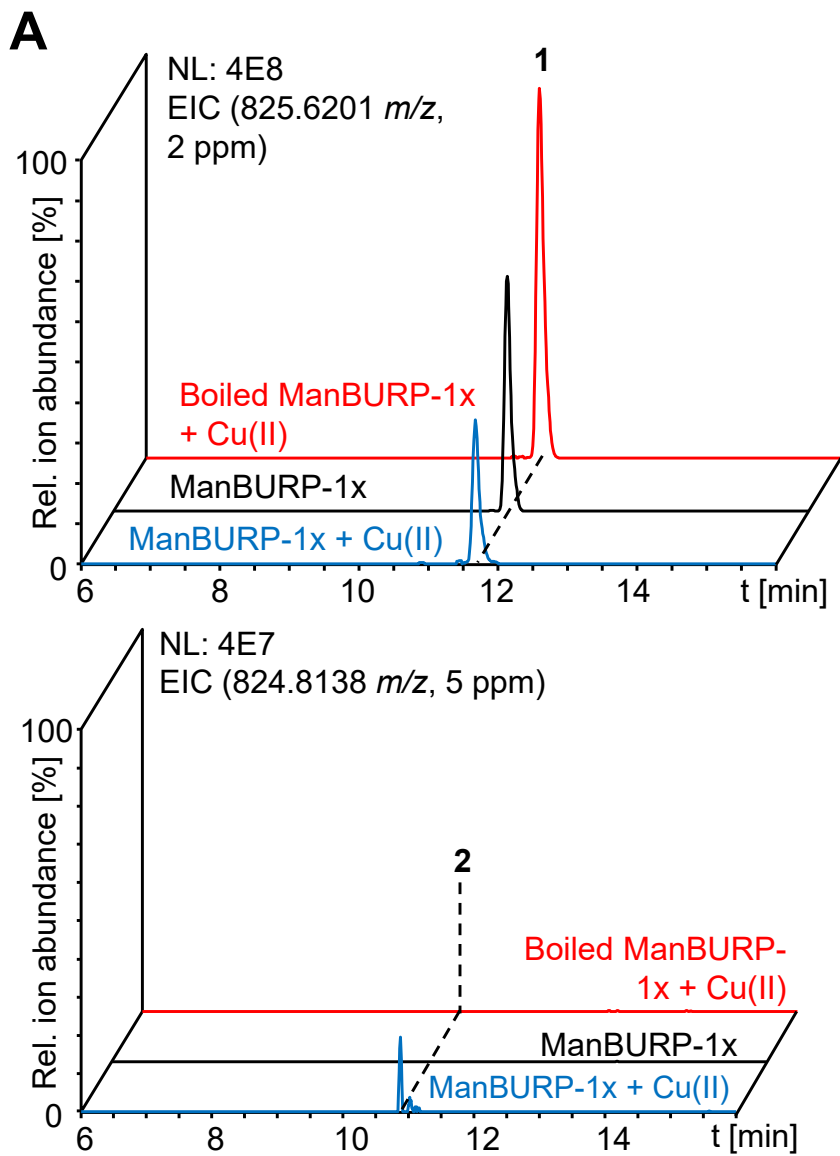
NOE (Phe⁴-H β -Tyr⁷-H3&H5 and Phe⁴-NH-Phe⁴-H2&H6) detected
NOE (Phe⁴-NH-Phe⁴-H β and Trp⁵-H α -Phe⁴-H2) not detected



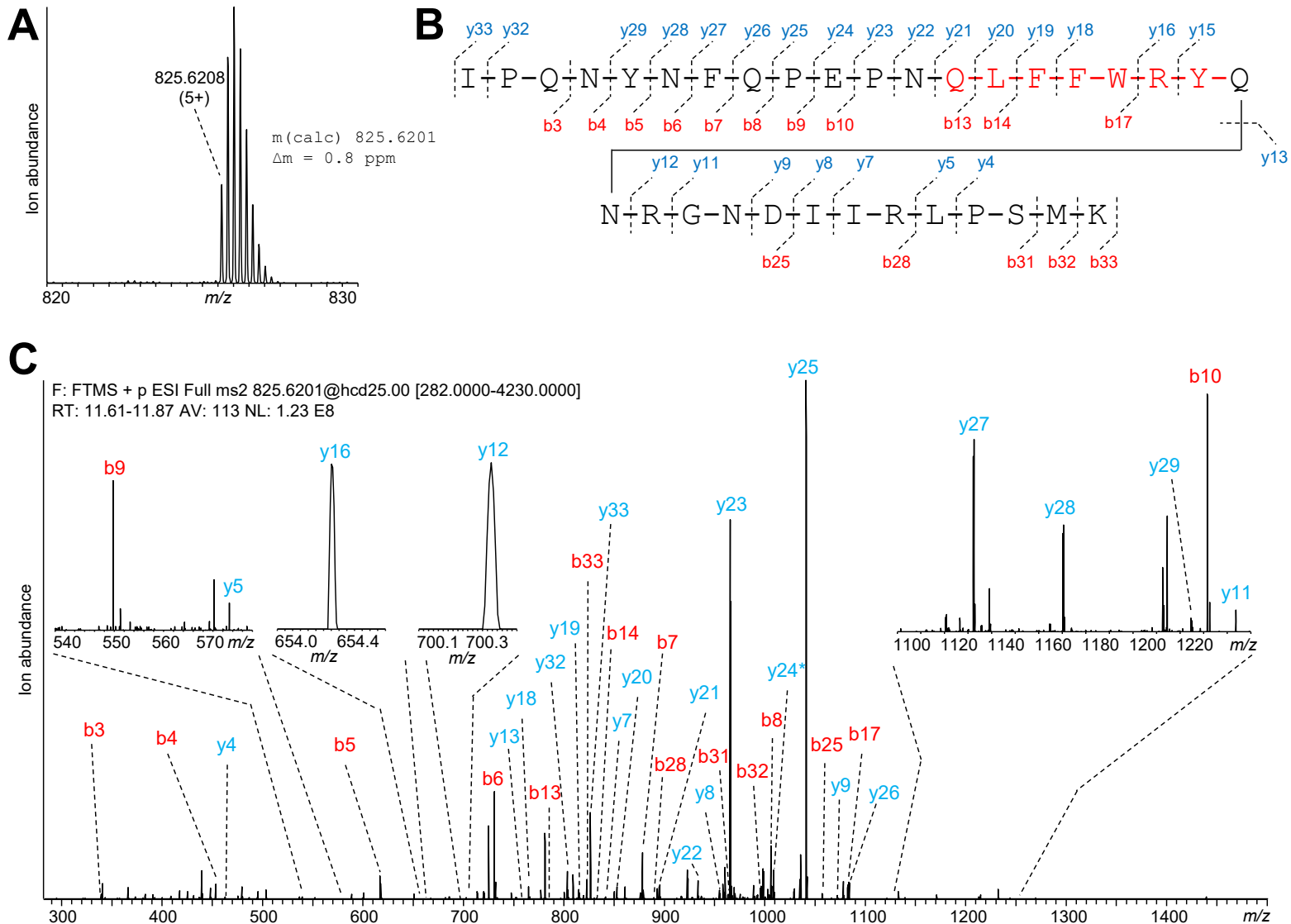
Supplementary Figure 31. Mercurialin geometry optimization and NOESY analysis of Phe4-C β assignment.



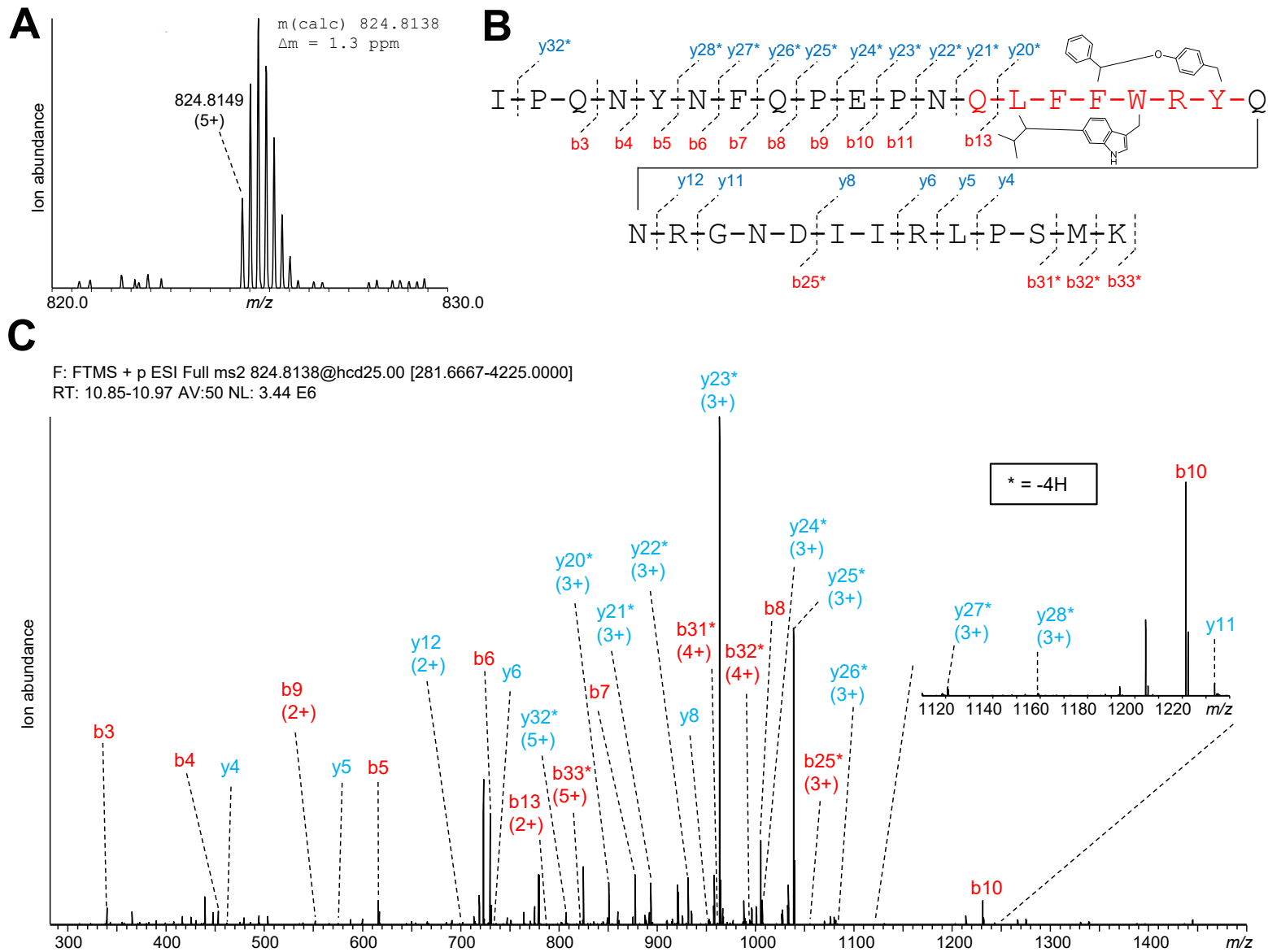
Supplementary Figure 32. Mercurialin cyclase purification. Non-reducing SDS-PAGE of ManBURP-1xQLFFWRY after denature-refolding purification. Source data are provided as a Source Data file.



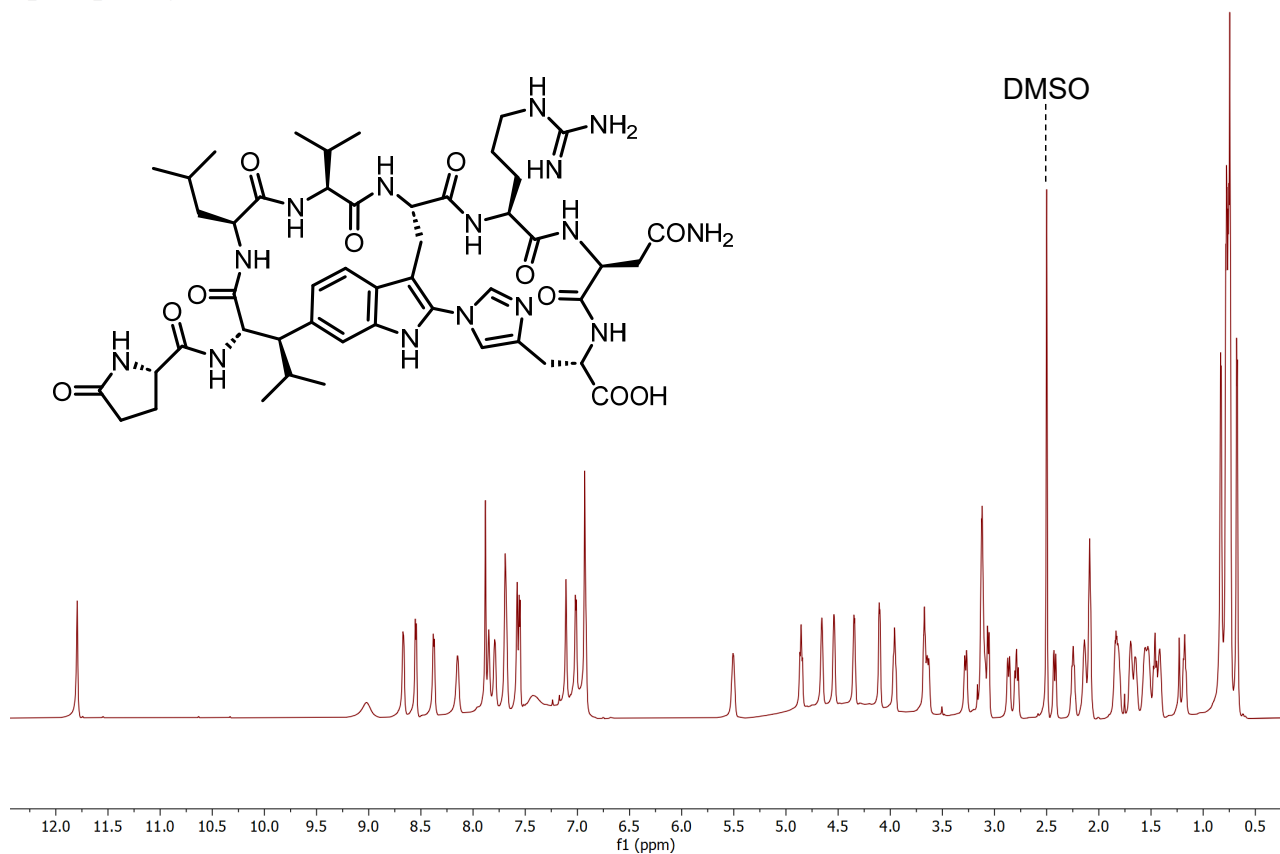
Supplementary Figure 33. LCMS analysis of *in vitro* reconstitution of mercurialin cyclase. Detection of bicyclic core peptide modifications in purified ManBURP-1xQLFFWRY after incubation with Cu(II) at pH 7. (A) EIC traces of linear and modified LysC-digested core peptide IPQNYNFQPEPNQLFFWRYQNRGNDIIRLPSMK. (B) MS spectra of linear and modified LysC-digested core peptide. Core peptide are highlighted in red. (C) Detection of mercurialin after incubation of modified bicyclic LysC-digested core peptide (peptide 2) with aminopeptidase (AP) and carboxypeptidase (CP).



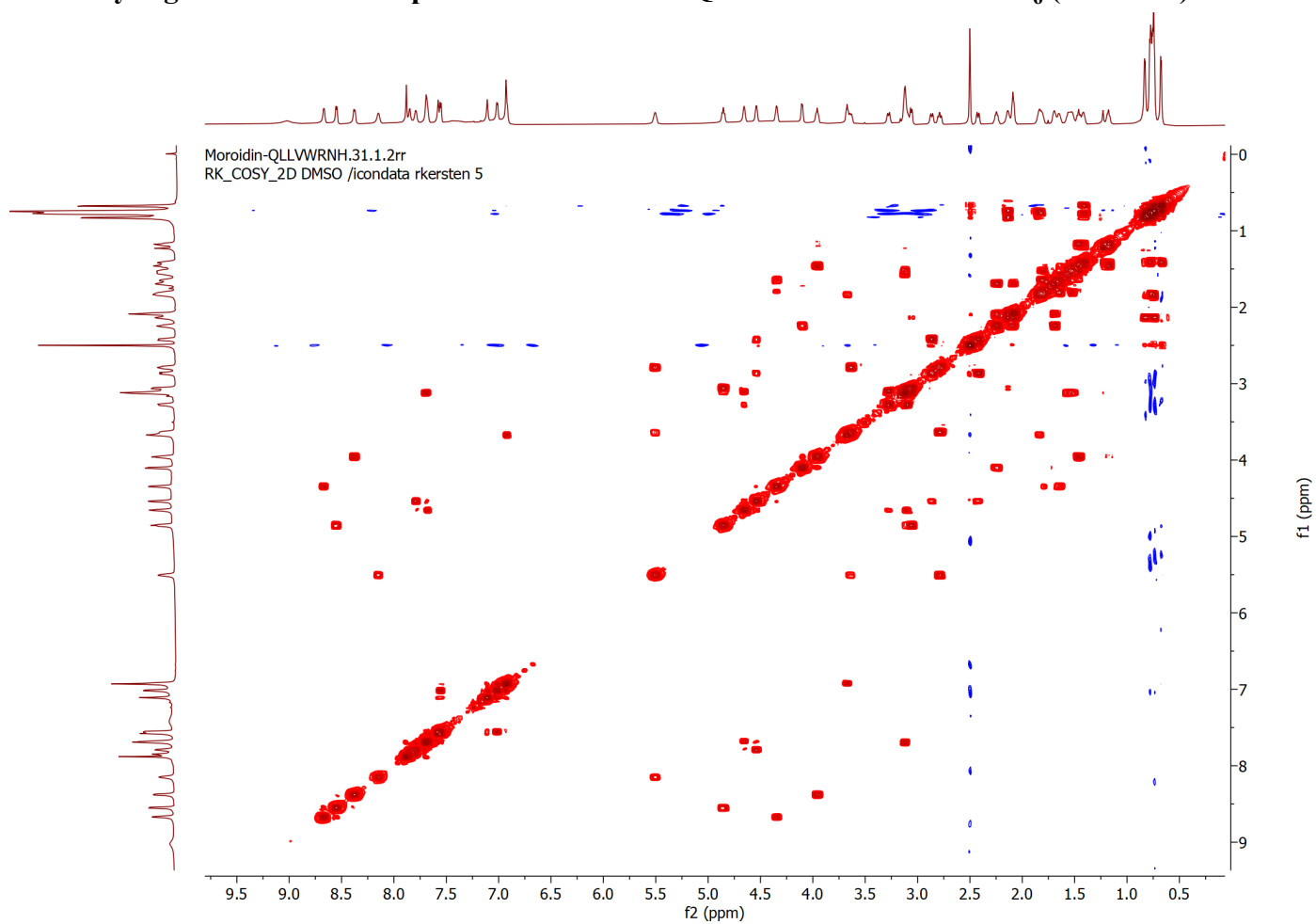
Supplementary Figure 34. MS and MS/MS analysis of unmodified core peptide from *in vitro* reconstitution of mercurialin cyclase. (A) MS1 signal of linear LysC-proteolytic core peptide. (B) Detected tandem MS fragments of linear LysC-proteolytic core peptide. (C) MS/MS analysis of LysC-proteolytic core peptide. Tandem MS fragmentation table of linear LysC-proteolytic core peptide is Supplementary Table 12.



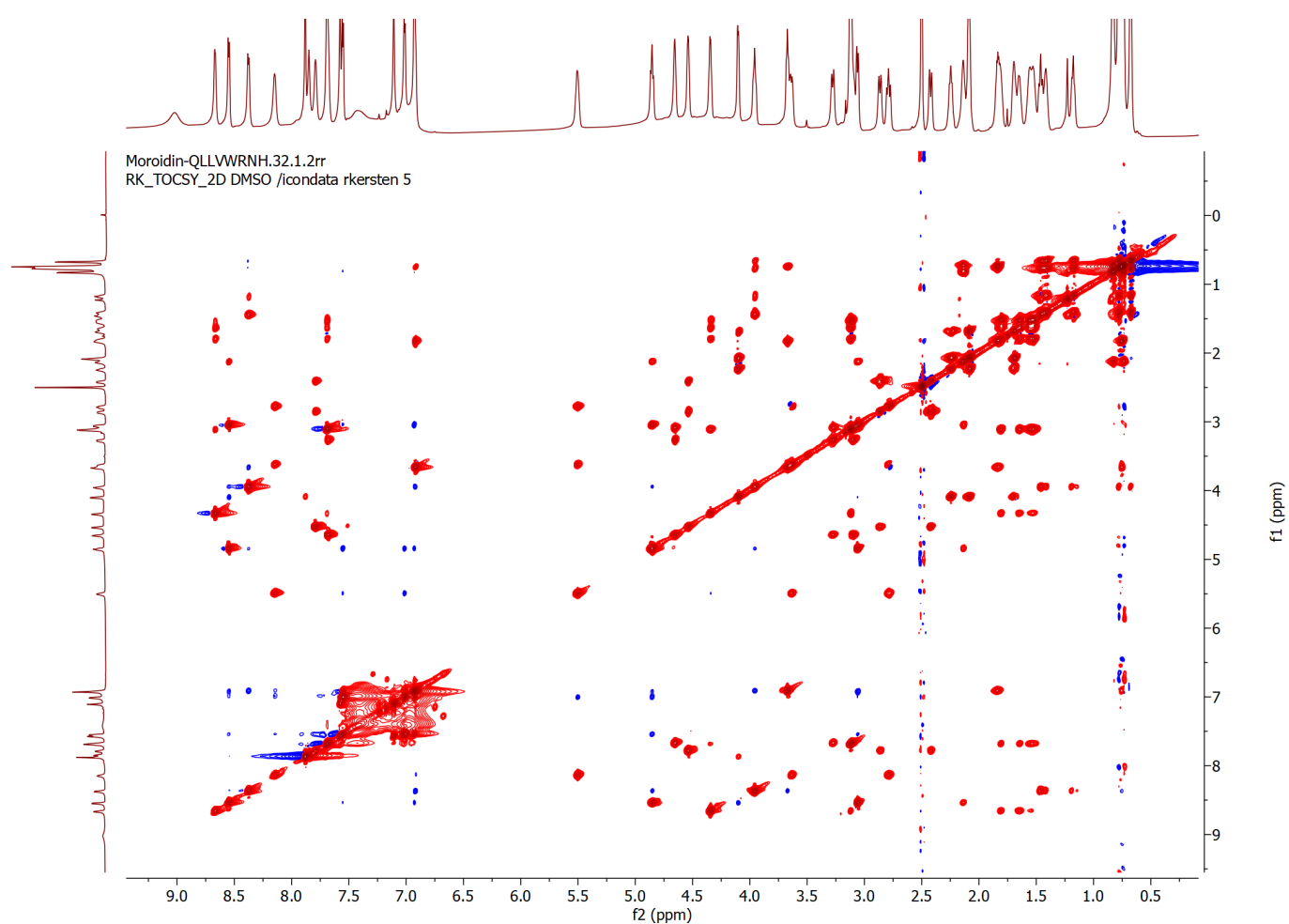
Supplementary Figure 35. MS and MS/MS analysis of modified core peptide from *in vitro* reconstitution of mercurialin cyclase. (A) MS1 signal of putative bicyclic LysC-proteolytic core peptide. (B) Detected tandem MS fragments of putative bicyclic LysC-proteolytic core peptide. (C) MS/MS analysis of putative bicyclic LysC-proteolytic core peptide. Tandem MS fragmentation table of putative bicyclic LysC-proteolytic core peptide is Supplementary Table 13.



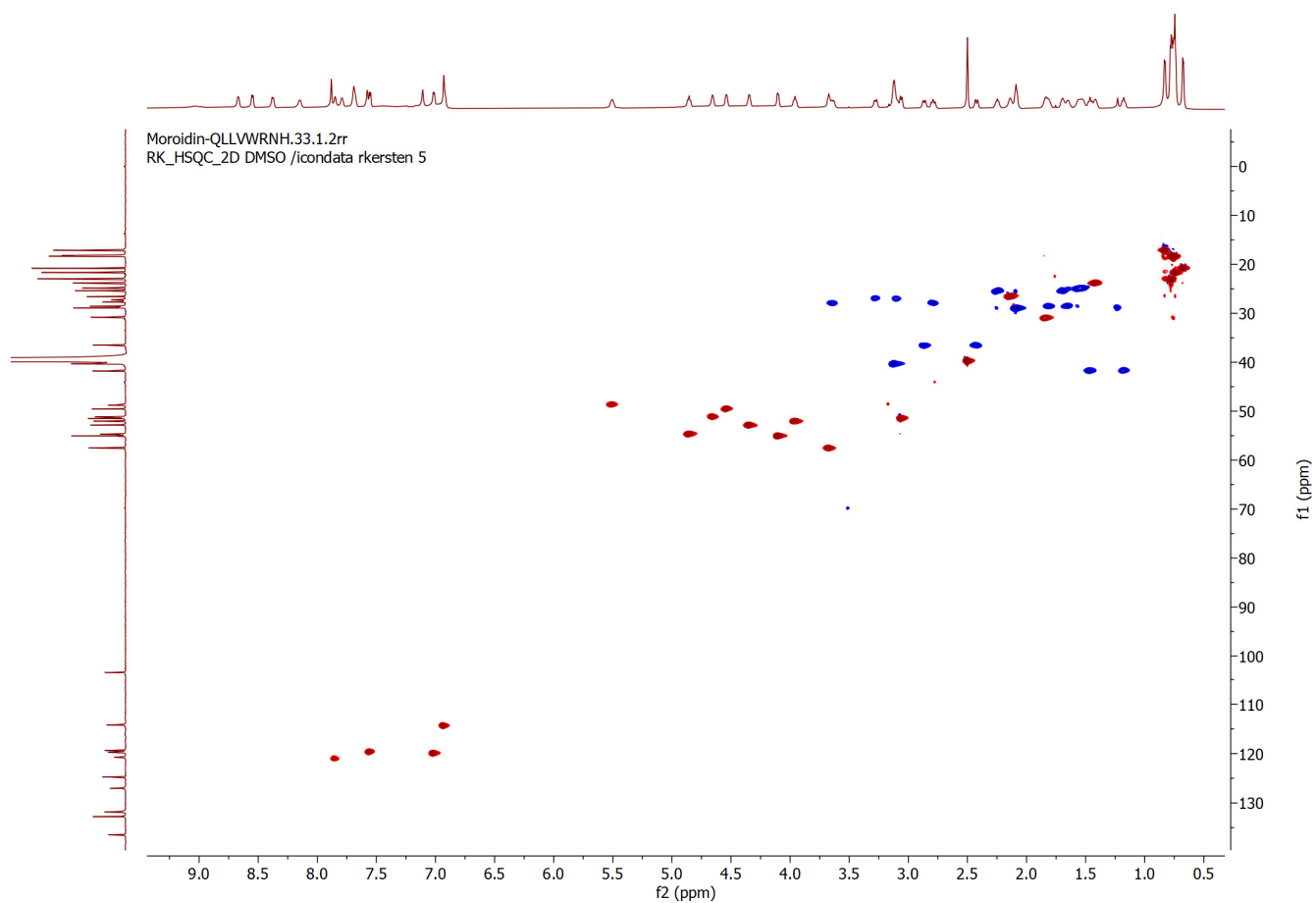
Supplementary Figure 36. ^1H -NMR spectrum of moroidin-QLLVWRNH in DMSO- d_6 (800 MHz).



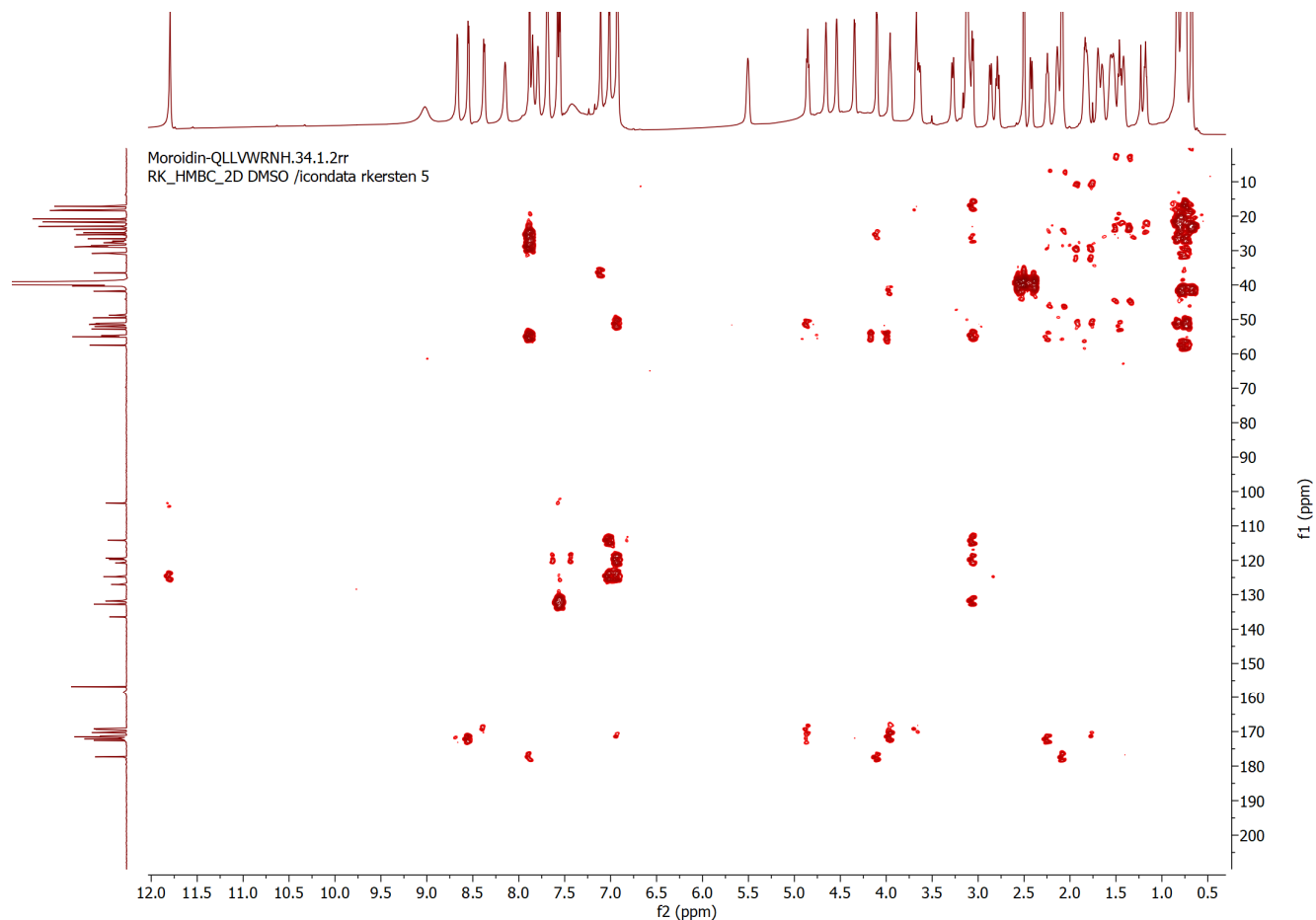
Supplementary Figure 37. ^1H - ^1H -COSY-NMR spectrum of moroidin-QLLVWRNH in DMSO- d_6 .



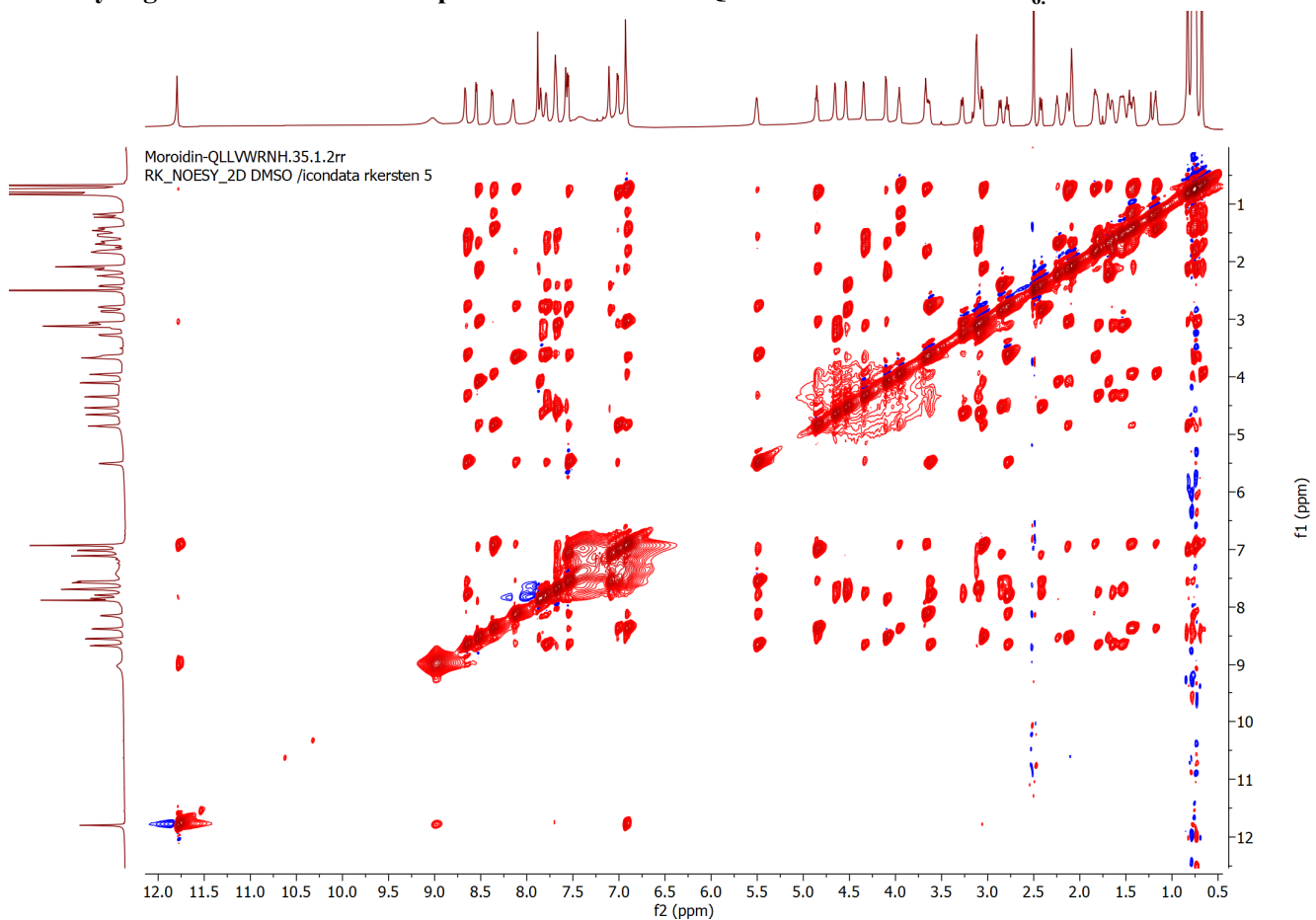
Supplementary Figure 38. ^1H - ^1H -TOCSY-NMR spectrum of moroidin-QLLVWRNH in DMSO-d_6 .



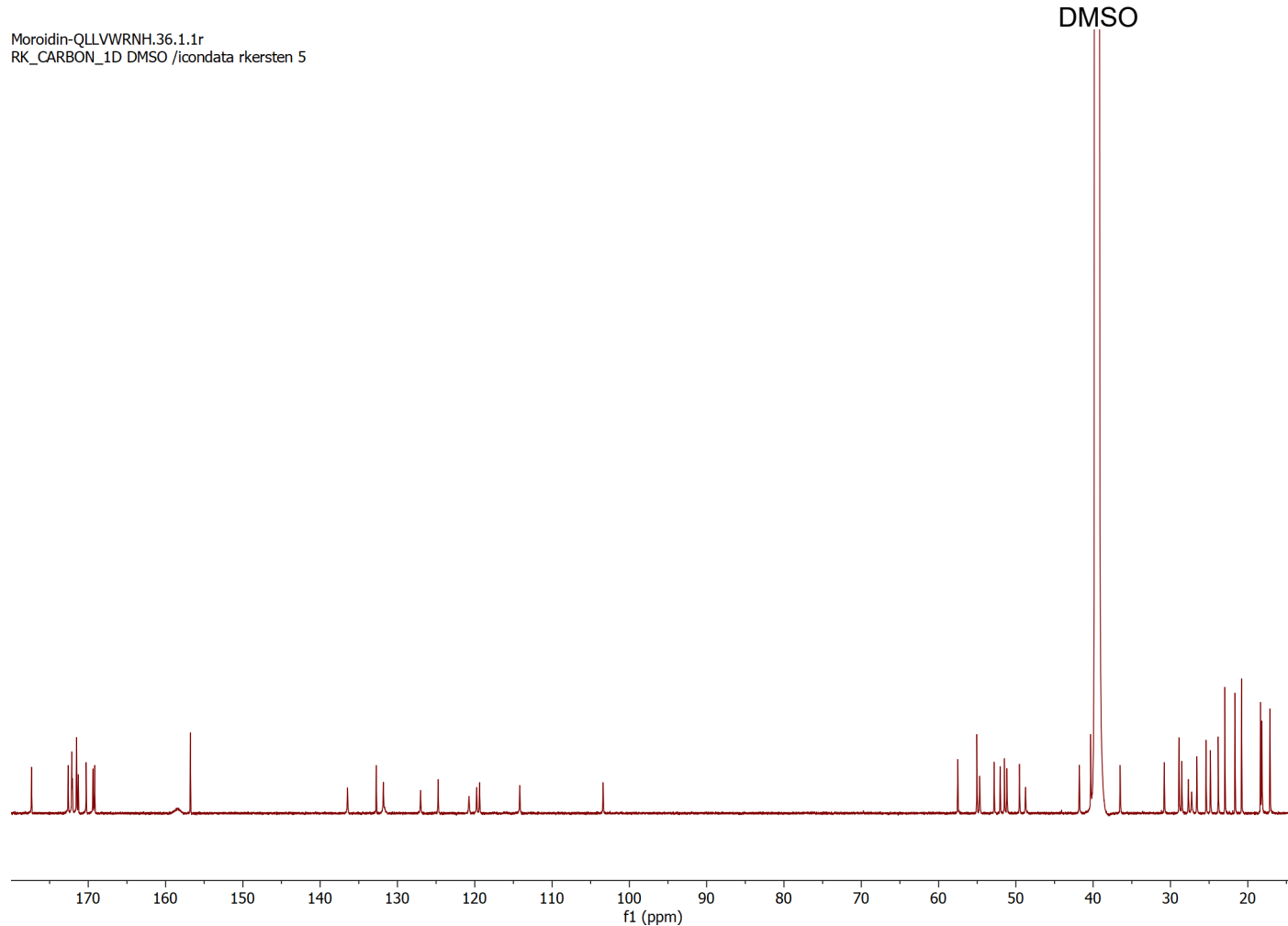
Supplementary Figure 39. HSQC-NMR spectrum of moroidin-QLLVWRNH in DMSO-d_6 .



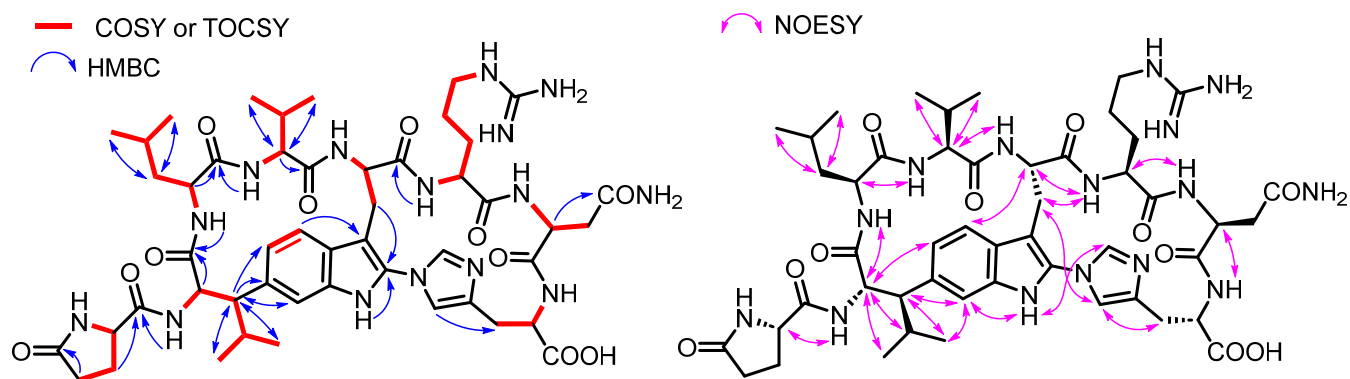
Supplementary Figure 40. HMBC-NMR spectrum of moroidin-QLLVWRNH in DMSO-d₆.



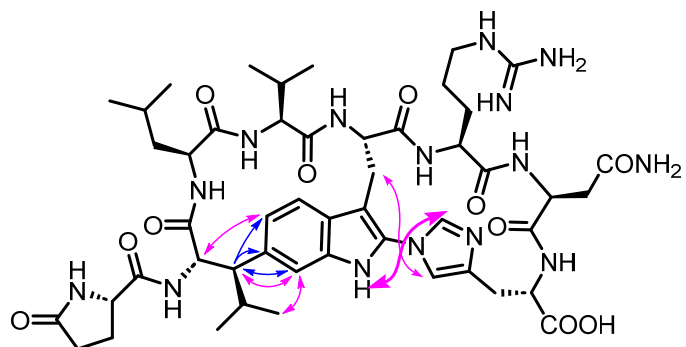
Supplementary Figure 41. NOESY-NMR spectrum of moroidin-QLLVWRNH in DMSO-d₆.



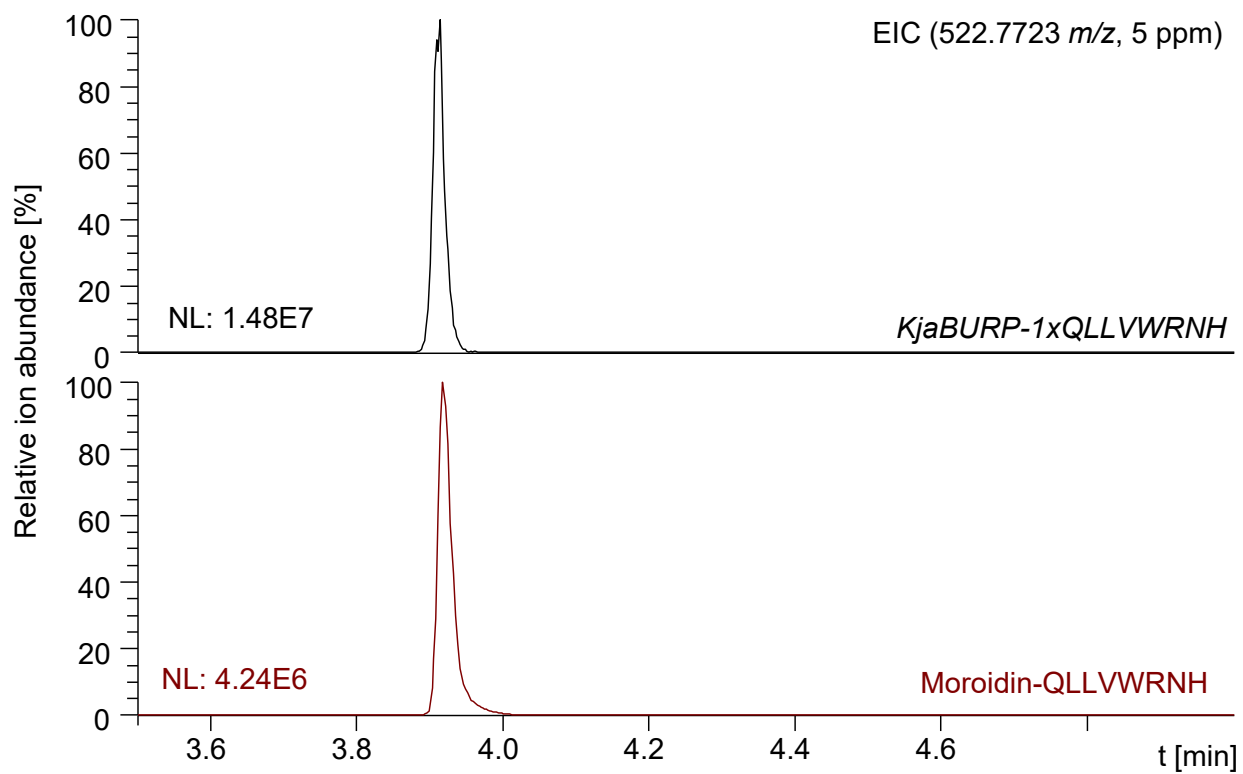
Supplementary Figure 42. ^{13}C -NMR spectrum of moroidin-QLLVWRNH in DMSO-d_6 .
2D NMR correlations in DMSO-d_6



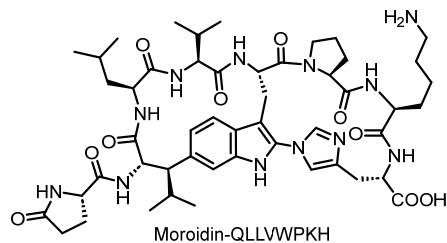
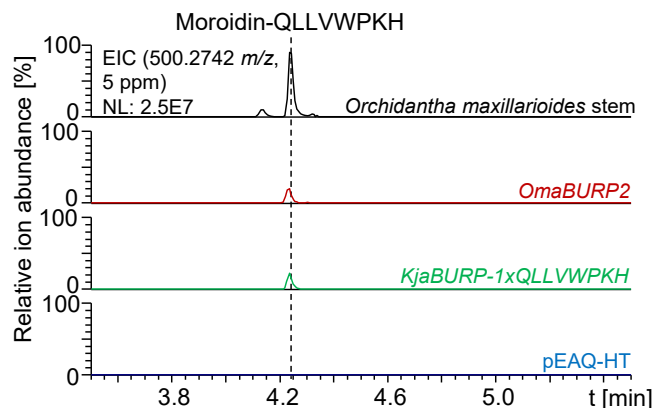
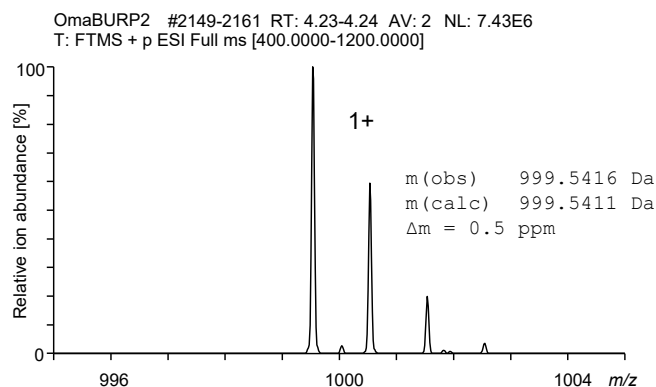
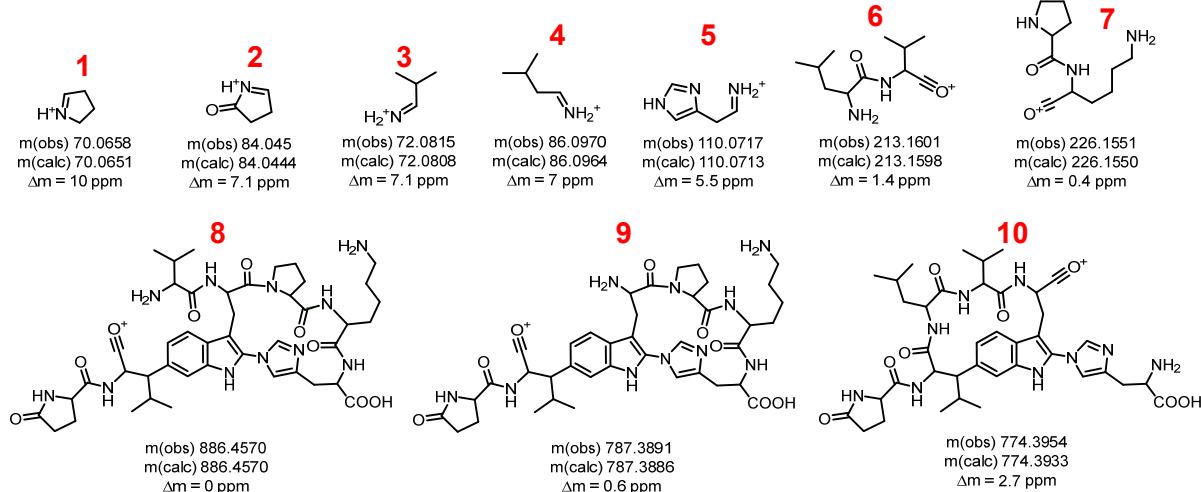
Crosslinks: Key **HMBC** and **NOESY**



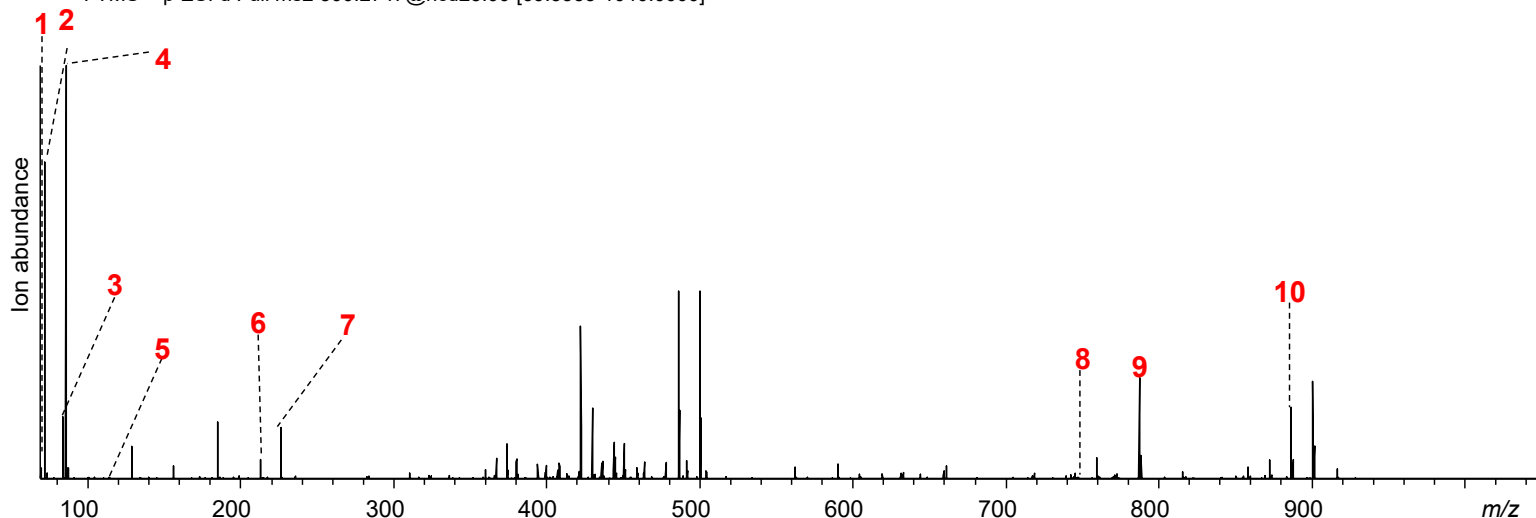
Supplementary Figure 43. Key 2D NMR correlations of moroidin-QLLVWRNH.



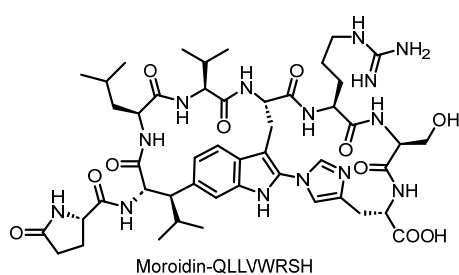
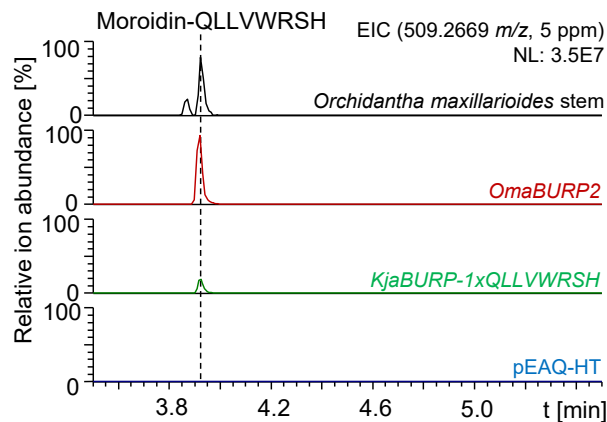
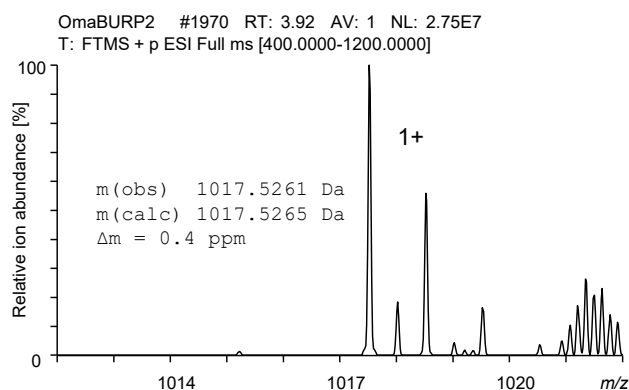
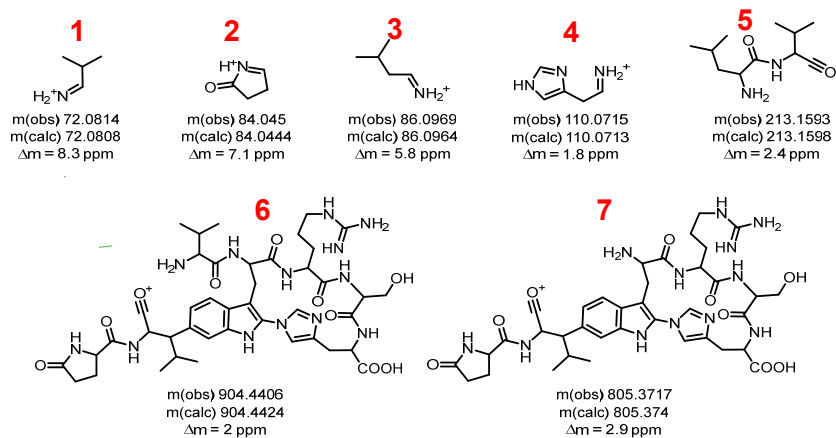
Supplementary Figure 44. Comparison of moroidin-QLLVWRNH standard to moroidin product of KjaBURP-1xQLLVWRNH expressed for 6 days transiently in *N. benthamiana*.

A**B****C****D**

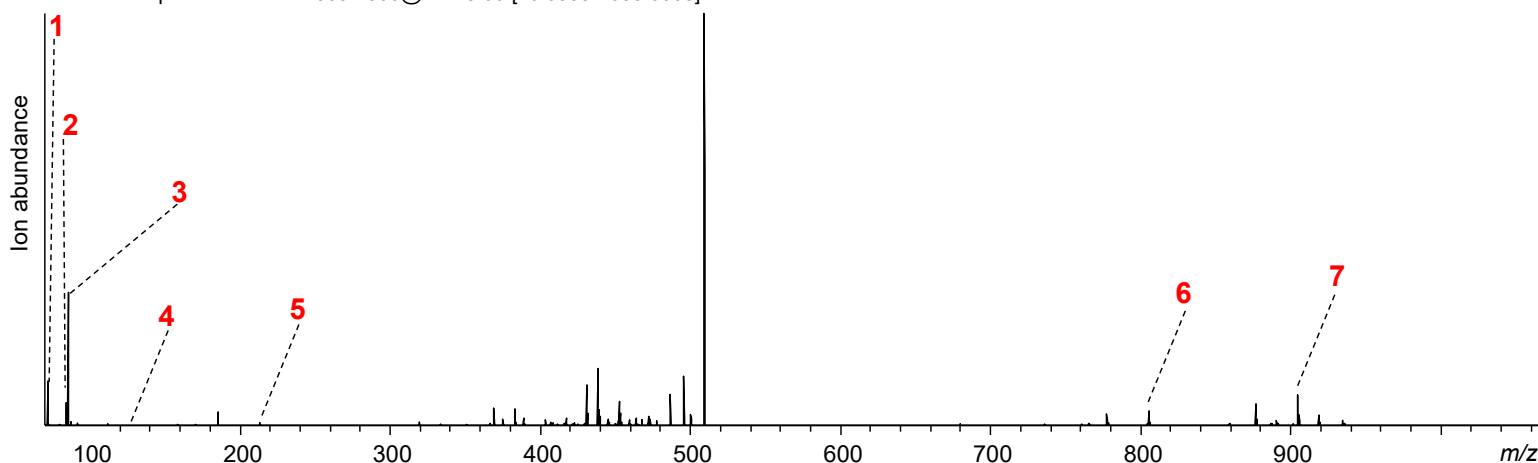
OmaBURP2 #2151-2173 RT: 4.22-4.26 AV: 3 NL: 3.64E6
F: FTMS + p ESI d Full ms2 500.2747@hcd25.00 [69.3333-1040.0000]



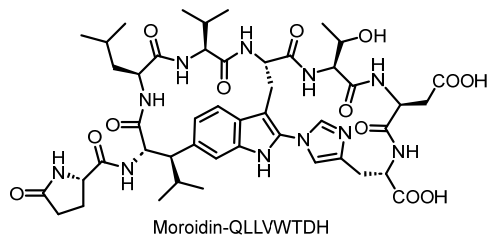
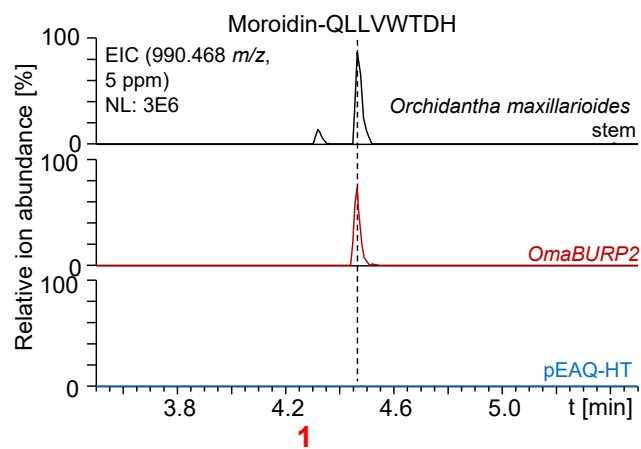
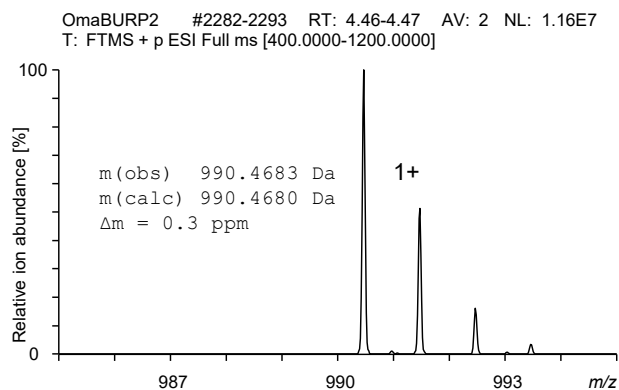
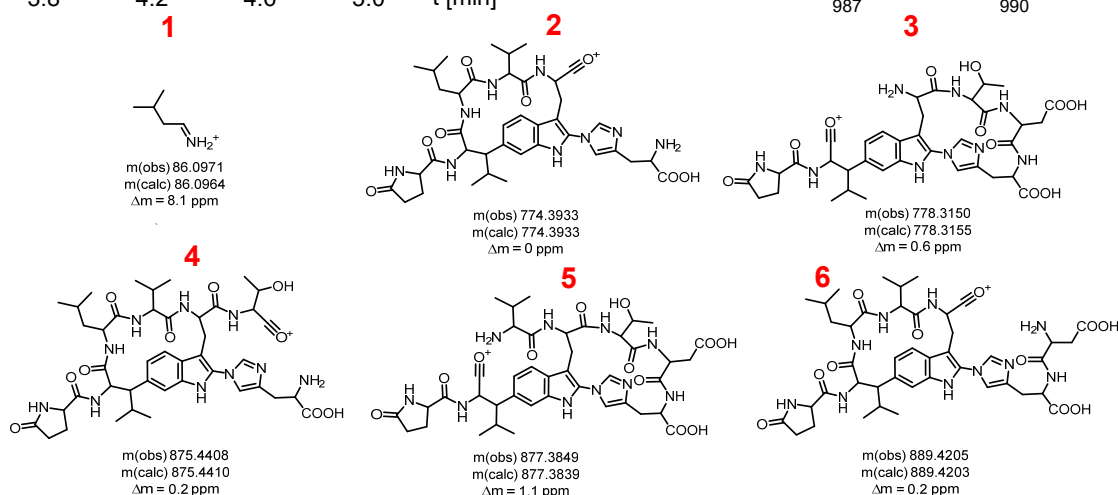
Supplementary Figure 45. Characterization of moroidin-QLLVWPKH biosynthesis in *Orchidantha maxillarioides*. (A) Predicted OmaBURP2 moroidin-QLLVWPKH product. (B) LCMS characterization of predicted moroidin-QLLVWPKH in methanolic extracts of *O. maxillarioides* leaf and transgenic *N. benthamiana* after six days of expression of OmaBURP2. Moroidin-QLLVWPKH was compared to KjaBURP-1xQLLVWPKH moroidin product, derived from methanolic extract of transgenic *N. benthamiana* expressing the KjaBURP construct for 6 days. (C) MS1 spectrum and (D) MS/MS spectrum of moroidin-QLLVWPKH with predicted MS/MS fragment ions from transgenic *N. benthamiana* leaves expressing OmaBURP2.

A**B****C****D**

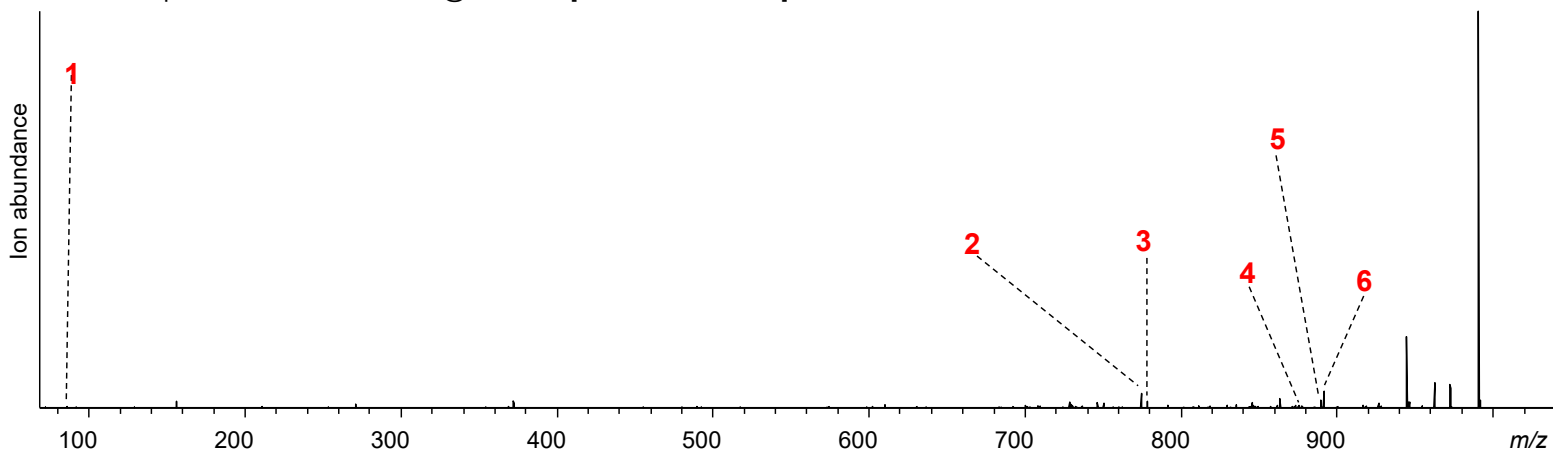
OmaBURP2 #1958-1996 RT: 3.90-3.94 AV: 3 NL: 4.43E7
F: FTMS + p ESI d Full ms2 509.2366@hcd25.00 [70.3333-1055.0000]



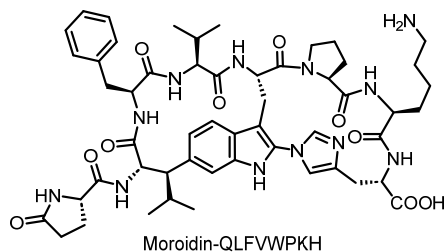
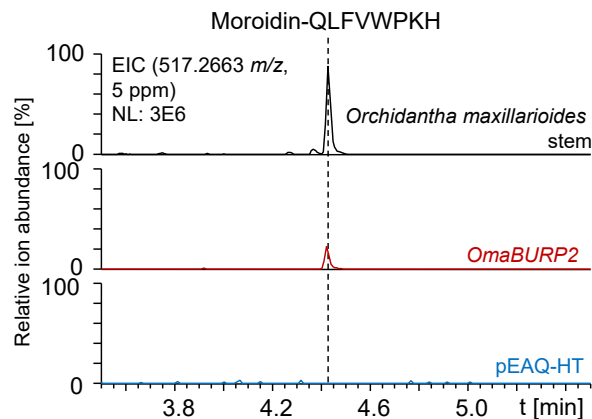
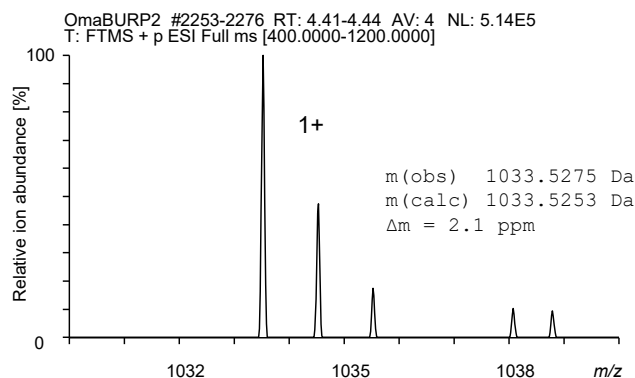
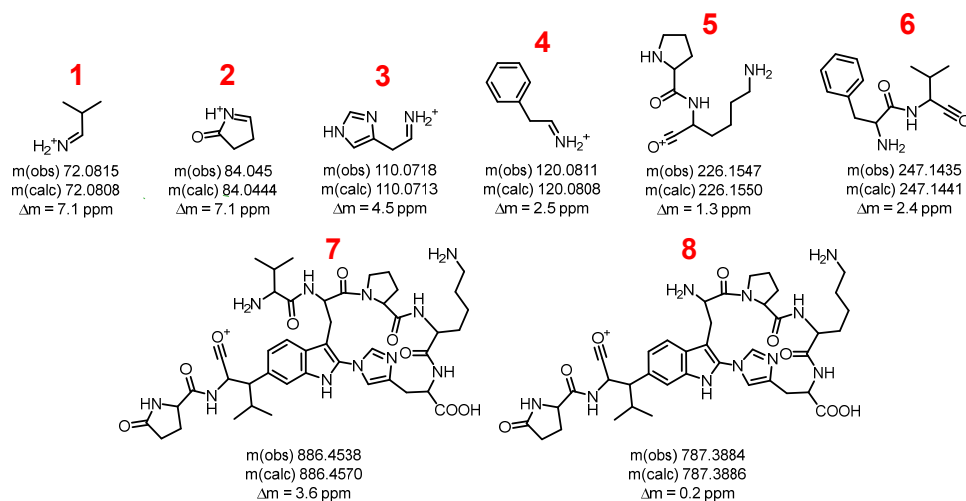
Supplementary Figure 46. Characterization of moroidin-QLLVWRSH biosynthesis in *Orchidantha maxillarioides*. (A) Predicted OmaBURP2 moroidin-QLLVWRSH product. (B) LCMS characterization of predicted moroidin-QLLVWRSH in methanolic extracts of *O. maxillarioides* leaf and transgenic *N. benthamiana* after six days of expression of OmaBURP2. Moroidin-QLLVWRSH was compared to KjaBURP-1xQLLVWRSH moroidin product, derived from methanolic extract of transgenic *N. benthamiana* expressing the KjaBURP construct for 6 days. (C) MS1 spectrum and (D) MS/MS spectrum of predicted moroidin-QLLVWRSH with predicted MS/MS fragment ions from transgenic *N. benthamiana* leaves expressing OmaBURP2.

A**B****C****D**

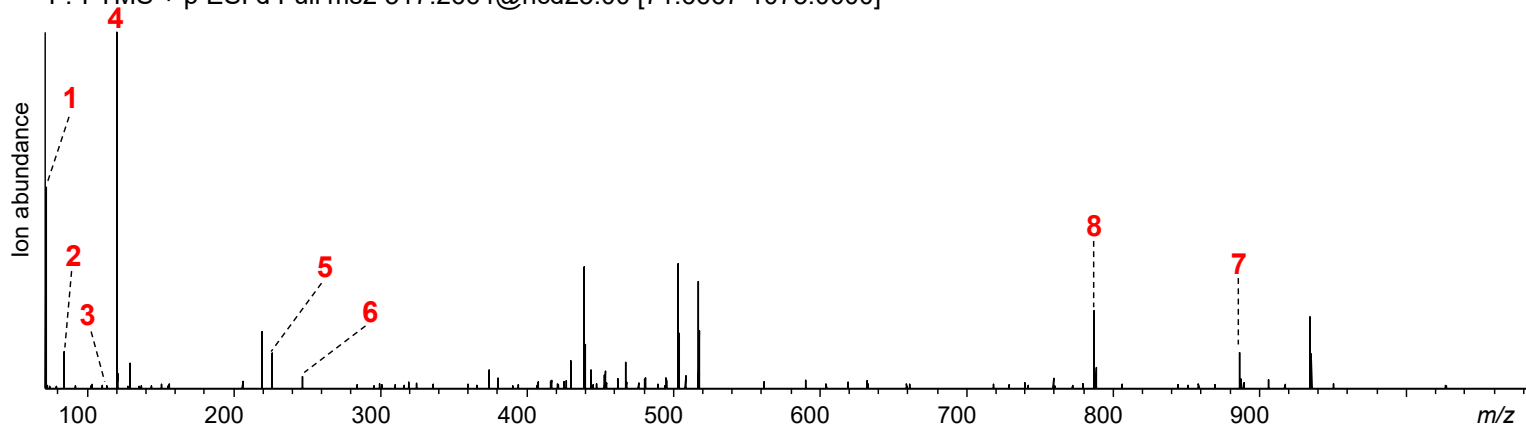
OmaBURP2 #2231-2361 RT: 4.45-4.47 AV: 2 NL: 4.96E6
F: FTMS + p ESI d Full ms2 990.4683@hcd25.00 [68.6667-1030.0000]



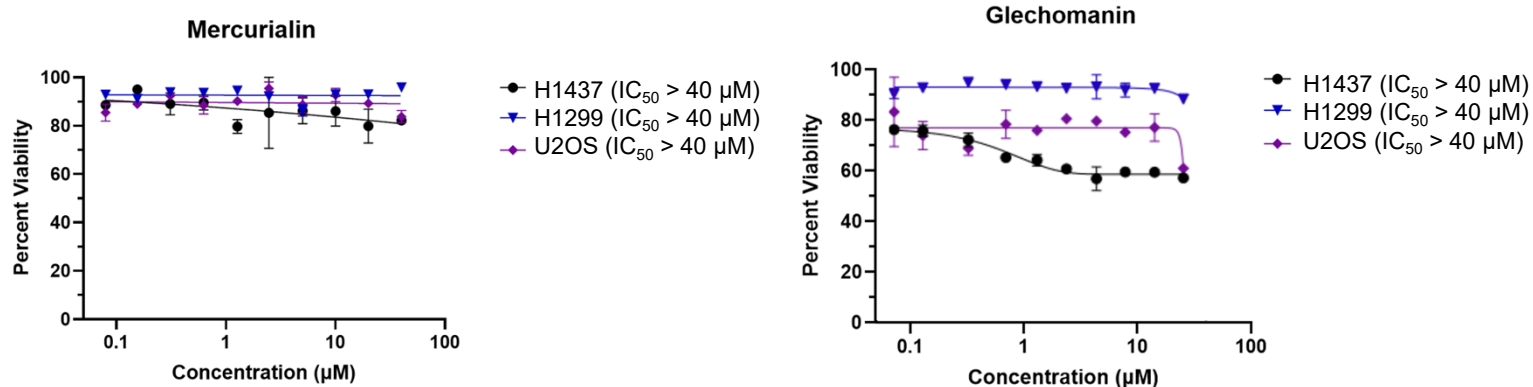
Supplementary Figure 47. Characterization of moroidin-QLLVWTDH biosynthesis in *Orchidantha maxillarioides*. (A) Predicted OmaBURP2 moroidin-QLLVWTDH product. (B) LCMS characterization of predicted moroidin-QLLVWTDH in methanolic extracts of *O. maxillarioides* leaf and transgenic *N. benthamiana* after six days of expression of OmaBURP2. Moroidin-QLLVWTDH was compared to KjaBURP-1xQLLVWTDH moroidin product, derived from methanolic extract of transgenic *N. benthamiana* expressing the KjaBURP construct for 6 days. (C) MS1 spectrum and (D) MS/MS spectrum of moroidin-QLLVWTDH with predicted MS/MS fragment ions from transgenic *N. benthamiana* leaves expressing OmaBURP2.

A**B****C****D**

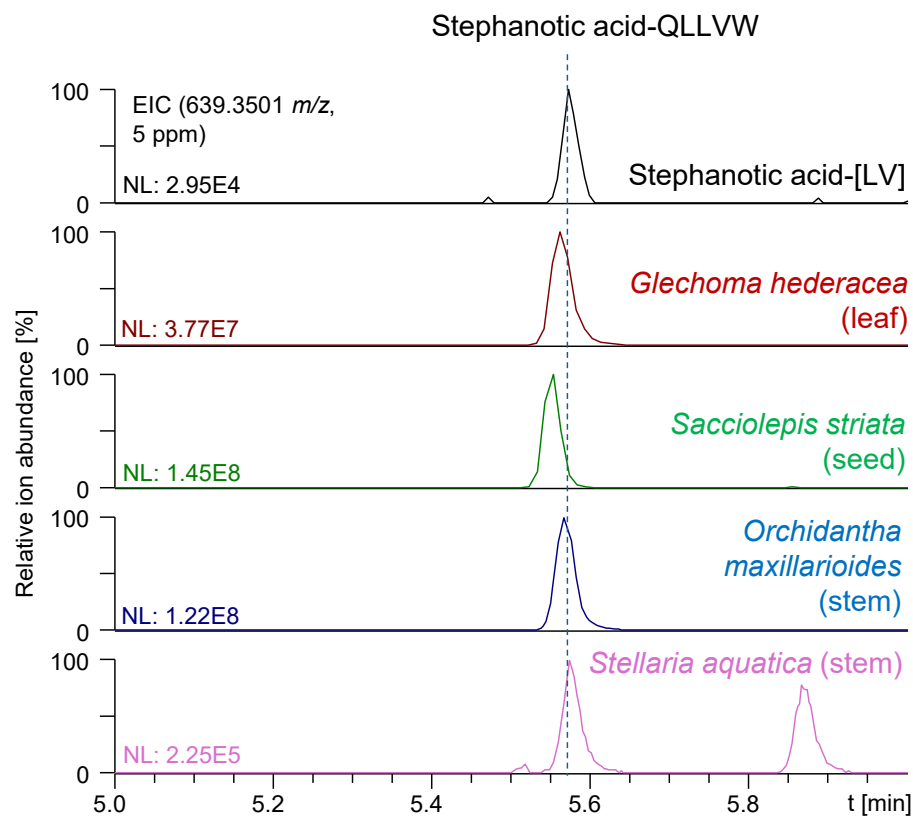
Orchidantha_maxillarioides_stem #1966-2200 RT: 4.42-4.44 AV: 2 NL: 3.78E5
F: FTMS + p ESI d Full ms2 517.2664@hcd25.00 [71.6667-1075.0000]



Supplementary Figure 48. Characterization of moroidin-QLFVWPKH biosynthesis in *Orchidantha maxillarioides*. (A) Predicted OmaBURP2 moroidin-QLFVWPKH product. (B) LCMS characterization of predicted moroidin-QLFVWPKH in methanolic extracts of *O. maxillarioides* leaf and transgenic *N. benthamiana* after six days of expression of OmaBURP2. Moroidin-QLFVWPKH was compared to KjaBURP-1x QLFVWPKH moroidin product, derived from methanolic extract of transgenic *N. benthamiana* expressing the KjaBURP construct for 6 days. (C) MS1 spectrum and (D) MS/MS spectrum of moroidin-QLFVWPKH with predicted MS/MS fragment ions from transgenic *N. benthamiana* leaves expressing OmaBURP2 and *O. maxillarioides* stem extract, respectively.

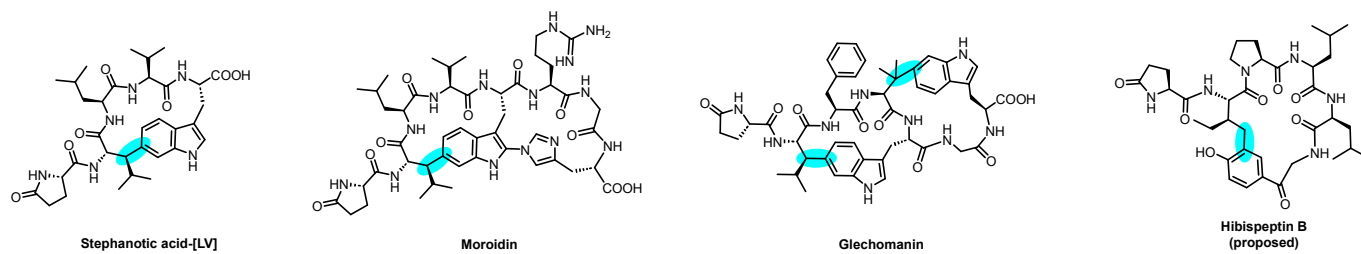


Supplementary Figure 49. *In vitro* anticancer assays of stephanotic acid-type peptides. The dose-response cytotoxicity of stephanotic acid-type peptides mercurialin and glechomanin was evaluated in three distinct cancer cell lines: H1437 (non-small cell lung adenocarcinoma), H1299 (non-small cell lung cancer), and U2OS (osteosarcoma). For all assays, the normalized cell viability was assessed across ten-point, two-fold serial dilutions, starting from a top concentration of 40 μM , in triplicate. Data points represent mean cell viability ($n=3$), with error bars denoting one standard error of the mean.

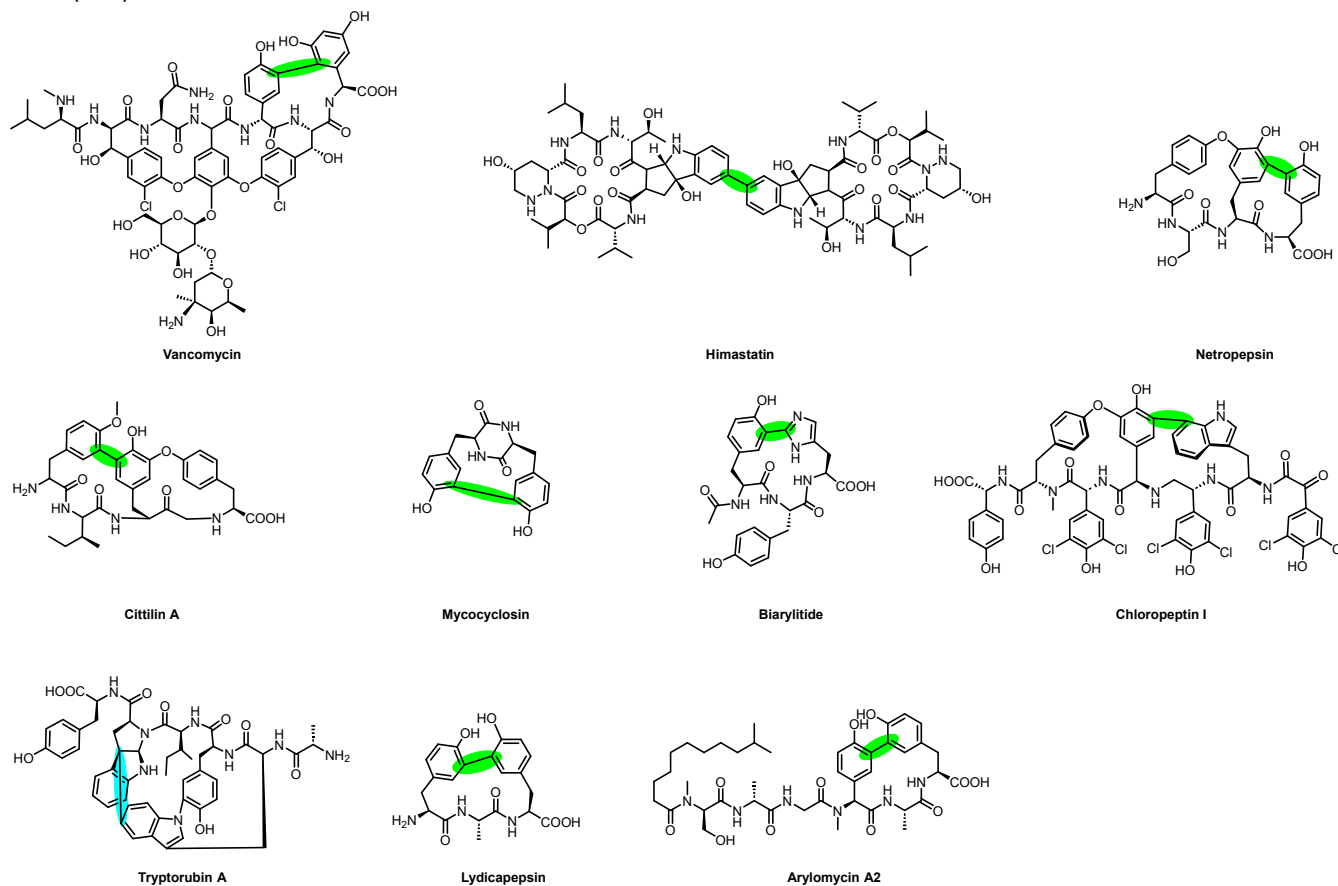


Supplementary Figure 50. Characterization of stephanotic acid-QLLVW in pinks, mints, grasses, and Lowiaceae.

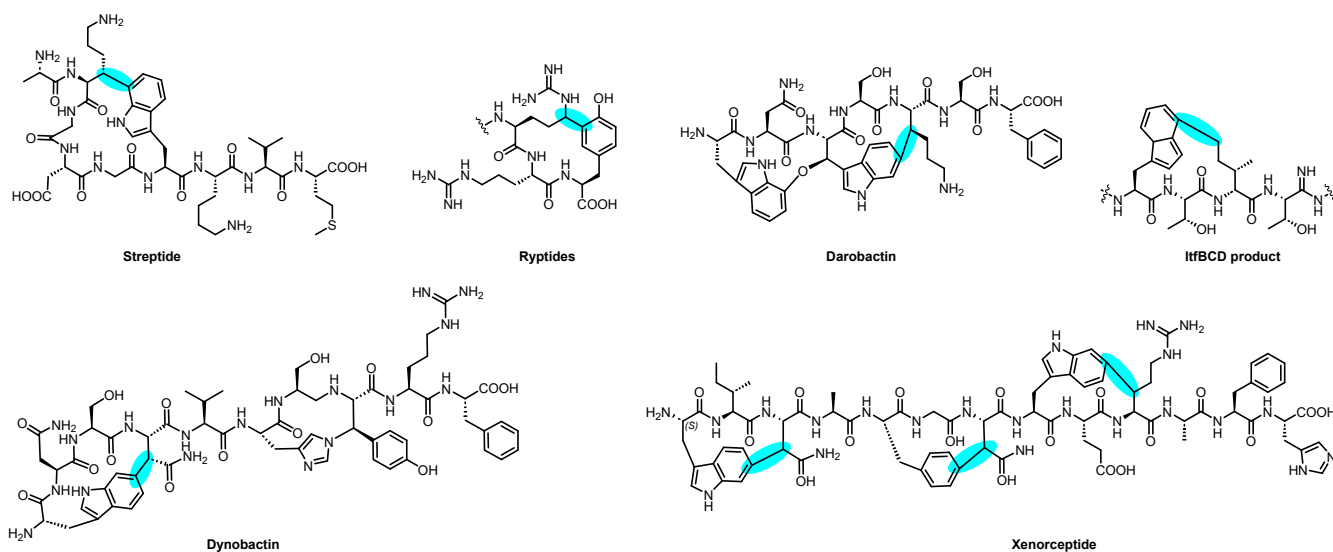
Plants (BURP domain)



Bacteria (P450)

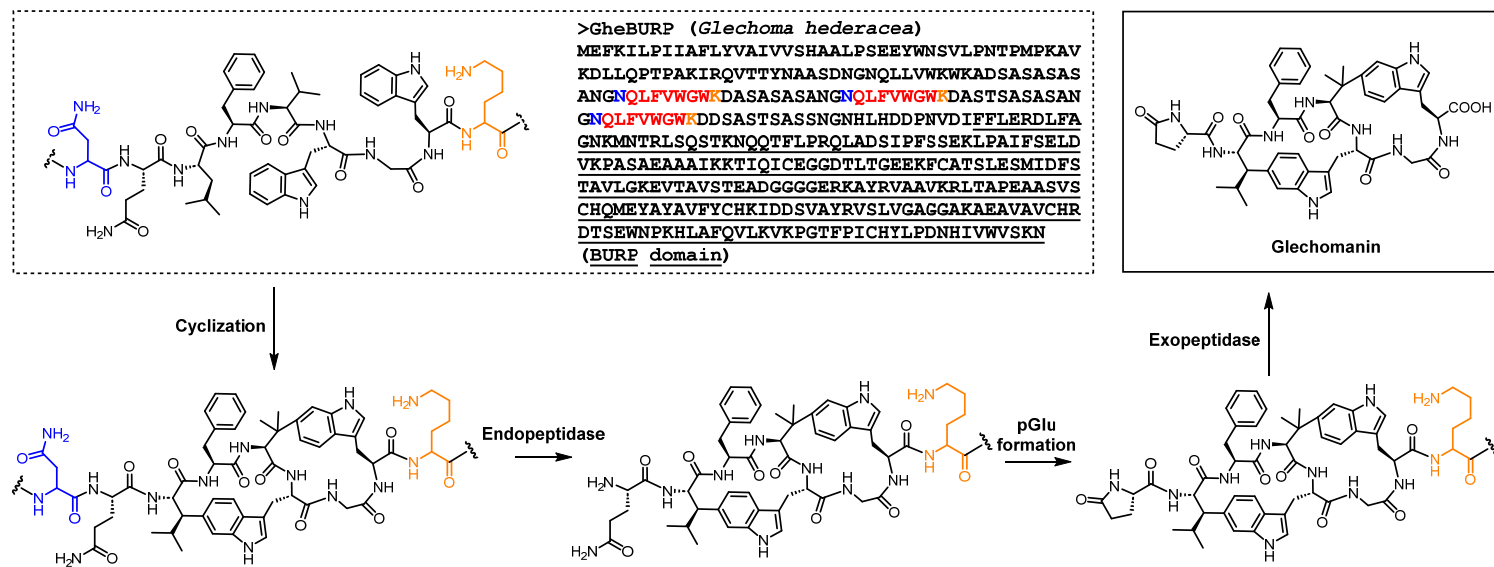


Bacteria (rSAM enzyme)

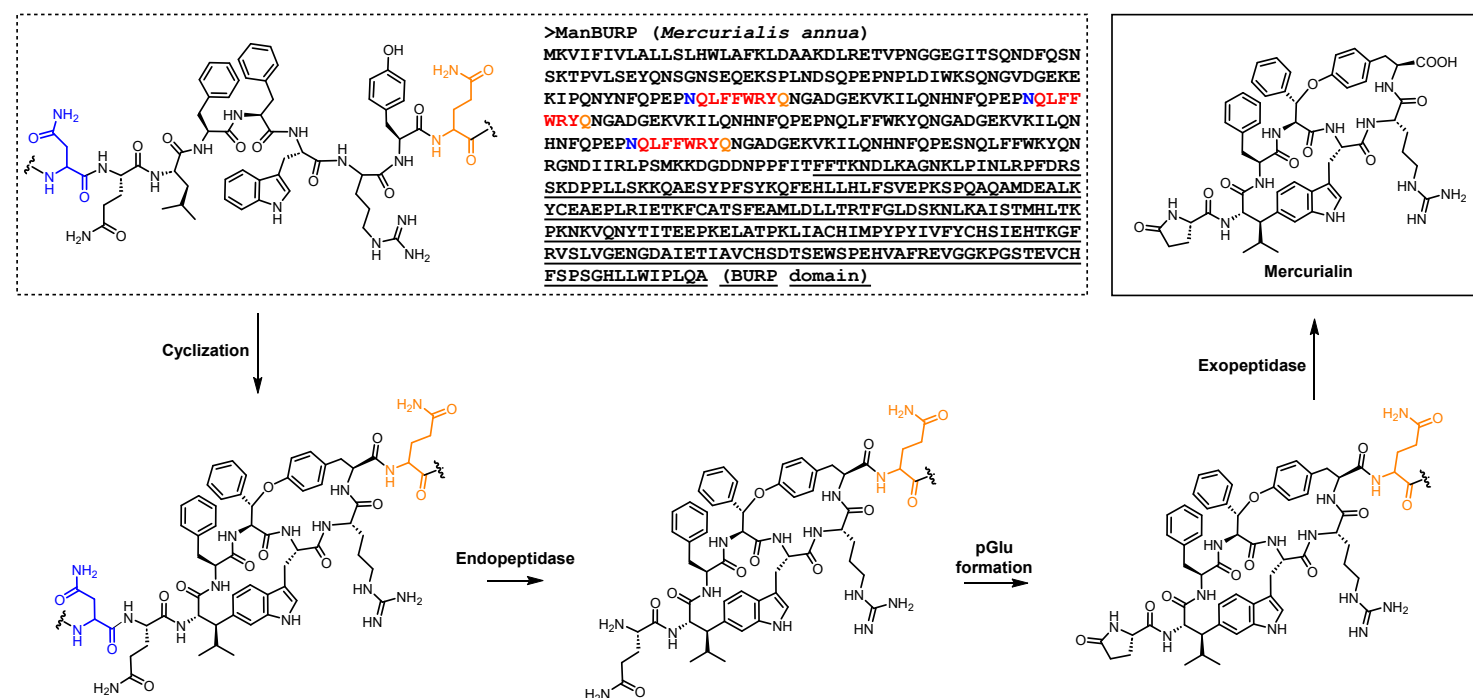


Supplementary Figure 51. Representative peptide natural products with C-C-macrocyclic crosslinks. Macrocyclic bonds are highlighted in blue ($C(sp^3)-C(sp^2)$) and green ($C(sp^2)-C(sp^2)$), respectively.

A



B



Supplementary Figure 52. Biosynthetic proposals of burpitides from this study. (A) Glechomanin biosynthesis in *Glechoma hederacea* and (B) Mercurialin biosynthetic proposal in *Mercurialis annua*.

Supplementary references

1. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
2. Sainsbury, F., Thuenemann, E. C. & Lomonossoff, G. P. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol. J.* **7**, 682–693 (2009).
3. Kersten, R. D. et al. Gene-Guided Discovery and Ribosomal Biosynthesis of Moroidin Peptides. *J. Am. Chem. Soc.* **144**, 7686–7692 (2022).
4. One Thousand Plant Transcriptomes Initiative. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**, 679–685 (2019).
5. Kriger, D., Pasquale, M. A., Ampolini, B. G. & Chekan, J. R. Mining raw plant transcriptomic data for new cyclopeptide alkaloids. *Beilstein J. Org. Chem.* **20**, 1548–1559 (2024).
6. Gomes, P. W. P. et al. plantMASST - Community-driven chemotaxonomic digitization of plants. *bioRxiv* (2024) doi:10.1101/2024.05.13.593988.