Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Structure Elucidation of the First Sex-Inducing Pheromone of a Diatom

*F. A. Klapper, C. Kiel, P. Bellstedt, W. Vyverman, G. Pohnert**

Table of Contents

Table of Figures

1 Experimental Procedures

Cultures and preparation of extracts

Sexually mature cells of isolates 85A and 84A of *Seminavis robusta* (DCG 0105, mating type + cells < 50 µm, DCG 0104, mating type - cells < 50 µm BCCM/DCG diatom culture collection, Ghent, Belgium) were cultivated in f/2 + Si medium[1] in T-25 culture flasks (Sarstedt, Nümbrecht, Germany) at 18 °C in a 12:12 h light:dark regime with fluorescent lamps at ~ 30 µmol photons m^{-2} s^{-1}. Cultures were reinoculated every week to keep them in exponential growth.

The sex inducing pheromone (SIP) was concentrated from cell free spent medium obtained after filtration under reduced pressure using GF/C filters (Whatman, UK) by solid phase extraction using hydrophilic lipophilic balanced (HLB) material (60 mg, Oasis®, Waters, Germany, 1 mL / 1 mg) following the instructions of the manufacturer. Elution was done with 30% MeOH and extracts were dried under nitrogen flow. All extracts were kept at -20 °C until further usage.

Bioassays

The bioactivity of the fractions was tested based on the induction of diproline production by SIP^+. Therefore, cultures of the MT^- mating type of *S. robusta* were prepared from exponentially growing stock cultures by inoculation of 2.5 mL into 22.5 mL fresh f/2 medium. Cultures were grown for 5 days and then dark-synchronized for 36 h. These cultures were treated with either the SIP+ containing medium or the purified fractions of the pheromone. Three concentrations corresponding to the initial concentration of the pheromone in MT+ cultures, a 1:10 dilution and a 1:100 dilution were tested in three replicates each. Deionized water was added to controls. Subsequently, the cultures were illuminated and incubated for 7 h. The cells were removed by filtration and 2.5 nmol caffeine dissolved in deionized water was added to each replicate as internal standard. The filtrate was extracted using solid phase extraction (HLB resin, 60 mg, Oasis HLB, Waters, Eschborn, Germany) as described in ref. 10 of the main manuscript. Extracts were dried under nitrogen flow, dissolved in 50 µL deionized water and analyzed by mass spectrometry as described in ref. 10 of the main manuscript. Concentration-dependent diproline detection was recorded as positive response indicative for the presence of SIP^+.

Labeling experiments

F/2 culture medium was prepared without nitrate and/or carbonate and filled in T-75 culture flasks (Sarstedt, Nümbrecht, Germany). $Na^{15}NO_3$ (sodium nitrate, ^{15}N, 98%+, Cambridge Isotope Laboratories, Inc., MA USA,) and/or $Na_2{}^{13}CO_3$ (sodium bicarbonate – ^{13}C, 99% Sigma Aldrich), were dissolved in aliquots of this medium and sterile filtered (0.2 µm pore size). These solutions were added to the culture flasks in amounts to reach concentrations as in the unlabeled medium. Filling the flasks to the neck to minimize CO_2 exchange increased the labelling degree. Cultures of 85A were inoculated and every five days, cells were transferred to new ^{13}C and/ or ^{15}N-enriched medium. *S. robusta* was cultured in labeled medium for 3–4 culturing cycles and medium samples of 60 mL each were taken at every time point to monitor the degree of labeling following HLB-extraction and MS-analysis as described above and below, respectively.

Incorporation of ^{13}C and ^{15}N in SIP was monitored by MS analysis of the isotopologues of *m*/*z* 842 u. The degree of labeling was calculated by manually extracting isotopologue intensities using the XCalibur software (Thermo Fisher Scientific, Waltham, MA, USA) and comparing them to computed distributions assuming the isotopologues to have a Bernoulli distribution[2]. The squared coefficient of variation between both was minimized in an iterative process.

Bernoulli equation:

$$\mathrm{I(m)} = \frac{n!}{m!\,(n-m)!} \cdot p^m (1-p)^{n-m}$$

For a metabolite with n carbon atoms the intensity of an isotopologue I(m) with m ^{13}C atoms is calculated using the degree of labelling p.

MS^n experiments

Dried samples of *S. robusta* extracts were dissolved in pure H_2O (LC/ MS grade, Chemsolute, Geyer GmbH, Germany) and measured on an UltiMate 3000 UHPLC (Thermo Fisher Scientific, Waltham, MA, USA) connected to a QEplus Orbitrap mass spectrometer (Thermo Fischer Scientific, Waltham, MA, USA) equipped with a HESI source.

Equipped with an C18 column (5x2.1 mm, 1.7 µm, Kinetex, Phenomenex) using H_2O + 2% MeCN + 0.1% FA (A) and MeCN + 0.1% FA (B) as mobile phase (LC/ MS grade, Chemsolute, Geyer GmbH, Germany), a 6.5 min method was run on the LC at a flow rate of 0.4 mL min^{-1}. The linear gradient started with 100% A for 0.2 min, increasing B to 48.7% till 4.0 min and reaching 100% B at 4.1 min. At 5.5 min B was decreased to 0%.

MS was operated in positive and negative ion mode with a resolving power of 280k. MS-MS experiments of (labeled) SIP$^+$ were performed within 2.2-3.0 min retention time applying 15, 25 and 35 normalized collision energies (NCE) (Table S 1). An *m*/*z* isolation window of 2.5 *m*/*z* with an offset of 1.0 *m*/*z* was chosen to detect the isotopologues. ^{15}N labeled SIP$^+$ was fragmented to determine the number of nitrogen atoms within the respective fragments and MS^3 data were acquired by in source fragmentation of SIP$^+$ with subsequent HCD of y_3–SO_3 and b_4 (Table S 1) to assign the smaller fragments to either the linear or the cyclic substructure.

Table S 1. Selected ions of SIP$^+$ (or fragments) for MS CID in positive and negative ion mode.

	m*/*z		**Energies**	
	Positive ([M+H–SO_3]$^+$)	Negative ([M–H]$^-$)	In source fragmentation (pos/neg eV)	HCD (NCE)
SIP$^+$	764.2801	842.2049	-	15, 25, 35
^{13}C SIP$^+$	775.2948	853.2388	-	15, 25, 35
^{15}N SIP$^+$	771.2543	849.1791	-	15, 25, 35
$y_3^{+/-}$-SO_3	376.1408	374.1205	60/80	15,20,30,40
$b_4^{+/-}$	389.2530	387.1167	60/80	15,20,30,40
red. SIP$^+$	-	844.2167		25, 35

Reduction of Disulfide Moiety

For verification of a disulfide bridge in SIP$^+$ a reduction with 1,4-dithiothreitol (DTT, molecular biology grade, VWR chemicals, USA) was performed. Therefore, an HLB extract of an 85A culture was treated with DTT (final concentration 50 mM) and reacted over night at room temperature before injection into the LC/MS.

Quantification of SIP$^+$

To estimate the amount of SIP$^+$ a calibration with five equidistant concentrations ranging from 2 nM to 10 nM (n = 3) was performed using a sulfated octapeptide ((Tyr[SO_3H]^{27}) Cholecystokinin fragment, 26-33 Amide, Sigma Aldrich, USA). We used this peptide assuming a similar ionization efficiency. 4-Fluorobenzoic acid was added as an internal standard (IS) with a concentration of 5 nM (Figure S1).

Triplicates of 85A cultures of *S. robusta* were grown for 4 days and three pictures of each replicate were taken with a Nikon DS-Fi2 CCD camera (Tokyo, Japan) attached to an inverted Leica DM IL LED light microscope (Heerbrugg, Switzerland) (10 x magnification) to determine cell density. Medium samples of 100 µL each were directly taken from the culture and filtered with 0.22 µm filter units (Millex-GV, Japan). After addition of IS to a final concentration of 5 nM, samples were analyzed without further purification. The calibration as well as the analyses of medium samples were performed on the LC-MS system (see "MS^n experiments") applying a UPLC method, starting with 2 min of pure water to ensure the elution of salts from the media before connecting to the MS. MS data were acquired in negative SIM.

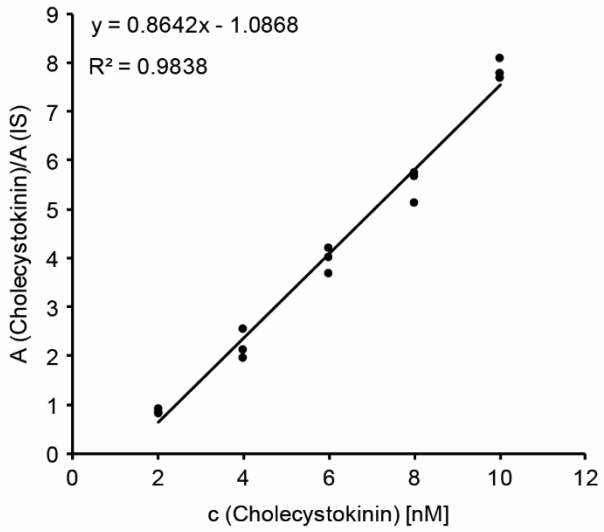

Figure S 1. Calibration curve for SIP$^+$ using Cholecystokinine as standard.

Further the minimal concentration of action of SIP$^+$ was determined by adding SIP$^+$ to cultures of mating type$^-$ from *S. robusta* (84A) and monitor the induced production of diproline by MS (*m*/*z* 195.1129 [M+H]$^+$ at 2.7 min) using the LC/MS method described above. Therefore, exponentially growing 84A (n = 4) were treated with SIP$^+$ to reach a final concentration of 100 fM, 10 fM, 1 fM and 100 aM. While the negative control was not treated, the positive control was performed by exchange of medium with 85A (mt$^+$) cell free spent medium (1:1, v/v). Medium samples were taken one day after treatment, filtered with 0.2 µm and directly analyzed with LC-MS.

Upscale and SIP$^+$ purification

A scale-up of *S. robusta* (85A, mt$^+$) culture was performed in 20 L culture bottles (Nalgene) with ^{15}N and ^{13}C enriched medium in 13 batches of 6 L (78 L in total). 400 mL of culture were inoculated, respectively and cultures were grown for 6 days. Subsequent filtration was done over GF/C filters (Whatman, VWR), the medium was acidified to 0.1% (v/v) formic acid (FA) and extracted using HLB cartridges (6 g, Oasis, Waters, Germany) following the instruction of the manufacturer. Cartridges were eluted with 150 mL 30% MeOH and dried extracts were kept at –20° C until further purification.

Extracts were dissolved in 4-6 mL H_2O to further purify SIP$^+$ on a Size Exclusion Chromatography system (ÄKTA pure, GE Healthcare, Sweden). Extracts were loaded on a peptide column (Superdex Peptide 10/300 GL, 10 x 310 mm, 24 mL bed volume, GE Healthcare; Uppsala, Sweden) and fractionated using pure H_2O as eluent with a flow rate of 0.25 mL min^{-1}. 300 µL of particle

free extracts were injected for each run and fractions of 1 mL were collected after 5.2 mL for ½ column volume (CV). Fractions 3+4+5 that contained the SIP-candidate were pooled, dried in a speedvac centrifuge and dissolved in 10 mL of H_2O. The column was eluted further with 1 CV (H_2O, 0.5 mL min^{-1}) and cleaned with 2 CV 20% EtOH every two days.
A final purification was performed on an UltiMate 3000 UPLC system (Thermo Fisher Scientific, Waltham, MA, USA) using a C18 column (Zorbax, Eclipse XDB-C18, 4.6x150 mm, 5 µm). A 15 min method using H_2O + 2% MeCN + 0.1% FA (A) and MeCN + 0.1% FA (B) with a flat gradient from 12 to 15% B in 7.5 min was applied to collect the pure SIP^+ fraction (4.0-4.5 min). An adjustable flow splitter was set to 1:20 (LC-MS : fraction collector, Analytical Scales & Services, Flanders, USA) to monitor the fractionation on a ISQ EC MS (Thermo Fisher Scientific, Waltham, MA, USA). All fractions containing SIP^+ were pooled and dried completely to yield < 200µg of labeled SIP^+ for NMR analysis.

NMR experiments

NMR data were acquired in D_2O in a 5 mm Shigemi tube on either a Bruker 500MHz Avance Neo equipped with a nitrogen cooled Prodigy BBO cryo probe or on a Bruker 600MHz Avance III equipped with a 5 mm triple resonance helium cooled cryo probe. The sample temperature was set to 297 K and all spectra were referenced to the residual solvent peak. To guide assignment of the ^{15}N, ^{13}C labeled substance, the following spectra were collected:

- 1D:
  - ^{1}H; ^{1}H with solvent suppression
  - $^{13}C\{^{1}H,^{15}N\}$
  - ^{13}C-APT
- 2D:
  - ^{15}N-^{1}H-HMBC
  - $^{1}H,^{13}C$-ASAP-HSQC-DEPT[3]
  - $^{1}H,^{1}H$-TOCSY with solvent suppression
  - $^{1}H,^{13}C$-HMBC
  - $^{1}H,^{13}C$-H2BC[4]

All NMR data were archived according to the FAIR guidelines using the LOGS scientific data management software (SIGNALS GmbH & Co KG, Frankfurt/Main, Germany).

Marfey's Analysis

Extracts of 18 L of spent medium from *S. robusta* were purified according to the methods described above. Less than 0.1 mg of SIP^+ were hydrolyzed with 100 µL of 6 N HCl at 110° C for four hours. The reaction was quenched by adding 30 µL of 1 M $NaHCO_3$. The dried sample was dissolved in 30 µL H_2O and 20 µL of $NaHCO_3$. The amino acid solution was treated with 100 µL 1-fluoro-2-4-dinitrophenyl-5-L-alanine amide (L-FDAA, 1% in acetone) and incubated for 2 hours at 37 °C. The reaction was quenched with 20 µL 1 M HCl. [5]
Standards of D-Ala, L-Ala, D-Leu, L-Leu, D-Cys, L-Cys, D-*cis*-Hyp, L -*cis*-Hyp, L -*trans*-Hyp, D-*trans*-Hyp, D/L-*threo*-β-OH-Asp and D/L-*erythro*-β-OH-Asp[1] were derivatized with L-FDAA according to the method described above.
Derivatives were analyzed on an UPLC-MS System (see "MS^n-experiments") using a C18 column (100x2.1 mm, 2.1 µm, Accucore C18, Thermo Scientific) with eluents H_2O + 2%MeCN + 0.1% FA (A) and MeCN (B) and a linear gradient: 10-50% B from 0-25 min, 50-100% B from 25-26 min,100% B from 26-28 min and 10% B from 28-30 min. The amino acids were assessed by UV

[1] Kindly provided by Ron Hermenau

(340 nm) and MS and the HR-mass of the standard and the respective free amino acid from SIP^+ was compared for the specific retention time.

2 Results

Labeling experiments

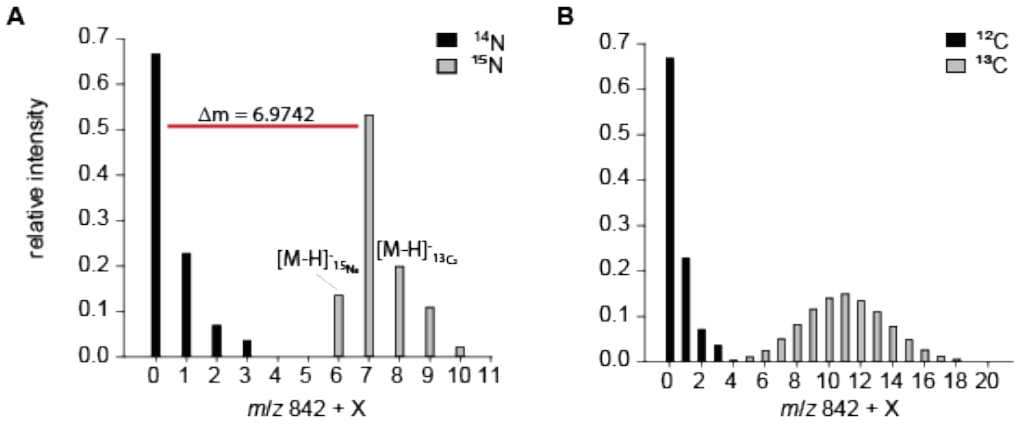

Figure S 2. Isotopic pattern of labelled SIP$^+$. A) SIP$^+$ shows almost complete labelling of nitrogen when cultured in 15N enriched medium, whereas B) ^{13}C labelling results in a highly but not quantitatively labelled SIP$^+$.

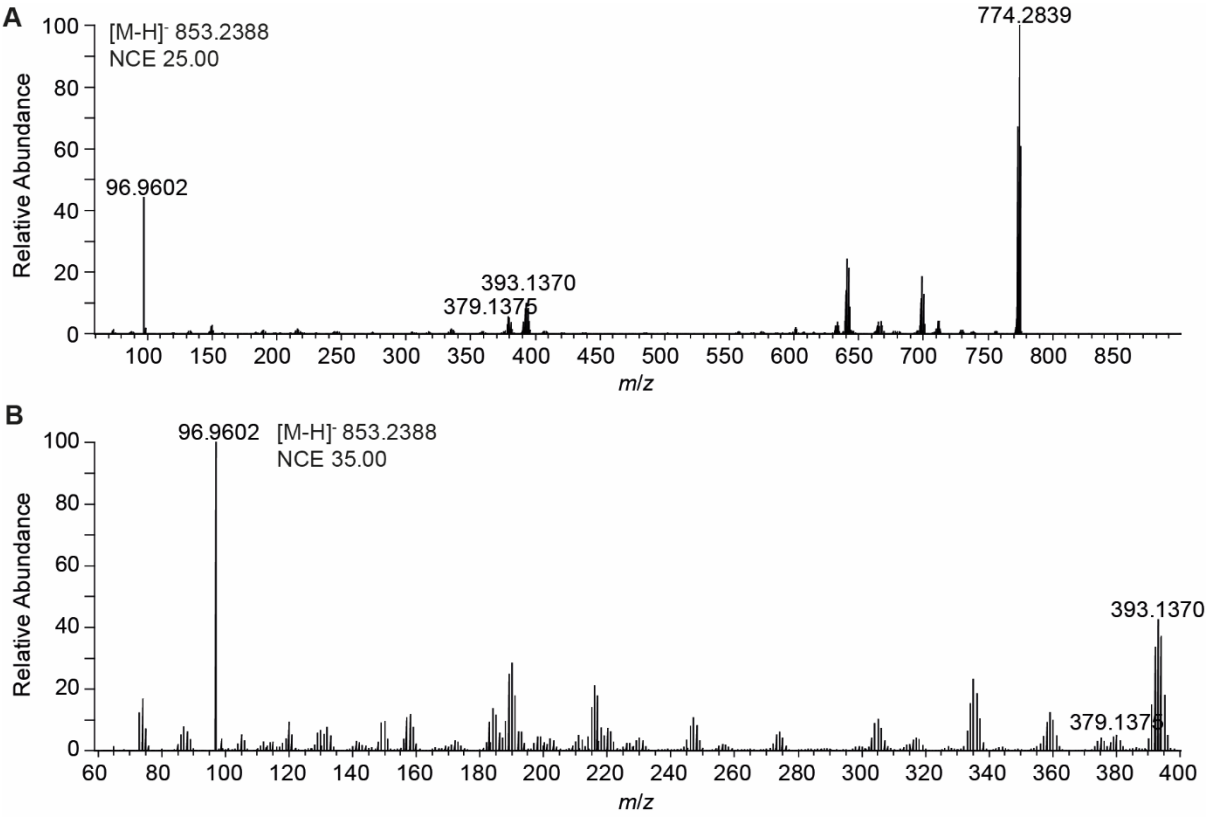

Figure S 3. ESI CID spectra of ^{13}C labeled SIP$^+$ measured in negative ion mode with different collision energies. *m/z* 853.2388 was fragmented with *m/z* window of 8.0 to detect the isotopic pattern of the fragments. The similarity of the isotopic patterns of the single fragments indicate an equal distribution of

^{13}C within SIP^+. A) 25 NCE. B) 35NCE.

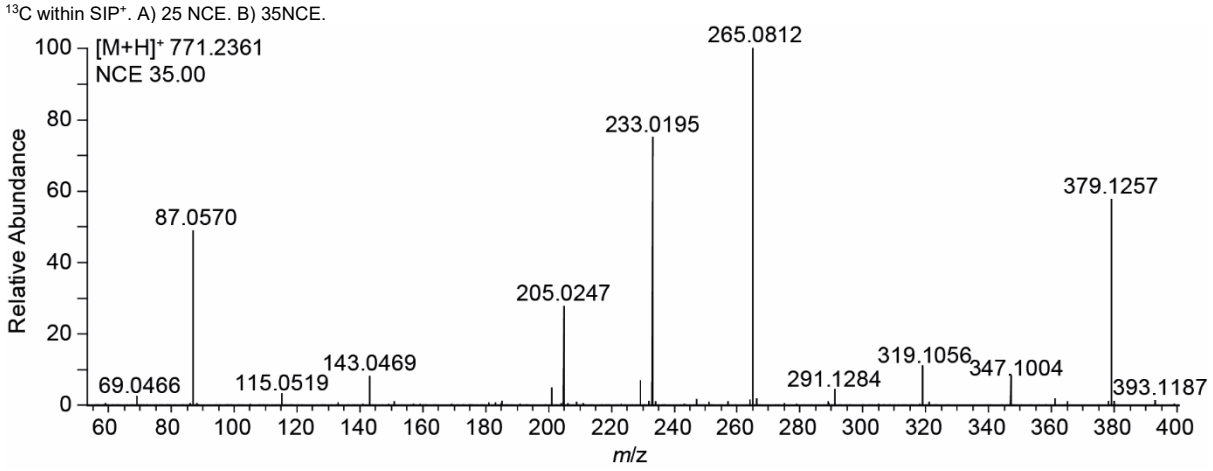

Figure S 4. ESI CID spectra of fully ^{15}N labeled SIP^+ measured in positive ion mode. The number of nitrogen atoms within the fragments was determined by the increase of the mass of the respective fragment.

MS^n experiments

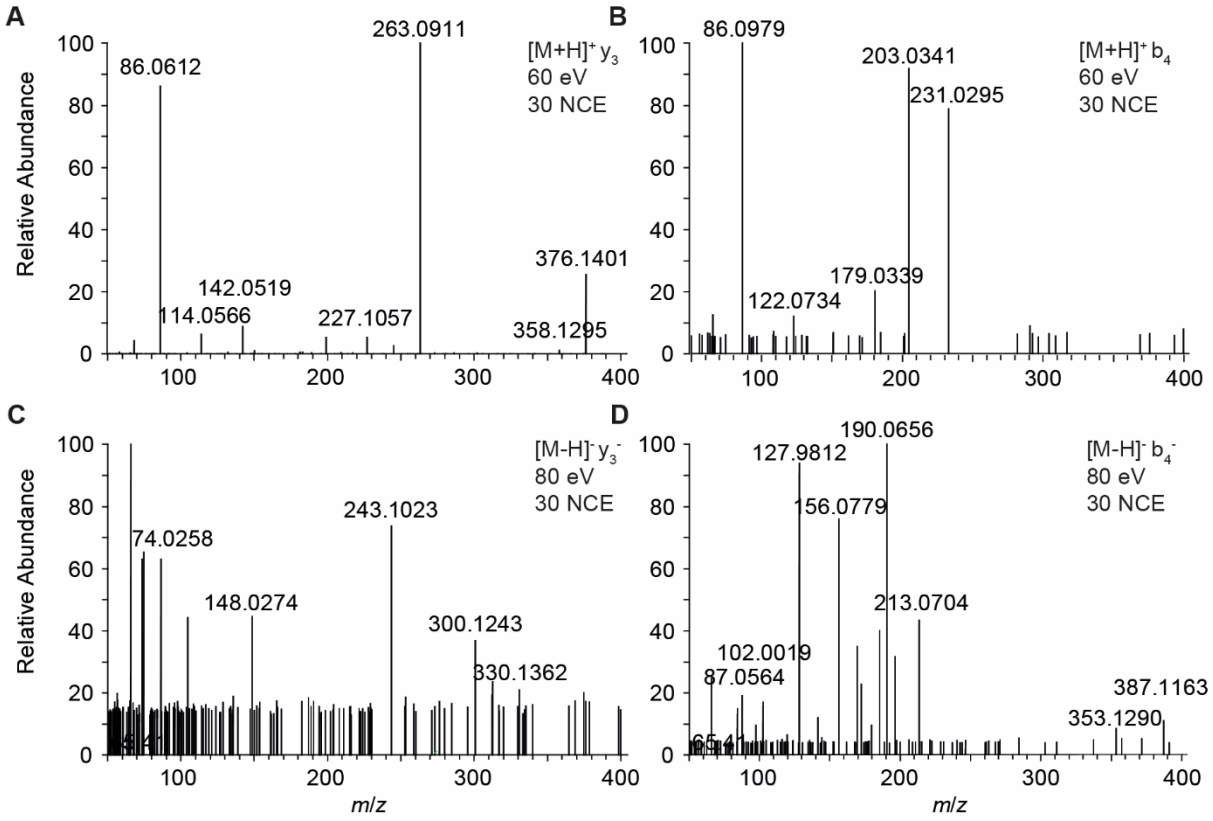

Figure S 5. Pseudo MS^3 spectra of SIP^+ conducted by in source fragmentation with subsequent CID of the respective fragments. A, B) y_3–SO_3 and b_4 were obtained in positive and C, D) negative ion mode to assign the smaller fragments to either the cyclic or the linear part of the peptide SIP^+. **Reduction**

of S-S bridge

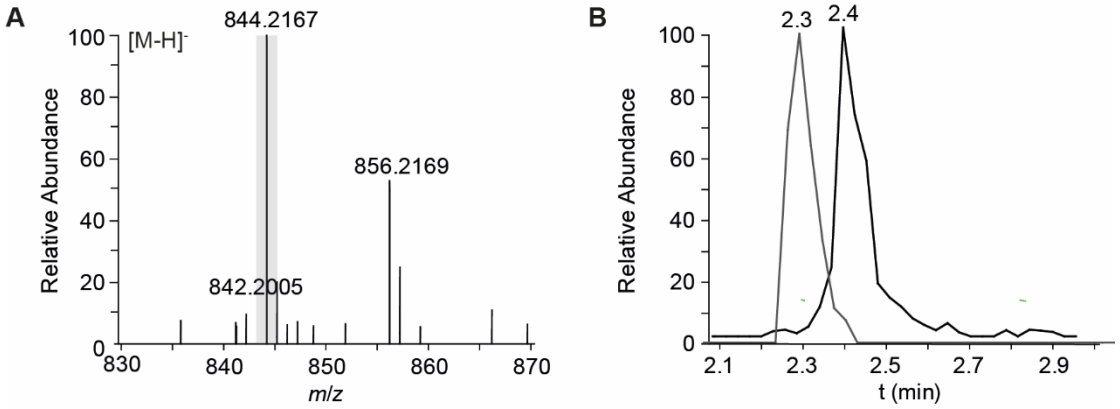

Figure S 6. Reduction of SIP⁺ with dithiotreithol (DTT) to prove the S-S bond in SIP⁺. A) The molecular mass of SIP⁺ increased by 2 Da while B) the retention time decreased (grey) due to a higher polarity of free cysteines.

NMR Experiments

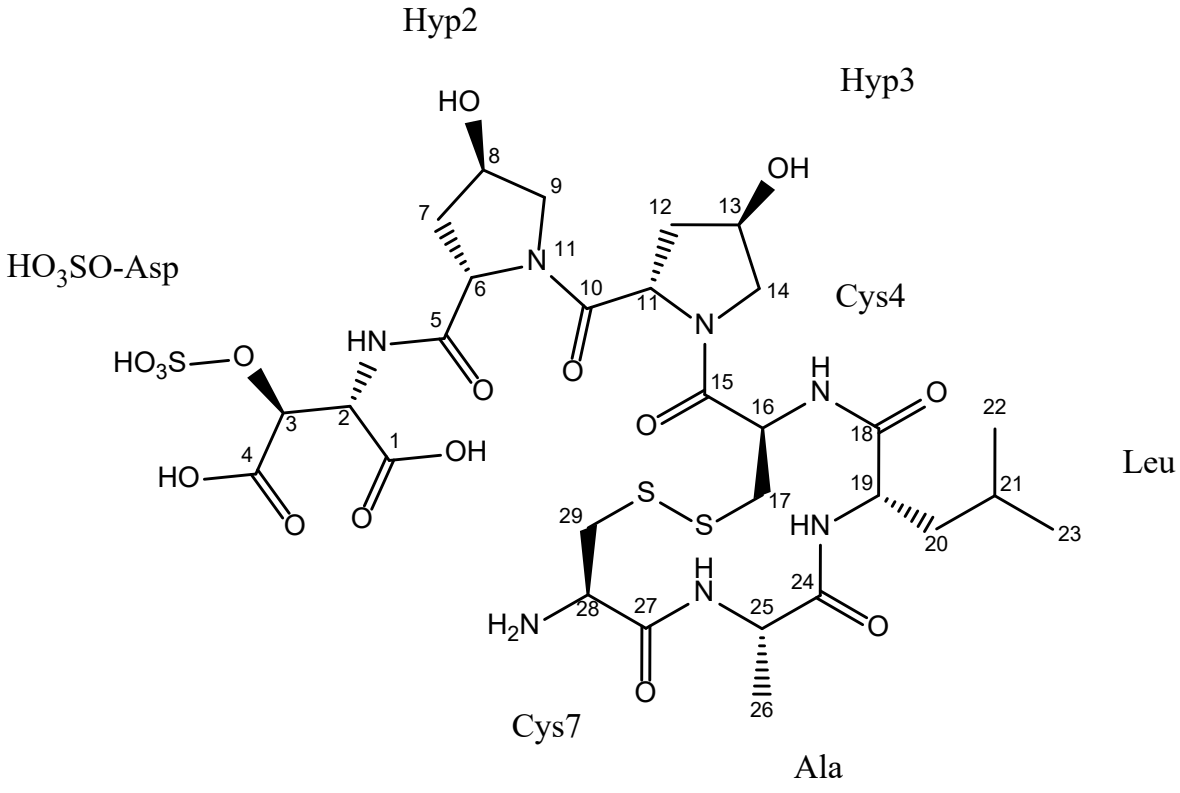

Figure S 7. Absolute configuration of SIP+.

Table S 2. ^{13}C (HSQC-DEPT, 600 MHz), ^{1}H (HSQC-DEPT, 600 MHz) and ^{15}N (HMBC, 600 MHz) NMR assignment for SIP⁺ in D_2O.

A. A. unit	position	δ_C [ppm]		δ_H [ppm]	δ_N [ppm]
Hya 1	1	176.3	C		116.2
	2	59.95	CH	4.6	

	3	81.1	CH	5.00	
	4	177.3	C		
	5	175.75	C		
	6	62.0	CH	4.65	
Hyp 2	7	39.6	CH_2	2.2/2.4	130.5
	8	72.7[a]	CH	4.64	
	9	57.9	CH_2	3.8/3.9	
	10	174.75	C		
	11	60.3	CH	4.9	
Hyp 3	12	38.9	CH_2	2.1/2.4	132.25
	13	72.7[a]	CH	4.64	
	14	57.1	CH_2	3.9	
	15	172.6	C		
Cys 4	16	57.4	CH	4.87	119.9
	17	37.9	CH_2	3.07/3.17	
	18	176.7	C		
	19	54.5	CH	4.45	
Leu 5	20	41.1	CH_2	1.7	115.9
	21	27.2	CH	1.6	
	22/23	23.2, 25.2	CH_3	0.9, 0.95	
	24	178.5	C		
Ala 6	25	54.6	CH	4.3	127.7
	26	18.8	CH_3	1.45	
	27	177.4	C		
Cys 7	28	55.3	CH	3.7	31.3, 31.7
	29	43.5	CH_2	3.35/3.15	

[a]signals were not resolved

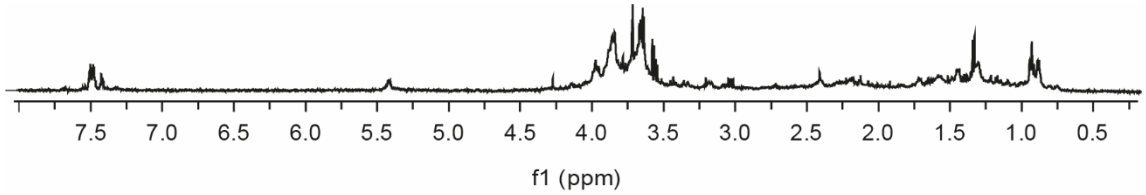

Figure S 8. ^{1}H NMR (500 MHz, D_2O, with solvent suppression) of SIP^+. The signal at 7.5 ppm is an impurity that is found in this fraction. Other fractions not containing the signal were active in bioassays as well, indicating that it is not relevant for biological activity. The signals around 3.5 to 4.0 ppm are introduced during purification through trace contaminations in the solvent or column. Blanks, containing this signal but not the SIP did not show activity, even in elevated concentrations.

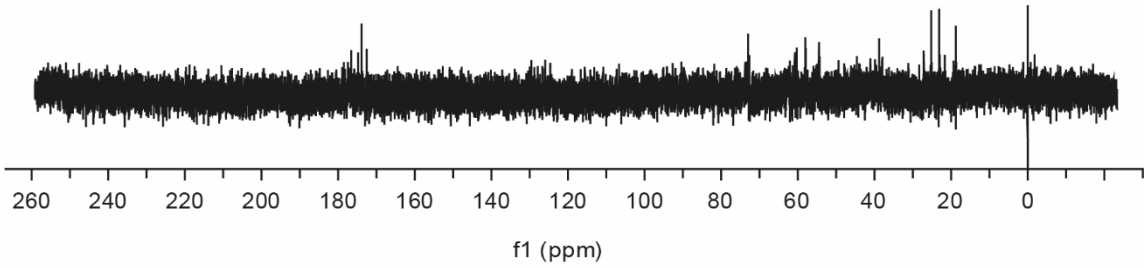

Figure S 9. ^{13}C {^{1}H} NMR (500 MHz, D_2O) of SIP^+.

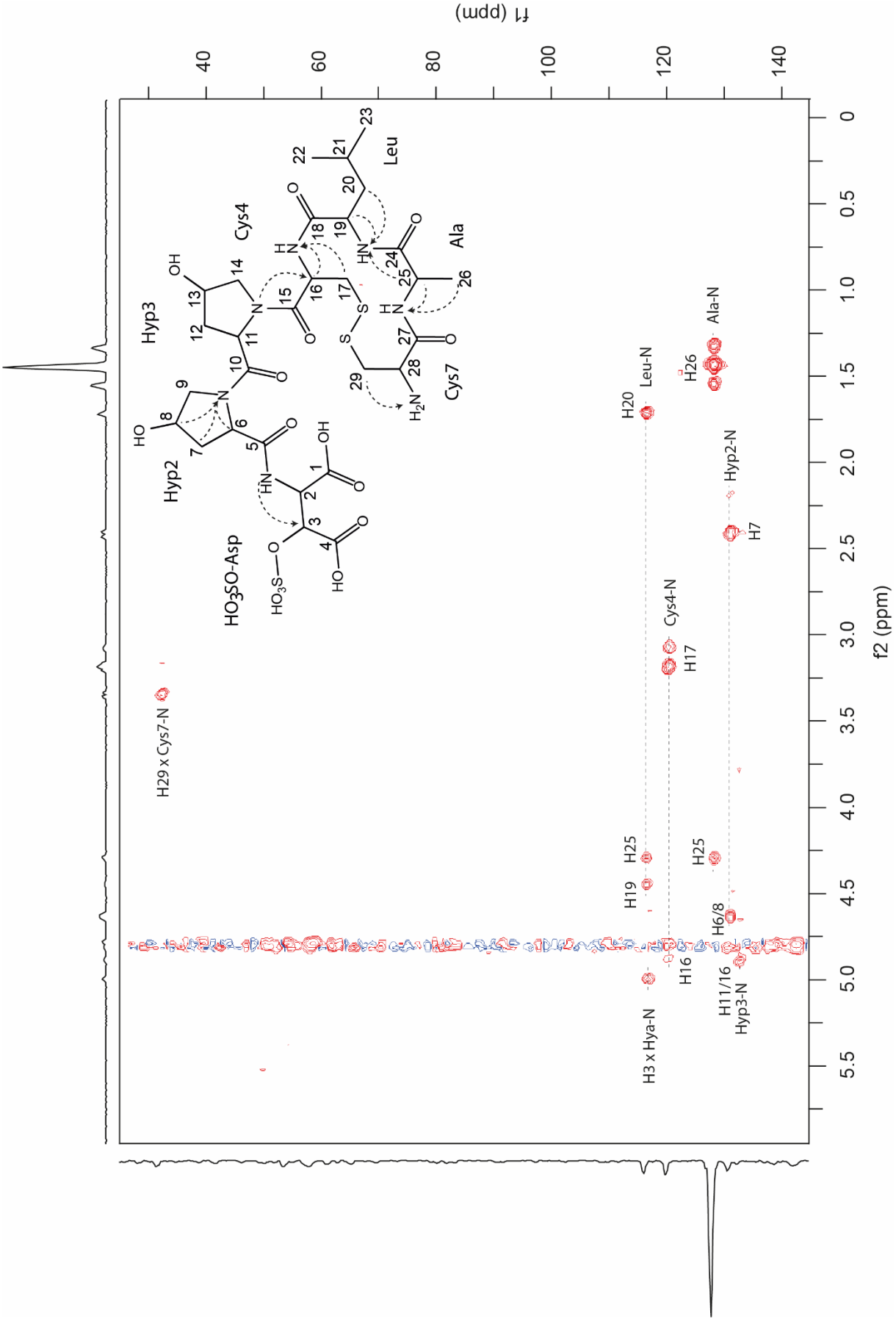

Figure S 10. $^{1}H,^{15}N$ HMBC (600 MHz, D_2O) spectrum of $SIP^{+\bullet}$.

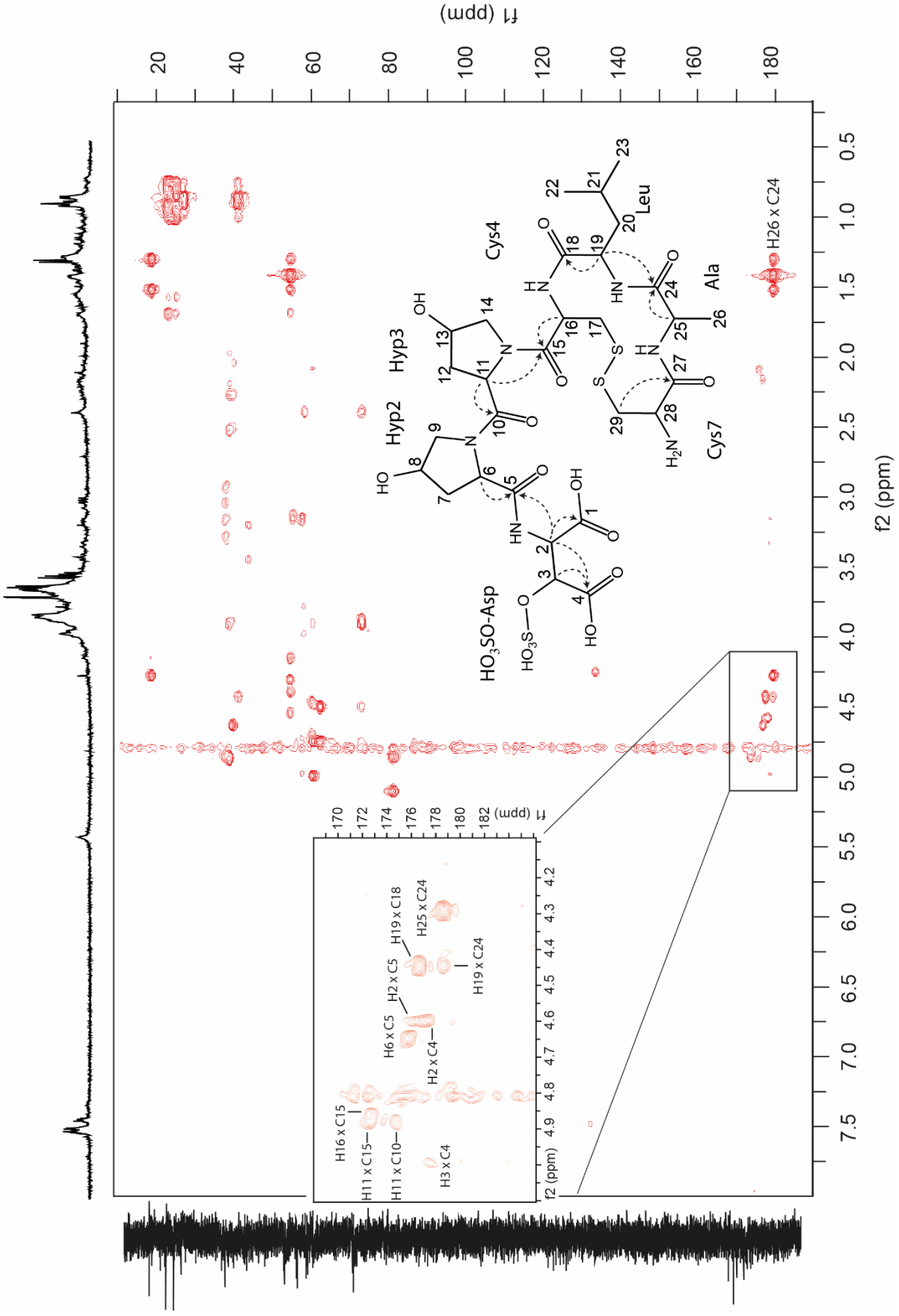

Figure S 11. $^{1}H,^{13}C$ HMBC NMR (600 MHz, D_2O) of SIP^{+}. Black box represents area for zoomed plot representing important correlations for structure elucidation.

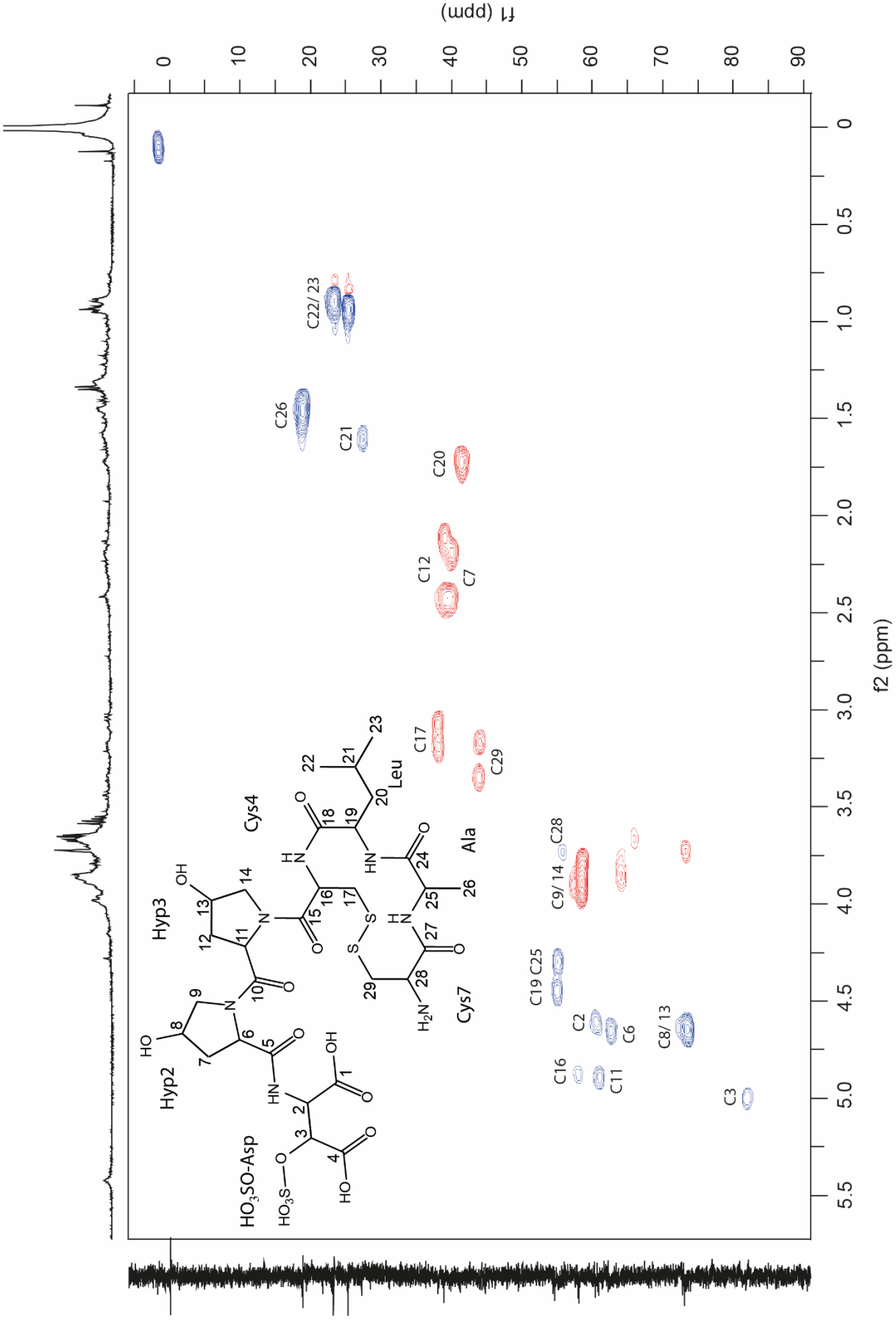

Figure S 12. $^{1}H,^{13}C$-ASAP-HSQC-DEPT NMR (600 MHz, D_2O) of SIP^{+}.

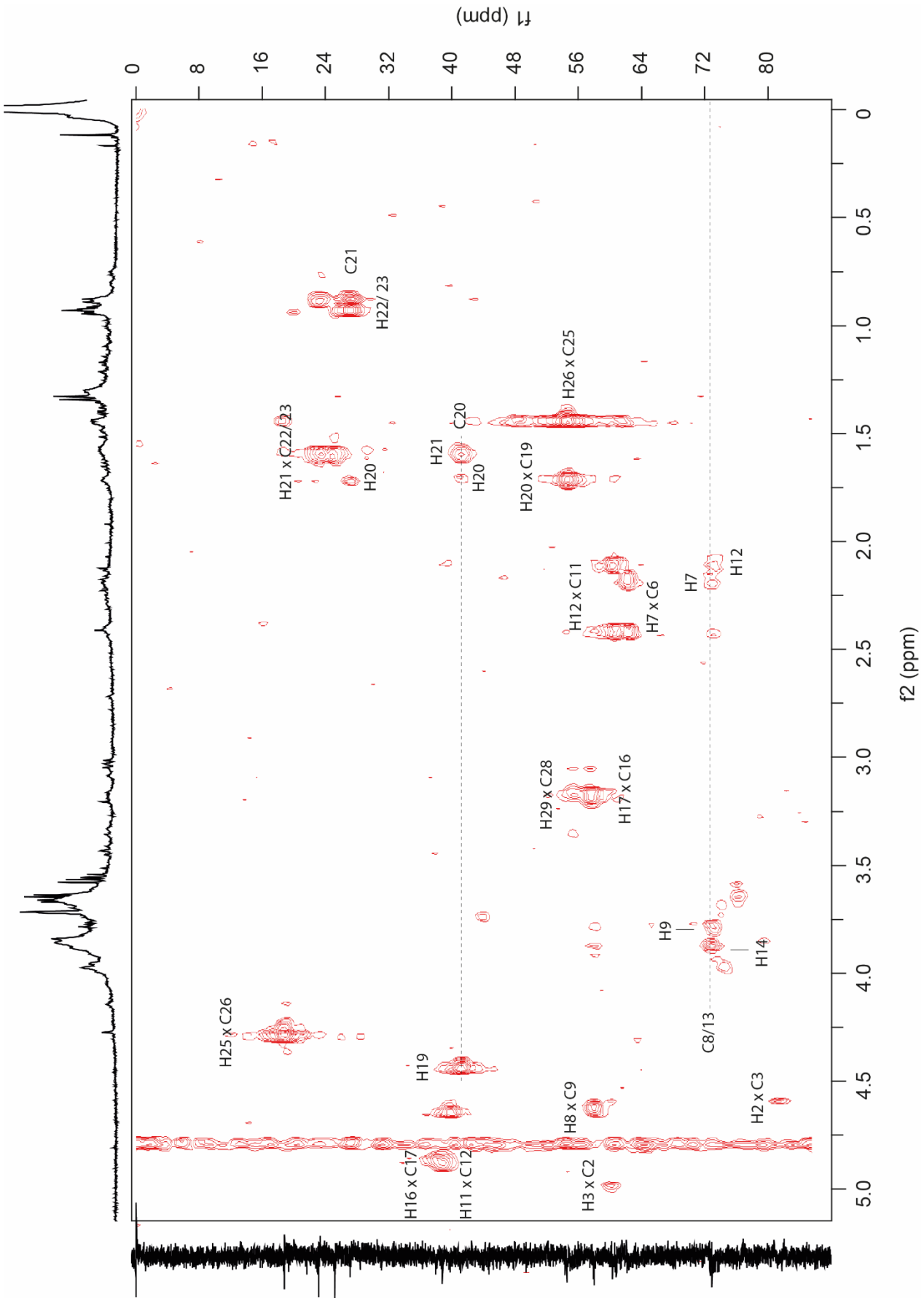
f1 (ppm)
0
8
16
24
32
40
48
56
64
72
80
f2 (ppm)
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
C21
H22/ 23
H26 x C25
H21 x C22/ 23
H20
H21
C20
H20
H20 x C19
H12 x C11
H7 x C6
H7
H12
H29 x C28
H17 x C16
H9
H14
C8/13
H25 x C26
H19
H8 x C9
H2 x C3
H16 x C17
H11 x C12
H3 x C2

Figure S 13. $^{1}H,^{13}C$ H2BC NMR (600 MHz, D_2O) of SIP^{+}

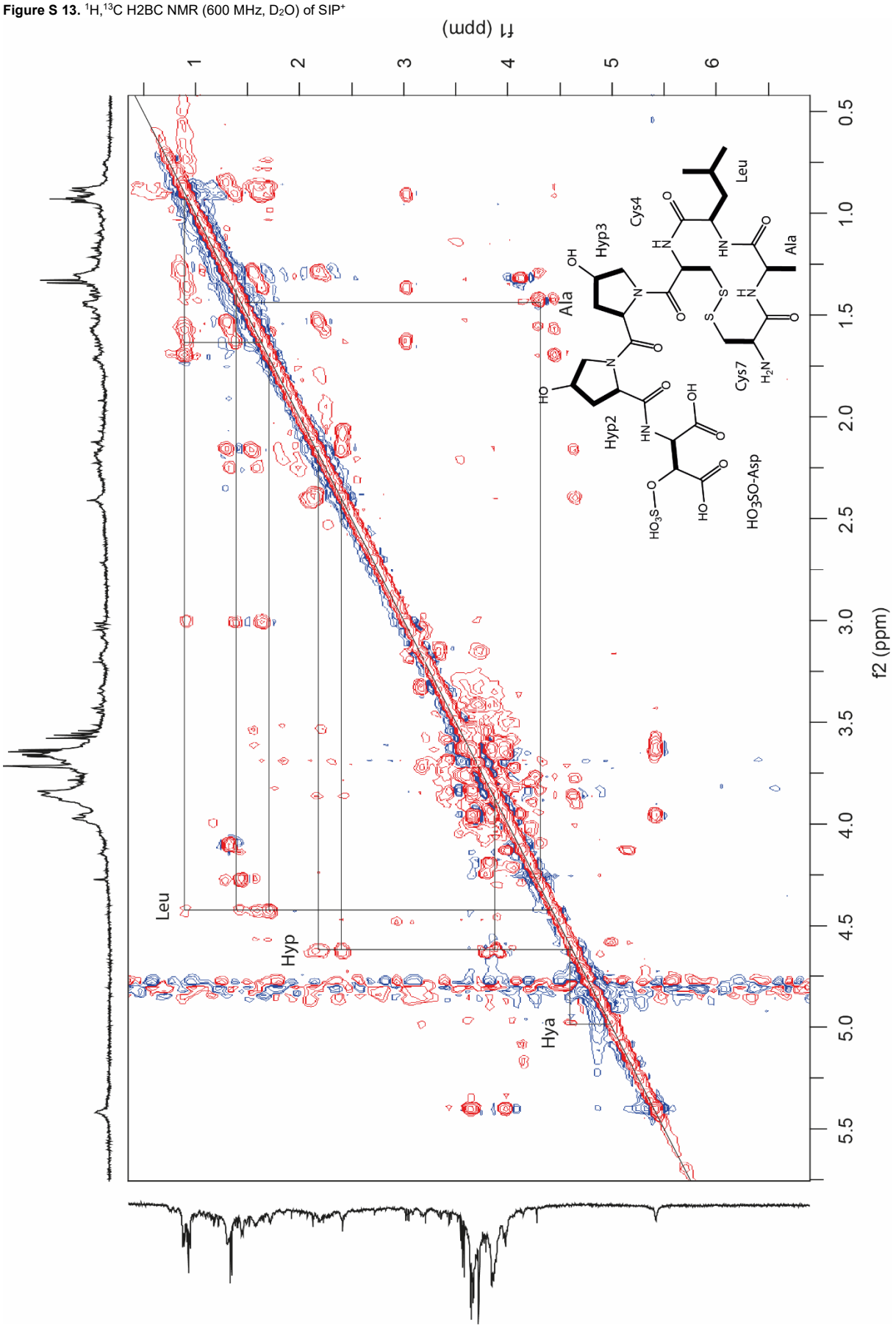

Figure S 14. $^{1}H,^{1}H$ TOCSY NMR (500 MHz, D_2O, with solvent presaturation) of SIP^+.

Marfey's analyses

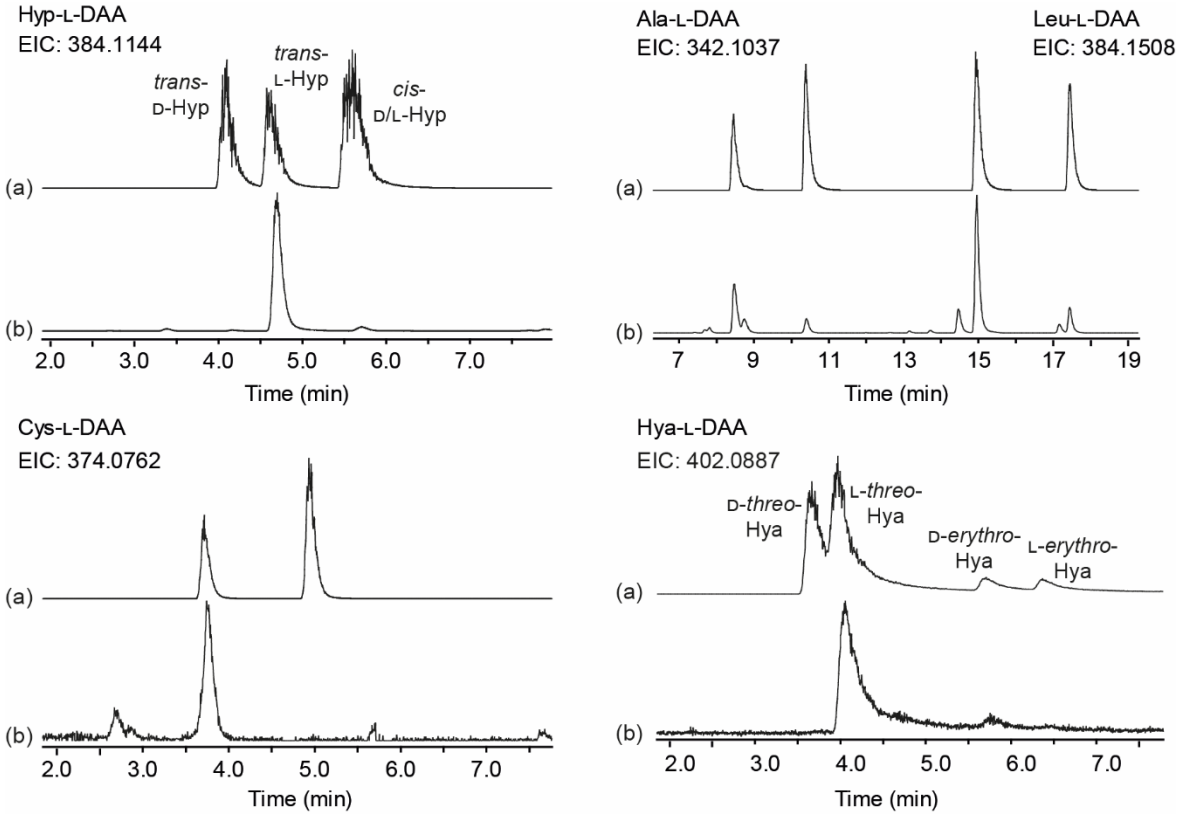

Figure S 15. Marfey's analysis for SIP^+. The LC-MS profiles depict the amino acid standards (a) and the respective extracted ion chromatogram of the free amino acids (b) gained from hydrolyzed SIP^+. If not stated otherwise the elution order of the amino acids are L -> D.

Bioactivity of SIP^+

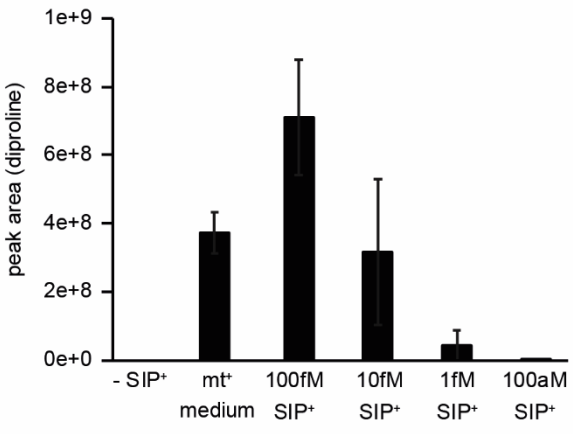

Figure S 16. Bioactivity of SIP^+. The minimal active concentration of SIP^+ was determined by adding SIP^+ to cultures of mt^- and detect the induced diproline production by LC-MS. The positive control was medium from a five day old mt^+ culture.

Acknowledgements

The authors appreciate Annika Brauns' and Alice Ormes' experimental assistance. We thank Sien Audoor and Darja Belišová for help with diatom culturing. Further, we thank Tim Baumeister for the valuable discussion on experimental design and Christoph Wiedemann and Nico Ueberschaar for advice on NMR and MS data interpretation. The authors are grateful to Johannes Rassbach for practical support and highly acknowledge Ron Hermenau for providing the standards of *threo* and *erythro* hydroxy aspartic acid.

References

[1] R. R. L. Guillard, in *Culture of Marine Invertebrate Animals: Proceedings — 1st Conference on Culture of Marine Invertebrate Animals Greenport* (Eds.: W. L. Smith, M. H. Chanley), Springer US, Boston, MA, **1975**, pp. 29-60.
[2] N. Meyer, A. Rydzyk, G. Pohnert, *Frontiers in Marine Science* **2022**, *9*.
[3] D. Schulze-Sünninghausen, J. Becker, M. R. M. Koos, B. Luy, *J. Magn. Reson.* **2017**, *281*, 151-161.
[4] B. O. Petersen, E. Vinogradov, W. Kay, P. Würtz, N. T. Nyberg, J. Ø. Duus, O. W. Sørensen, *Carbohydr. Res.* **2006**, *341*, 550-556.
[5] K. Fujii, Y. Ikai, T. Mayumi, H. Oka, M. Suzuki, K.-i. Harada, *Analytical Chemistry* **1997**, *69*, 3346-3352.

Author Contributions

Franziska Klapper:	conception of studies, experimental planning, data collection, data analysis and interpretation, manuscript writing
Christine Kiel:	experimental planning
Peter Bellstedt:	data collection, help with NMR data interpretation
Wim Vyverman:	conception of the studies, correction of the manuscript
Georg Pohnert:	conception of studies, experimental planning, writing of the manuscript