Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

SUPPORTING INFORMATION

The Structure of Tagetitoxin

Abil E. Aliev,[,a] Kersti Karu,[a] Robin E. Mitchell,[b] Michael J. Porter[a]*

[a]Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, U.K.

[b]Plant and Food Research, Private Bag 92619, Auckland 1142, New Zealand

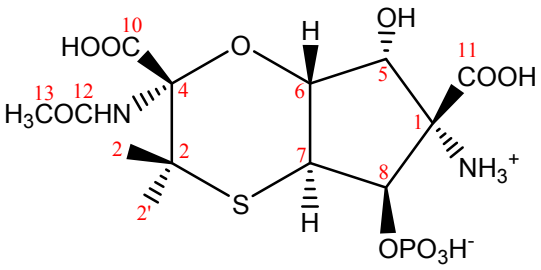


Table S1. Proton and carbon chemical shifts of tagetitoxin in D_2O. ^{1}H-^{13}C, ^{13}C-^{31}P and ^{1}H-^{31}P couplings are also shown. The measured ^{31}P chemical shift was 1.75 ppm (relative to 85% H_3PO_4).

Position	δ_H / ppm	δ_C / ppm	$^{1}J_{CH}$ / Hz	J_{CP} / Hz	J_{HP} / Hz
1	-	71.20		4.9	
2	3.25	33.07	142.8		
2’	2.98				
4	-	85.56			
5	4.46	72.92	158.4		
6	5.13	79.76	150.5		
7	3.48	43.20	141.3	2.8	
8	4.73	76.96	152.9	4.8	11.5
10	-	174.43			
11	-	171.16			
12	-	173.84			
13	2.01	22.91			
NH	8.47	-			

Table S2. Experimental values of ^{1}H-^{1}H coupling constants of tagetitoxin in D_2O. The calculated values averaged over two conformations, **4**-chair (75%) and **4**-twisted-chair (25%), are also shown.

	Exper. J_{HH} / Hz	*Calc.* J_{HH} / Hz[a]	*Calc.* J_{HH} / Hz[b]
2-2′	-13.7	-12.94	-
5-6	4.1	3.27	3.26
6-7	12.2	10.35	10.85
7-8	7.8	9.81	7.23

[a]DFT B3LYP/6-311+G(2d,p) IEFPCM(H_2O); [b]Using modified Karplus equation of [Haasnoot, C.A.G.; DeLeeuw, F.A.A.M.; Altona, C. *Tetrahedron* **1980**, *36*, 2783.]

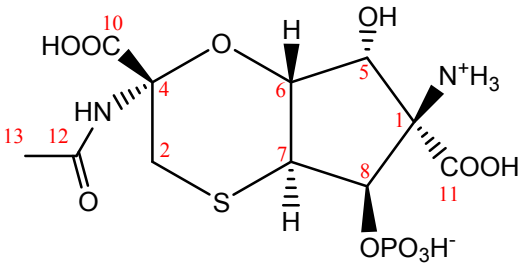
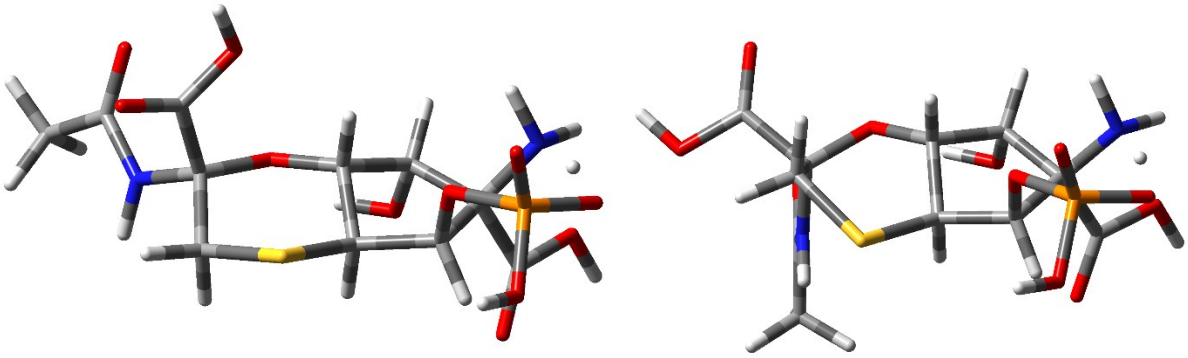


Table S3. Predicted values of $^{2,3,4}J_{CH}$ couplings for an alternative orientation of substituents in position 1 (structure **5** shown above). The calculated values [at the DFT B3LYP/6-311+G(2d,p) IEFPCM(H_2O) level of theory] for individual conformers, as well as averaged values over two conformers, **5**-chair (78%) and **5**-twisted-chair (22%), are shown.

	Exper. $^2J_{CH}$ / Hz	*Exper.* $^3J_{CH}$ / Hz	*Exper.* $^4J_{CH}$ / Hz	**5**-chair *Calc.* J_{CH}/Hz	**5**-tw.-chair *Calc.* J_{CH}/Hz	**5** *Calc.av.* J_{CH}/Hz
C1-H7		1.3		0.01	0.13	0.04
C1-H8	2.4			0.23	-0.20	0.14
C1-H5	~0			-0.71	-0.64	-0.69
C1-H6		~0		-0.05	0.02	-0.04
C2-H7		1.5		2.61	-0.02	2.06
C4-H2	2.3			-2.26	-1.86	-2.18
C4-H2′	4.1			-4.06	-2.03	-3.63
C4-H13			0.8	0.54	0.50	0.53
C4-H6		1.4		0.41	-0.03	0.32
C5-H7		(+)1.1		1.72	1.80	1.74
C5-H8		(+)0.9		3.69	4.39	3.84
C5-H6	(+)0.2			1.38	2.26	1.57
C6-H5	(-)0.7			0.48	0.74	0.54
C6-H7	(-)5.6			-4.95	-5.17	-5.00
C6-H8		(+)8.0		8.11	8.06	8.10
C7-H2		5.0		6.53	0.04	5.17
C7-H2′		3.0		2.28	5.77	3.01
C7-H5		(+)6.1		6.67	6.71	6.68
C7-H6	(-)2.7			-2.43	-2.17	-2.38
C7-H8	(-)1.1			-1.26	-1.10	-1.23
C8-H2			1.4	1.41	-0.13	1.09
C8-H5		(+)6.2		3.57	3.11	3.48
C8-H6		(+)0.3		0.95	0.99	0.96
C8-H7	(-)0.4			-0.94	0.23	-0.70
C10-H2		1.2		1.27	1.40	1.30
C10-H2′		5.0		7.03	1.73	5.92
C11-H8		1.5		3.66	1.37	3.60
C11-H5		2.7		0.26	0.51	0.31
C12-H13	6.0			-5.56	-5.59	-5.56
C12-NH	3.7			4.56	5.04	4.66
rms_J/Hz				1.24	1.81	1.17



5-chair **5**-twisted-chair

Figure S1. Geometries of **5**-chair and **5**-twisted-chair conformations derived from DFT M06-2X/def2-TZVP calculations with the alternative orientation of substituents at C1 compared to **4** (Figures 4 and 5 in the main text). One of the N^+H_3 protons is delocalized between N^+ and OPO_3H^- groups in both conformations, with the N^+…H distance of 1.14 Å.

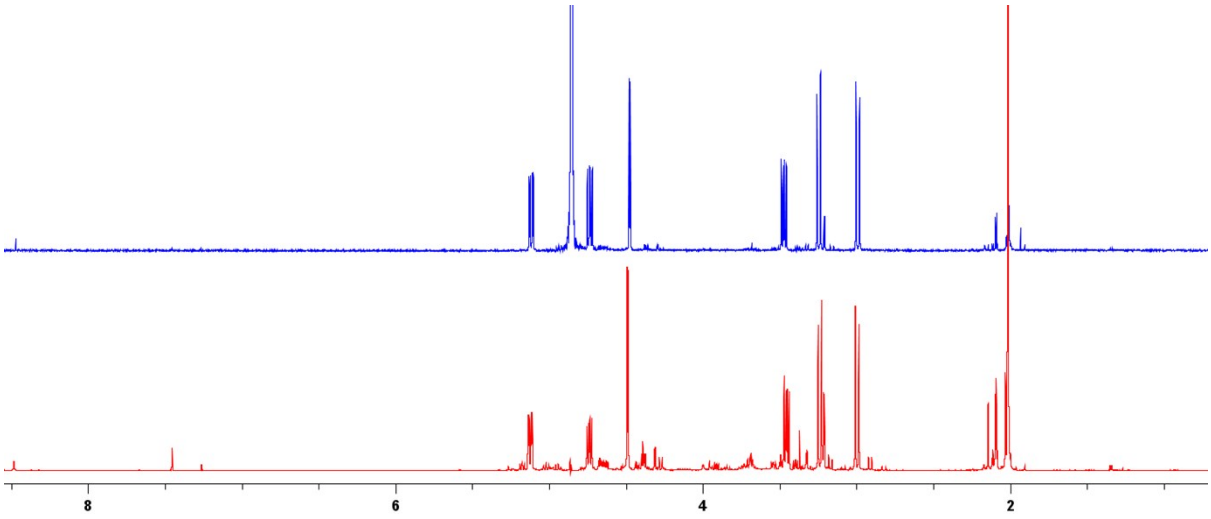


Figure S2. Experimental ^{1}H NMR spectra of tagetitoxin (293 K, 600 MHz). Top: freshly prepared sample; bottom: after 6 weeks in D_2O (with solvent suppression). The residual amide proton N**H**COMe of tagetitoxin is observed at 8.47 ppm in both spectra.

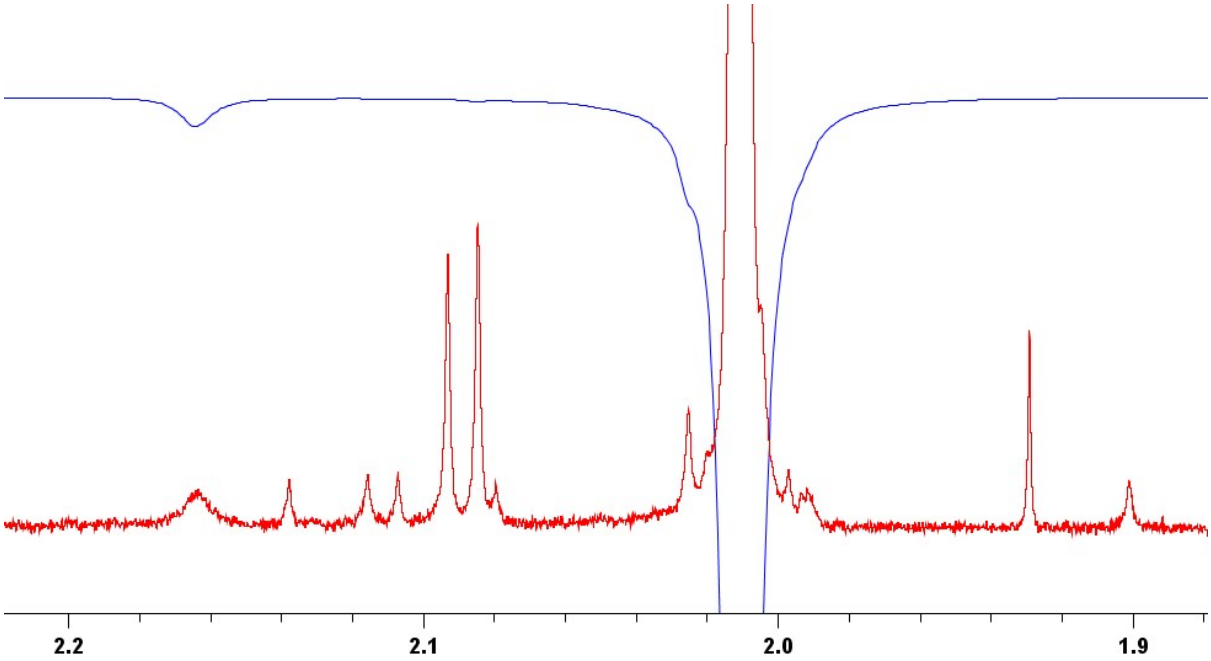


Figure S3. 1D NOESY experiment with selective excitation of methyl protons at 2.01 ppm (in blue) and ^{1}H NMR spectrum (in red). The negative enhancement observed at 2.16 ppm is in favour of chemical exchange between two rotamers with the estimated population ratio of 68 (2.01 ppm) : 1 (2.16 ppm).

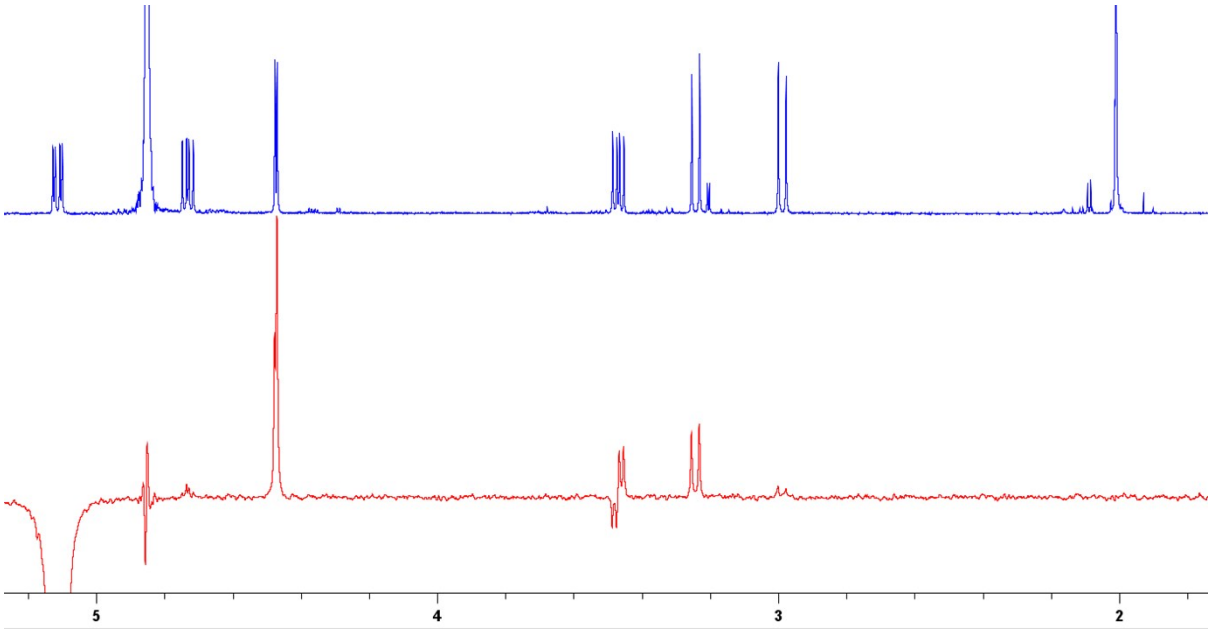


Figure S4. 1D NOESY experiment with selective excitation of H6 (in red) and ^{1}H NMR spectrum (in blue).

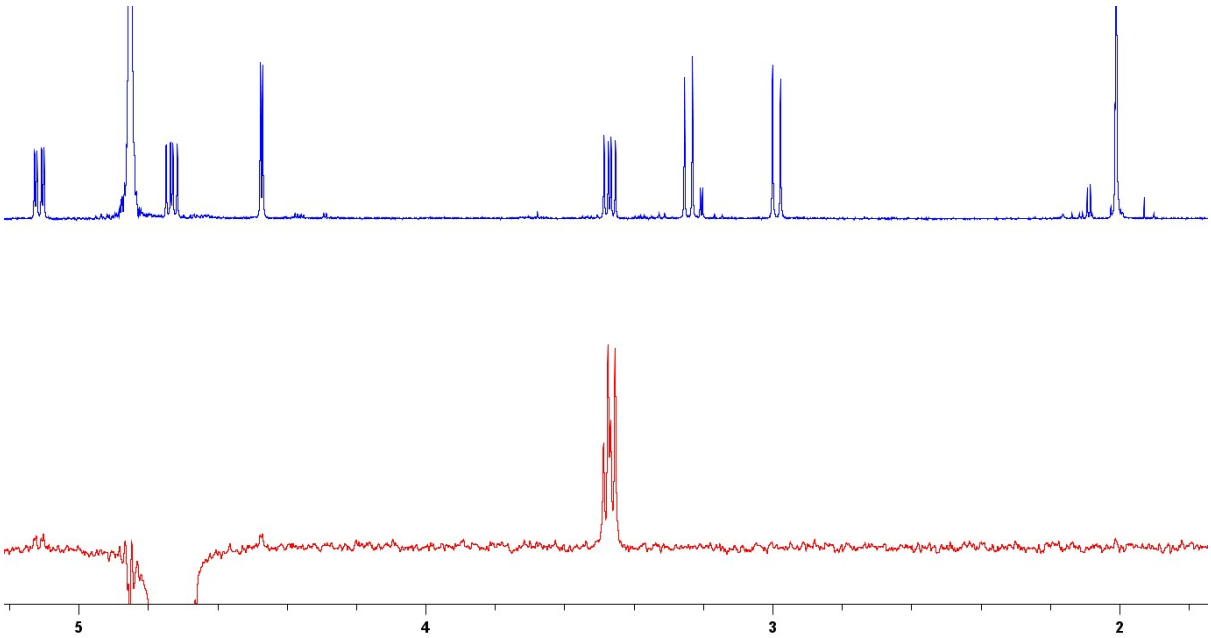


Figure S5. 1D NOESY experiment with selective excitation of H8 (in red) and ^{1}H NMR spectrum (in blue).

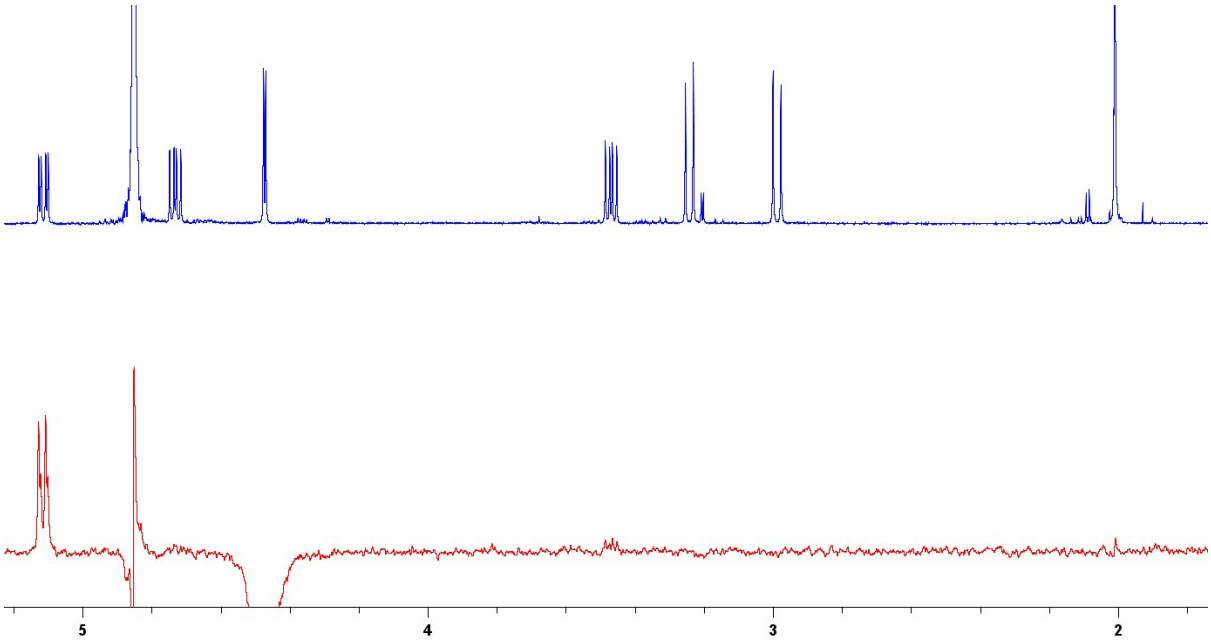


Figure S6. 1D NOESY experiment with selective excitation of H5 (in red) and ^{1}H NMR spectrum (in blue).

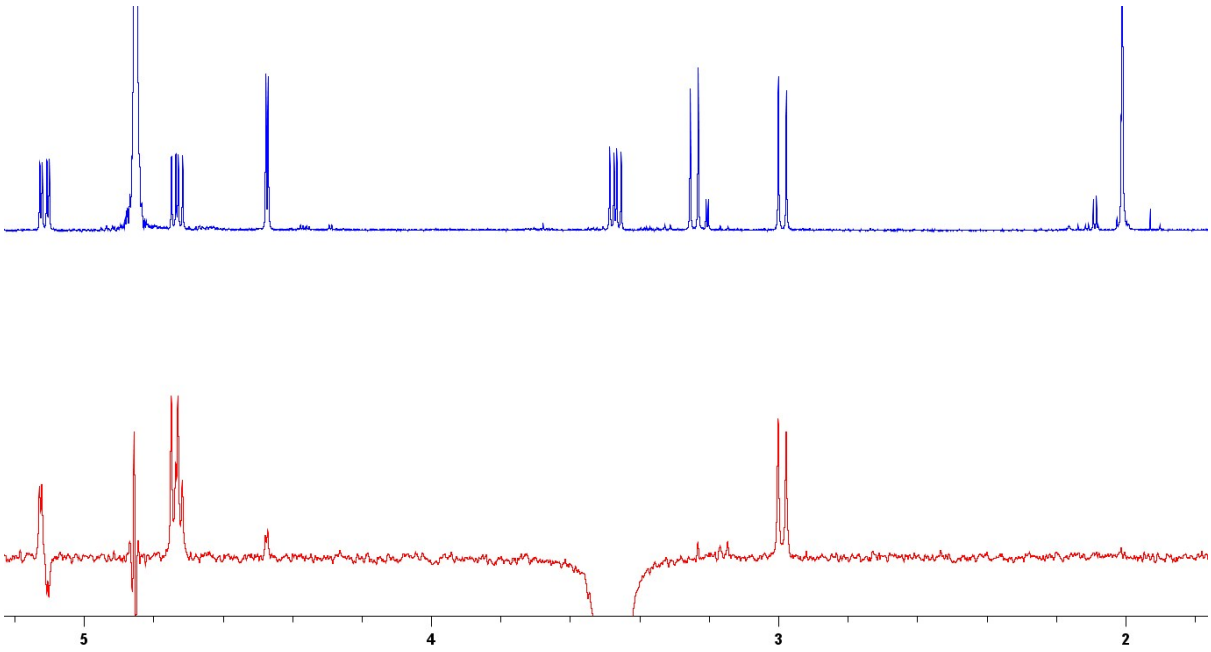


Figure S7. 1D NOESY experiment with selective excitation of H7 (in red) and ^{1}H NMR spectrum (in blue).

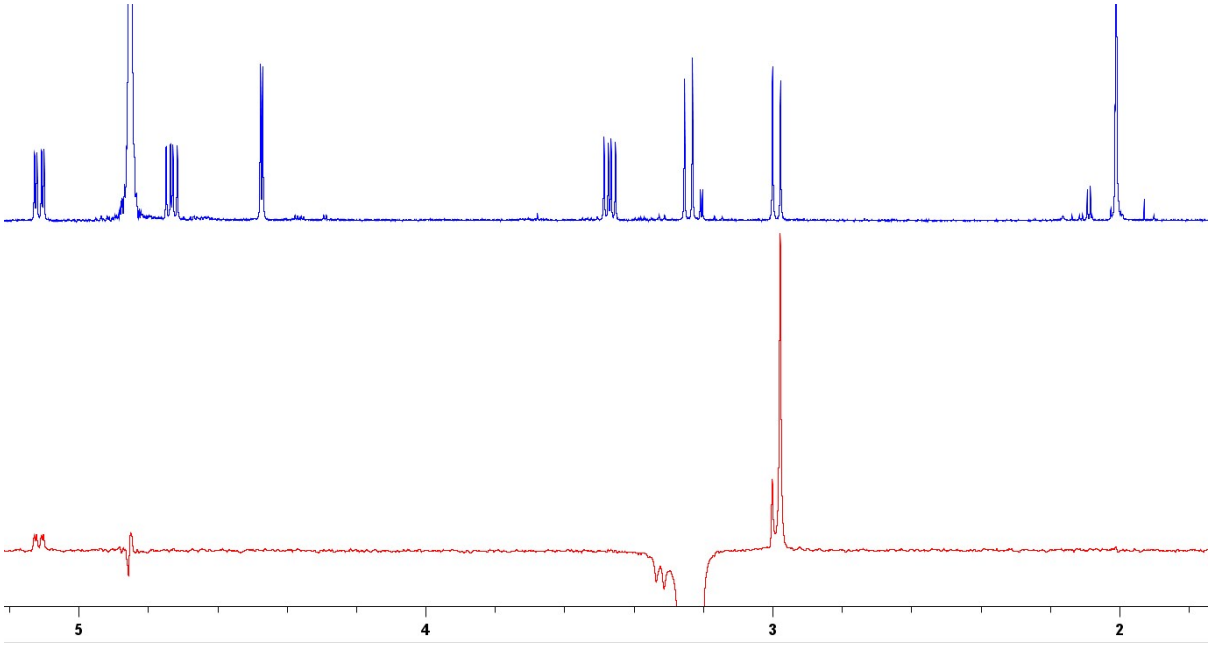


Figure S8. 1D NOESY experiment with selective excitation of H2 (in red) and ^{1}H NMR spectrum (in blue).

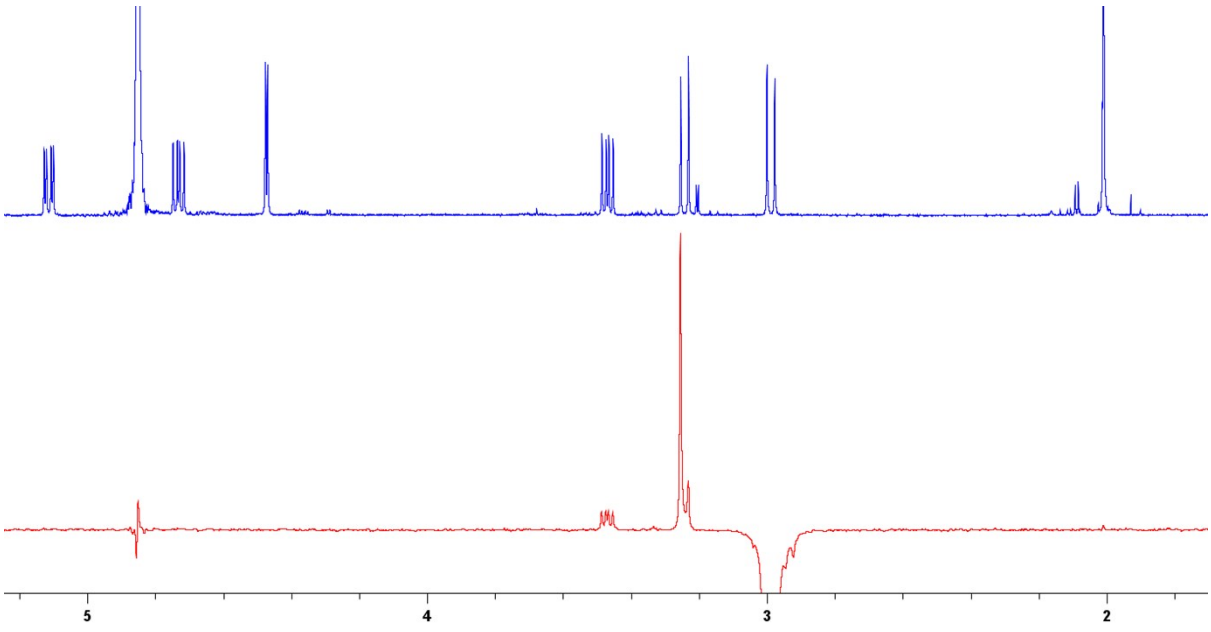


Figure S9. 1D NOESY experiment with selective excitation of H2′ (in red) and ^{1}H NMR spectrum (in blue).

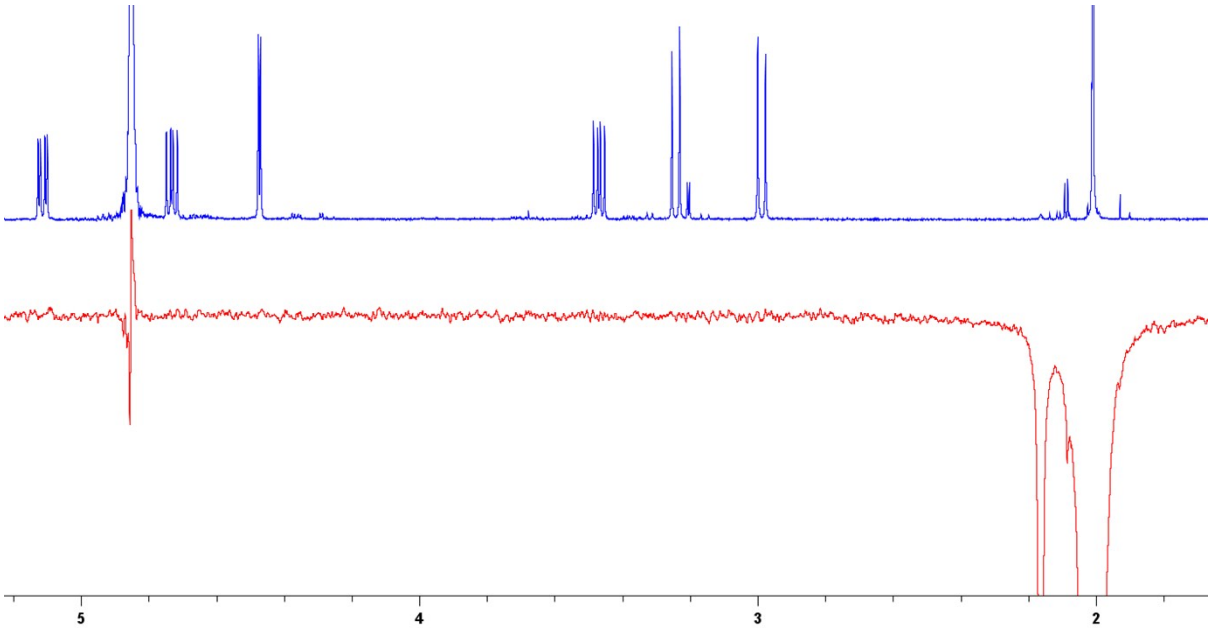


Figure S10. 1D NOESY experiment with selective excitation of methyl protons (in red) and ^{1}H NMR spectrum (in blue).

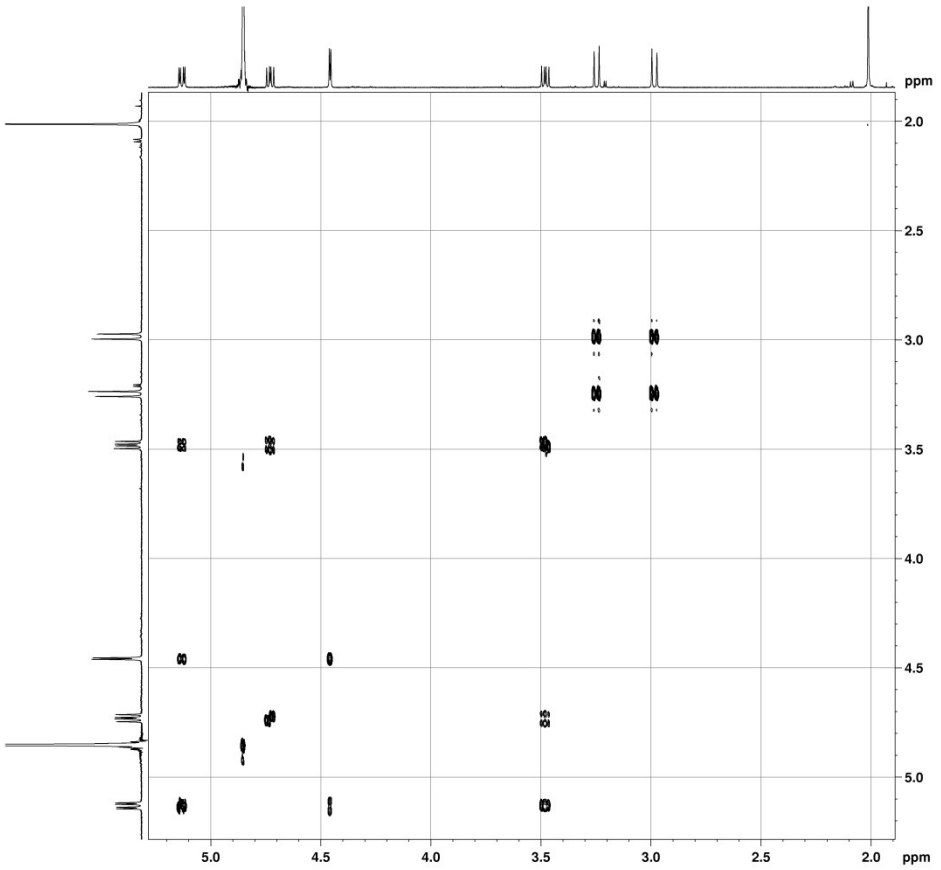


Figure S11. The ^{1}H-^{1}H COSY spectrum of tagetitoxin in D_2O.

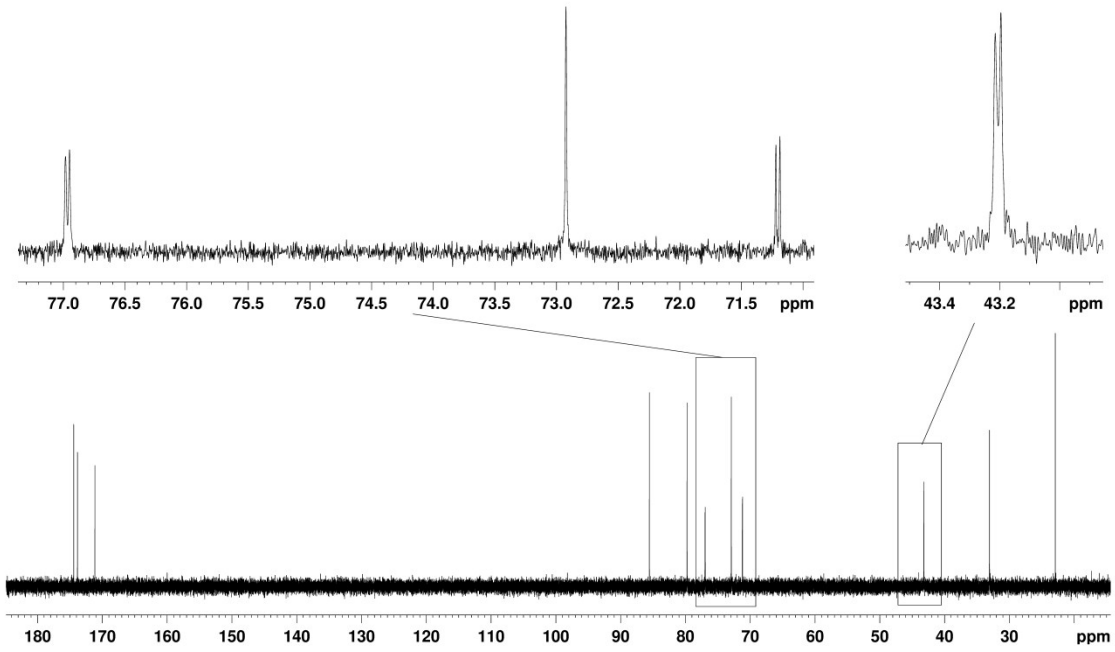


Figure S12. The ^{13}C{^{1}H} spectrum of tagetitoxin in D_2O. Splittings observed are due to J_{CP} couplings.

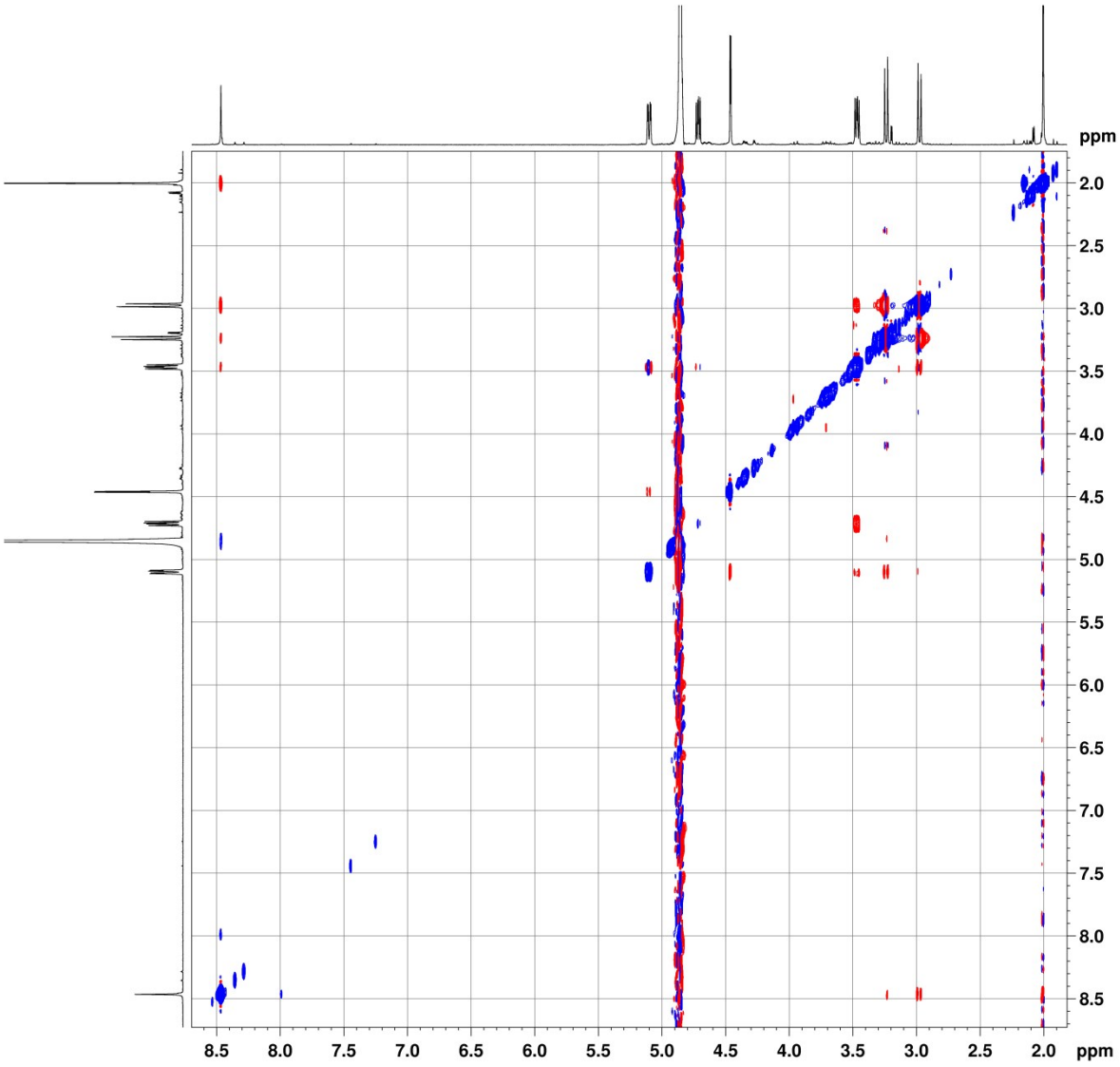


Figure S13. The ^{1}H-^{1}H NOESY spectrum of tagetitoxin in H_2O:D_2O (9:1) with mixing time of 600 ms and water suppression using excitation sculpting.

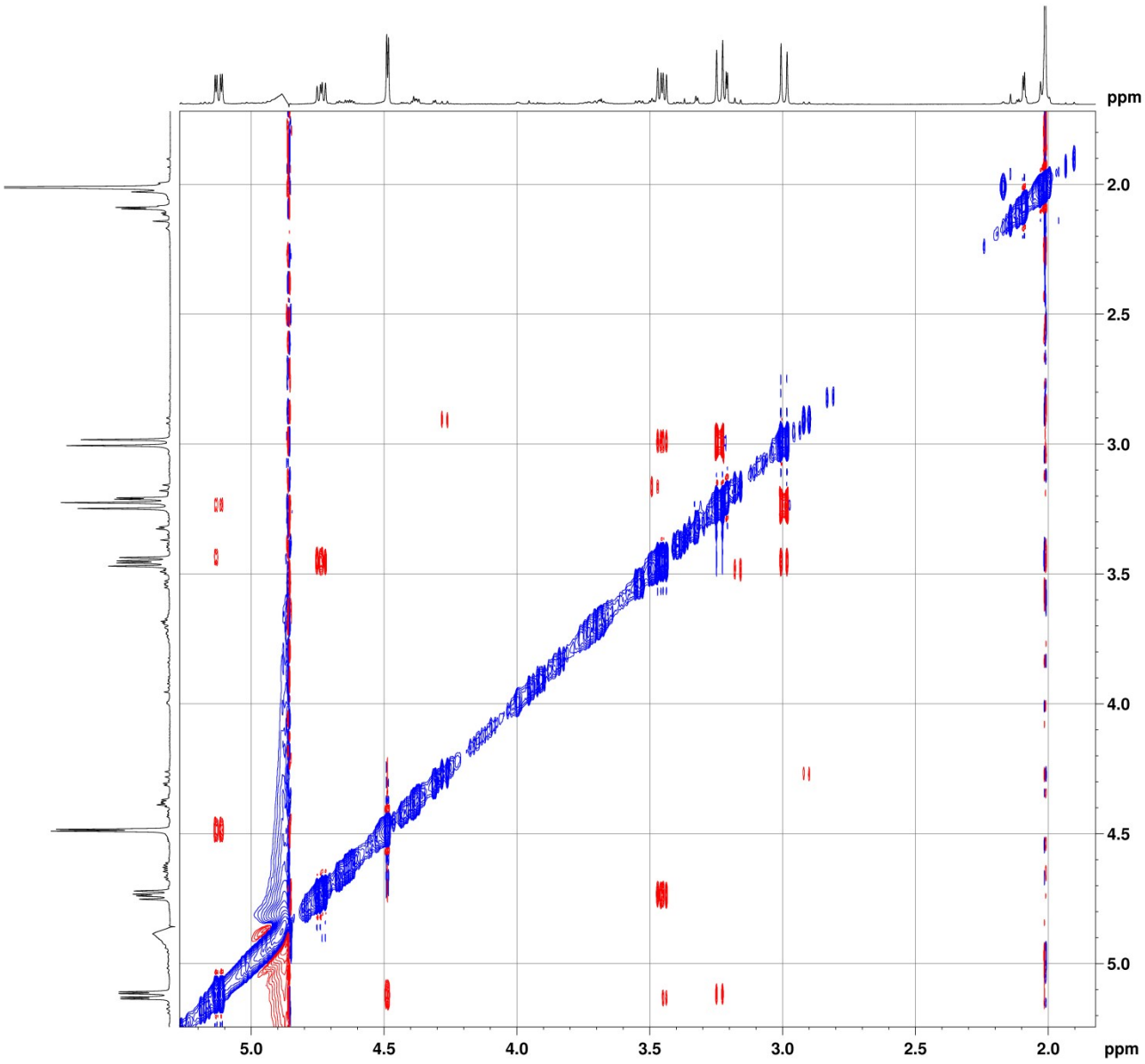


Figure S14. The ^{1}H-^{1}H EASY-ROESY spectrum of tagetitoxin in D_2O with a 300 ms spinlock pulse.

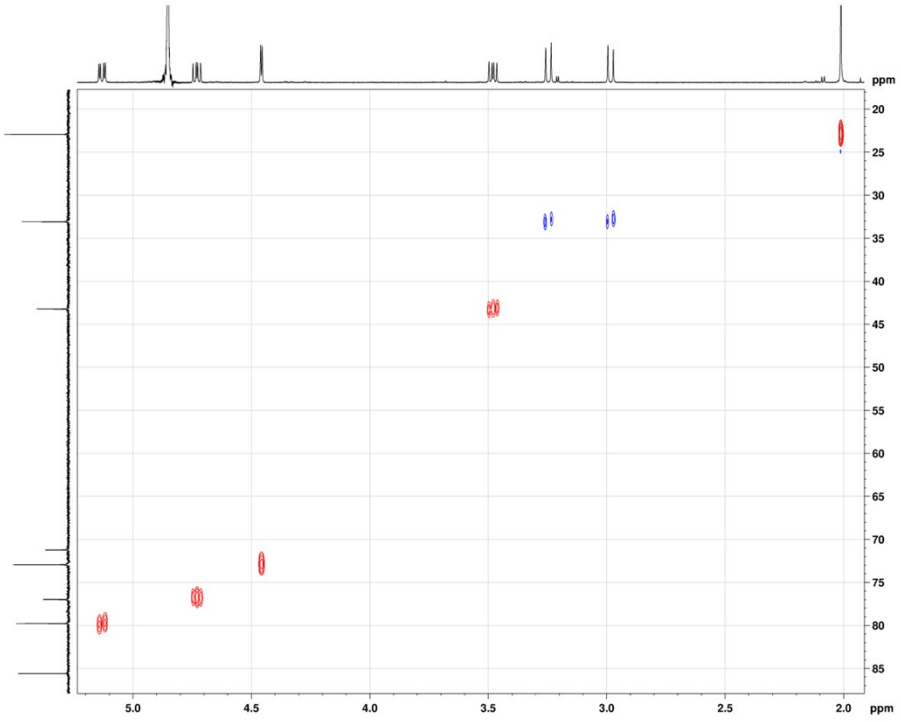


Figure S15. The ^{1}H-^{13}C HSQC spectrum of tagetitoxin in D_2O.

Figure S16. The ^{1}H-^{13}C HMBC-JC spectrum of tagetitoxin in D_2O.

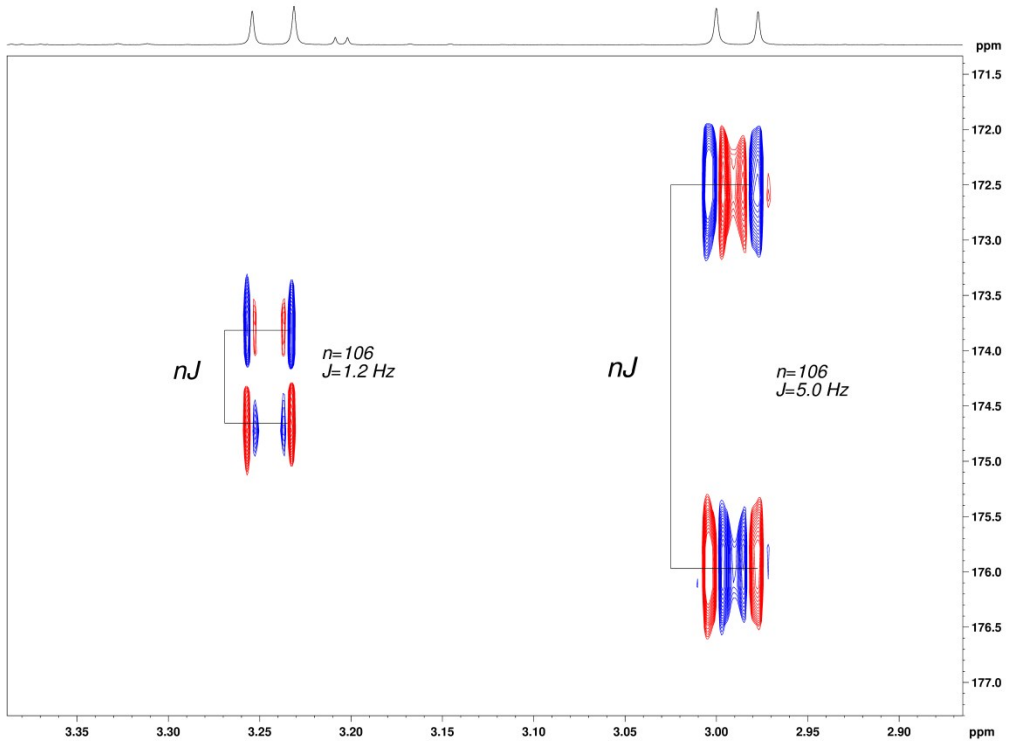


Figure S17. Expansion of the ^{1}H-^{13}C HMBC-JC spectrum of tagetitoxin in D_2O illustrating measurements of $^{3}J_{C10H2}$ (left) and $^{3}J_{C10H2'}$ (right) couplings (n = 106 is the scaling factor).

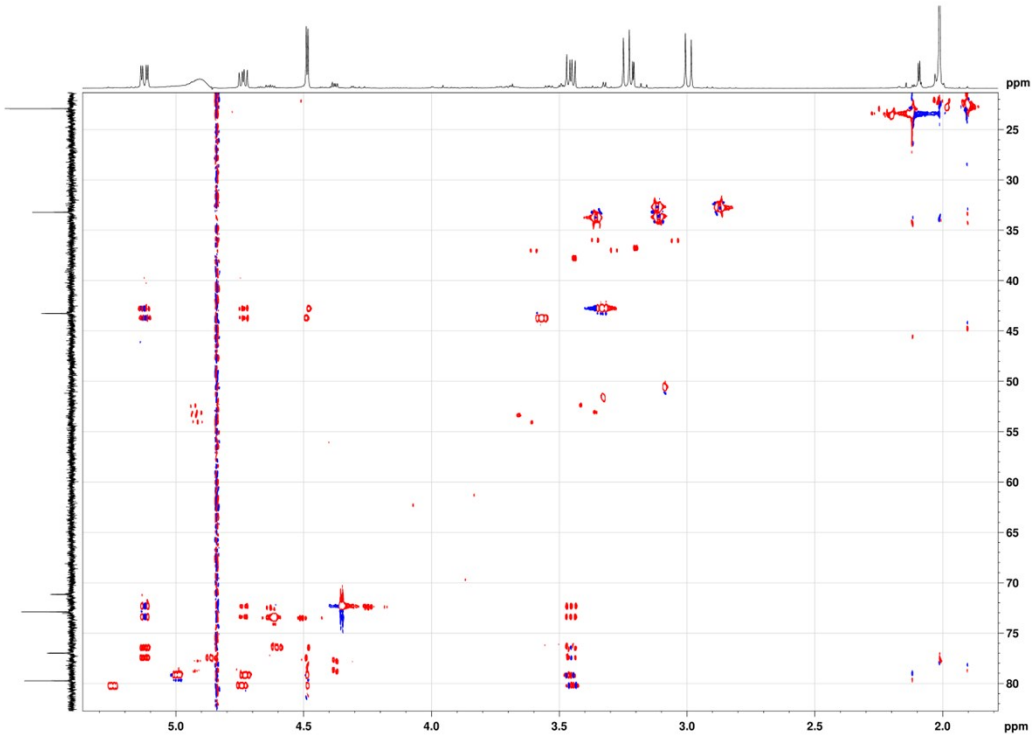


Figure S18. The ^{1}H-^{13}C HSQC-HECADE spectrum of tagetitoxin in D_2O.

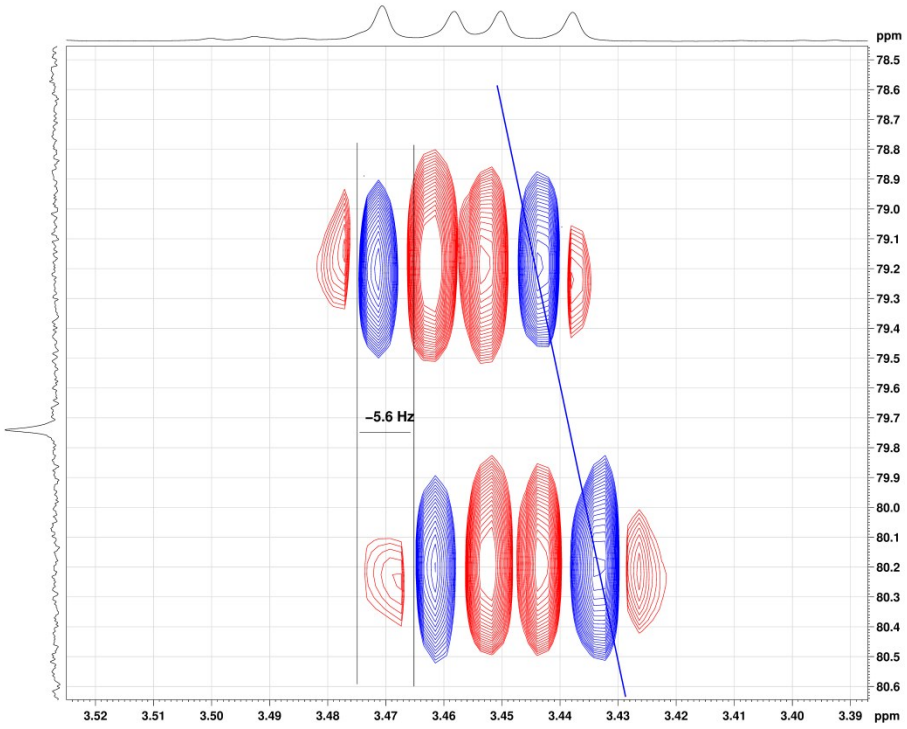


Figure S19. Expansion of the ^{1}H-^{13}C HSQC-HECADE spectrum of tagetitoxin in D_2O illustrating the negative "\" tilt observed for the J_{C6H7} coupling.

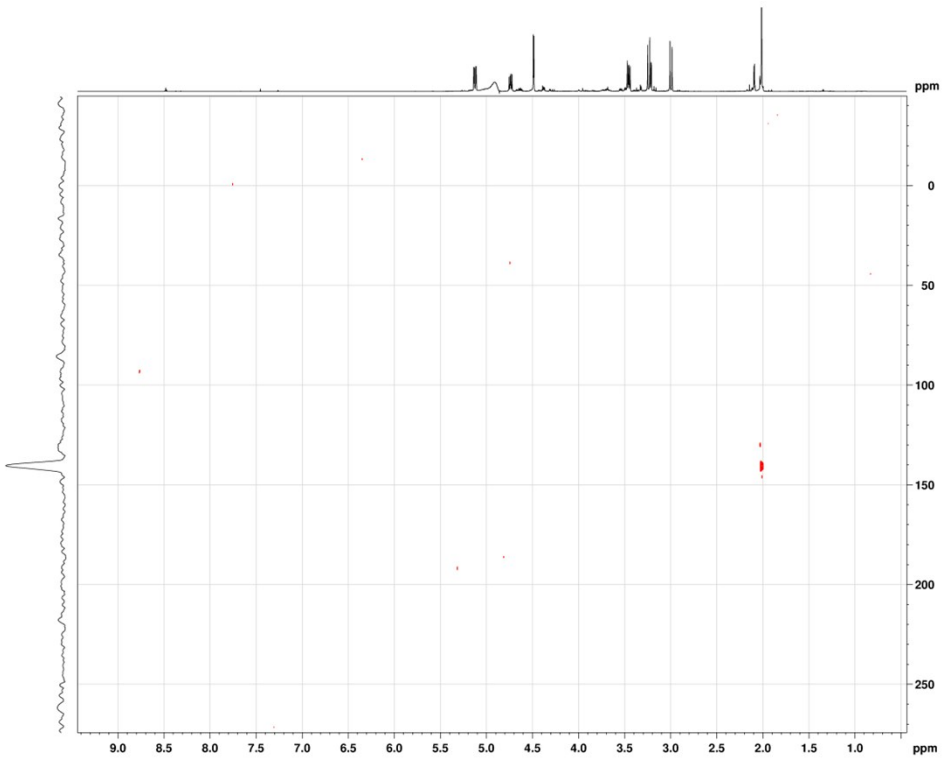


Figure S20. The ^{1}H-^{15}N HMBC spectrum of tagetitoxin in D_2O.

Mass spectrometry analysis

Capillary-LC-MS/MS analysis used for a structure elucidation of tagetitoxin

Chromatographic separation was performed on an Accela HPLC system utilising a Hypersil GOLD reversed-phase column (1.9 µm particles, 50 x 2.1 mm) from Thermo Electron Corp (San Jose, CA). Mobile phase A consisted of 1% methanol containing 0.1% formic acid, and mobile phase B consisted of 50% methanol and 50% acetonitrile containing 0.1% formic acid. After 0.5 min at 5% B, the proportion of B was raised to 70% B over the next 8 min. This followed by increase to 95% B in 0.5 min, and remaining at 95% B for 2.5 min, before returning to 5% B in 6 s and re-equilibration for a further 2 min 50 s, giving a total run time of 13 min. The flow rate was maintained at 200 µL min and eluent directed to the capillary ESI source of a LTQ mass spectrometer. Mass spectrometry analysis was performed on a Finnigan LTQ (Thermo Electron Corp, San Jose, CA) with a linear ion-trap (LIT) analyser. Ions from the ESI source, operated in the positive ESI mode, were analysed in both MS and MS^2 modes. Full mass spectra was acquired at *m/z* 100 - 1000. MS and MS^2 scans consisted of three averaged "microscans" each with a maximum injection time of 200 ms. For the acquisition of MS^2 spectra the collision energy setting was 35 using argon as the collision gas. The MS^2 isolation width was set at 2 so as to allow the selection of monoisotopic precursor ions. Collision gas was helium.

For the analysis of tagetitoxin, around 100 µg of tagetitoxin was solubilised in 1 mL methanol-water (1:1, v/v) and the stock solution was further diluted in 1:1000 methanol-water (1:1, v/v). Two µL of 100 pg/µL was injected onto the reversed-phase column and eluted into the LTQ at a flow-rate of 200 µL/min. MS and MS^2 spectra were recorded.

Accurate mass measurement

Two accurate mass measurement analyses were used to determine the accurate and precise value of *m/z* for a gas phase ion of tagetitoxin with a view to suggesting elemental composition.

Ten µL of the sample in methanol was injected using a Waters 1525µ binary HPLC pump system connected to a Waters Sample Manager autosampler. An isocratic mode was utilised using 50%

methanol - 50% water, 0.1% formic acid with the flow rate of 200 µL/min. The effluent was continuously directed into a Waters LCT Premier XE. Exact mass measurements were carried out automatically using the LCT Premier's integral LockSpray dual-electrospray source. The LockSpray dual-electrospray source was operated in both positive- and negative-ion ESI mode with full MS scans over 2 min time period of the direct infusion isocratic run. The LCT Premier mass spectrometer was calibrated prior an accurate mass measurement with sulfadimethoxine $[M+H]^+$ with *m/z* value of 311.0814 and leucine encephalin $[M+H]^+$, *m/z* value of 556.2771. The LCT Premier XE was operated in W-mode with the following parameters: the source capillary was 2300 V, sample cone 10 V, desolvation temperature 400 °C, source temperature 80 °C, cone gas flow 50, and desolvation gas flow 450 L/hr. The analyser parameters were TOF flight tube 7200 V, reflectron 1800 V, pusher voltage 895, MCP 2500. Data was acquired in centroid mode.

The second accurate mass measurement was performed on an Agilent 6510 Q-TOF LC-MS system. Three µL of 20 pg/µL of tagetitoxin in water-methanol (1:1, v/v) was injected on a Chromolith Performance RP-18 endcapped, 100 mm x 2mm column (Merck, USA). The column temperature was 30 °C. Mobile phase A was 0.1% formic acid and B 95% acetonitrile in 0.1% formic acid. The gradient was as follows: 1% B for 0.5 min with increase to 3% B over 0.2 min and gradient increase to 5% B over 3 min and sharp increase to 95% B in 0.2 min and held at 95% B for 3.4 min and rapid decrease to 1% B hold for 3 min at 1% B. Post-LC analysis was set up for 3 min. The effluent was directed through dual ESI source to an Agilent 6510 Q-TOF mass spectrometer. An accurate-mass Q-TOF mass spectrometer operated in high resolution positive ion mode at standard 3200 *m/z* with an extended dynamic range. The Q-TOF was calibrated externally with HP121/HP921 solution (Agilent, UK) and mass accuracy was better than 5 ppm. Source parameters were optimised with sulfadimethoxine $[M+H]^+$ with *m/z* value of 311.0814 and reserpine $[M+H]^+$ with *m/z* value of 609.2813. These standards were prepared at 20 pg/µL in water-methanol (1:1, v/v, 0.1% formic acid). Drying gas temperature was 325 °C and the flow rate was 10 L/min. Nebuliser was set at 20; Vcap was 3500; fragmentor was 175 V; skimmer at 50 V; OCT1 RF voltage was 750V. The acquisition mass range was

from *m/z* 50 to 2000 with the rate 3 spectra/s. For targeted MS/MS, set at *m/z* 417.0363 ± 5 ppm, the acquisition mass range was *m/z* 60 to 1000 with the rate of 5 spectra/s. MS/MS was acquired at collision energies of 10, 20 and 30. For MS/MS experiments, N_2 was used as the collision gas. Reference mass ions were at *m/z* 121.05087 and 922.00979 over LC chromatographic run. Data was acquired in centroid mode.

Results

Sixty pg of tagetitoxin was injected on a reversed phase monolithic column and the effluent was directed into the Agilent 6510 Accurate-Mass Q-TOF. Total ion chromatogram is shown in Figure S21a. The ESI spectrum of chromatographic peak at 1.5 min give $[M+H]^+$ ion at *m/z* 417 (Figure S21b).

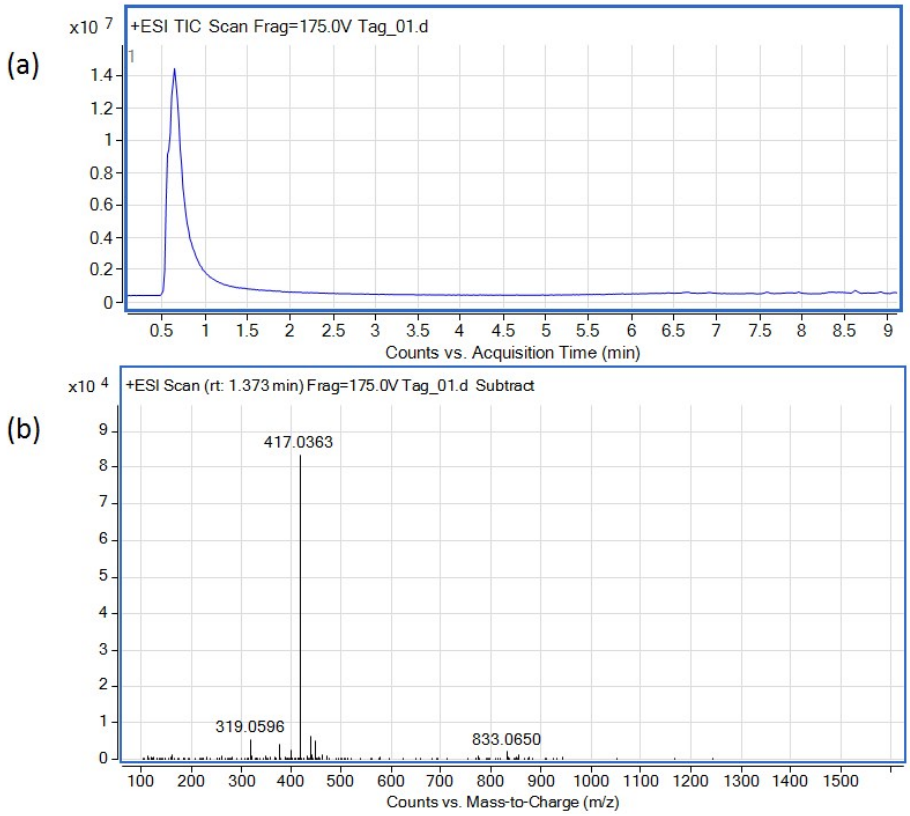


Figure S21. (a) Total ion chromatogram of 60 pg of tagetitoxin injected on a Chromolith Performance RP-18 endcapped, 100 mm x 2 mm column using an Agilent LC. (b) ESI mass spectrum of the peak at 1.5 min. Spectra were recorded on an Agilent 6510 Q-TOF.

Spectra were obtained based on an internal mass calibration. We used Agilent Molecular formula generation (MFG) software as part of Agilent Mass Hunter software to generate possible elemental composition for this *m/z* value consisting of carbon, nitrogen, sulphur, phosphor and oxygen atoms.

The *m/z* 417.0363 with the mass error of 0.09 ppm and the isotope abundance match generated the elemental composition formula $C_{11}H_{18}N_2O_{11}PS$ (Figure S22). As tagetitoxin generated $[M+H]^+$ ions at *m/z* 417.0 in the ESI process leading to the determination of molecular weight of 416.29. For confirmation of an identity, *m/z* 417.0363 was selected and analysed using collision induced

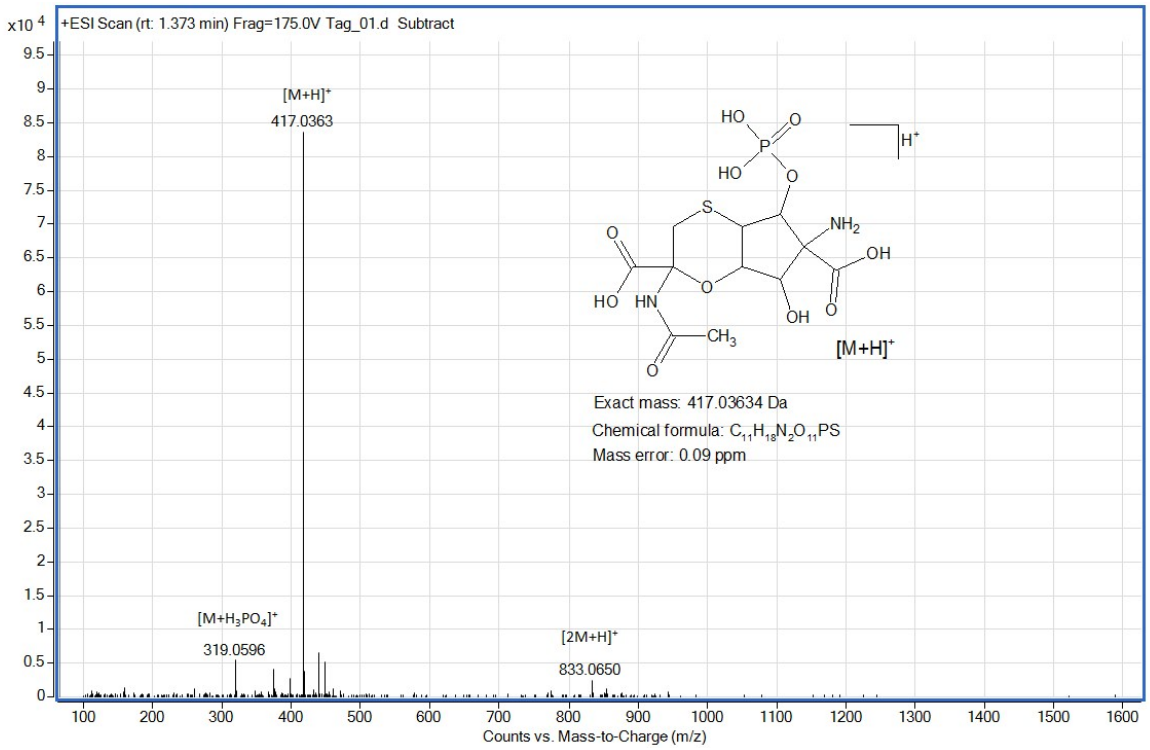


Figure S22. Agilent Mass Hunter software an elemental composition determination. ESI mass spectrum of the chromatographic peak eluting at 1.5 min identified as tagetitoxin acquired with the Agilent 6510 Q-TOF.

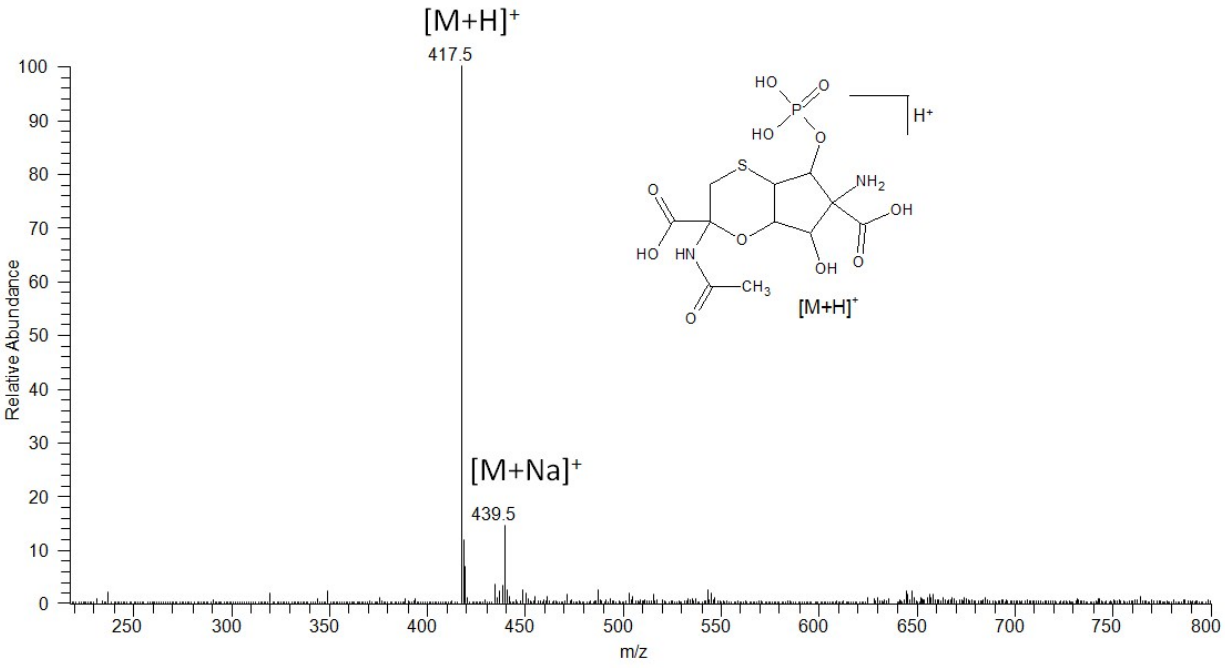


Figure S23. ESI mass spectrum of the chromatographic peak eluting at 2.9 min identified as tagetitoxin using an Acella LC connected to the linear ion trap mass spectrometer.

dissociation (CID) in Q-TOF MS/MS mode. Also, we used a second LC-MS (MS^n) system, 80 pg of tagetitoxin was injected onto a C18 column the effluent was directed to the linear ion trap (LTQ) mass spectrometer, where MS, -MS^2 spectra were continuously recorded over the chromatographic run. MS^2 experiment was performed as follows: a precursor ion, $[M+H]^+$ was selected followed by activation of the $[M+H]^+$ ion, and recording of its fragmentation pattern, when tagetitoxin eluted from the column it was subjected to MS^2 ($[M]^+\rightarrow$). MS spectrum shows $[M+H]^+$ ion at *m/z* 417 and $[M+Na]^+$ ion at *m/z* 439 corresponding to tagetitoxin (Figure S23). This is in agreement with mass spectra obtained from the Agilent Q-TOF mass spectrometer. MS^2 (417→) spectrum gave a wealth of information on the structure of tagetitoxin. The MS^2 (417→) spectrum shows a rapid loss of water from the precursor ion at *m/z* 417 generating $[M+H-H_2O]^+$ ions at *m/z* 399. Two routes of tagetitoxin gas-phase fragmentation occurred with: (a) loss of H_3PO_4 and (b) loss of C_2H_2O from the precursor ion $[M]^+$ (Figure S24).

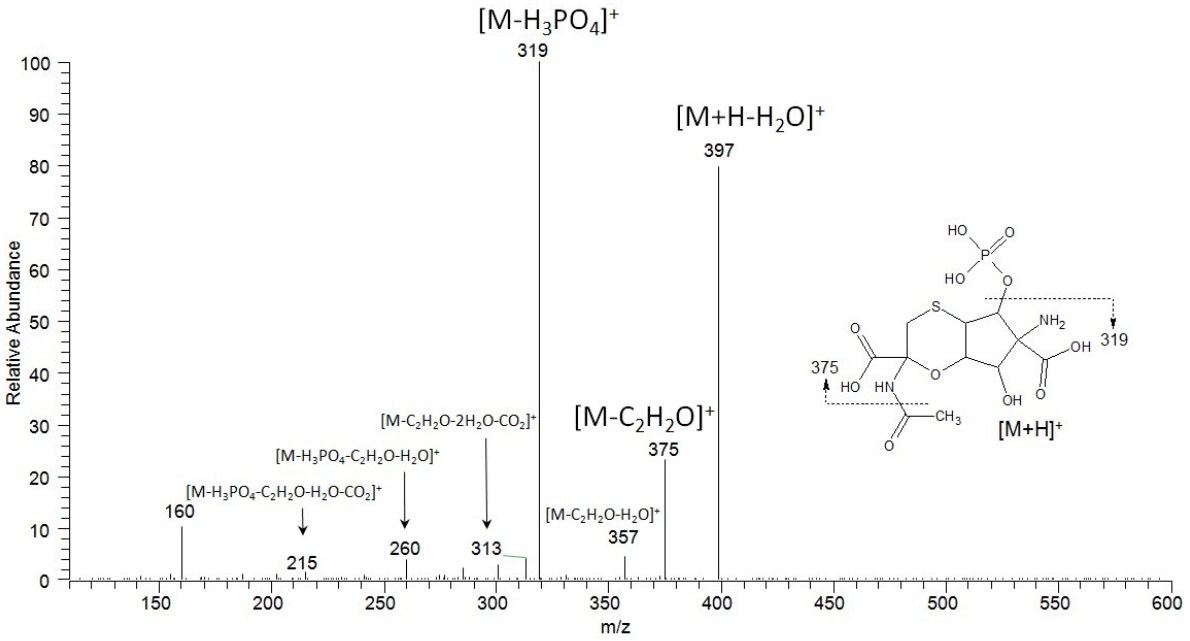


Figure S24. MS^2 (417→) spectrum of the chromatographic peak eluting at 2.9 min identified as tagetitoxin using an Acella LC connected to the linear ion trap mass spectrometer.

The most abundant peak at *m/z* 319 generated due to loss of a H_3PO_4 radical (Δm = 98 Da, H_3PO_4, requires 98 Da). The *m/z* 375 ion relative intensity amounted to 25% and displaced in mass by 42 Da

from the precursor ion $[M+H]^+$ corresponding to loss of C_2H_2O. The MS^2 (417→) spectrum shows a $[M-C_2H_2O-H_2O]^+$ ion observed at *m/z* 357 which corresponds to loss of water molecule and C_2H_2O from the precursor ion, $[M]^+$. This fragment at *m/z* 357 is sequentially losing 44 Da corresponding to loss of CO_2 generating $[M-C_2H_2O-H_2O-CO_2]^+$ ions at *m/z* 313 (with relative intensity 4%). Further loss of 98 Da was observed from the fragment ion at *m/z* 313 giving $[M-H_3PO_4-C_2H_2O-H_2O]^+$ ions at *m/z* 260. This fragment at *m/z* 260 has to lose CO_2 moiety and proton re-arrangement occurred with the loss of proton to generate the peak at *m/z* 215 (RA 2%). This was followed by further loss of 44 Da and 17 Da and further re-arrangements to generate the fragment at *m/z* 160. Figure S25 summarises possible tagetitoxin gas-phase fragment ions at collision energy 35% in the linear ion trap mass spectrometer.

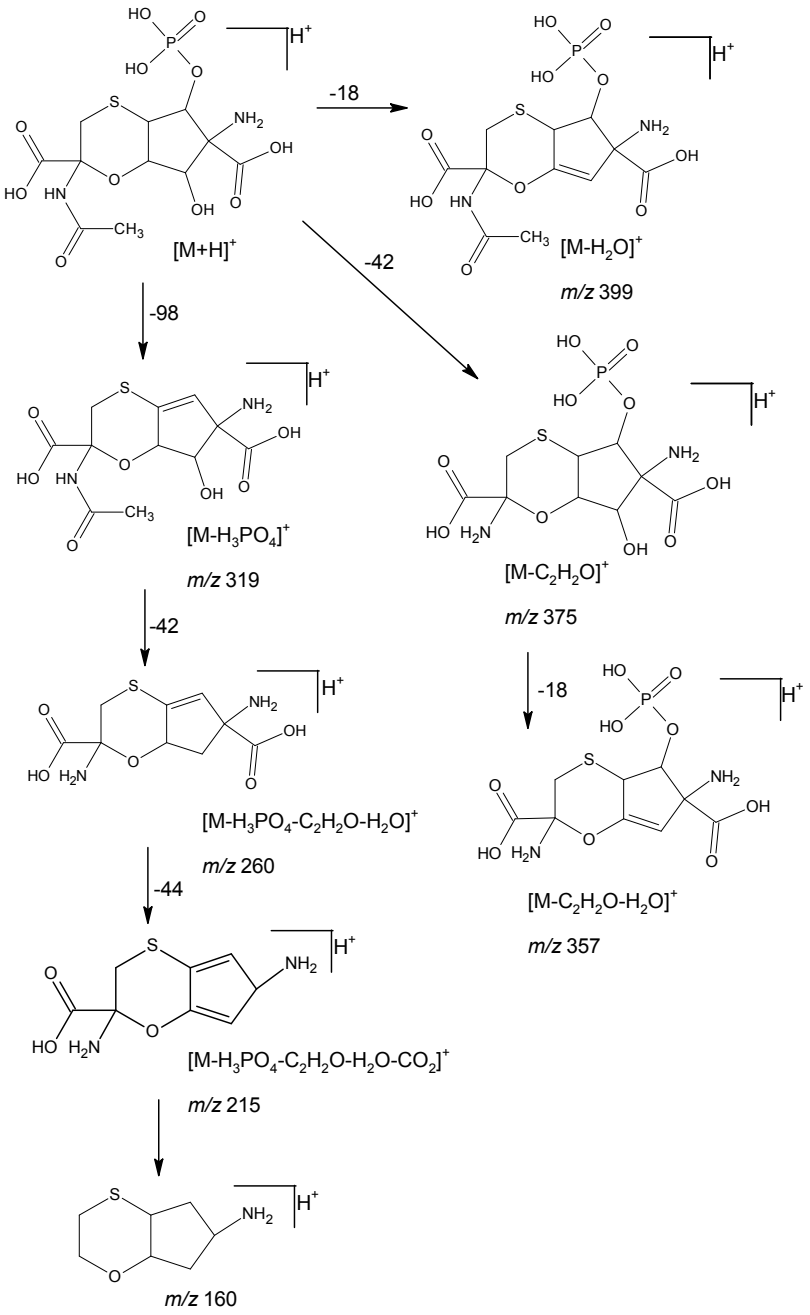


Figure S25. Possible ways of tagetitoxin fragmentation in gas phase using a Thermo LTQ.

Figure S26 shows the CID mass spectrum of tagetitoxin (*m/z* 417.0363), as observed with the Agilent 6510 Q-TOF. The precursor ion at *m/z* 417.0363 was still observed at collision energy 10 and 20 (Figure S26 a,b), but was not present at 30. The most abundant fragment ions were identical to data acquired on the LTQ. Three product ions, $[M- C_2H_2O]^+$ at *m/z* 375, $[M-H_3PO_4]^+$at *m/z* 319 and $[M-H_3PO_4-C_2H_2O-2H_2O-2CO_2-NH]^+$ at *m/z* 160 were observed (Figure S26a). The same fragmentation behaviour was observed for most fragment ions in the LTQ.

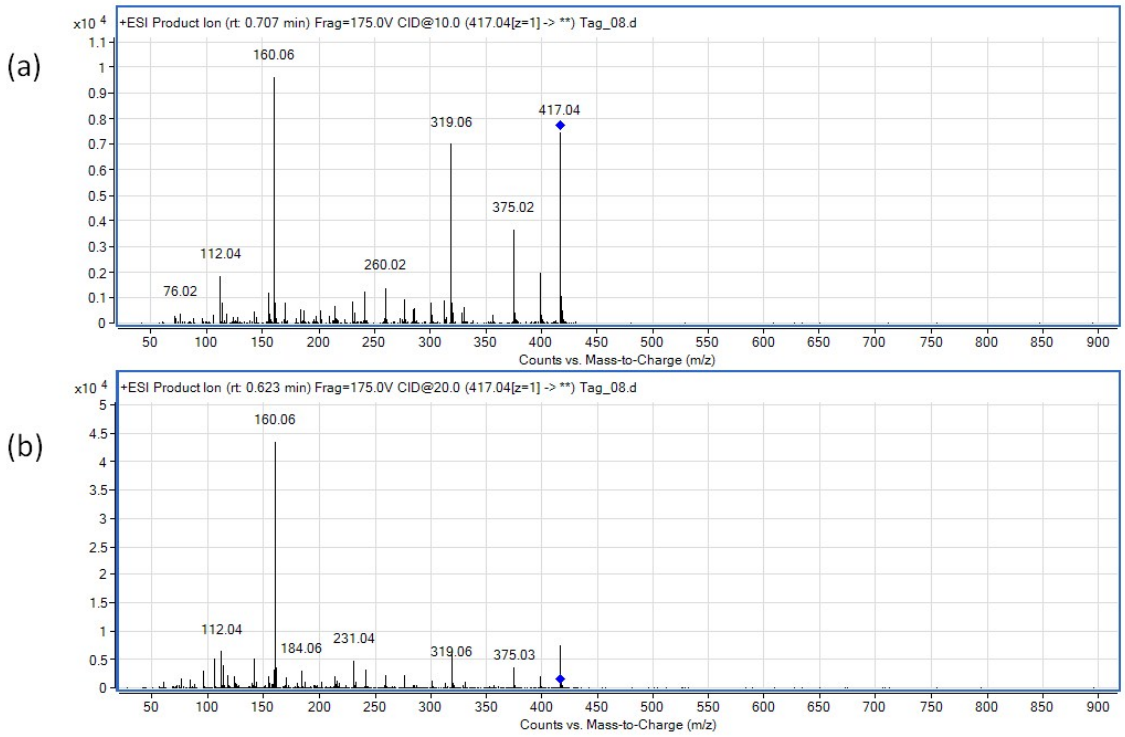


Figure S26. CID mass spectrum of tagetitoxin *m/z* 417.0363 at chromatographic peak eluting at 1.5 min (a) with collision energies (a) 10 and (b) 20 acquired an Agilent 6510 Q-TOF.

In the third experiment, we acquired a negative and positive ion mode ESI mass spectra simultaneously with an automatic lock-mass accurate mass measurement setting on a Waters LCT Q-TOF. The sample was injected through a direct infusion LC system into the LCT Q-TOF. The positive ion mode ESI mass spectrum shows peaks at *m/z* 417, 439 and 449 corresponding to $[M+H]^+$ and $[M+Na]^+$ and $[M+H+CH_3OH]^+$ ions of tagetitoxin. This is an agreement with data from the Thermo LTQ and Agilent 6510 Q-TOF. Negative ESI mode shows the peak at *m/z* 414 corresponding to $[M-3H]^-$ ions with relative intensity of 100% (Figure S27). The peaks at *m/z* 435 and 445 correspond to $[M-3H+Na]^-$ and $[M-3H+CH_3OH]^-$, respectively. The [M-3H]- species, which was observed to be generated as a major product only with the phosphate and carboxylic anions can nominally correspond to the loss of three hydrogen atoms from two carboxylic groups and a phosphate group under negative ESI mode. These data suggests that tagetitoxin possesses two carboxylic groups and a phosphate group.

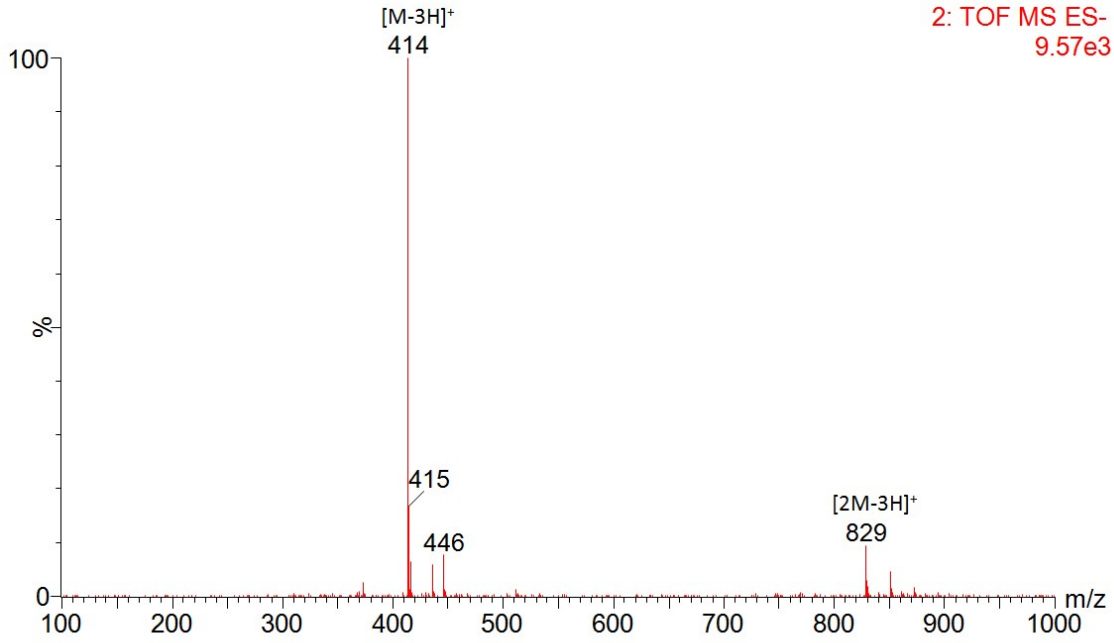


Figure S27. Negative mode ESI mass spectrum of tagetitoxin acquired on the Waters LCT Q-TOF mass spectrometer.

Furthermore, the isotopic pattern of this signal confirmed the presence of P and S atoms. As a proof, the isotopic pattern at *m/z* 414 is compared with the theoretical patterns as calculated for assigned ionic formula [$C_{11}H_{18}N_2O_{11}PS$] (Figure S28 a,b).

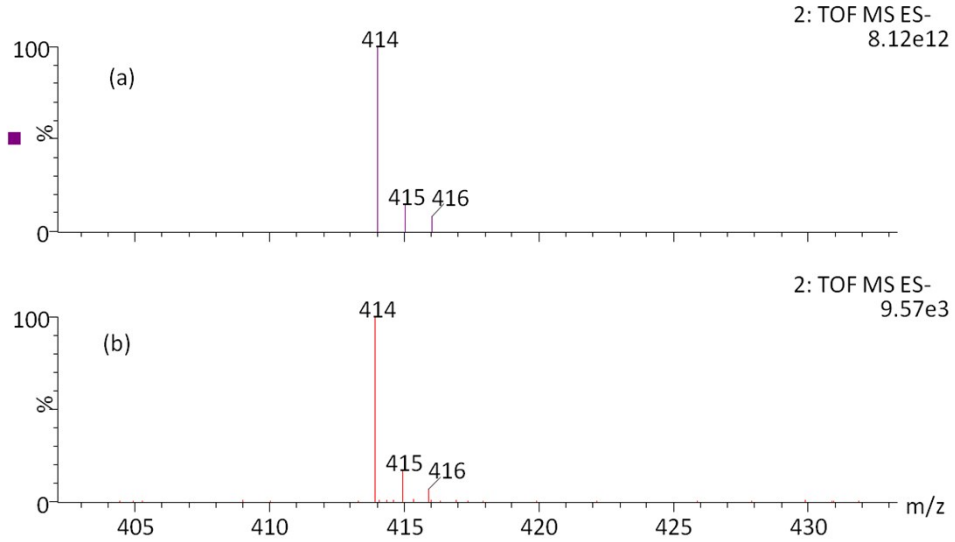


Figure S28. (a) The isotopic pattern of the signal at *m/z* 414 as compared to (b) the theoretical isotopic pattern of $C_{11}H_{18}N_2O_{11}PS$. Negative mode ESI mass spectrum acquired on the Waters LCT Q-TOF.