Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Biosynthesis of Nodulisporic Acids: A Multifunctional Monooxygenase Delivers a Complex and Highly Branched Array

A. T. Richardson, R. C. Cameron, L. J. Stevenson, A. J. Singh, Y. Lukito, D. Berry, M. J. Nicholson, E. J. Parker**

Table of Contents

Table of Contents **2**
General **4**
Molecular Biology and Fungal Work **5**
Table S1: Fungal transformant list 6
Table S2: Components of DNA constructs used in this study 7
Sequences **9**
Promoters used in this study 9
Coding sequences used in this study (gDNA) 11
Genes used in this study (cDNA) 16
Terminators used in this study 20
LC-MS Analysis of Fungal Transformants **22**
Identification of Nodulisporic Acids by Tandem Mass Spectrometry **43**
Figure S1. Proposed fragmentation of nodulisporic acids, NAD_4 and NAD_4-OMe 43
Table S3: HR-ESI-MS^2 analysis of nodulisporic acids 45
Isolation of Nodulisporic Acids **66**
Isolation of NAD_4 67
Table S4. NMR Data for NAD_4 and NAD_4-OMe, $CDCl_3$ 68
Isolation of NAD & NAD_1 69
Table S5: NMR Data for NAD, $CDCl_3$. 70
Table S6: NMR Data for NAD_1, $CDCl_3$. 71
Isolation of NAD_2 72
Table S7: NMR Data for 18-23 alkene-NAD_2, $CDCl_3$. 73
Isolation of DH-NAA_4 & DH-NAC_4 74
Table S8: NMR Data for DH-NAA_4, $CDCl_3$. 75
Table S9: NMR Data of DH-NAC_4, $CDCl_3$. 76
Isolation of NAC 77
Table S10. NMR Data for NAC, $CDCl_3$. 78
Isolation of NAB_4 79
Table S11: NMR Data for NAB_4 and its 2'-epimer, $CDCl_3$. 79
Isolation of NAA_1 and NAB 80
Table S12: NMR Data for NAA_1, CD_3OD. 81
Table S13: NMR Data for NAB, $CDCl_3$. 82
Isolation of NAD_6 83
Table S14: NMR Data for NAD_6, $(CD_3)_2CO$. 84
Isolation of NAD_5 85
Table S15: NMR Data for NAD_5, $(CD_3)_2CO$ 86
NMR Spectra **87**
NAD NMR Spectra 87
NAD_1 NMR Spectra 91

NAD$_4$ NMR Spectra ... 96
NAD$_4$-OMe NMR Spectra ... 102
NAD$_2$ NMR Spectra ... 107
NAC NMR Spectra ... 111
DH-NAC$_4$ NMR Spectra ... 115
NAB$_4$ NMR Spectra ... 119
DH-NAA$_4$ NMR Spectra ... 122
NAA$_1$ NMR Spectra ... 126
NAB NMR Spectra ... 131
NAD$_6$ NMR Spectra ... 135
NAD$_5$ NMR Spectra ... 139
***In Vivo* Feeding Studies ... 143**
Table S16. NodJ *in vivo* feeding studies ... 143
Figure S58. LC-MS analysis of in vivo feeding studies. ... 144
Figure S59. NAD$_{5/1}$ feeding study ... 145
Non-enzymatic NAD$_6$ elimination ... 145
Proposed NodJ P450 Monooxygenase Bifurcation Mechanism ... 146
Figure S60. Proposed mechanism for NodJ ... 146
Phylogenetic Classification of NodX ... 147
Figure S61. Phylogenetic classification of NodX ... 147
Nodulisporic Acid Biosynthesis Scheme ... 148
Figure S62. Nodulisporic acid biosynthesis ... 148
Author Contributions ... 149
References ... 149

General

NMR spectroscopy was conducted on a JEOL JNM-ECZ600R with a nitrogen cooled 5 mm SuperCOOL cryogenic probe (600 MHz for ^{1}H nuclei and 150 MHz for ^{13}C nuclei). The residual solvent peak was used as an internal reference for ^{1}H [δ_H 3.31, CD_3OD; 7.26, $CDCl_3$; 2.05, $(CD_3)_2CO$] and ^{13}C [δ_C 49.00, CD_3OD; 77.16, $CDCl_3$; 29.84, $(CD_3)_2CO$] chemical shifts.

Flash chromatography was carried out on a Buchi Reveleris X2 using Buchi FlashPure 20 μm silica cartridges and a $CHCl_3$: MeOH gradient (1–100%).

Semi-preparative HPLC was carried out on an Agilent 1260 Infinity II HPLC system with DAD. In all cases the mobile phase was A: H_2O, B: MeCN both containing 0.1% formic acid. Purification was achieved using either a Phenomenex Luna C18 250 × 10 mm 100 Å 5 μ column (column #1), or an Agilent Zorbax SB-C18 50 × 9.4 mm 5 μm column (column #2).

Preparative HPLC was performed on an Agilent 1260 Infinity II Preparative HPLC system, equipped with multi-wavelength and evaporative light-scattering detection, using A: H_2O and B: MeCN as the mobile phases both containing 0.1% formic acid. Purification was performed using Column #1 (see above).

Reversed-Phase Prefractionation HP-20 poly(styrene-divinylbenzene) (Supelco, Sigma) was used for metabolite capture and reversed-phase pre-fractionation. Acetone and MeOH were analytical grade or higher. H_2O was purified by reverse osmosis. Solvent compositions are reported as % v/v unless otherwise stated.

Thin-Layer Chromatography (TLC) was carried out using Merck KGaA Silicagel 60 F_{254} plates and a mobile phase of $CHCl_3$:MeOH (9:1). Indole diterpenes were identified by staining with Ehrlich's reagent (*p*-dimethylaminobenzaldehyde).

Liquid Chromatography-Mass Spectrometry (LC-MS) was completed on an Agilent 1260 Infinity II LC-MS system with DAD and electrospray ionisation. A Phenomenex C18 Kinetex column (2.6 μ, 100 Å, 50 × 2.1 mm) equipped with a Phenomenex C18 guard cartridge and maintained at 40 °C was eluted with a mobile phase of A: H_2O and B: MeCN, both containing 0.1% formic acid. An injection volume of 10 μL and flow rate of 0.4 mL/min were used. The gradient was as follows: 0–1 min 40% B, 1–5 min 40–60% B, 5–24 min 60–90% B, 24–25 min 90–100% B, 25–27 min 100% B.

High-resolution mass spectrometric data and MS/MS spectra were obtained with an Agilent 6530 Accurate Mass Q-TOF fitted with an electrospray ion source and equipped with an Agilent 1260 Infinity II LC system. Chromatography was carried out using an Agilent Accucore C18 2.0 μm 50 × 2.1 mm column eluted with a mobile phase of A: H_2O and B: MeCN, both containing 0.1% formic acid. The flow rate was 0.3 mL/min and injection volume 5 μL. The gradient used was as follows: 0–1 min 40% B, 1–30 min 40–100% B, 30–35 min 100% B. The mass spec parameters used were: positive ion mode, mass range 100–1000 Da, acquisition rate 2 scans/s, capillary temperature 300 °C, capillary voltage 3500 V, fragmentor voltage 175 V, drying gas flow 8 L/min, sheath gas temp 350 °C, sheath gas flow 11 L/min, and a nebulizer pressure of 35 psi. MS/MS data were acquired for selected masses using CID with an isolation width of *M* 1.6 and collision energy of 30 eV.

Molecular Biology and Fungal Work

The following protocols were completed as described in van Dolleweerd *et al* 2018.[1]

- Molecular biology.
- Bacterial and fungal strains.
- Construction of DNA constructs using the MIDAS cloning system.
- Protocols for MIDAS Level-1 module cloning.
- Protocols for MIDAS Level-2 TU assembly.
- Protocols for MIDAS Level-3 multigene assembly.
- Media and reagents used for fungal work.
- Fungal Protocols – Protoplast Preparation.
- Fungal Protocols – Transformation of *P. paxilli.*

Synthetic DNA MIDAS Level-1 constructs. MIDAS Level-1 plasmids containing cDNA versions of NOD genes were synthesised, cloned into MIDAS Level-1 vector pML1, and sequence verified by Twist Biosciences.

Indole diterpene production and extraction. Fungal transformants were grown in 25 mL of CDYE medium (recipe below) with trace elements (recipe below) for 7 days at 28 °C in shaker cultures (≥ 200 rpm), in 125 mL Erlenmeyer flasks capped with cotton wool. Mycelia were isolated from fermentation broths by filtration. IDTs were extracted from an 850 mg sample of mycelia via homogenisation with 500 µL EtOAc in a bead beating apparatus (MPBio FastPrep24 5G bead beater grinder and lysis system, 40 s, 6m/s). The extract was recovered by centrifuging the homogenised sample at 17,000 rcf for 10 minutes and solvent removed in vacuo using a speedvac (Labconco Centrivap DNA concentrator). Extracts were resuspended in MeCN (150 µL), filtered using 0.2 µm syringe filters prior to LC-MS analysis

Large scale fungal growths. Fungal transformants were grown in a YEPGA starter culture (25 mL media in a 125 mL Erlenmeyer flask) for 15 h and subsequently used to inoculate (4% v/v) CDYE production cultures (400 mL media in a 2 L Erlenmeyer flask) which were grown for 7 days at 28 °C in shaker cultures (≥ 200 rpm). Mycelia was isolated from the fermentation broths by vacuum filtration.

CDYE media with trace elements prepared as follows: Czapek dox (34 g/L, Oxoid, CM0095), yeast extract (5 g/L, Oxoid LP0021) and trace elements (5 mL/L).

Trace elements prepared as follows: $FeSO_4.7H_2O$ (1.7 mM), $ZnSO_4.7H_2O$ (1.73 mM), $MnSO_4.H_2O$ (0.59 mM), $CuSO_4.5H_2O$ (0.2 mM), $CoCl_2.6H_2O$ (0.17 mM), HCl (0.8 M).

YEPGA media prepared as follows: Yeast extract (10 g/L, Oxoid LP0021), mycological peptone (20 g/L, Oxoid LP0040), glucose (20 g/L, Sigma Aldrich G8270), Agar (4 g/L, Invitrogen 30391-023), trace elements (20 mL/L), and adjusted to pH 6.0 using NaOH.

RNAseq Methodology. Crushed mycelial discs taken from colonies of *H. pulicicidum* strain MF5954 (ATCC 74245) grown on YM agar (3 g/L yeast extract (Oxoid), 3 g/L malt extract, 10 g/L dextrose, 5 g/L peptone, 20 g/L agar) were used to inoculate a seed culture in a 250 mL Erlenmeyer flask containing 50 mL SL3 medium (50 g/L glucose, 10 g/L L-glutamic acid monosodium salt, 2 g/L Amicase® casein acid hydrolysate (Sigma), 3 g/L NH_4Cl, 1 g/L K_2HPO_4, 2 g/L lactic acid, 20 g/L 2-(*N*-morpholino)ethanesulfonic acid, 0.5 g/L $MgSO_4{\bullet}7H_2O$, 1 g/L $CaCO_3$, and 20 mL/L of a 50× trace element solution (40 mg/L $CoCl_2{\bullet}6H_2O$, 50 mg/L $CuSO_4{\bullet}5H_2O$, 500 mg/L $FeSO_4{\bullet}7H_2O$, 100 mg/L $MnSO_4{\bullet}H_2O$, 500 mg/L $ZnSO_4{\bullet}7H_2O$)). The seed culture was grown in total darkness on an orbital shaker at 220 RPM for 2 days at 29°C, then a 2 mL aliquot was used to inoculate a 250 mL Erlenmeyer flask containing 50 mL of LSFM medium (18.7 g/L glycerol, 40 g/L glucose, 5 g/L yeast autolysate (Sigma), 2 g/L $(NH_4)_2SO_4$, 5 g/L soybean meal, 5 g/L tomato paste, 2 g/L sodium citrate, pH adjusted to 7.0) or FFL medium (70 g/L glucose, 100 g/L glycerol, 3 g/L NH_4Cl, 10 g/L L-glutamic acid monosodium salt, 8 g/L Amicase® casein acid hydrolysate (Sigma), 0.7 g/L L-tryptophan, 1 g/L K_2HPO_4, 1 g/L $CaCO_3$, 0.5 g/L $MgSO_4{\bullet}7H_2O$, 20 g/L 2-(*N*-morpholino)ethanesulfonic acid, 5 mL of 85% lactic acid solution, 20 mL of 50× trace element solution, pH to 6.8). LSFM and FFL cultures were grown in total darkness on an orbital shaker at 220 RPM for 15 days at 29°C. *H. pulicicidum* cultures were strained through cloth, isolated mycelia were placed in RNAlater® (Sigma-Aldrich), and samples were stored at −80°C. Total RNA was extracted from approximately 100 mg mycelia per sample using the RNeasy Plant Mini Kit (Qiagen), including the optional RLC buffer and on-column digestion with RNase-free DNase (Qiagen), as per the manufacturer's instructions. Sample RNA integrity was confirmed by analysis on a LabChip® GX Touch HT Nucleic

Acid Analyzer (PerkinElmer). RNA samples were sent to Novogene (Singapore) for poly-A enriched cDNA library preparation using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (New England Biolabs) and 150 bp paired-end sequencing using a HiSeq X instrument (Illumina). Sequencing data generated for this project was submitted to NCBI and is available under Bioproject PRJNA770096. RNAseq reads were aligned to the *H. pulicicidum NOD* cluster sequence (GenBank accession MG182145.1) using the Geneious mRNA mapper (Geneious R11.1, Biomatters, Auckland, New Zealand). These alignments identified errors in the originally-published[2] coding sequences for genes *nodD2*, *nodR*, *nodY1*, *nodY2*, and *nodZ*. Our RNAseq-verified coding sequences of all *NOD* genes are included below (pg 11-20), and have also been updated on GenBank.

Table S1: Fungal transformant list

Code*	**IDT Genes**	***P. paxilli* base strain**	**Genomic context**	**Plasmid one**	**Plasmid two**
Assignment of the 4-Series					
RC63-5	*G,C,M,B,W + D1*	CY2 (ΔPAX)	-	pSK81 (*paxG, nodM,B,C,W*)	pRC63 (*nodD1*)
RC121-3 (**Ia**)	*G,C,M,B,W,D1 + O*	CY2 (ΔPAX)	-	pRC121 (*paxG,nodM,B,C,W,D1,O*)	-
RC128-1 (**IIa**)	*G,C,M,B,W,D1,O + D2*	CY2 (ΔPAX)	-	pRC121 (*paxG,nodM,B,C,W,D1,O*)	pRC128 (*nodD2*)
RC166-8 (**IIIa**)	*G,C,M,B,W,D1,O, D2 + R*	CY2 (ΔPAX)	-	pRC121 (*paxG,nodM,B,C,W,D1,O*)	pRC166 (*nodD2,R*)
RC167-8 (**IVa**)	*G,C,M,B,W,D1,O, D2,R + Z*	CY2 (ΔPAX)	-	pRC121 (*paxG,nodM,B,C,W,D1,O*)	pRC167 (*nodD2,R,Z*)
LS201-3 (**Ib**)	*G,C,M,B,W,D1,O + Y2*	CY2 (ΔPAX)	-	pRC121 (*paxG,nodM,B,C,W,D1,O*)	pLS201 (*nodY2*)
LS202-9 (**IIb**)	*G,C,M,B,W,D1,O, D2 + Y2*	CY2 (ΔPAX)	-	pRC121 (*paxG,nodM,B,C,W,D1,O*)	pLS202 (*nodD2,Y2*)
LS203-8 (**IIIb**)	*G,C,M,B,W,D1,O, D2,R + Y2*	CY2 (ΔPAX)	-	pRC121 (*paxG,nodM,B,C,W,D1,O*)	pLS203 (*nodD2,R,Y2*)
RC173-5 (**IVb**)	*G,C,M,B,W,D1,O, D2,R,Z + Y2*	CY2 (ΔPAX)	-	pRC121 (*paxG,nodM,B,C,W,D1,O*)	pRC173 (*nodD2,R,Z,Y2*)
Assignment of NodJ function					
LS148-7 (**V**)	*G,C,M,B,W,D1,O + J*	PN2257 (Δ*paxM*)	*paxG,C,B*	pLS148 (*nodM,W,D1,O,J*)	-
LS146-1 (**VI**)	*G,C,M,B,W,D1,O, D2,R,Z,Y2 + J*	PN2257 (Δ*paxM*)	*paxG,C,B*	pLS133 (*nodM,W,D1,O*)	pLS146 (*nodD2,R,Z,Y2,J*)
Assignment of NodX function					
LS196-4	*G,C,M,B,W,D1,O + X*	PN2257 (Δ*paxM*)	*paxG,C,B*	pLS196 (*nodM,W,D1,O,X*)	-
LS170-3 (**VII**)	*G,C,M,B,W,D1,O, J + X*	PN2257 (Δ*paxM*)	*paxG,C,B*	pLS170 (*nodM,W,D1,O,J,X*)	-
LS197-3	*G,C,M,B,W,D1,O, D2,R,Z,Y2 + X*	PN2257 (Δ*paxM*)	*paxG,C,B*	pLS133 (*nodM,W,D1,O*)	pLS197 (*nodD2,R,Z,Y2,X*)
LS147-3 (**VIII**)	*G,C,M,B,W,D1,O, D2,R,Z,Y2,J + X*	PN2257 (Δ*paxM*)	*paxG,C,B*	pLS133 (*nodM,W,D1,O*)	pLS147 (*nodD2,R,Z,Y2,J,X*)
In vivo feeding studies					
LS29-1	*J*	CY2 (ΔPAX)	-	pLS29	-

**Numeral in brackets refers to the trace number used to identify this transformant in Figure 3 of the manuscript.*

SUPPORTING INFORMATION

Table S2: Components of DNA constructs used in this study

Code	Promoter	Gene	gDNA/cDNA	Original MIDAS level 1 plasmid	Terminator
RC63	P*trpC*	*natR*	gDNA	pRC1	T*trpC*
RC63	P*janD*	*nodD1*	gDNA	pRC25	T*paxP*
RC121	P*trpC*	*neoR*	gDNA	pSK16	T*TtrpC*
RC121	P*paxG*	*paxG*	gDNA	pSK2	T*paxG*
RC121	P*paxM*	*nodM*	gDNA	pSK18	T*paxM*
RC121	P*paxB*	*nodB*	gDNA	pSK19	T*paxB*
RC121	P*paxC*	*nodC*	gDNA	pSK20	T*paxC*
RC121	P*janD*	*nodW*	gDNA	pKV45	T*trpC*
RC121	P*janO*	*nodD1*	gDNA	pRC25	T*paxB*
RC121	P*janD*	*nodO*	gDNA	pSK42	T*paxP*
RC128	P*trpC*	*natR*	gDNA	pRC1	T*trpC*
RC128	P*janD*	*nodD2*	gDNA	pKV61	T*paxG*
RC166	P*trpC*	*natR*	gDNA	pRC1	T*trpC*
RC166	P*janD*	*nodD2*	gDNA	pKV61	T*paxG*
RC166	P*janO*	*nodR*	gDNA	pKV58	T*paxG*
RC167	P*trpC*	*natR*	gDNA	pRC1	T*trpC*
RC167	P*janD*	*nodD2*	gDNA	pKV61	T*paxG*
RC167	P*janO*	*nodR*	gDNA	pKV58	T*paxG*
RC167	P*janO*	*nodZ*	gDNA	pRC156	T*paxG*
RC173	P*trpC*	*natR*	gDNA	pRC1	T*trpC*
RC173	P*janD*	*nodD2*	gDNA	pKV61	T*paxG*
RC173	P*janO*	*nodR*	gDNA	pKV58	T*paxG*
RC173	P*janO*	*nodZ*	gDNA	pRC156	T*paxG*
RC173	P*janO*	*nodY2*	gDNA	pKV49	T*paxG*
LS201	P*trpC*	*natR*	gDNA	pRC1	T*trpC*
LS201	P*janP*	*nodY2*	cDNA	pLS58	T*paxP*
LS202	P*trpC*	*natR*	gDNA	pRC1	T*trpC*
LS202	P*janD*	*nodD2*	gDNA	pKV61	T*paxG*
LS202	P*aceB*	*nodY2*	cDNA	pLS58	T*aceB*
LS203	P*trpC*	*natR*	gDNA	pRC1	T*trpC*
LS203	P*janD*	*nodD2*	gDNA	pKV61	T*paxG*
LS203	P*janO*	*nodR*	gDNA	pKV58	T*paxG*
LS203	P*janP*	*nodY2*	cDNA	pLS58	T*paxP*
LS148	P*janO*	*neoR*	gDNA	pSK16	T*trpC*
LS148	P*janO*	*nodM*	cDNA	pLY52	T*trpC*
LS148	P*janO*	*nodW*	cDNA	pYL54	T*trpC*
LS148	P*janO*	*nodD1*	cDNA	pYL55	T*trpC*
LS148	P*janO*	*nodO*	cDNA	pYL56	T*trpC*
LS148	P*aceB*	*nodJ*	cDNA	pLS59	T*aceB*
LS146	P*aceB*	*natR*	gDNA	pRC1	T*aceB*

LS146	*PaceB*	*nodD2*	cDNA	pYL57	T*aceB*
LS146	*PaceB*	*nodR*	cDNA	pLS56	T*aceB*
LS146	*PaceB*	*nodZ*	cDNA	pLS57	T*aceB*
LS146	*PaceB*	*nodY2*	cDNA	pLS58	T*aceB*
LS146	*PaceB*	*nodJ*	cDNA	pLS59	T*aceB*
LS196	*PjanO*	*neoR*	gDNA	pSK16	T*trpC*
LS196	*PjanO*	*nodM*	cDNA	pLY52	T*trpC*
LS196	*PjanO*	*nodW*	cDNA	pYL54	T*trpC*
LS196	*PjanO*	*nodD1*	cDNA	pYL55	T*trpC*
LS196	*PjanO*	*nodO*	cDNA	pYL56	T*trpC*
LS196	*PaceB*	*nodX*	cDNA	pLS50	T*aceB*
LS170	*PjanO*	*neoR*	gDNA	pSK16	T*trpC*
LS170	*PjanO*	*nodM*	cDNA	pLY52	T*trpC*
LS170	*PjanO*	*nodW*	cDNA	pYL54	T*trpC*
LS170	*PjanO*	*nodD1*	cDNA	pYL55	T*trpC*
LS170	*PjanO*	*nodO*	cDNA	pYL56	T*trpC*
LS170	*PaceB*	*nodJ*	cDNA	pLS59	T*aceB*
LS170	*PaceB*	*nodY2*	cDNA	pLS58	T*aceB*
LS197	*PaceB*	*natR*	gDNA	pRC1	T*aceB*
LS197	*PaceB*	*nodD2*	cDNA	pYL57	T*aceB*
LS197	*PaceB*	*nodR*	cDNA	pLS56	T*aceB*
LS197	*PaceB*	*nodZ*	cDNA	pLS57	T*aceB*
LS197	*PaceB*	*nodY2*	cDNA	pLS58	T*aceB*
LS197	*PaceB*	*nodX*	cDNA	pLS50	T*aceB*
LS147	*PaceB*	*natR*	gDNA	pRC1	T*aceB*
LS147	*PaceB*	*nodD2*	cDNA	pYL57	T*aceB*
LS147	*PaceB*	*nodR*	cDNA	pLS56	T*aceB*
LS147	*PaceB*	*nodZ*	cDNA	pLS57	T*aceB*
LS147	*PaceB*	*nodY2*	cDNA	pLS58	T*aceB*
LS147	*PaceB*	*nodJ*	cDNA	pLS59	T*aceB*
LS147	*PaceB*	*nodX*	cDNA	pLS50	T*aceB*
LS29	*PjanO*	*nodJ*	gDNA	pKV47	T*paxG*

Sequences

Promoters used in this study

>P***trpC***
GAATTCATGCCAGTTGTTCCCAGTGATCTTCGTTTCGAAGATGGACACTCCCAATTTGTGCAAGTTATTCGGC
CTACCTGGCTGTGGCCGAGGCGCGTTATCATGACCGTCGCTGTTCAAAGATAAGGCGAGAAGTTTGCGGGC
TGTCTTGACGATATGGCTTCGTTCAGACAGATATAGTTCCCGGAGTCGCAGGCGTCTATTCTTCTCCGAAAC
AAACTCGGCTGCACTGTTTCCATCACCGGGTCTGGCGTTGAGGATGTCAGCGAAACTCGGCCCGGCAAGTG
ACACCCGAAAAGTATCGACTCCGGCTGCCCGTTTCAAGCTAGTGGCTTCCTCATCAGCGAGTCGGCCAAGC
AGACGTGAAGCAGGACGGGTTTGCCATTCCAAGACCGTGATCCGAAGCGCGTTGATTTCATCAATCCCAGC
CTTTTCGCTCAACCAAAGAGCATCGGCTTTGATTTCCTTCAGGTCATACGAGGCTTGTGCAATGGTCTgCGCA
TGGATCGCTGCTGTTCTCCTATCAAACTCGGATTTTGTCTTAGGGGATGGCGTAGGAAAGACGCTGCCGCG
GTTCAGAAGCACCTCGATGCTATCAGGATGTGACAAAAACGACTCGAAAACCCGGGATTCATCGGTGATGCT
TTCGGGATCGCAAGCGTAAAGAAAGACTCTCTTCCAAGACCTAGAAGTATAGCAAAATCAGCAGCAGACCAT
CAATGTATAGCGAATGCGCCCATACAAAAGCTGAACGTCCCCGGAGAAGCACTTGTCCAGGGACGGGAAAT
AGGCTTCCGGAACGGGAGCCATTGGCAGCACAGCTATATCATTCTAAGTAAACAAATGTAATGAGCAAGCG
GACGGAGTGCTGAAACCTCCGTATGCCTGAAGCCGACGAAAGCGCGTTGGATTAGAGGTCGACAGAAGAT
GATATTGAAGGAGCACTTTTTGGGCTTGGCTGGAGCTAGTGGAGGTCAACAATGAATGCCTATTTTGGTTTA
GTCGTCCAGGCGGTGAGCACAAAATTTGTGTCGTTTGACAAGATGGTTCATTTAGGCAACTGGTCAGATCAG
CCCCACTTGTAGCAGTAGCGGCGGCGCTCGAAGTGTGACTCTTATTAGCAGACAGGAACGAGGACATTATT
ATCATCTGCTGCTTGGTGCACGATAACTTGGTGCGTTTGTCAAGCAAGGTAAGTGAACGACCCGGTCATACC
TTCTTAAGTTCGCCCTTCCTCCCTTTATTTCAGATTCAATCTGACTTACCTATTCTACCCAAGCATCGAT

>P***janD***
CTTACCCGTCCTTTTTTCTCAACGAAAACTCAGAGATCGGGCTTAAATAGATGTCAAGGGCACGAGGAAGGC
TGTAGGGGGTTCAAGTAATGCTTATTAACAGAAAGAAAAAAGAAACTGAAGTACTTGCATTCAGCCAGCAAAT
TTGTTTGAGTGGTATTCCAAGGAAGTGAAAAGATAGTGGGCAAGGCGAAGTGATATAGGTATTTAAGTGTTC
ATGAGCCCTACTGGCCAAGGAGATTCTCCTATCAGCTACCTATTGACCCCAAAAACATAAGGTAAAATTACG
CAATAAGGAAACTGAACACCTCCTAGGTGCTGCACAGCTCTGGACCTGTCTCACCTCTACCTCATCGGTAAA
GACTCGATTCTCTAGATCAGCTCTCGTGTTCATGTTGATTGCAACGTTTCACTGCAGTCCACATTAGCCCATG
CAGACGCAATCTCAATCACTATGTTTCCTGCATTGCAGCTATAGTTTATGCACTCGAATATCAGATGTAATTTA
GCTACGTGTCACTTGGCATAGAAAAGAAATGCAGATTCTCGGCGCAGGTAGCCCATCAAACTATTTCCGCCT
CGCATAACGCTATGAGCTAGTAATTGGCATATGCATTCGGGCAGATTGGGCTTCACTCAACGTGTAACCCTA
CACGTTCTGTACGGTATGGAATACACAATACCACTAGCTGTATGAATCTCGTCCTATTTTCCTGCACAGGAGT
CGGGATTCTTATCATGGCTTCGAAGGTAATTGATAGATCACAGGGGCCCACCGCACATTCTGATTTGCGTAT
AGAGCTGCTATACATCCGGGCTGACTGTAGTTCTTCGACTTTTCATCCTGCGAAAGTGGGGTAATACGACAT
GCATGATACCGTTGACCAGATACCAGATGTCAATCATGGTCCTTTGATCTTCATCAATAGATACTGTAATTTG
TCAACATCACTCACCCTTCTTTCTCATATAAATCTGATTAGTCGTGGGGTGCACTCAGCGAGCACCTTGTTTG
ACCAAACCTAATCTTTCCACTTCATCAAGCTCAAGATCATCTTAAACT

>P***paxG***
ATTCACGACCTGTGACTAGTCAAGGTTATCAATTTTAAAACTTACTGGGTGTTGTGTTGAATACTATGCACGA
GATCTGGCACGTATAATCTATGGAGTTTACTTCCGTACACAGAATTACATCTCGCTTTGTGAATACGAATCAC
CCAGTCATGGATGTATTACAACGCAACCATAATCGTTGCATCATGTCATCCATGTTCGCACGCCTTGTCTATA
ATACGGAGTTCTCAAGTGGGGGCAGAGGGATTTATTTTCGAGAATTTGAAACATGACTGTACCGCAGTTGTA
ATTATAGAGAGTTATTTAACTGGGGTAGTTGGGAACGGATTTTTAGCGGCATAAATCAATACAAGGATCCGGT
GTCGTTTACTCTCGAACCCTACCCATACCCCGGAATTTCCTTCATCGCAATTTTGCAATCCTACTCTGCAAAA
TAGAAAAGAATTTCTTCCAGTCCCATTTCTTTCATCATAGGGTTATAAAAATAGTGCTCTTGAATTCAGGTTCA
CGCCAGAGAAAACTTCATCAAGTTCGACGCC

>P***paxM***
GTTGTTGGCATGGGAGTAGGATGGATCATAAACTACATCGGAATGGTCTACATCTCGTTCAAACATGAAACTT
ACGGAATGTCGATTATGCCCCTCTGTTGCAACATCGCGTGGGAACTGGTCTACTGTCTGGTCTTTCCCTCGA
AAAGCCCGGTGGAGAGGGGTGTATTCTGGATGGGCCTCCTCATCAACTTTGGTGTCATGTACGCAGCAATA
ACGTTTTCATCCCGAGAGTGGGGCCATGCGCCGCTTGTTGAGCGCAATATTTCCTTGATATTTTTCGTCGCG
ACCATGGGCTTTCTCTCTGGGCATGTTGCATTGGCCCTTGAAATCGGGCCTGCGCTGGCATATTCATGGGG

GGCCGTCATTTGTCAACTTCTTCTAAGCGTTGGTGGATTGAGTCAACTGTTGTGTCGAGGCAGTACGCGTGG
CGCATCCTATACTCTATGGTATGTCTGATCTCATTTCATGCCCGATGAATCGGAATTTCTTCCCCAAAGTGAC
AGAAATACAAGTACTCATCTTCTGTACTATAGGGCTTCCCGCTTCTTAGGATCAACATGTACAGTTGGGTTTG
CCGGCTTACGTTGGATGTATTGGTCGGAAGCATTCGGTTGGCTGAATAGTCCTTTGGTACTGTGGAGTCTTG
TGGTGTTTTTATCAATCGATGGGTTTTATGGAATTTGTTTCTGGTACGTCGACCGTAACGAAAAGTCACTTGG
CATAAGCGGGCCGAAAAAAGCAAATTGATCGGCAGTTGAGGGTGAAACATGTAACTAGCTTATTCTGCATAC
AATGCATTATCTATGGAGGGATGAGAAAAAGTAAAATGAAATTGAAGAAAAGAAAAGAAAATTAGAAACAAAA
ATACACGTTTACTTTTTGAAATAACCTTGGAAACTAAGATCCGAAAAAAGAAATCTTTTCTTCATGCATTTCTG
CCAGTGCTAGATTGAAATTGAAGAAAAGTACATACGATGTACTTTGCTTACGATACGCGTAACGTGCATGTCC
AAGGCGACCCAATCCATGATATAATTGGCACGTTTTAGTAGTAAGCAATGATCATTTGTGGCTGACACTATGC
AATCTGTATCATATGATATTAGTATGCAAAGGTTGATTCTGAGTTTGAATTTCAACCACTGTTCATCGTTCCTC
ATTGTTAACCTTATTCATATTTCACCTTTTGCTTTATTCTTTTCATGTATCTTTAAGATTCAGAAACC

>**P*paxB***
GGAGAGAATCGAAAGAGTATTCATGTTGACTTGGAACAGAGAATTTGGTATTGACTAACGTATTCTCGAGTTC
TCTCAAGGAAAGGAAAAATGAAATGAGTCGAGTAGCTCTGCTTGGTAGGCTGTTGAGATTCGGCGCAATACA
TCCCCTCCATATAATAAGTAATGCAAGGAGTATACTTTCCCAGGTGTCCAATCGGCGGCTGCAGCTATCACA
ATTAGAAAAAGCAGTGGGATTCATCGCGTTCAGCATTTTCGCCCCACAATTGTCAGCTTTCAGGCGTGATCC
CAAAAGTGATACATGTTACTAAATTTACGGATTGACCTCATTCAGCGATGTCTACGATGAGCAACGAGGTAGT
AGTATCAAGAGGGGCTGTCGGTACCTTAATAATTCCGGTTTCTCCGTGATGCAGATATCAAACTATGCAGTA
GTAATTAATCTTTTATTTTTTCTTCATACCAGCCCCTGTAATGACAACATCACGCCTAAATAGGTTAAGCGCAG
CAACGCCTCAAGATCCTTCATGCAGTGAGAAAAATAAGAATTATAATATTAATTAAAAAAGAGGGAAAGTAAC
GAAACTTTTGCCCTGATCCACATCGACGCCTCGATTTTCCAGGCGCACCGTACTCAAATCGGCTTTCATTCC
TTTGAGCCATTTTCTTCACTTGGCAGTAAGCCTAAACTCTGCCTATCCAAAGACAGCTAATTGTTGCGACGCA
GTTCGTCTGAATTCCATGGCAAGAATGACTGCATACTAGCTTTCGTTGGCGAAATAGACTAGTTTAGATCTTG
CAGGATCTCCTCGTGGACTGCATCTACAGAGATTTCCGGGAGTGTCCATCTCCCGCTGCGGGCTAACCGAT
GAGATAATGCAGTTTCATTAACAATTCAATATTCCCATTTCATCCACAGGGCATAAATAGACGTACATTTAGCC
GTTAATACGCATCATGGTCAAATCCTTCAATCTATAGCTATTCAACGAGTGACTCTAATTTCTTTTTCCCACGT
CAACCTTAGAAACT

>**P*paxC***
ACAACAAAAAGATCAGCCAATGGCTTAATCGCCTGGTACTCCGGAGGAGCTTGGGAAACATCAAAACCGTC
CATAGTTTCTAAGGTTGACGTGGGAAAAAGAAATTAGAGTCACTCGTTGAATAGCTATAGATTGAAGGATTTG
ACCATGATGCGTATTAACGGCTAAATGTACGTCTATTTATGCCCTGTGGATGAAATGGGAATATTGAATTGTT
AATGAAACTGCATTATCTCATCGGTTAGCCCGCAGCGGGAGATGGACACTCCCGGAAATCTCTGTAGATGCA
GTCCACGAGGAGATCCTGCAAGATCTAAACTAGTCTATTTCGCCAACGAAAGCTAGTATGCAGTCATTCTTG
CCATGGAATTCAGACGAACTGCGTCGCAACAATTAGCTGTCTTTGGATAGGCAGAGTTTAGGCTTACTGCCA
AGTGAAGAAAATGGCTCAAAGGAATGAAAGCCGATTTGAGTACGGTGCGCCTGGAAAATCGAGGCGTCGAT
GTGGATCAGGGCAAAAGTTTCGTTACTTTCCCTCTTTTTTAATTAATATTATAATTCTTATTTTTCTCACTGCAT
GAAGGATCTTGAGGCGTTGCTGCGCTTAACCTATTTAGGCGTGATGTTGTCATTACAGGGGCTGGTATGAAG
AAAAAATAAAAGATTAATTACTACTGCATAGTTTGATATCTGCATCACGGAGAAACCGGAATTATTAAGGTAC
CGACAGCCCCTCTTGATACTACTACCTCGTTGCTCATCGTAGACATCGCTGAATGAGGTCAATCCGTAAATTT
AGTAACATGTATCACTTTTGGGATCACGCCTGAAAGCTGACAATTGTGGGGCGAAAATGCTGAACGCGATGA
ATCCCACTGCTTTTTCTAATTGTGATAGCTGCAGCCGCCGATTGGACACCTGGGAAAGTATACTCCTTGCATT
ACTTATTATATGGAGGGGATGTATTGCGCCGAATCTCAACAGCCTACCAAGCAGAGCTACTCGACTCATTTC
ATTTTTCCTTTCCTTGAGAGAACTCGAGAATACGTTAGTCAATACCAAATTCTCTGTTCCAAGTCAACATGAAT
ACTCTTTCGATTCTCTCCAACACAGCCTTTCAGGGTGTAGGTCCCATTTT

>**P*janO***
AGTTTAAGATGATCTTGAGCTTGATGAAGTGGAAAGATTAGGTTTGGTCAAACAAGGTGCTCGCTGAGTGCA
CCCCACGACTAATCAGATTTATATGAGAAAGAAGGGTGAGTGATGTTGACAAATTACAGTATCTATTGATGAA
GATCAAAGGACCATGATTGACATCTGGTATCTGGTCAACGGTATCATGCATGTCGTATTACCCCACTTTCGCA
GGATGAAAAGTCGAAGAACTACAGTCAGCCCGGATGTATAGCAGCTCTATACGCAAATCAGAATGTGCGGT
GGGCCCCTGTGATCTATCAATTACCTTCGAAGCCATGATAAGAATCCCGACTCCTGTGCAGGAAAATAGGAC
GAGATTCATACAGCTAGTGGTATTGTGTATTCCATACCGTACAGAACGTGTAGGGTTACACGTTGAGTGAAG
CCCAATCTGCCCGAATGCATATGCCAATTACTAGCTCATAGCGTTATGCGAGGCGGAAATAGTTTGATGGGC
TACCTGCGCCGAGAATCTGCATTTCTTTTCTATGCCAAGTGACACGTAGCTAAATTACATCTGATATTCGAGT
GCATAAACTATAGCTGCAATGCAGGAAACATAGTGATTGAGATTGCGTCTGCATGGGCTAATGTGGACTGCA
GTGAAACGTTGCAATCAACATGAACACGAGAGCTGATCTAGAGAATCGAGTCTTTACCGATGAGGTAGAGGT
GAGACAGGTCCAGAGCTGTGCAGCACCTAGGAGGTGTTCAGTTTCCTTATTGCGTAATTTTACCTTATGTTTT

TGGGGTCAATAGGTAGCTGATAGGAGAATCTCCTTGGCCAGTAGGGCTCATGAACACTTAAATACCTATATC
ACTTCGCCTTGCCCACTATCTTTTCACTTCCTTGGAATACCACTCAAACAAATTTGCTGGCTGAATGCAAGTA
CTTCAGTTTCTTTTTTCTTTCTGTTAATAAGCATTACTTGAACCCCCTACAGCCTTCCTCGTGCCCTTGACATC
TATTTAAGCCCGATCTCTGAGTTTTCGTTGAGAAAAAAGGACGGGTAAG

>**P*janP***
GACTTCGGTTTCCATTGGCAGTGCTCGTGTCTATACTCAGGAAGACTTGACCGTGACCGTCTAGTTAGTTTT
CTTTGCGATACCGTGCAAAAAAAAAGAAGAGCCGGGCAGTGCAAAATAGCTTCCTACTCGGTCCCTAGAAAC
GTTTGCTGGATGGATAGACTTCGTTTCTTTTCCTTCTATTGCCTCGCTCCCAAAGGTACAATAATTTACGTATA
TTGAGTCCAGTATTTATGGTCAGACCTATACGCGTACATACCTTCAGTAAATTTAGAGTACACTTCCTGGCCA
AGTCATTTCTGATATTCCACCGTAGTGTCCAATGGTACCTGGGTCTTTCAATCAATCTTACCCCTCTATGATG
CATGGTAGTAAAACCCTCTTTAAGTACCGTAAACTTTTGAAAATTCAATAATGCAATGATTATGCAATGTGGG
GCTATCACGGTCTGGCCGGAGACTATATTGTGGTAATGATACGACATGCAGAATACCAAAGCCAGGCACGAT
GGAACAGTGAATACATACAGGTCATCCAGGCGCCAACGAGGTTCATCCAGGTGTGGACCTACTCCACTTTAA
GCTAACCCCGGTGTGCTCAGTAATAAAATTGCGTAAAACTCCAGTATACATTATTTAACTCGTTATAAATCTTT
TACAACTACATTATCTCTACTAATAGTTTCATTATCATTTCGATAGAGGTCTTTAGGGTTTTCTAGTACTCTATA
AGCCTAAAATATTATATTTTGTACTAGTTGATTATATTATTTTATTTTAAGACTTATTTCACTACTTTTAGACTAG
AGTAAGAATAAAATATCTTAATTACAGTGATAATATTATTATATCACAGACCCTGTATTTCTTTGTTTTAAAATA
CAAGCCGACCGGTAAAGATACGCCCTGTAATAAACACGGGACTGGGTGGGTCGCTTAGGAAAAGTCCCACA
CCTGGGTTAAGCTCGTTGGCACCTGGATGACCTGTACTCTGGATGGAACATGCTTTTGTCCAAAGATGTTCT
GCAGAAATGGCTCATATGCTCAATATAACAGATCATGCATAATGGCAACCGATAAGAAACTAAAAACAAGCCA
ATCATATTGAAAATTACATAAAATTCTAAAGCTACATCCACGGTGCCTAAATTATGTCTCAGGCACCCTCGTAT
AGCCATGGATAGGGTATGATTCCAGTTACCCTGTGAAAAGCCGCAAGTCCTGAAAATGTTTCCCAAAAATGT
GATAACCAGCTCACTAGCCCCTTGAAATTTGTTTGCTTCCTACACCATCCACTTCTATCTTGCTTTCTGCTGC
ATTCCACAGGACA

>**P*aceB***
CTCGGGAGGCCCAGGACTGGGCTGGCACCCCGAAGACACATTTCTGGCAGATTTGCAATGAGGCTCATGG
CGTAGAAGTGATCTTGATATGAATTCCAGTGATGGAGAGGACGATGCGTAGGAAGACAAAATCGGCTTCAAT
ATCCAACAGCAGCACTTTCTTCAAGAATTCACCTCCGTCATATCAGCGGGCCGCACTAGCACCCGCTGCTCC
TCGCGCCATTCTGGATATATCGGGAACCCTCTAAAAATCCACATTTTTGCCTATCTACTGTACTAACATGGGT
GCTACCTCGTTCAGGCGTGCCCGATACAGCGATTTGCAGTGGTCAACGTCAGCCGGTCTTCAGCCGTAGCG
GTCAAGGATGACGAATGATCTAATTTGTGGGGCTCCGAAAGTATACTTAGTGGTATCAGTATGCTAATCAAGT
ATTTCCCTCGTTAATGTCCTGCTTTCTCATCACAAAGTGCATGCAGGCGTGGCATCCCATGGCAAAAAAAGC
AAGATGCATCCTCATATCCTGCAGGCTTATCTAAATCACGGATGTGACAGTATCAGTGTCCCGTGTGGTATG
AGCTCATCCCGCTGGAATTGCAAAAAACAGTATTCATTTTGAATTACAATACATACAATAACTCCATTCACCTT
AGTGACCCTAATGCATGGCCTGCATTTTTAATATAATTTATTATGAAATGGGTTCAGCTGTTATTGAATACTTA
CCCATAGCTACATGGCAAATCGGAAAGCACTTACTCTCAGGGCTAATCAGAAACGGTCATCTCGCGCGATCC
CTCATCCTTTACGCAAACCATAAATAGGCAGGCGCTTGAGAGGCCTTAGTCCTGGACTCAACAGTATACGAC
TCCTTCTGCACTTACGTGGATACGCCATCAATGTGAGAC

Coding sequences used in this study (gDNA)

>***paxG*_gDNA**
ATGTCCTACATCCTTGCAGAAGCTCTGAACTTCGTTCGTCGAGGAATATCTCATCTTAATTATTGGGGAGCTT
CACACTCATTATCTGCTGATAATTACTGGGAATCAAACTTCCAAGGATTCCCAAGACTCCTGAGTGATTCCAG
CAAAGCACCAAGCACGATTAGAACAGTACAGGTGCTAGAAGATGACGTTGACGATATTGCTATTCAGTATAA
CAAGGTCAGTTCTGTCATTGCTCTTGGGATTCCCCTTGTGTTTCATTATACCAATCGTTTGTTCTTCTTAGATT
GTTCGTGGACCTCTAGACTATCTTCTAGCGATTCCGGGAAAAGACATCCGAAGCAAGCTGATCGACTCCTTC
AACATATGGCTCCAGCTTCCAGAGGAGAAGTTGTCTATTGTGAAAGATATAATAAACTTACTTCACACTGCAT
CCCTTCTGTGAGATTTCCAACCCCAGCGCTCCTGTTCTATTCTAACATCATGGTGATTATAGTATCGATGATA
TTCAAGATGCATCACGGCTCCGGAGAGGAAAGCCTGTGGCTCATGATGTATACGGTGTAGCACAAACTATCA
ACTCAGCCAACTATGCCTACTATCTCCAACAAGCAAGATTGAAAGAGATTGGTGACCCCCGCGCCTTCGAAA
TATTTACAAGGTCACTGCTAGATCTACATCTTGGCCAGGGCATGGACCTATATTGGCGTGATATGGTGGTTT
GTCCGACCGAGGAGGAATACACCCGGATGGTTATGTACAAAACTGGAGGTCTTTTCAACCTGGCCCTCGATT
TGATGCGTATCCAATCCCGCAAGAATACGTAAGTTTCTAATCCATAACGGTCACTTCGCAGCAATGCTAATTA
GTATTTTACAGTGACTTTTCTAAACTTGTCGAGCTATTGGGTGTTATATTCCAAATTCGTGATGACTATATGAA
TCTGCAGAGCGGACTCTATGCTGAGAAAAAAGGGCTGATGGAAGATTTGACGGAGGGGAAGTTCTCCTATC
CTATCATACACAGCATCCGTGCATCTCCAGAGAGCTCCGAGCTTCTCGATATTCTAAAGCAACGCACAGAGG

ACGAAGCAGTCAAAATCCGGGCTGTGAAGATCATGGAGTCGACGGGGAGCTTCCAATATACTAGGGAAACC
CTAAGTCGACTTAGTGCGGAGGCTCGTGGCTATGTAAAGAAGCTGGAGACTTCCTTAGGGCCCAATCCTGG
AATTCATAAGATTCTCGATCTACTTGAAGTGGAGTACCCTACTAATGAGAAAGGAAGAGTTTAA

>*nodM*_gDNA
ATGTCTACCCCTGAGTTCAAGGTGATAATCGTTGGCGGCTCCCTTGCGGGTCTAACATTGGCCCATTGCCTC
CTCCGCGCCGGAATTTCCCACATCGTTCTTGAGAGGCGTAGTGTAATCGCACCAGAGGAAGGTGCCTCAAT
AGGGATTCTTCCTAATGGCGCTCGTGTCTTGGATCAACTAGGCATCTATGAACATATTCAAGACACTACTGAG
CCATTGAGTACAGCTCATATACGATATCCAGATGGGTTCTATTTCAGCAGTCGCTATCCCGAGATCATAAAAG
AAAGGTGAGCACTCCCATGAACATATGCCTATCGTTTGCGCAATGGCTGACCCCAGCATGGCCGTTCGTAG
GTTTGGGTATCCGATAGCCTTTTTACCCCGGAGGAGACTTTTGGAGATTCTCTATACATCGTACCCTGACCAT
TCTAACATCTATACCAACAAGAATGTGATCAAGGTGCAAAGCCATGATAATCAAGTTTCTGTCTTGACAGAGG
ATGGAAATACATATCGAGGGGATCTTGTGGTGGGCGCCGACGGTGTCAACAGTCGAGTTCTATCTGAAATTT
GGAAATTAGCAGGTAATCCCTCTCTCACGAAACGGGAAGGAAGAGGTCAGTCGAGATCTTATAATATTCGCA
TAGCGACTCATACAGACGTTGATAATAAACTTAGGAAGGACCATCGAGTACGCTTGTGTGTTTGGGATATCC
TCGCCAATCTCAGACCTCAAGCCCGGCGAGCAGGTTAACGCGTTCTACGACGGACTCACAATCGTCACGAT
CCATGGTAGAAACGGAGAGATATTCTGGTTTTTTATCAAGAAACTCTCGCGTCGCTATATCTATCCCGATCTC
ATCAGATTACAGCAGAAAGACGCAGAAGGGATCTGTGAAGAAGCCAAGTCACTAACTGTCTGGAAAGGTGTT
ACATTCGGTGACATTTGGGAAAGAAGAGAAACAGCATCATTGACCGTTTTGGATGAGTTCTTGCATCACACTT
GGAGTTGGGACAGATCGGTTTGTGTGGGCGACAGTATCCACAAGGTATATGCCTCCTATGGTCTGAAGCAC
ATTAAGGGGTTGGCTAACGCTGTCATATAGATGACTCCAAATTTTGGCCAGGGTGCGAATACTGCTATCGAG
GATTCTGCCGCTCTAGCCAATCTACTGCACAGTTTGATCAAGGAGAAACGAGCCGAAAAGCCGACTGACTCA
GATATATCACTGCTTCTGAGGCAGTTCAAATCACAGCGTCTTCGACGCGTTCAGAAGATATATAAAATGTCAA
GGTTTGTTACGCGGCTTCAAGCGCGCGATGGGTTGTTGAATACTCTTCTGGGACGCCATTACGCCCCATATG
CAGCAGATCTTCCAGCTAAGATTGCCTCTGGATGCATCGCTGGTGCGGAAGTTCTGGATTACCTCCCATTAC
CCAAGGTAACTGGGGCGGGCTGGAATAGGGGCCATCGTAGATCTACCATGTACATCCTGCTAGGATTCACA
GGGGTATTCACCTCTGCTCTGGCGATGGTAGTGTTACTCCATATTAGAGACATCGCTTCATAG

>*nodB*_gDNA
ATGGATGGATTCGATCGTTCCAATGCACCCGTTGAATATCAGCGCGTCGAATGGATTTCGGACATCTTCGTC
TTCGGAATGGGCGTTTGTTGGCTCATAAACTACGCCGGCATGATCTACACCTCGCTCCAGGAACAGACTTAC
AGCATGGCACCGCTGGCTCTTTGCTGCAACTTTGCTTGGGAAATGGTATATGGTCTTATCTACCCTTCAAAAA
GCCGCATCGAGCAAGGTGTGTTTCTAGCTGGACTGGTAGTCAATCTCGGAGTCATGTACACCGCGATCAGG
TTTGCGCCGAATGAATGGGCCCATGCGCCTCTTGTGATGAACAACATCACACTCATTTTCGCACTAGGTGTC
TTGGGCTCCTTGACGGGCCATCTAGCCCTAGCAGCAGAGATTGGGCCAGCACTTGGATATTCTTGGGGTGC
AGTGGCTTGCCAGCTACTGCTAAGCGTCGGGGGGTTCTGCCAGCTTCTAGGGAGAAGTAGTAGTCGTGGAG
CTTCATATACCTTATGGTAGGTAATCCATCTTCTTGCTCCGAAGTTATGCAACAGAACTTTGAGGACATATAAA
TAACAAGGTATAGGTTATCTCGCTTTATAGGATCCGGTTGTGTGGTTGGGTTTGCCATATTGAGGTACATGTA
CTGGTCGGAGGCATTCAACTGGCTCAATAGTCCTCTGGTACTATGGAGCCTCGGTGTATTCATTGCGGTAGA
CAGCCTCTATGGAATTTGTCTGTGGAATGTGAAGAAATACGAACATGGCCAAGAACGAAGCAATGCGCGGA
AGGCTCAATAA

>*nodC*_gDNA
ATGTCCTTAGGTTTACAGTGCTTGGCGGCAGTGTTGTTTTCGGCTTTGTTTTCACTTGGGGTCATCCTAGTTC
ATCTTCCATGGCGCGCCTTGAAGTCAAAGGACCCGCGTGAGCGAATATTAGGTTCGCCCAAAGAACTGGTT
CCAACATGCCCTTACGAATATATTCGAAATATATACGGGCGTCATCATTGGGCGCCCTTTGTGGCCAAGTTA
GCACCGAATCTCAAAGAAAGTGATTCAGACAGGTACACAATGGTACTTGAAATCATGGACTGCATACACCTA
TGCCTGATTATGGTCGATGATGTACGTTCATCTTATTATACGCCTCTTGTTGGTGATAGAGGTGAGAGTTGTA
AATAACTAACCCGGACTCTCAGATTACAGATGACAGCGACTATCGTAAAGGCCGCCCAGCGGCCCATATCAT
CTATGGCCGTTCGGAGACAGCTAACCGTGCTTATCTTCGTGTCAGTCAGATTATAAACAAGACAACTCAGGA
CTTCCCGCGGCTCGCCCCGTGGGTCACACAGAGTTTGGCAGAGATTCTAGAGGGCCAGGACATCTCGCTG
GTTTGGCGACGTGACGGCCTCACTAGCTTTCCAAAAGCTCACGACGAGCGCGTGATTGCTTATCGGTGCAT
GTCATCTTTGAAGACTGGCGCGCTTTTTAGGTTGCTAGGGAGGCTTGTCTTGGAAAATCGTTCCATGGATGA
CACATTGAGTCAGGTTGGGTAAGTGAACTGTATCCATCCTGAACTCGACCCCTTTCACAGTGGCGGAAGCCA
ATGCTAAAACGTCTAGATACTATTCACAATTACAAAATGACTGCAAAAATGTTTTCTCATCCGAGTACGCAAA
GGCAAAAGGTACTTTAGCTGAGGACTTACGGAACCGAGAGCTGACGTATCCTATCATCTTGGCCCTCAATGA
GCCTGAAGGATTTTATATTGAGAAAGCCTTTGAGTCTGGCTCCCCTCGTGACATACAAAATGCAATCGGTGT
AATACAGAGTGAAAACGTATACCGTGCTTGTTTGGACGAGTTGAAACAATATGAATCGAACGTCAGAGAGTG
GGTTACACTATGGGGTAGAAAGGAGAAACTCGATCTTACGCATTGA

>***nodW*_gDNA**
ATGACTTTAGCTATTTTAGGCATCAGTTGCCTGGCAGCCGTCCTATCTCATACTATTTTATTCATCCGAGGAG
AGTGGGACAAGCACGCCCCGGTCCTGTTCCATTGCTTCATTACTTGCCCCATTGCGCTCATATGCTCGTTGT
TGGCAGTGGGATTATATGGGATCACTTTTATGTTCATCCAAGTTTGGCTATGCTATATCTCTTCTATCTGTCTT
TCAATCATTATATATCGGTTGTTTCTGCATCCCTTAAGAGATTTCCCCGGCAAGGTATCCGCCAGGCTGAGCA
CGTTGAGATGGATGAAAACAGCAGCCGTAGATAGAAAATGGCACCTTGAAGTACAGAATATGCATAGTGTCT
ATGGCGATTTCGTCCGAATTAGTAAGTTCTAGCCTATTTCCAACTCTGCCATAAACTCGAAGTTGATAGACTA
TGTGTCAGGGCCTCGTGAGATATCCGTCAACAACTTCAATGCCATTCGCGATTTGGCCAAGTGTGCCAGGG
GTCCATTCTACGACGTGAATTATCCCCACAAGTCCCTTCAAATGACAAGGGACAGGAATTTCCACCTCCGAC
GGAAGAAAATTTGGCAAAAGGCCTTCAGTCCTAATGGTGAGTATCACACAACCCTGGCGTGGGCGATTCATG
CTGATAGAGCATACAGCTATGCTAGCTCATGAAGGCCATGTCCGGCAATGTTGCCAAGAACTCATGGGCTAT
CTCTCATCTAGAGGAGGTCAACCGGTAGAAGTAACTGAATTGATGCAAAGAGTCACATTTGAATCAATGGGT
TGGCTTGTATTTGGAAAGGCCTTCAACATGATTCAAAAAGGCCTTTCGCATCCCGGCTTCACGCAACTGAGG
TCTGTCAAGCAACTAGGGGGGTTACTTCTGTGGGCGCCTTGGACGTTGATCCTTCTCCGCAATCTACCGTTC
TTGAGGTCCAAGGCAACAAGTTGGCTACAGTGGTGCTCACAGGAGGTAGTGGAAAGGAAGCAAGTAGGTAA
TACTTAATTACATACTAATCAGGATACCTGACTGACGCACAGTGGAGAATCCAACTATGACTCACGATCTATT
CAGCCATCTCTTAGACGACGAGTCTAAAGATACGGATGACCTCGTCTATGATTCTGAGCTGGCTGTCGTGGC
CGGTAGCGACTCTACAGCATCGACCCTTACTGCCATCTTGTTTCTTCTGGCAACCCATCCAGACAAGCAGGC
AATTCTTCAAAAAGAGGTTGATGACATGAATATAGGTGGACTGCCTATGCCAACCTCCACCCTAGTCAATATG
CCTTACCTCAACAGTTGCATCAACGAAGCACTCCGTTTATATCCAGCAGTGATGAGCGGCTCGCAGCGGGA
GTCCGGTCCTGGTGGAGTGATCCTTGATGGCCAGTTTATACCAGAGCATATGCTCGTATCCATGCCCACTTA
CACGATACACAGGGGTTAGTATCGATTCGATACAGTCATCCTTTGAGCTTAATTTTAACTCTGCCGTCTTTAG
ATAAGAGAAATTTCGTCAGGCCAGATGAGTTCATCCCGGAGCGATGGACGAGCCAATCACACTTGACTCTCA
ACCCAGATGCATTCATCCCTTTCTCTGTGGGTGTGTACGGATGCCTCGGTAAACCTTTCGCGATGATGGAGA
TGCGTATGGTAATTTCCTCTATAATGTCAAGATTCAATCTGAGACTTCTGCCTGGGCGGCCGATGAGCCATG
AAGCCTCTGCGCCAGCATACGATTGCTATATAATTCATATACCAGCGTTTTCTCTTGTTTTCGAACCCCGGGC
TGGAAATATCTGGGCATAA

>***nodD1*_gDNA**
ATGGATGCTGCTTCAACTCTTACGCATGCACCGGTATCTCAGCCATGGCAGTCCCTAGCTCAAGGGTTGGG
GTTCGTCAATGAGCATGAAGGGTACTGGTGGTCTAAGCTTGGACCTCCTCTCGGTAAAATGATGAACTGGGC
TCGATACTCGACATCGGAACAGTACAGAGTCCTAGCATTCCTTTACAAATATCTTCTCCCTGCCTGCGGCCC
AAAGCCTGGTGATGATGGTGAGCTGTTCTGGAAGGTTTTCATCAGCTATGATTACACGCCCATTCAGCTCAG
TCTCAATTTTCACAATGGCAAAATGACGCTGCGCACCGCGAACATACCAATTAGCGATAAATCGGGAACCGC
AGACGACCCAATCAACCAACAAGCTTCGGTAGACGCCATAATCCGCCAGGAACGAGTGTTGCCATCCCAGG
ATCTACGTTGGTTCAACCACTTTGCATCCCAGTACTTCTTCGACAAGGACACGGCAGCCTCTCTAAAGACCA
AGGTCGATAAGCTCCGAGTCCAGCAGGGAGTTCAGTGTATGCTGAGCCACGACTTTCCTGAGCGTGATGTC
CAATGCAAAGTGGCTTTCTGCCCGCTTTGGAAAGCCGTCGCTACAGGTCTTTCCAACAAGGAGATCATCTGG
GATTCGATTCTAGGGCTCGGAGATGACATCATCCCATACAAGCGAGCGCTTGCTGTCCTTGAACAGTACACA
TCGTCCGAAAATGCAGCGAAAGCAGGAGTGCGACCTGTATTCTTCGCTTTCGATACGGTGTTAAAAGATAAT
TACAAGAGCTCTCGTATCAAGATCTACTACCTCACAACACGGACAGCCTTCAACTCTATGGTCGACATCTACA
CACTTGGAGGCCTGCTAAAAGGGCCTGATATCCAAAAAGGAGTAGAGGCCCTTGAAGTGCTCTGGAAAGCC
GTCCTCAACGTCCCCGAGGGGTGGCCCGATGATAAAGATCTACCCATGAATCCACACCGTTGTGCGGCAGT
AATCTTCAATTTTGAGCTGTGGCCAGGAGCAGAGTTTCCGAGTCCCAAGGCCTATCTCCCAGCCCATTATTA
TGGCCGGCCTGATTTGGAGATAGCTGATGGTATGGACTACTTCTTCAAGCAGCAAGGGTTGGATGGGGTAT
ATGGTTCTTACAAGGAGAACTATTTGAAGTGCTTGTACGACCTCCTCACCTCACTCACATCTACCAAGCTATG
GCTAATAGATCAGAAGTACGAACTCACAGAACCAACTCACAGCCCTCCACCATGATATTTCTTTTTCATTCAA
AGGGTCCAATGCCTACGTTACGGTGTACTACAAGCCCGAGCTATCTCTAGATACCGAGTAG

>***nodO*_gDNA**
ATGGAGGTTTCCTCAAGTGAAACACTCCCCATTCTATGGAGAACGAGTCCTCCCTCAAACGACTATGAAAAT
GCAAGATGCAGAGTGTTCAATGGCAGGCAGCCTGAGCACTTCCCCCTTGCAATAGTCAAGGCAAACAAGGT
CGAGCATATTGTAGCCGCCGTGAAACTGGCAGCAGAACTAGACGCCTGTATCGCCGTTCGTTCGGGCGGCC
ACAGCCTCTCTTGTTGGACTATACGCCATGGGGCGATCCTTATTGACCTCGAAGATTATCAGCACTTGAGCT
ATGACGACGAGATCCATGAAGTGCAAGCTTCACCCAGTACACTTGGTGCGGACCTACTTACGTTCCTTGCGA
AGAAAAAGAGGTTCTTTCCCGTAGGTCACTCTGGAGACATTGGCTTAGGTGGCTATCTCCTTCAGGGCGGAA
TTGGGCTCAATTCTCGGGTATGAACGTTGACCCTACATACTCAGTAGCTTTCGAATACTGACGATATCTCTAG
GGATATGGGTATGCCTGTGAATACATTACCGGGCTTGATATCATCACCGCTGATGGTGAAATCAAGCATTGT
GATAAGACGGAAAACTCTGACCTATACTGGGCTGCTCGTGGAGCTGGACCGGGTAAGTTCCTCAAAGCTGT
CTGTACCAGTGTGTAAGAGTTCTAATAAAGCAGAATTCCCTGCAATCGTTATACGGTTCTTCCTGAAGACGTG

CCCCCTCTTGCCGGTGTGTAAGCGGAGTAGATACGTCTGGCCGGCGGCCATGTATGGGAAGGTTTTTAAGT
GGCTTGAAGAGGTAAGTTTAATCCCAAGGATCTAGGACGGCAGGGGTTAATCTGATCCGCGTCTTTACAGCT
ATTGAATTCCTTGAGCGAGGACGTCGAGATTGCCGTTTTCGGGTTTGTATTGCCCCGACTCAACCAGCCAGG
CTTAGTCCTCCATGCAACAGCATTTGGTGACTCTAGCGAGAATGTCCGGGAAAAGCTCACGCCTATCATCAA
AAATCATCCTCCAGGGACTTTCTTGGCTGAAGATTTCGTGAGCACCAACTTCCCCGAAGACTACGACTTAGG
TAAGGATACCATGCCGCGGGGCGCTCGCTATTTTACCGACAGCGTCTTTCTCAAGCCCGGTATCGACTTCGT
CGCAACCTGCAAAGGAATGTTTACGGAGCTCAAACATCCCCGTGCATTGGCGTACTGGCAACCGATGAAGA
CCAATATTGATCGCATCCTTCCCGATATGGCGATGAGCATCCATAGTCATCATTACGTGTCACTACTTGCCAT
CTATGAAGACCCCAGTGAGGACCAACAGCAAATATCCTGGATCATAGATCGTATGAAAAGTCTTGAGCCGGC
AATCCTGGGAACTTTCATAGGGGATGCACACCCAGTGGAAAGGCCATCTAATTATTGGTCCGAGGAAGCTGA
AGAACGGGTGATCACTATTGGGCGGAAGTGGGACCCTAGTAGTAGAATTCGGGGTATTGTGTTGAGCGATG
CCTAA

>*nodD2*_gDNA
ATGGACTCTACCGGGAGACAGCCTCTCAGCCAGGACGGGCAGCCATGGCAAGCTCTTGCCTCTGGCTTGG
GATTTCCAGATGAAGACCAGAAGTATTGGTGGTCAGTCATGGCGCCCCTATTAGGGGAACTGATGAAATGG
GCCAACTATCCTGTTGATAAACAGTATCTTGTTCTTGCGTTCTGCCATGAATACATACTTCCATTTTGCGGAC
CTCGTCCAACGGCCGAAGGCGGTATCTTCTGGCCCACGTTGATCACCAAGGATGGTACTCCGTTCGAGCCG
AGCCTTAACTTCTACAAAAATAAAGCCACTCTACGAGTTGGATATGCGCCTGCATGCGAGCTTTCAGGAAGC
AATGACGACCCTATCAATCAACGAGCCCCAATTGCGGCATTGGAGCACCAAAAGAGAGTCCTACCACAGCA
GAATCTCAAATGGGTGGATAATTTCAAGAAAGCATGGTTCATCGATAATGACGACGCAGTAGCCTTGAAGGC
ACGCGTACATAACGAGCTCTTTGAACAGGCAGCCGTCCAATGCTTGATAGGTTATGTGTTTTCCGACTACAC
GCAGGTCAAGGTTGCCATGAGTCCTCTTTGGAAGTCAGTCCAGACAGGCCAGCAAATTTCGCGAGTGATCT
GGGACACCTTTCGCCAGCTTGGTGACGATGCTTCCAGTTACCTAGATTGCCTATCAGTGCTAGAGGAGTATA
CTGAGTCTAAACAGGCCAAACTAGCACAAGTACAGCCTTCCTTCGTCAATTTCGATGTGAATCTGAAGGGGG
ACTACCAGCAGTCGCGACTTAAGGTATACTATGCTACACCATGCACTGCATTCGACACGATGGTCCAGGTTT
TCACTTTGGGCGGAAGGCTCAAGGGCCCTGAAGTTGACCATGCGATTGAATGCCTTCGTGTCCTATGGCCC
AGCGTGCTTGCAGTTCCTGAAAACCATCCGGACGACCAAGACTTGCCTCGACGGTACCACTCTGTAGCAGT
GACCCAGTTTAACTTCGAGCTCTGGCCGGGAGCGAAACTGCCTGTTCCCCAGATCTATCTCCCGACTAACTT
TTACGGTCGTGATGAATTGGAAATTGCGGAGGGCCTAGAGGGATTTTTCAAGACTCTTGGCTGGAGTGAGC
CTTTCCATGCCTATAAACAGAATTACATTGCAACATGGTTAGTCTATCGCCTTTCTCTCGCTCATTGCCCTGAT
CCAATGAAGTATTTTACTCACACGATGTAGTGCCACGCCCGAAGGGAAATGGAAAGCGATCCAACACGACG
TATCATTCTCGTTCAAGGACTCAAGCCCATATGTGTCGGTTTATTATAAACCAGAGCTCTCGGTCCTATCCTC
ACCGTCGTAG

>*nodR*_gDNA
ATGTTTGATATTGATTTTGGCATTCTATTTCCAATTTCCTGGGAACAATCCCCTATCTTTCTGGCTGTTGGGCT
AATATTCGCATTTGCGACCTTATCTCCGTGGCTCCGTTCCGGGGAACGCCTGATCAATGGTCGCGAAGGCTT
CGAAATACTGTGGACGAATGCGAAGAAGCGATATCAAACGAAGGGGCGATCTGTTATGGAGGCAGGCTTTT
CAAAGGTAGGATACATCATTGTCTCCAAATGAAAGACATCAAATCGTGCCGGGTCATATGTGCTGATACTTG
GAGATAGTACAACGATTCTTTCTACATGATGACGGATAGTGGTACCGAAATGGTCCTACACCCGAGATACGT
CGACGAGATCCGGAACGACCCAAGGTTGGACTTTCACAGATATATGAAAACTGTAATAGAATTCGATCCATTT
TCTATATATCCGGGCTGACGTCTACGTACGCGAGGTCCTACATGGGCTCAGACAGAAAGGTGGCGAGATAA
ACCGAGATCTTATCCAAACCAAATTGACTAGATCTGTGGGTAAGGGGTTGTCGTCCTCTAGAGAATGCAGGC
AAAGGGAAGGACAAGAGTAAACTAATGCCCATCCTTGTATAGGTAGGCTCATTGGGTCCATATCGGCTGAAA
TCGAGGACGCCCTTCACAATCGATGGGAAGAAGGCGAAGGTGAATATGAAGAGTACCTTCTAGTGCTTCCC
CCTGCTGAAATGCATACAGAATGGCATGAGATTGTATTGCTCTCCGTAATGATCCCGGTTGTTGCGCAAGGA
GTATCCAAAATGTTTGTCGGAGATCCATTGTGCCGTAACAAGGATTGGATAGGTATGATCTTAAGGCATACAA
GGAGTGTCCAGACAGCACTGCGGTCCCTTCGATTGTGGCCTTACTTTCTTAGACCACTTGCAGCCAGATTTC
TGCCAACCTGTCGACAAGTCACTGCAGAAATAGAAGAGGCCAGGCGCATCATCAATCCAGTACTTGAGAAG
AAGCGTGCCGAGAAGCTCGCAATGATTCAGAAGGGAGAAAAACCTCCAGAGCCCAACACCTACATGGATTG
GTTAGAAGAATCCGACAAGGATGAATTCTATGACCCAGTTGTTGCTCAGCTGAAGATTTCCATGGCAGCTAT
ACACGCTACCTCGGACCTCTTATCGCAGACAATTTTCAGCCTATGTGACAGCCCAGAGCTAGTCAAGGAACT
GCGAGCCGAAGCTGTGTCTGTTATAGGAGCGTACGGCTGGGGGAAGGAAGCAATCTACAACCTGAAGCTGA
TGGACAGCGTCCTGAAGGAGACTCAACGCCTAAAACCTATGCAAATCAGTAAGCGCCGGTGCCGCCGTGTG
GACTTGCTATCGCTAATCGCTTTGCCTTTGCTTGCTCTCTAGACTTGACGCGACTAGCTCTAGATCGCATAAA
ACTTTCCGATGGCACGGTTATCCCTAGGGGCTCCAAAGTTCTTATCTCCTGCCACAACATGTGGGATTCGAA
CGTCTACCCCAACGCAAACCAGTATGACGGTCATCGCTTCTATAAACTTCGGCAGCGCGCTGGAATGGAGA
ATTCCGCACAGCTATCAACACCTAGTCCGGATCATCTGGGCTTCGGACTGGGAATGTATGCCTGTCCCGGTA
GGCATATCGCATCGACCGTTATGAAAGTCACACTATGCCATATCCTCCTGAAGTATGACTTCGAATTGGCGG

AGGGCTGCACCCCTAGAGTGATTGAATATGGTAGTTTCTTGTTAGCTGATCCAACGGCGAGAGTTTCTATCA
GACGGCGGAAAGAAGAGATTCAACTGTAA

>***nod*Z_gDNA**
ATGATAACAGCATCCACTTCCGTGTTTGGTGGCTTAATCCTTGCCTTTATCTTCTCACTTTTATATAAGAATAA
GAAGACCCGTATCCCAGCAGAAATCGACCGCGTAAGGACGGGAGGTTTCCTAGCACACATTAGAGCCTTTG
GGTGCAGGTTAGGTACGCGAGTTGATGATATTCGCAATGGATATAATAAGGTTAGTTGCTCCTATAGCTGCT
CACAACAAATACTCAACACATGATATCTCAATACGCAGTTCAACAAAAACGGTAAGCCGTTTGTGATCCAGGA
TTCTACCTTTATCCCACAGGTCGTCATACCACCTCAATACCTGGGATGGTTGAAGGAGCAACCAGAGAAGGC
TCTTTCTGCAGAAACCGTGAGGCTAGAACAGCTTGGACTTCGCTATTTGGTCCCTAGCTCAGATCCCGAGAT
GGTCCATTTGTTGACAGACGTCGTGTGTCGCTATCTTACCCGCAATTTTCAAAGAGTACAAGAGCGTTTATAT
GAGGAGCTCCATATGAGTACGGATGAAATCATGGGGCTGGAGGCAACTGAGTGGCGTCAGATTTGTCTCCA
TGAAGCGATGGAAACGATTCTTCGAAGGATGATTAGTTGCGTCCTGATTGGACTCCCATGGTGTCGAGATGA
GGAATGCTTAAAATCGTGGACCGGATTTCTTCACTGCATGGCCATTGCAGGAACTATTCTAGGGGCAGTAAC
ACCTTGGTTCCTACGGCCACTCCTCGGACTACTGCTCAAGCCGCCCGTTGGCTATATGCGCAGAAGATCAC
TGCGTTATCTTACTCCGATCTTTACTGAAAGATGGAAAAAAATTGAGAAGCATGAGAAGAGCTCACTGACGA
CGCGTGAGCTGCCGGATGACTTCGTAACATGGTGTATTCAGGAGGTACGGAATGGTAATATGAAAGAAAAA
CACGATGCTCTTAGCCTTGCGTCGGAATTTCTATTCTTTGTGAGTCACTCTACGCGTTGGTATCATGGTTTCC
CATATTCAATACGCGATGCGATGGATCGAGGCTCTAACGAGAACGCAGTCGATCGCATTCTTTGACGCGCCT
ATAGGAGCTGCAGAAGTCACCATGCTCGATCTTCTAAGTGCGGATCCCACAATAGGTTACTGGGAAAAGCTT
GTGGAAGAAGCCACTACTGCATTCAGAACAGACGAGGACTGGATTCATGCAGGCACTGTGTCAAAATTGGC
ATACACGGACAGCGCAATCAGGGAGTCTCTACGCCGAAACCCTTTCAGCATCCGAAACGTGACCCGAGAGG
TGATAGGAAAAGATGGGCTAACACTGCCATCCGGTACGCGTCTGCCACAGGGCACTTGGATCACGACCGCT
CTCGCCAATATACATCATGATGCAAGATTTTACTCAAATCCCACCGAATACCAGCCTTTCCGTTTTGTGGCCA
GAGAAGCGTTTCACACAGAGGGGAAAGAGGGTAGCGAAAAGGTTTTACAGCCCTCTGAAGCTATCTTGACG
AGCACAATTGATGAAAGGCTCTTGACATTCGGATATGGGCGTCGAGCATGGTACGAAATCCAACTGTCTCCT
AGAGTCTCAAACTCTTATATCCTGTCTCCAGTCGAACCTACCTAGGTACTAATGCTTTTGCAATTACTGTAGC
CCCGGCCGATGGTTCGCTTCACACATACTGAAAATGTTGATCGCGTATATCACGATCAATTATGACATACAG
CCCTTGACGGGACCGCCGAAAAAAGTCAAATTCGCAGATTTCACCGTGCCATCGCCGAGTATCAAAATCATT
GTGCGCCGGAAGAATCTCGCTTACCTTAGGCAACGTGAGCGTTGA

>***nod*Y2_gDNA**
ATGGCGCCTGATAGACTTGGTCCCGAGGGCACAGCACGCCCCAATTCCAGTGGCATCTCAGTTATCGTGGT
CGGCCTTGGAATTGCTGGCCTAACAGCTGCAATTGAGTGCCACCGGAAGGGCCATTCAGTCATTGCCTTCG
AGCGAATGAAGGATGTCGAACCCTTTGGTGAGTTTGGCTTCGAATCCCCATATTGGTTACTCCGCCAACCTA
GTAAAGGCGATAGCATCATCATCGGCAGCAACGGCGGCCGTATTTTCGGCAAATGGGGCCGCGGTGAGGT
ACGCAATGCTATGCAAGCCTGGCGATATACGCCTACCCATGCCGACATTTACGATACCGCCGGAAGATTCAT
GGCCCAGTCTGAGATTCCCAAAGCTGCGGATGACATGTACTTCACTCTTCGAGGCAGACTAGCGAAGACCT
TCTACGAACACGCACAAAGCCTCGGCATTGATATGAGGATGGGGTCGAAAGTGACTGAATTTTGGGAAGAC
AGCAATCGGGCTGGAATCGTTGTGGAAGGAGAGAGGTTTGAGGCCGACTGTGTTATTTGCGCTGATGGCAT
ACACAGCAAGTCTCGCTCCTTGTTCACTTCTCTAAACGCTCAACCATTTCGTTCTGGTTTCTCTATTTTCAGG
GGGAAAGCGGACGCTAATGCGATTATTGCCGATCCCGATGCGAAATGGATCCTTGACCAGACAGAGAACAC
CGATCAGTTCAAAGTGTTTCTGGGGAAGGAGATCTGTGTTGTCATAATTACCTGCGGGCTAGGCCGTGCAGT
GGTCTGCAGTGCTATGCATAGGGTGAGCAATACCTATCTCCCAGATACCAATAATATATAGCAGGTAGGCTT
TGGACTATATCTTCCTGTTCTGCTCACACAAGACTGCCCCAGGACCTAAATGAAGCGGAACAGTCGTGGTCG
ACCCATGCCAACCCGGATGATCTATTGGACGCCATCAAAGACTGGCCGTGCAGGCGCCAGATCGAACCAAT
CGTTCGGAAGATATCCGAAGACCAGTTCATCGACTATCCCCTTCTAACTGTGTCTCCACTGGACACGTGGGT
ATCCCAGCACGGGCGGATGATTCTAATAGGAGATGCTGCTCATCCATTCTTTCCGACTTCCGGACAAGGAG
GCGCACAAGCCATGGAGGATGCAGCTGTGCTTGCAATTTGCCTCGAGTTGGCAGGGAAAGGAAACATCCCC
CTGGCTCTTCATGCAACAGAAAAGATCAGGTCTAGTTTCCCGGCCCTTCGGGCTACTCTCCCTTTCCATTTTA
TCCGTTTTCCTTACTTTCGCTAATACCACGTAACTTCCAGAAAGAGCCGAGCTTCAGTCCTCCAACTAAACAG
GACGTATTCAGAAGGGGTTCAACTAGCACCAGCGCTGCCGAAATCCAAAGACAGTATGTCTGTTCCAAATGT
TCCAGTAATGGATTGGATCTGGCATCACTGCTGCCAGTCCTACGCATATGATGAGTTCGACAAGGTAGCGGA
GGCGATTCAAAGCGGGAGTGAATACATTCCACATAATCTTCCAGAAGATGGTACGTAG

>***nodJ*_gDNA**
ATGGAACTTATCGTTATTATCATAACTCTCGCTTTTTGCATCCTGCTGTACGGTACTCGATGGAGAGCTGCTC
TCGACCCCCGCGAGCCCCGTTTGATATCCCCTACGGTACCACTCATCGGGCACATACTTGGTATCGCGACT
GACGGCTTTGGGTATTTTTCCAAGCTCAAGTATGGTTAAGACTTCTACATGTTGCTTAGTTGAAGCGAGGAAG
ATATGCTGTATATAGTACATAGACTAACGGTGTAGATAGTGACAAATATGGACTTCCCGCCTTCTCGCTCCAG

ATGCCACTATCAAGGCTTTATGTGATCACGTCTAGCGAGCTAGTCCCGGCTATTCAAAGACAGTCCCAGAAT
ATACGATTCGATACGTTCGAATTTACACTGGCCGCAGAACGTGTAGGTGGTGTCAGTGGGCCTGGACTTAAA
TTGCTTAAAGGTAGTATCGTGGATGAGCTACAACATGCTATGCATCACGCACTTATAGGACATGGCTTGGAT
GCCATGAATCTTTCCATGATCGAGGCTATAAAGCCCTCAATAGACGAACTCCAAAGCCAGAAACAAGCGGCA
TTCGACTTATTTGCATGGTGTAAACGTTCTATCACAATGGCGAGCACAGACTCGGTTTACGGTCCTATGAATC
CATTCCGATCCATAGAAGTTGAAAGGGCTTTTTGGTAGGTCAAAGCACAGAATAGAACCCCAATGCTGTGTA
AGAATTGCTGAGCTCATACGTGGTGTAGGGATTTCGCGTCGAATATGAATATGATAATCCTCAATGTCCTTCC
CTTTCTCACAGCACGAAAGTCCCTAGATGATCGTCGTAAGGTTGTGGACGCTCTAACAGAGTATTACAACCT
GGGAGGCCACGAAAACTCCTCGGAAATGACGTATGGTCGATGGGAGGTCCAATATAACAAGGGAATCACCA
CACAAGATATAGCTCGAATGGAAATTGTTAACGCGATAGGTGTCCTATCGAATACTGCGCCTTCCACATTTTG
GACCCTTTTCGAAATTTACTCACGCCCGAGTTTGCTGAGGGATCTTCGACAAGAGCTCGTGGCTTCTGCTGT
GTACACGCATCCAGGTACCGATGGAGGGATAGTTCGTACTATTGATCTCAGCGCTGTACGCGCCAAGTGCA
GTTTGCTTCTCGGTACCTTTCAGGAGGTCCTCCGTATGAGGTCTAATGCTATAGTCACGCGCATGGTCCATG
AGGATACCATCCTCAACGAGCGAGTTCTATTCAAGAAGGGGTCCGTGATCGTGATCCCTGCAAGATGTGTTA
ATCGCGAAAAGTCAGTATGGGGCGAAACAGGGGATAGCTTCGACGCATATCGATATCTCGGCCAGGGGAAG
AGTGGGACATCGCGAAGTGGCCGTAACGTCACACGCGTTGCGTTCCAGTCTTTCGGGACAGCACCTAACAT
CTGTCCTGGGCGCCACTTCGCATCTGGGGAGATATTAGCGGTAGTAGCCATGGTGATTTTACGGTTTGACAT
GGAGCCCGTGCTCGGAGAGTGGATTCCTCCCAAGGCCAACGTGAAAACACTTGCAAGTTCCATACAGACAC
CGGCAGGGGAGTTCATGGTGACTCTCAGAGAGAGAAAGGAATTCAAGGATGACAAATGGGATTTCAGGGTC
ACAGAAGGCGACTCGAAGTTTCCACTGCTAGTTGGGTAA

>***natR***
ATGACCACTCTTGACGACACGGCTTACCGGTACCGCACCAGTGTCCCGGGGGACGCCGAGGCCATCGAGG
CACTGGATGGGTCCTTCACCACCGACACCGTCTTCCGCGTCACCGCCACCGGGGACGGCTTCACCCTGCG
GGAGGTGCCGGTGGACCCGCCCCTGACCAAGGTGTTCCCCGACGACGAAAGCGACGACGAATCGGACGC
CGGGGAGGATGGCGACCCGGACTCCCGGACGTTCGTCGCGTACGGGGACGACGGCGACCTGGCGGGCTT
CGTGGTTGTCTCGTACTCCGGCTGGAACCGCCGGCTGACCGTCGAGGACATCGAGGTCGCCCCGGAGCAC
CGTGGGCACGGGGTCGGGAGAGCGTTGATGGGGCTCGCGACTGAGTTCGCTCGCGAGCGAGGCGCCGGG
CACCTCTGGCTGGAGGTCACCAACGTCAACGCACCGGCGATCCACGCGTACCGGCGGATGGGGTTCACCC
TCTGCGGCCTGGACACCGCCCTGTACGACGGCACCGCCTCGGACGGCGAGCAGGCGCTCTACATGAGCAT
GCCCTGCCCCTGA

>***neoR***
ATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTG
GGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTT
TTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGG
CCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATT
GGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTG
ATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATC
GAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGC
TCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCA
TGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGC
TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAA
TGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCT
TCTTGACGAGTTCTTCTGAG

Genes used in this study (cDNA)

>*nodM*_cDNA
ATGTCTACCCCTGAGTTCAAGGTGATAATCGTTGGCGGCTCCCTTGCGGGTCTAACATTGGCCCATTGCCTC
CTCCGCGCCGGAATTTCCCACATCGTTCTTGAGCGCCGTAGTGTAATCGCACCAGAGGAAGGTGCCTCAAT
AGGGATTCTTCCTAATGGCGCTCGTGTCTTGGATCAACTAGGCATCTATGAACATATTCAAGACACTACTGAG
CCATTGAGTACAGCTCATATACGATATCCAGATGGGTTCTATTTCAGCAGTCGCTATCCCGAGATCATAAAAG
AAAGGTTTGGGTATCCGATAGCCTTTCTCCCCCGGAGGAGACTTTTGGAGATTCTCTATACATCGTACCCTG
ACCATTCTAACATCTATACCAACAAGAATGTGATCAAGGTGCAAAGCCATGATAATCAAGTTTCTGTCTTGAC
AGAGGATGGAAATACATATCGAGGGGATCTTGTGGTGGGCGCCGACGGTGTCAACAGTCGAGTTCTATCTG
AAATTTGGAAACTGGCAGGTAATCCCTCTCTCACGAAACGGGAAGGAAGAGGAAGGACCATCGAGTACGCT
TGTGTGTTTGGGATATCCTCGCCAATCTCAGACCTCAAGCCCGGCGAGCAGGTTAACGCGTTCTACGACGG
ACTCACAATCGTCACGATCCATGGTAGAAACGGAGAGATATTCTGGTTTTTTATCAAGAAACTCTCGCGTCGC

TATATCTATCCCGATCTCATCAGACTGCAGCAGAAAGACGCAGAAGGGATCTGTGAAGAAGCCAAGTCACTA
ACTGTCTGGAAAGGTGTTACATTCGGTGACATTTGGGAAAGAAGAGAAACAGCATCATTGACCGTTTTGGAT
GAGTTCTTGCATCACACTTGGAGTTGGGACAGATCGGTTTGTGTGGGCGACAGTATCCACAAGATGACTCCA
AATTTTGGCCAGGGTGCGAATACTGCTATCGAGGATTCTGCCGCTCTAGCCAATCTACTGCACAGTTTGATC
AAGGAGAAACGAGCCGAAAAGCCGACTGACTCAGATATATCACTGCTTCTGAGGCAGTTCAAATCACAGCGT
CTTCGACGCGTTCAGAAGATATATAAAATGTCAAGGTTTGTTACGCGGCTTCAAGCGCGCGACGGGTTGTTG
AATACTCTTCTGGGACGCCATTACGCCCCATATGCAGCAGATCTTCCAGCTAAGATTGCCTCTGGATGCATT
GCTGGTGCGGAAGTTCTGGATTACCTCCCACTTCCCAAGGTAACTGGGGCGGGCTGGAATAGGGGCCATC
GTAGATCTACCATGTACATCCTGCTAGGATTCACAGGGGTATTCACCTCTGCTCTGGCCATGGTAGTGTTGC
TCCATATTAGAGACATTGCTTCATAG

>***nodW*_cDNA**
ATGACTCTCGCTATTCTGGGCATCAGTTGCCTGGCAGCCGTCCTATCTCATACTATTCTGTTCATCCGAGGA
GAGTGGGACAAGCACGCCCCGGTCCTGTTCCATTGCTTCATTACTTGCCCCATTGCGCTCATATGCTCGTTG
TTGGCAGTGGGATTATATGGGATCACTTTTATGTTCATCCAAGTTTGGCTATGCTATATCTCTTCTATCTGTCT
TTCAATCATTATATATCGGTTGTTTCTGCATCCCTTGAGAGATTTCCCCGGCAAGGTATCCGCCAGGCTGAG
CACGTTGAGATGGATGAAAACAGCAGCCGTAGATAGAAAATGGCACCTTGAAGTACAGAATATGCATAGTGT
CTATGGCGATTTCGTCCGAATTAGGCCTCGTGAGATATCCGTCAACAACTTCAATGCCATTCGCGATTTGGC
CAAGTGTGCCAGGGGTCCATTCTACGACGTGAATTATCCCCACAAGTCCCTTCAAATGACAAGGGACAGGAA
TTTCCACCTCCGACGGAAGAAAATTTGGCAAAAGGCCTTCAGTCCTAATGCTATGCTAGCTCATGAAGGCCA
TGTCCGGCAATGTTGCCAAGAACTCATGGGCTATCTCTCATCTAGAGGAGGTCAACCGGTAGAAGTAACTGA
ATTGATGCAAAGAGTCACATTTGAATCAATGGGTTGGCTTGTATTTGGAAAGGCCTTCAACATGATTCAAAAA
GGCCTTTCGCATCCCGGCTTCACGCAACTGAGGTCTGTCAAGCAACTAGGGGGGCTCCTTCTGTGGGCGCC
TTGGACGTTGATCCTTCTCCGCAATCTACCGTTCTTGAGGTCCAAGGCAACAAGTTGGCTACAGTGGTGCTC
ACAGGAGGTAGTGGAAAGGAAGCAAAATCCAACTATGACTCACGATCTATTCAGCCATCTCCTGGACGACGA
GTCTAAAGATACGGATGACCTCGTCTATGATTCTGAGCTGGCTGTCGTGGCCGGTAGCGACTCTACAGCATC
GACCCTTACTGCCATCTTGTTTCTTCTGGCAACCCATCCAGACAAGCAGGCAATTCTTCAAAAAGAGGTTGAT
GACATGAATATAGGTGGACTGCCTATGCCAACCTCCACCCTAGTCAATATGCCTTACCTCAACAGTTGCATC
AACGAAGCACTCCGTCTTTATCCAGCAGTGATGAGCGGCTCGCAGCGGGAAACCGGTCCTGGTGGAGTGAT
CCTTGATGGCCAGTTTATACCAGAGCATATGCTCGTATCCATGCCCACTTACACGATACACAGGGATAAGAG
AAATTTCGTCAGGCCAGATGAGTTCATCCCGGAGCGCTGGACGAGCCAATCACACTTGACTCTCAACCCAG
ATGCATTCATCCCTTTCTCTGTGGGTGTGTACGGATGCCTCGGTAAACCTTTCGCCATGATGGAGATGCGTA
TGGTAATTTCCTCTATAATGTCAAGATTCAATCTGAGACTTCTGCCTGGGCGGCCGATGAGCCATGAAGCCT
CTGCGCCAGCATACGATTGCTATATAATTCATATACCAGCGTTTTCTCTTGTTTTCGAACCCCGGGCTGGAAA
TATCTGGGCATAA

>***nodD1*_cDNA**
ATGGATGCTGCTTCAACTCTTACGCATGCACCGGTATCTCAGCCATGGCAGTCCCTAGCTCAAGGGTTGGG
GTTCGTCAATGAGCATGAAGGGTACTGGTGGTCTAAGCTTGGACCTCCTCTCGGTAAAATGATGAACTGGGC
TCGATACTCGACATCGGAACAGTACAGAGTCCTAGCATTCCTTTACAAATATCTTCTCCCTGCCTGCGGCCC
AAAGCCTGGTGATGATGGTGAGCTGTTCTGGAAGGTTTTCATCAGCTATGATTACACGCCCATTCAGCTCAG
TCTCAATTTTCACAATGGCAAAATGACGCTGCGCACCGCGAACATACCAATTAGCGATAAATCGGGAACCGC
AGACGACCCAATCAACCAACAAGCTTCGGTAGACGCCATAATCCGCCAGGAACGAGTGTTGCCATCCCAGG
ATCTACGTTGGTTCAACCACTTTGCATCCCAGTACTTCTTCGACAAGGACACGGCAGCCTCTCTAAAGACCA
AGGTCGATAAGCTCCGAGTCCAGCAGGGAGTTCAGTGTATGCTGAGCCACGACTTTCCTGAGCGTGATGTC
CAATGCAAAGTGGCTTTCTGCCCGCTTTGGAAAGCCGTCGCTACAGGTCTTTCCAACAAGGAGATCATCTGG
GATTCGATTCTAGGGCTCGGAGATGACATCATCCCATACAAGCGAGCGCTTGCTGTCCTTGAACAGTACACA
TCGTCCGAAAATGCAGCGAAAGCAGGAGTGCGACCTGTATTCTTCGCTTTCGATACGGTGTTAAAAGATAAT
TACAAGAGCTCTCGTATCAAGATCTACTACCTCACAACACGGACAGCCTTCAACTCTATGGTCGACATCTACA
CACTTGGAGGCCTGCTAAAAGGGCCTGATATCCAAAAAGGAGTAGAGGCCCTTGAAGTGCTCTGGAAAGCC
GTCCTCAACGTCCCCGAGGGGTGGCCCGATGATAAAGATCTACCCATGAATCCACACCGTTGTGCGGCAGT
AATCTTCAATTTTGAGCTGTGGCCAGGAGCAGAGTTTCCGAGTCCCAAGGCCTATCTCCCAGCCCATTATTA
TGGCCGGCCTGATTTGGAGATAGCTGATGGTATGGACTACTTCTTCAAGCAGCAAGGGTTGGATGGGGTAT
ATGGTTCTTACAAGGAGAACTATTTGAAGTGCTTTACGAACTCACAGAACCAACTCACAGCCCTCCACCATGA
TATTTCTTTTTCATTCAAAGGGTCCAATGCCTACGTTACGGTGTACTACAAGCCCGAGCTATCTCTAGATACC
GAGTAG

>***nodO*_cDNA**
ATGGAGGTTTCCTCAAGTGAAACACTCCCCATTCTATGGAGAACGAGTCCTCCCTCAAACGACTATGAAAAT
GCAAGATGCAGAGTGTTCAATGGCAGGCAGCCTGAGCACTTCCCCCTTGCAATAGTCAAGGCAAACAAGGT

CGAGCATATTGTAGCCGCCGTGAAACTGGCAGCAGAACTAGACGCCTGTATCGCCGTTCGTTCGGGCGGCC
ACAGCCTCTCTTGTTGGACTATACGCCATGGGGCGATCCTTATTGACCTCGAAGATTATCAGCACTTGAGCT
ATGACGACGAGATCCATGAAGTGCAAGCTTCACCCAGTACACTTGGTGCGGACCTACTTACGTTCCTTGCGA
AGAAAAAGAGGTTCTTTCCCGTAGGTCACTCTGGAGACATTGGCTTAGGTGGCTATCTCCTTCAGGGCGGAA
TTGGGCTCAATTCTCGGGGATATGGGTATGCCTGTGAATACATTACCGGGCTTGATATCATCACCGCTGATG
GTGAAATCAAGCATTGTGATAAGACGGAAAACTCTGACCTATACTGGGCTGCTCGTGGAGCTGGACCGGAA
TTCCCTGCAATCGTTATACGGTTCTTCCTGAAGACGTGCCCCCTCTTGCCGGTGTGTAAGCGGAGTAGATAC
GTCTGGCCGGCGGCCATGTATGGGAAGGTTTTTAAGTGGCTTGAAGAGCTATTGAATTCCTTGAGCGAGGA
CGTCGAGATTGCCGTTTTCGGGTTTGTATTGCCCCGACTCAACCAGCCAGGCTTAGTCCTCCATGCAACAGC
ATTTGGTGACTCTAGCGAGAATGTCCGGGAAAAGCTCACGCCTATCATCAAAAATCATCCTCCAGGGACTTT
CTTGGCTGAAGATTTCGTGAGCACCAACTTCCCCGAAGACTACGACCTTGGTAAGGATACCATGCCGCGGG
GCGCTCGCTATTTTACCGACAGCGTCTTTCTCAAGCCCGGTATCGACTTCGTCGCAACCTGCAAAGGAATGT
TTACGGAGCTCAAACATCCCCGTGCATTGGCGTACTGGCAACCGATGAAGACCAATATTGATCGCATCCTTC
CCGATATGGCCATGAGCATCCATAGTCATCATTACGTGTCACTACTTGCCATCTATGAAGACCCCAGTGAGG
ACCAACAGCAAATATCCTGGATCATAGATCGTATGAAAAGTCTTGAGCCGGCAATCCTGGGAACTTTCATAG
GGGATGCACACCCAGTGGAAAGGCCATCTAATTATTGGTCCGAGGAAGCTGAAGAACGGGTGATCACTATT
GGGCGGAAGTGGGACCCTAGTAGTAGAATTCGGGGTATTGTGTTGAGCGACGCCTAA

>*nodD2*_cDNA
ATGGACTCTACCGGGAGACAGCCTCTCAGCCAGGACGGGCAGCCATGGCAAGCTCTTGCCTCTGGCTTGG
GATTTCCAGATGAAGACCAGAAGTATTGGTGGTCAGTCATGGCGCCCCTACTCGGGGAACTGATGAAATGG
GCCAACTATCCTGTTGATAAACAGTATCTTGTTCTTGCGTTCTGCCATGAATACATACTTCCATTTTGCGGAC
CTCGTCCAACGGCCGAAGGCGGTATCTTCTGGCCCACGTTGATCACCAAGGATGGTACTCCGTTCGAGCCG
AGCCTTAACTTCTACAAAAATAAAGCCACTCTACGAGTTGGATATGCGCCTGCATGCGAGCTTTCAGGAAGC
AATGACGACCCTATCAATCAACGAGCCCCAATTGCGGCATTGGAGCACCAAAAGAGAGTCCTACCACAGCA
GAATCTCAAATGGGTGGATAATTTCAAGAAAGCATGGTTCATCGATAATGACGACGCAGTAGCCTTGAAGGC
ACGCGTACATAACGAGCTCTTTGAACAGGCAGCCGTCCAATGCTTGATAGGTTATGTGTTTTCCGACTACAC
GCAGGTCAAGGTTGCCATGAGTCCTCTTTGGAAGTCAGTCCAGACAGGCCAGCAAATTTCGCGAGTGATCT
GGGACACCTTTCGCCAGCTTGGCGACGATGCTTCCAGTTACCTAGATTGCCTATCAGTGCTAGAGGAGTATA
CTGAGTCTAAACAGGCCAAACTAGCACAAGTACAGCCTTCCTTCGTCAATTTCGATGTGAATCTGAAGGGGG
ACTACCAGCAGTCGCGACTTAAGGTATACTATGCTACACCATGCACTGCATTCGACACGATGGTCCAGGTTT
TCACTTTGGGCGGAAGGCTCAAGGGCCCTGAAGTTGACCATGCGATTGAATGCCTTCGTGTCCTATGGCCC
AGCGTGCTTGCAGTTCCTGAAAACCATCCGGACGACCAAGACTTGCCTCGACGGTACCACTCTGTAGCAGT
GACCCAGTTTAACTTCGAGCTCTGGCCGGGAGCGAAACTGCCTGTTCCCCAGATCTATCTCCCGACTAACTT
TTACGGTCGTGATGAATTGGAAATTGCGGAGGGCCTAGAGGGATTTTTCAAGACTCTTGGCTGGAGTGAGC
CTTTCCATGCCTATAAACAGAATTACATTGCAACATGTGCCACGCCCGAAGGGAAATGGAAAGCGATCCAAC
ACGACGTATCATTCTCGTTCAAGGACTCAAGCCCATATGTGTCGGTTTATTATAAACCAGAGCTCTCGGTCCT
ATCCTCACCGTCGTAG

>*nodR*_cDNA
ATGTTTGATATTGATTTTGGCATTCTATTTCCAATTTCCTGGGAACAATCCCCTATCTTTCTGGCTGTTGGGCT
AATATTCGCATTTGCGACCTTATCTCCGTGGCTCCGTTCCGGGGAACGCCTGATCAATGGTCGCGAAGGCTT
CGAAATACTGTGGACGAATGCGAAGAAGCGATATCAAACGAAGGGGCGATCTGTTATGGAGGCAGGCTTTT
CAAAGTACAACGATTCTTTCTACATGATGACGGATAGTGGTACCGAAATGGTCCTACACCCGAGATACGTCG
ACGAGATCCGGAACGACCCAAGGTTGGACTTTCACAGATATATGAAAACTGTCCTACATGGGCTCAGACAGA
AAGGTGGCGAGATAAACCGAGATCTTATCCAAACCAAATTGACTAGATCTGTGGGTAGGCTCATTGGGTCCA
TATCGGCTGAAATCGAGGACGCCCTTCACAATCGATGGGAAGAAGGCGAAGAATGGCATGAGATTGTATTG
CTCTCCGTAATGATCCCGGTTGTTGCGCAAGGAGTATCCAAAATGTTTGTCGGAGATCCATTGTGCCGTAAC
AAGGATTGGATAGGTATGATCTTAAGGCATACAAGGAGTGTCCAGACAGCACTGCGGTCCCTTCGATTGTGG
CCTTACTTTCTTAGACCACTTGCAGCCAGATTTCTGCCAACCTGTCGACAAGTCACTGCAGAAATAGAAGAG
GCCAGGCGCATCATCAATCCAGTACTTGAGAAGAAGCGTGCCGAGAAGCTCGCAATGATTCAGAAGGGAGA
AAAACCTCCAGAGCCCAACACCTACATGGATTGGTTAGAAGAATCCGACAAGGATGAATTCTATGACCCAGT
TGTTGCTCAGCTGAAGATTTCCATGGCAGCTATACACGCTACCTCGGACCTCTTATCGCAGACAATTTTCAG
CCTATGTGACAGCCCAGAGCTAGTCAAGGAACTGCGAGCCGAAGCTGTGTCTGTTATAGGAGCGTACGGCT
GGGGGAAGGAAGCAATCTACAACCTGAAGCTGATGGACAGCGTCCTGAAGGAAACGCAACGCCTAAAACCT
ATGCAAATCAACTTGACGCGACTAGCTCTAGATCGCATAAAACTTTCCGATGGCACGGTTATCCCTAGGGGC
TCCAAAGTTCTTATCTCCTGCCACAACATGTGGGATTCGAACGTCTACCCCAACGCAAACCAGTATGACGGT
CACCGCTTCTATAAACTTCGGCAGCGCGCTGGAATGGAGAATTCCGCACAGCTATCAACACCTAGTCCGGAT
CATCTGGGCTTCGGACTGGGAATGTATGCCTGTCCCGGTAGGCATATCGCATCGACCGTTATGAAAGTCAC

ACTATGCCATATCCTCCTGAAGTATGACTTCGAATTGGCGGAGGGCTGCACCCCTAGAGTGATTGAATATGG
TAGTTTCTTGTTAGCTGATCCAACGGCGAGAGTTTCTATCAGACGGCGGAAAGAAGAGATTCAACTGTAA

>***nodZ*_cDNA**
ATGATAACAGCATCCACTTCCGTGTTTGGTGGCTTAATCCTTGCCTTTATCTTCTCACTTTTATATAAGAATAA
GAAGACCCGTATCCCAGCAGAAATCGACCGCGTAAGGACGGGAGGTTTCCTAGCACACATTAGAGCCTTTG
GGTGCAGGTTAGGTACGCGAGTTGATGATATTCGCAATGGATATAATAAGTTCAACAAAAACGGTAAGCCGT
TTGTGATCCAGGATTCTACCTTTATCCCACAGGTCGTCATACCACCTCAATACCTGGGATGGTTGAAGGAGC
AACCAGAGAAGGCTCTTTCTGCAGAAACCGTGAGGCTAGAACAGCTTGGACTTCGCTATTTGGTCCCTAGCT
CAGATCCCGAGATGGTCCATTTGTTGACAGACGTCGTGTGTCGCTATCTTACCCGCAATTTTCAAAGAGTAC
AAGAGCGTTTATATGAGGAGCTCCATATGAGTACGGATGAAATCATGGGGCTGGAGGCAACTGAGTGGCGT
CAGATTTGTCTCCATGAAGCCATGGAAACGATTCTTCGAAGGATGATTAGTTGCGTCCTGATTGGACTCCCA
TGGTGTCGAGATGAGGAATGCTTAAAATCGTGGACCGGATTTCTTCACTGCATGGCCATTGCAGGAACTATT
CTAGGGGCAGTAACACCTTGGTTCCTACGGCCACTCCTCGGACTACTGCTCAAGCCGCCCGTTGGCTATAT
GCGCAGAAGATCACTGCGTTATCTTACTCCGATCTTTACTGAAAGATGGAAAAAAATTGAGAAGCATGAGAA
GAGCTCACTGACGACGCGTGAGCTGCCGGATGACTTCGTAACATGGTGTATTCAGGAGGTACGGAATGGTA
ATATGAAAGAAAAACACGATGCTCTTAGCCTTGCGTCGGAATTTCTATTCTTTTCGATCGCATTCTTTGACGC
GCCTATAGGAGCTGCAGAAGTCACCATGCTCGATCTTCTAAGTGCGGATCCCACAATAGGTTACTGGAAAA
GCTTGTGGAAGAAGCCACTACTGCATTCAGAACAGACGAGGACTGGATTCATGCAGGCACTGTGTCAAAATT
GGCATACACGGACAGCGCAATCAGGGAGTCTCTACGCCGAAACCCTTTCAGCATCCGAAACGTGACCCGAG
AGGTGATAGGAAAAGATGGGCTAACACTGCCATCCGGTACGCGTCTGCCACAGGGCACTTGGATCACGACC
GCTCTCGCCAATATACATCATGATGCAAGATTTTACTCAAATCCCACCGAATACCAGCCTTTCCGTTTTGTGG
CCAGGGACGCGTTTCACACAGAGGGGAAAGAGGGTAGCGAAAAGGTTTTACAGCCCTCTGAAGCTATCTTG
ACGAGCACAATTGATGAAAGGCTCTTGACATTCGGATATGGGCGTCGAGCATGCCCCGGCCGATGGTTCGC
TTCACACATACTGAAAATGTTGATCGCGTATATCACGATCAATTATGACATACAGCCCTTGACGGGACCGCC
GAAAAAAGTCAAATTCGCAGATTTCACCGTGCCTTCGCCGAGTATCAAAATCATTGTGCGCCGGAAGAATCT
CGCTTACCTTAGGCAACGTGAGCGTTGA

>***nodY2*_cDNA**
ATGGCGCCTGATAGACTTGGTCCCGAGGGCACAGCACGCCCCAATTCCAGTGGCATCTCAGTTATCGTGGT
CGGCCTTGGAATTGCTGGCCTAACAGCTGCAATTGAGTGCCACCGGAAGGGCCATTCAGTCATTGCCTTCG
AGCGAATGAAGGATGTCGAACCCTTTGGCGATAGCATCATCATCGGCAGCAACGGCGGCCGTATTTTCGGC
AAATGGGGCCGCGGTGAGGTACGCAATGCTATGCAAGCCTGGCGATATACGCCTACCCATGCCGACATTTA
CGATACCGCCGGAAGATTCATGGCCCAGTCTGAGATTCCCAAAGCTGCGGATGACATGTACTTCACTCTTCG
AGGCAGACTAGCGAAGACCTTCTACGAACACGCACAAAGCCTCGGCATTGATATGAGGATGGGGTCGAAAG
TGACTGAATTTTGGGAAGACAGCAATCGGGCTGGAATCGTTGTGGAAGGAGAGAGGTTTGAGGCCGACTGT
GTTATTTGCGCTGATGGCATACACAGCAAGTCTCGCTCCTTGTTCACTTCTCTAAACGCTCAACCATTTCGTT
CTGGTTTCTCTATTTTCAGGGGGAAAGCGGACGCTAATGCGATTATTGCCGATCCCGATGCGAAATGGATCC
TTGACCAGACAGAGAACACCGATCAGTTCAAAGTGTTTCTGGGGAAGGAGATCTGTGTTGTCATAATTACCT
GCGGGCTAGGCCGTGCAGTGGTCTGCAGTGCTATGCATAGGGACCTAAATGAAGCGGAACAGTCGTGGTC
GACCCATGCCAACCCGGATGATCTATTGGACGCCATCAAAGACTGGCCGTGCAGGCGCCAGATCGAACCAA
TCGTTCGGAAGATATCCGAAGACCAGTTCATCGACTATCCCCTTCTAACTGTGTCTCCACTGGACACGTGGG
TATCCCAGCACGGGCGGATGATTCTAATAGGAGATGCTGCTCATCCATTCTTTCCGACTTCCGGACAAGGAG
GCGCACAAGCCATGGAGGATGCAGCTGTGCTTGCAATTTGCCTCGAGTTGGCAGGGAAAGGAAACATCCCC
CTGGCTCTTCATGCAACAGAAAAGATCAGAAAGAGCCGAGCTTCAGTCCTCCAACTAAACAGGACGTATTCA
GAAGGGGTTCAACTAGCACCAGCGCTGCCGAAATCCAAAGACAGTATGTCTGTTCCAAATGTTCCAGTAATG
GATTGGATCTGGCATCACTGCTGCCAGTCCTACGCATATGATGAGTTCGACAAGGTAGCGGAGGCGATTCA
AAGCGGGAGTGAATACATTCCACATAATCTTCCAGAAGATGCGGAGGACGAGATCAATAGGGCTTTTGGACT
CCAGAAGTCTCCCCCTACGTGA

>***nodJ*_cDNA**
ATGGAACTTATCGTTATTATCATAACTCTCGCTTTTTGCATCCTGCTGTACGGTACTCGATGGAGAGCTGCTC
TCGACCCCCGCGAGCCCCGTTTGATATCCCCTACGGTACCACTCATCGGGCACATACTTGGTATCGCGACT
GACGGCTTTGGGTATTTTTCCAAGCTCAATGACAAATATGGACTTCCCGCCTTCTCGCTCCAGATGCCACTAT
CAAGGCTTTATGTGATCACGTCTAGCGAGCTAGTCCCGGCTATTCAAAGACAGTCCCAGAATATACGATTCG
ATACGTTCGAATTTACACTGGCCGCAGAACGTGTAGGTGGTGTCAGTGGGCCTGGACTTAAATTGCTTAAAG
GTAGTATCGTGGATGAGCTACAACATGCTATGCATCACGCACTTATAGGACATGGCTTGGATGCCATGAATC
TTTCCATGATCGAGGCTATAAAGCCCTCAATAGACGAACTCCAAAGCCAGAAACAAGCGGCATTCGACTTAT
TTGCATGGTGTAAACGTTCTATCACAATGGCGAGCACAGACTCGGTTTACGGTCCTATGAATCCATTCCGAT
CCATAGAAGTTGAAAGGGCTTTTTGGGATTTCGCGTCGAATATGAATATGATAATCCTCAATGTCCTTCCCTT

TCTCACAGCACGAAAGTCCCTAGATGATCGTCGTAAGGTTGTGGACGCTCTAACAGAGTATTACAACCTGGG
AGGCCACGAAAACTCCTCGGAAATGACGTATGGTCGATGGGAGGTCCAATATAACAAGGGAATCACCACAC
AAGATATAGCTCGAATGGAAATTGTTAACGCGATAGGTGTCCTATCGAATACTGCGCCTTCCACATTTTGGAC
CCTTTTCGAAATTTACTCACGCCCGAGTTTGCTGAGGGATCTTCGACAAGAGCTCGTGGCTTCTGCTGTGTA
CACGCATCCAGGTACCGATGGAGGGATAGTTCGTACTATTGATCTCAGCGCTGTACGCGCCAAGTGCAGTT
TGCTTCTCGGTACCTTTCAGGAGGTCCTCCGTATGAGGTCTAATGCTATAGTCACGCGCATGGTCCATGAGG
ATACCATCCTCAACGAGCGAGTTCTATTCAAGAAGGGGTCCGTGATCGTGATCCCTGCAAGATGTGTTAATC
GCGAAAAGTCAGTATGGGGCGAAACAGGGGATAGCTTCGACGCATATCGATATCTCGGCCAGGGGAAGAG
TGGGACATCCCGAAGTGGCCGTAACGTCACACGCGTTGCGTTCCAGTCTTTCGGGACAGCACCTAACATCT
GTCCTGGGCGCCACTTCGCATCTGGGGAGATATTAGCGGTAGTAGCCATGGTGATTTTACGGTTTGACATG
GAGCCCGTGCTCGGAGAGTGGATTCCTCCCAAGGCCAACGTGAAAACACTTGCAAGTTCCATACAGACACC
AGCAGGGGAGTTCATGGTGACTCTCAGAGAGAGAAAGGAATTCAAGGATGACAAATGGGATTTCAGGGTCA
CAGAAGGCGACTCGAAGTTTCCACTGCTAGTTGGGTAA

>*nodX*_cDNA
ATGGAGTCAAGTGCGGATAATACAGTTCCCAAGCAAGCGGAATCTGTATTCATACGCGAAATTCTAGAGAAT
CCGCTAATGCCAAATCTTCCCCCTGAGCTGGCCGAAATCGCCAAATTTGTCTCCTTCGAGGGAAACGTCAAG
CCTAGCATTCCTGTTAACTGGAGGCTCGCAGAGAGCATCTCAGCTTTGAAAGCCTTTGAAGCAACGTTCCTG
AATTATCTGGCCCATCGGAAGTATGGTGTACGCCCTAGCAATGTCTCAATAAACACAGACCACGCAACCTTG
TTCCTCATGTCCCCCATGCTCACCCAAATCGAGGATAGAGGAGAAGTCCGTCCTTTCTCGCCATTTGATACA
CTGACAGCAGAACTTTTTCCGAACCGAGATAAACACCGCGCCAACGCATCTCTACAACGGGCATTGATTACC
AATATCTATAAAACAAAGGATGGCCGTTTCTACCATACACACGGGGGAATGAATCCAGAGCCCACCCTACAG
GCACTAGGGTTGCCCGTCGAAGGCGAGGTTCAAGACACTTCTGAGGTCGTAACCGACAGGATCCAACAAAG
ACTCTCAAAGTATGACGCCACACATCTAGACCATTTGCTAAATGAAGAACACCGCCAAGCCGGTACCATTGT
CTACAGCTCCGCTGAGTACTTTGCCAGCGAGCACGGCCAGAAAAATGGCAAGGTGGGACTGTATGAGGTCA
TTAGAGATCCAAATTCTTCTCAGCCAGCAGCCTGGTGGCCTGATCATCCTAACGCTCCATCGAGCCCCAAGA
GGCCGCTGGCCGGGCTGAAGGTGGTGGACTTAACGCGCGCTATAGCGGGTCCGACGATAACACGAAGTCT
TGCGGAGATGGGGGCTAGTGTAATGCGTGTTACTTCCCCAGACATCACGGACATGTCGGTGTTACACCAAG
ACTTGAATTGGGGCAAATGGAACTCATGGCTACGGTTGGATGTCGAATCTGACCGCCAAAAGCTGAGGGAC
TTGATATTGGATGCAGATGTGGTTGTAGATAGCTATCGGCCAGGTGTGATGAAACGTTTTGGCTTTGGCTAC
GACGAGGTGTTCAATCTGGTGCGGGACCGCAGTCGTGGTATTATTTACGTCAGGGAGAACTGCTATGGTTG
GTATGGACCCTGGAGCCATAGAAGCGGCTGGCAGCAGATTAGCGACGCGTGCTGCGGGATTTCTACTTCGT
ACGGCCATGCTATGGGTCTTGAAGAGCCTGTAACTCCTGCTCTGTTCAATTCTGATTATTGTACCGGAATATG
CGGGTCCACGGCTGTCTTGGACGCCCTTGTCCAAAGAGCCGAGAAAGGGGGAAGCTACAGGGTTGATGTA
TCAATCAACTATTACAATCAATGGCTTGTTCGATCAGTCGGGACATACGACGAGCATACCTGGACTGACCTG
TTCCGCCGTCACGACGCACCTGTTTTCAGACACTACCACTCCATGCAATACATGCTGCCTAAGCTGCTAACG
GCATTGTACAAATTCGACGGCGAGATTCTCTTTCAAAGGGAATTTTTCGGCCCATTCCGGTCCGGCGCGTTG
AACACGACGTTCATACAAGTAAGGCCTATCGCTCGATTCAAGGACGACGCGATTGAACTCAAGTATAATGTG
GGCACGAGAGGGAACGGAGTCGATGCGCCAACTTGGCCGGCAGATCTACGCCGTGAGATTGTGAGAGATG
AGGATGAGCAGGGGTCCGGATTCAGGTCAGGGTCAGGGTTAGGGTCAGGGTCCATTGCAGATTAG

Terminators used in this study

>T*paxG*
TCACTCCCGAGCAATATTGCTAAAATAAAGTGGGTCGGTGATCGTTTTACCGAATGAAAATGTCGCATTTCCA
CTTTAAGAGGTAAACCTCGGCATTTTTCTTGTGTCACTTGGTTAAAAGCACTATCCGTATATCTAACAACACAT
CAAGCTCACGAACTTGAAAATAGTTCATAGAATAACTTGAACCTCAACGACAATATTGATAAATGCTACTAATA
TTTGGAGCATAAACTCTGCATTTTCAGAGTATCGTTAAGTCCATGTCATTTAATTTTATGACCGCCAAGGTTAA
CTACTGAATGGTGAGACTTGATACCAACTTCCTTGCCAAAACATTTGAATGATCGCGGTCCGAGGTACATTTG
CTTCACTATTGCGAGTTCTATGCAAAAATTCGATAAATGGAAAATATGGGTTGAAAAACGCCTGGAAAATATT
TCATTCATAAAGGAATCTCTATCTCCGCGATACAGCGATCCTGGAAGAGAAATGCAGCACGATTGA

>T*paxC*
TTGGCCTTGTGAAATATGGGACTACGAAGTGAATCAGAAGTGAATATTGAGTATTTGAGTCTAGATAGTTCAC
AATTTTACTTGCAATTACCAAACTTTGCAGAATTCTACCAAGCTAGACTACATGCGGGGTGTATCCTTCTTAG
ACTCATGTCTGCAGTTCAGTATACAACGAGGATGCCTTCTATAGGATATTGAATCTACGTTGAAGTATTCCTT
ATAGAGGTTGCAAAGTGCCCAGCAACATCTGATATCCGACATGACAGAGAT

>T*paxM*

ACCATTGGAGCAATTTTTGGTTTTCCTTTCTTTCCAGTGTCTAGCTGTTGATTGTCACCTATCTTGCTTTCAAG
AAACATCACTAGATTTGAGTTCAATTAGTACTATATCGTTAAGCCGCTTTGTTTTAAACCGTCTTCAGGCTCCC
TGCGAGAGCATTCAAGCGATCTCAAATGGGCAGCGAATACTCACAGCTTGAGTAAGCCATACTATGAAAATT
CCATACACTATCCATACCTACTTTTTCTACGAAATATTTCTACTGCAGAGGGGTACTCTGACTGACCAGTACG
ATGGAAGATAGAACTTTCGAAATTGACAATCGACATTCCTTCACCACCTTCTACTACACTGGGTCAACTCCGA
AAACGAAATTCTATATTAAATGTGGCAAATCAAGGCCGGACAGTGTCTCATAATATATAAGCAATGTTCCAAG
GTTAGTATCATAATGCAAGTCCTAATCCACCTCTACGACAGAAGAATTACCTGGGTGGCCGTAGACCCACAA
GCAATCGAGTCCAGCCAGCGAACTTGGTAATTTCGTCAGCTGTAATGGATGTCAGTTACACCGATGAAAGAT
CCAATTTCCTCCTAGCTATAATTCCTCGTCTTTTTGGGCCATGTTTCATTCTATCGGGGATTTAAGAAGGGAG
ATCTCACCAGGCTATTGTATCAGATGCCCATTGTTTTGCAACTCTATTCTCGGAAATCGGCCAATTAGCAAAA
CAGTTAAACTGAAGTAAGCCAAAA

>T***paxB***
GGTTAACAATGAGGAACGATGAACAGTGGTTGAAATTCAAACTCAGAATCAACCTTTGCATACTAATATCATA
TGATACAGATTGCATAGTGTCAGCCACAAATGATCATTGCTTACTACTAAAACGTGCCAATTATATCATGGATT
GGGTCGCCTTGGACATGCACGTTACGCGTATCGTAAGCAAAGTACATCGTATGTACTTTTCTTCAATTTCAAT
CTAGCACTGGCAGAAATGCATGAAGAAAAGATTTCTTTTTTCGGATCTTAGTTTCCAAGGTTATTTCAAAAAGT
AAACGTGTATTTTTGTTTCTAATTTTCTTTTCTTTTCTTCAATTTCATTTTACTTTTTCTCATCCCTCCATAGATA
ATGCATTGTATGCAGAATAAGCTAGTTACATGTTTCACCCTCAACTGCCGA

>T***paxP***
ATAGGGATACGTTGCCGGTCGCTTTCGTTGTCTTTGAACGCAGTGGTGGGCAGTGGACGTTCACTGGAGAT
AGATTTTACATCTTAGAATGAATTGAATTATGCTTTCTGTTTATTAATCCTCAGGAATGTCTTCTATTCTCTGTA
ACAGTGAGTCGGATCAGAATACCTGAGTGCATTAACCTTGTTGGCATCCTTTTCTGGTTCTTCTGTAAAGGGA
CATTGCCTTGAGATTGAATGGCCGCCCTCGCAAAATTGTAATCCATGCTC

>T***paxQ***
TCTAGTTTCAAATTCGCTGGGTTGGAGTAATCATTTTAACAATTGATGAAAGCCCAATCGCAAGTCCTACAGC
TCATTATCGTGCAAGCAGCTGGCTCACCCCACCTCTGGGGTTGTAGTGGCGAACTATGTGAGACTAAGGAG
CCCCGATCCTGCTAGTGTTTAACCTGCCCCATTTTTCATGCCCACTTTCACATCTTGCCTTATTGTGGCGCAG
TCTAGAGTTGCAAACTTTGAGATACTTGGTAAGCGAGCAGGAGAGATACCATGCGCCTCAGATACTGTTTCA
TTTCATTGATAGTAGTAGAGTTAGCTAGACTGGCGAGTCTTCTTCTGGAACTACAGGATAGTCCAGATTCTGA
GTAGTAACAAACAATAAGCTCGTAGTAGTACATGTTGAGAAAGCCTTGAAATTTTGATACTATGGTTGCTTGA
ATAAGGCGATCCGATTCCACGGATGTCAAATTTATCGAGTGTAATCTATTTCCAAATTTCCCT

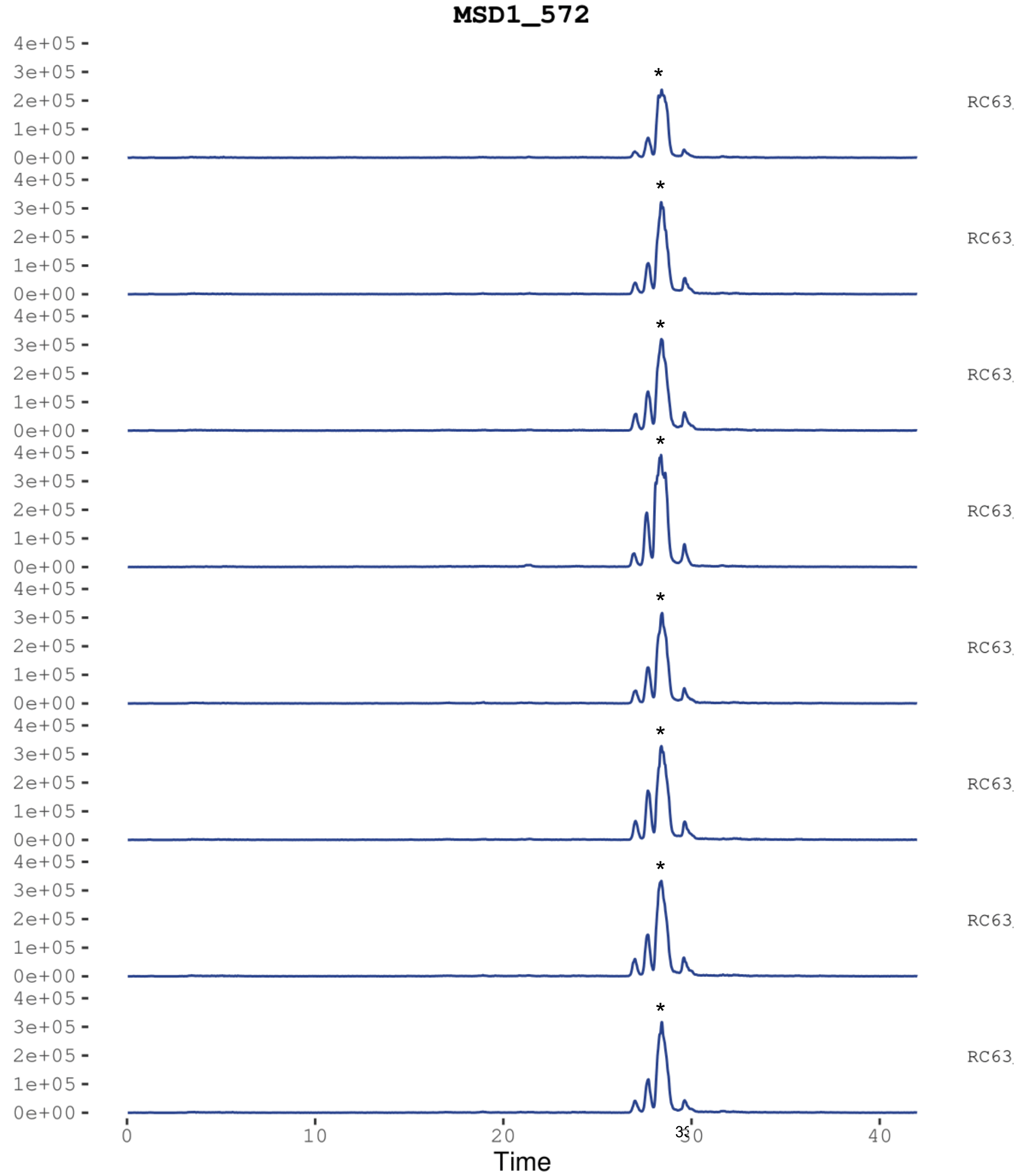
>T***trpC***
GATCCACTTAACGTTACTGAAATCATCAAACAGCTTGACGAATCTGGATATAAGATCGTTGGTGTCGATGTCA
GCTCCGGAGTTGAGACAAATGGTGTTCAGGATCTCGATAAGATACGTTCATTTGTCCAAGCAGCAAAGAGTG
CCTTCTAGTGATTTAATAGCTCCATGTCAACAAGAATAAAACGCGTTTCGGGTTTACCTCTTCCAGATACAGC
TCATCTGCAATGCATTAATGCATTGGACCTCGCAACCCTAGTACGCCCTTCAGGCTCCGGCGAAGCAGAAGA
ATAGCTTAGCAGAGTCTATTTTCATTTTCGGCAGACGAGATCAAGCAGATCAACGGTCGTCAACAGACCTAC
GAGACTGAGGAATCCGCTCTTGGCTCCACGCGACTATATATTTGTCTCTAATTGTACTTTGACATGCTCCTCT
TCTTTACTCTGATAGCTTGACTATGAAAATTCCGTCACCAGCCCCTGGGTTCGCAAAGATAATTGCACTGTTT
CTTCCTTGAACTCTCAAGCCTACAGGACACACATTCATCGTAGGTATAAACCTCGAAAATCATTCCTACTAAG
ATGGGTATACAATAGTAACCATGGTTGCCTAGTGAATGCTCCGTAACACCCAATACGCCGGCCGAAACTTTT
TTACAACTCTCCTATGAGTCGTTTACCCAGAATGCACAGGTACACTTGTTTAGAGGTAATCCTTCTTTCTAGA

>T***aceB***
GCTTTACACTGCGGTGATACAGAGATAGATGAGAAAAGTATCTTAAAAATAAAACTATTGAGCGGATGAAGAA
AAAGTAAACATTTGTATGGTTGACCAACGCGAAATGTAACGTTAGGTACGCAACTCTGCAGGAAAATATATG
GGTCAGCTGAAACCTACAGGTCCTAAACCTTGGCGAAACCAGGGTTTGGATTTCCAGTCGGCGTCCGTCCA
AATGCAAATGGTGTGATTTTACAGACACCTCTTTTTTTGCCGCTCCACTTTAGCTCGTGAATCGGAGGTAGAT
CGCCATCAGAGAGCGTCGAGACTCTGCGTGGACACGCAGTTGGATTCATTAAGTACTGCTAGCTTCCTATAG
GGCCTGTGTGCCTTGAATCTCCCCGTCCAACCGTGCAAG

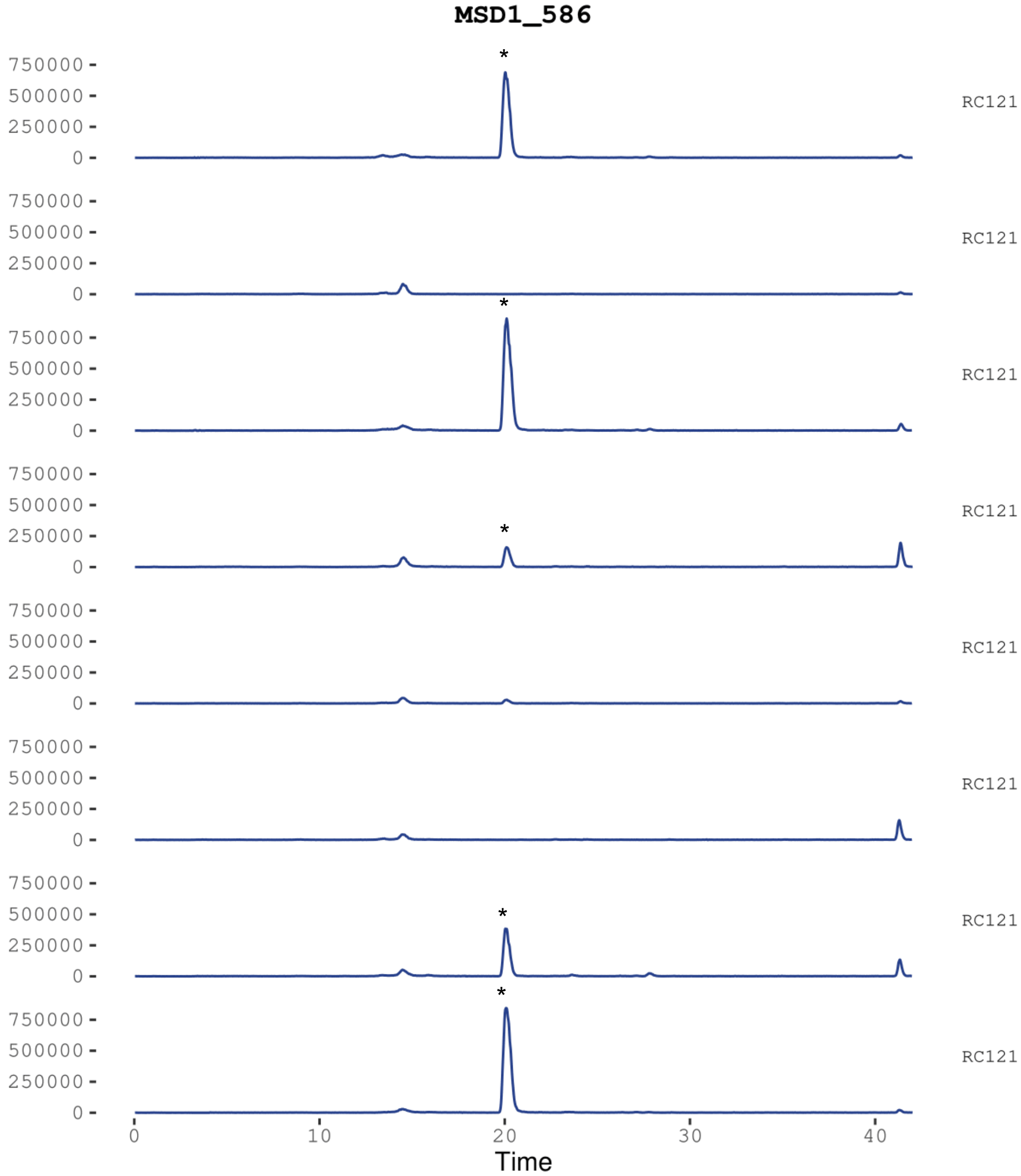
LC-MS Analysis of Fungal Transformants

The analysis of fungal transformants by LC-MS was conducted as described above (pg 4–5). For each of the transformants in this study extracted ion chromatograms (EICs) for key nodulisporic acid products have been reported. In each case the peak representing the metabolite of interest has been marked with an asterisk after confirmation by HR-ESI-MS^2 (Table S3, and pgs 43–65), or isolation and NMR analysis (pg 66–142). When no asterisk is shown the metabolite was either not produced, was below the limit of detection for the instrument, and/or was not confirmed by HR-ESI-MS^2.

Strain: RC63
Metabolite: NAE*
LC-MS Trace: EIC (*M* 572.4)

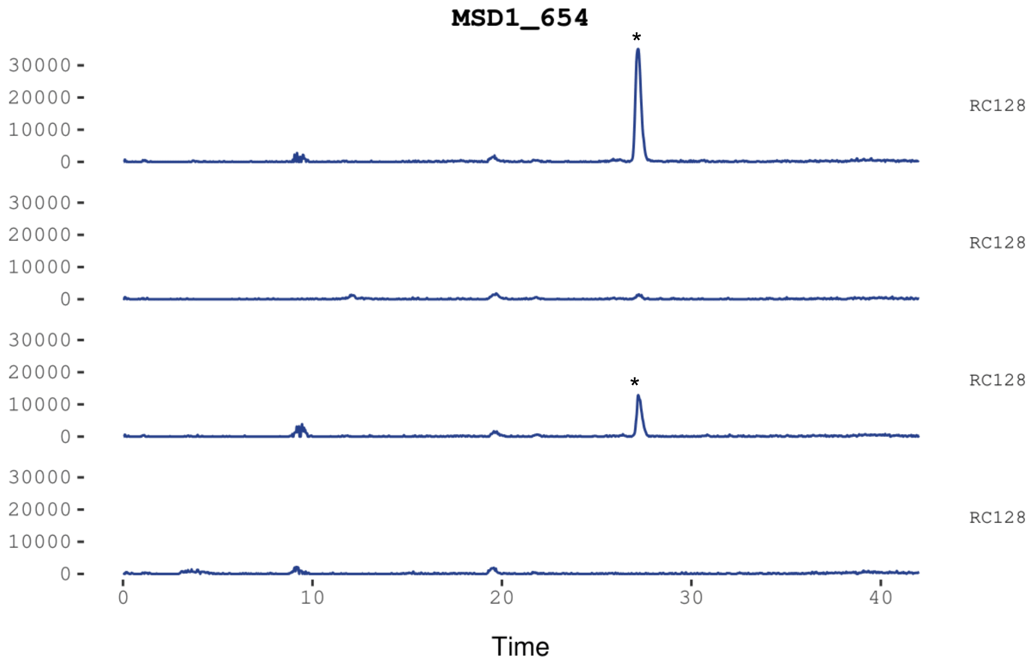


Strain: RC121
Metabolite: NAD_4*
LC-MS Trace: EIC (586.4 *M*)

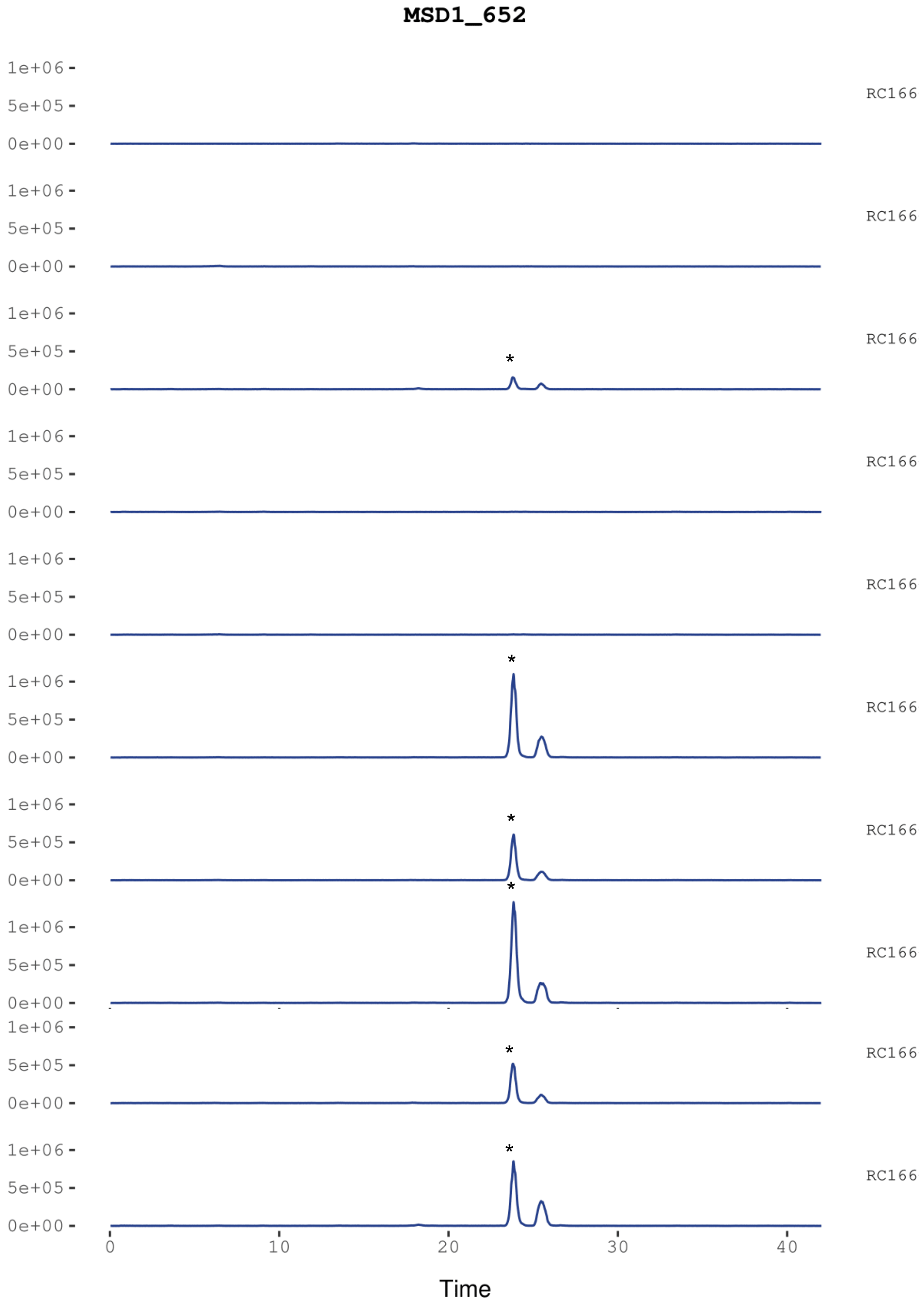


SUPPORTING INFORMATION

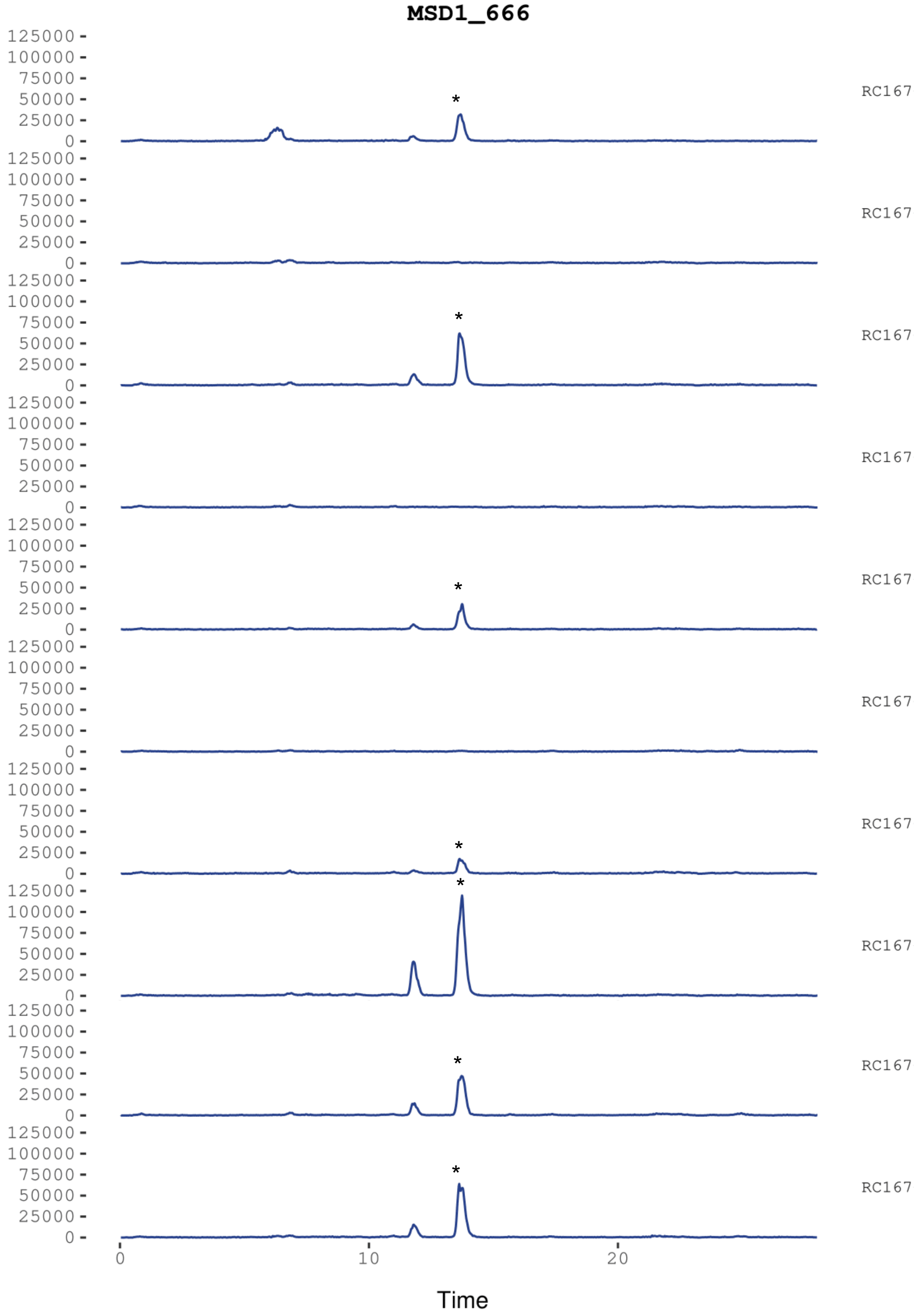
Strain: RC128
Metabolite: DH-NAC$_4$*
LC-MS Trace: EIC (654.4 *M*)



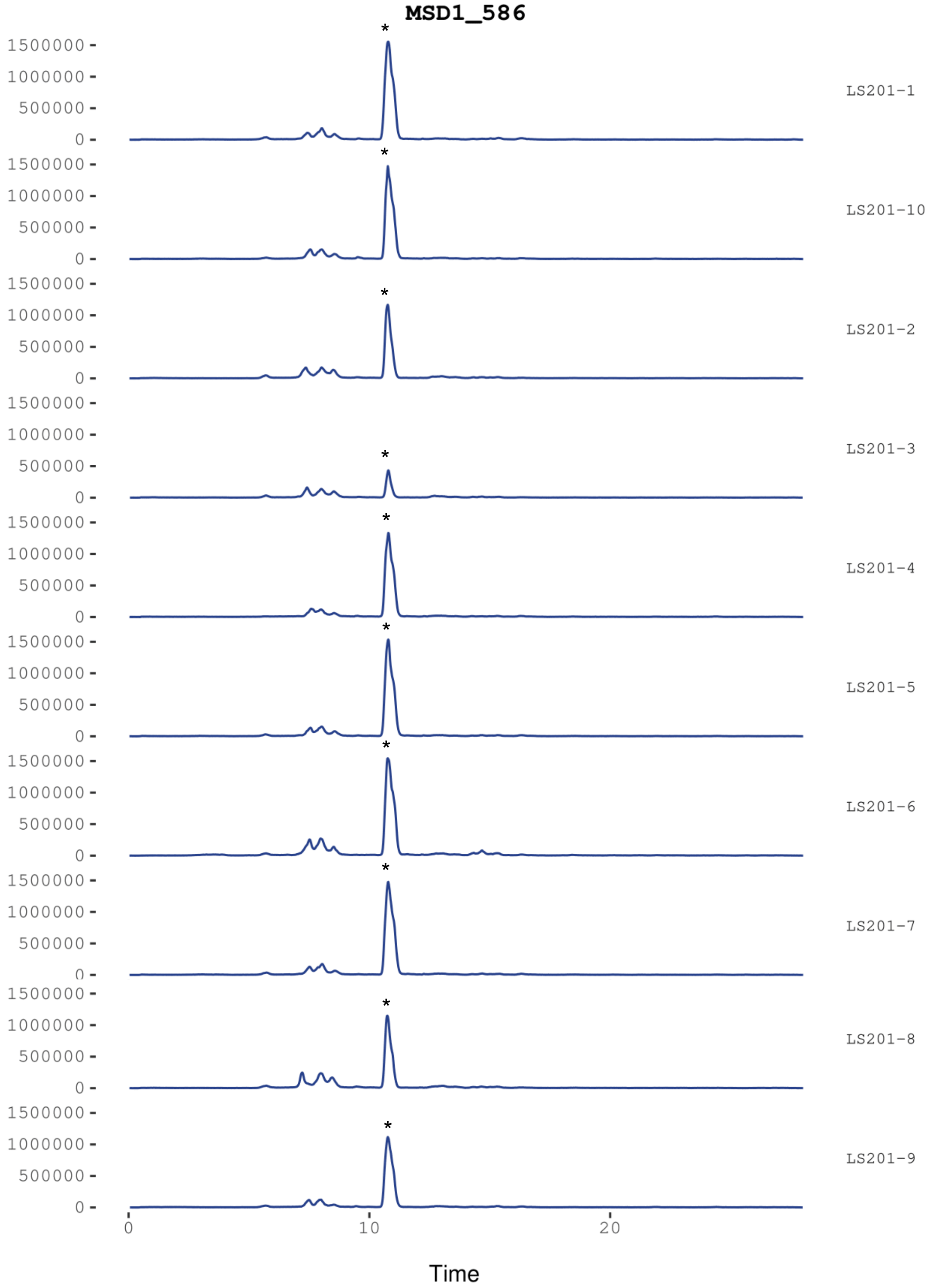
Strain: RC166
Metabolite: DH-NAB$_4$*
LC-MS Trace: EIC (652.4 *M*)



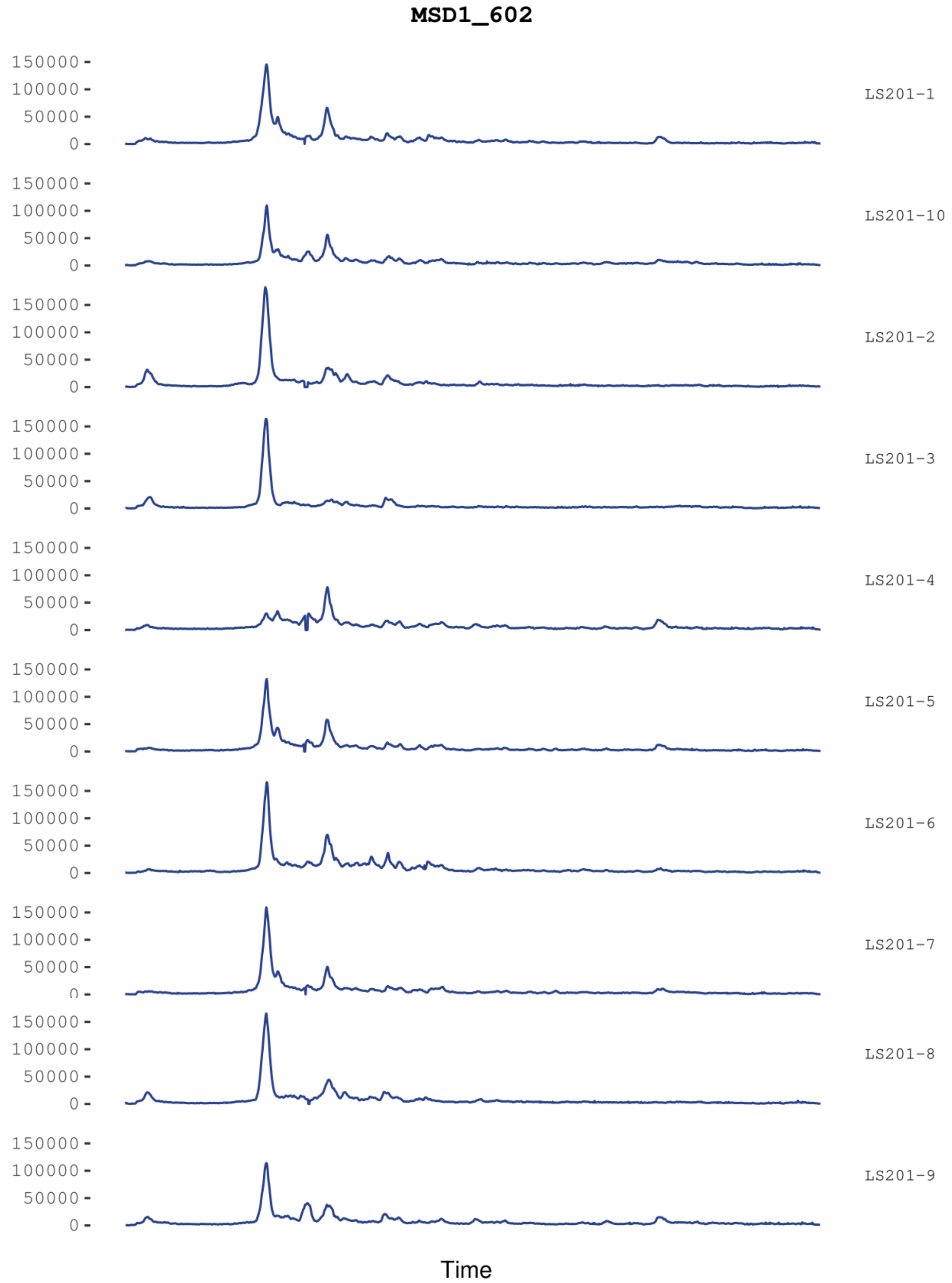
Strain: RC167
Metabolite: DH-NAA$_4$*
LC-MS Trace: EIC (666.4 *M*)



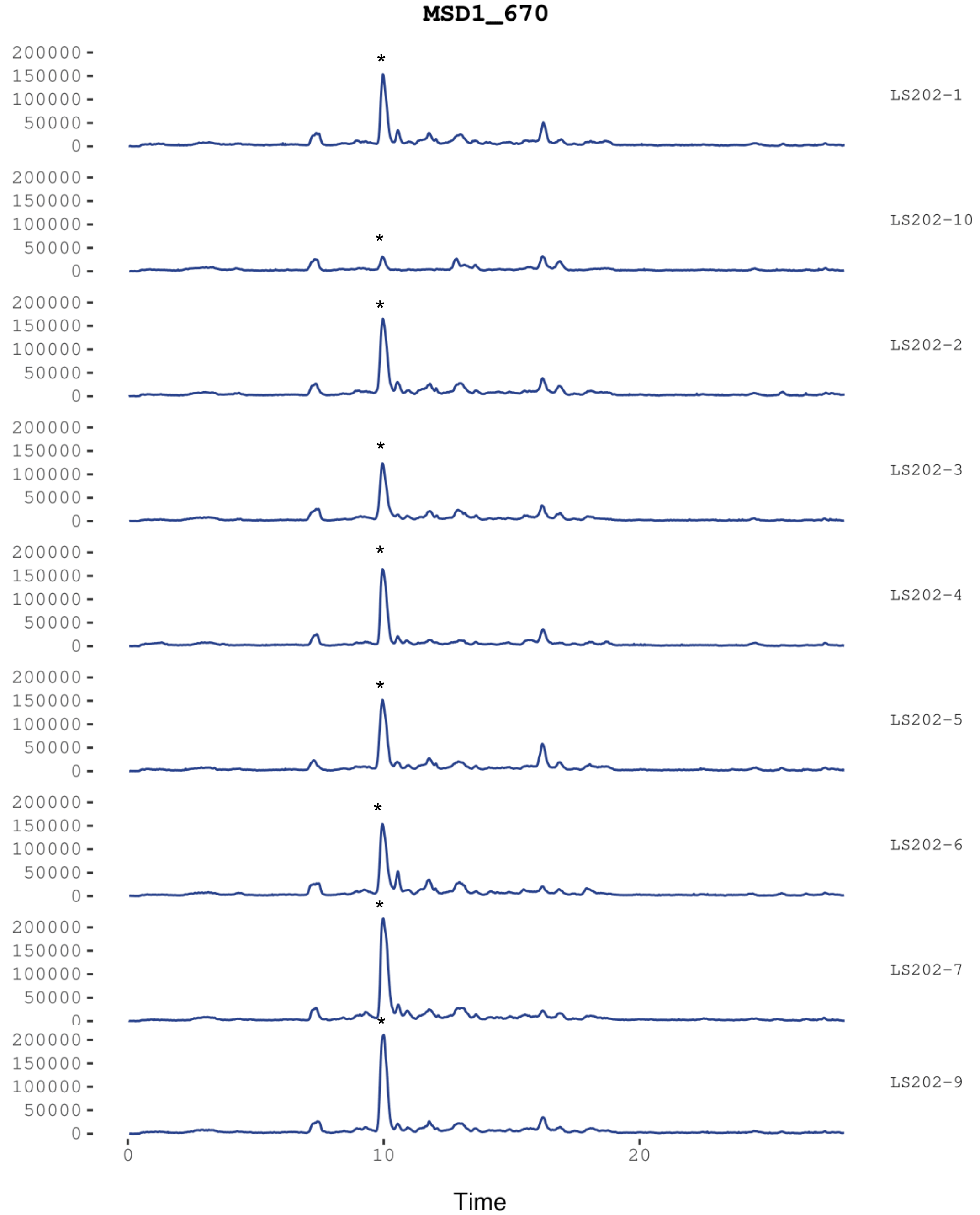
Strain: LS201
Metabolite: NAD_4*
LC-MS Trace: EIC (586.4 *M*)



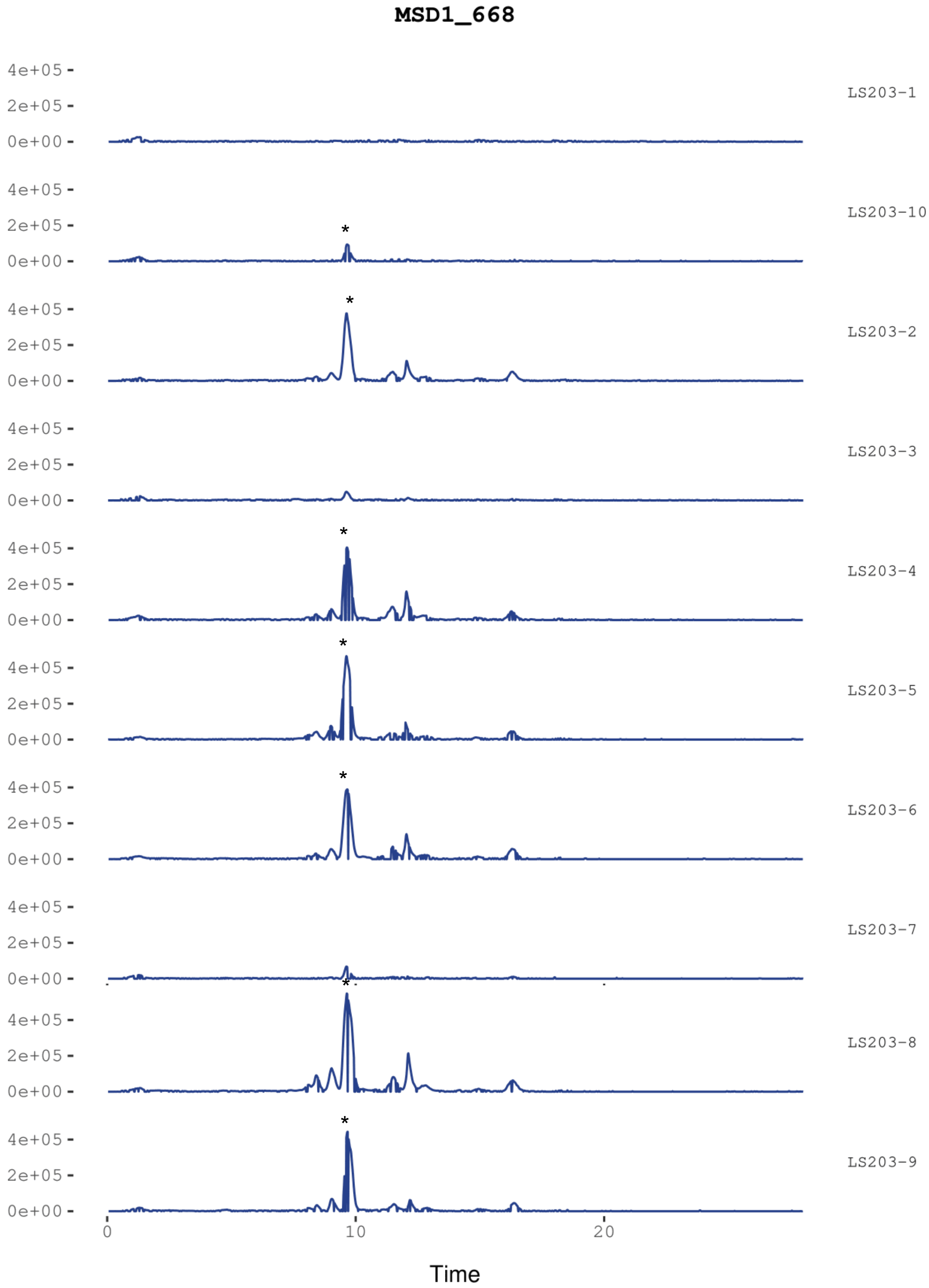
Strain: LS201
Metabolite: 24-hydroxy-NAD_4 (*metabolite not observed*)
LC-MS Trace: EIC (602.4 *M*)



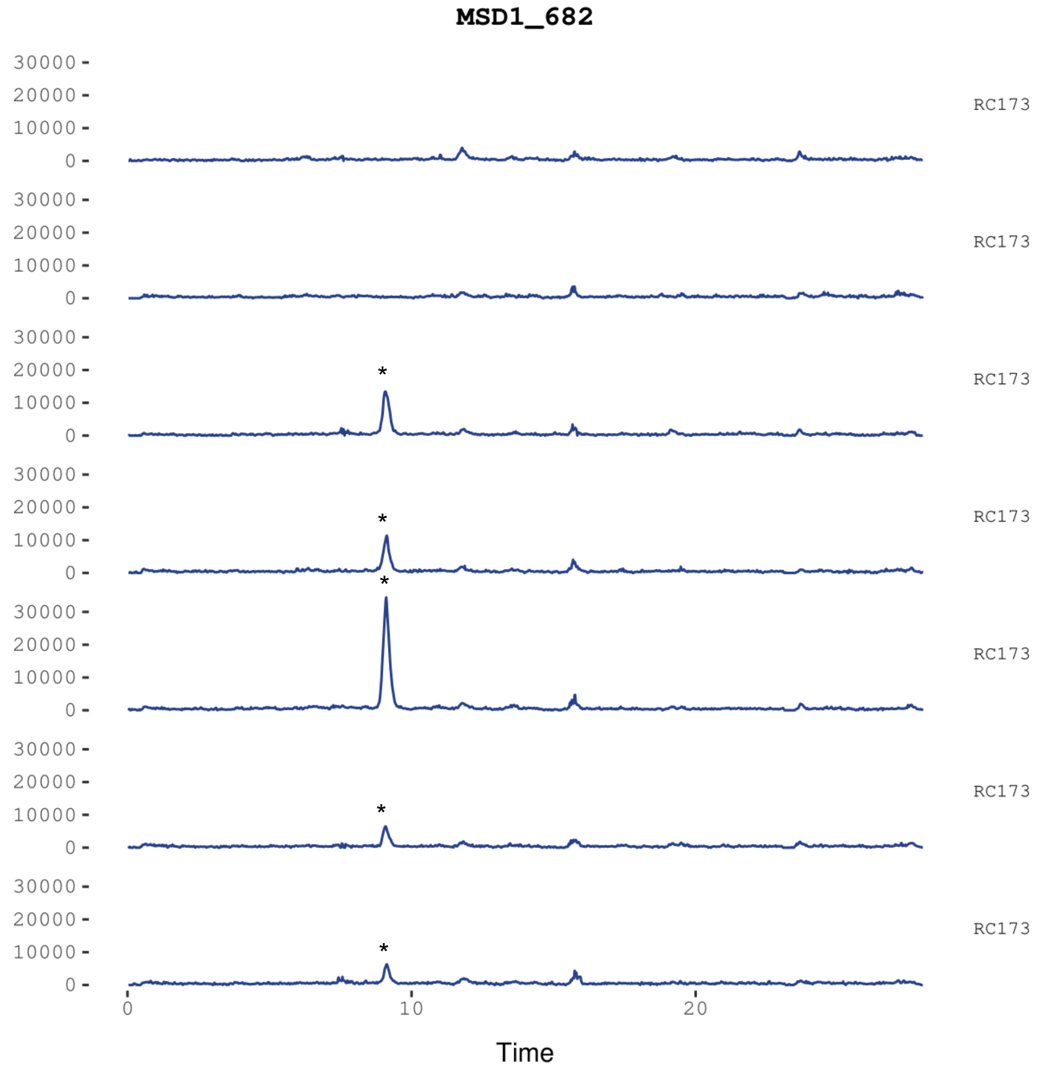
Strain: LS202
Metabolite: NAC_4*
LC-MS Trace: EIC (670.4 *M*)



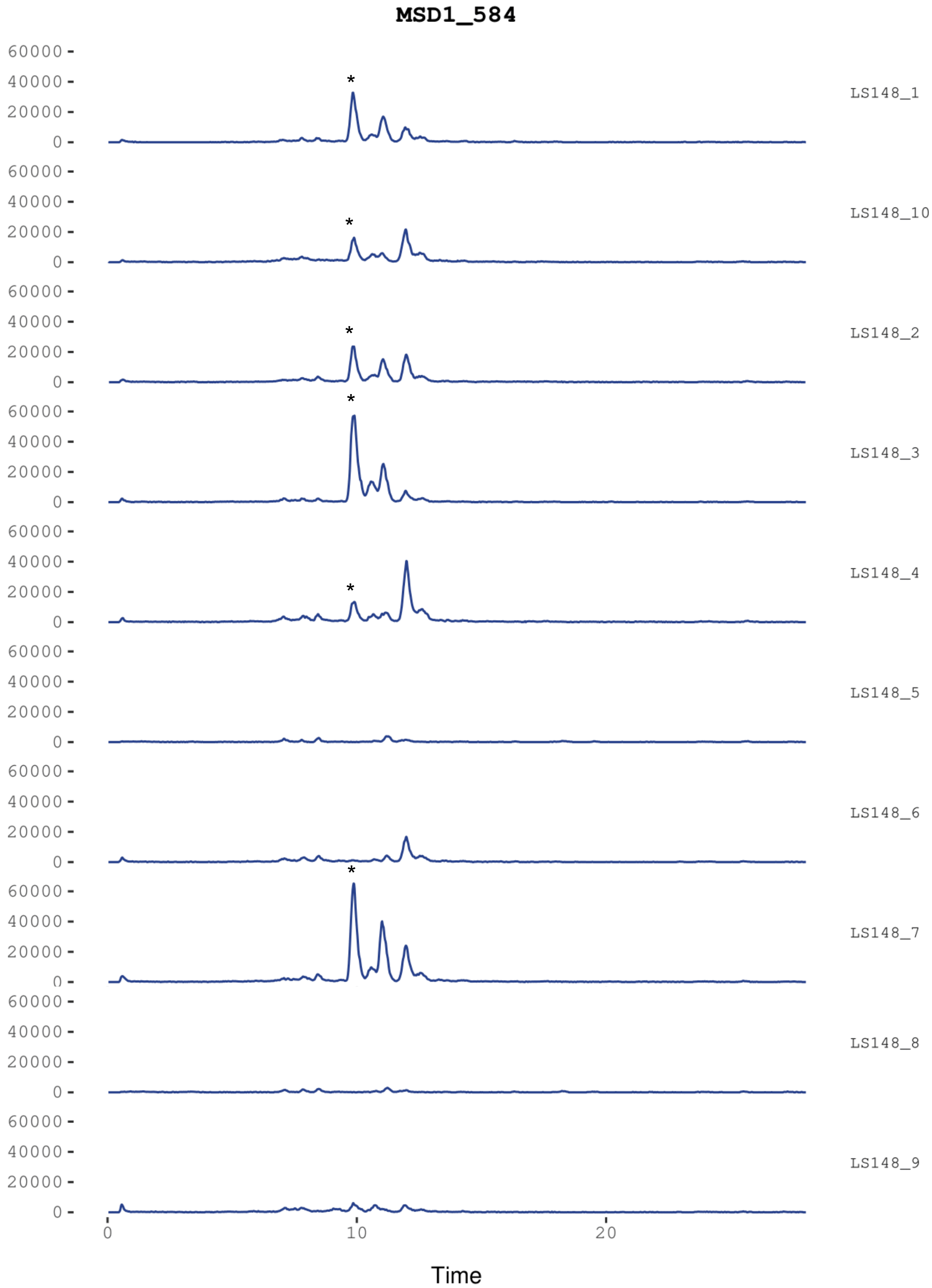
Strain: LS203
Metabolite: NAB_4*
LC-MS Trace: EIC (668.4 *M*)



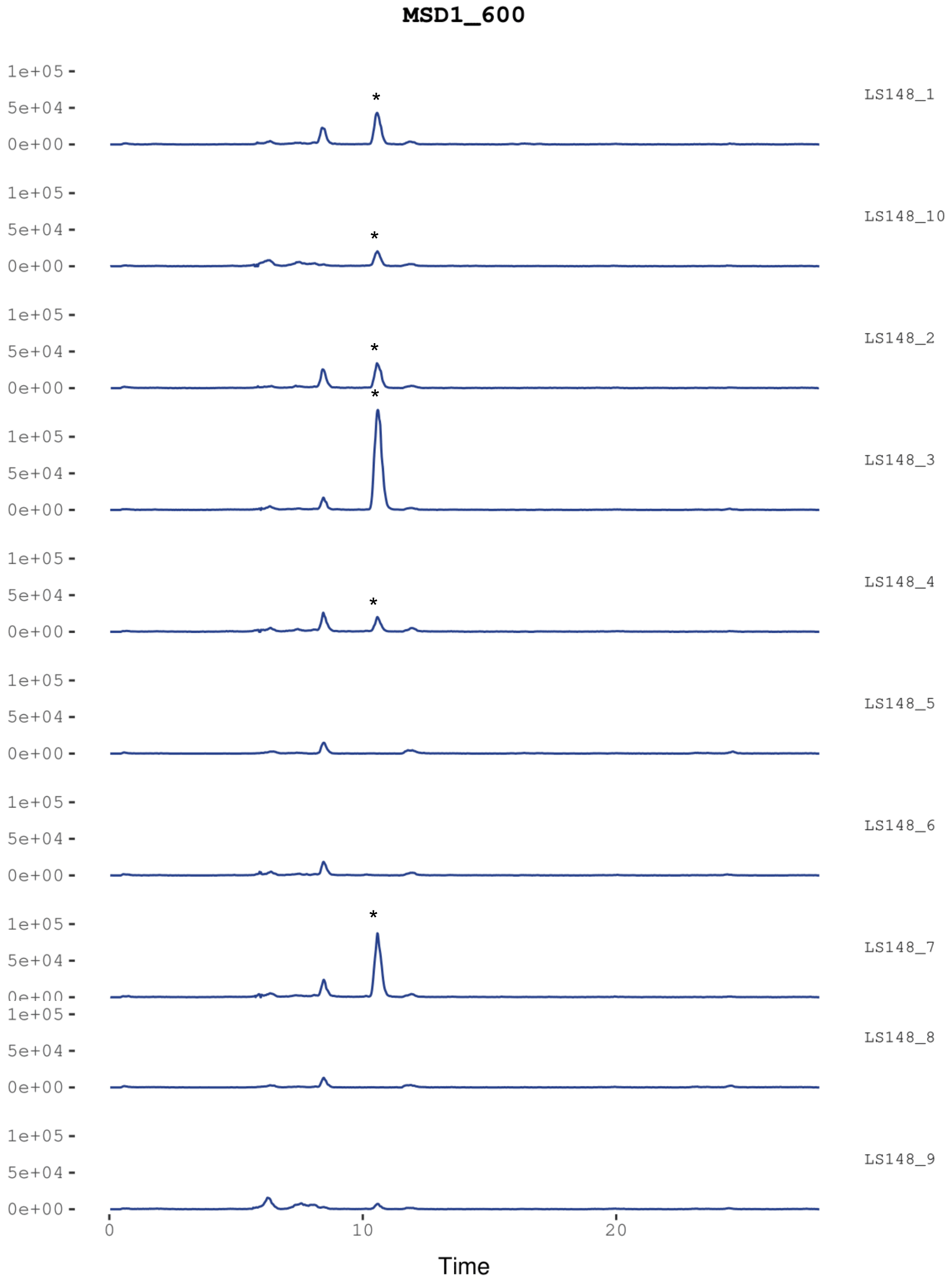
Strain: RC173
Metabolite: NAA_4*
LC-MS Trace: EIC (682.4 *M*)



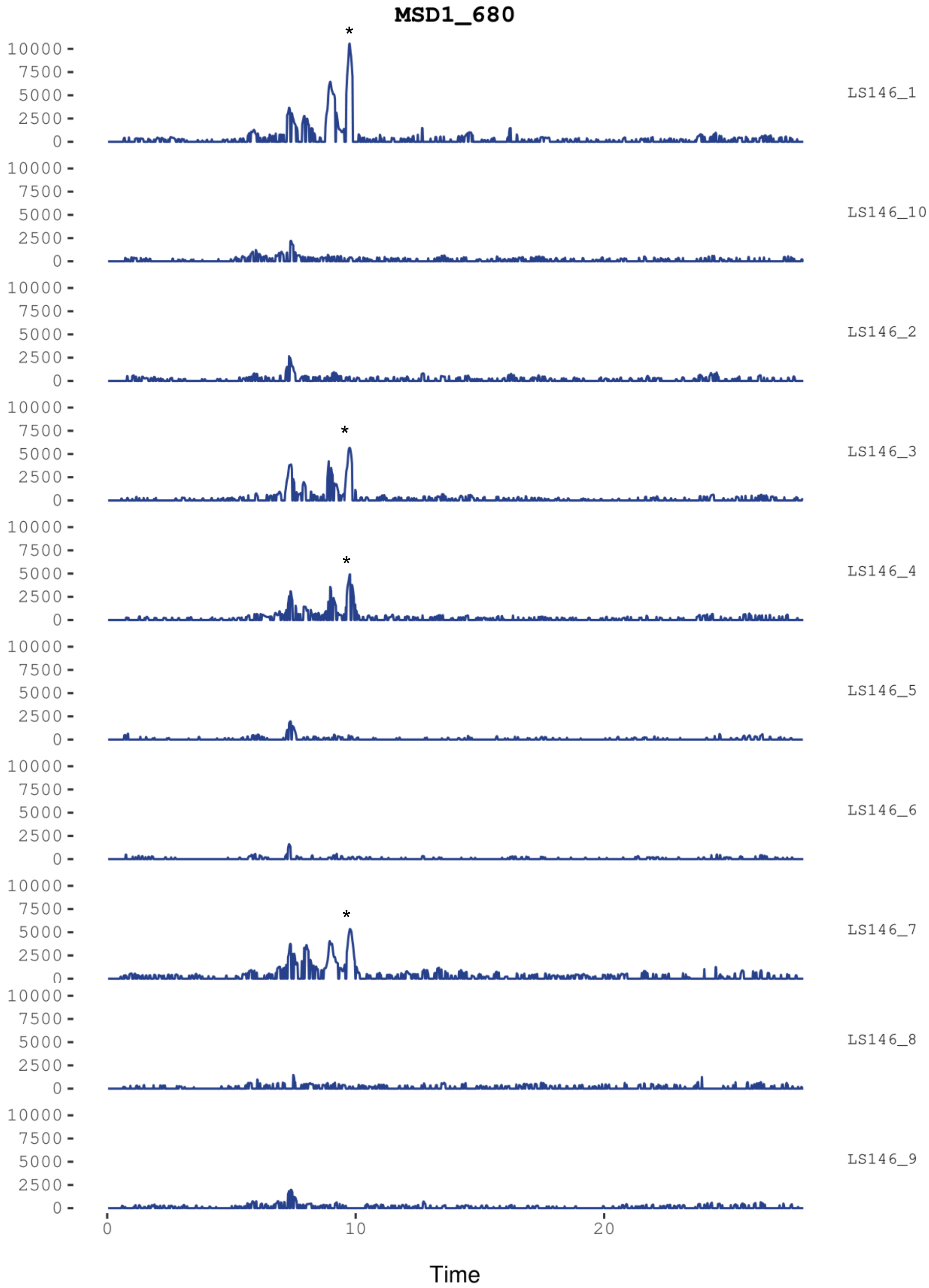
Strain: LS148
Metabolite: NAD*
LC-MS Trace: EIC (584.4 *M*)



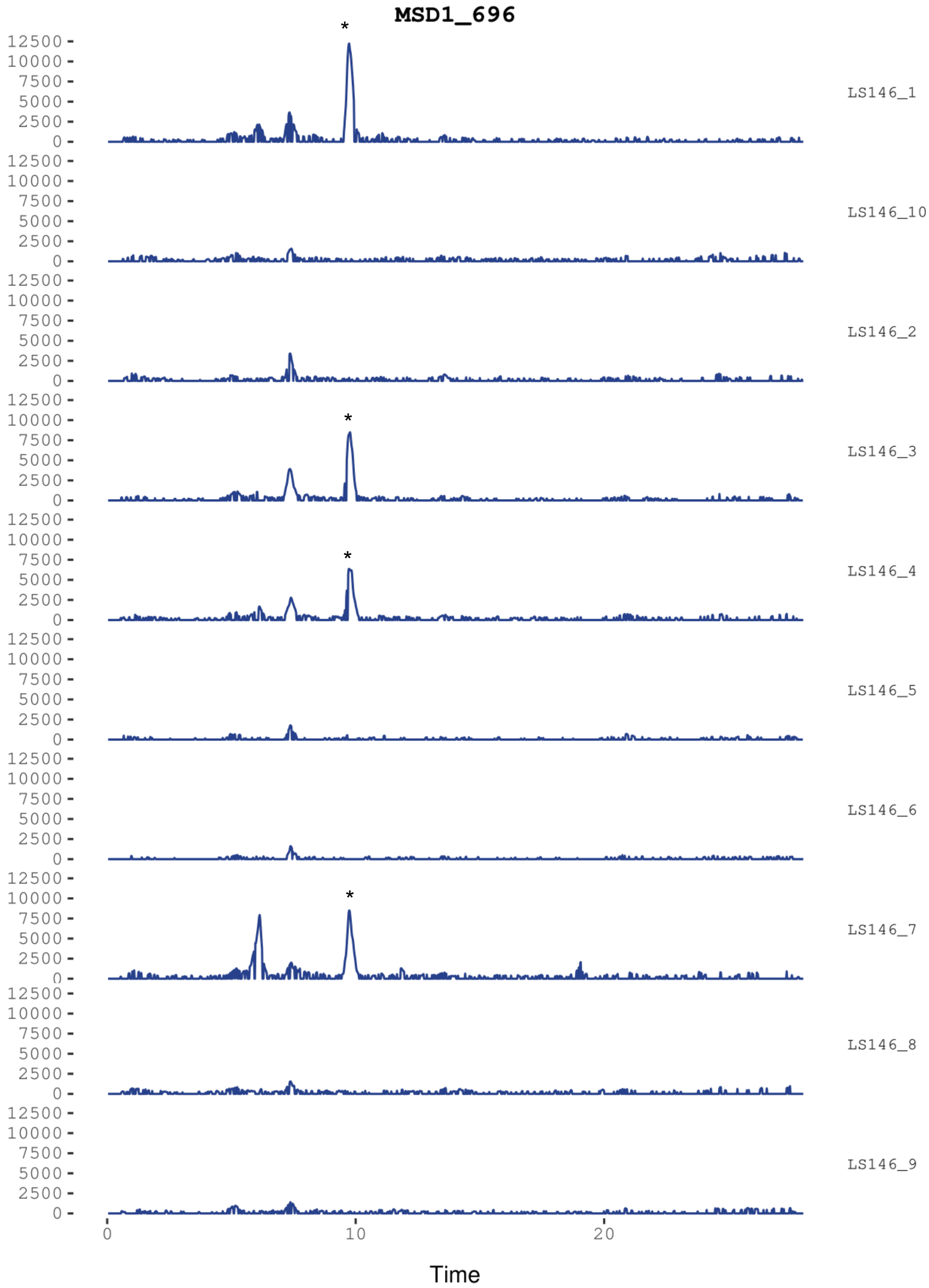
Strain: LS148
Metabolite: NAD_1*
LC-MS Trace: EIC (600.4 *M*)



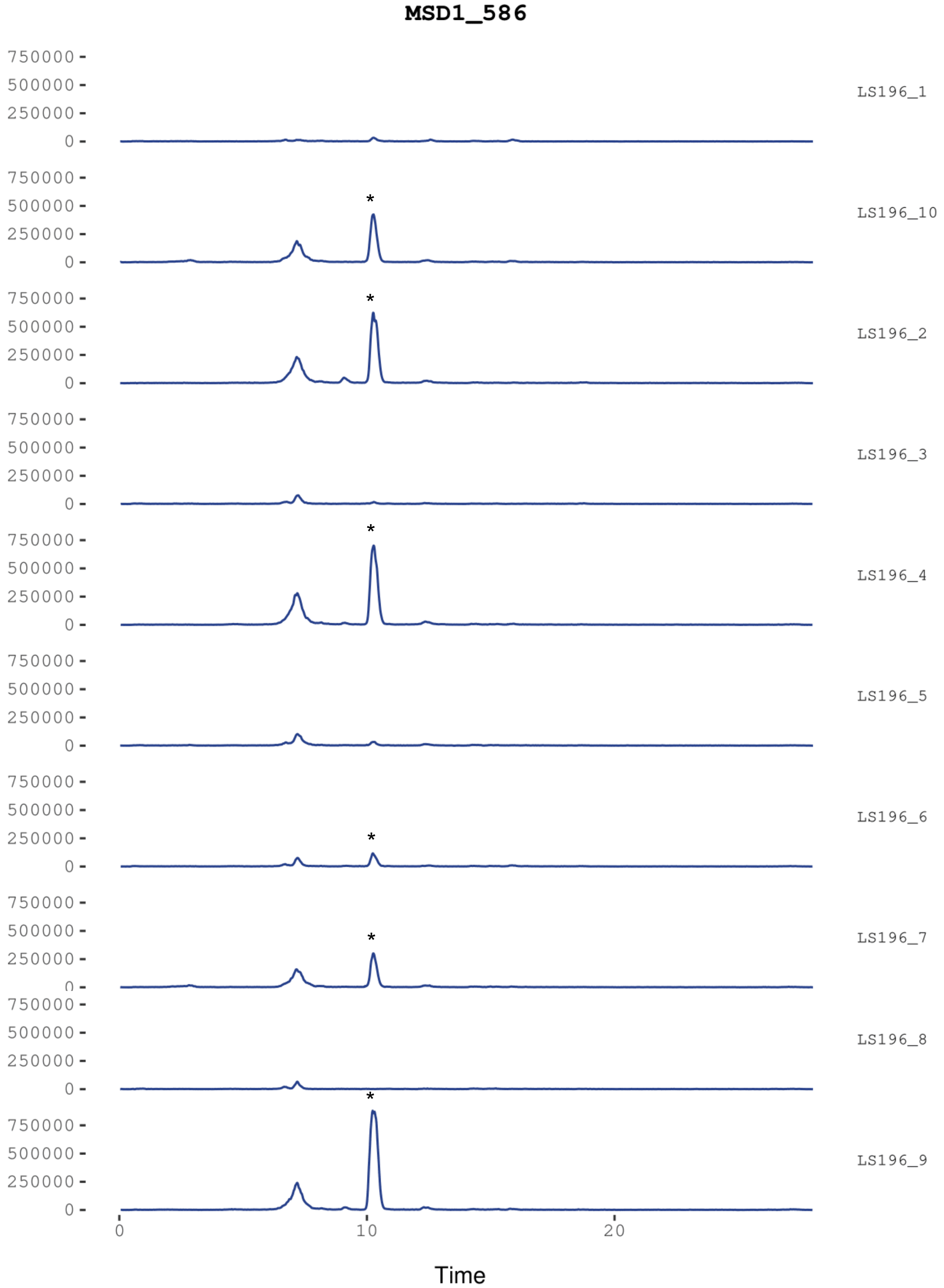
Strain: LS146
Metabolite: NAA*
LC-MS Trace: EIC (680.4 *M*)



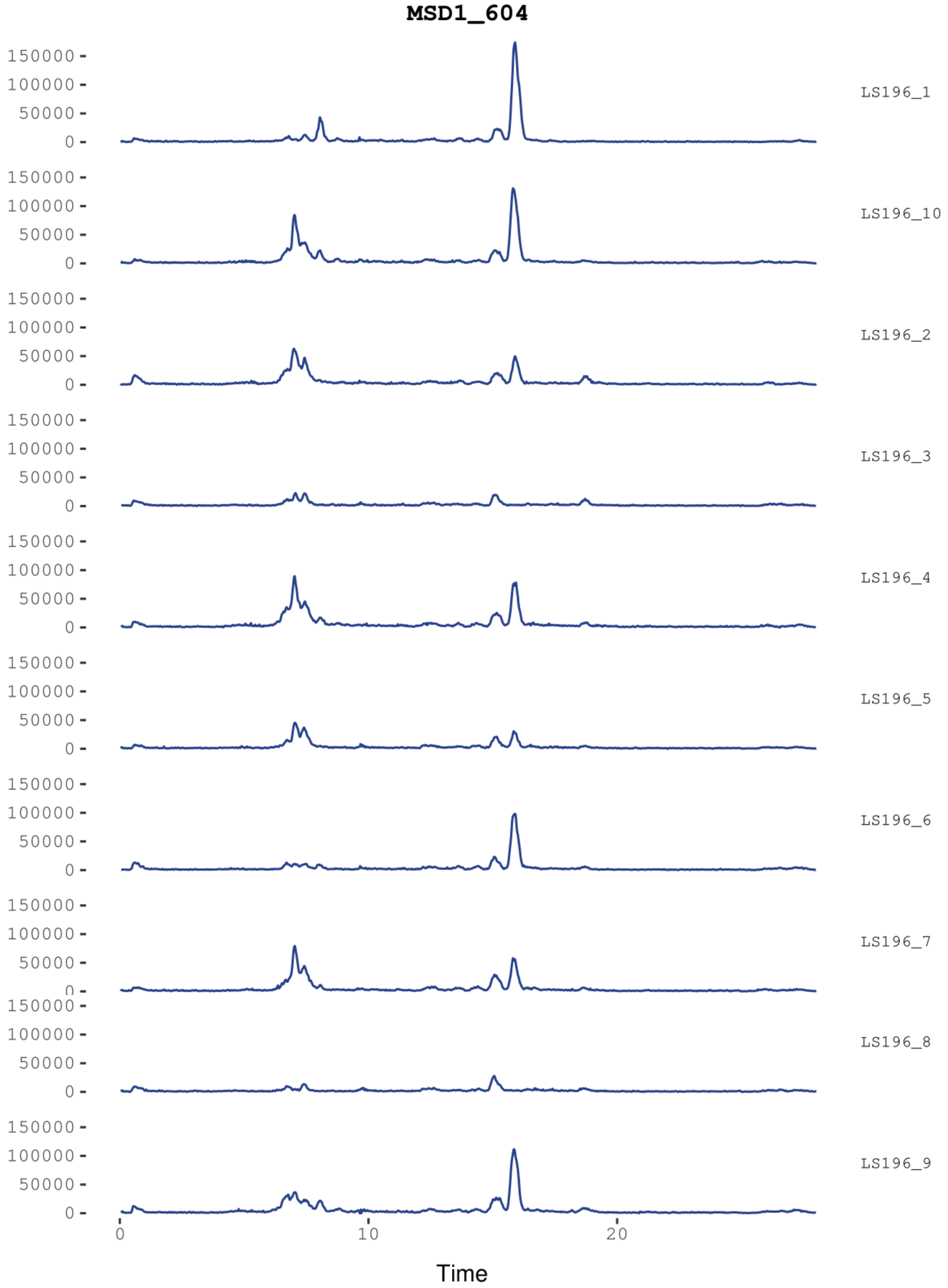
Strain: LS146
Metabolite: NAA_1*
LC-MS Trace: EIC (696.4 *M*)



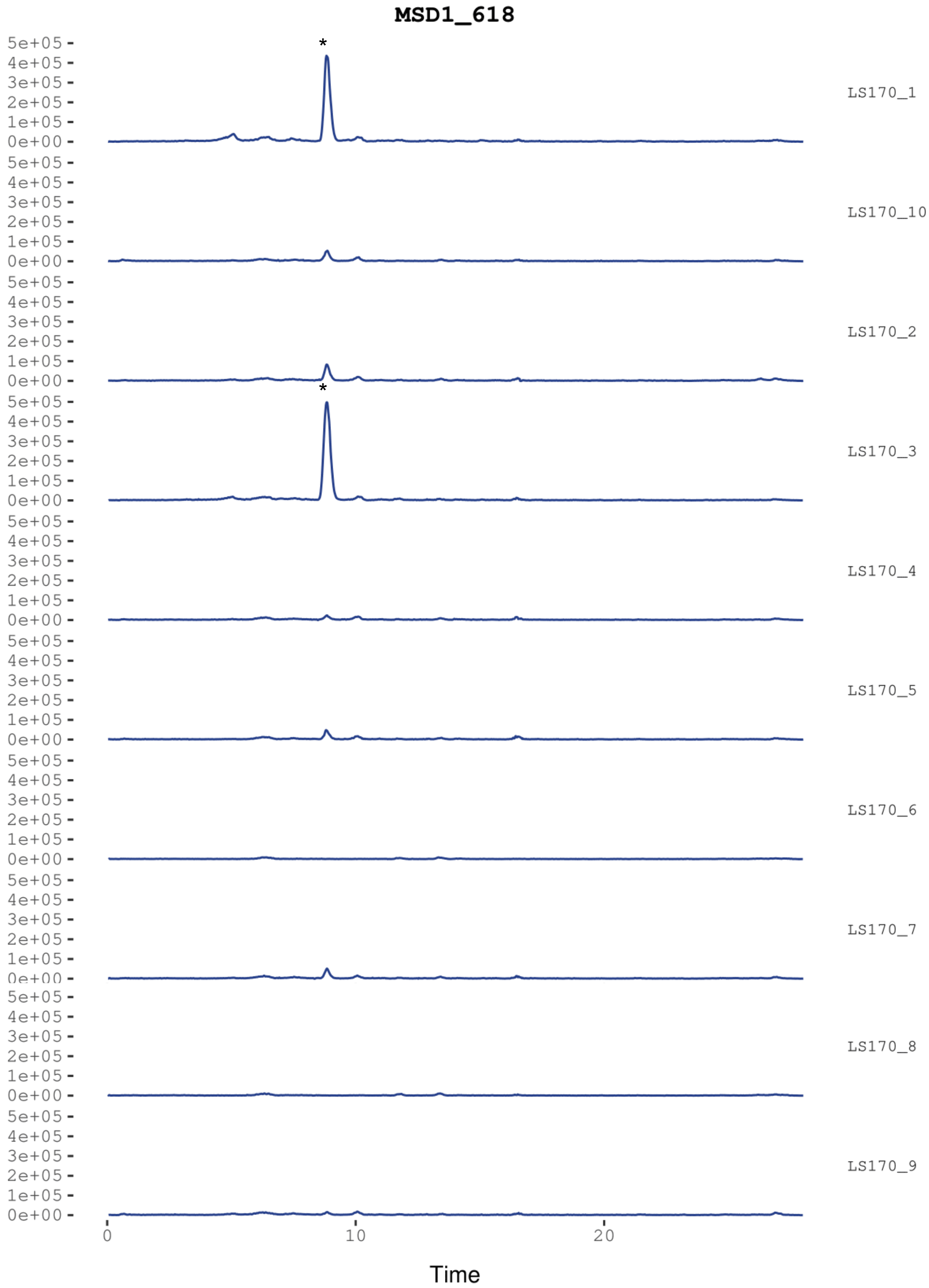
Strain: LS196
Metabolite: NAD_4*
LC-MS Trace: EIC (586.4 *M*)



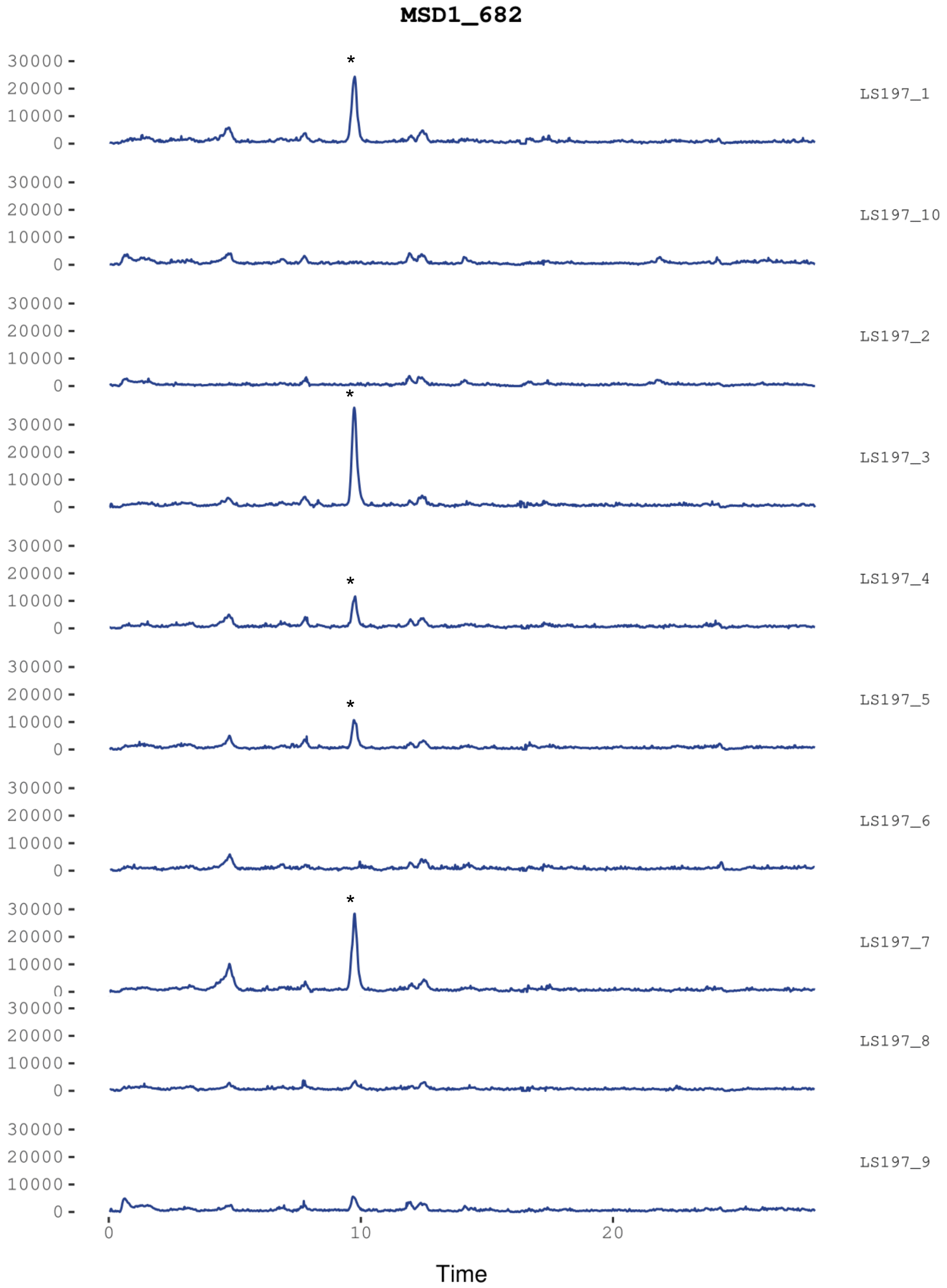
Strain: LS196
Metabolite: 3“-hydroxy-NAD_4 (*Metabolite not observed*)
LC-MS Trace: EIC (604.4 *M*)



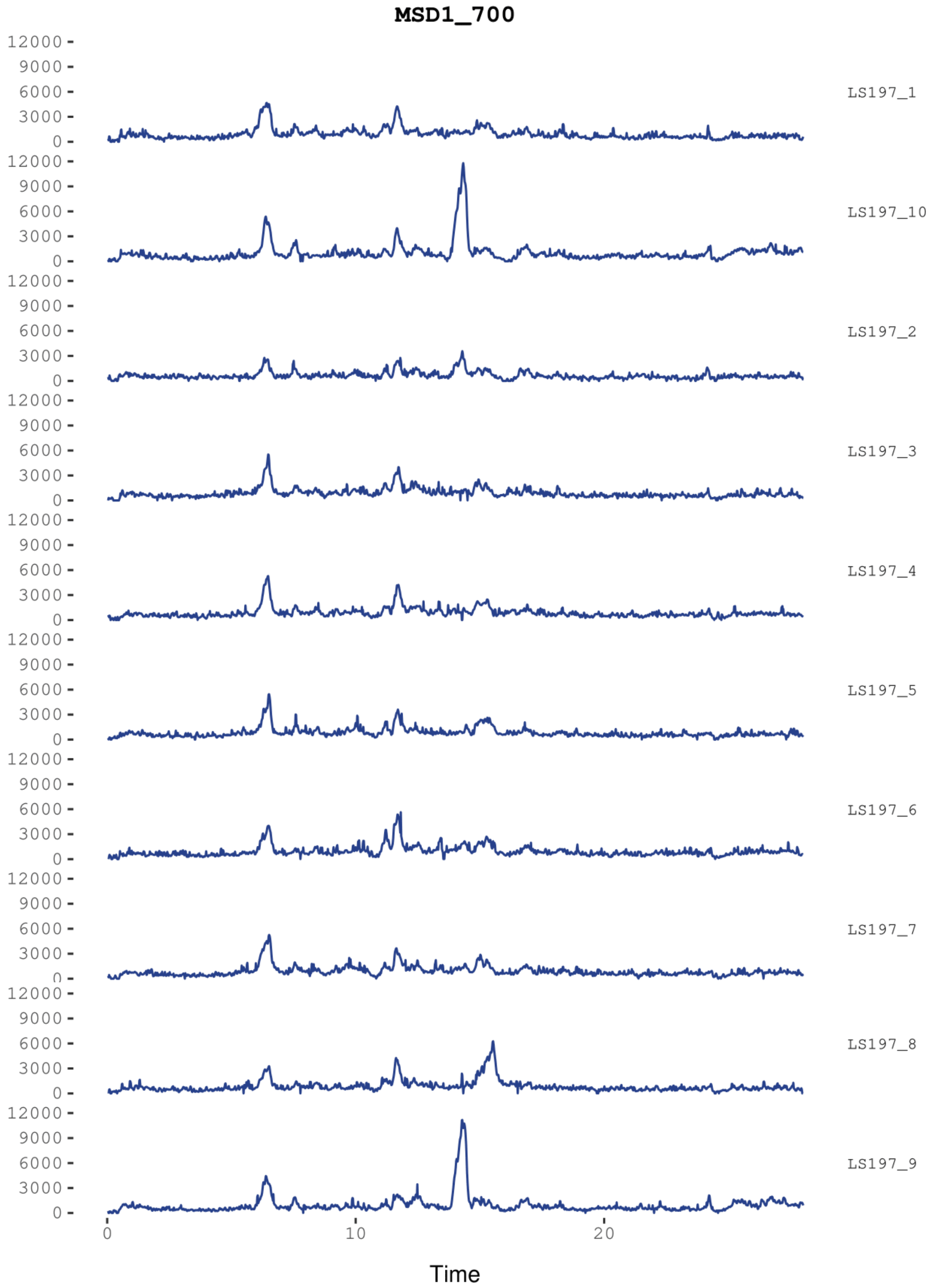
Strain: LS170
Metabolite: NAD_2*
LC-MS Trace: EIC (618.3 *M*)



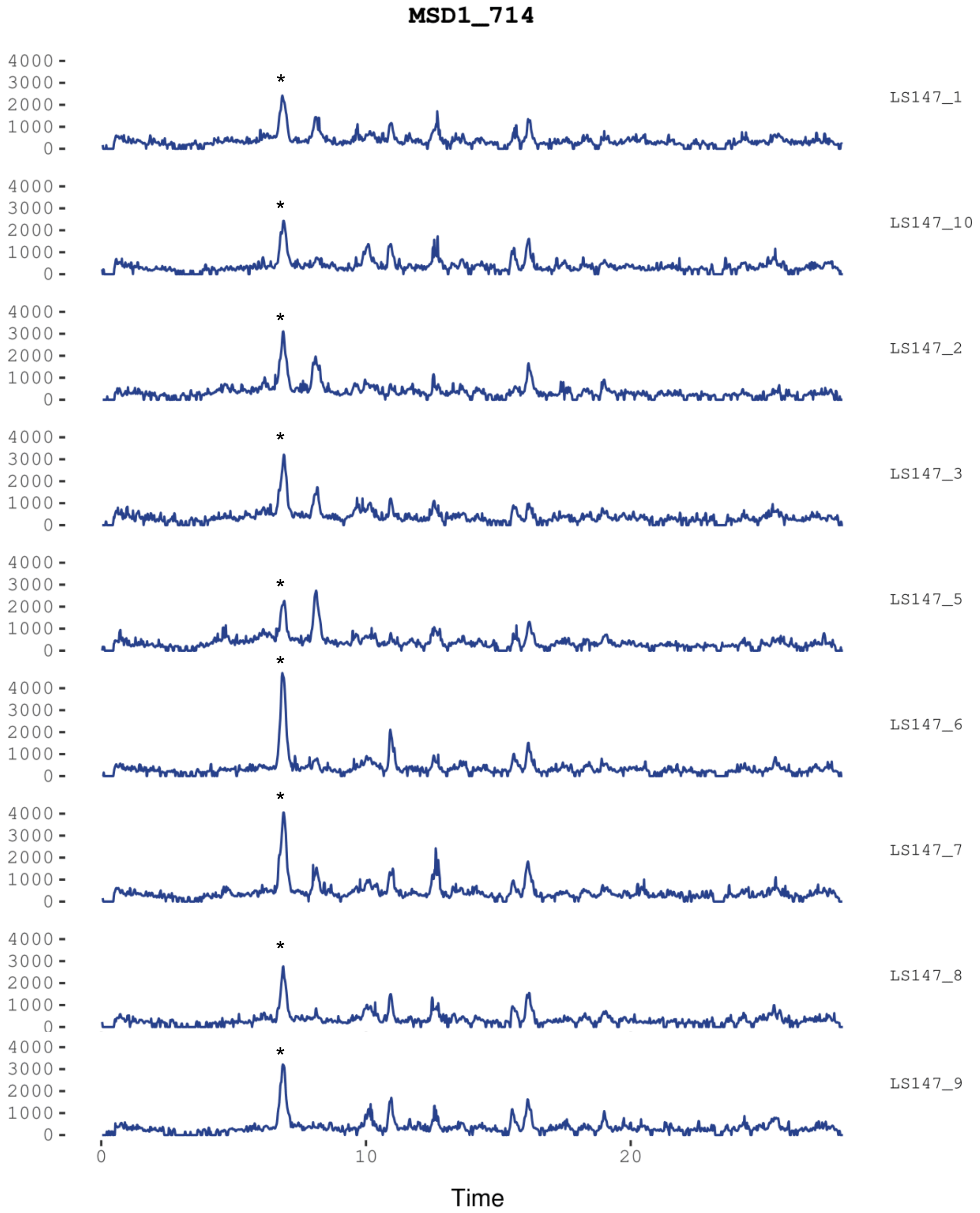
Strain: LS197
Metabolite: NAA_4 *
LC-MS Trace: EIC (682.4 *M*)



Strain: LS197
Metabolite: 3“-hydroxy-NAA$_4$ (*Metabolite not observed*)
LC-MS Trace: EIC (700.4 *M*)



Strain: LS147
Metabolite: NAA_2*
LC-MS Trace: EIC (714.4 *M*)



Identification of Nodulisporic Acids by Tandem Mass Spectrometry

Nodulisporic acids were identified primarily by HR-ESI-MS[2] using a method and information gained from the analysis of purified nodulisporic acids and their precursors. These included examples from each of the major series and at each biosynthetic step Emindole SB, NAF, NAE, NAD_4, NAD_4-OMe, NAD, NAD_1, NAD_2, DH-NAC_4, NAC, NAB_4, NAB, DH-NAA_4, and NAA_1. Comparison of HR-ESI-MS[2] spectra collected from these compounds allowed for fragmentation patterns to be rationalised and the information utilised to predict, and subsequently identify, MS[2] spectra of nodulisporic acids present in fungal transformants.

Both emindole SB and NAF have only one major fragment, 130 *M*, which is characteristic of the indole moiety having lost the terpenoid portion of the molecule. NAE also has only one major fragment, 266 *M*, which is consistent with a diprenyl indole moiety. No 130 *M* fragment was observed for NAE suggesting these prenyl groups are less readily lost than the original diterpene skeleton. For NAD_4 a more complex fragmentation pattern was observed. Loss of the terpene skeleton and retention of the western portion of the molecule explained the peak at 280 *M*, but two other fragments (524 *M* and 222 *M*) were also observed. The loss of a second portion of the molecule, with a mass of 58 *M*, would explain both these fragments. It was hypothesised that this second portion could be the ring oxygen and neighbouring geminal dimethyl carbon, $OC(CH_3)_2$, from the A ring. Analysis of the methyl ester, NAD_4-OMe, showed that only one of the fragments was affected by the esterification as 524 *M* was replaced by 542 *M*. This was consistent with the proposed fragmentation pattern of the molecule as shown in figure S1. NAD, NAD_1, and NAD_2 all showed MS[2] spectra consistent with the fragmentation pattern observed for NAD_4.

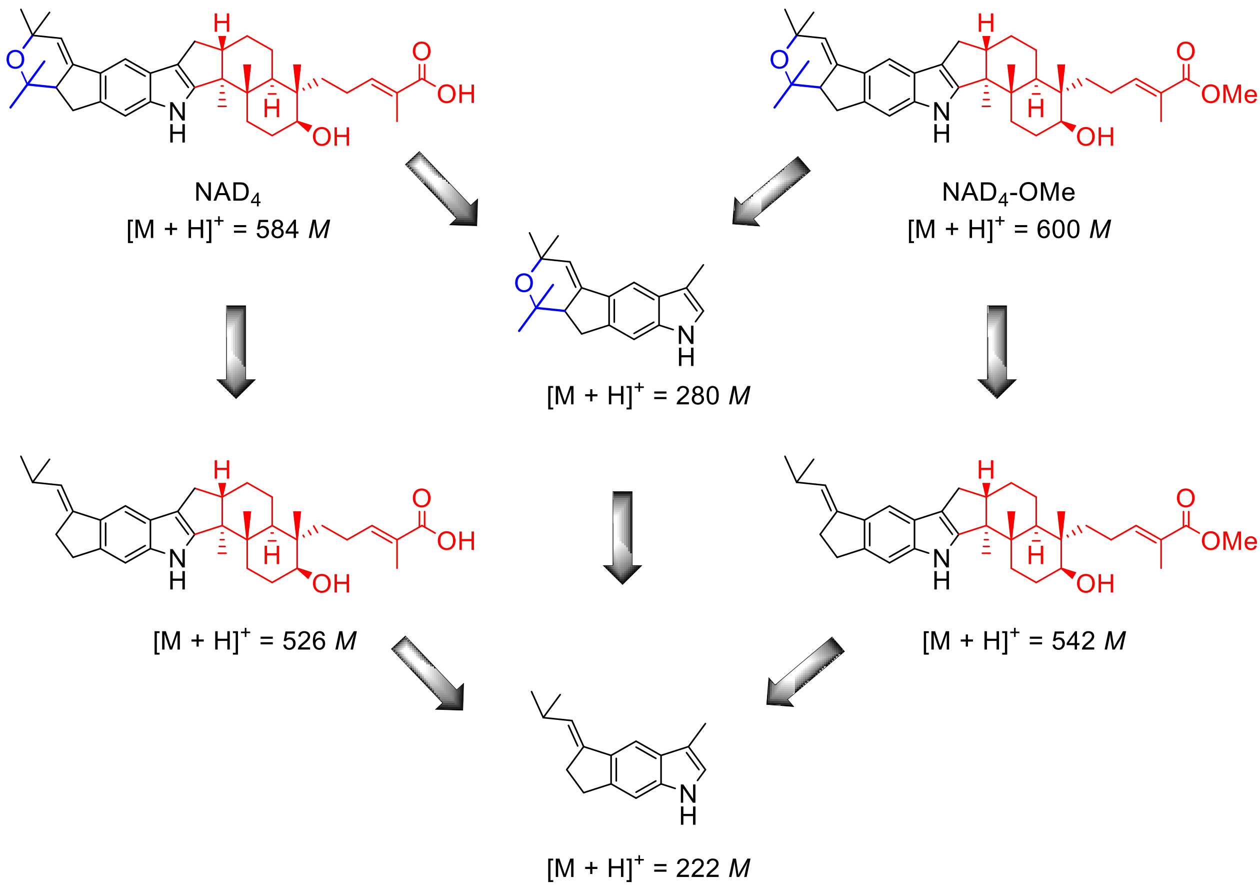


Figure S1. Proposed fragmentation of nodulisporic acids, NAD_4 and NAD_4-OMe

The major fragments in the HR-ESI-MS[2] spectra of DH-NAC_4 and NAC were consistent with successive loss of the eastern terpenoid skeleton and $OC(CH_3)_2$ of the A ring, indicating the same fragmentation pattern as for the D-series metabolites. An additional fragment for C-series compounds, suggesting a loss of *M* 68, was attributed to the C-26 prenyl group. Further confirmation was found in the fact that this fragment was not observed for NAB and NAB_4 in which the prenyl group is cyclised and therefore less readily lost during CID. The HR-ESI-MS[2] spectra of

DH-NAA$_4$ and NAA$_1$ were also consistent with the fragmentation patterns observed for all other nodulisporic acids analysed.

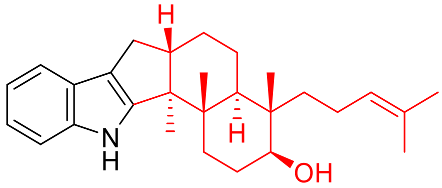
Given the consistent and predictable way in which the nodulisporic acids were observed to behave when analysed by HR-ESI-MS^2 it was possible to predict fragmentation patterns for biosynthetic intermediates. This allowed for rapid and accurate identification of nodulisporic acids in crude extracts where purification was not possible due to instability or low abundance. High resolution masses, key fragments, and annotated HR-ESI-MS^2 spectra for all nodulisporic acids identified are reported below.

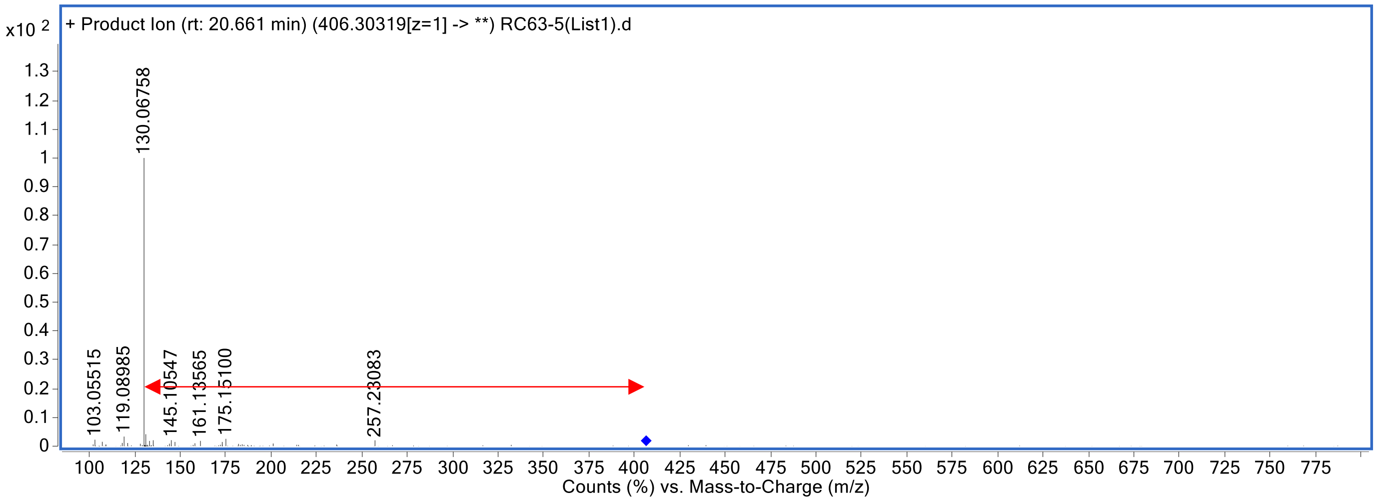
Table S3: HR-ESI-MS^2 analysis of nodulisporic acids

Name	Formula	Exact Mass	Calc. for $[M+H]^+$	HRMS Obs. (Δppm)	MS^2 Fragments (%relative intensity)
Emindole SB	$C_{28}H_{39}NO$	405.3032	406.3104	406.3093 (2.81)	130 (100)
NAF	$C_{28}H_{37}NO_3$	435.2773	436.2846	436.2856 (-2.25)	130 (100)
NAE	$C_{38}H_{53}NO_3$	571.4025	572.4098	572.4098 (0.91)	266 (100)
NAD	$C_{38}H_{49}NO_4$	583.3662	584.3734	584.3743 (-1.48)	526 (60), 280 (40), 222 (100)
NAD_1	$C_{38}H_{49}NO_5$	599.3611	600.3684	600.3685 (-0.25)	542 (60), 280 (60), 222 (100)
NAD_2	$C_{38}H_{51}NO_6$	617.3716	618.3789	618.3799 (-1.59)	560 (70), 280 (30), 222 (100)
NAD_4	$C_{38}H_{51}NO_4$	585.3818	586.3891	586.3887 (0.66)	528 (70), 280 (40), 222 (100)
NAD_4-OMe	$C_{39}H_{53}NO_4$	599.3975	600.4047	600.4064 (-2.77)	542 (100), 280 (65), 222 (40)
NAD_5	$C_{38}H_{49}NO_4$	583.3662	584.3734	584.3737 (-0.45)	526 (70), 280 (30), 222 (100)
NAD_6*	$C_{38}H_{51}NO_5$	601.3767	602.3840	602.3829 (-1.83)	544 (60), 280 (45), 222 (100)
DH-NAC	$C_{43}H_{57}NO_4$	651.4288	652.4360	652.4346 (2.2)	594 (70), 348 (90), 290 (100)
DH-NAC_1	$C_{43}H_{57}NO_5$	667.4237	668.4310	668.4311 (-0.22)	610 (60), 348 (100), 290 (95), 222 (15)
DH-NAC_2	$C_{43}H_{59}NO_6$	685.4342	686.4415	686.4415 (0.0)	628 (40), 348 (100), 290 (80), 222 (5)
DH-NAC_4	$C_{43}H_{59}NO_4$	653.4444	654.4517	654.4516 (0.13)	596 (85), 348 (90), 290 (100), 222 (10)
DH-NAC_5	$C_{43}H_{57}NO_4$	651.4288	652.4360	652.4373 (-1.94)	594 (90), 348 (70), 290 (100), 222 (10)
DH-NAC_6	$C_{43}H_{59}NO_5$	669.4393	670.4466	670.4452 (2.09)	612 (70), 348 (80), 290 (100), 222 (20)
NAC	$C_{43}H_{57}NO_5$	667.4237	668.4310	668.4328 (-2.77)	610 (20), 364 (100), 306 (50), 288 (30)
NAC_1	$C_{43}H_{57}NO_6$	683.4186	684.4259	668.4259 (2.29)	626 (20), 364 (100), 306 (40), 288 (25)
NAC_2	$C_{43}H_{59}NO_7$	701.4292	702.4364	702.4346 (2.6)	644 (30), 364 (100), 306 (40), 288 (30)
NAC_4	$C_{43}H_{59}NO_5$	669.4393	670.4466	670.4466 (0.3)	612 (30), 364 (100), 306 (50), 288 (30)
NAC_5	$C_{43}H_{57}NO_5$	667.4237	668.4310	668.4324 (-2.17)	610 (40), 364 (70), 306 (100), 288 (30)
NAC_6	$C_{43}H_{59}NO_6$	685.4342	686.4416	686.4424 (-1.29)	628 (15), 364 (100), 306 (70), 288 (50)
DH-NAB	$C_{43}H_{55}NO_4$	649.4131	650.4204	650.4208 (-0.64)	592 (20), 346 (100), 288 (40)
DH-NAB_1	$C_{43}H_{55}NO_5$	665.4080	666.4153	666.4137 (2.4)	608 (20), 346 (100), 288 (35)
DH-NAB_2	$C_{43}H_{57}NO_6$	683.4186	684.4259	684.4259 (2.29)	626 (15), 346 (100), 288 (30)
DH-NAB_4	$C_{43}H_{57}NO_4$	651.4288	652.4360	652.4368 (-1.17)	594 (30), 346 (100), 288 (40)
DH-NAB_5	$C_{43}H_{55}NO_4$	649.4131	650.4204	650.4192 (1.82)	592 (30), 346 (100), 288 (50)
NAB	$C_{43}H_{55}NO_5$	665.4080	666.4153	666.4163 (-1.5)	608 (40), 362 (85), 304 (100)
NAB_1	$C_{43}H_{55}NO_6$	681.4029	682.4103	682.4102 (0.02)	624 (30), 362 (100), 304 (90)
NAB_2	$C_{43}H_{57}NO_7$	699.4135	700.4208	700.4208 (-0.17)	642 (50), 362 (100), 304 (75)
NAB_4	$C_{43}H_{57}NO_5$	667.4237	668.4310	668.4300 (1.42)	610 (40), 362 (95), 304 (100)
NAB_5	$C_{43}H_{55}NO_5$	665.4080	666.4153	666.4163 (-1.5)	608 (10), 362 (100), 304 (20)
DH-NAA	$C_{43}H_{53}NO_5$	663.3924	664.3997	664.3977 (2.94)	606 (100), 360 (70), 302 (40)
DH-NAA_1	$C_{43}H_{53}NO_6$	679.3873	680.3946	680.3940 (0.83)	622 (90), 360 (100), 302 (70)
DH-NAA_4	$C_{43}H_{55}NO_5$	665.4080	666.4153	666.4155 (-0.3)	608 (100), 360 (100), 302 (50)
NAA	$C_{43}H_{53}NO_6$	679.3873	680.3946	680.3946 (-0.35)	622 (70), 604 (65), 564 (60), 376 (100), 318 (60), 290 (60)
NAA_1	$C_{43}H_{53}NO_7$	695.3822	696.3895	696.3882 (1.84)	638 (50), 620 (100), 376 (80), 318 (45), 290 (30)
NAA_2	$C_{43}H_{55}NO_8$	713.3928	714.4001	714.3998 (0.34)	656 (60), 638 (100), 376 (40), 318 (45), 290 (30)
NAA_4	$C_{43}H_{55}NO_6$	681.4029	682.4102	682.4117 (-2.18)	624 (70), 606 (100), 376 (90), 318 (35), 290 (40)

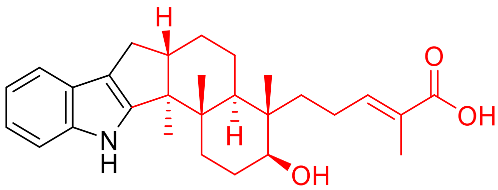
**As noted in Figure 3 of the manuscript an EIC-MS^2 trace for NAD_6 (602 M → 280 M, 222 M) will also show a peak for NAD_1 given ^{13}C isotopes of NAD_1 (i.e. 600.36 + 2) will closely match the parent mass of NAD_6 and also give the diagnostic 280, 222 M fragments which are common to both NAD_1 and NAD_6.*

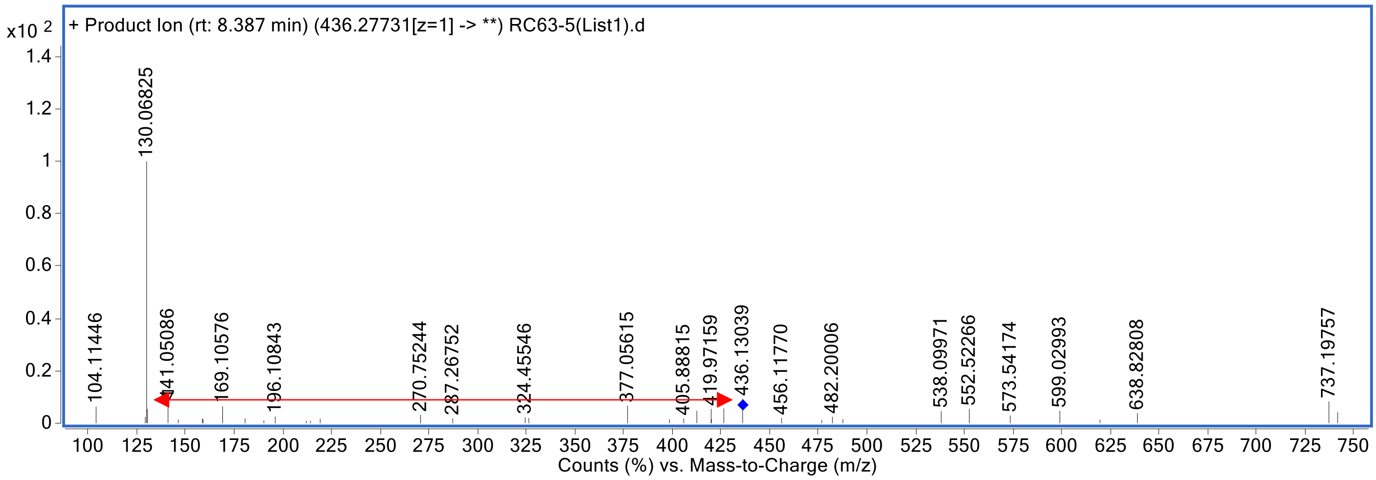
Emindole SB



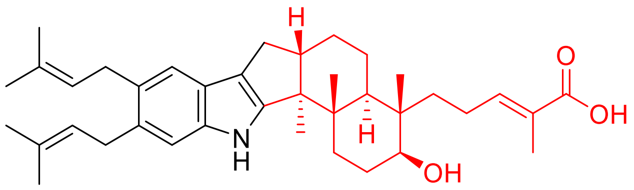


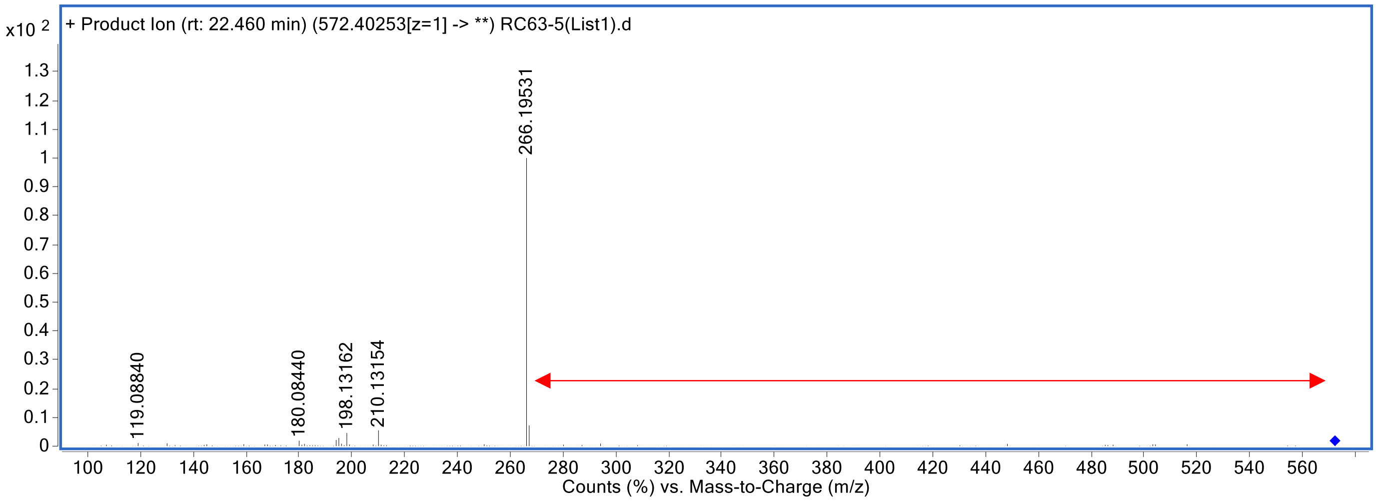
NAF



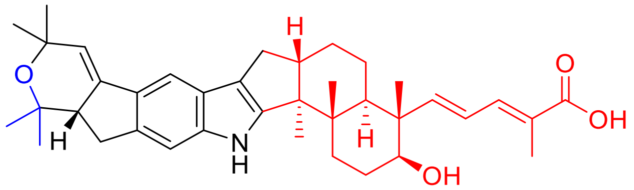


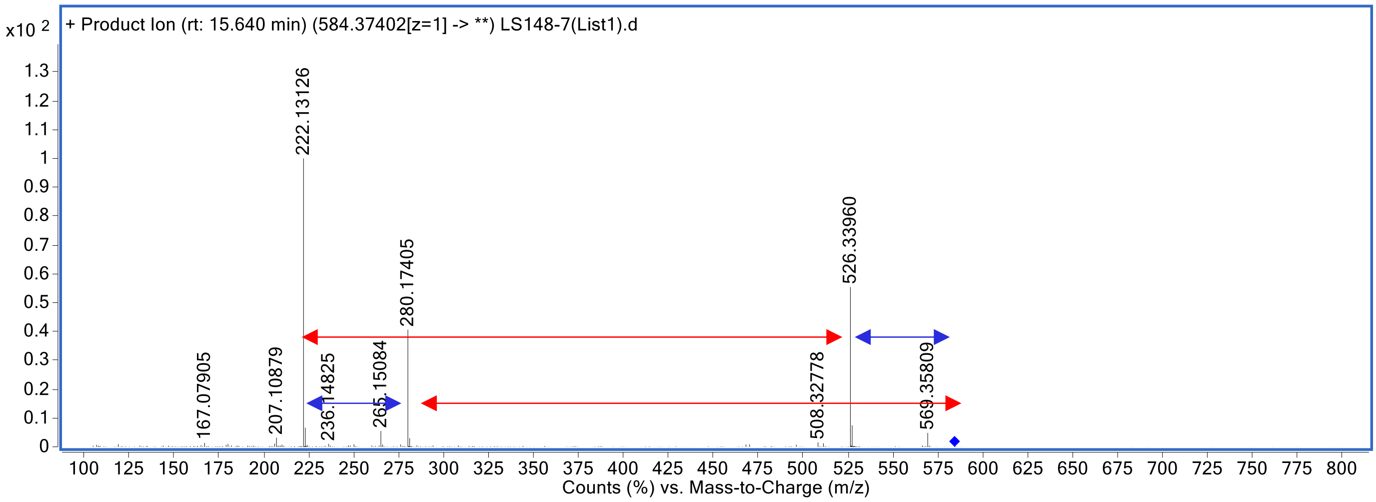
NAE


H
O
OH
N
H
H
OH

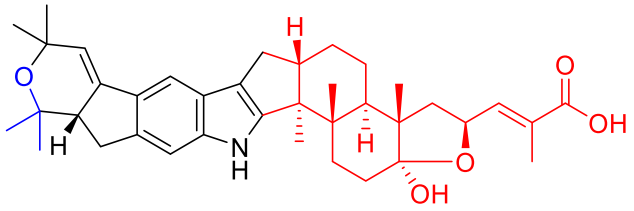

x10 2
+ Product Ion (rt: 22.460 min) (572.40253[z=1] -> **) RC63-5(List1).d
266.19531
119.08840
180.08440
198.13162
210.13154
Counts (%) vs. Mass-to-Charge (m/z)

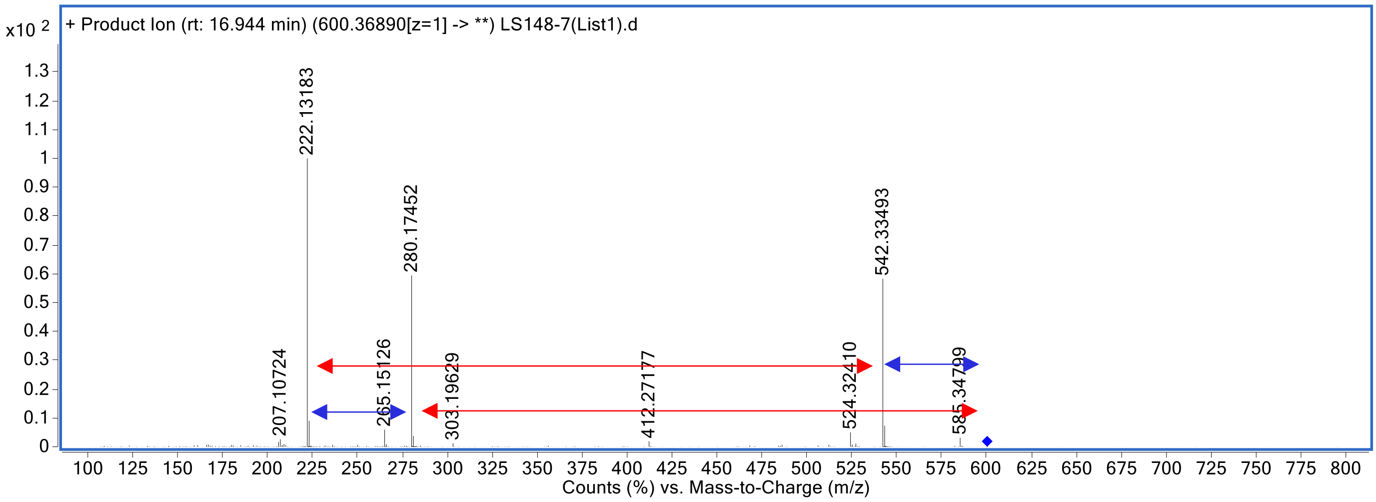
NAD


H
O
O
OH
H
N
H
H
OH

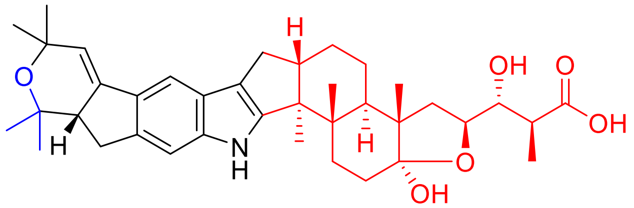

x10 2
+ Product Ion (rt: 15.640 min) (584.37402[z=1] -> **) LS148-7(List1).d
222.13126
526.33960
280.17405
167.07905
207.10879
236.14825
265.15084
508.32778
569.35809
Counts (%) vs. Mass-to-Charge (m/z)

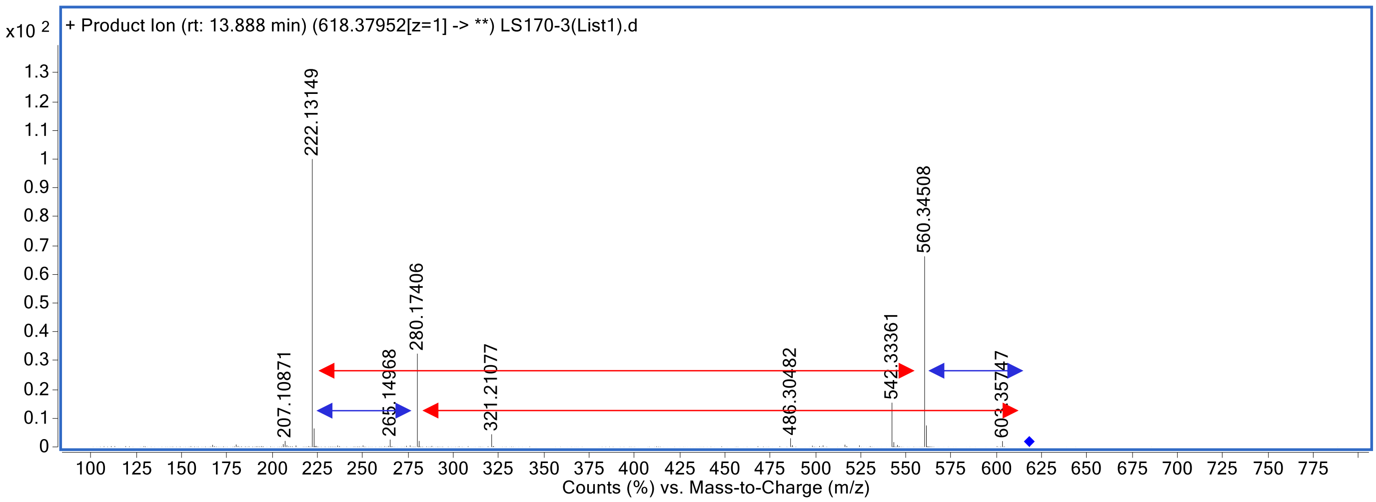
NAD1


H
O
H
N
H
H
O
OH
O
OH

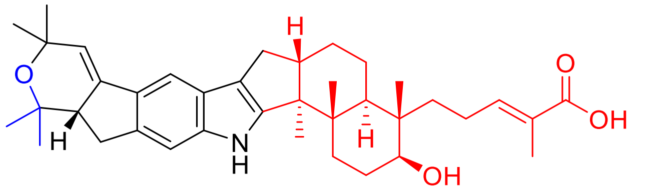

+ Product Ion (rt: 16.944 min) (600.36890[z=1] -> **) LS148-7(List1).d
x10 2
222.13183
280.17452
542.33493
207.10724
265.15126
303.19629
412.27177
524.32410
585.34799
Counts (%) vs. Mass-to-Charge (m/z)

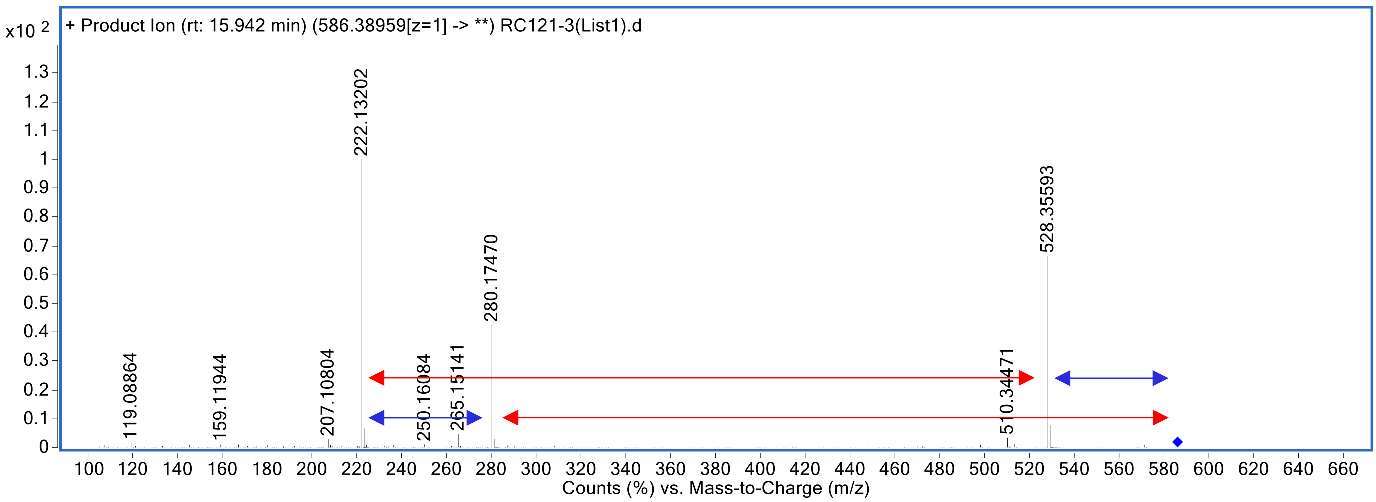
NAD2


H
O
H
N
H
H
OH
O
OH
O
OH

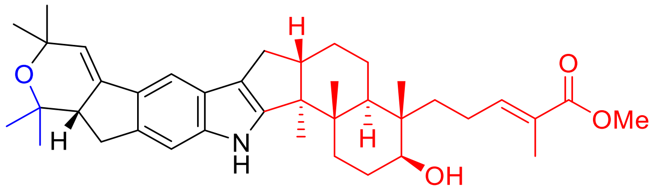

+ Product Ion (rt: 13.888 min) (618.37952[z=1] -> **) LS170-3(List1).d
x10 2
222.13149
560.34508
280.17406
207.10871
265.14968
321.21077
486.30482
542.33361
603.35747
Counts (%) vs. Mass-to-Charge (m/z)

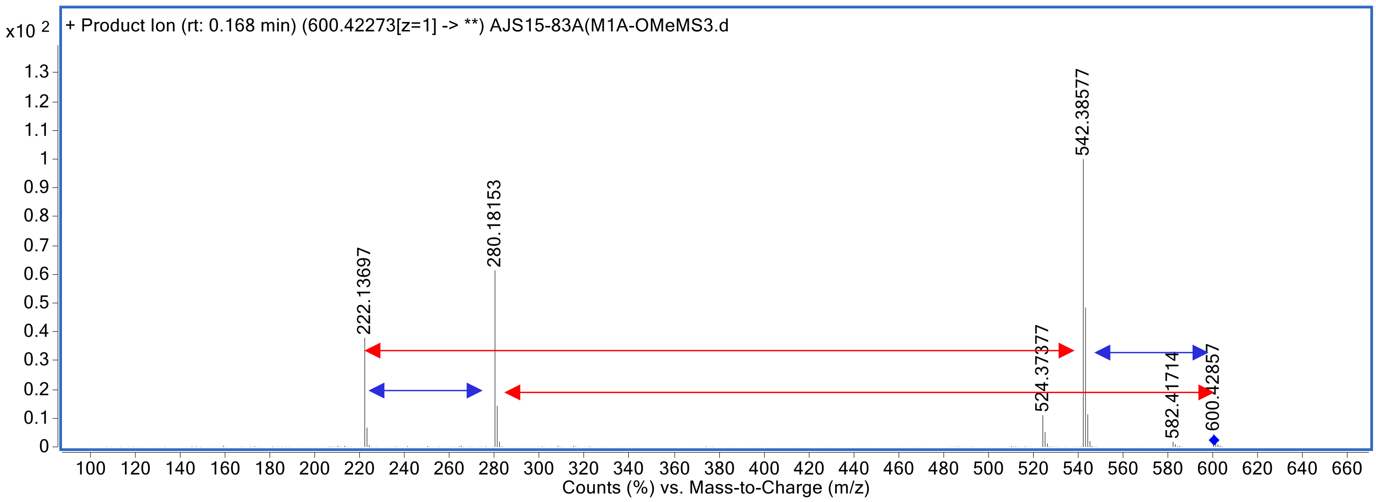
NAD4



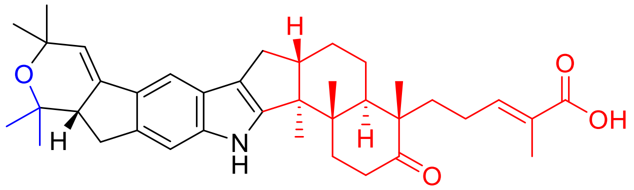


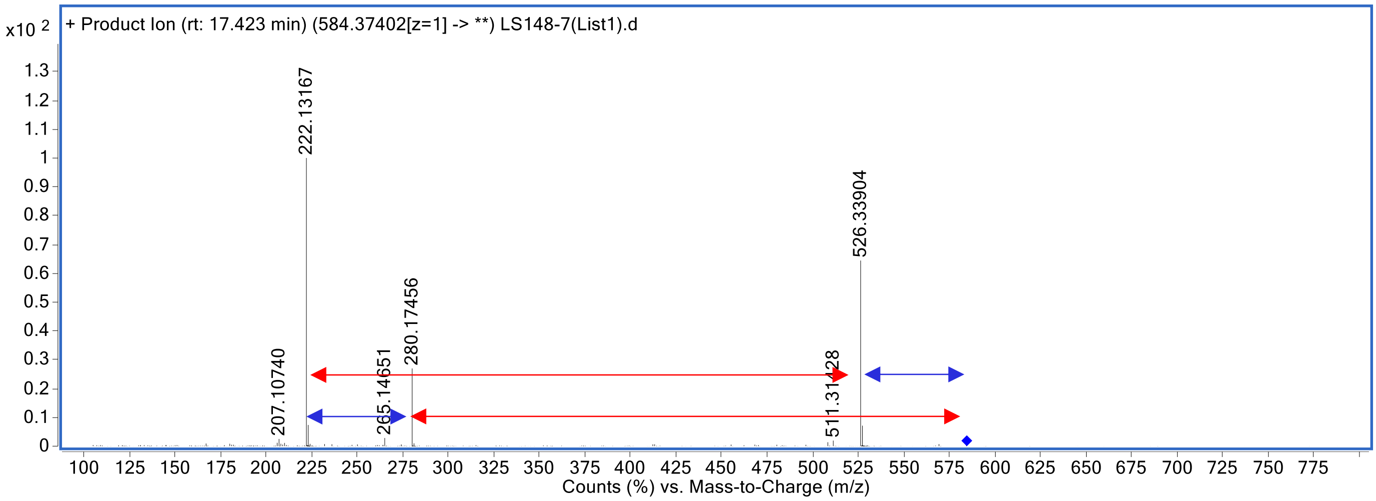
NAD4-OMe



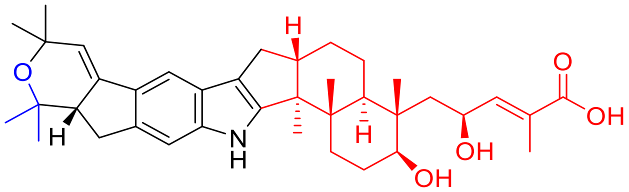


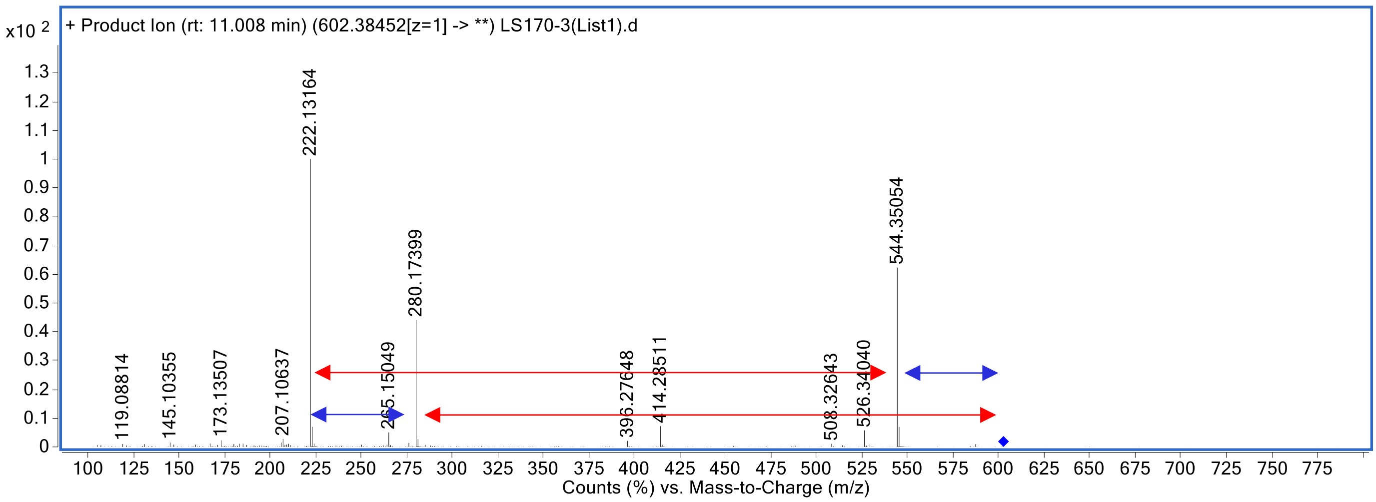
NAD5



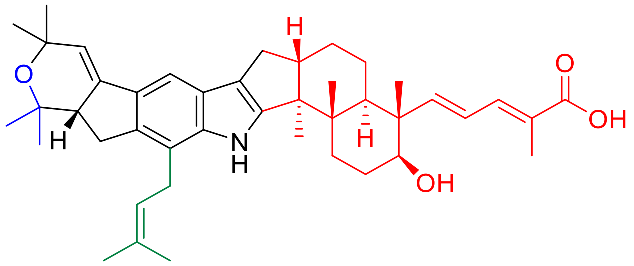

+ Product Ion (rt: 17.423 min) (584.37402[z=1] -> **) LS148-7(List1).d
x10 2
222.13167
526.33904
280.17456
207.10740
265.14651
511.31428
Counts (%) vs. Mass-to-Charge (m/z)

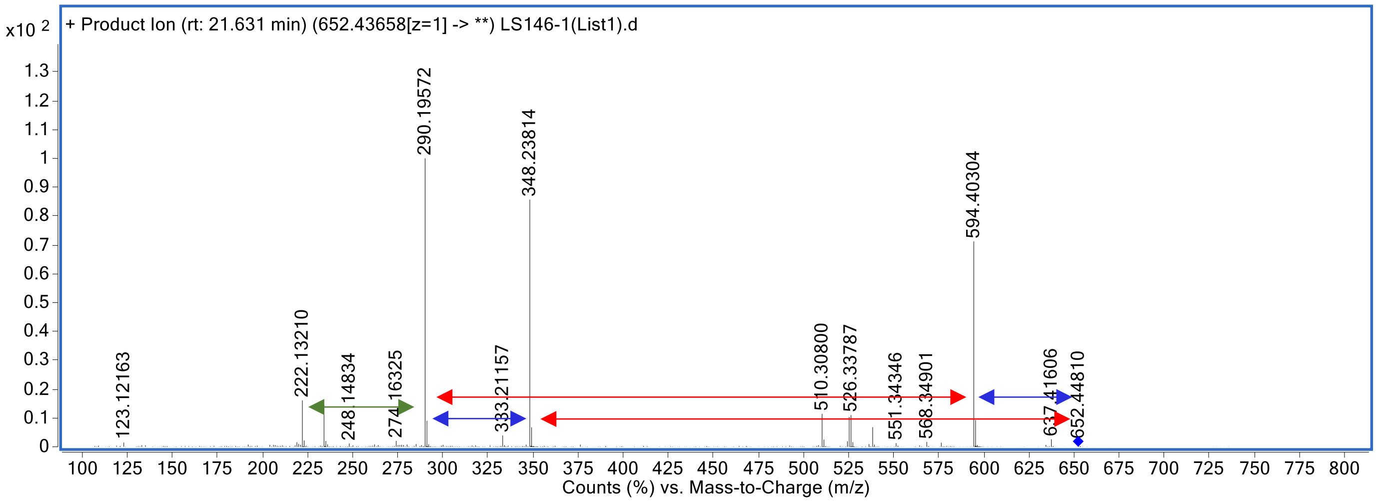
NAD6



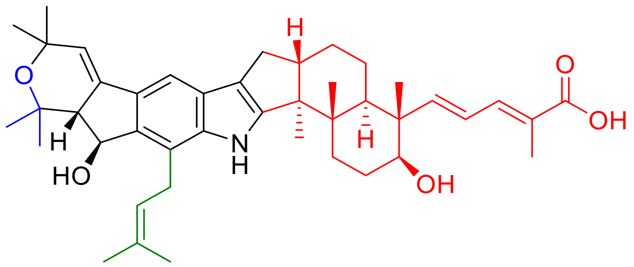

+ Product Ion (rt: 11.008 min) (602.38452[z=1] -> **) LS170-3(List1).d
x10 2
222.13164
544.35054
280.17399
119.08814
145.10355
173.13507
207.10637
265.15049
396.27648
414.28511
508.32643
526.34040
Counts (%) vs. Mass-to-Charge (m/z)

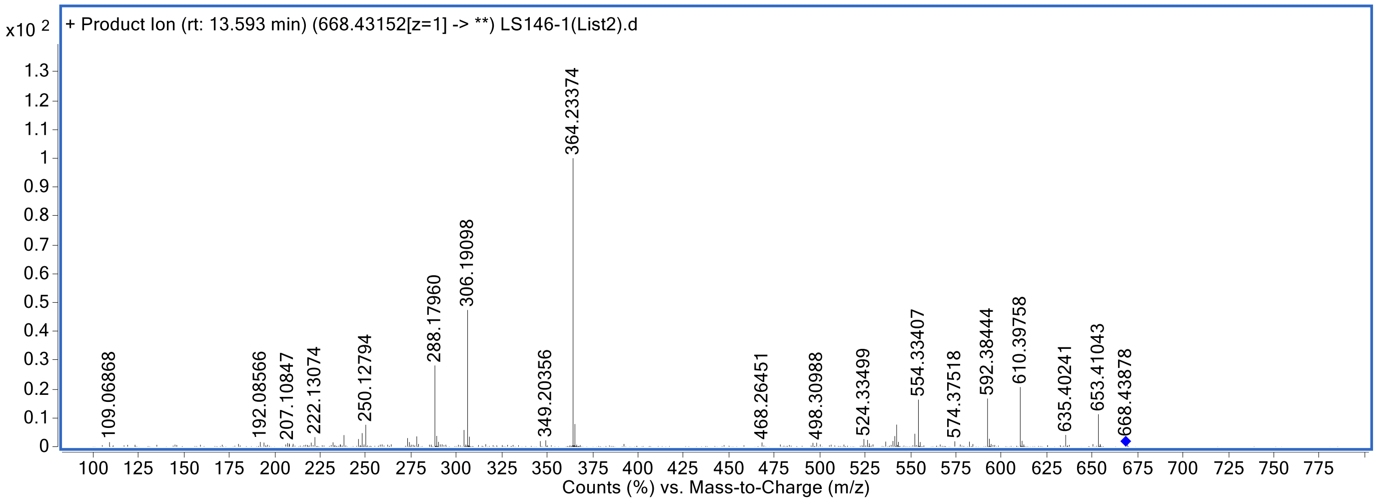
DH-NAC



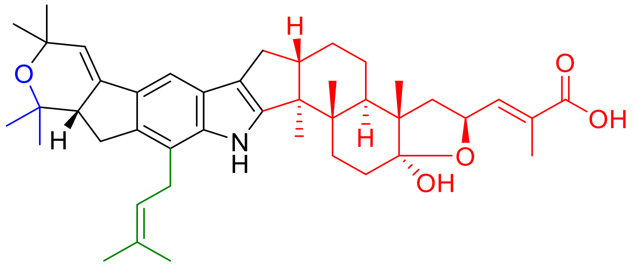

x10 2
+ Product Ion (rt: 21.631 min) (652.43658[z=1] -> **) LS146-1(List1).d
123.12163
222.13210
248.14834
274.16325
290.19572
333.21157
348.23814
510.30800
526.33787
551.34346
568.34901
594.40304
637.41606
652.44810
Counts (%) vs. Mass-to-Charge (m/z)

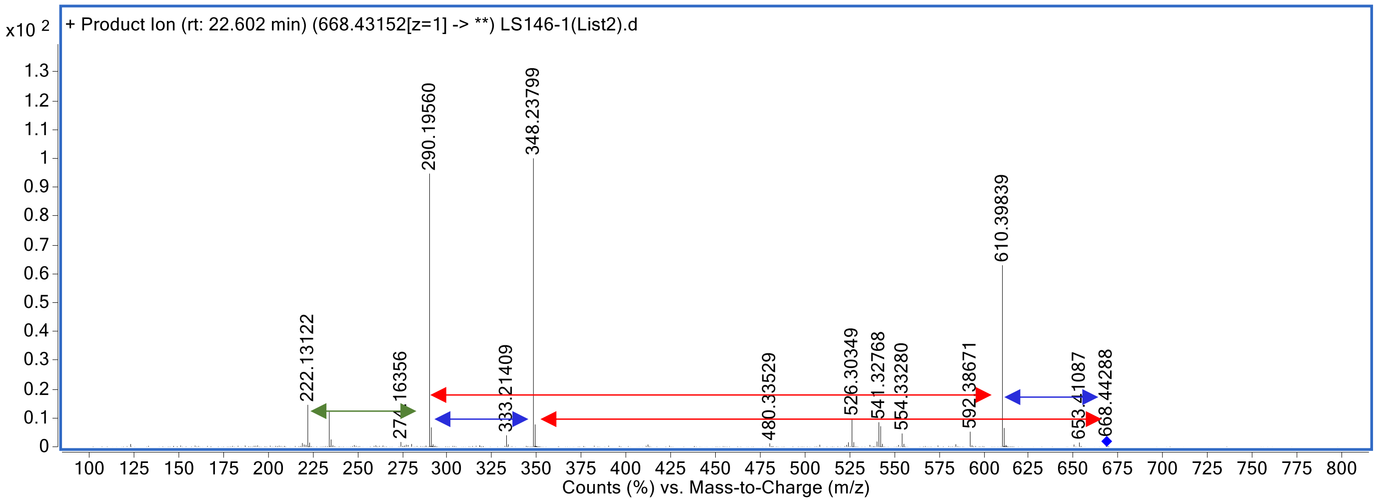
NAC



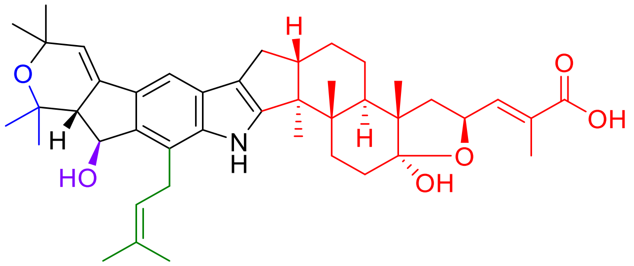

x10 2
+ Product Ion (rt: 13.593 min) (668.43152[z=1] -> **) LS146-1(List2).d
109.06868
192.08566
207.10847
222.13074
250.12794
288.17960
306.19098
349.20356
364.23374
468.26451
498.30988
524.33499
554.33407
574.37518
592.38444
610.39758
635.40241
653.41043
668.43878
Counts (%) vs. Mass-to-Charge (m/z)

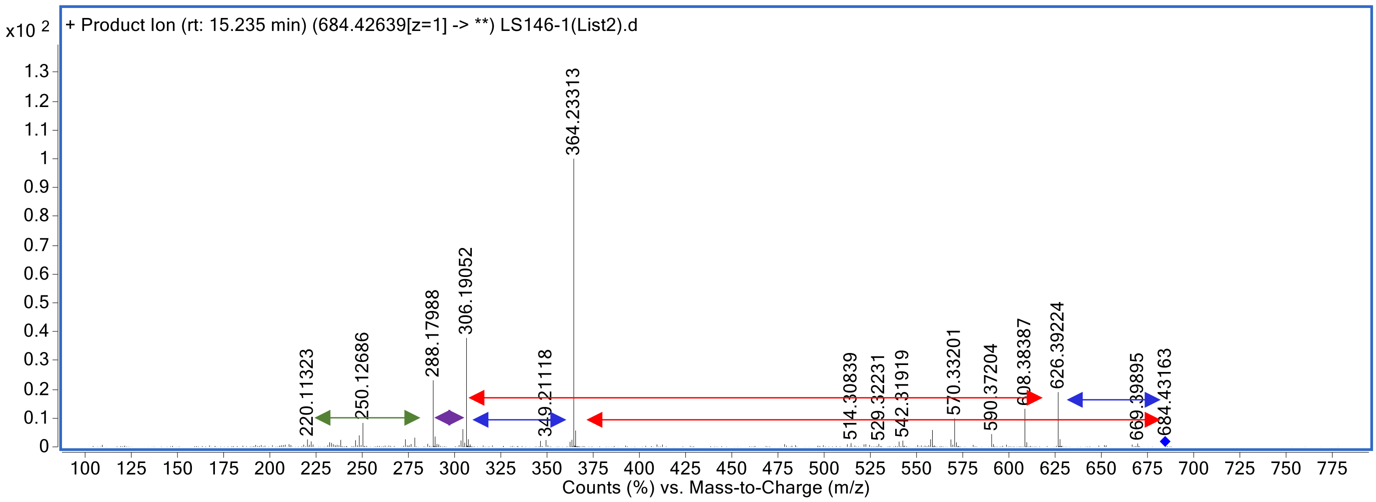
DH-NAC1


H
O
OH
H
N
H
H
O
OH

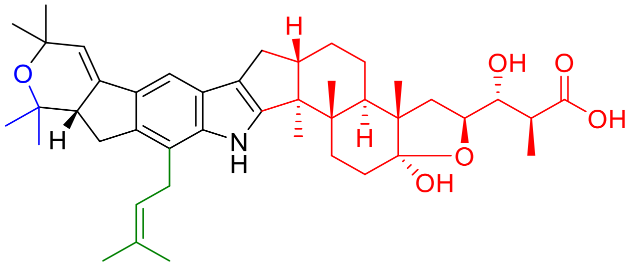

x10 2
+ Product Ion (rt: 22.602 min) (668.43152[z=1] -> **) LS146-1(List2).d
222.13122
274.16356
290.19560
333.21409
348.23799
480.33529
526.30349
541.32768
554.33280
592.38671
610.39839
653.41087
668.44288
Counts (%) vs. Mass-to-Charge (m/z)

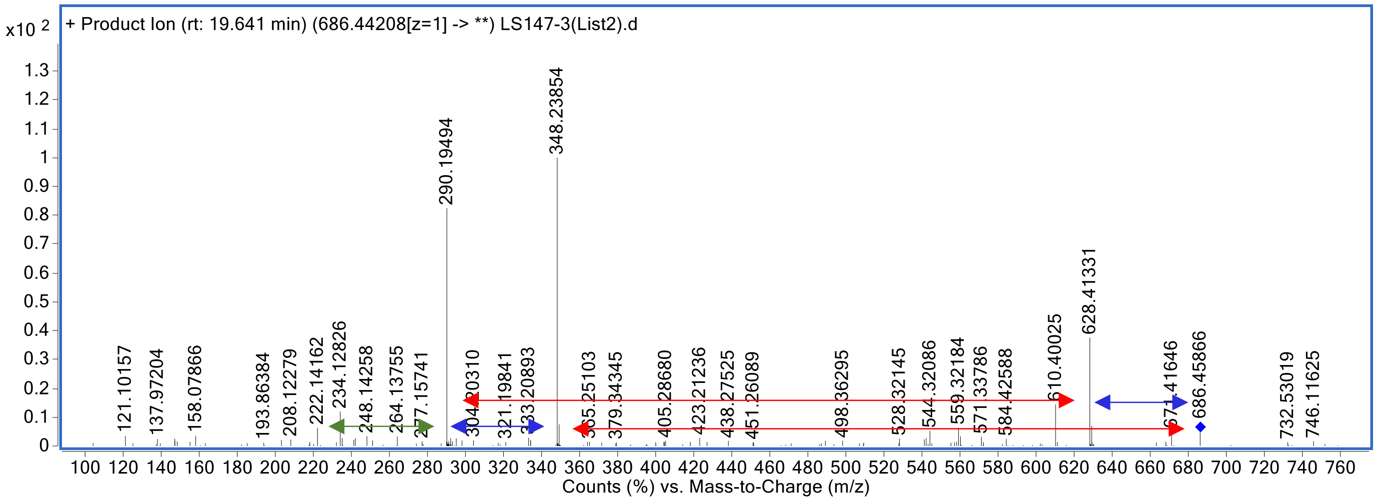
NAC1


H
O
OH
H
HO
N
H
H
O
OH

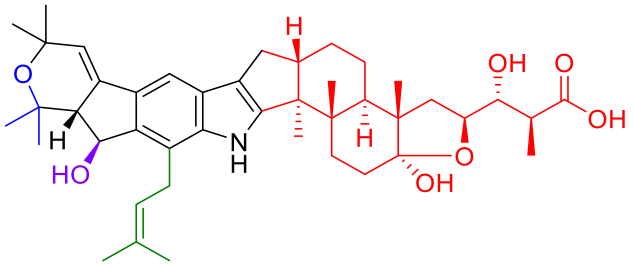

x10 2
+ Product Ion (rt: 15.235 min) (684.42639[z=1] -> **) LS146-1(List2).d
220.11323
250.12686
288.17988
306.19052
349.21118
364.23313
514.30839
529.32231
542.31919
570.33201
590.37204
608.38387
626.39224
669.39895
684.43163
Counts (%) vs. Mass-to-Charge (m/z)

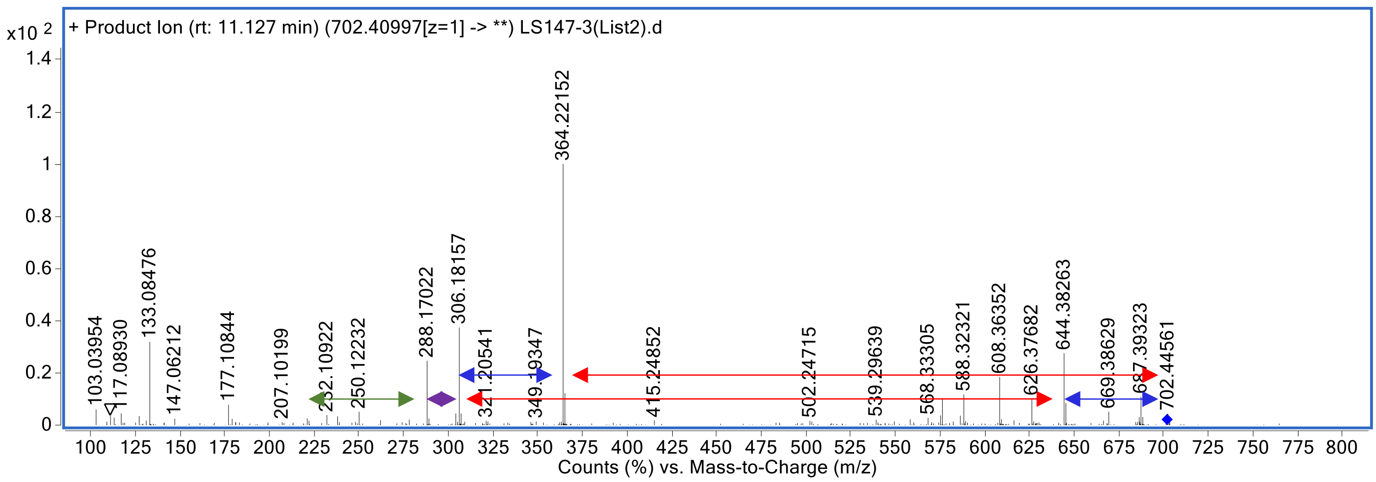
DH-NAC2



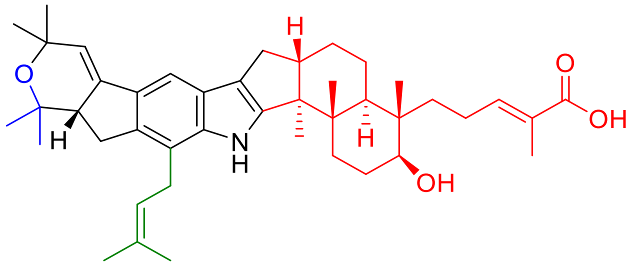

x10 2
+ Product Ion (rt: 19.641 min) (686.44208[z=1] -> **) LS147-3(List2).d
348.23854
290.19494
628.41331
610.40025
686.45866
Counts (%) vs. Mass-to-Charge (m/z)

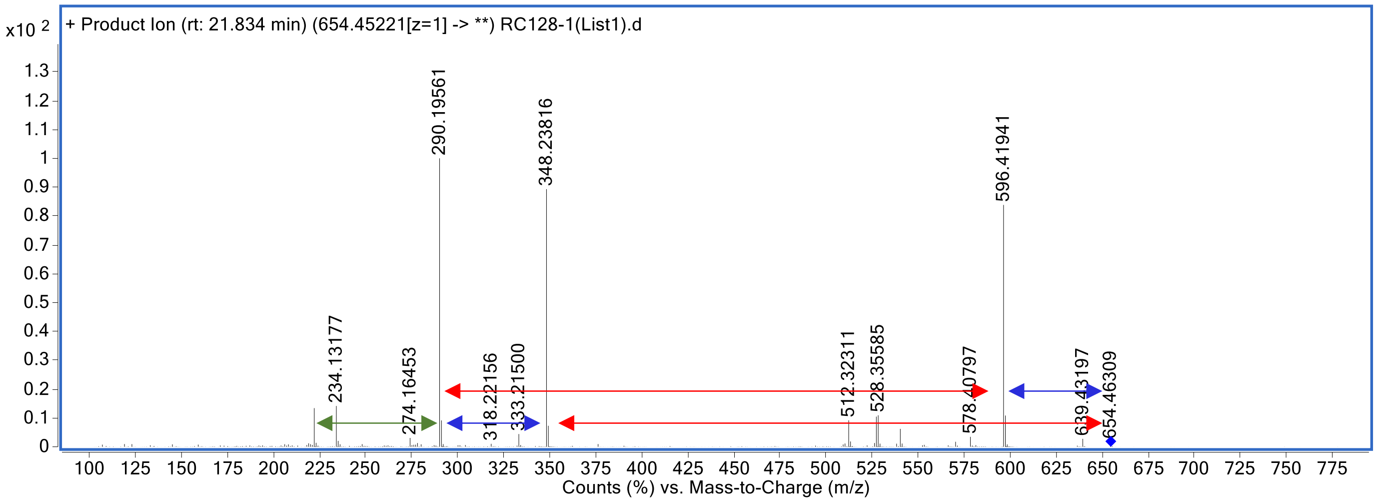
NAC2



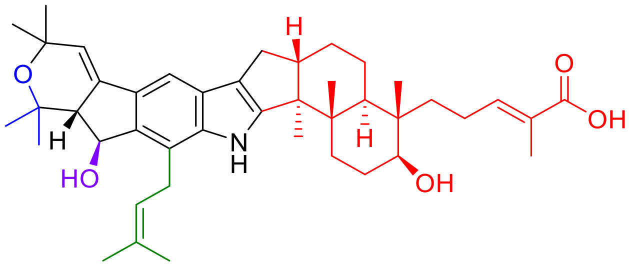

x10 2
+ Product Ion (rt: 11.127 min) (702.40997[z=1] -> **) LS147-3(List2).d
364.22152
306.18157
288.17022
133.08476
644.38263
702.44561
Counts (%) vs. Mass-to-Charge (m/z)

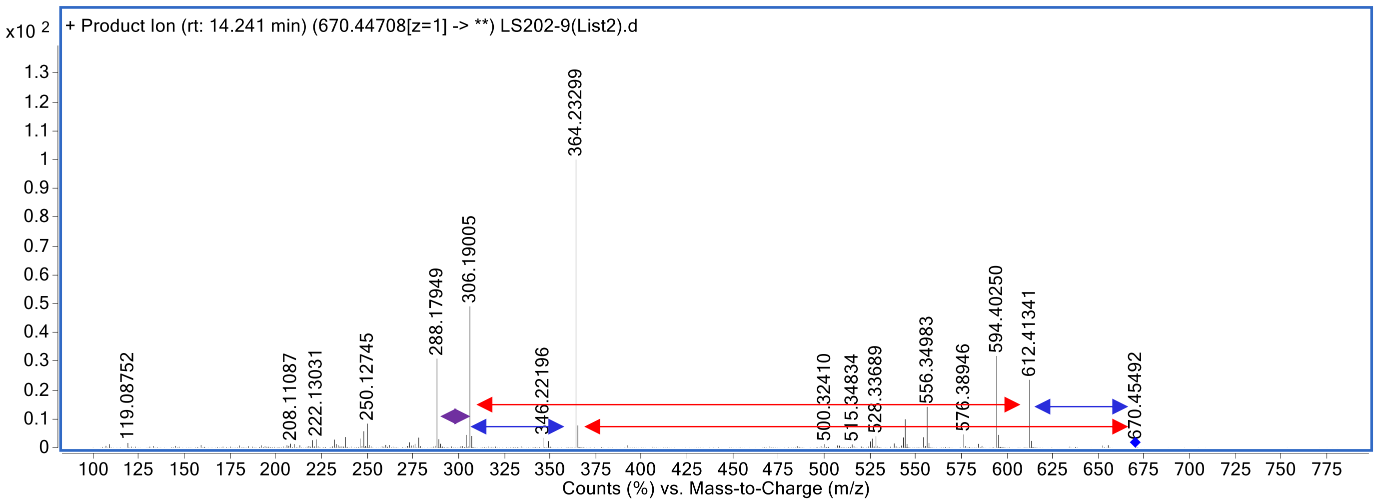
DH-NAC4


H
O
OH
H
N
H
H
OH

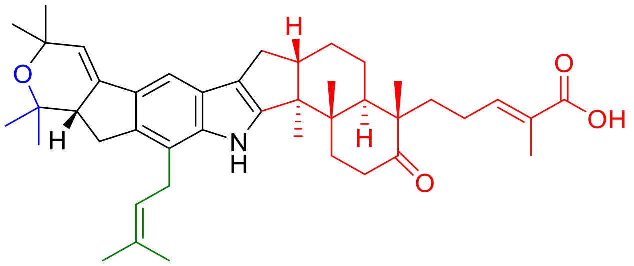

x10 2
+ Product Ion (rt: 21.834 min) (654.45221[z=1] -> **) RC128-1(List1).d
290.19561
348.23816
596.41941
234.13177
274.16453
318.22156
333.21500
512.32311
528.35585
578.40797
639.43197
654.46309
Counts (%) vs. Mass-to-Charge (m/z)

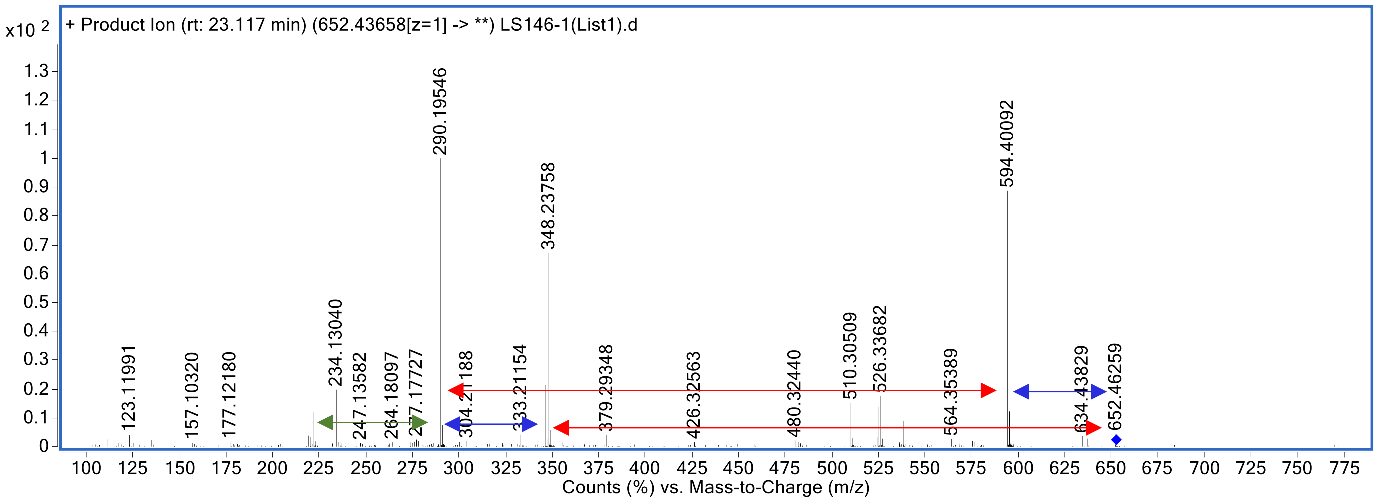
NAC4


H
O
OH
H
HO
N
H
H
OH

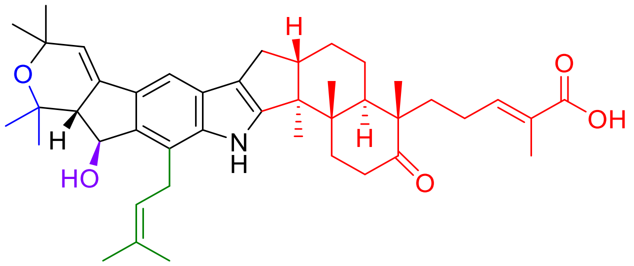

x10 2
+ Product Ion (rt: 14.241 min) (670.44708[z=1] -> **) LS202-9(List2).d
364.23299
306.19005
288.17949
119.08752
208.11087
222.13031
250.12745
346.22196
500.32410
515.34834
528.33689
556.34983
576.38946
594.40250
612.41341
670.45492
Counts (%) vs. Mass-to-Charge (m/z)

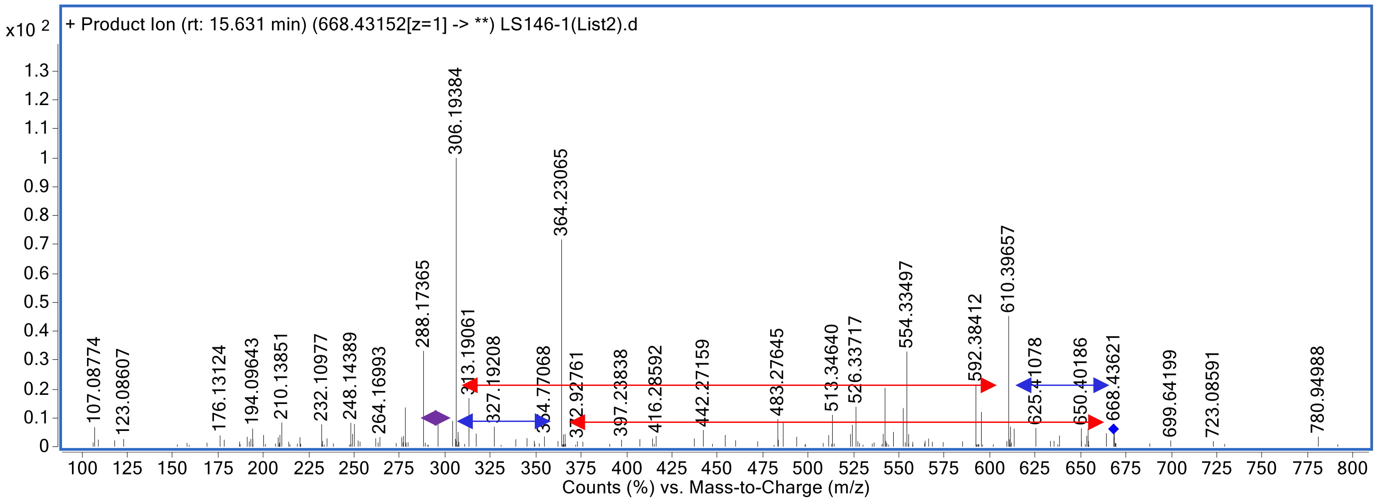
DH-NAC5


O
H
N
H
H
H
O
O
OH

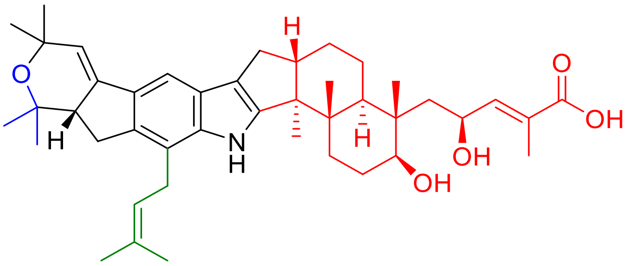

x10 2
+ Product Ion (rt: 23.117 min) (652.43658[z=1] -> **) LS146-1(List1).d
123.11991
157.10320
177.12180
234.13040
247.13582
264.18097
277.17727
290.19546
304.21188
333.21154
348.23758
379.29348
426.32563
480.32440
510.30509
526.33682
564.35389
594.40092
634.43829
652.46259
Counts (%) vs. Mass-to-Charge (m/z)

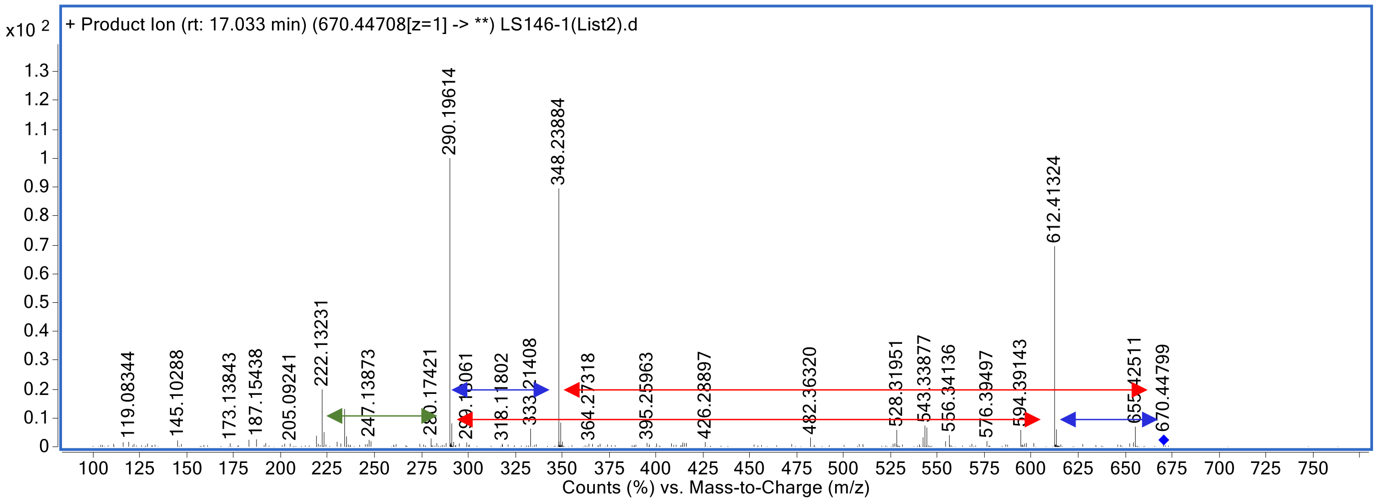
NAC5


O
H
HO
N
H
H
H
O
O
OH

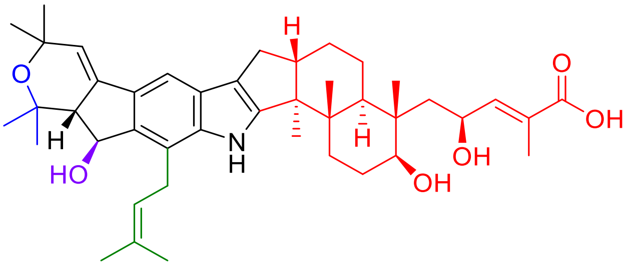

x10 2
+ Product Ion (rt: 15.631 min) (668.43152[z=1] -> **) LS146-1(List2).d
107.08774
123.08607
176.13124
194.09643
210.13851
232.10977
248.14389
264.16993
288.17365
306.19384
313.19061
327.19208
354.77068
364.23065
372.92761
397.23838
416.28592
442.27159
483.27645
513.34640
526.33717
554.33497
592.38412
610.39657
625.41078
650.40186
668.43621
699.64199
723.08591
780.94988
Counts (%) vs. Mass-to-Charge (m/z)

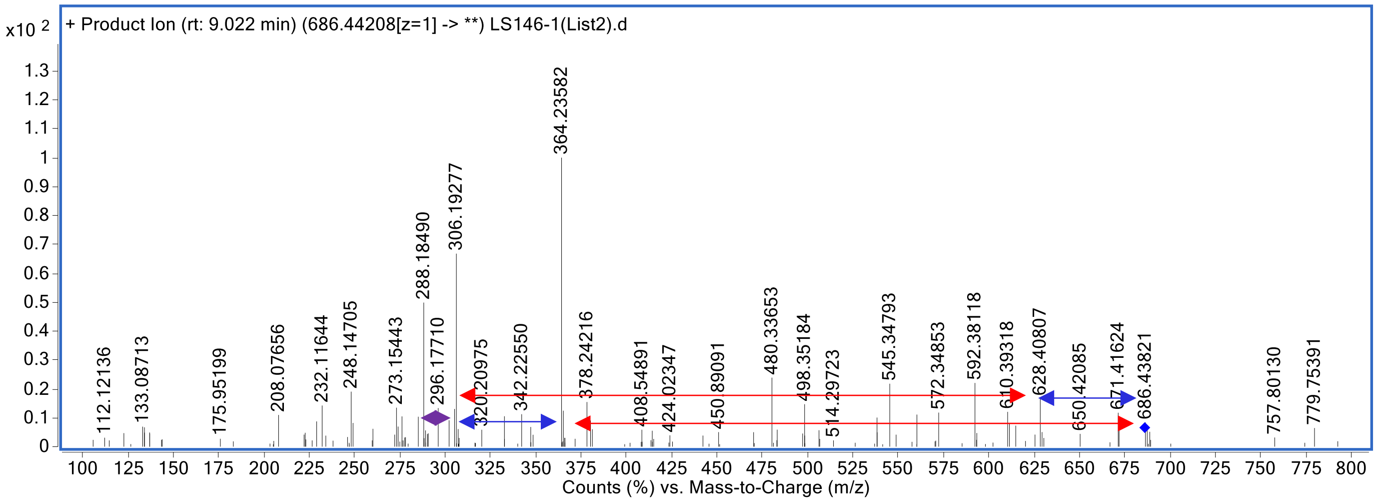
DH-NAC$_6$


H
O
OH
OH
OH
H
N
H
H

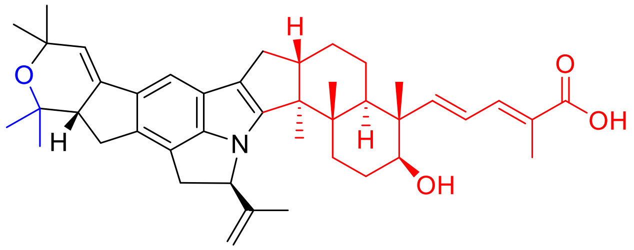

+ Product Ion (rt: 17.033 min) (670.44708[z=1] -> **) LS146-1(List2).d
x10 2
119.08344
145.10288
173.13843
187.15438
205.09241
222.13231
247.13873
280.17421
290.19614
318.11802
333.21408
348.23884
364.27318
395.25963
426.23897
482.36320
528.31951
543.33877
556.34136
576.39497
594.39143
612.41324
655.42511
670.44799
Counts (%) vs. Mass-to-Charge (m/z)

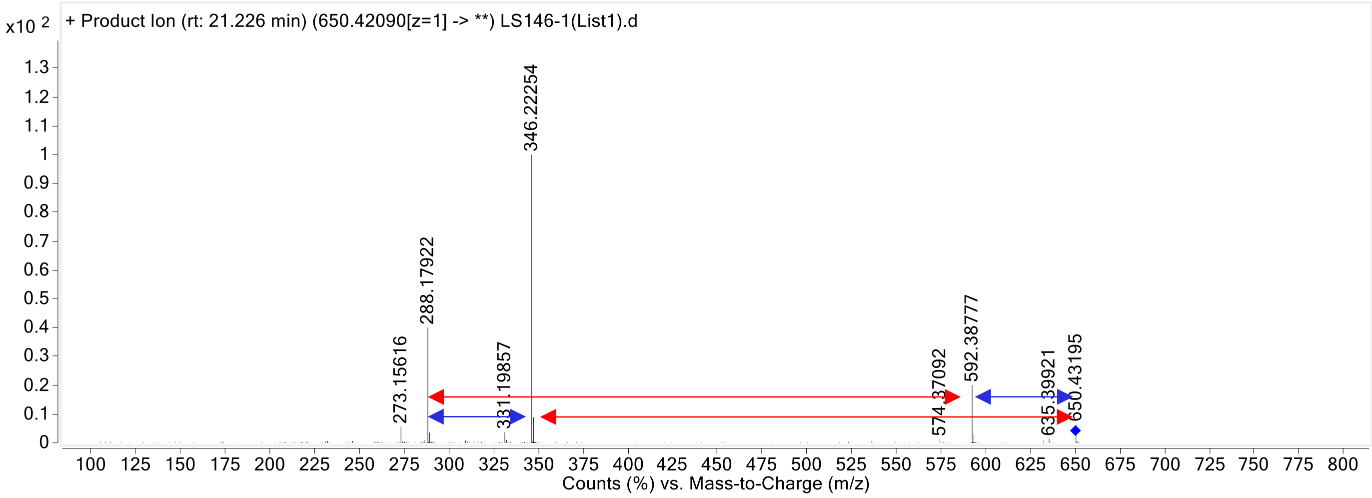
NAC$_6$


H
O
OH
OH
OH
HO
H
N
H
H

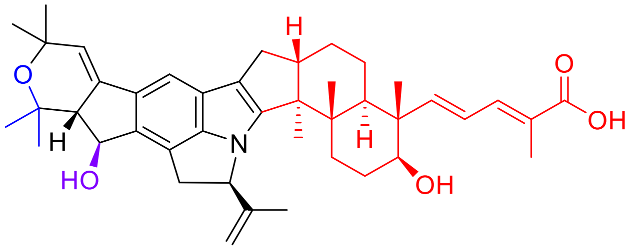

+ Product Ion (rt: 9.022 min) (686.44208[z=1] -> **) LS146-1(List2).d
x10 2
112.12136
133.08713
175.95199
208.07656
232.11644
248.14705
273.15443
288.18490
296.17710
306.19277
342.22550
364.23582
378.24216
408.54891
424.02347
450.89091
480.33653
498.35184
514.29723
545.34793
572.34853
592.38118
628.40807
650.42085
686.43821
757.80130
779.75391
Counts (%) vs. Mass-to-Charge (m/z)

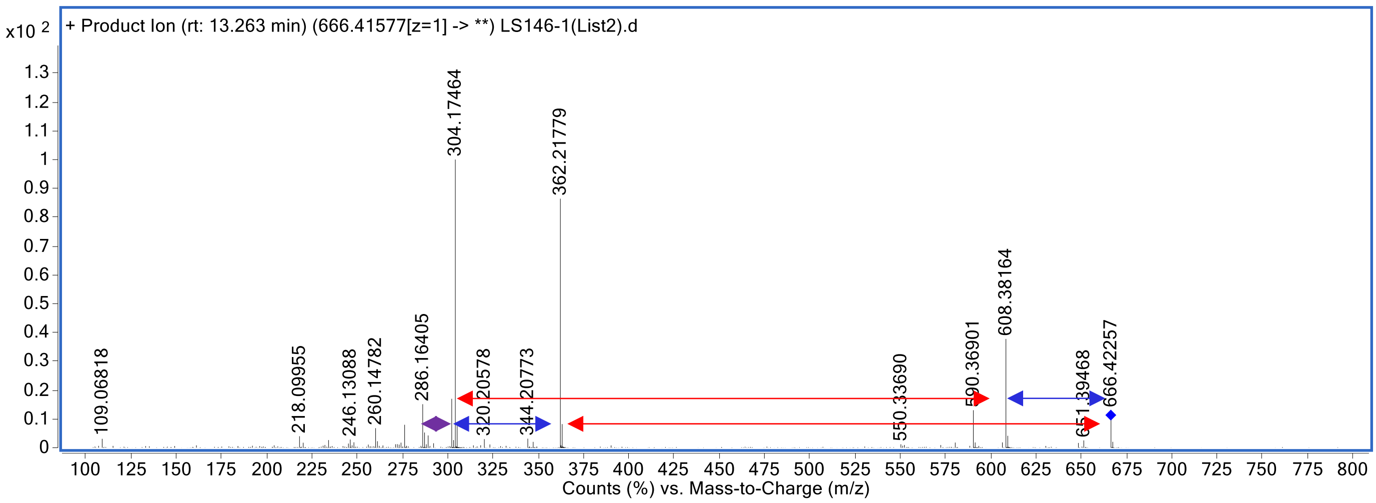
DH-NAB



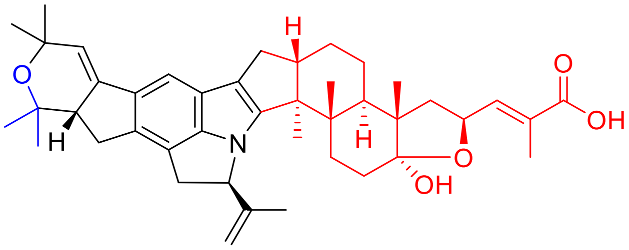

+ Product Ion (rt: 21.226 min) (650.42090[z=1] -> **) LS146-1(List1).d
x10 2
273.15616
288.17922
331.19857
346.22254
574.37092
592.38777
635.39921
650.43195
Counts (%) vs. Mass-to-Charge (m/z)

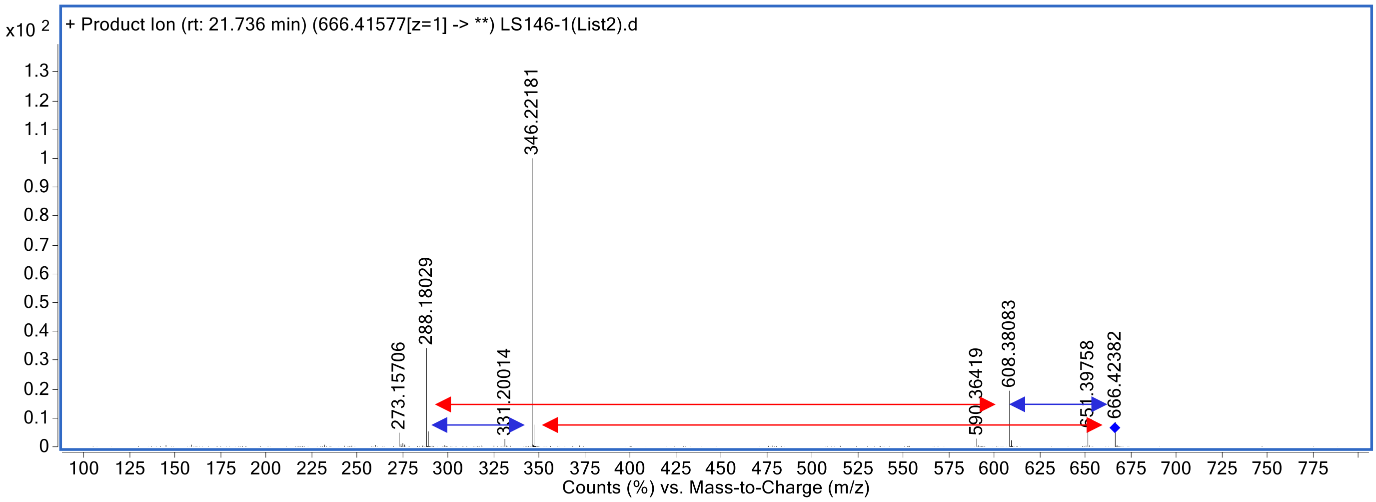
NAB



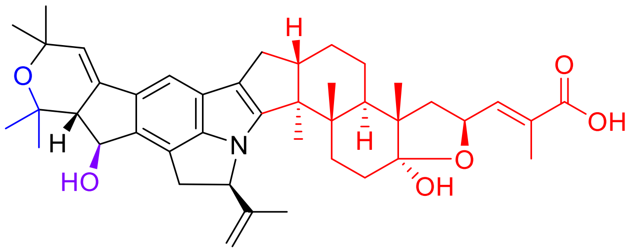

+ Product Ion (rt: 13.263 min) (666.41577[z=1] -> **) LS146-1(List2).d
x10 2
109.06818
218.09955
246.13088
260.14782
286.16405
304.17464
320.20578
344.20773
362.21779
550.33690
590.36901
608.38164
651.39468
666.42257
Counts (%) vs. Mass-to-Charge (m/z)

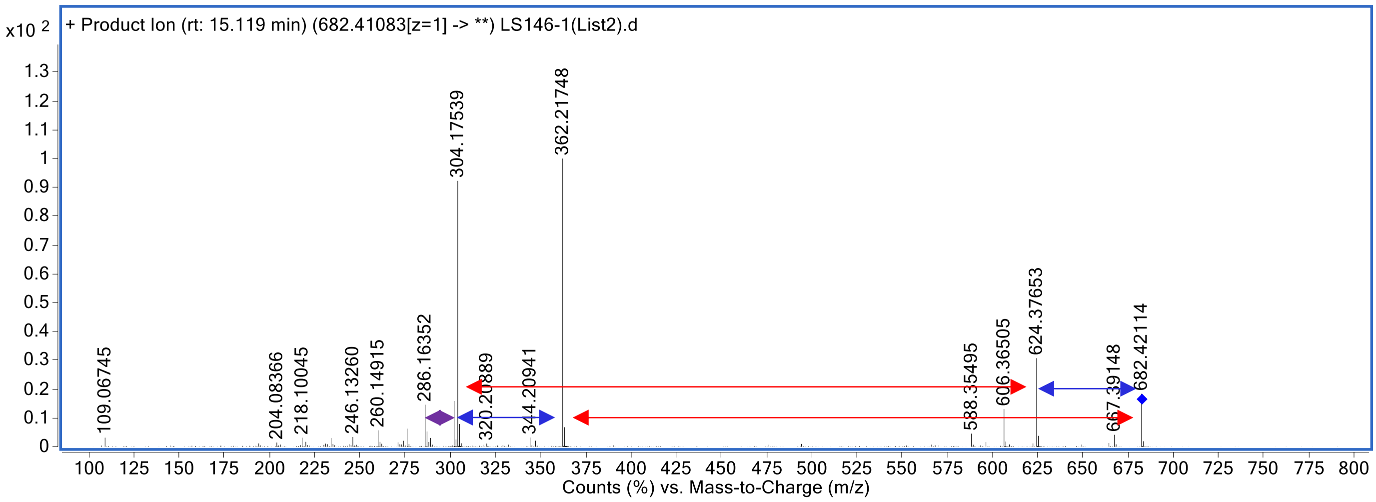
DH-NAB1


H
O
H
N
H
O
OH
O
OH

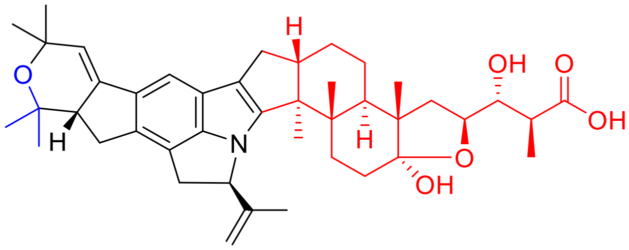

x10 2
+ Product Ion (rt: 21.736 min) (666.41577[z=1] -> **) LS146-1(List2).d
273.15706
288.18029
331.20014
346.22181
590.36419
608.38083
651.39758
666.42382
Counts (%) vs. Mass-to-Charge (m/z)

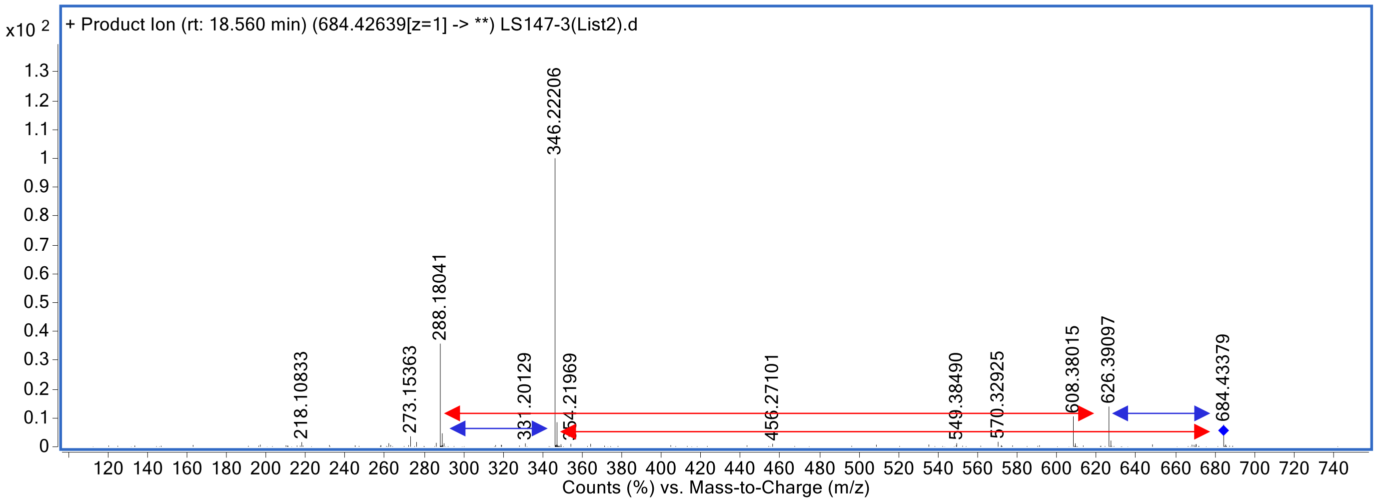
NAB1


H
O
H
HO
N
H
O
OH
O
OH

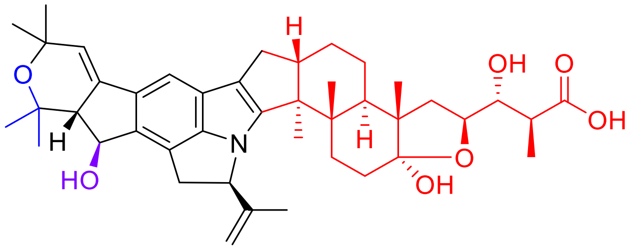

x10 2
+ Product Ion (rt: 15.119 min) (682.41083[z=1] -> **) LS146-1(List2).d
109.06745
204.08366
218.10045
246.13260
260.14915
286.16352
304.17539
320.20889
344.20941
362.21748
588.35495
606.36505
624.37653
667.39148
682.42114
Counts (%) vs. Mass-to-Charge (m/z)

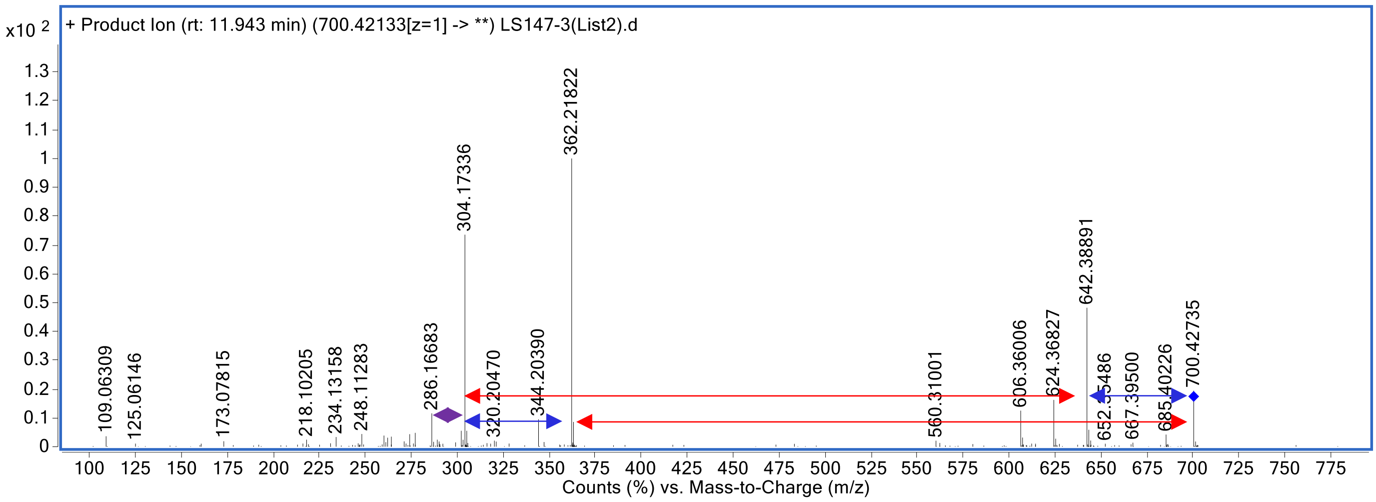
DH-NAB2


O
H
N
H
H
OH
O
OH
O
OH

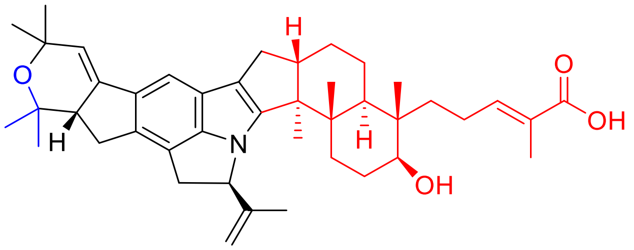

x10 2
+ Product Ion (rt: 18.560 min) (684.42639[z=1] -> **) LS147-3(List2).d
346.22206
288.18041
218.10833
273.15363
331.20129
354.21969
456.27101
549.38490
570.32925
608.38015
626.39097
684.43379
Counts (%) vs. Mass-to-Charge (m/z)

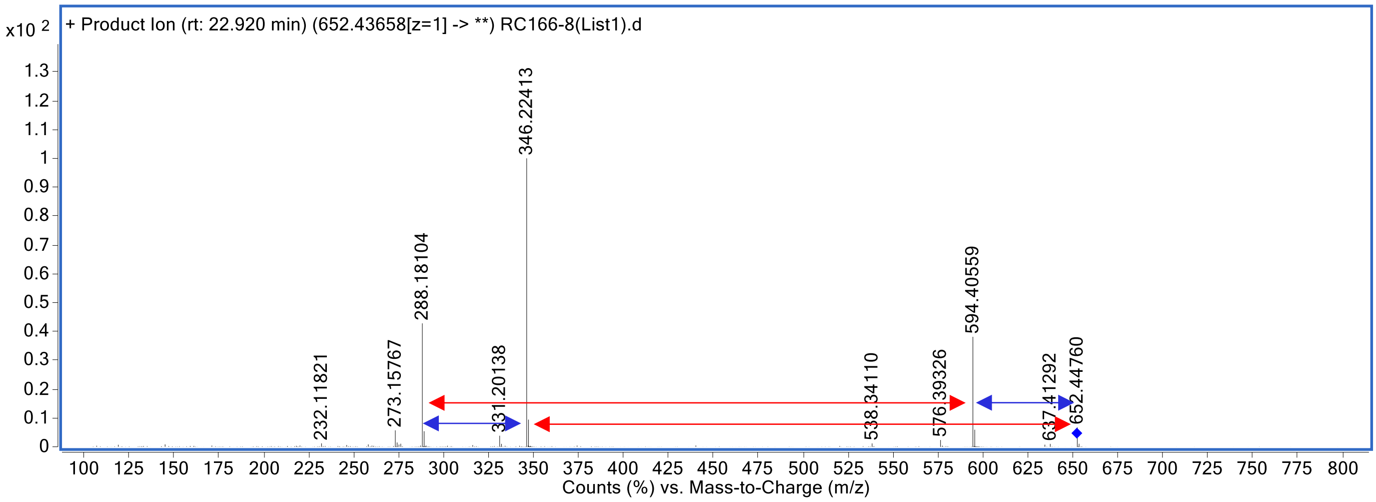
NAB2


O
H
HO
N
H
H
OH
O
OH
O
OH

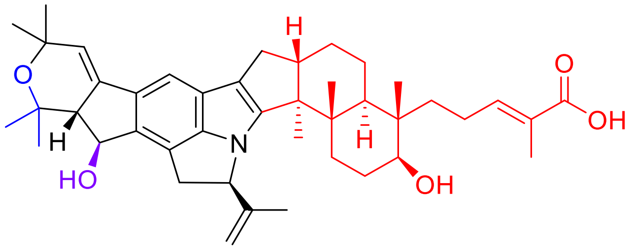

x10 2
+ Product Ion (rt: 11.943 min) (700.42133[z=1] -> **) LS147-3(List2).d
362.21822
304.17336
642.38891
109.06309
125.06146
173.07815
218.10205
234.13158
248.11283
286.16683
320.20470
344.20390
560.31001
606.36006
624.36827
652.35486
667.39500
685.40226
700.42735
Counts (%) vs. Mass-to-Charge (m/z)

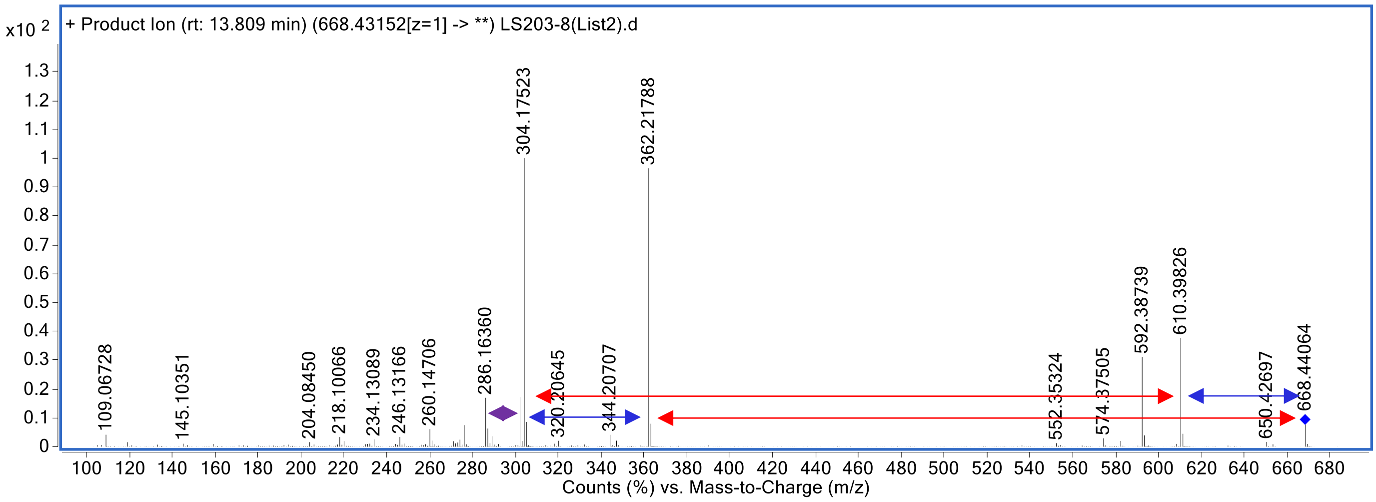
DH-NAB4


O
H
N
H
H
O
OH
OH

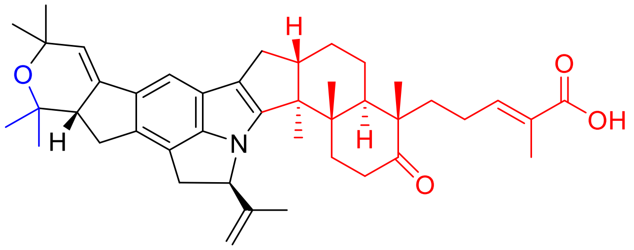

x10 2
+ Product Ion (rt: 22.920 min) (652.43658[z=1] -> **) RC166-8(List1).d
232.11821
273.15767
288.18104
331.20138
346.22413
538.34110
576.39326
594.40559
637.41292
652.44760
Counts (%) vs. Mass-to-Charge (m/z)

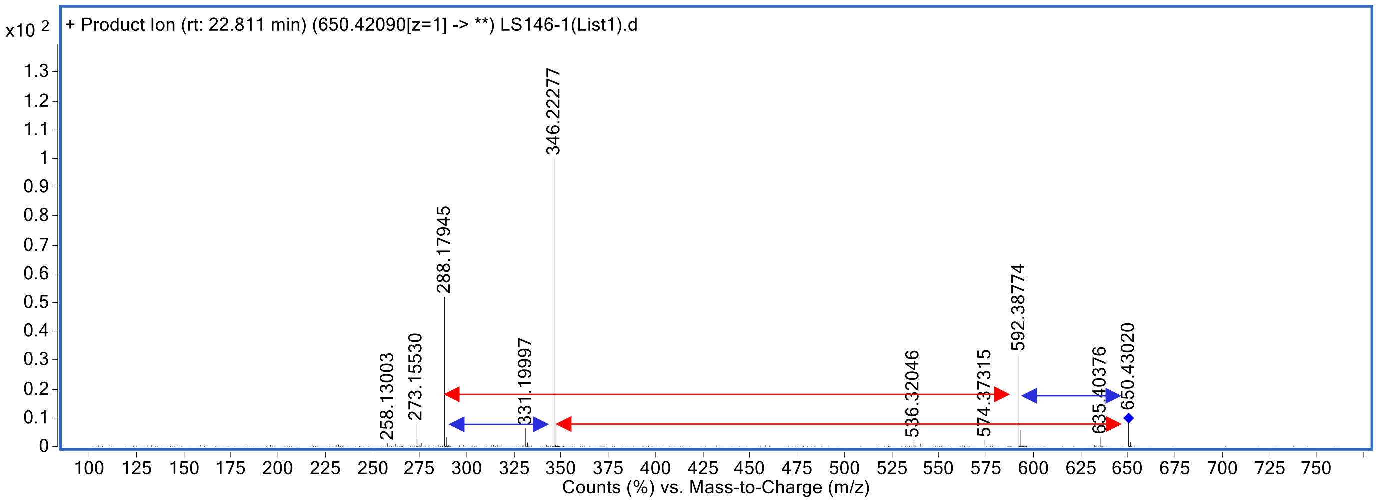
NAB4


O
H
HO
N
H
H
O
OH
OH

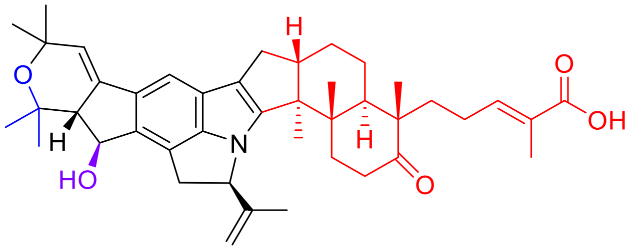

x10 2
+ Product Ion (rt: 13.809 min) (668.43152[z=1] -> **) LS203-8(List2).d
109.06728
145.10351
204.08450
218.10066
234.13089
246.13166
260.14706
286.16360
304.17523
320.20645
344.20707
362.21788
552.35324
574.37505
592.38739
610.39826
650.42697
668.44064
Counts (%) vs. Mass-to-Charge (m/z)

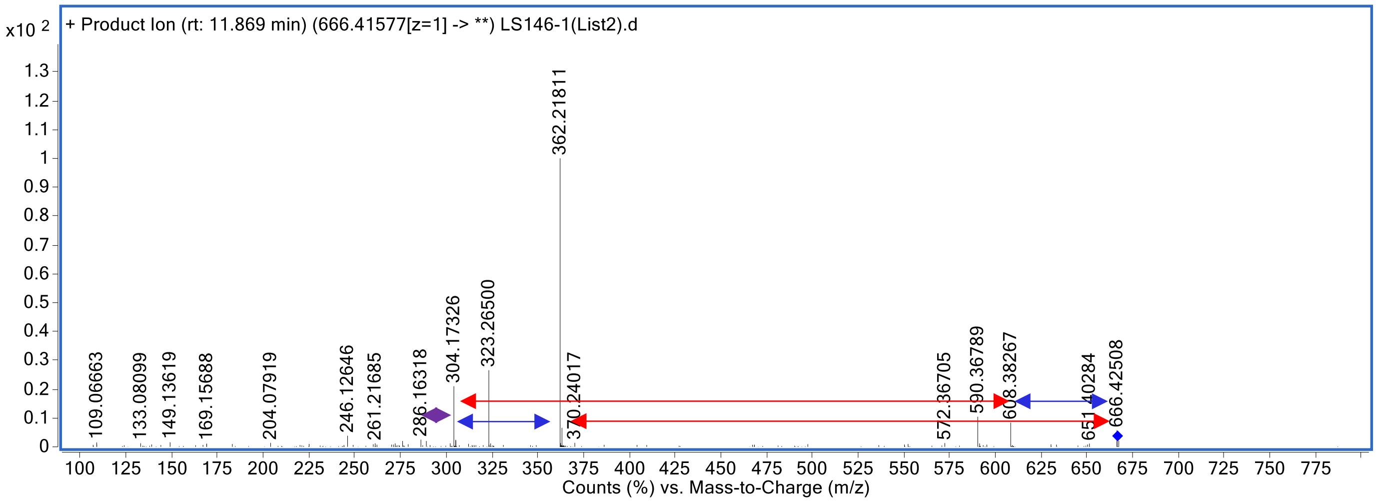
DH-NAB5


H
O
OH
H
H
N
O

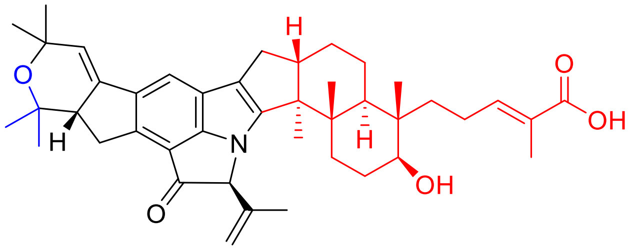

x10 2
+ Product Ion (rt: 22.811 min) (650.42090[z=1] -> **) LS146-1(List1).d
346.22277
288.17945
592.38774
258.13003
273.15530
331.19997
536.32046
574.37315
635.40376
650.43020
Counts (%) vs. Mass-to-Charge (m/z)

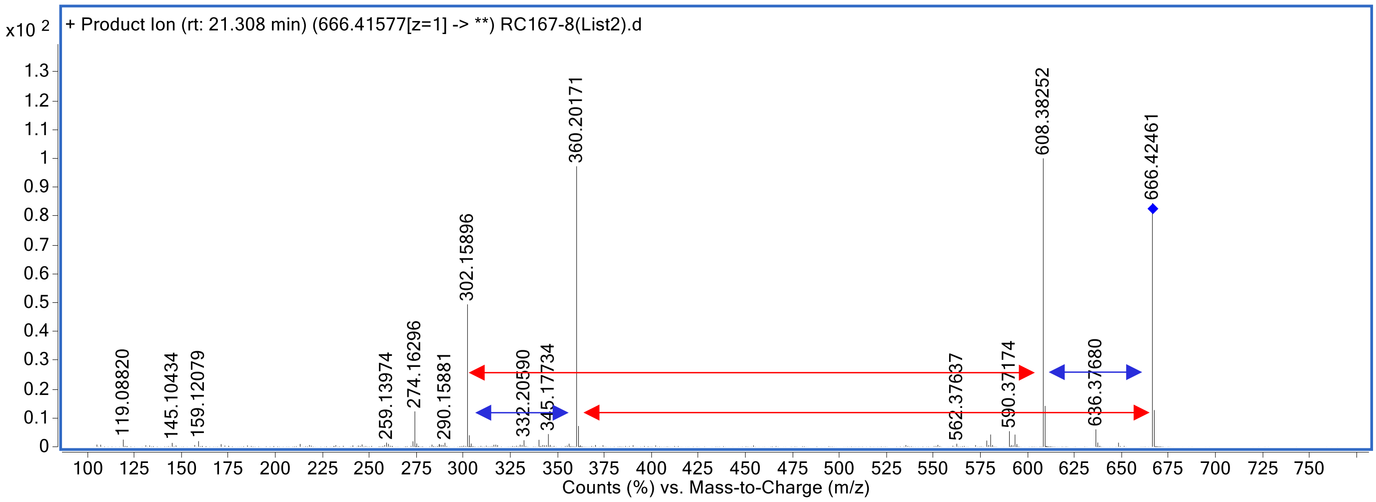
NAB5


H
O
OH
H
HO
H
N
O

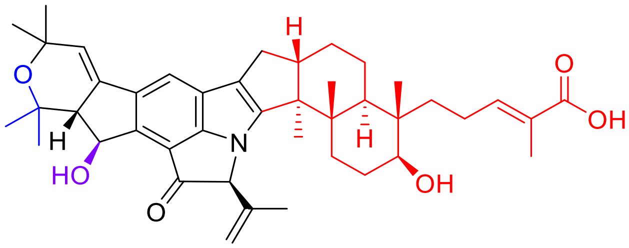

x10 2
+ Product Ion (rt: 11.869 min) (666.41577[z=1] -> **) LS146-1(List2).d
362.21811
109.06663
133.08099
149.13619
169.15688
204.07919
246.12646
261.21685
286.16318
304.17326
323.26500
370.24017
572.36705
590.36789
608.38267
651.40284
666.42508
Counts (%) vs. Mass-to-Charge (m/z)

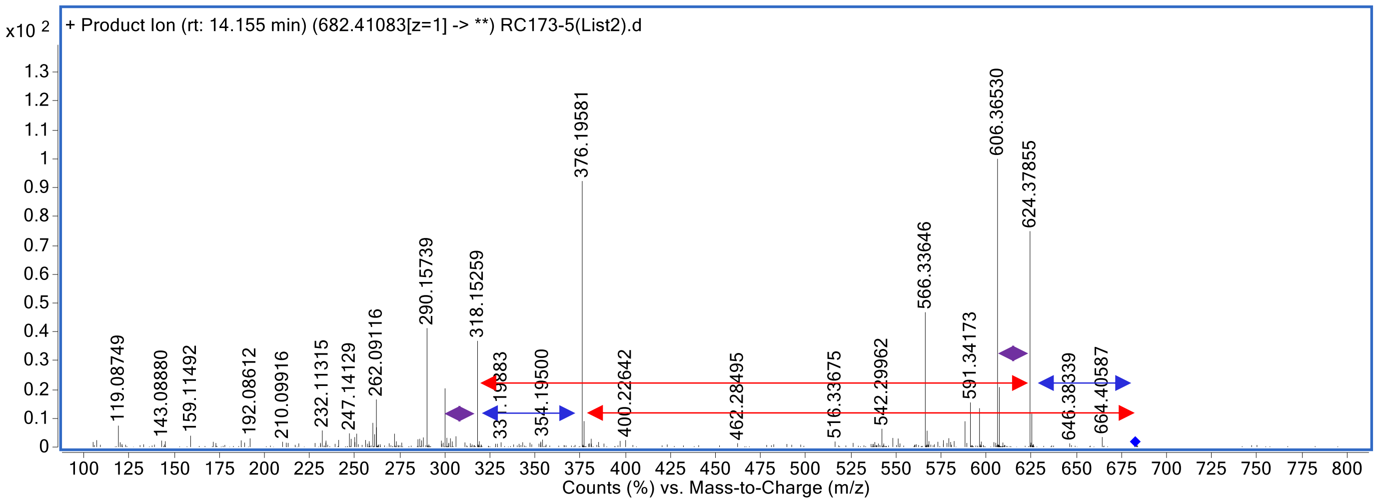
DH-NAA4



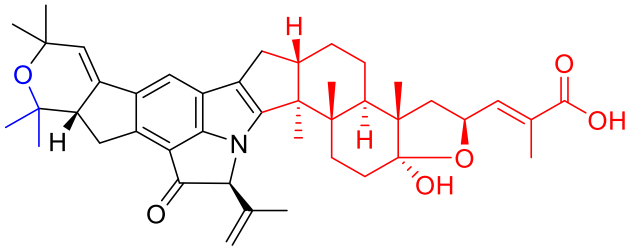

x10 2
+ Product Ion (rt: 21.308 min) (666.41577[z=1] -> **) RC167-8(List2).d
119.08820
145.10434
159.12079
259.13974
274.16296
290.15881
302.15896
332.20590
345.17734
360.20171
562.37637
590.37174
608.38252
636.37680
666.42461
Counts (%) vs. Mass-to-Charge (m/z)

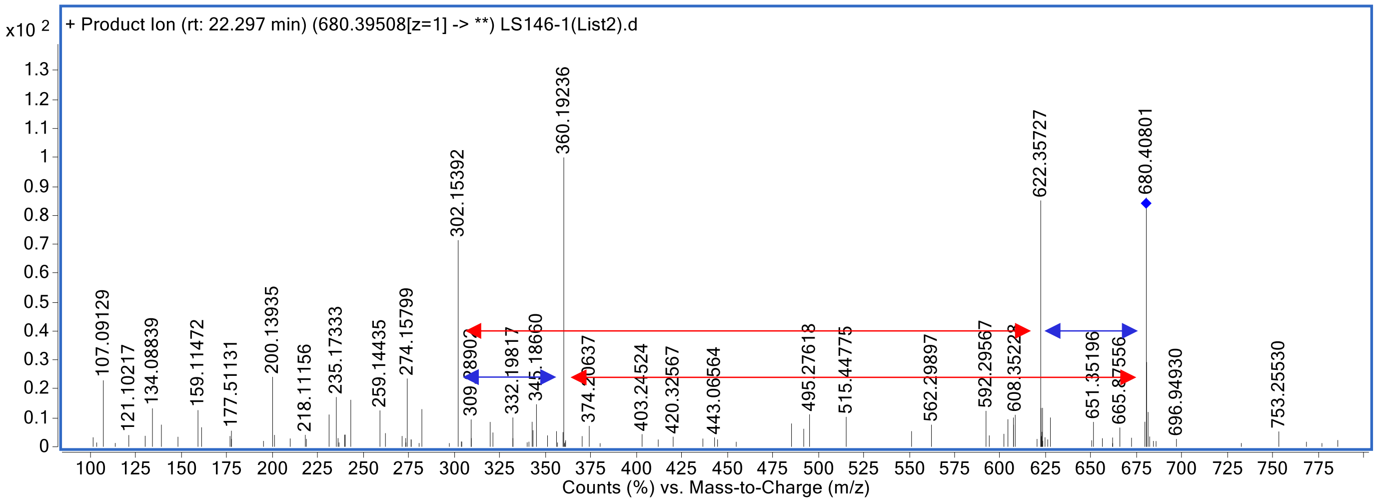
NAA4



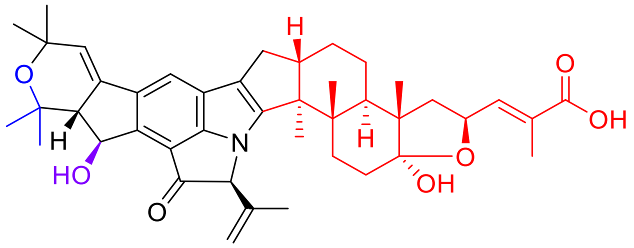

x10 2
+ Product Ion (rt: 14.155 min) (682.41083[z=1] -> **) RC173-5(List2).d
119.08749
143.08880
159.11492
192.08612
210.09916
232.11315
247.14129
262.09116
290.15739
318.15259
336.19883
354.19500
376.19581
400.22642
462.28495
516.33675
542.29962
566.33646
591.34173
606.36530
624.37855
646.38339
664.40587
Counts (%) vs. Mass-to-Charge (m/z)

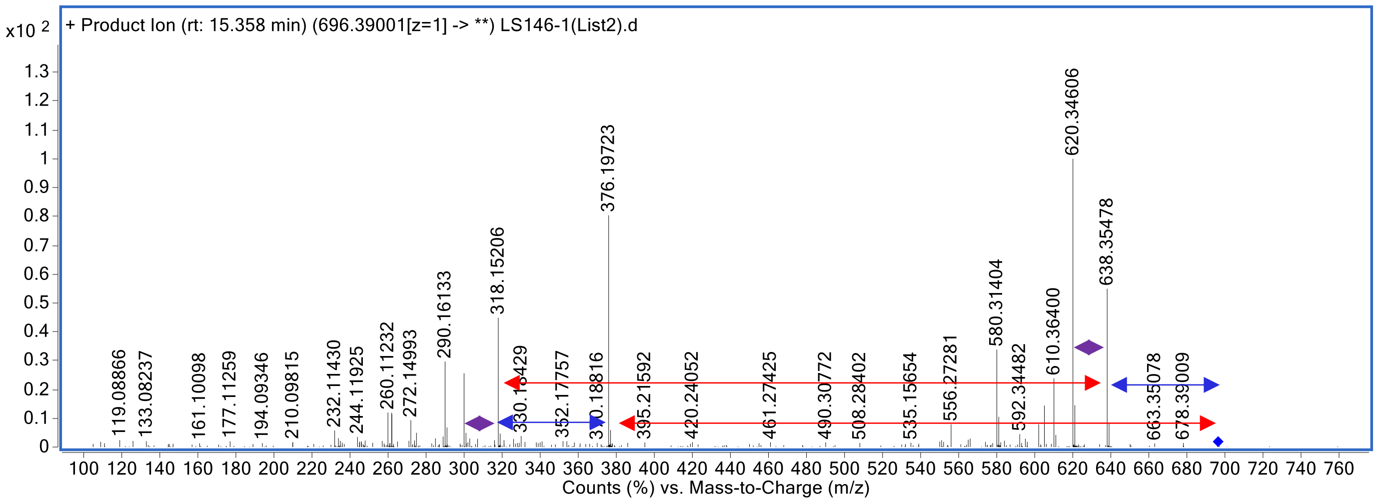
DH-NAA1


H
O
H
N
H
O
OH
O
OH
O

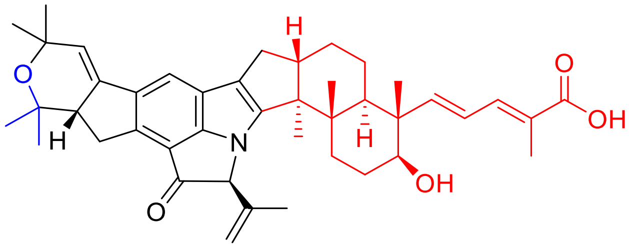

x10 2
+ Product Ion (rt: 22.297 min) (680.39508[z=1] -> **) LS146-1(List2).d
107.09129
121.10217
134.08839
159.11472
177.51131
200.13935
218.11156
235.17333
259.14435
274.15799
302.15392
309.18903
332.19817
345.18660
360.19236
374.20637
403.24524
420.32567
443.06564
495.27618
515.44775
562.29897
592.29567
608.35222
622.35727
651.35196
665.87556
680.40801
696.94930
753.25530
Counts (%) vs. Mass-to-Charge (m/z)

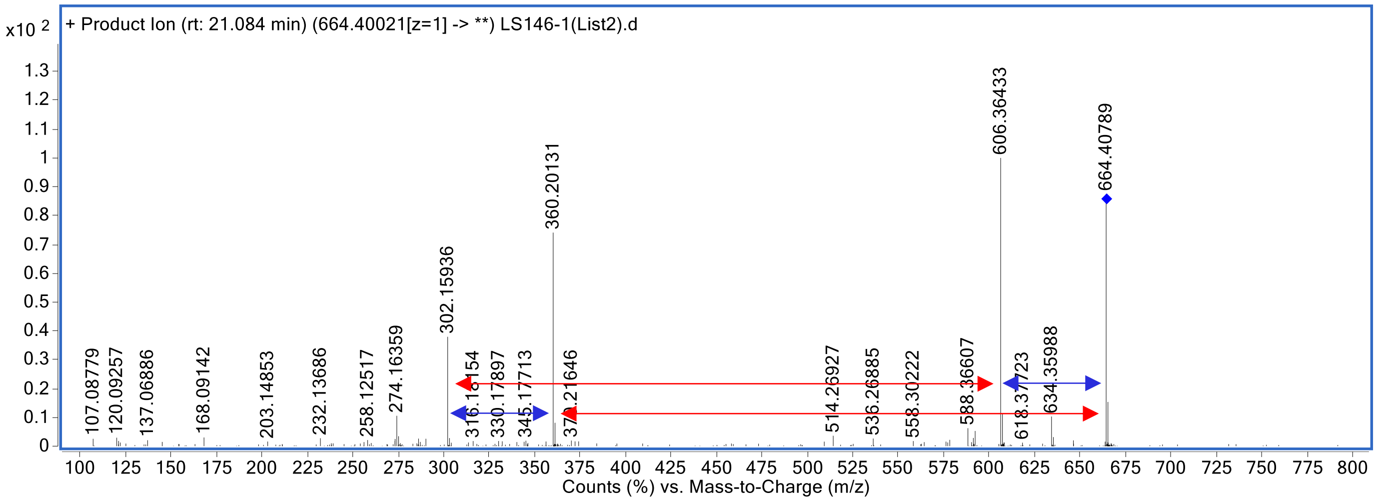
NAA1


H
O
H
HO
N
H
O
OH
O
OH
O

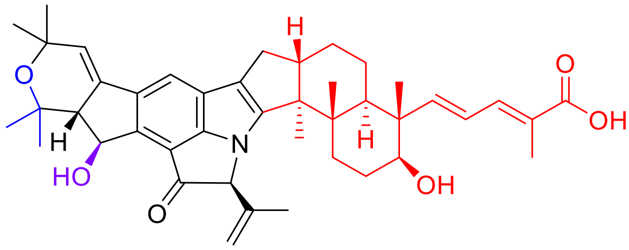

x10 2
+ Product Ion (rt: 15.358 min) (696.39001[z=1] -> **) LS146-1(List2).d
119.08866
133.08237
161.10098
177.11259
194.09346
210.09815
232.11430
244.11925
260.11232
272.14993
290.16133
318.15206
330.15429
352.17757
370.18816
376.19723
395.21592
420.24052
461.27425
490.30772
508.28402
535.15654
556.27281
580.31404
592.34482
610.36400
620.34606
638.35478
663.35078
678.39009
Counts (%) vs. Mass-to-Charge (m/z)

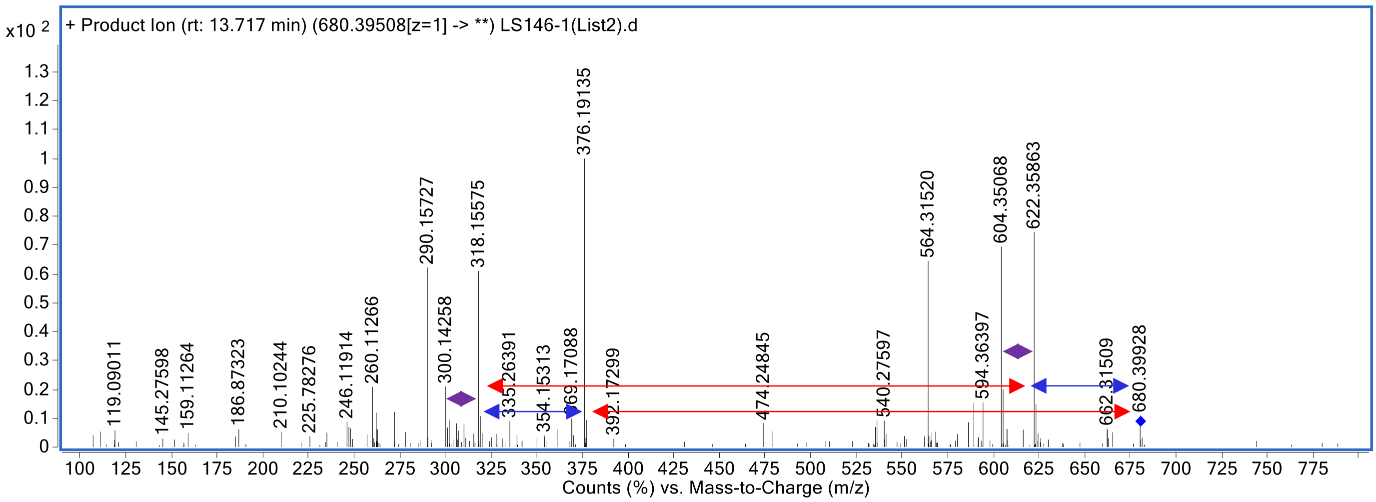
DH-NAA


O
H
N
O
H
H
O
OH
OH

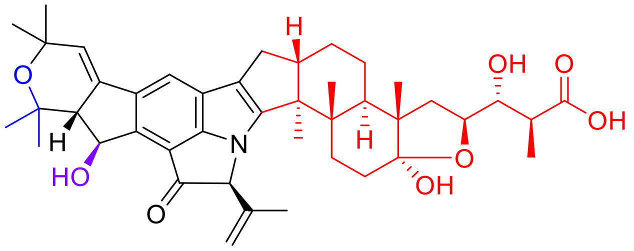

x10 2
+ Product Ion (rt: 21.084 min) (664.40021[z=1] -> **) LS146-1(List2).d
107.08779
120.09257
137.06886
168.09142
203.14853
232.13686
258.12517
274.16359
302.15936
330.17897
360.20131
514.26927
536.26885
558.30222
588.36607
606.36433
634.35988
664.40789
Counts (%) vs. Mass-to-Charge (m/z)

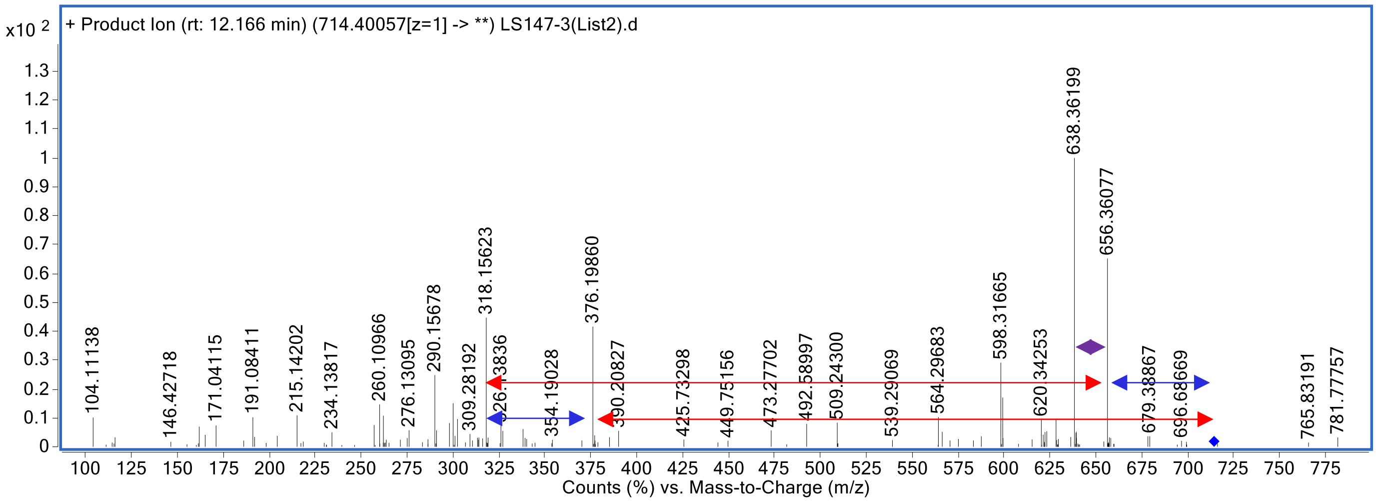
NAA


O
H
HO
N
O
H
H
O
OH
OH


x10 2
+ Product Ion (rt: 13.717 min) (680.39508[z=1] -> **) LS146-1(List2).d
119.09011
145.27598
159.11264
186.87323
210.10244
225.78276
246.11914
260.11266
290.15727
300.14258
318.15575
376.19135
474.24845
540.27597
564.31520
594.36397
604.35068
622.35863
680.39928
Counts (%) vs. Mass-to-Charge (m/z)

NAA2


H
O
H
HO
N
O
H
OH
O
OH
O
OH


x10 2
+ Product Ion (rt: 12.166 min) (714.40057[z=1] -> **) LS147-3(List2).d
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
104.11138
146.42718
171.04115
191.08411
215.14202
234.13817
260.10966
276.13095
290.15678
309.28192
318.15623
326.13836
354.19028
376.19860
390.20827
425.73298
449.75156
473.27702
492.58997
509.24300
539.29069
564.29683
598.31665
620.34253
638.36199
656.36077
679.38867
696.68669
765.83191
781.77757
100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675 700 725 750 775
Counts (%) vs. Mass-to-Charge (m/z)

Isolation of Nodulisporic Acids

Isolation of nodulisporic acids was carried out following procedures adapted from those described in the literature.[3] To provide sufficient structural evidence for the nodulisporic acids identified, an example of each major series (main-series, 1-series, 2-series, 4-series) and each biosynthetic step (D, C, B, and A) was isolated and characterised. These metabolites were either known compounds or closely related to those previously reported allowing for assignment of NMR data to be confirmed by comparison to the literature.[3] Analysis of these purified nodulisporic acids by HR-ESI-MS^2 provided further evidence for their identity and allowed for related intermediates to identified as described above (pg 43).

Isolation of NAD_4

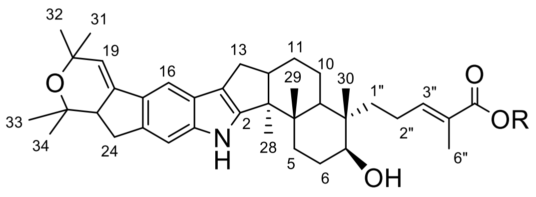
Mycelia from an NAD_4 producing strain (230 g) was extracted with MeOH (2 × 250 mL). The second, then first extracts were passed through a column of HP-20 (80 mL, pre-equilibrated in MeOH). The combined eluents were diluted with H_2O (500 mL) and recycled back through the same column. The loaded column was washed with H_2O (250 mL) and eluted with 250 mL portions of i) 60% acetone/ H_2O (fraction A), ii) 80% acetone/ H_2O (fraction B), and iii) acetone (fraction C). One-third (approx. 200 mg) of fraction B (606.7 mg) was chromatographed on silica gel (40 g, eluting with 3 CV each of 9:1 PE/DCM, 20:1 DCM/EtOAc, EtOAc, 5:1 EtOAc/MeOH, MeOH), from which fractions eluting from EtOAc afforded two samples (13.0 mg and 2.4 mg) enriched in NAD_4.

NMR data see Table S4; HR-ESI-MS see Table S3.

NAD_4-OMe

A sample enriched in NAD_4 (13 mg) was dissolved in 2 mL 1:1 MeOH/DCM, and cooled to 0 °C. To this, 300 µL $TMSCHN_2$ (2 M in Et_2O) was added and stirred for at 0 °C for 1 h. The reaction mixture was quenched with 2% $AcOH_{(aq)}$ and warmed to room temperature over 30 min. The resulting DCM layer was separated by microextraction, dried *in vacuo* and purified by semi-preparative HPLC (C18, 100% MeCN, t_R = 9.8 min; then 90% $MeCN/H_2O$, t_R = 17.2–18.5 min) to produce NAD_4-OMe (2.2 mg). Evidence of *O*-methylation at the C-5'' carboxyl terminus was confirmed by HR-ESI-MS^2 and 2D-NMR experiments.

NMR data see Table S4; HR-ESI-MS see Table S3



R = H = NAD_4
R = Me = NAD_4-OMe

Table S4. NMR Data for NAD_4 and NAD_4-OMe, $CDCl_3$

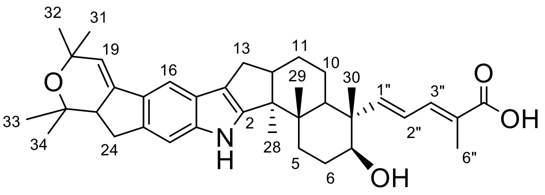
	NAD_4					NAD_4-OMe	
Position	^{13}C	^{1}H (int., mult., *J*)	COSY	HMBC	NOESY[a]	^{13}C	^{1}H (int., mult., *J*)
NH-1	-	7.59 (1H, s)	-	2, 14, 15, 27	28(w)	-	7.59 (1H, s)
2	151.2	-	-	-	-	151.2	-
3	53.4	-	-	-	-	53.5	-
4	39.4	-	-	-	-	39.4	-
5	33.6	1.55 (1H, m)	5b, 6	-	5b	33.6	1.56[c] (1H, m)
		1.90 (1H, m)	5a, 6, 7(w)	-	5a, 28		1.91[c] (1H, m)
6	27.9	1.84 (2H, m)	5a, 5b, 7	-	7, 29, 30	27.9	1.84 (2H, m)
7	73.5	3.56 (1H, dd, *J* = 6.9, 6.9 Hz)	5b(w), 7	30	6, 9, 2"b	73.5	3.55 (1H, appt t, *J* = 7.4 Hz)
8	41.4	-	-	-	-	41.4	-
9	40.2	1.70 (1H, m)	10a, 10b	-	7, 28	40.2	1.71 (1H, m)
10	22.9	1.46 (1H, m)	9, 10b	-	-	22.9[b]	1.45[c] (1H, m)
		1.61 (1H, m)	9, 10a	-	-		1.60[c] (1H, m)
11	25.2	1.61 (1H, m)	11b, 12	-	-	25.2	1.59[c] (1H, m)
		1.80 (1H, m)	11a, 12	-	-		1.79[c] (1H, m)
12	48.6	2.74 (1H, m)	11a, 11b, 13a, 13b	-	29	48.7	2.74 (1H, m)
13	27.6	2.31 (1H, dd, *J* = 13.3, 10.7 Hz)	12, 13b	2(w), 12(w), 14	28	27.6	2.30[c] (1H, dd, *J* =12.9, 10.4 Hz)
		2.66 (1H, dd, *J* =13.4, 6.5 Hz)	12, 13a	2, 3, 12, 14	-		2.66[c] (1H, dd, *J* =13.3, 6.4 Hz)
14	118.6	-	-	-	-	118.7	-
15	125.0	-	-	-	-	124.8	-
16	109.8	7.49 (1H, s)	26	14, 17, 25, 27	-	109.8	7.49 (1H, s)
17	139.2	-	-	-	-	139.1[d]	-
18	132.8	-	-	-	-	132.6[d]	-
19	119.8	5.95 (1H, d, *J* =3.0 Hz)	23	18, 20, 23	31, 32	119.6[c]	5.95 (1H, d, *J* = 3.0 Hz)
20	72.8	-	-	-	-	72.8	-
O-21	-	-	-	-	-		-
22	74.6	-	-	-	-	74.6	-
23	48.7	2.92 (1H, ddd, *J* =10.7, 7.8, 3.0 Hz)	19, 24a, 24b	17, 22	33	48.6	2.93 (1H, m)
24	33.5	2.69 (1H, dd, *J* =16.0, 8.4 Hz)	23, 24b, 26	22, 23(w), 25	24b, 34	33.5	2.68[c] (1H, dd, *J* =16.2, 8.0 Hz)
		3.13 (1H, dd, *J* = 16.0, 9.3 Hz)	23, 24a, 26	17, 18, 23, 25, 26(w)	24a		3.13[c] (1H, dd, *J* =16.0, 9.1 Hz)
25	138.2	-	-	-	-	138.0[d]	-
26	107.6	7.14 (1H, s)	16, 24a, 24b	15, 18, 24	-	107.6	7.13 (1H, s)
27	141.4	-	-	-	-	141.5	-
28	14.6	0.99 (3H, s)	-	2, 3, 4, 12	NH-1(w), 5b, 9, 13a	14.6	0.99 (3H, s)
29	19.4	1.12 (3H, s)	-	3, 4, 5, 9	6, 12	19.3	1.12 (3H, s)
30	16.4	0.85 (3H, s)	-	7, 8, 9, 1"	6, 1“a	16.4	0.85 (3H, s)
31	30.1[b]	1.35[b] (3H, s)	-	19, 20, 32	19	30.1[b]	1.35[b] (3H, s)
32	30.2[b]	1.33[b] (3H, s)	-	19, 20, 31	19	30.2[b]	1.34[b] (3H, s)
33	32.1	1.34 (3H, s)	34	22, 23, 34	23	32.1	1.33 (3H, s)
34	22.3	1.09 (3H, s)	33	22, 23, 33	24a	22.3[b]	1.09 (3H, s)
1"	36.0	1.41 (1H, m)	1”b, 2”a, 2”b	-	30, 3”(w)	36.1	1.41 (1H, m)
		1.70 (1H, m)	1”a, 2”a, 2”b	-	3”(w)		1.69 (1H, m)
2"	22.7	2.10 (1H, m)	1“a, 1“b, 2“b, 3"	-	-	22.5[b]	2.06[c] (1H, m)
		2.20 (1H, m)	1“a, 1“b, 2“a, 3"	-	7		2.16[c] (1H, m)
3"	144.9	6.90 (1H, tq, *J* = 7.4, 1.6 Hz)	2"a, 2"b, 6"	5", 6"	1“a(w), 1“b(w)	142.7	6.78 (1H, tq, *J* = 7.4, 1.3 Hz)
4"	126.8		-	-	-	127.6	-
5"	170.9		-	-	-	168.7[d]	-
6"	12.3	1.88 (3H ,s)	3"	3", 4", 5"	-	12.5	1.87 (3H ,s)
OMe	-	-	-	-	-	51.9	3.74 (3H, s)

[a]Mixing time = 0.5 s. Selected correlations. [b]Interchangeable within column. [c]obtained from HSQC experiment. [d]obtained from HMBC experiment

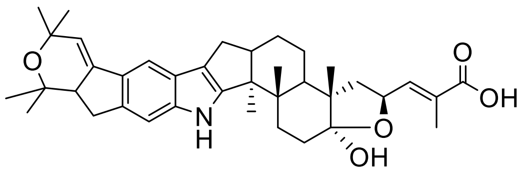
Isolation of NAD & NAD_1

Mycelia from an NAD producing strain (320 g wet weight approx.) was extracted with EtOAc (500 mL approx.) overnight, filtered, and extracted for a second time overnight. The combined extracts were washed with brine and solvent removed in vacuo to give a crude brown oil (3.6 g) which was dissolved in DCM:MeOH (9:1) and fractionated using flash chromatography. Fraction B (15 mg) was further purified by semi-preparative HPLC (column#2, 80% B, isocratic, 3.0 mL/min) which afforded NAD (0.8 mg) as a white solid. Fractions D and E were combined and purified by semi-preparative HPLC using the same conditions as described for NAD, this yielded NAD_1 (1.1 mg) as a white powder.

NMR Data see Tables S5 and S6, HR-ESI-MS Data see Table S3



NAD



NAD_1

Table S5: NMR Data for NAD, $CDCl_3$.

Position	^{13}C*	^{1}H	COSY	HMBC
1 (N)	-	7.60 (1H, s)	-	27
2	150.7	-	-	-
3	53.2	-	-	-
4	38.9	-	-	-
5a	33.2	n.o.	-	-
5b		n.o.	-	-
6a	25.7	1.86 (1H, m)	7	-
6b		1.86 (1H, m)	7	-
7	76.6	3.44 (1H, dd, J = 11.3, 4.6 Hz)	6a, 6b	-
8	47.2	-	-	-
9	44.5	1.68 (1H, m)	-	-
10a		n.o.	-	-
10b		n.o.	-	-
11a		n.o.	-	-
11b		n.o.	-	-
12	48.7	2.73 (1H, m)	-	13a, 13b
13a	27.3	2.65 (1H, dd, J = 13.3, 6.4 Hz)	13b	2, 3, 12, 14
13b		2.30 (1H, dd, J = 13.3, 10.6 Hz)	13a	12
14	118.5	-	-	-
15	124.8	-	-	-
16	109.6	7.49 (1H, s)	-	14, 17, 18, 27
17	138.0	-	-	-
18	139.2	-	-	-
19	119.7	5.95 (1H, d, J = 3 Hz)	23	-
20	72.7	-	-	-
21 (O)		-	-	-
22	74.4	-	-	-
23	48.5	2.93 (1H, m)	19, 24a, 24b	-
24a	33.2	3.13 (1H, dd, J = 15.9, 9.3 Hz)	23, 24b	-
24b		2.69 (1H, dd, J = 15.9, 7.8 Hz)	23, 24a	-
25	132.7	-	-	-
26	107.5	7.14 (1H, s)	-	25, 15
27	141.5	-	-	-
28	14.4	1.01 (3H, s)	-	2, 3, 4, 12
29	19.2	1.16 (3H, s)	-	3, 9, 4, 5
30	10.9	1.09 (3H, s)	-	1", 7, 8, 9
31	29.9	1.35 (3H, s)	-	19, 20, 32
32	31.8	1.33 (3H, s)	-	19, 20, 31
33	22.1	1.09 (3H, s)	-	22, 23, 34
34	30.9	1.34 (3H, s)	-	22, 23, 33
1"	153.4	5.92 (1H, d, J = 15.4 Hz)	2"	-
2"	125.6	6.40 (1H, dd, J = 15.4, 11.2 Hz)	1", 3"	-
3"	139.9	7.32 (1H, d, J = 11.2 Hz)	2"	-
4"	124.8	-	-	-
5"	169.9	-	-	-
6"	12.7	1.98 (3H, s)	-	3", 4", 5"

n.o. = not observed, signal obscured or not resolved. *Chemical shift obtained from HSQC and HMBC experiments

Table S6: NMR Data for NAD_1, $CDCl_3$.

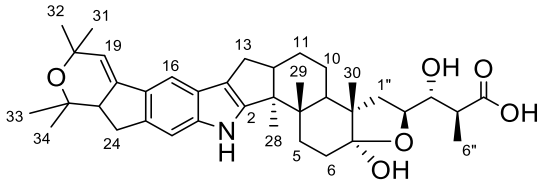
Position	^{13}C	^{1}H	COSY	HMBC
1 (N)	-	7.61 (1H, s)	-	-
2	150.9	-	-	-
3	52.7	-	-	-
4	39.4	-	-	-
5a	30.1	2.04 (1H, m)	5b	-
5b		1.94 (1H, m)	5a	-
6a	30.6	2.07 (1H, m)	-	-
6b		2.10 (1H, m)	-	-
7	106.5	-	-	-
8	49.2	-	-	-
9	40.4	n.o.	-	-
10a	25.7	1.83 (1H, m)	11b, 10b	-
10b		1.76 (1H, m)	11b, 10a	-
11a	24.9	1.66 (1H, m)	12, 11b	-
11b		1.50 (1H, m)	11a, 10a, 10b	-
12	48.7	2.80 (1H, m)	13a, 13b, 11a	13
13a	27.5	2.33 (1H, m)	12, 13b	2, 14
13b		2.68 (1H, m)	12, 13a	2, 3, 12, 14
14	118.3	-	-	-
15	124.9	-	-	-
16	109.7	7.50 (1H, s)	-	26, 27, 17, 18, 14, 15
17	138.2	-	-	-
18	139.1	-	-	-
19	119.7	5.95 (1H, d, J = 3.0 Hz)	23	23, 25, 20, 31
20	72.6	-	-	-
21 (O)	-	-	-	-
22	74.5	-	-	-
23	48.9	2.93 (1H, ddd, J = 9.2, 8.6, 3.0 Hz)	19, 24a, 24b	18, 19, 22, 24, 33, 34
24a	33.3	2.68 (1H, m)	23, 24b, 26	17, 23, 22
24b		3.13 (1H, dd, J = 15.9, 9.2 Hz)	23, 24a, 26	17, 18, 23, 25, 26
25	132.7	-	-	-
26	107.8	7.14 (1H, s)	24a, 24b	25, 24, 15
27	141.3	-	-	-
28	14.5	0.99 (3H, s)	-	12, 2, 3, 4
29	17.6	1.15 (3H, s)	-	3, 4, 5, 9
30	16.4	1.08 (3H, s)	-	7, 1", 8, 9
31	29.9	1.35 (3H, s)	-	20, 32
32	31.8	1.34 (3H, s)	-	20, 31
33	22.0	1.09 (3H, s)	-	22, 23, 34
34	30.1	1.34 (3H, s)	-	22, 23, 33
1"a	44.1	1.77 (1H, m)	2”, 1”b	-
1"b		2.33 (1H, m)	2”, 1”a	7
2"	72.7	4.88 (1H, ddd, J = 8.1, 8.1, 8.1 Hz)	3”, 1”a, 1”b	-
3"	146.3	6.95 (1H, d, J = 8.1 Hz)	2”	-
4"	126.8	-	-	-
5"	170.9	-	-	-
6"	12.7	1.89 (3H, s)	-	3", 4", 5"

n.o. = not observed, signal obscured or not resolved.

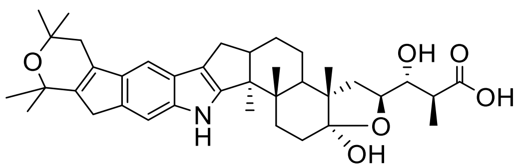
Isolation of NAD_2

Mycelia from an NAD_2 producing strain (380 g wet weight approx.) was extracted with EtOAc (750 mL approx.) overnight, filtered, and extracted for a second time overnight. The combined extracts were washed with brine and solvent removed in vacuo to give a dark brown oil (3.6 g) which was dissolved in DCM: MeOH (9:1) and fractionated using flash chromatography. Fractions containing NAD_2 were identified by LC-MS and TLC, combined (60 mg approx.) and purified by semi-preparative HPLC (column#1, 90% B, isocratic, 3.5 mL/min). Collections of the major peak were freeze dried to give a white powder (4.9 mg). Analysis of this sample by NMR spectroscopy showed the major product contained an eastern hemisphere consistent with that of a 2-series nodulisporic acid. The characteristic H-19 signal was not present in the spectrum and further analysis identified the metabolite as an isomerised product of NAD_2 containing an 18-23 alkene instead of the expected 18-19-alkene. The migration of this double bond has been reported in previous isolations of nodulisporic acids and is a product of degradation during workup rather than an alternative biosynthetic product.[3a]

NMR Data see Table S7. HR-ESI-MS see Table S3.



NAD_2



18-23-alkene-NAD_2

Table S7: NMR Data for 18-23 alkene-NAD_2, $CDCl_3$.

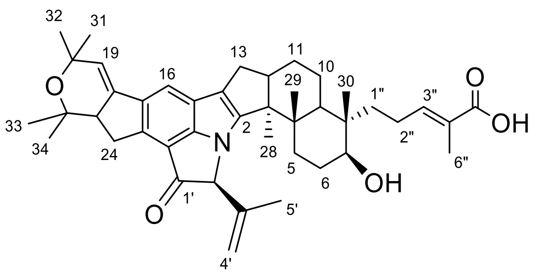
Position	^{13}C*	^{1}H	COSY	HMBC
1 (N)	-	7.66 (1H, s)	26	14, 2, 15, 17
2	149.6	-	-	-
3	52.3	-	-	-
4	39.5	-	-	-
5a	31.6	1.58 (1H, m)	5b	-
5b		2.20 (1H, m)	5a	-
6	n.o.	n.o.	-	-
7	112.9	-	-	-
8	48.8	-	-	-
9	46.5	1.80 (1H, m)	10a, 10b	3, 8, 10, 30, 1"
10a	23.2	1.49 (1H, m)	9, 10b, 11a, 11b	8, 9
10b		1.41 (1H, m)	9, 10a, 11a, 11b	-
11a	25.3	1.79 (1H, m)	11b, 12, 10a, 10b	-
11b		1.59 (1H, m)	11a, 12, 10a, 10b	3, 12
12	49.0	2.77 (1H, m)	13a, 13b, 11a, 11b	-
13a	27.5	2.34 (1H, m)	13b, 12	2, 3, 11, 12, 14
13b		2.69 (1H, m)	13a, 12	2, 3, 12, 14
14	118.2	-	-	-
15	123.7	-	-	-
16	107.4	7.35 (1H, s)	26, 19	26, 15, 17, 19, 27
17	138.3	-	-	-
18	141.7	-	-	-
19	36.4	3.37 (2H, m)	16, 24	18, 17, 25, 23
20	73.5	-	-	-
21 (O)	-	-	-	-
22	71.4	-	-	-
23	136.2	-	-	-
24	34.3	2.45 (2H, m)	19	18, 17, 25
25	131.8	-	-	-
26	107.2	7.22 (1H, s)	1, 16	16, 14, 15, 17, 25, 23
27	143.0	-	-	-
28	14.4	1.00 (3H, s)	-	2, 3, 4, 12
29	17.4	1.10 (3H, s)	-	3, 4, 5, 9
30	16.4	1.12 (3H, s)	-	7, 8, 9, 1"
31/32	30.4	1.42 (6H, s)	-	18, 20, 31/32
33/34	29.6	1.32 (6H, s)	-	22, 24, 33/34
1"a	40.0	1.74 (1H, m)	1"b, 2"	2", 3", 8, 9
1"b		1.49 (1H, m)	1"a	3", 7, 8, 30, 9
2"	78.4	4.60 (1H, dd, J = 5.5, 3.1 Hz)	3", 1"a, 6"	3", 8, 7
3"	80.6	4.03 (1H, m)	2", 6", 4"	5", 2", 4"
4"	40.2	2.70 (1H, m)	3", 6"	5", 6", 3"
5"	176.8	-	-	-
6"	13.2	1.13 (1H, m)	4", 2", 3"	3", 4", 5"

n.o. = not observed, signal obscured or not resolved. *Chemical shift obtained from HSQC and HMBC experiments

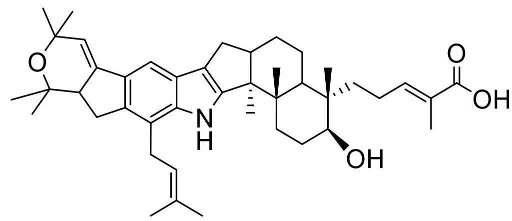
Isolation of DH-NAA$_4$ & DH-NAC$_4$

Mycelia from a DH-NAA$_4$ producing strain (260 g wet weight approx.) was extracted with EtOAc (500 mL approx.) overnight, filtered, and extracted for a second time overnight. The combined extracts were washed with brine and solvent removed in vacuo to give a buttery yellow solid (2.1 g) which was dissolved in DCM: MeOH (9:1) and fractionated using flash chromatography. Fractions containing the target nodulisporic acids were identified by LC-MS and TLC, combined (250 mg) and purified by semi-preparative HPLC (column#1, 95% B, isocratic, 3.5 mL/min). Collections of the two major peaks were freeze dried and analysed by NMR identifying them as DH-NAA$_4$ (0.2 mg) and DH-NAC$_4$ (0.1 mg).

NMR Data see Tables S8 and S9, HR-ESI-MS see Table S3



DH-NAA$_4$



DH-NAC$_4$

Table S8: NMR Data for DH-NAA$_4$, $CDCl_3$.

Position	^{13}C*	^1H	COSY	HMBC
1 (N)	-	-	-	-
2	153	-	-	-
3	55.8	-	-	-
4	39.1	-	-	-
5a	32.2	1.86 (1H, m)	5b	-
5b		1.68 (1H, m)	5a	-
6a	27.5	1.70 (1H, m)	6b, 7	-
6b		1.75 (1H, m)	6a, 7	-
7	73.4	3.53 (1H, dd, J = 9.6, 5.8 Hz)	6a, 6b	8
8	46.9	-	-	-
9	41.3	1.62 (1H, m)	-	-
10	n.o.	n.o.	-	-
11	25.3	n.o.	-	-
12	47.4	2.84 (1H, m)	13a	-
13a	27.6	2.30 (1H, dd, J = 13.8, 10.7 Hz)	13b, 12	2
13b		2.74 (1H, dd, J = 13.8, 6.5 Hz)	13a	2, 3, 15
14	121.7	-	-	-
15	122	-	-	-
16	116.3	7.70 (1H, s)	-	25, 27, 18, 14
17	n.o.	-	-	-
18	136.1	-	-	-
19	120.8	6.01 (1H, d, J = 3.0 Hz)	23	18, 20
20	72.6	-	-	-
21 (O)	-	-	-	-
22	74.4	-	-	-
23	49.2	3.00 (1H, ddd, J = 9.3, 7.7, 3.0 Hz)	19, 24a, 24b	-
24a	31.8	2.8 (1H, dd, J = 17.6, 7.7 Hz)	24b, 23	18, 22, 26
24b		3.41 (1H, dd, J = 17.6, 9.3 Hz)	24a, 23	18
25	138.5	-	-	-
26	113.3	-	-	-
27	161.6	-	-	-
28	15	0.91 (3H, s)	-	2, 3, 4, 12
29	19.4	1.10 (3H, s)	-	3, 4, 5, 9
30	16.6	0.84 (3H, s)	-	7, 9, 1"
31	29.9	1.36 (3H, s)	-	32, 19, 20
32	31.9	1.34 (3H, s)	-	31, 19, 20
33	22.2	1.08 (3H, s)	-	23, 34, 22
34	30	1.36 (3H, s)	-	23, 33, 22
1'	197.3	-	-	-
2'	75.9	4.97 (1H, s)	-	27, 1', 3'
3'	140	-	-	-
4'a	122.3	4.93 (s, br)	-	-
4'b		5.15 (s, br)	-	-
5'	n.o.	n.o.	-	-
1"a	35.9	1.44 (1H, m)	-	8
1"b		1.71 (1H, m)	-	-
2"a	22.8	2.09 (1H, m)	3", 2"b	-
2"b		2.19 (1H, m)	3", 2"a	-
3"	144.9	6.89 (1H, dd, J = 6.4, 6.4 Hz)	2"a, 2"b, 6"	-
4"	126.4	-	-	-
5"	169.6	-	-	-
6"	12.2	1.87 (3H, s)	3"	4", 5"

n.o. = not observed, signal obscured or not resolved. *Chemical shift obtained from HSQC and HMBC experiments

Table S9: NMR Data of DH-NAC4, $CDCl_3$.

Position	^{13}C*	^{1}H	COSY	HMBC
1 (N)	-	7.69 (1H, brs)	-	-
2	150.7	-	-	-
3	53.0	-	-	-
4	39.2	-	-	-
5	33.4	n.o.	-	-
6	n.o.	n.o.	-	-
7	73.4	3.51 (1H, m)	-	-
8	35.8	-	-	-
9	40.0	1.68 (1H, m)	-	-
10	n.o.	n.o.	-	-
11	n.o.	1.61 (2H, m)	12	-
12	48.7	2.71 (1H, m)	13a, 13b	-
13a	n.o.	2.30 (1H, dd, J = 13.2, 10.7 Hz)	12, 13b	-
13b		2.65 (1H, dd, J = 13.2, 6.4 Hz)	12, 13a	-
14	n.o.	-	-	-
15	n.o.	-	-	-
16	107.8	7.39 (1H, s)	-	-
17	n.o.	-	-	-
18	n.o.	-	-	-
19	119.5	5.93 (1H, d, J = 3.0 Hz)	23	-
20	72.7	-	-	-
21 (O)	-	-	-	-
22	74.4	-	-	-
23	48.5	2.92 (1H, ddd, J = 9.4, 7.5, 3.0 Hz)	19, 24a, 24b	-
24a	31.6	3.13 (1H, dd, J = 15.8, 9.4 Hz)	23, 24b	-
24b		2.57 (1H, dd, J = 15.8, 7.5 Hz)	23, 24a	-
25	n.o.	-	-	-
26	n.o.	-	-	-
27	n.o.	-	-	-
28	14.5	0.99 (3H, s)	-	2, 3, 4, 12
29	19.2	1.11 (3H, s)	-	3, 4, 5, 9
30	16.3	0.86 (3H, s)	-	7, 8, 9, 1"
31	30.1	1.35 (3H, s)	-	19, 20, 32
32	31.9	1.33 (3H, s)	-	19, 20, 31
33	30.0	1.35 (3H, s)	-	34, 22
34	22.1	1.09 (3H, s)	-	22, 23, 33
1'a	28.2	3.55 (1H, m)	5', 1'b	-
1'b		3.51 (1H, m)	1'a, 2', 4'	-
2'	122.5	5.35 (1H, m)	4', 5', 1'b	-
3'	132.7	-	-	-
4'	25.6	1.76 (3H, s)	2'	2', 3', 5'
5'	18.1	1.86 (3H, s)	1'a, 2'	2', 3', 4'
1"a	41.2	1.71 (1H, m)	1"b, 2"a, 2"b	-
1"b		1.45 (1H, m)	1"a, 2"a, 2"b	-
2"a	n.o.	2.19 (1H, m)	1"a, 1"b, 2"b, 3"	-
2"b		2.10 (1H, m)	1"a, 1"b, 2"a, 3"	-
3"	144.7	6.89 (1H, dd, J = 6.4, 6.4 Hz)	2"a, 2"b, 6"	-
4"	126.2	-	-	-
5"	168.9	-	-	-
6"	12.2	1.88 (3H, s)	3"	3", 4", 5"

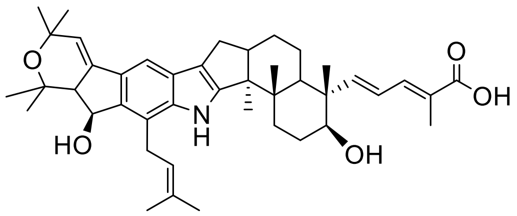
n.o. = not observed, signal obscured or not resolved. *Chemical shift obtained from HSQC and HMBC experiments

Isolation of NAC

Mycelia from an NAC producing strain (120 g) was extracted with MeOH (2 × 250 mL). The second, then first extracts were passed through a bed of HP-20 beads (80 mL, pre-equilibrated in MeOH). The combined eluents were diluted with 750 mL of H_2O to a final concentration of 40% MeOH/ H_2O and cycled back through the same HP-20 column. The loaded beads were washed with H_2O (250 mL) and then eluted with 250 mL portions of i) 50% acetone/ H_2O (fraction A), ii) 80% acetone/ H_2O (fraction B) and iii) acetone (fraction C). Fractions B and C, verified by TLC and LC-MS analysis to contain nodulisporic acid compounds, were combined (667 mg) and chromatographed over silica gel (40 g, MeOH/$CHCl_3$, 1–100%, fractions D–H). Fraction E (5–6% MeOH/$CHCl_3$, 378 mg) was further chromatographed over silica gel (24 g, EtOAc/*n*-hexane, 0–100%, 17 CV; EtOAc, 7 CV, fractions I–L). The MeOH-soluble portion of fraction L (EtOAc, 32.4 mg) was purified using semipreparative HPLC (column#1, 85% acetonitrile/ H_2O, 3.3 mL/min) to give NAC (t_R = 5.9 min).

Fraction F (6–7% MeOH/$CHCl_3$) from the first silica column was purified by semipreparative HPLC (C18, 80% acetonitrile/H_2O, 3.3 mL/min) to afford NAC (t_R = 9.7 min) and combined with the earlier HPLC fraction to give a total yield of 0.6 mg. NMR spectra of our sample of NAC recorded in $CDCl_3$ were consistent with literature data.[3b]

NMR data see Table S10; HR-ESI-MS see Table S3



NAC

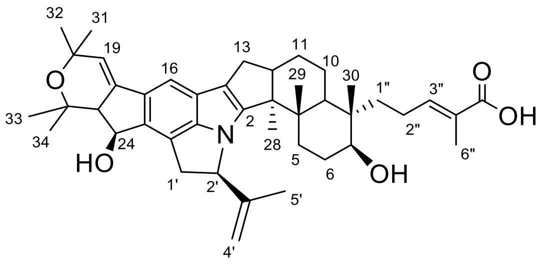
Table S10. NMR Data for NAC, $CDCl_3$.

Position	^{13}C*	^{1}H	COSY	HMBC
1 (N)		7.81 (1H, br s)	-	2, 14, 15, 27
2	152.0	-	-	-
3	53.1	-	-	-
4	39.1	-	-	-
5a	33.2	1.52 (1H, m)	5b, 6a	-
5b		1.99 (1H, m)	5a	-
6a	25.6	1.86 (1H, m)	5a, 6b, 7	-
6b		1.90 (1H, m)	6a, 7	-
7	76.5	3.44 (1H, dd, *J* = 10.8, 4.2 Hz)	6a, 6b	-
8	47.4	-	-	-
9	44.5	1.68 (1H, m)	-	-
10	24.8	1.46 (2H, m)	11a, 11b	-
11a	25.1	1.57 (1H, m)	10, 11b, 12(w)	-
11b		1.72 (1H, m)	10, 11a	-
12	48.9	2.72 (1H, m)	11a(w), 13a, 13b	-
13a	27.4	2.30 (1H, dd, *J* = 13.3, 10.8 Hz)	12, 13b	-
13b		2.65 (1H, dd, *J* = 13.3, 6.4 Hz)	12, 13a	2, 3, 14
14	118.7	-	-	-
15	126.7	-	-	-
16	107.4	7.40 (1H, s)	-	14, 15, 17, 25, 26, 27
17	135.4	-	-	-
18	131.9	-	-	-
19	120.2	5.96 (1H, d, *J* = 2.8 Hz)	23	18, 20, 23
20	72.7	-	-	-
21 (O)	-	-	-	-
22	73.9	-	-	-
23	60.1	2.69 (1H, dd, *J* = 4.6, 2.8 Hz)	19, 24	-
24	76.5	5.08 (1H, d, *J* = 4.6 Hz)	23	22, 25
25	136.1	-	-	-
26	121.9	-	-	-
27	141.1	-	-	-
28	14.4	1.02 (3H, s)	-	2, 3, 4, 12
29	19.1	1.14 (3H, s)	-	3, 4, 5, 9
30	11.0	1.09 (3H, s)	-	7, 8, 9, 1''
31	32.0	1.34 (3H, s)	-	19, 20, 32
32	29.85	1.32 (3H, s)	-	19, 20, 31
33	29.85	1.50 (3H, s)	-	22, 23, 34
34	22.6	1.07 (3H, s)	-	22, 23, 33
1'	27.0	3.79 (2H, m)	2'	-
2'	123.1	5.38 (1H, br t, *J* = 6.7 Hz)	1', 4', 5'	-
3'	133.0	-	-	-
4'	25.6	1.76 (3H, s)	2', 5'	2', 3', 5'
5'	18.1	1.90 (3H, s)	2', 4'	2', 3', 4'
1''	153.5	5.91 (1H, d, *J* = 15.3 Hz)	2''	-
2''	125.9	6.39 (1H, dd, *J* = 15.3, 11.5 Hz)	1'', 3''	-
3''	139.9	7.33 (1H, d, *J* = 11.5 Hz)	2'', 6''	-
4''	125.2	-	-	-
5''	171.1	-	-	-
6''	12.6	1.98 (3H, br s)	3''	3'', 4'', 5''

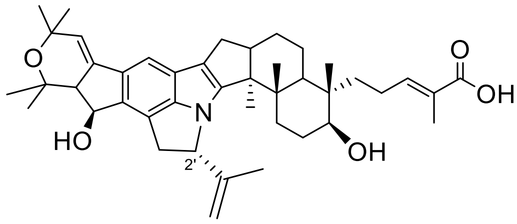
n.o. = not observed, signal obscured or not resolved. *Chemical shifts obtained from HSQC and HMBC experiments.

Isolation of NAB4

Mycelia from an NAB4 producing strain (240 g wet weight approx.) was extracted with EtOAc (500 mL approx.) overnight filtered and extracted for a second time overnight. The combined extracts were washed with brine and solvent removed in vacuo to give a buttery yellow solid (2.8 g) which was dissolved in DCM: MeOH (9:1) and fractionated using flash chromatography. Fractions containing indole diterpenes were identified by TLC, combined, and subjected to a second round of flash chromatography. Fractions containing NAB4 were identified by LC-MS and combined (200 mg). Semi-preparative HPLC (column#1, 95% B, isocratic, 3.5 mL/min) was used to purify the major peak which was freeze dried to give a white powder (12 mg). 2D NMR Showed that this sample contained two indole diterpene products which had signals consistent with 2' epimers of NAB4 which are known to be common breakdown products of nodulisporic acids.[4] Further purification of this sample led to degradation of the target metabolites.



NAB$_4$



2'-epi-NAB$_4$

Table S11: NMR Data for NAB4 and its 2'-epimer, CDCl3.

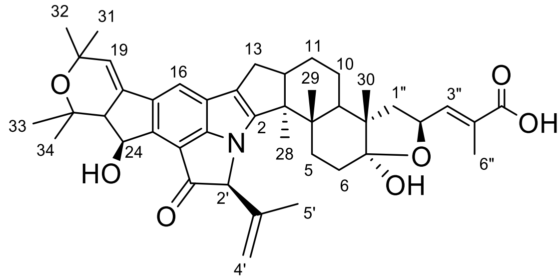
Position	^{13}C*	^{1}H	COSY	HMBC
Epimer 1				
1'a	n.o.	3.30 (1H, d, J = 15.8 Hz)	1'b, 2'	26, 3', 27
1'b		4.06 (1H, m)	1'a, 2'	2', 26, 3', 27
2'	68.6	5.20 (1H, d, J = 7.0 Hz)	1'a, 1'b	5', 4', 27
3'	146.3	-	-	-
4'	112.3	4.76 (m)	5'	16.7
5'	16.7	1.22 (3H, s)	4'	4', 3'
26	115.3	-	-	-
27	153.8	-	-	-
Epimer 2				
1'a	n.o.	3.47 (1H, d, J = 17.6 Hz)	1'b, 2'	26, 3', 27
1'b		4.04 (1H, m)	1'a, 2'	2', 26, 3', 27
2'	68.6	5.20 (1H, d, J = 7.0 Hz)	1'a, 1'b	5', 4', 27
3'	145.0	-	-	-
4'	113.6	4.85 (m)	5'	16.7
5'	16.7	1.29 (3H, s)	4'	4', 3'
26	123.9	-	-	-
27	152.5	-	-	-

n.o. = not observed, signal obscured or not resolved.
*Chemical shift obtained from HSQC and HMBC experiments

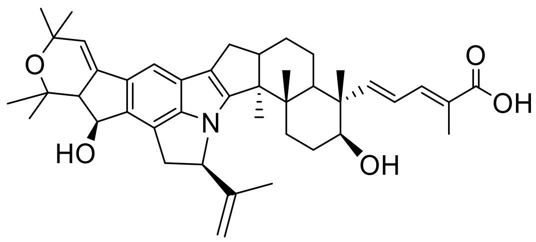
Isolation of NAA_1 and NAB

Mycelia from an NAA_1 producing strain (500 g wet weight approx.) was extracted with MeOH (1 L). The extract was passed through a bed of HP-20 beads (250 mL, pre-equilibrated in MeOH). The eluent was diluted with H_2O (1:1) and cycled back through the same HP-20 beads. The loaded beads were washed with 500 mL portions of i) H_2O, ii) 50% acetone/H_2O, iii) 80% acetone/H_2O, and iv) acetone. Target nodulisporic acids were identified by LC-MS in the 50% and 80% fractions which were subsequently combined and subjected to silica gel chromatography ($CHCl_3$: MeOH). Fractions containing NAA_1 were identified by LC-MS and the target purified using semi-preparative HPLC (Column#1, 80–90% B over 20 min, 3.5 mL/min) to give NAA_1 as a yellow powder (2.8 mg) and NAB as a white powder (0.1 mg).

NMR data see Table S12 and S13; HR-ESI-MS see Table S3.



NAA_1



NAB

Table S12: NMR Data for NAA_1, CD_3OD.

Position	^{13}C	^{1}H	COSY	HMBC
1 (N)	-	-	-	-
2	155.7	-	-	-
3	56.5	-	-	-
4	40.3	-	-	-
5a	31.1	1.73 (1H, m)	6, 5b	-
5b		1.91 (1H, m)	5a, 6	4
6	30.0	1.86 (2H, m)	5a, 5b	4
7	107.5	-	-	-
8	50.3	-	-	-
9	42.1	1.81 (1H, dd, J = 2.6, 12.8 Hz)	10a, 10b	29, 30
10a	25.7	1.58 (1H, m)	9, 10b, 11a, 11b	-
10b		1.72 (1H, m)	9, 10a, 11b	12, 9
11a	26.6	1.71 (1H, m)	12, 11b, 10a	12, 9
11b		1.88 (1H, m)	12, 11a, 10a, 10b	3
12	49.0	2.96 (1H, m)	13a, 13b, 11a, 11b	-
13a	28.2	2.37 (1H, m)	12, 13b	12, 2, 14
13b		2.77 (1H, dd, J = 13.8, 6.5 Hz)	12, 13a	12, 2, 14
14	123.0	-	-	-
15	123.1	-	-	-
16	117.0	7.76 (1H, s)	24	27, 26, 17, 25, 18, 14, 15
17	137.0	-	-	-
18	135.2	-	-	-
19	121.9	6.12 (1H, d, J = 3.0 Hz)	23	18, 20, 23, 31
20	74.1	-	-	-
21 (O)	-	-	-	-
22	75.7	-	-	-
23	60.2	2.76 (1H, m)	19, 24	18
24	75.4	5.16 (1H, d, J = 5.9 Hz)	23, 16	22, 23, 18, 25, 26
25	138.4	-	-	-
26	114.6	-	-	-
27	163.2	-	-	-
28	15.2	1.00 (3H, s)	-	12, 2, 3, 4
29	18.1	1.15 (3H, s)	-	2, 3, 4, 9, 5
30	17.1	1.05 (3H, s)	-	7, 9, 1", 8
31	29.9	1.36 (3H, s)	-	19, 20, 32
32	32.0	1.34 (3H, s)	-	19, 20, 31
33	23.2	1.13 (3H, s)	-	22, 23, 34
34	30.2	1.45 (3H, s)	-	33, 23, 22
1'	197.7	-	-	-
2'	77.2	5.22 (1H, s)	-	1', 3', 4', 5', 27
3'	141.6	-	-	-
4'a	117.4	5.18 (1H, m)	4'b, 5'	1', 3', 5'
4'b		5.00 (1H, m)	4'a	-
5'	18.3	1.48 (3H, s)	4'a	-
1"a	45.3	1.69 (1H, m)	2", 1"b	8, 9, 30, 2", 3"
1"b		2.35 (1H, m)	2", 1"a	8, 9, 7
2"	74.0	4.9 (1H, m)	3", 1"a, 1"b	7
3"	146.1	6.90 (1H, dq, J = 8.0, 1.4 Hz)	2", 6"	6", 5"
4"	128.7	-	-	-
5"	171.8	-	-	-
6"	12.7	1.83 (3H, d, J = 1.4 Hz)	3"	5", 3", 4"

n.o. = not observed, signal obscured or not resolved.

Table S13: NMR Data for NAB, $CDCl_3$.

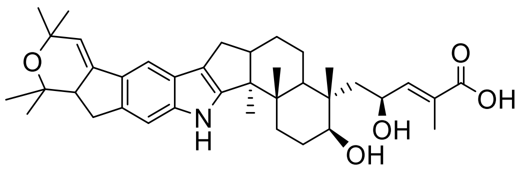
Position	^{13}C*	^{1}H	COSY	HMBC
1 (N)	-	-	-	-
2	151.3	-	-	-
3	55.3	-	-	-
4	38.5	-	-	-
5	31.8	n.o.	-	-
6	25.5	1.77 (2H, m)	7	-
7	76.6	3.41 (1H, m)	6	-
8	47.5	-	-	-
9	44.9	n.o.	-	-
10	n.o.	n.o.	-	-
11	n.o.	n.o.	-	-
12	47.0	2.73 (1H, m)	13a, 13b	-
13a	28.0	2.23 (1H, dd, J = 13.6, 2.8 Hz)	13b, 12	-
13b		2.65 (1H, dd, J = 13.6, 6.3 Hz)	13a, 12	3
14	n.o.	-	-	-
15	n.o.	-	-	-
16	107.4	7.31 (1H, s)	-	27, 17, 18
17	133.2	-	-	-
18	135.4	-	-	-
19	119.9	5.96 (1H, d, J = 2.8 Hz)	23	20
20	72.5	-	-	-
21 (O)	-	-	-	-
22	73.6	-	-	-
23	60.4	2.69 (1H, dd, J = 5.7, 2.8 Hz)	19, 24	-
24	75.9	4.98 (1H, d, J = 5.7 Hz)	23	-
25	n.o.	-	-	-
26	n.o.	-	-	-
27	154.6	-	-	-
28	15.0	0.91 (3H, s)	-	2, 3, 4, 12
29	19.3	1.09 (3H, s)	-	3, 4, 5, 9
30	11.0	1.06 (3H, s)	-	7, 8, 9, 1"
31	32.0	1.32 (3H, s)	-	32, 20
32	29.8	1.33 (3H, s)	-	31, 20, 19
33	23.1	1.12 (3H, s)	-	22, 23, 34
34	30.1	1.48 (3H, s)	-	22, 23, 33
1'a	40.6	3.50 (1H, s, J = 16.4 Hz)	1'b, 2'	-
1'b		4.07 (1H, dd, J = 16.4, 8.0 Hz)	1'a, 2'	-
2'	69.0	5.24 (1H, d, J = 8.0 Hz)	1'a, 1'b	-
3'		-	-	-
4'a	112.9	4.8 (1H, s br)	4'b, 5'	-
4'b		4.72 (1H, s br)	4'a	-
5'	17.0	1.25 (3H, s)	4'a	-
1"	153.8	5.91 (1H, d, J = 15.3 Hz)	2"	-
2"	125.7	6.38 (1H, dd, J = 15.3, 13.0 Hz)	1", 3"	-
3"	n.o.	7.31 (1H, m)	2", 6"	-
4"	n.o.	-	-	-
5"	n.o.	-	-	-
6"	12.7	1.98 (3H, s)	3"	-

n.o. = signal not observed or obscured. *Signals obtained from HSQC and HMBC experiments

Isolation of NAD_6

Mycelia from an NAD_6 producing strain (264 g wet weight) was extracted with EtOAc (500 mL approx.) overnight filtered and extracted for a second time overnight. The combined extracts were washed with brine and solvent removed in vacuo to give a brown solid (2.0 g) which was dissolved in $CHCl_3$ and fractionated using flash chromatography. Fractions containing NAD_6 were identified by TLC and LC-MS and combined (40 mg). Semi-preparative HPLC (column#1, 70% B, isocratic, 3.5 mL/min) was used to purify the major peak which was freeze dried to give NAD_6 as a white powder (1.6 mg).

NMR data see Table S14; HR-ESI-MS see Table S3.



NAD_6

Table S14: NMR Data for NAD_6, $(CD_3)_2CO$.

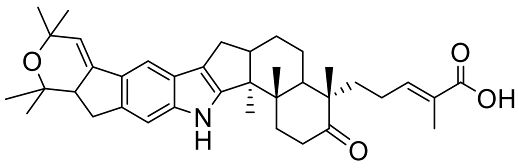
Position	^{13}C*	^1H	COSY	HMBC
1 (N)	-	9.67 (1H, s br)	-	-
2	152.4	-	-	-
3	54.2	-	-	-
4	40.3	-	-	-
5a	33.5	1.86 (1H, m)		
5b		1.82 (1H, m)		
6a	27.6	1.70 (1H, m)	7	
6b		1.85 (1H, m)	7	
7	74.4	3.90 (1H, dd, J = 10.8, 3.8 Hz)	6a, 6b	
8	42.0	-	-	-
9	42.9	n.o.		
10a	n.o.	1.51 (1H, m)		
10b		1.62 (1H, m)	11b	
11a	25.6	1.66 (1H, m)	12, 11b	
11b		1.77 (1H, m)	11a, 10b	
12	49.6	2.76 (1H, obscured)	11a, 13a, 13b	
13a	27.8	2.29 (1H, dd, J = 13.0, 10.9 Hz)	12, 13b	
13b		2.65 (1H, m)	12, 13a	14
14	118.2	-	-	-
15	n.o.	-	-	-
16	109.9	7.43 (1H, s)	26	14, 17, 18
17	138.2	-	-	-
18	143.0	-	-	-
19	119.3	5.98 (1H, d, J = 3.0 Hz)	23	
20	72.9	-	-	-
21 (O)	-	-	-	-
22	74.6	-	-	-
23	49.7	2.83 (1H, obscured)	19, 24a, 24b	
24a	33.6	2.65 (1H, m)	23, 24b, 26	
24b		3.09 (1H, dd, J = 15.8, 9.3 Hz)	23, 24a, 26	
25	132.7	-	-	-
26	108.2	7.18 (1H, s)	16, 24a, 24b	25
27	n.o.	-	-	-
28	14.6	1.04 (3H, s)		2, 3, 4, 12
29	19.5	1.14 (3H, s)		3, 4, 5, 9
30	16.7	0.88 (3H, s)		7, 8, 9, 1"
31	30.4	1.31 (3H, s)		19, 20, 32
32	32.1	1.26 (3H, s)		19, 20, 31
33	30.0	1.27 (3H, s)		22, 23, 34
34	22.2	1.05 (3H, s)		22, 23, 33
1"a	45.5	1.49 (1H, m)	1"b, 2"	
1"b		1.89 (1H, m)	1"a, 2"	
2"	65.3	4.77 (1H, td, J = 8.4, 4.6 Hz)	1"a, 1"b, 3"	
3"	145.8	6.75 (1H, d, J = 8.8 Hz)	2", 6"	
4"	126.6	-	-	-
5"	169.4	-	-	-
6"	12.5	1.89 (3H, s)	3", 2"	3", 4", 5"

n.o. = signal not observed or obscured. *Signals obtained from HSQC and HMBC experiments

Isolation of NAD_5

Mycelia from an NAD_5 producing strain (208 g wet weight) was extracted with EtOAc (500 mL approx.) overnight filtered and extracted for a second time overnight. The extracts were combined, and solvent removed in vacuo to give a brown-yellow oil (1.05 g) which was dissolved in $CHCl_3$ and fractionated using flash chromatography. Fractions containing NAD_5 were identified by TLC and LC-MS and combined (67 mg). Preparative HPLC (column#1, 85–100% B over 20 min, 5 mL/min) was used to purify NAD_5 which was freeze dried to give a white powder (<0.1 mg).

NMR data see Table S15; HR-ESI-MS see Table S3.



NAD_5

Table S15: NMR Data for NAD_5, $(CD_3)_2CO$

Position	^{13}C*	^1H	COSY**	HMBC
1 (N)	-	9.82 (1H, s)		
2	151.9	-	-	-
3	53.3	-	-	-
4	39.6	-	-	-
5a	33.2	2.11 (1H, m)	5b, 6a, 6b	
5b		2.25 (1H, dd, J = 15.0, 6.8 Hz)	5a, 6a, 6b	
6a	n.o.	2.47 (1H, ddd, J = 16.4, 6.8, 3.3 Hz)	5a, 5b, 6b	
6b		2.67 (1H, m)	5a, 5b, 6a	
7	214.9	-	-	-
8	51.8	-	-	-
9	43.0	n.o.	-	-
10	n.o.	n.o.	-	-
11	n.o.	n.o.	-	-
12	49.7	2.85 (obscured)	13a, 13b	
13a	n.o.	2.70 (1H, dd, J= 13.2, 6.0 Hz)	13b, 12	
13b		2.35 (1H, dd, J = 13.2, 10.7 Hz)	13a, 12	
14	n.o.	-	-	-
15	n.o.	-	-	-
16	109.9	7.45 (1H, s)		
17	n.o.	-	-	-
18	n.o.	-	-	-
19	119.4	5.99 (1H, d, J = 3 Hz)	23	
20	72.9	-	-	-
21 (O)	-	-	-	-
22	74.6	-	-	-
23	49.5	2.81 (obscured)	19, 24a, 24b	
24a	n.o.	2.66 (1H, dd, J = 16.0, 7.5 Hz)	23, 24b	
24b		3.09 (1H, dd, J = 16.0, 9.6 Hz)	23, 24a	
25	n.o.	-	-	-
26	108.2	7.18 (1H, s)		
27	n.o.	-	-	-
28	14.7	1.07 (3H, s)		2, 3, 4, 12
29	17.7	1.25 (3H, s)		3, 4, 5, 9
30	21.2	1.07 (3H, s)		7, 8, 9, 1"
31	30.3	1.31 (3H, s)		19, 20, 32
32	32.1	1.26 (3H, s)		19, 20, 31
33	30.1	1.27 (3H, s)		22, 23, 34
34	22.2	1.05 (3H, s)		22, 23, 34
1"a	38.0	1.87 (1H, m)	1"b, 2"	
1"b		1.51 (1H, m)	1"a, 2"	
2"	n.o.	2.05 (2H, obscured)	1"a, 1"b, 3"	
3"	142.7	6.76 (1H, td, J = 7.6, 1.7 Hz)	2", 6"	
4"	128.6	-	-	-
5"	169.4	-	-	-
6"	12.3	1.81 (3H, d, J = 1.7 Hz)	3"	3", 4", 5"

n.o. = signal not observed or obscured. *Signals obtained from HSQC and HMBC experiments
**Some correlations confirmed using selective 1D TOCSY (mixing time 50 ms)

NMR Spectra

NAD NMR Spectra

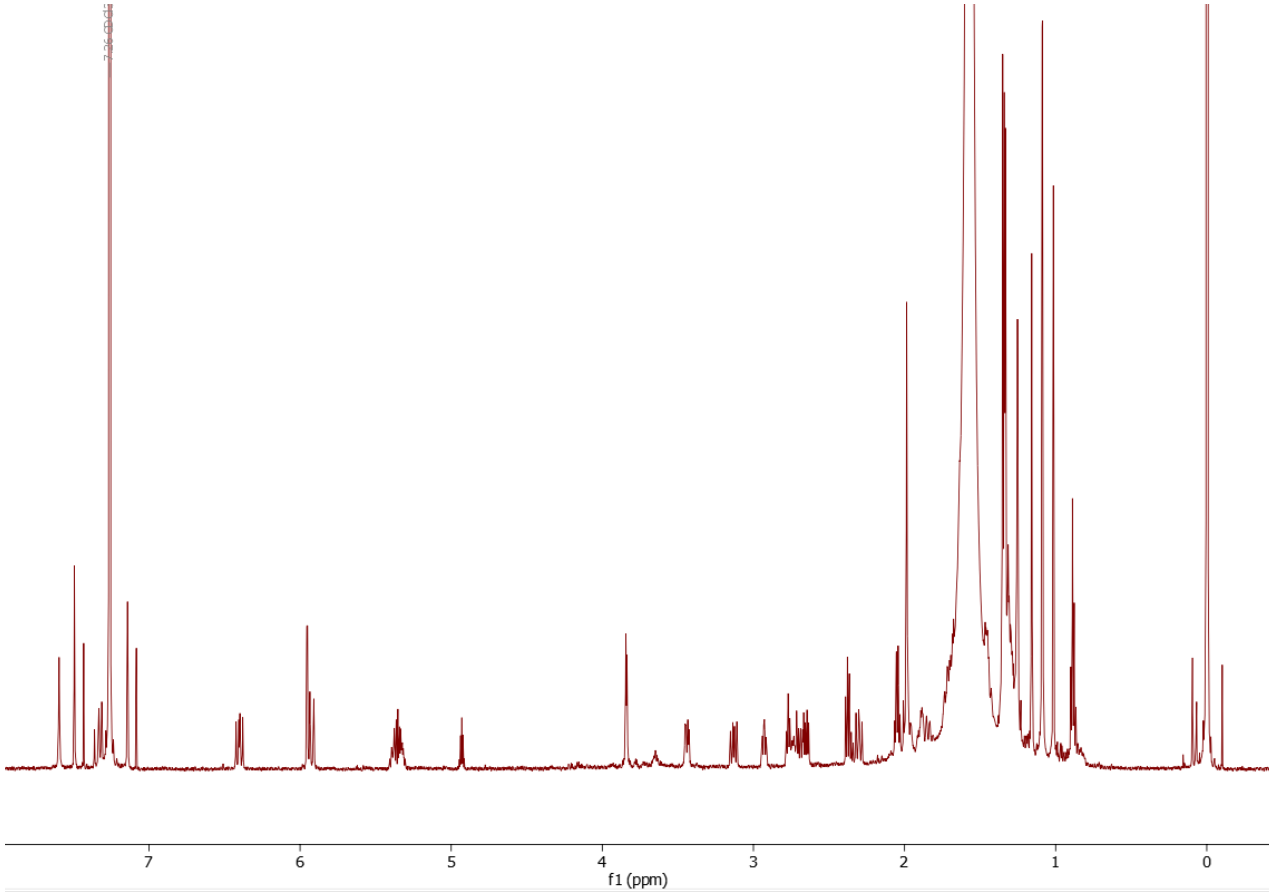


Figure S2. ^{1}H NMR spectrum of NAD

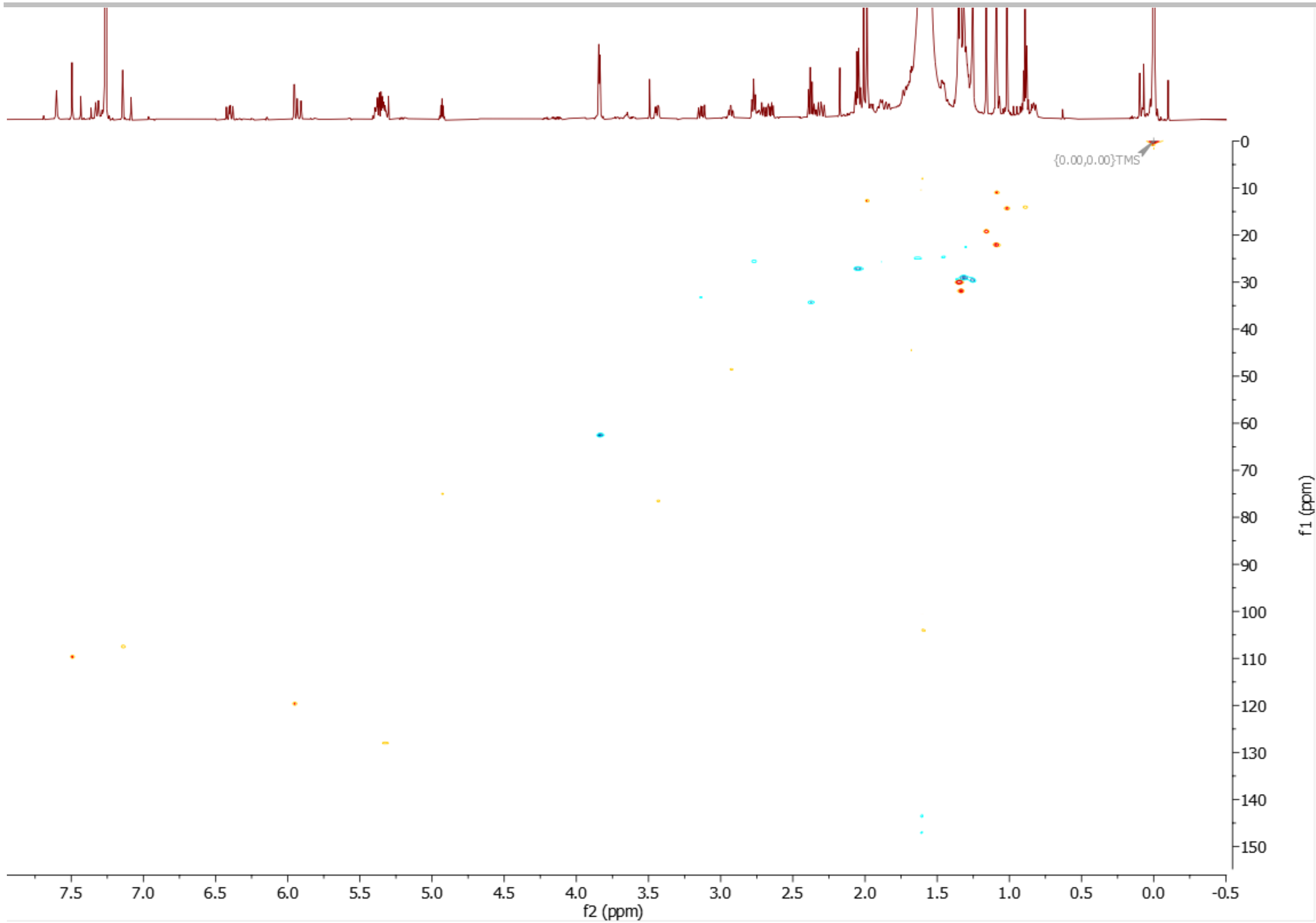


Figure S3. HSQC Spectrum of NAD

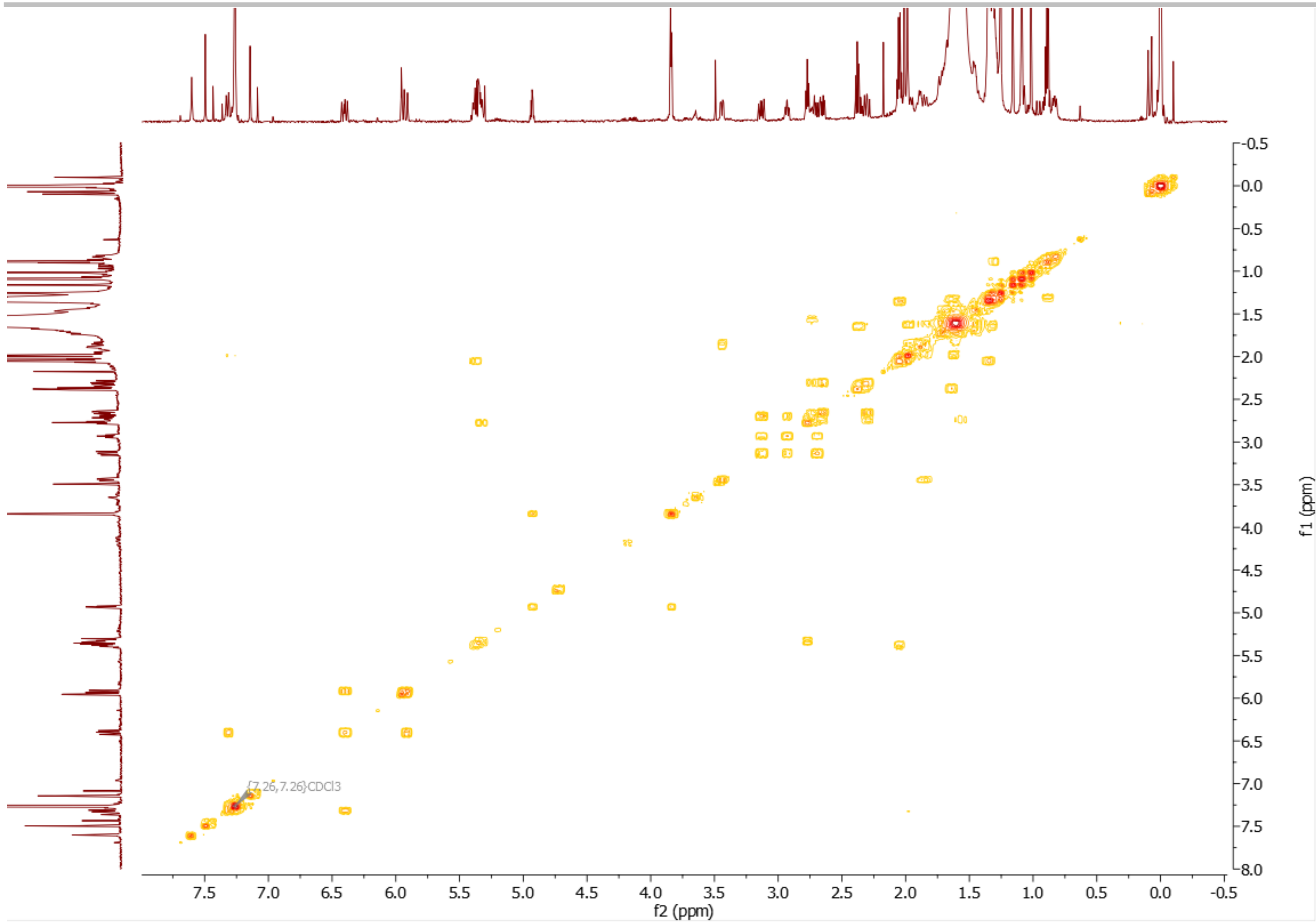


Figure S4. COSY Spectrum of NAD

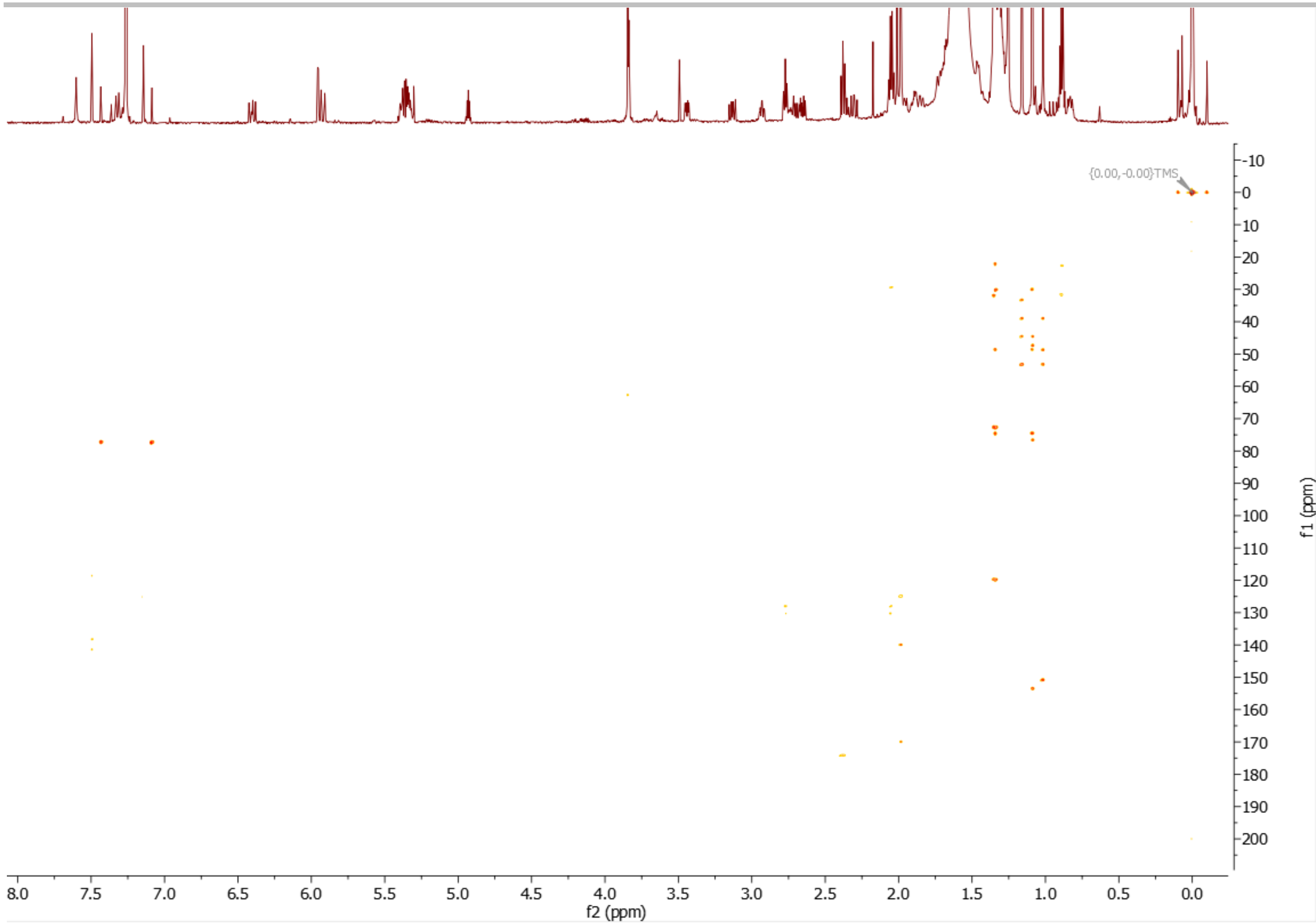


Figure S5. HMBC Spectrum of NAD

NAD_1 NMR Spectra

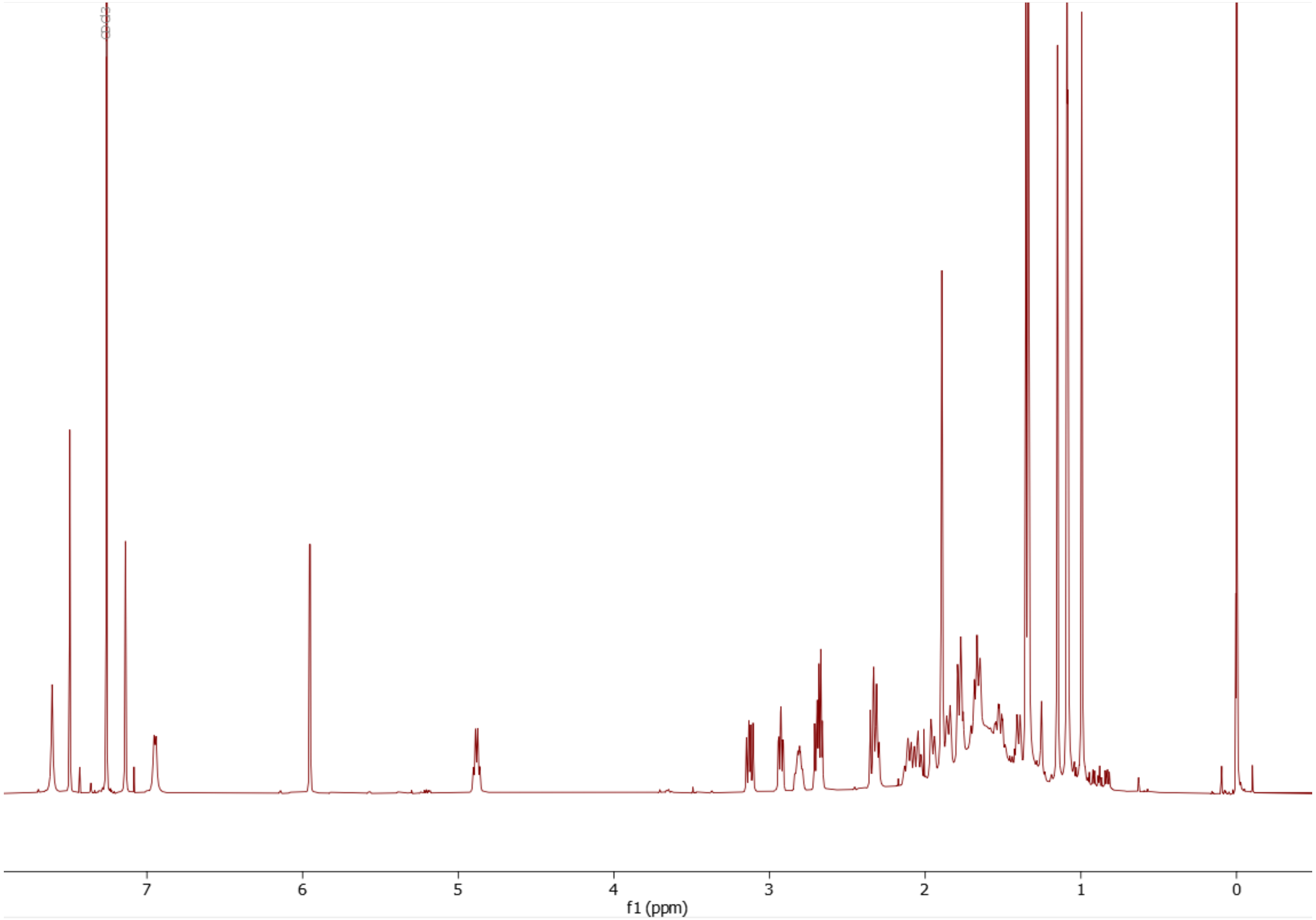


Figure S6. ^{1}H NMR Spectrum of NAD_1

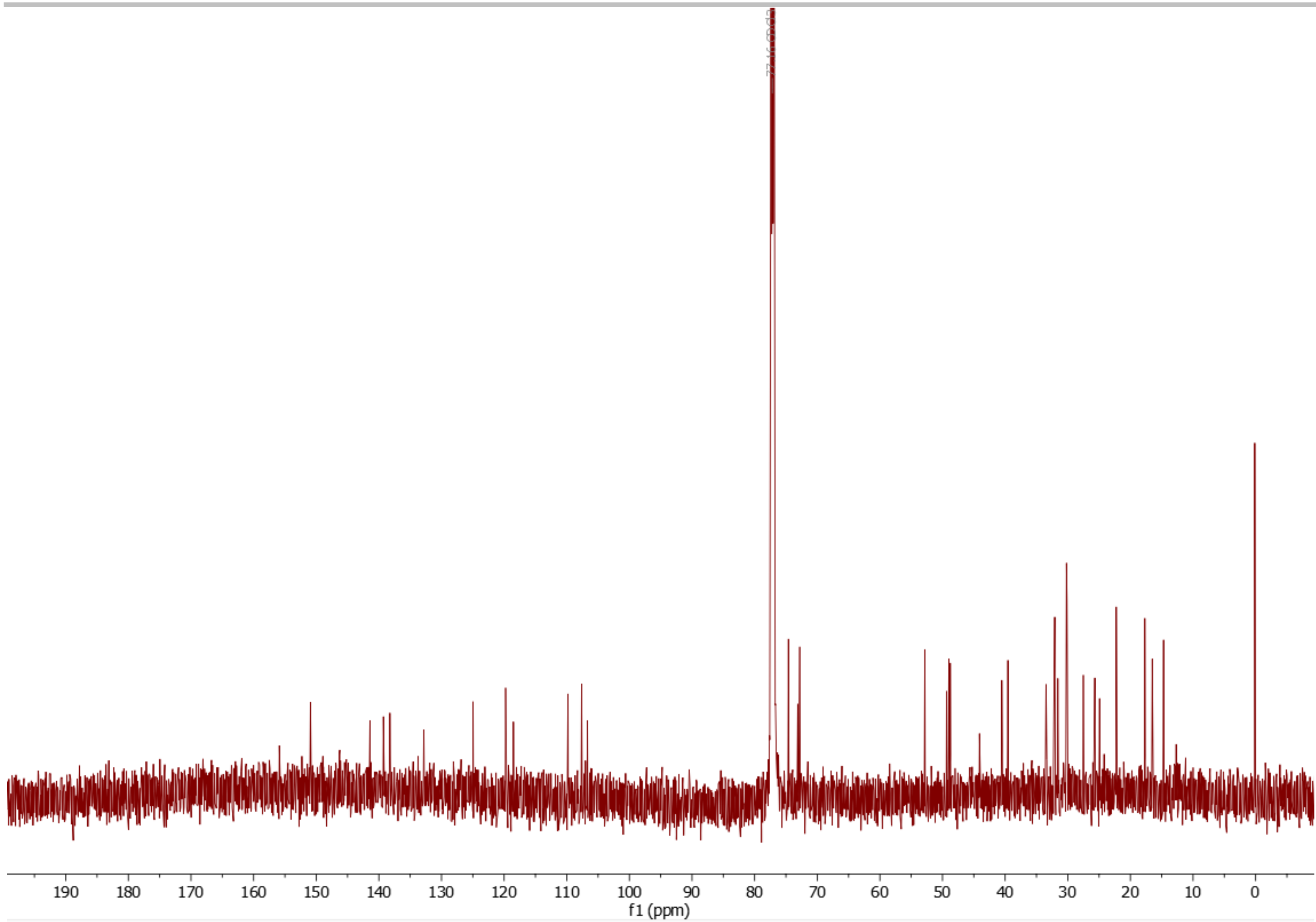


Figure S7. ^{13}C NMR Spectrum of NAD_1

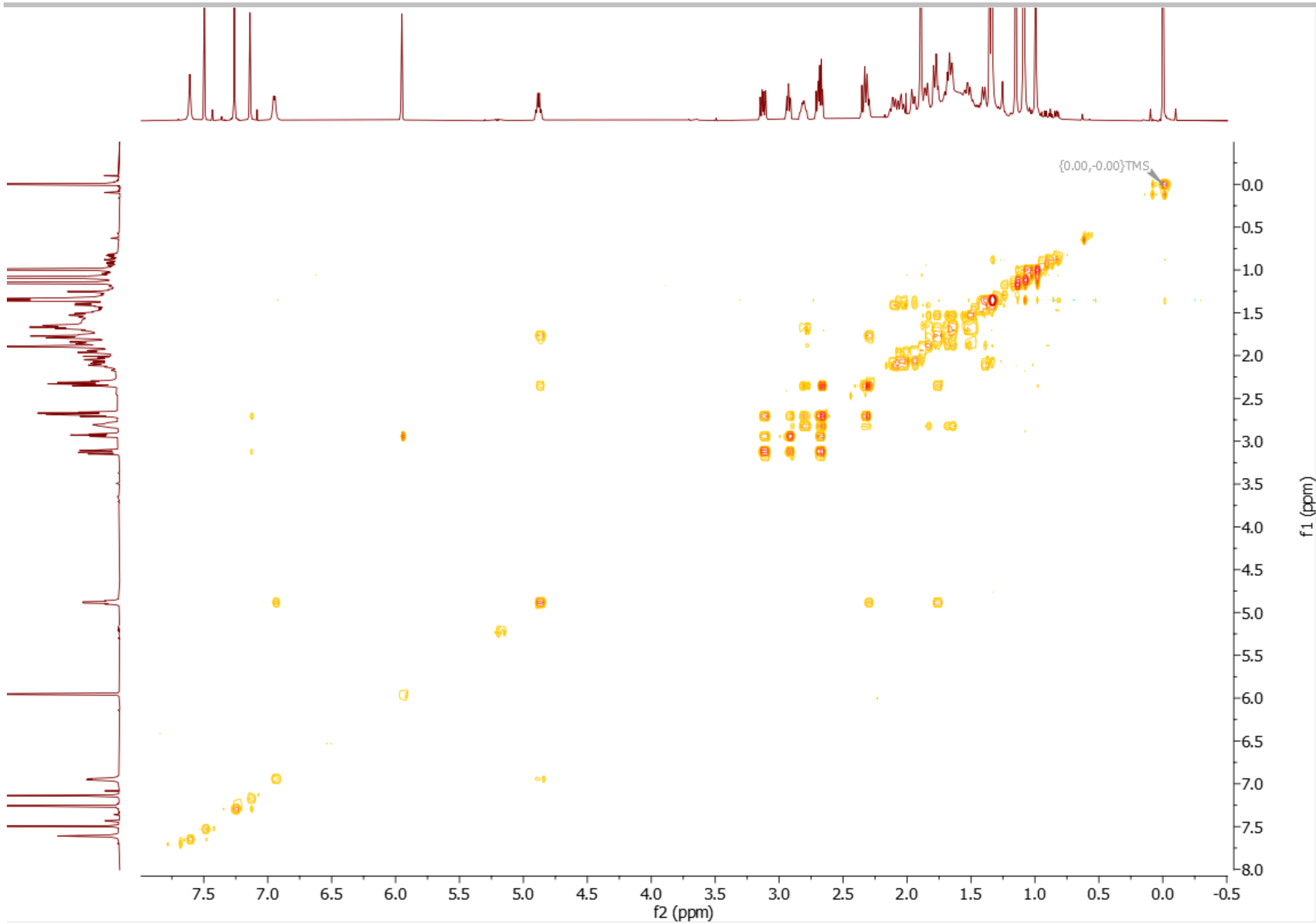


Figure S8. COSY Spectrum of NAD_1

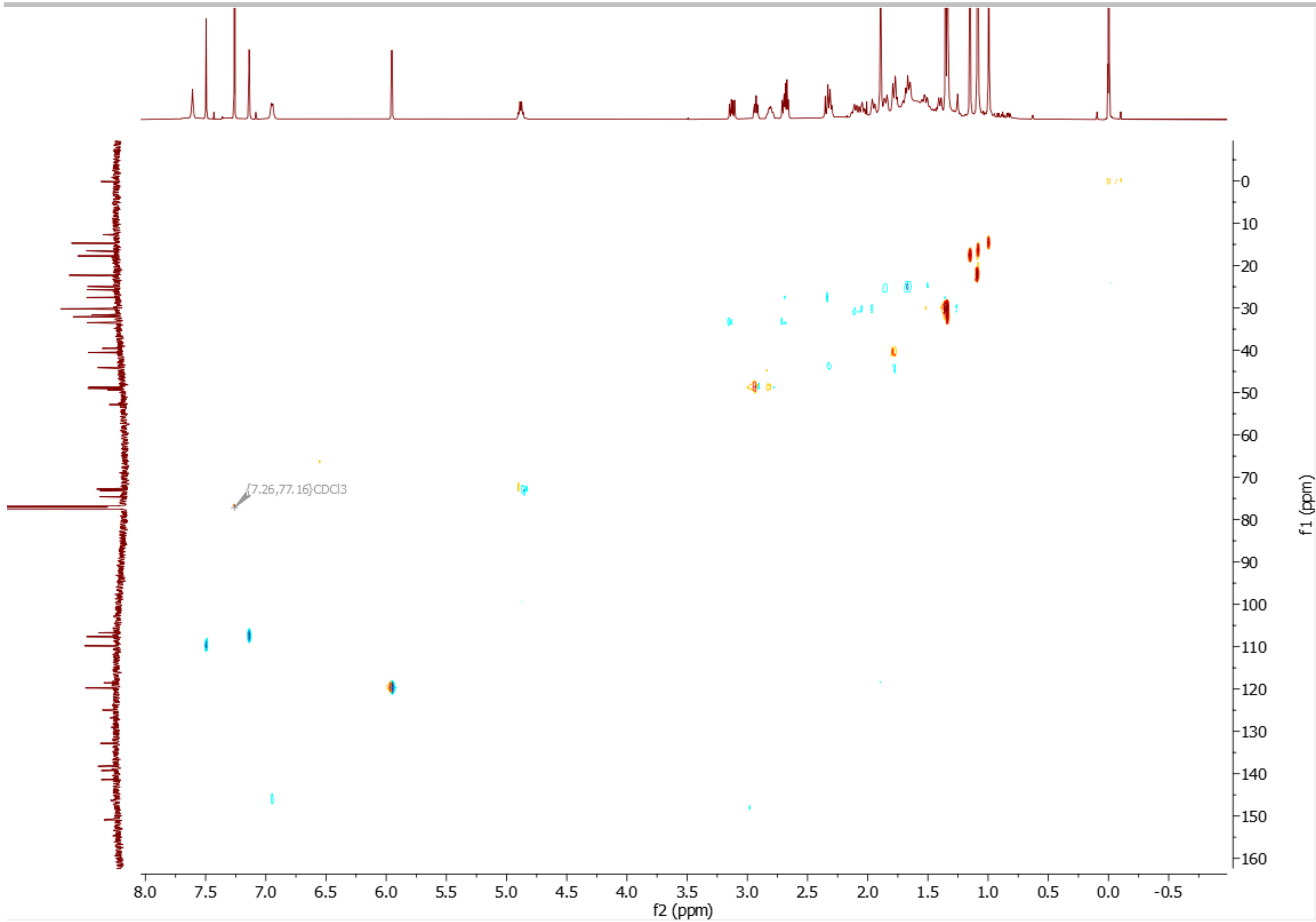


Figure S9. HSQC Spectrum of NAD_1

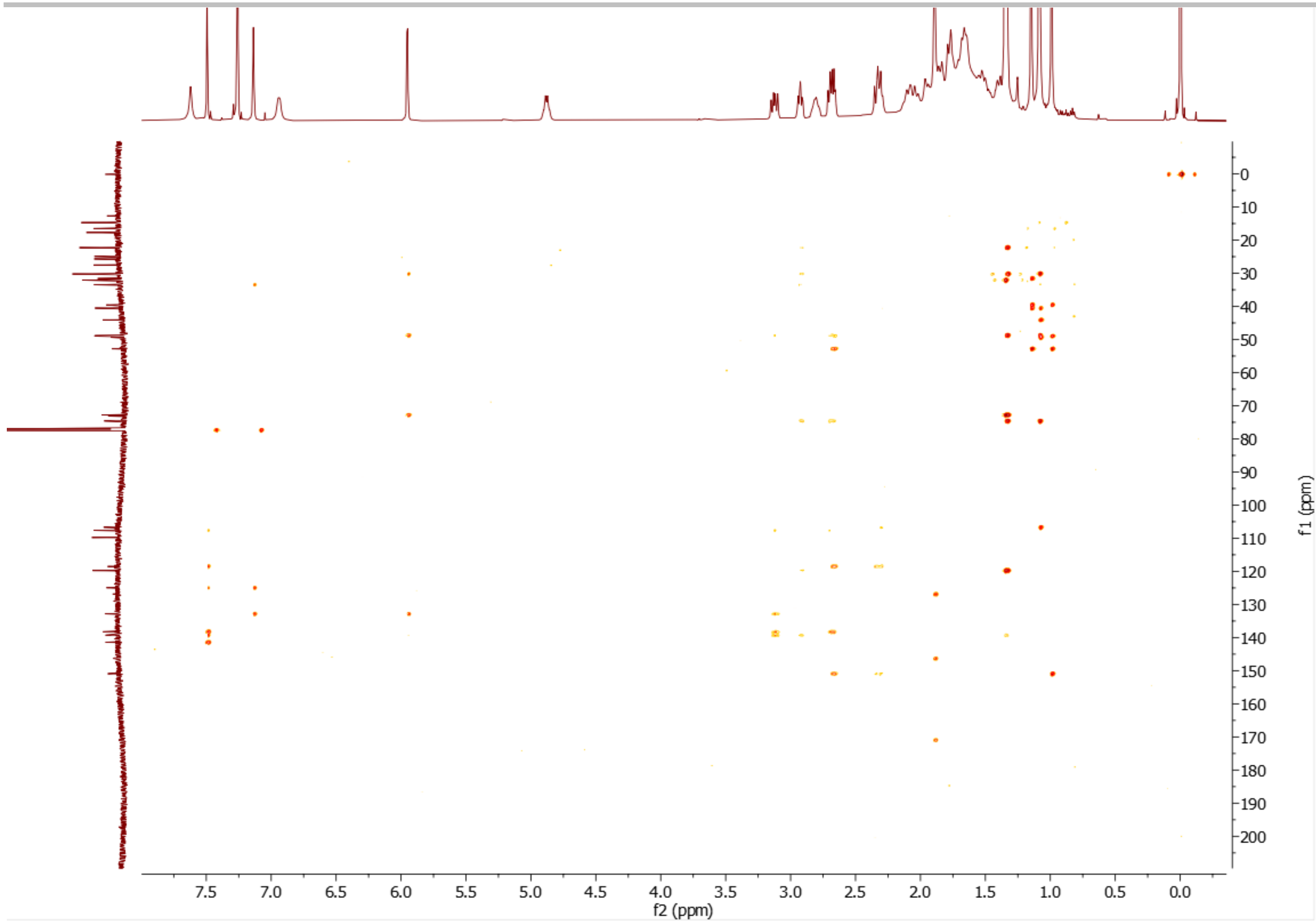


Figure S10. HMBC Spectrum of NAD_1

NAD4 NMR Spectra

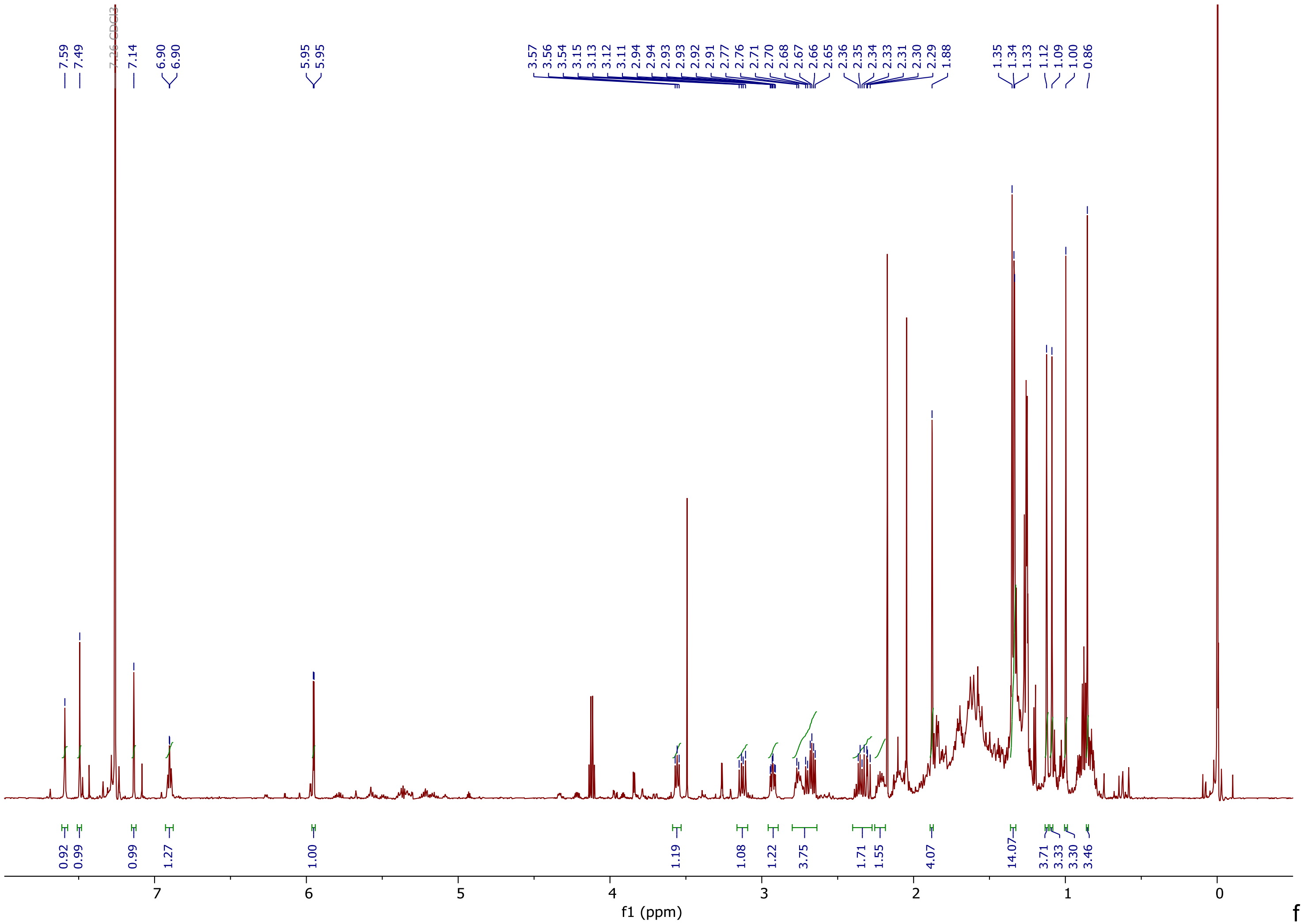


Figure S11. ^{1}H NMR Spectrum of NAD_4

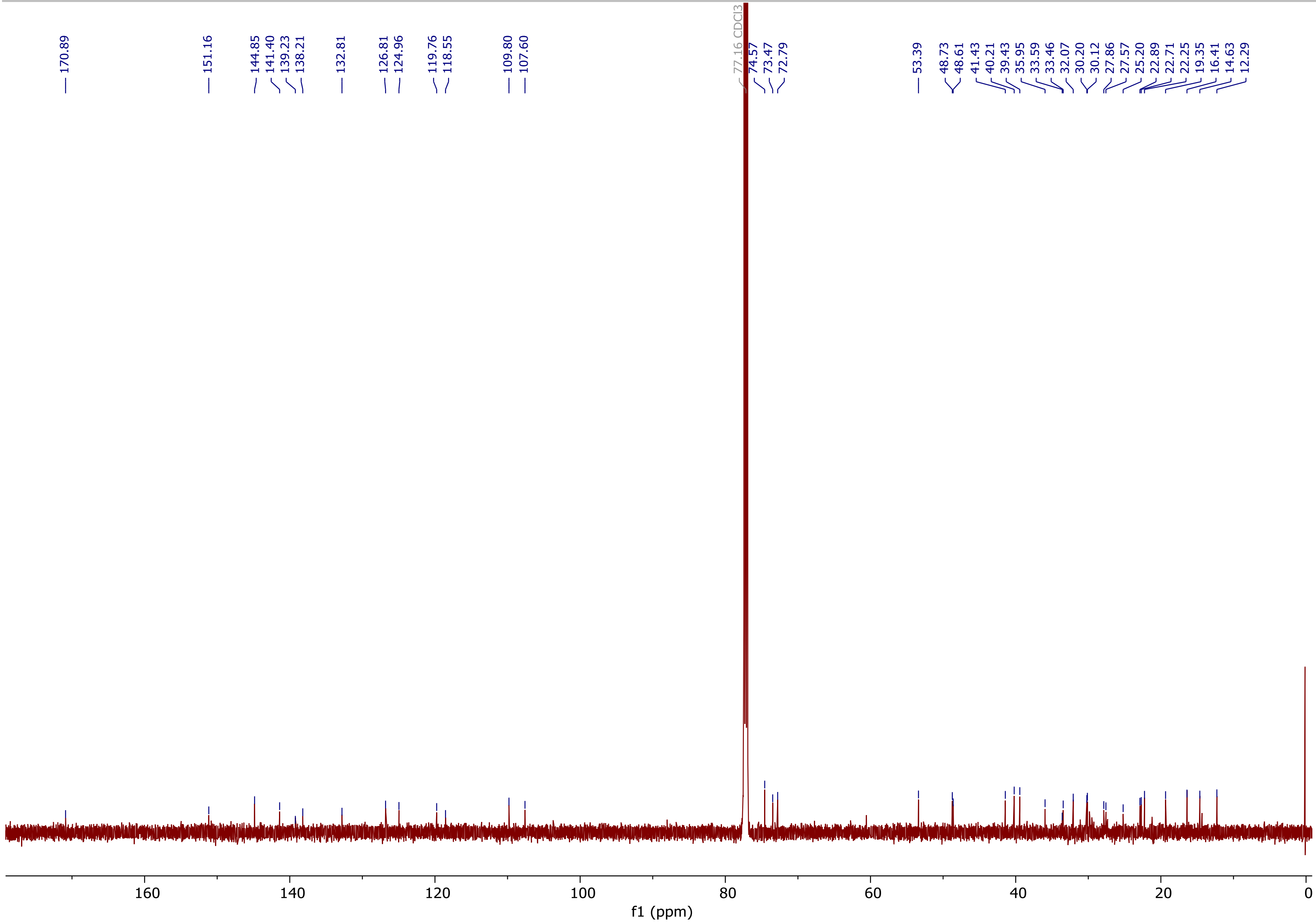


Figure S12. ^{13}C NMR Spectrum of NAD_4

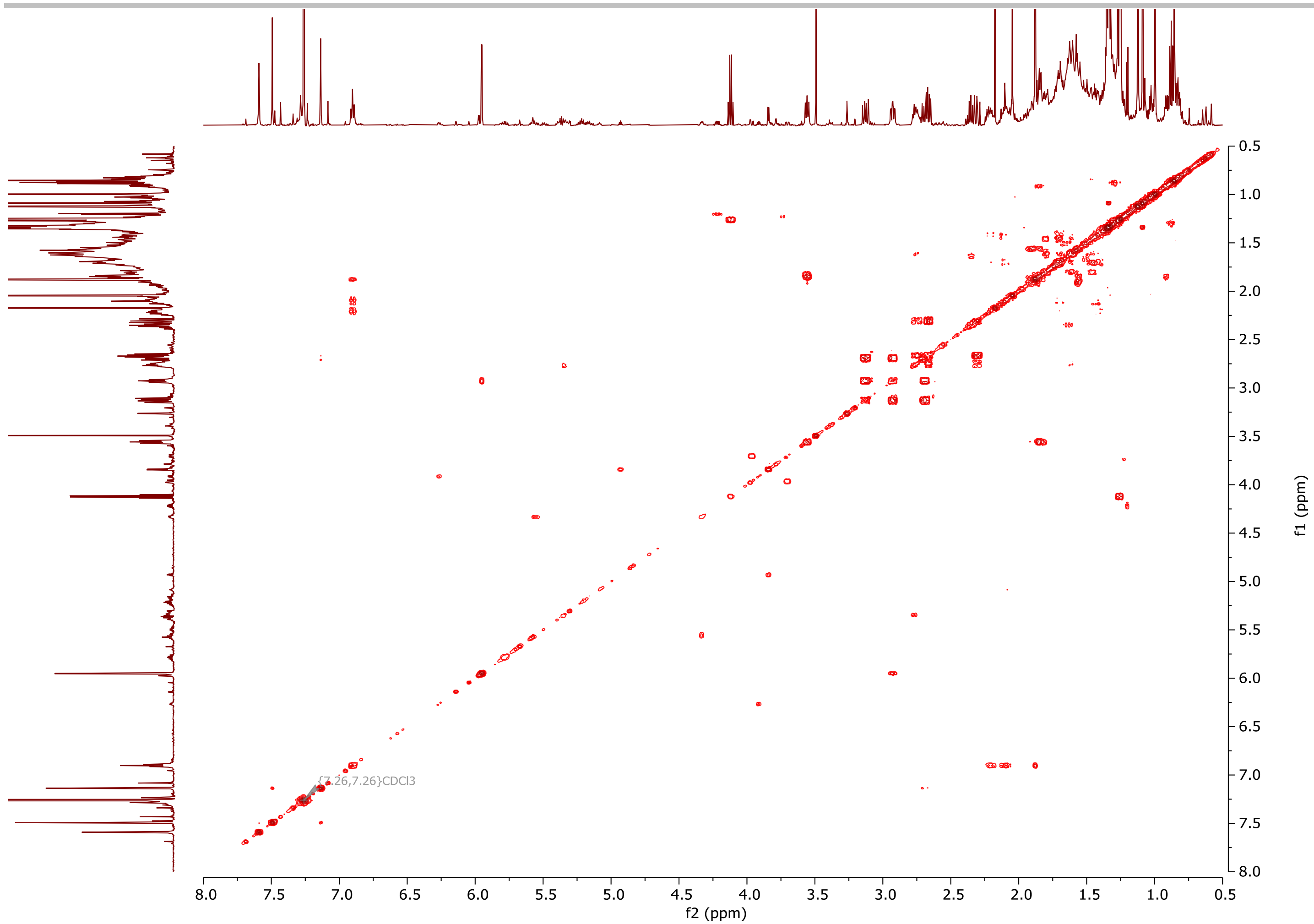


Figure S13. COSY Spectrum of NAD_4

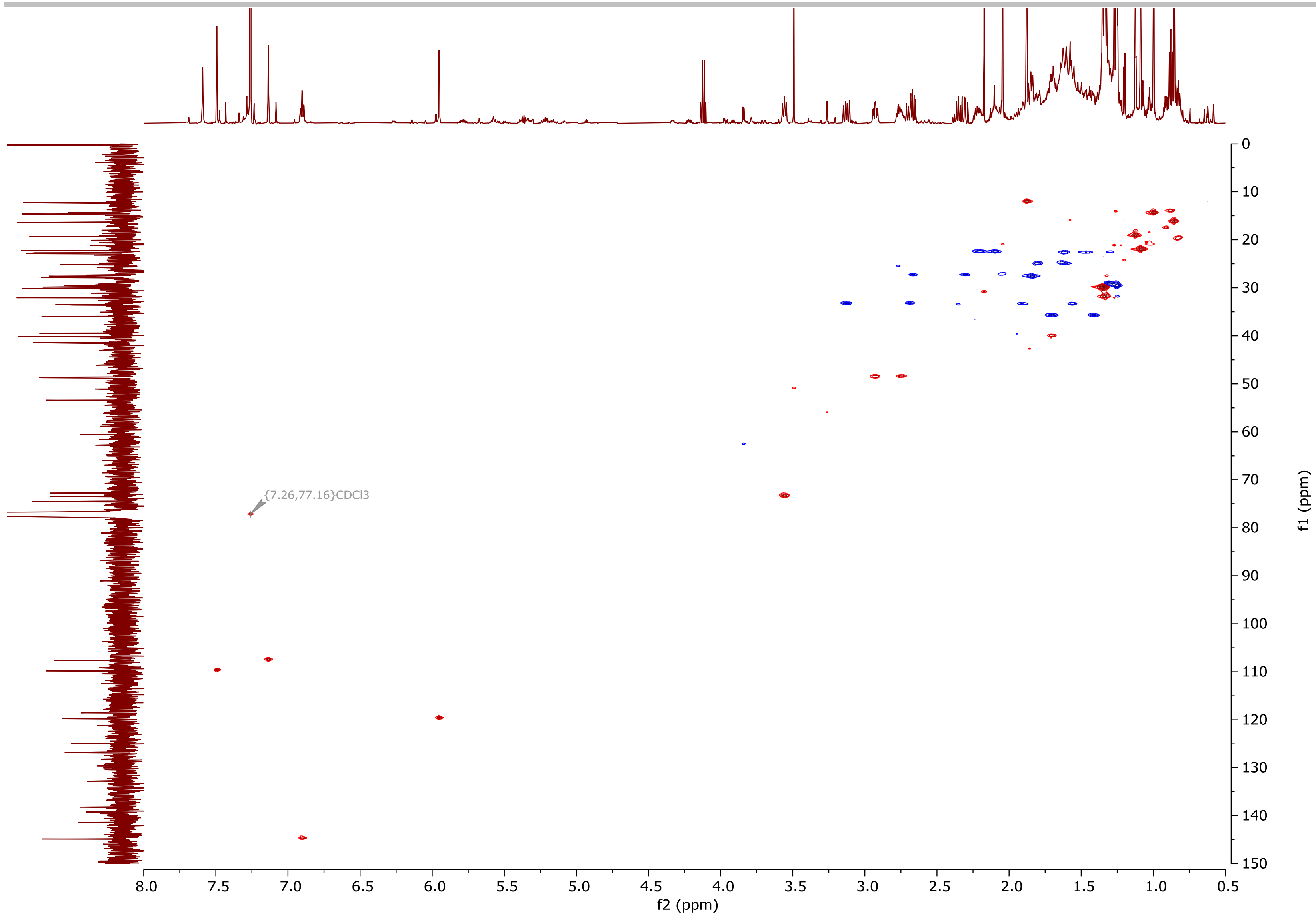


Figure S14. HSQC Spectrum of NAD_4

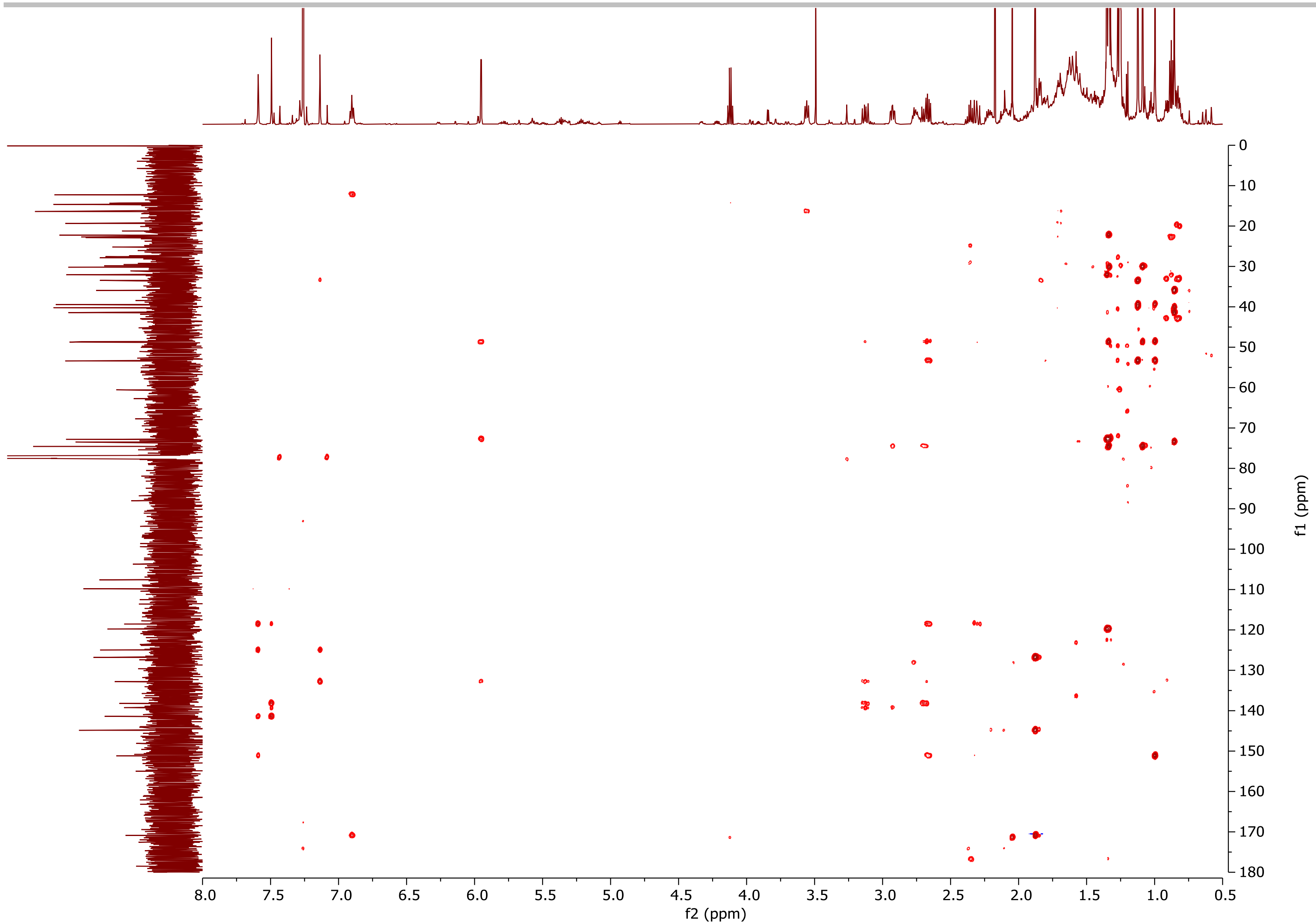


Figure S15. HMBC Spectrum of NAD_4

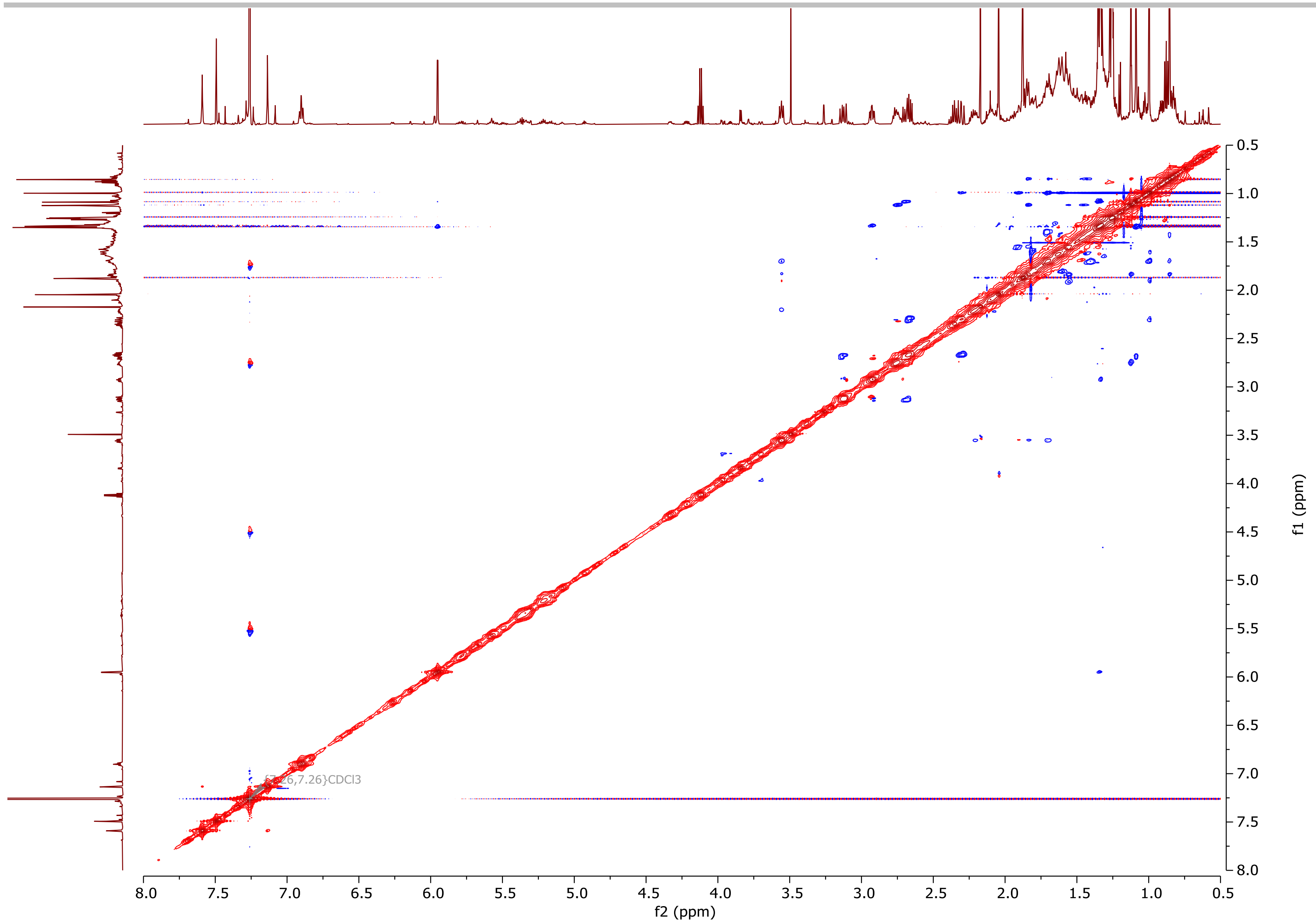


Figure S16. NOESY (mixing time = 0.5 s) NMR Spectrum of NAD_4

NAD4-OMe NMR Spectra

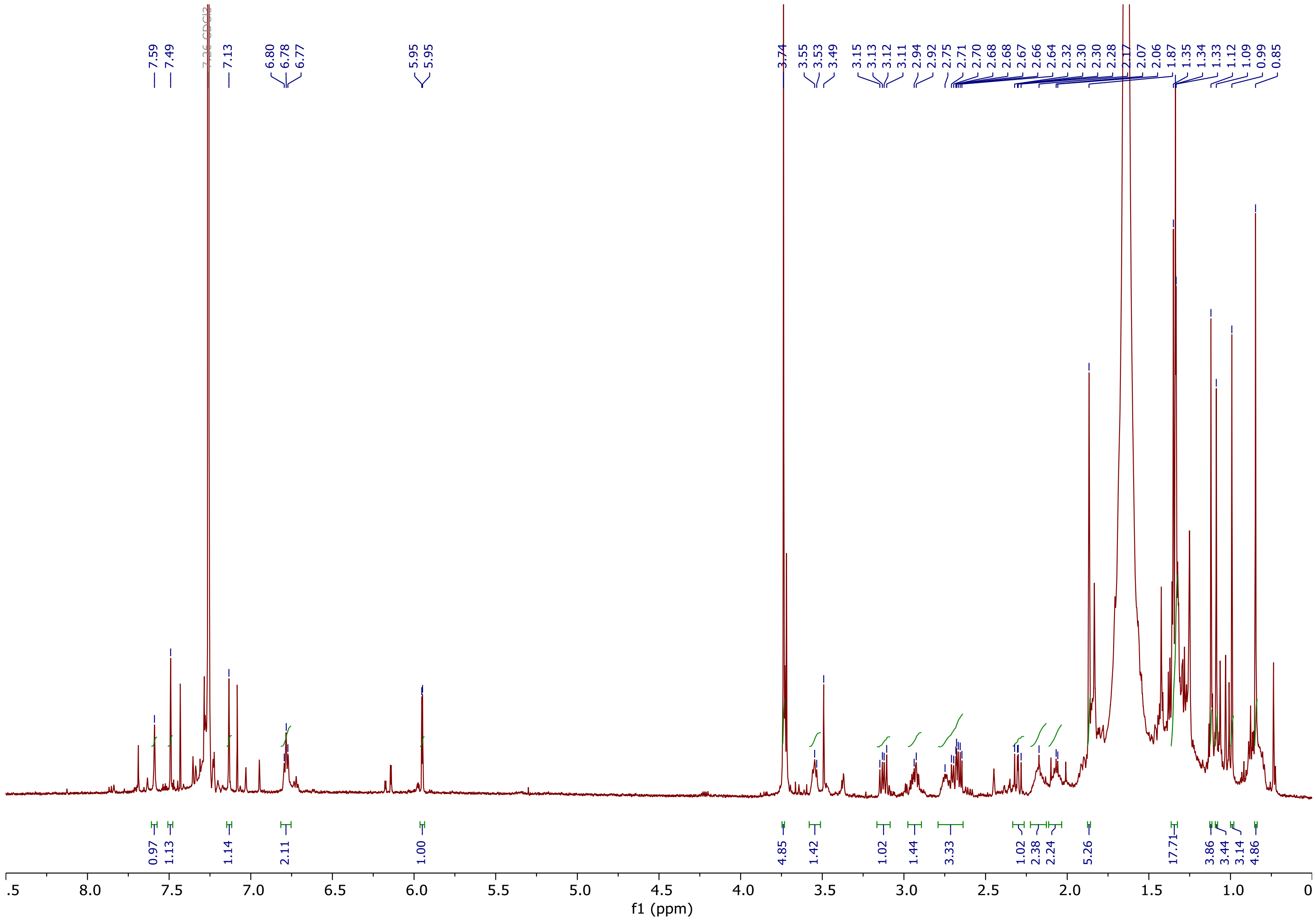


Figure S17. ^{1}H NMR Spectrum of NAD4-OMe

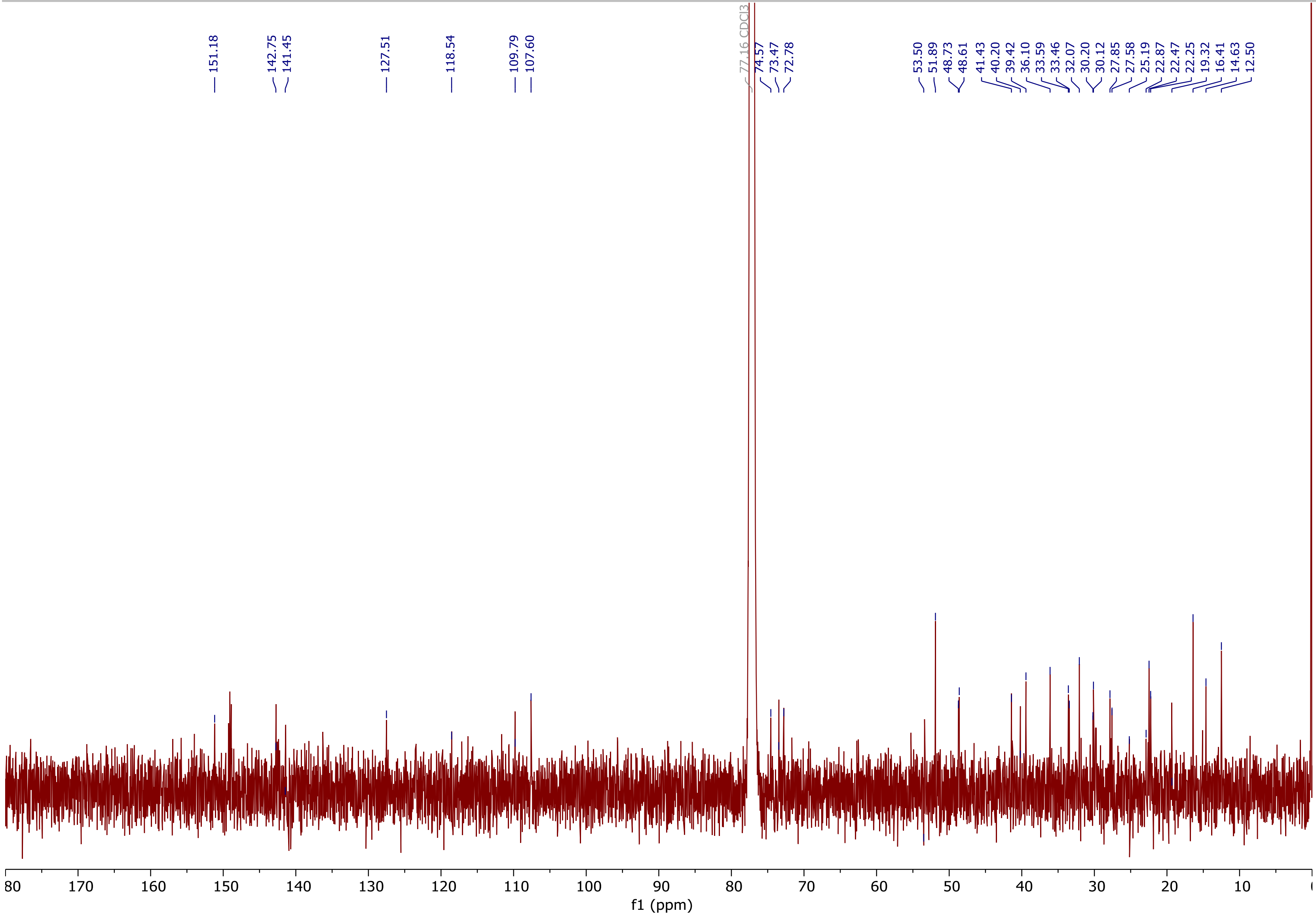


Figure S18. ^{13}C NMR Spectrum of NAD_4-OMe

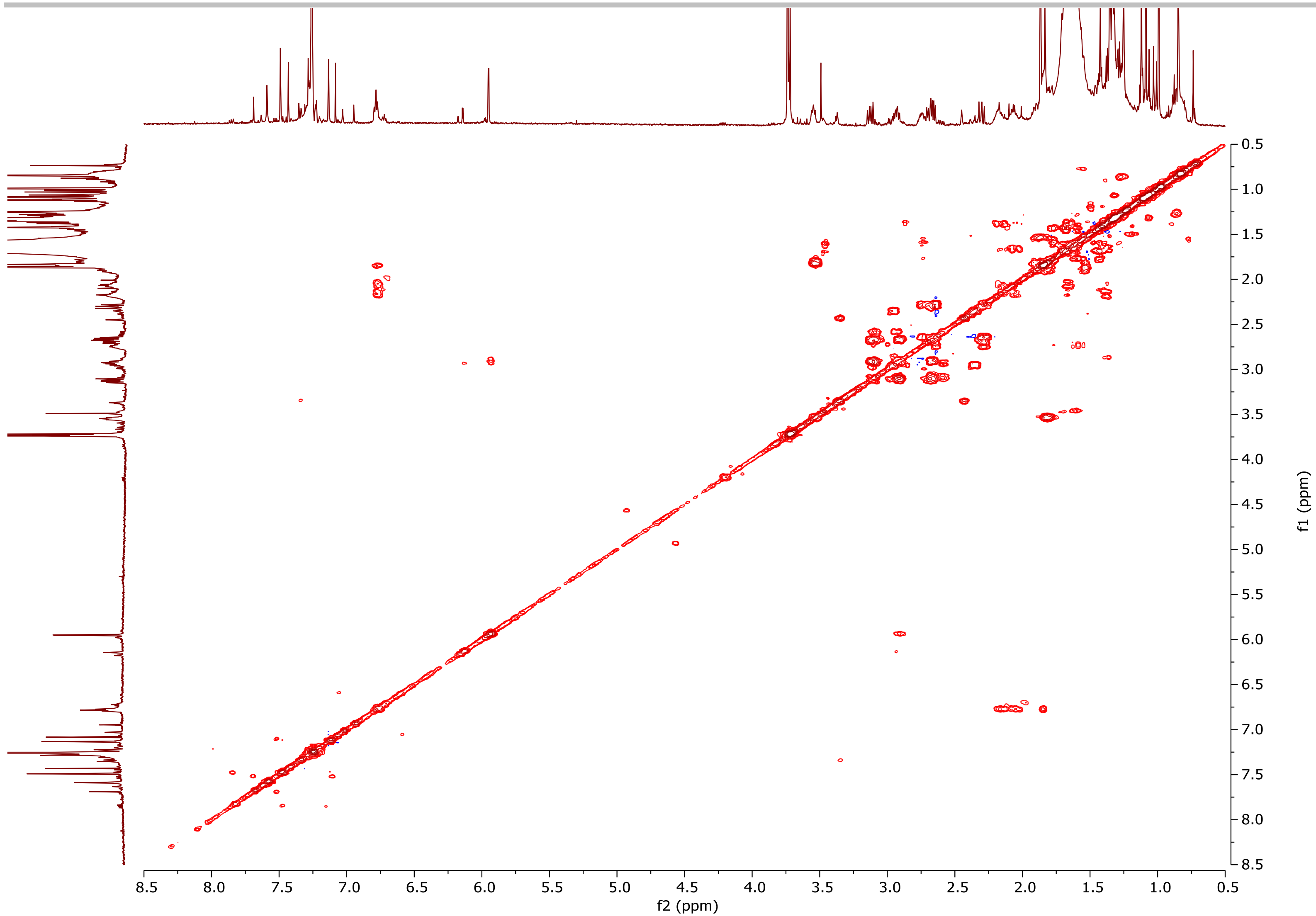


Figure S19. COSY Spectrum of NAD_4-OMe

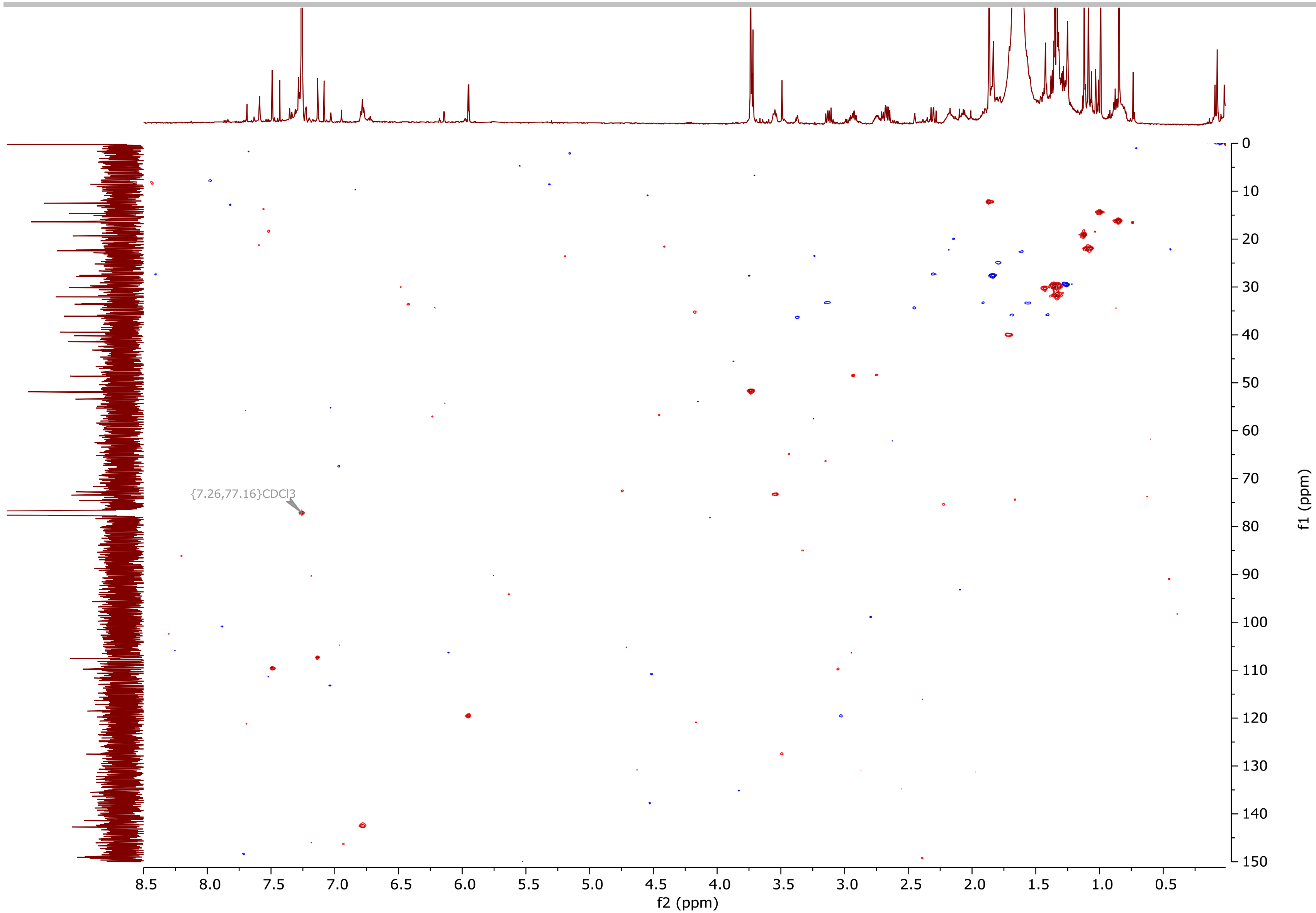


Figure S20. HSQC Spectrum of NAD_4-OMe

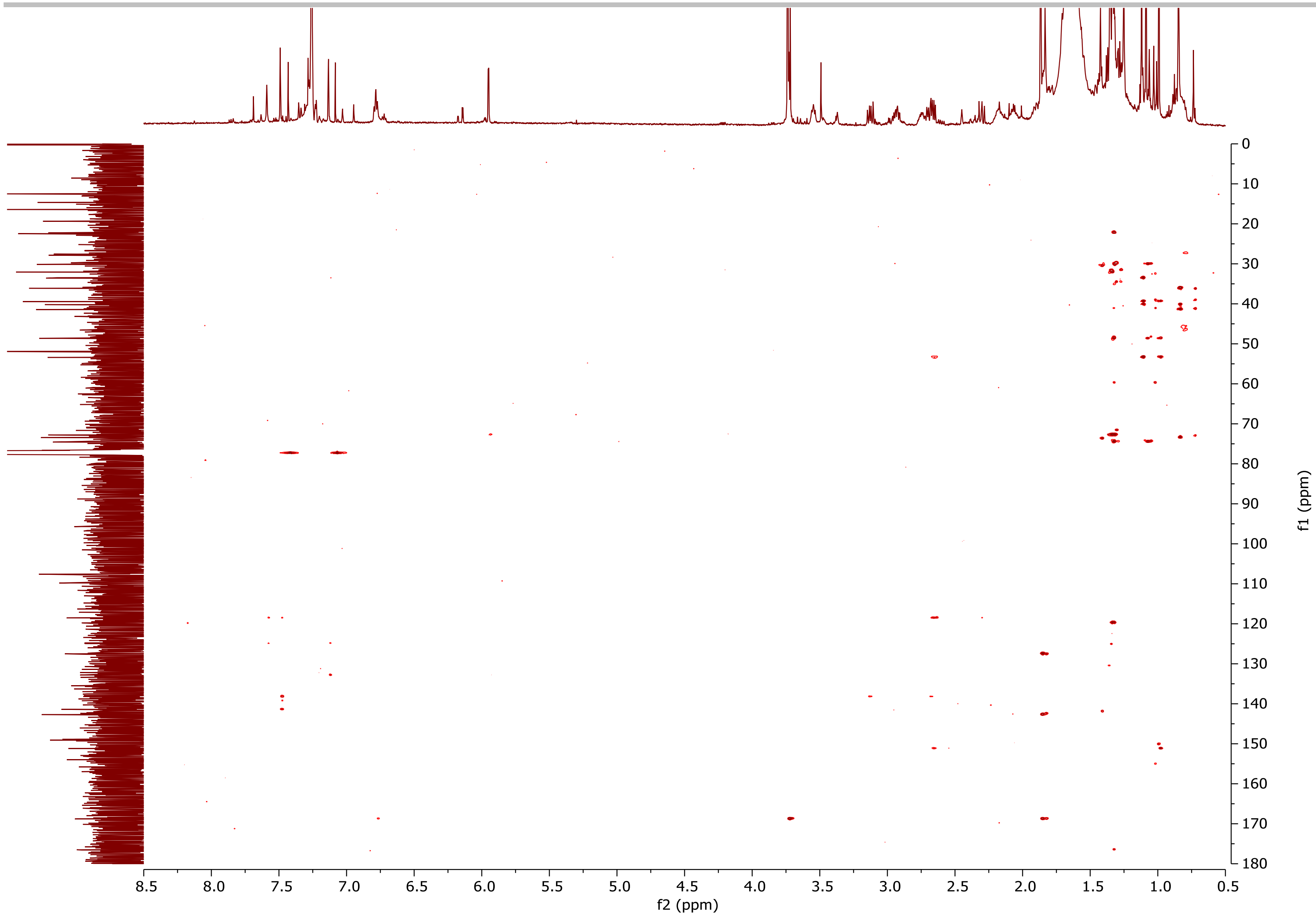


Figure S21. HMBC Spectrum of NAD_4-OMe

NAD$_2$ NMR Spectra

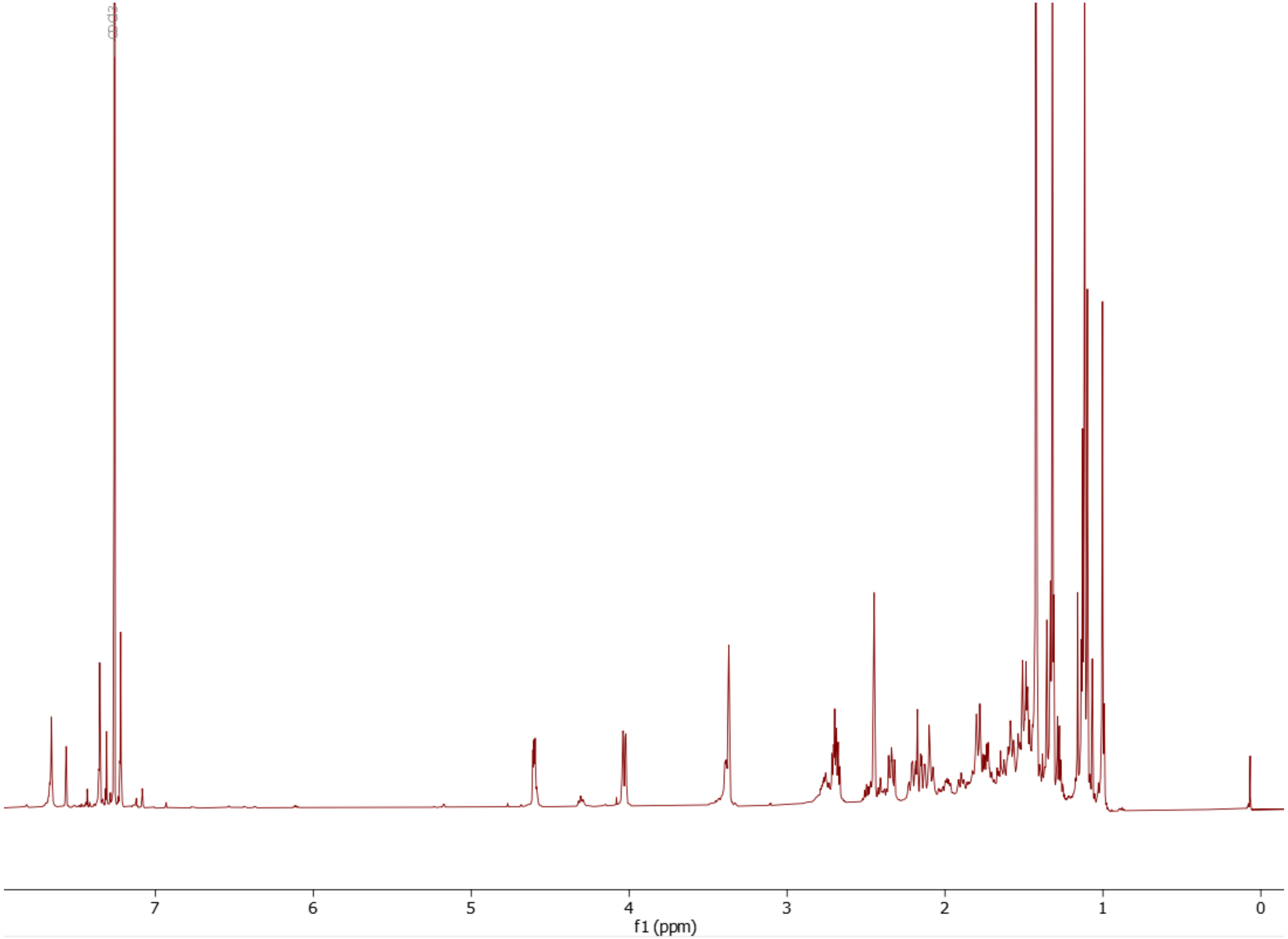


Figure S22. ^1H NMR Spectrum of NAD$_2$

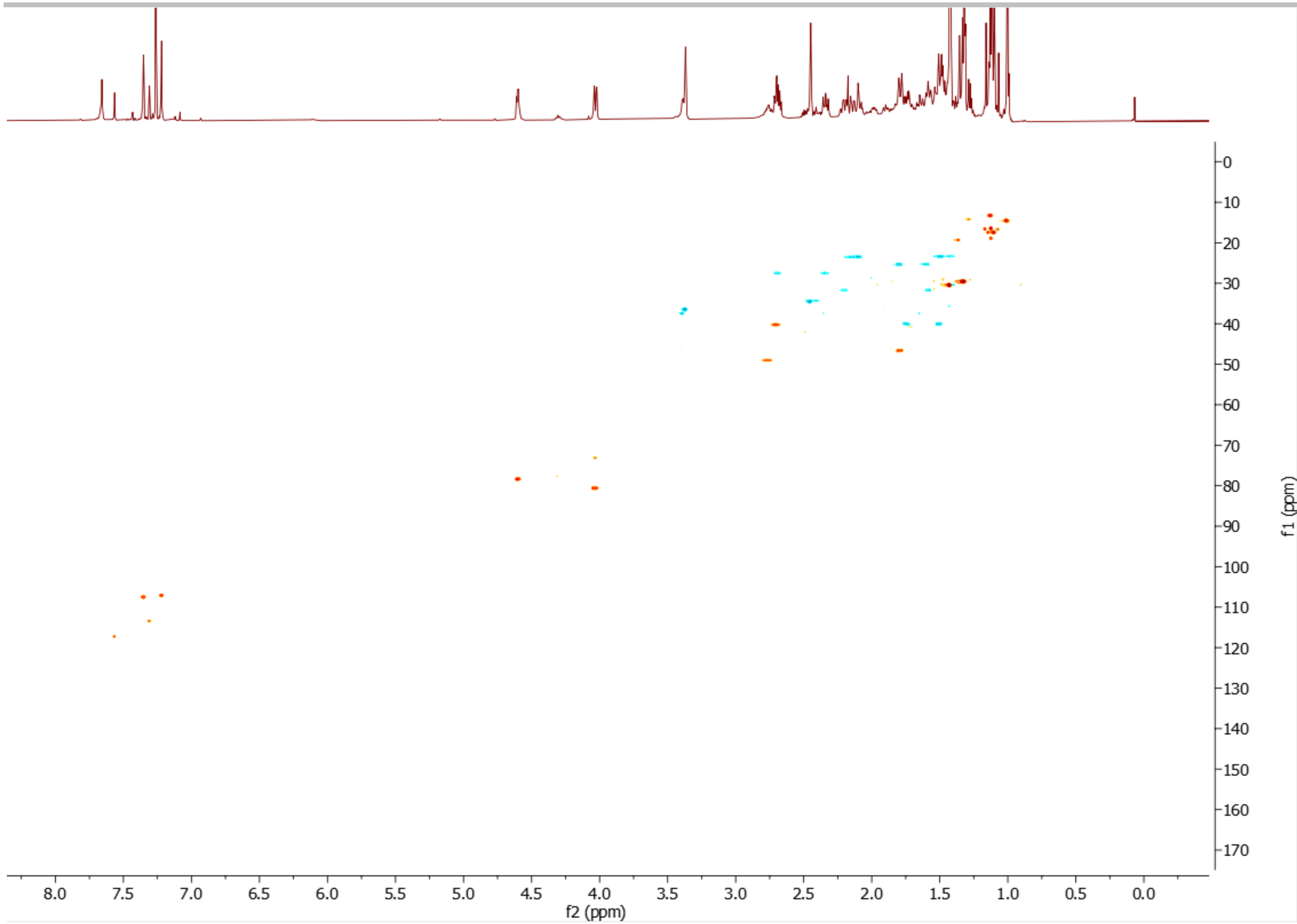


Figure S23. HSQC Spectrum of NAD_2

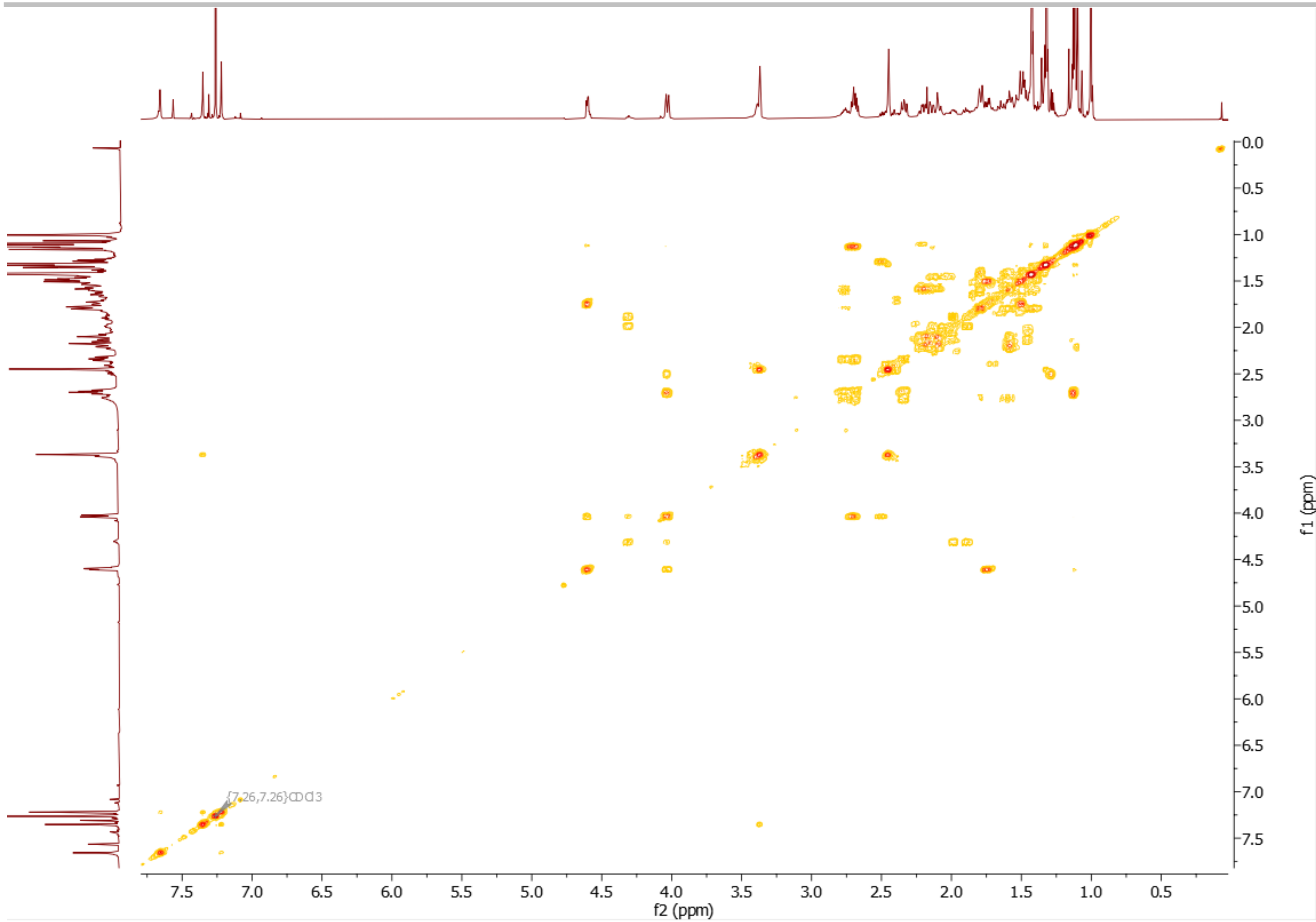


Figure S24. COSY Spectrum of NAD_2

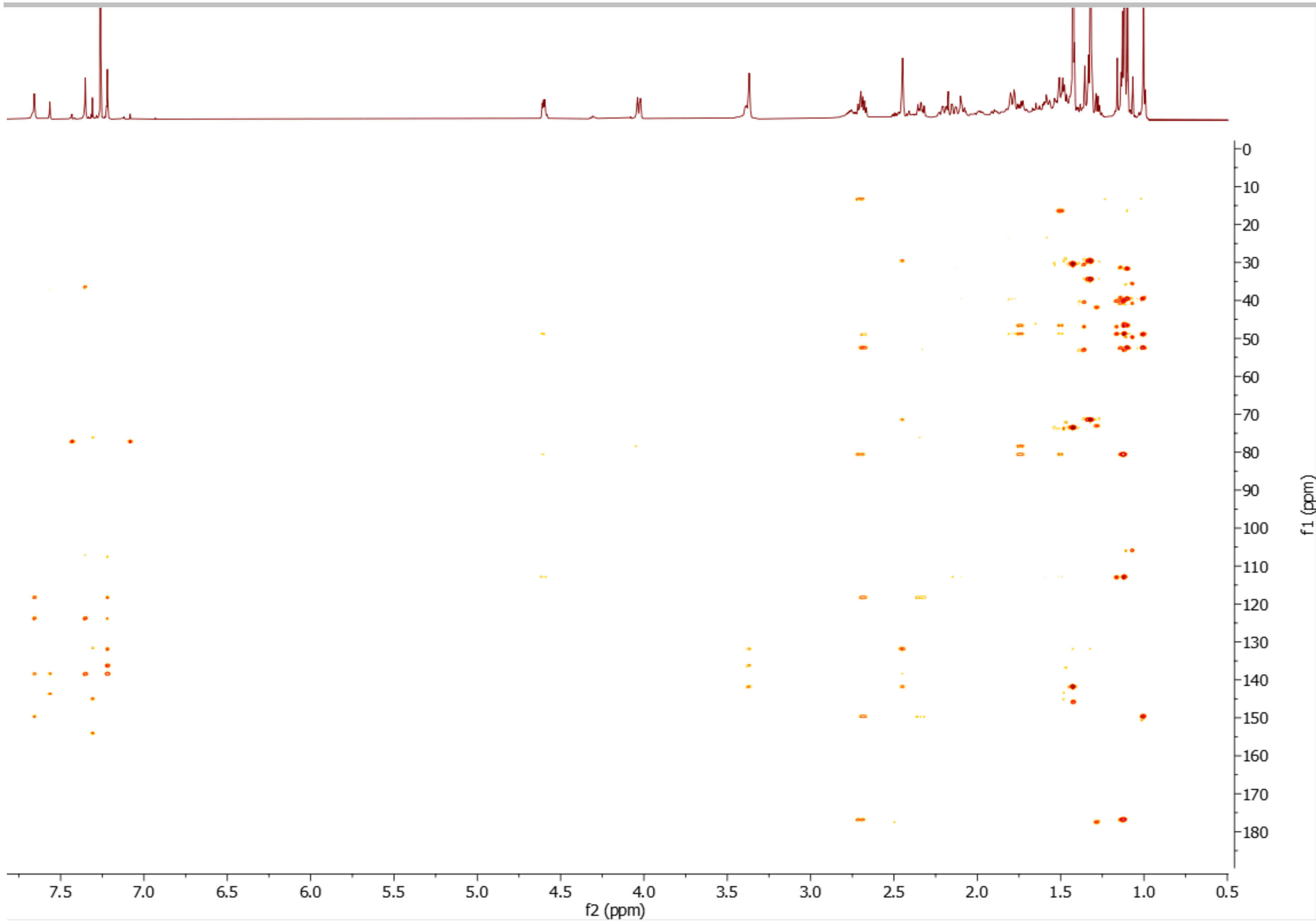


Figure S25. HMBC Spectrum of NAD_2

NAC NMR Spectra

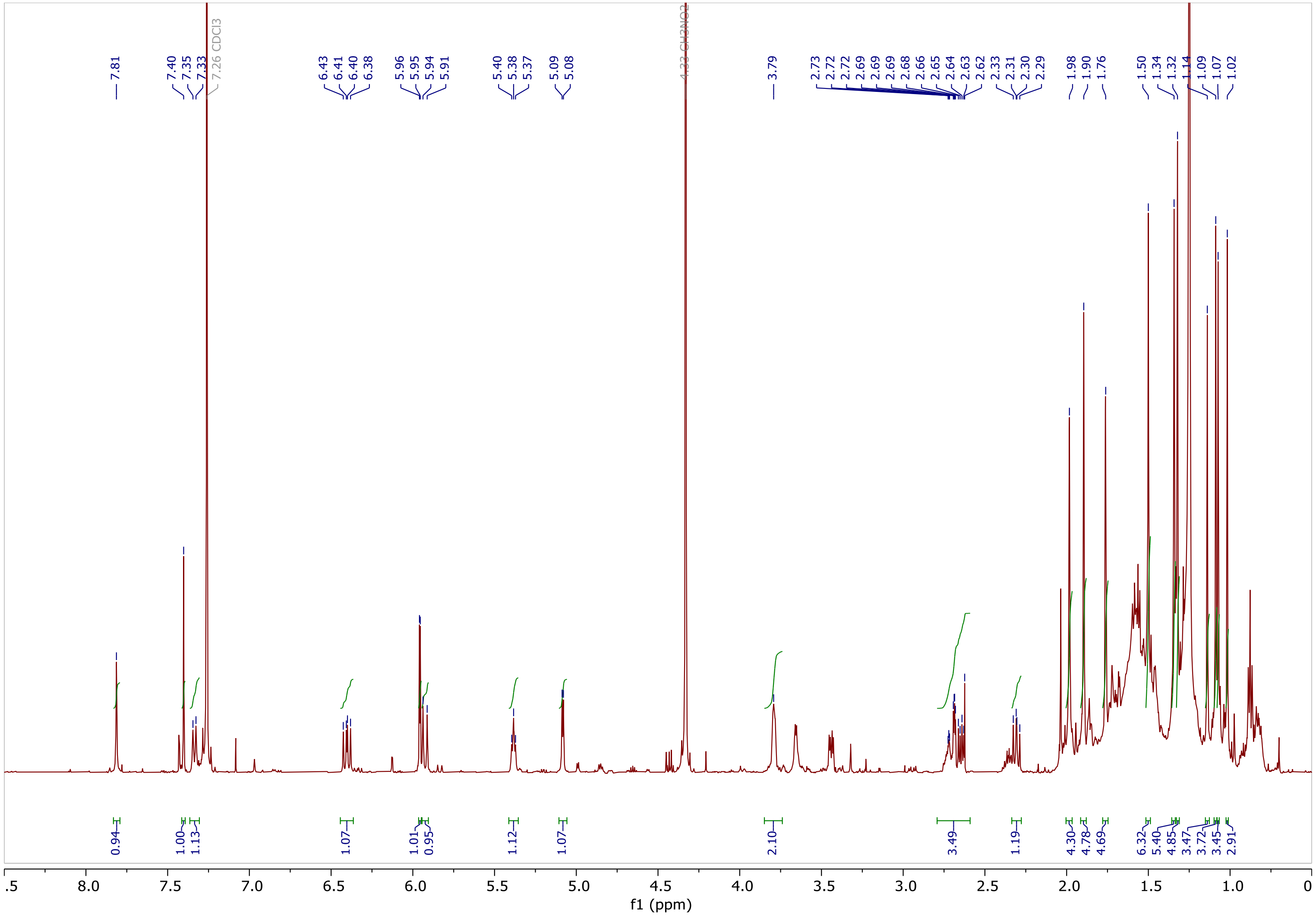


Figure S26. ^{1}H NMR Spectrum of NAC

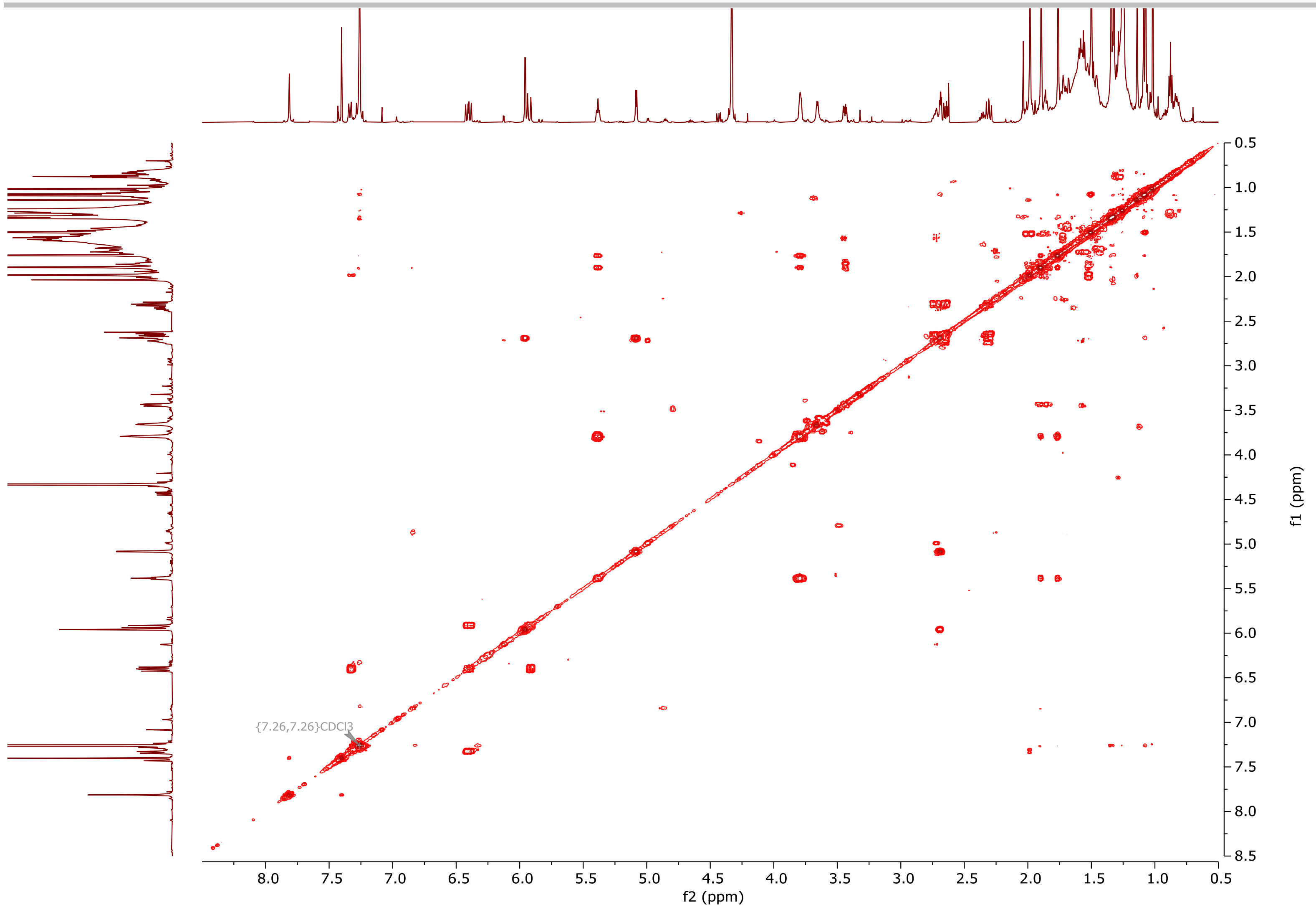


Figure S27. COSY Spectrum of NAC

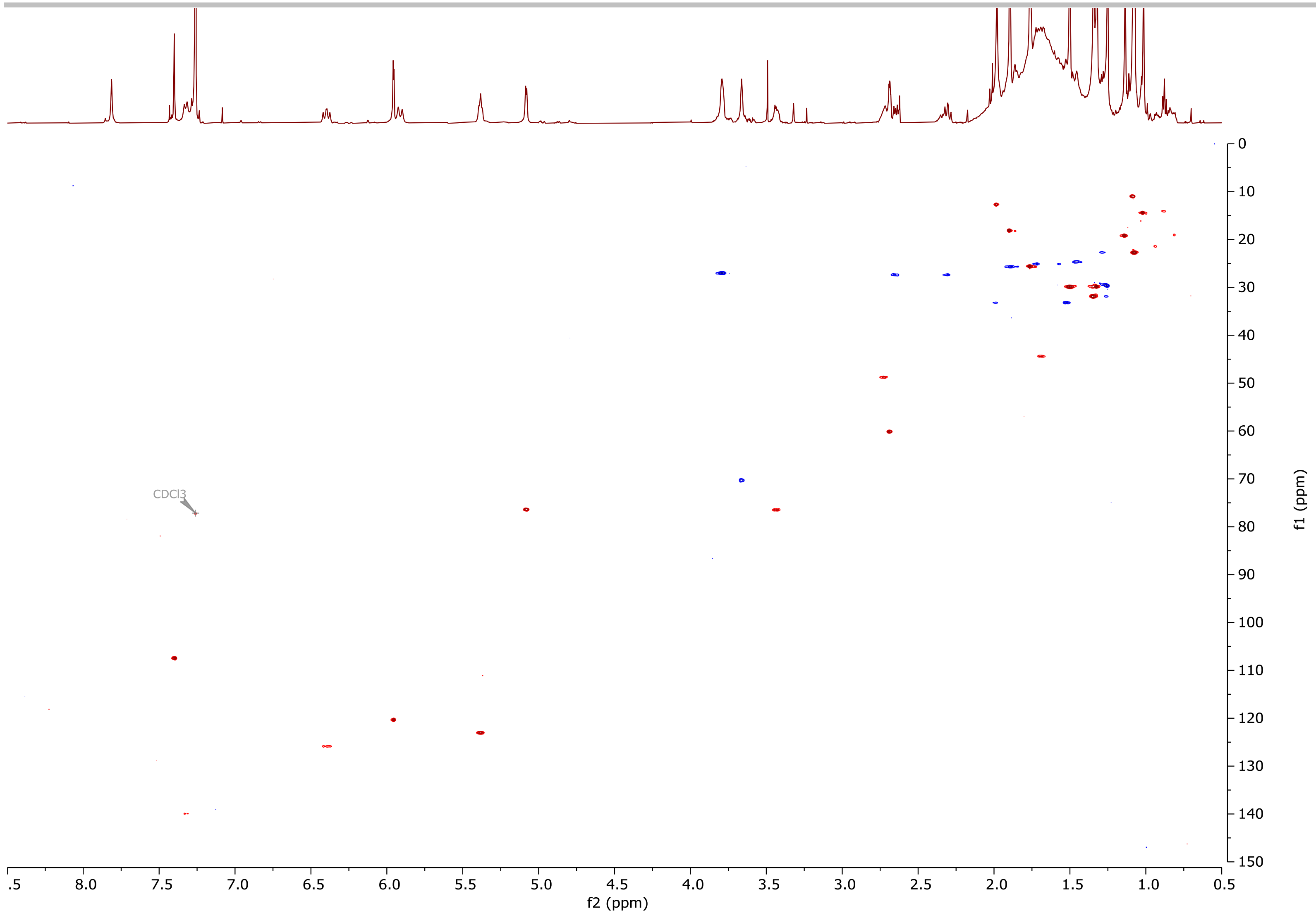


Figure S28. HSQC Spectrum of NAC

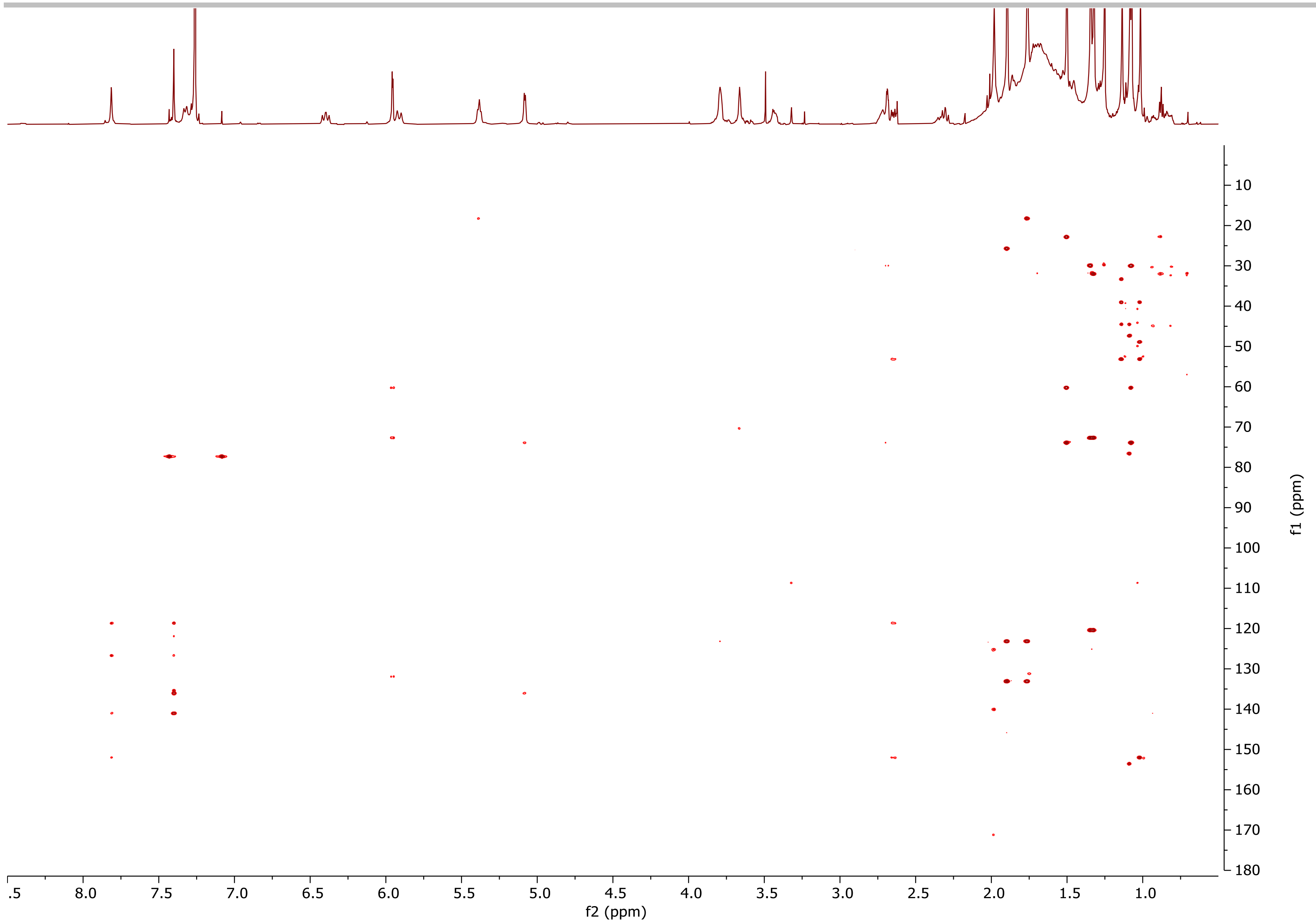


Figure S29. HMBC Spectrum of NAC

DH-NAC4 NMR Spectra

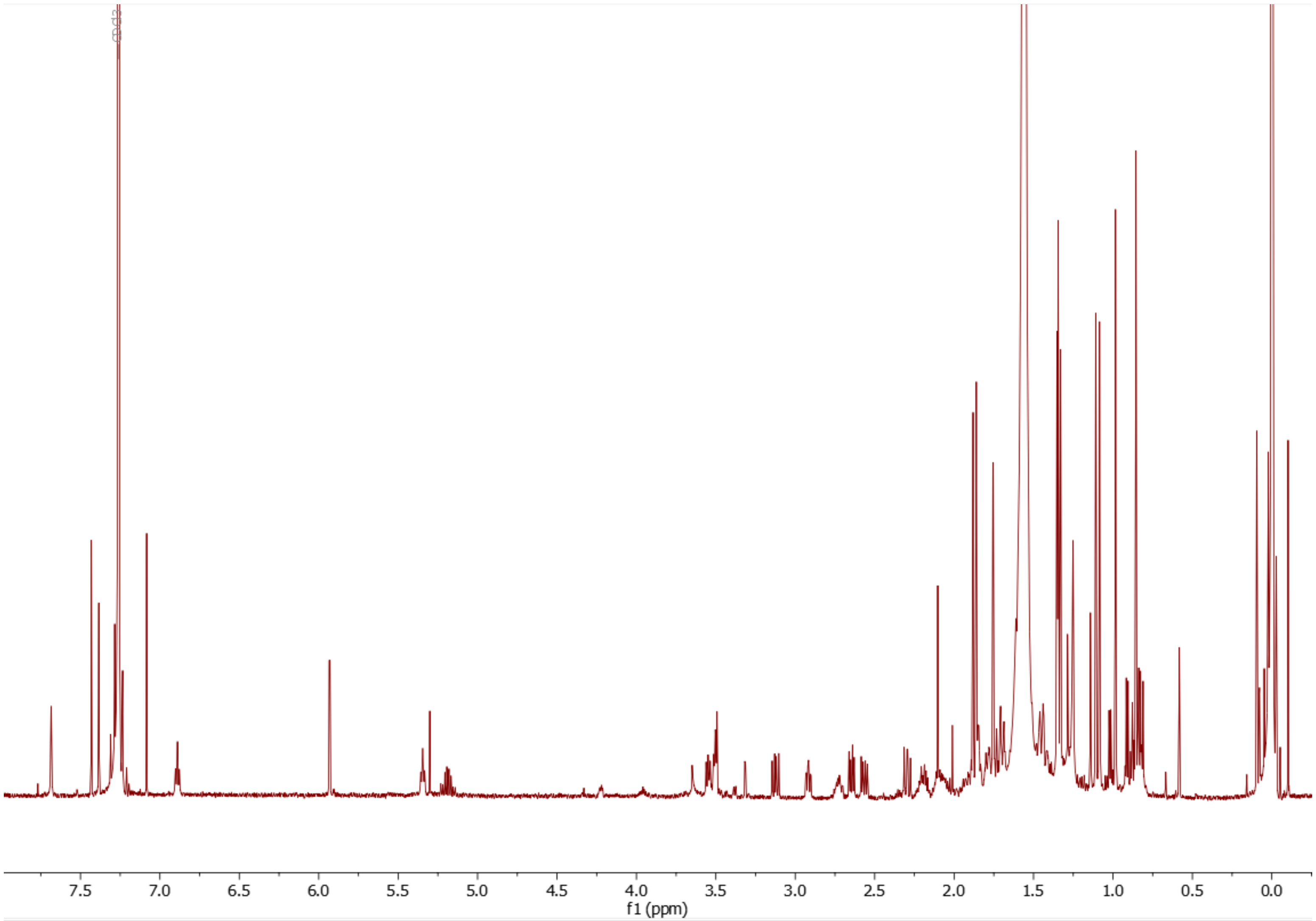


Figure S30. ^{1}H NMR Spectrum of DH-NAC$_4$

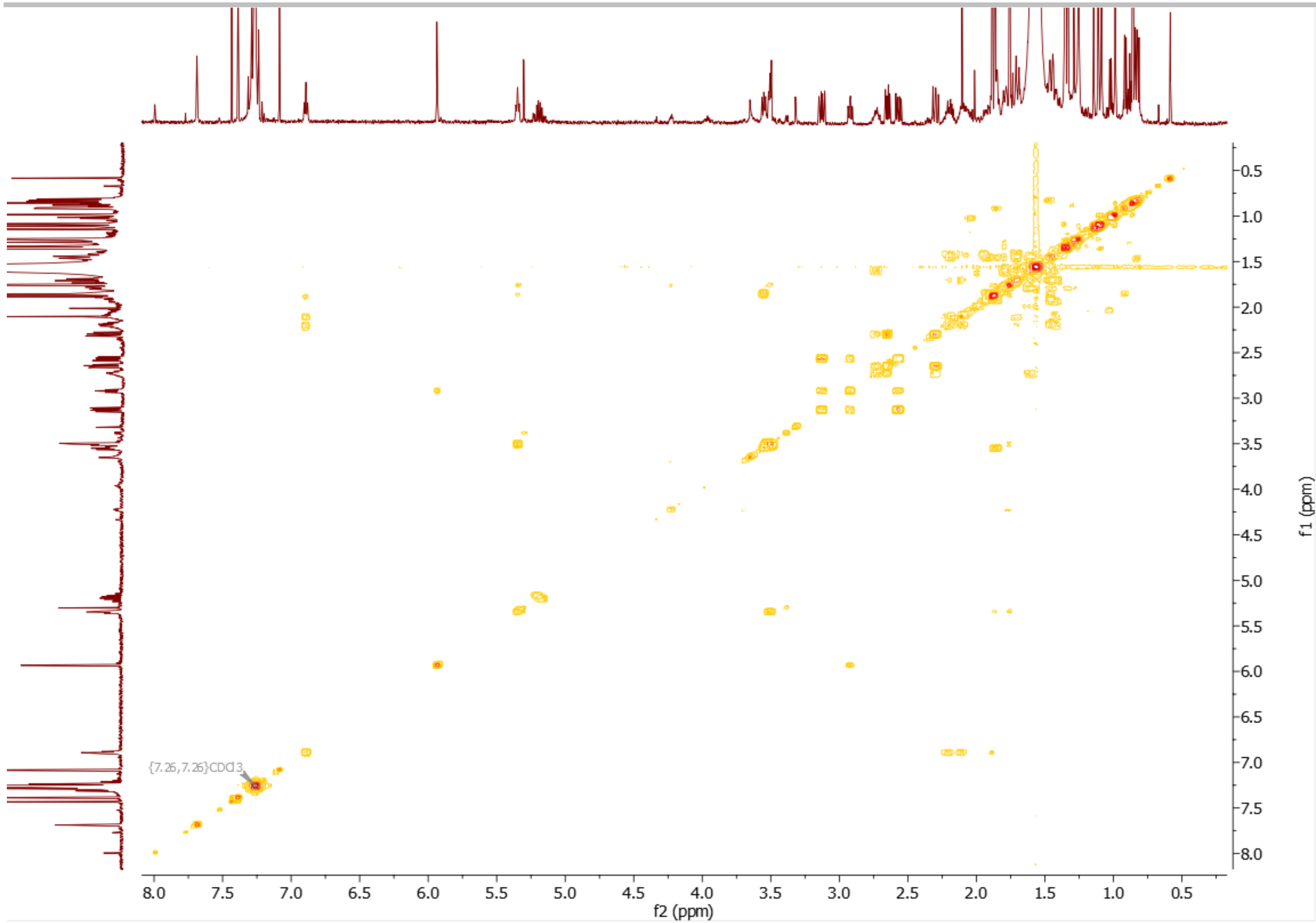


Figure S31. COSY Spectrum of DH-NAC4

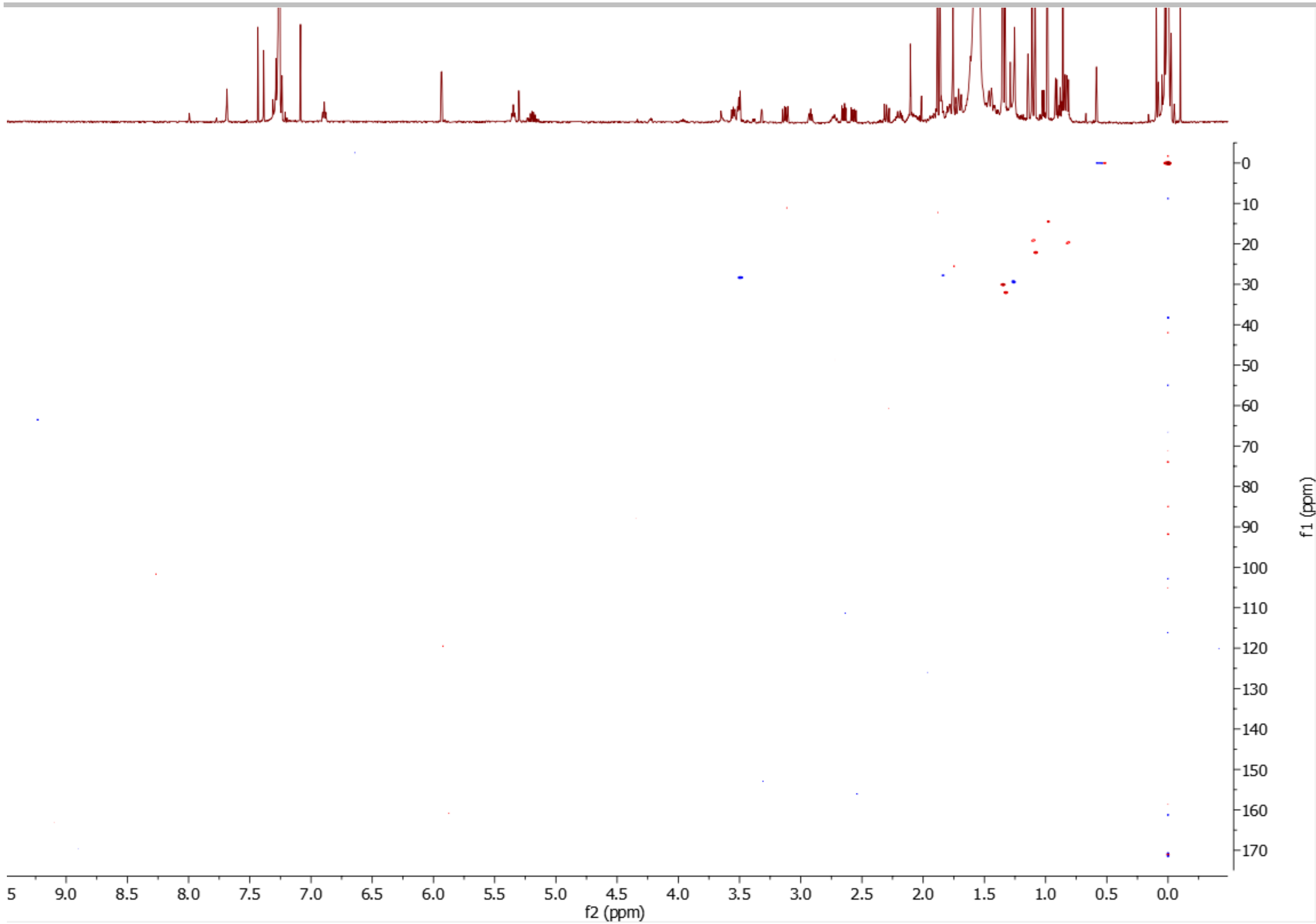


Figure S32. HSQC Spectrum of DH-NAC4

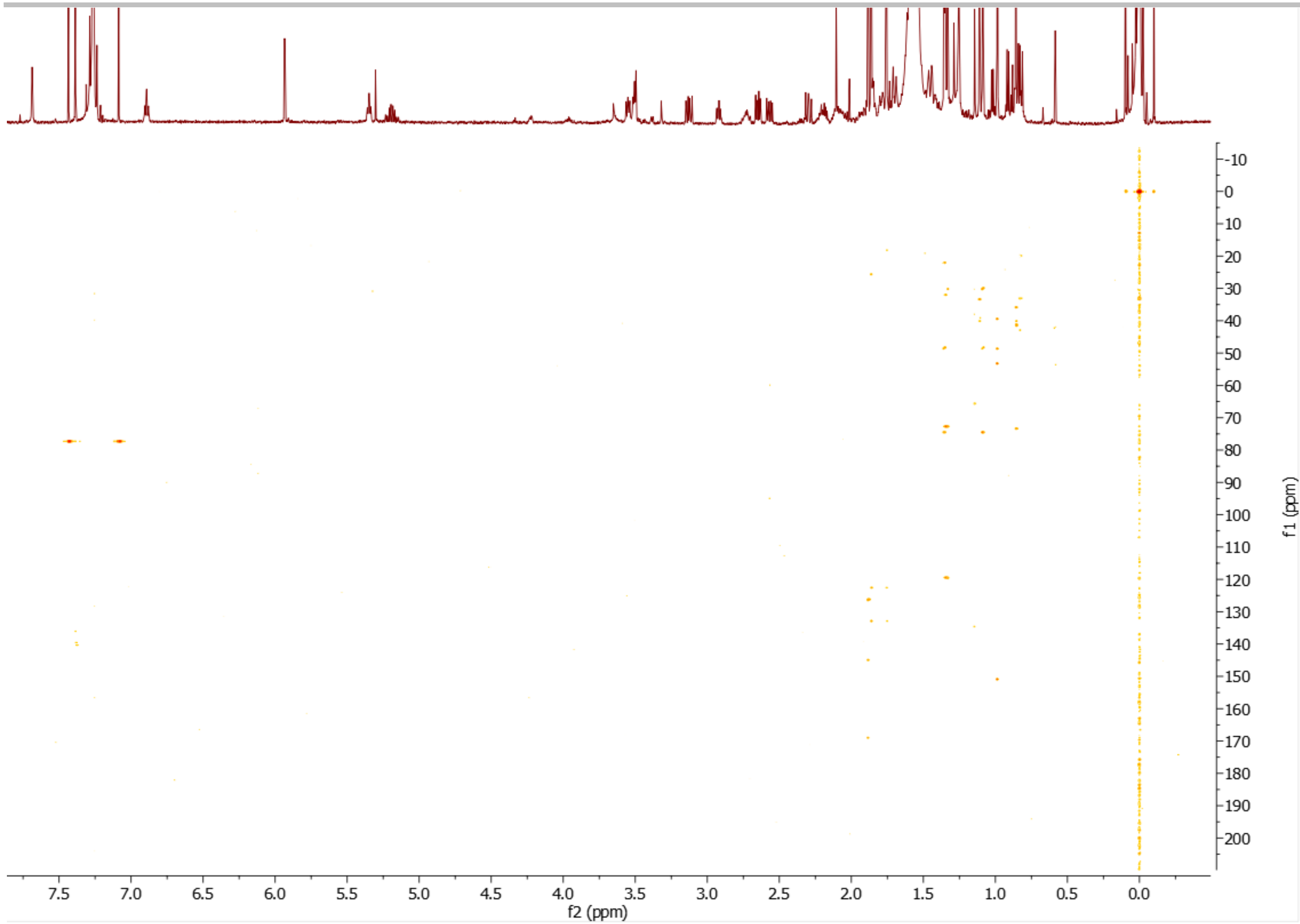


Figure S33. HMBC Spectrum of DH-NAC4

NAB$_4$ NMR Spectra

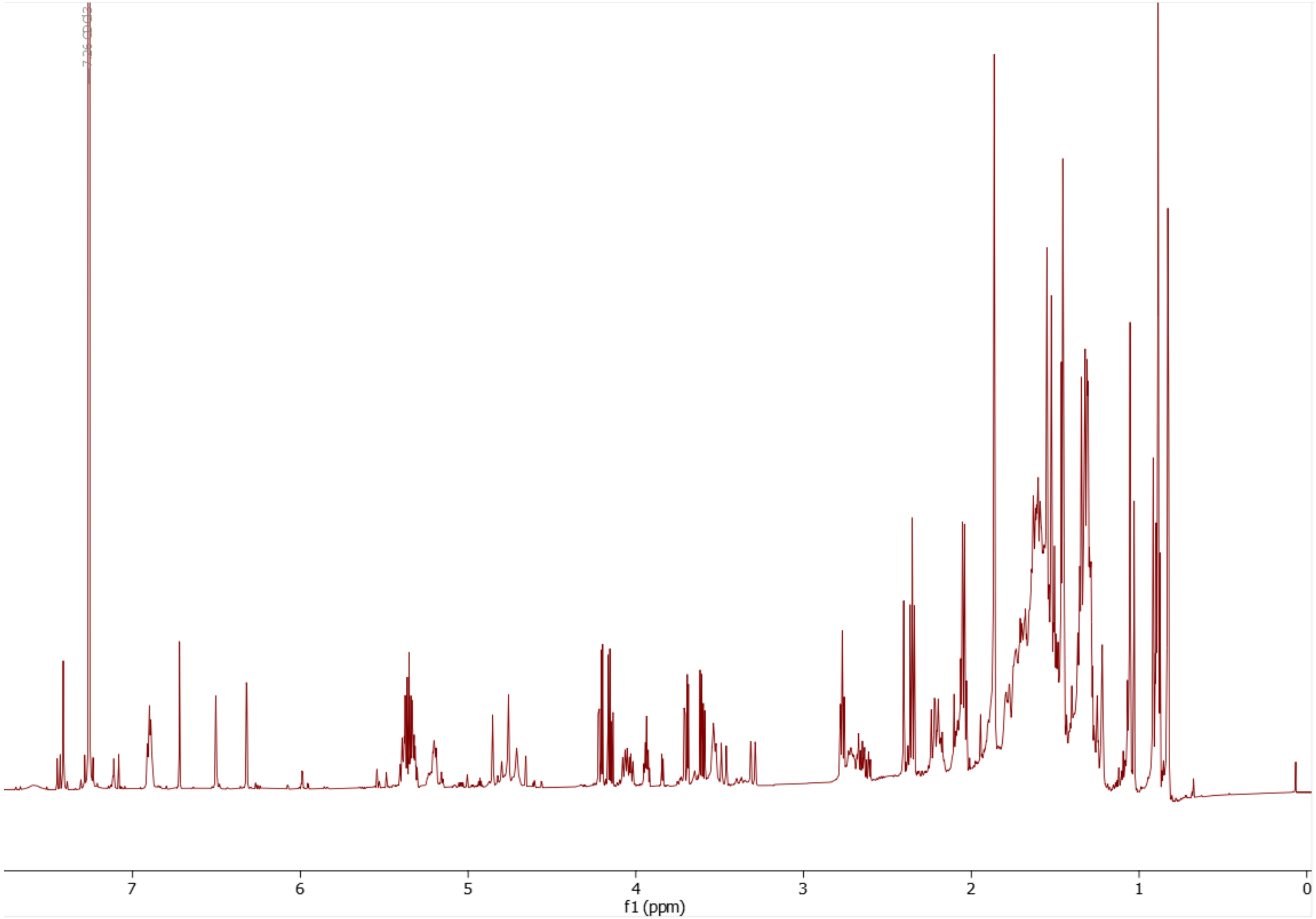


Figure S34. ^{1}H NMR Spectrum of NAB$_4$

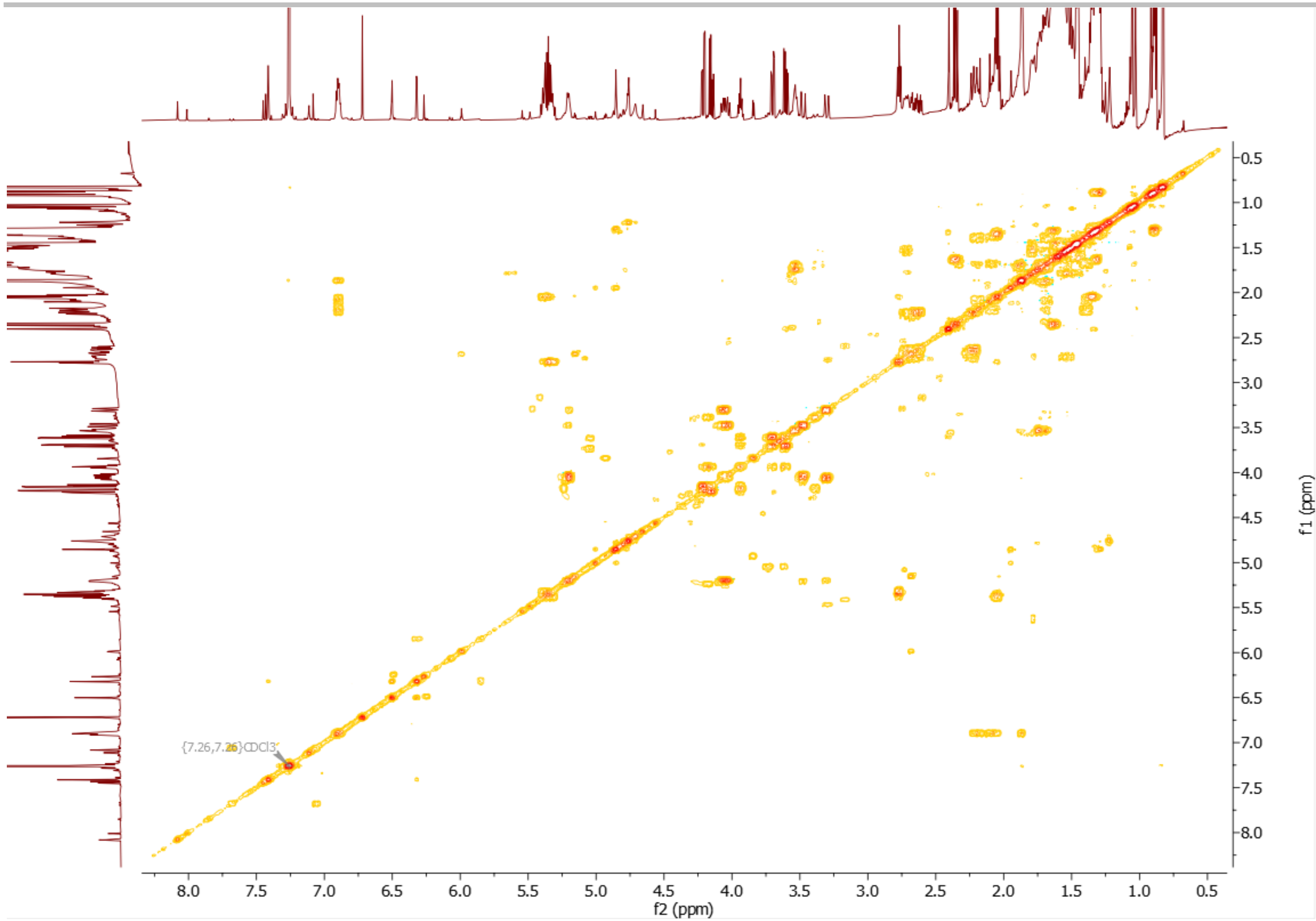


Figure S35. COSY NMR Spectrum of NAB_4

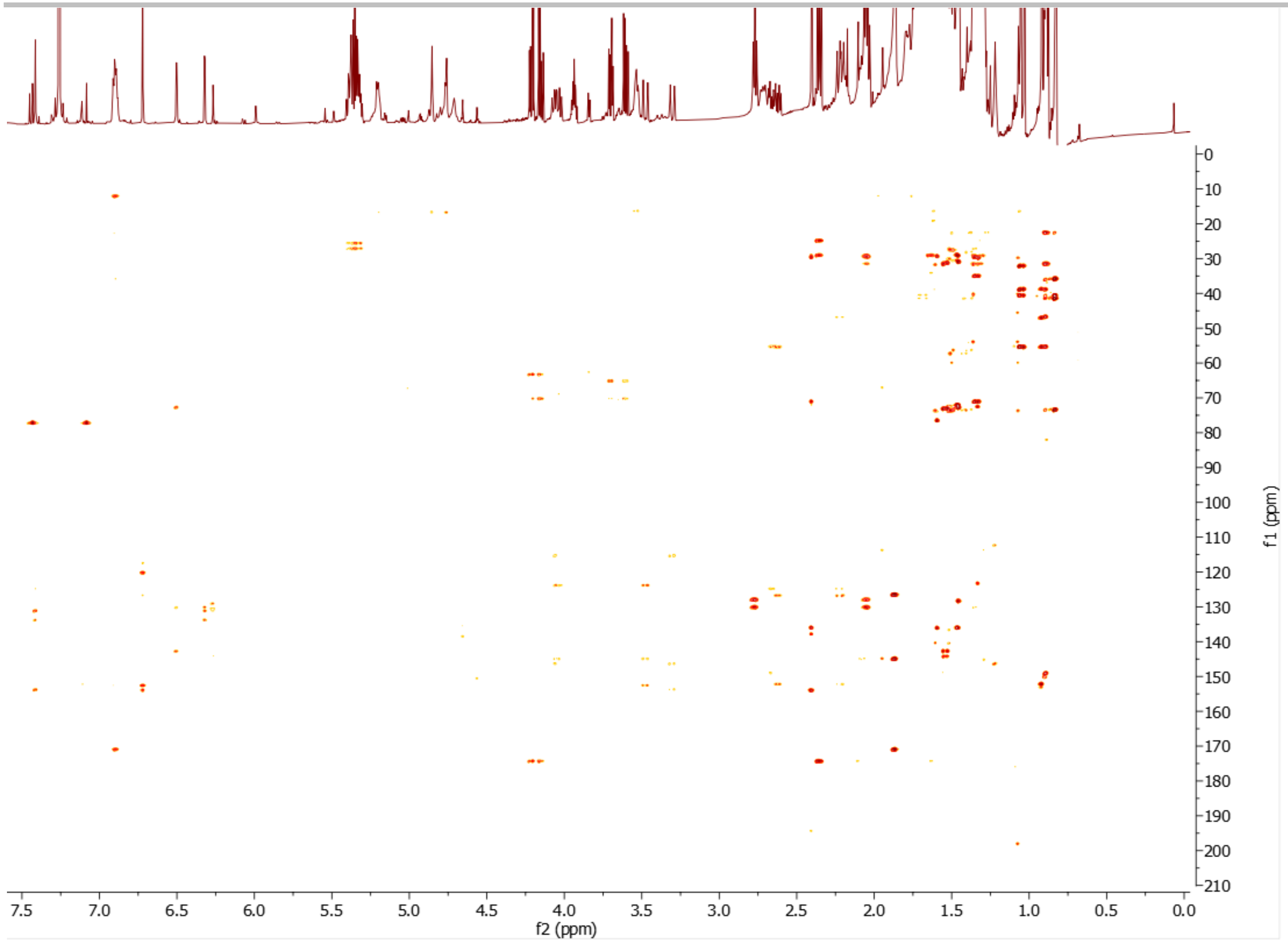


Figure S36. COSY NMR Spectrum of NAB_4

DH-NAA$_4$ NMR Spectra

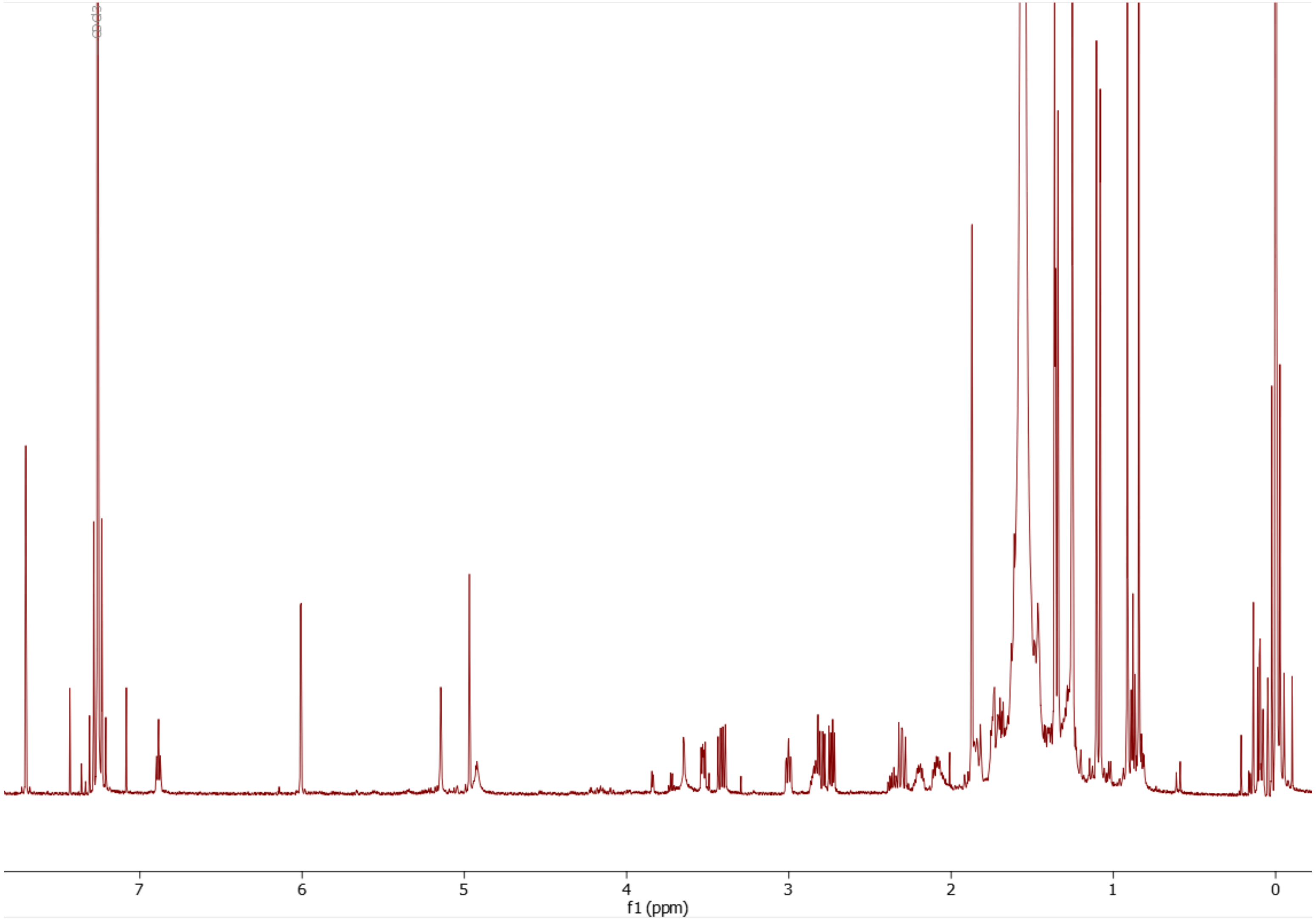


Figure S37. ^{1}H NMR Spectrum of DH-NAA$_4$

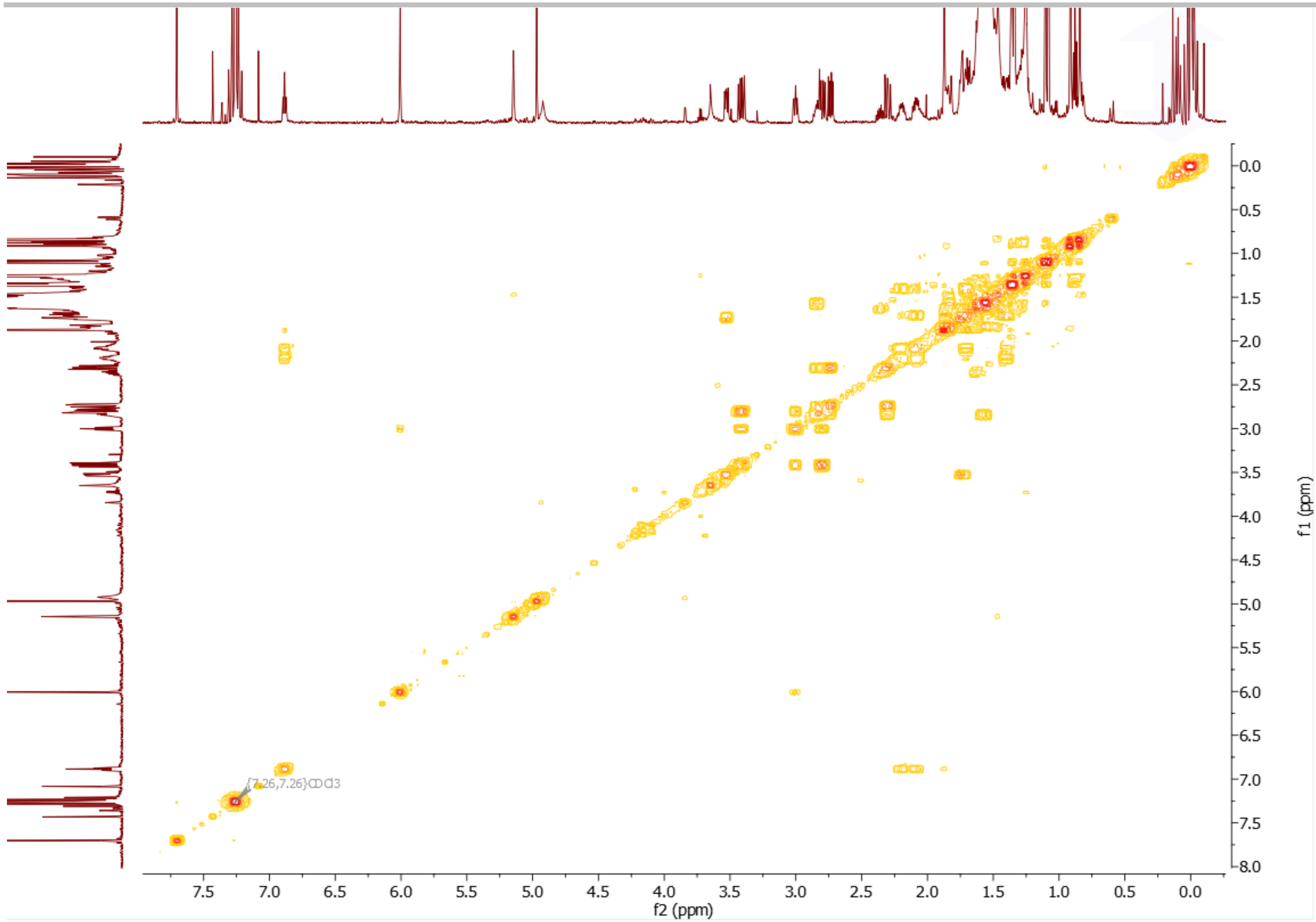


Figure S38. COSY Spectrum of DH-NAA_4

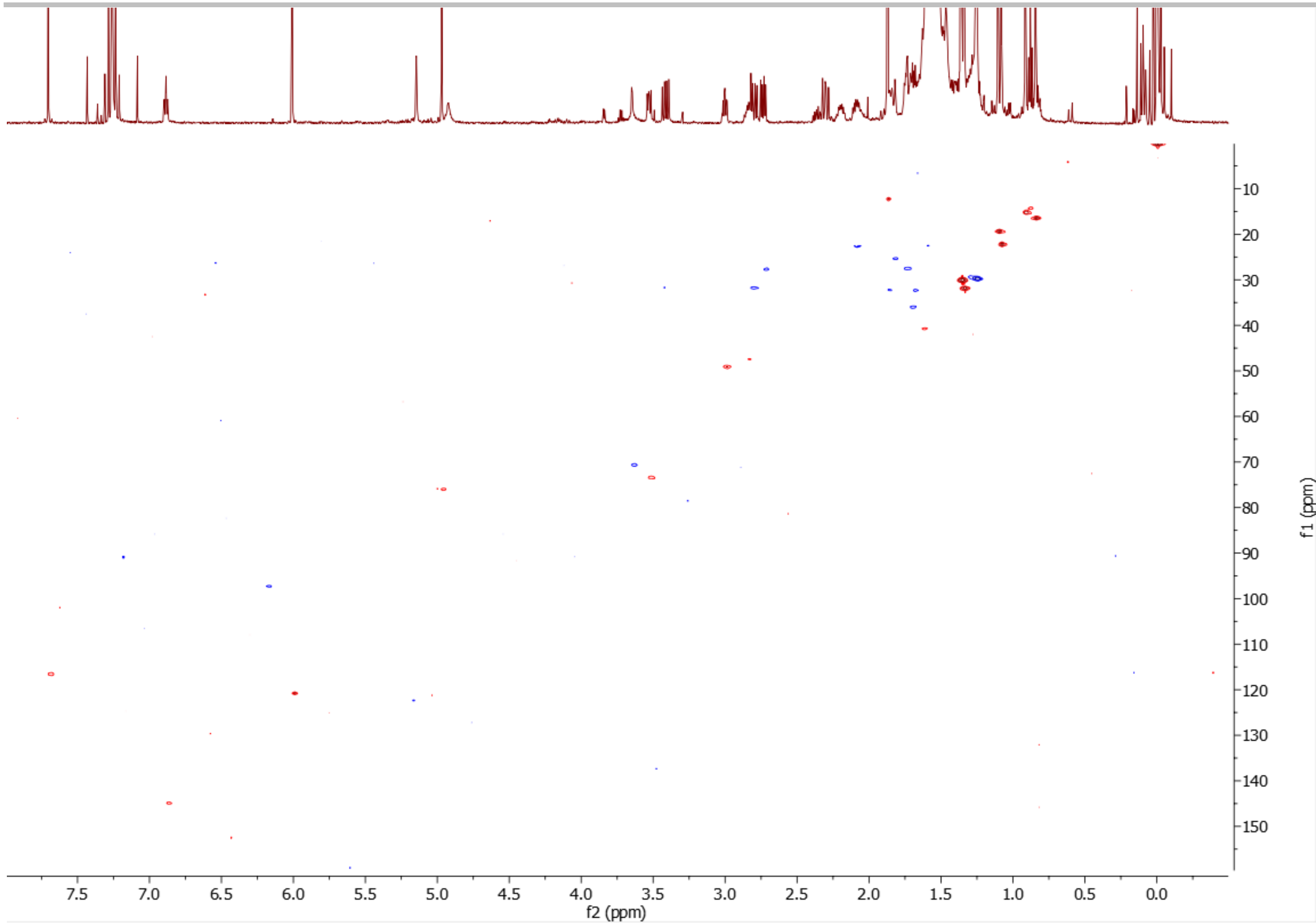


Figure S39. HSQC Spectrum of DH-NAA_4

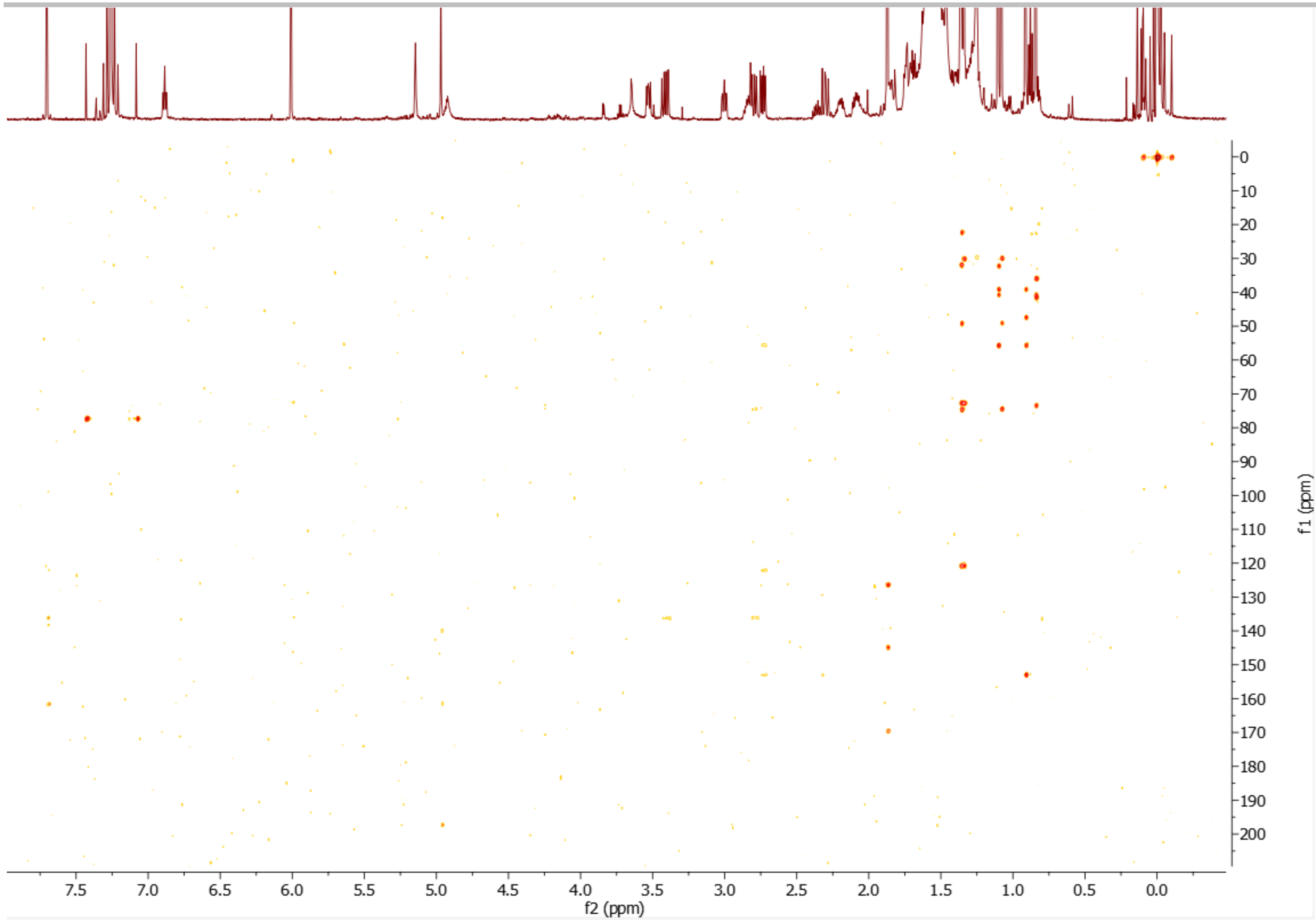


Figure S40. HMBC Spectrum of DH-NAA$_4$

NAA$_1$ NMR Spectra

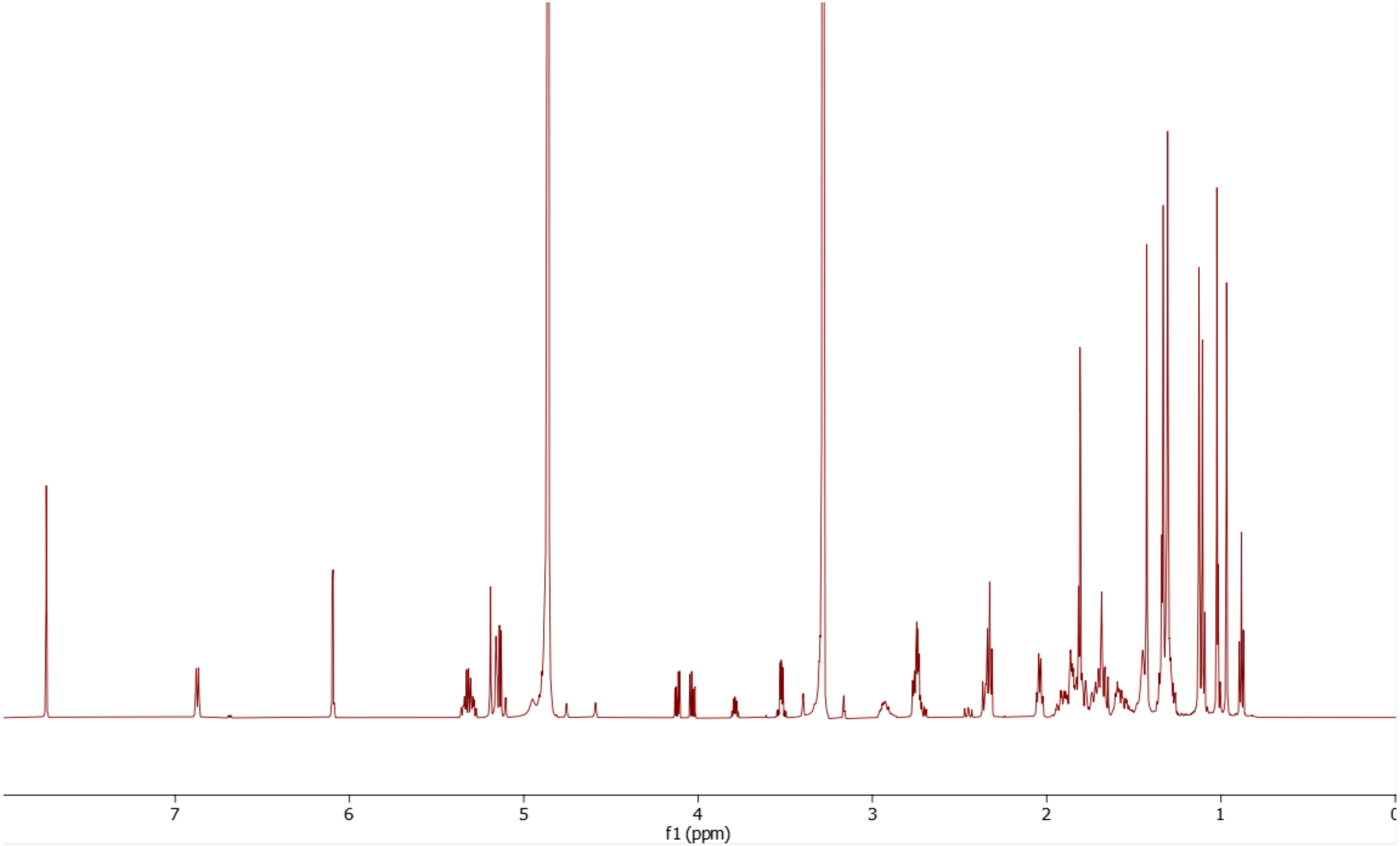


Figure S41. ^{1}H NMR spectrum of NAA$_1$

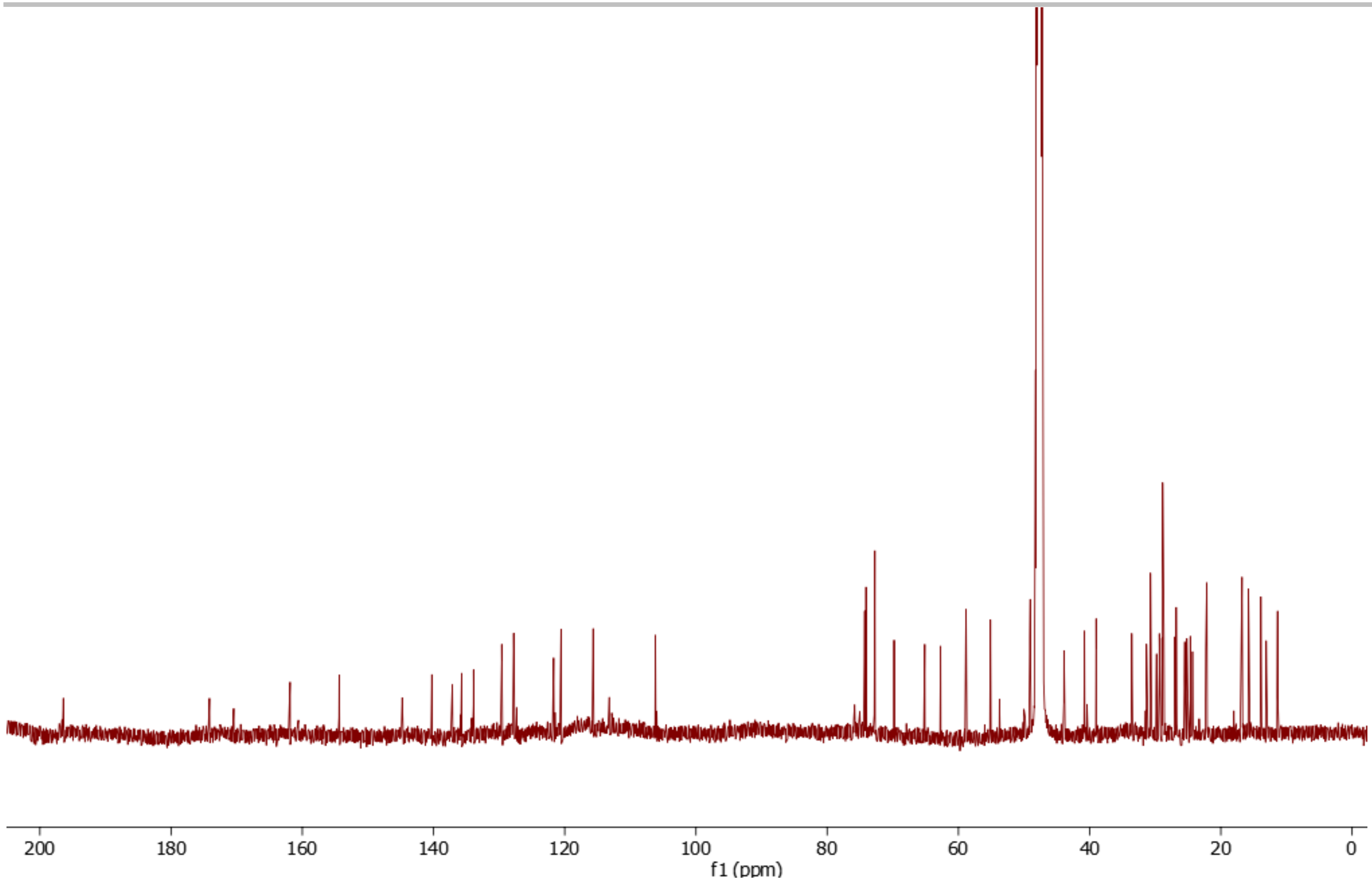


Figure S42. ^{13}C NMR spectrum of NAA_1

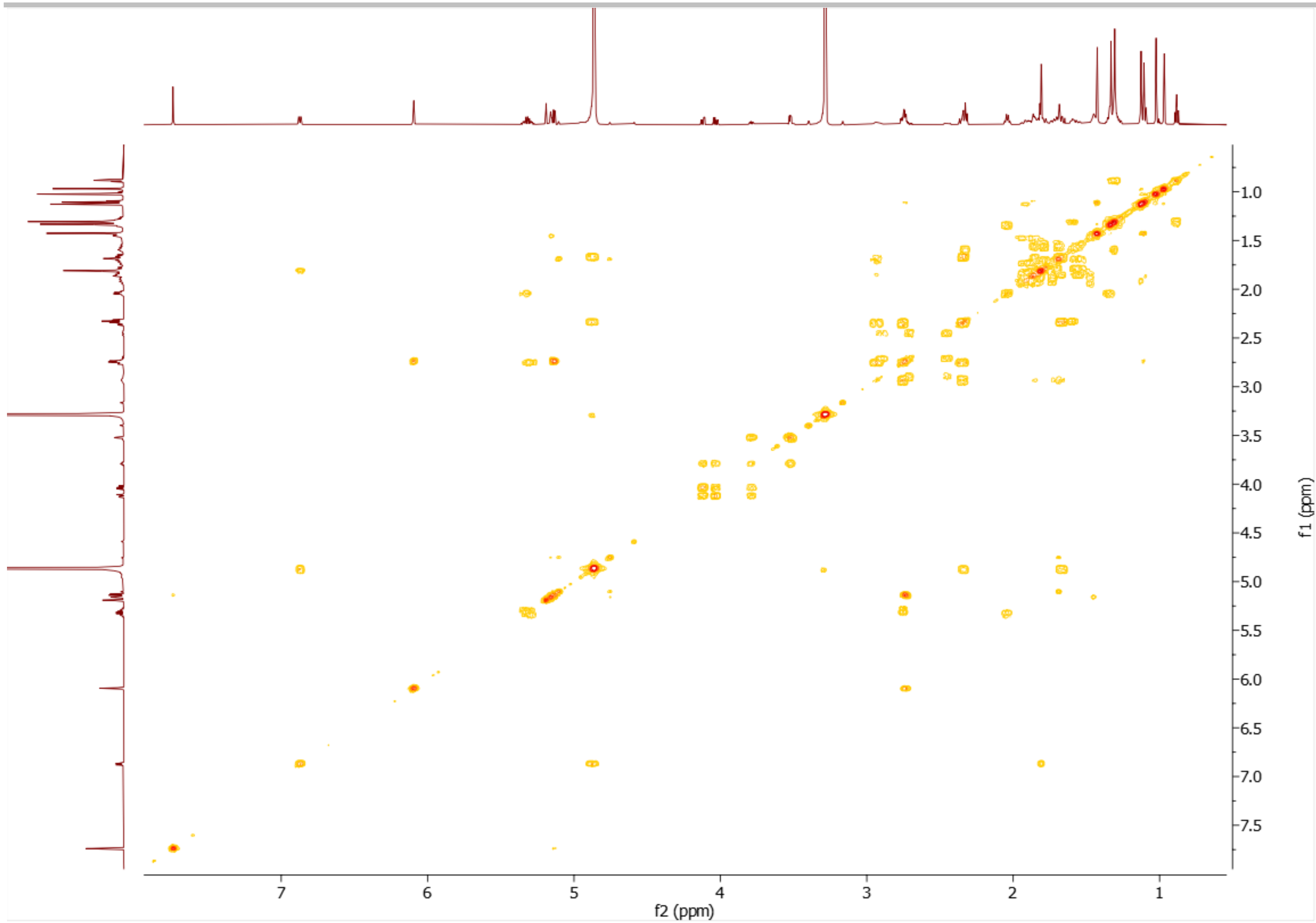


Figure S43. COSY spectrum of NAA_1

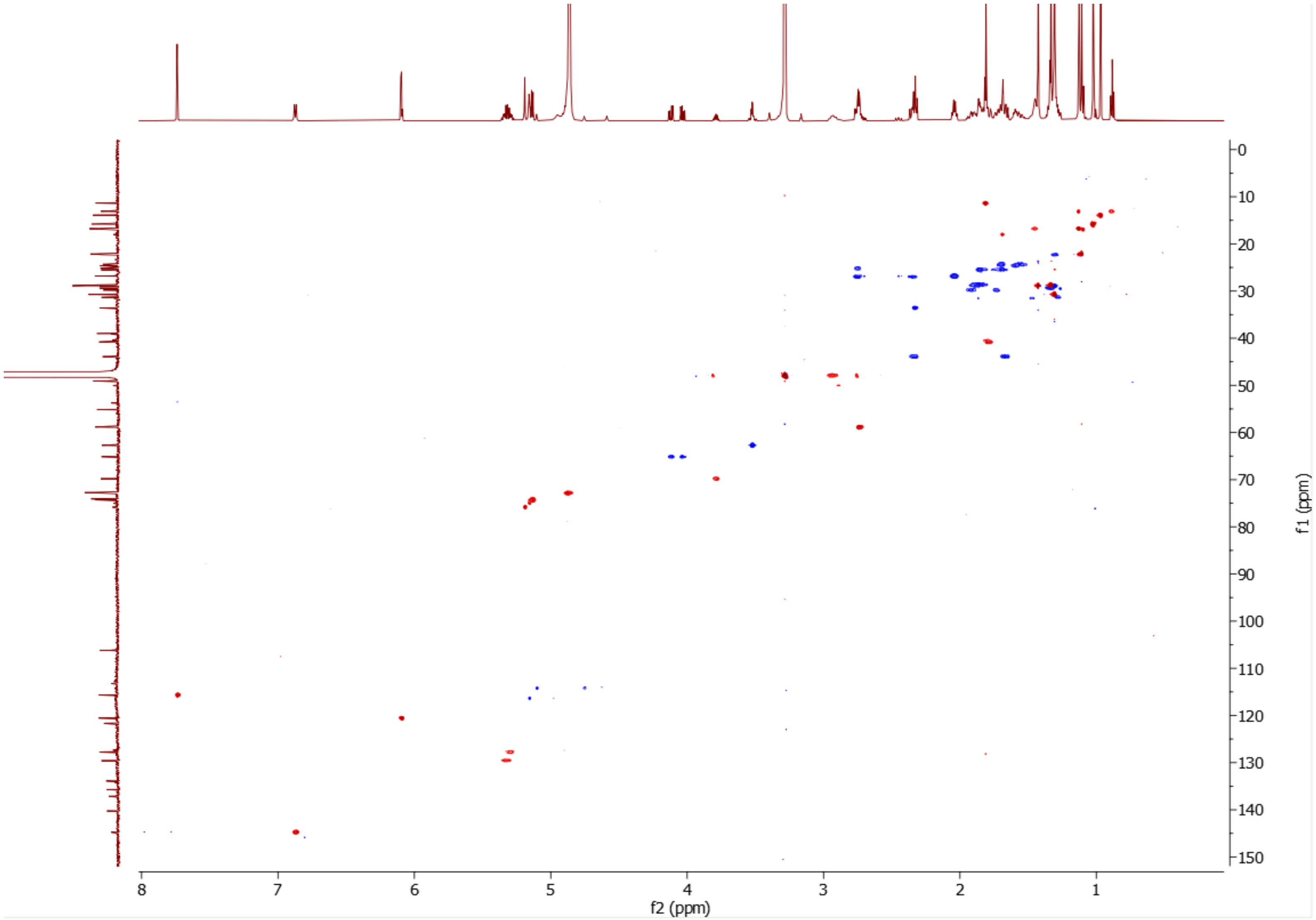


Figure S44. HSQC Spectrum of NAA_1

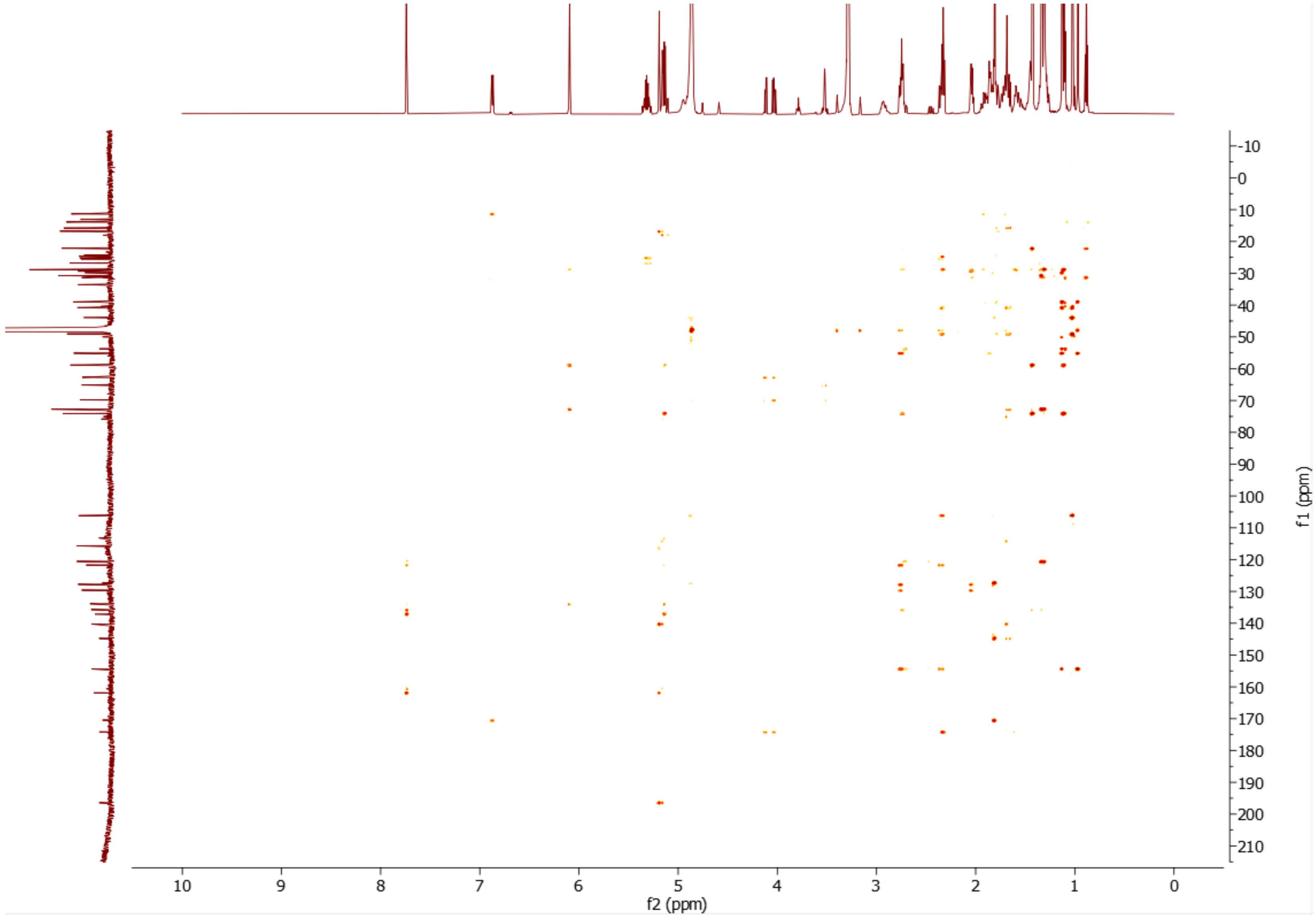


Figure S45. HMBC Spectrum of NAA_1

NAB NMR Spectra

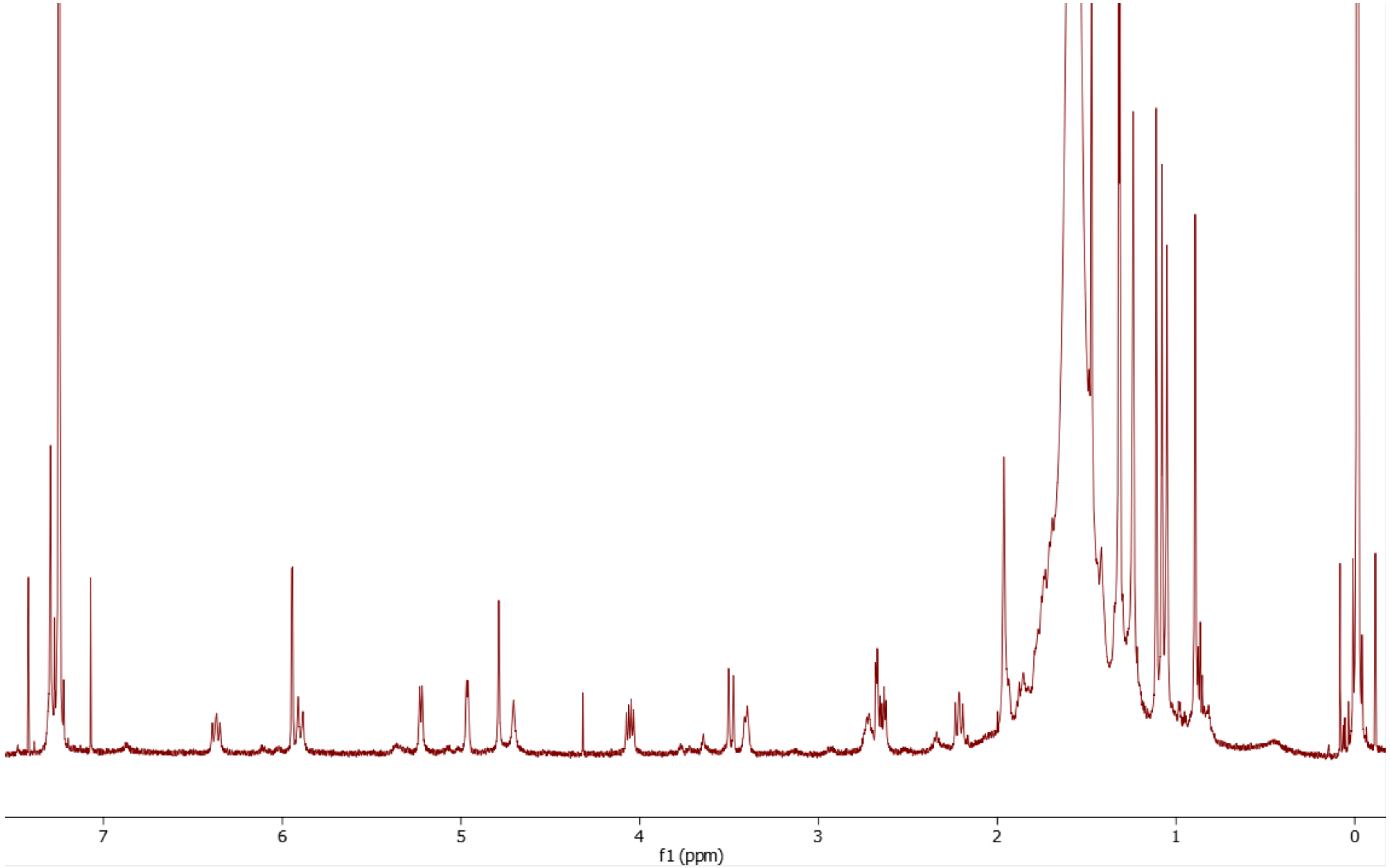


Figure S46. ^{1}H NMR Spectrum of NAB

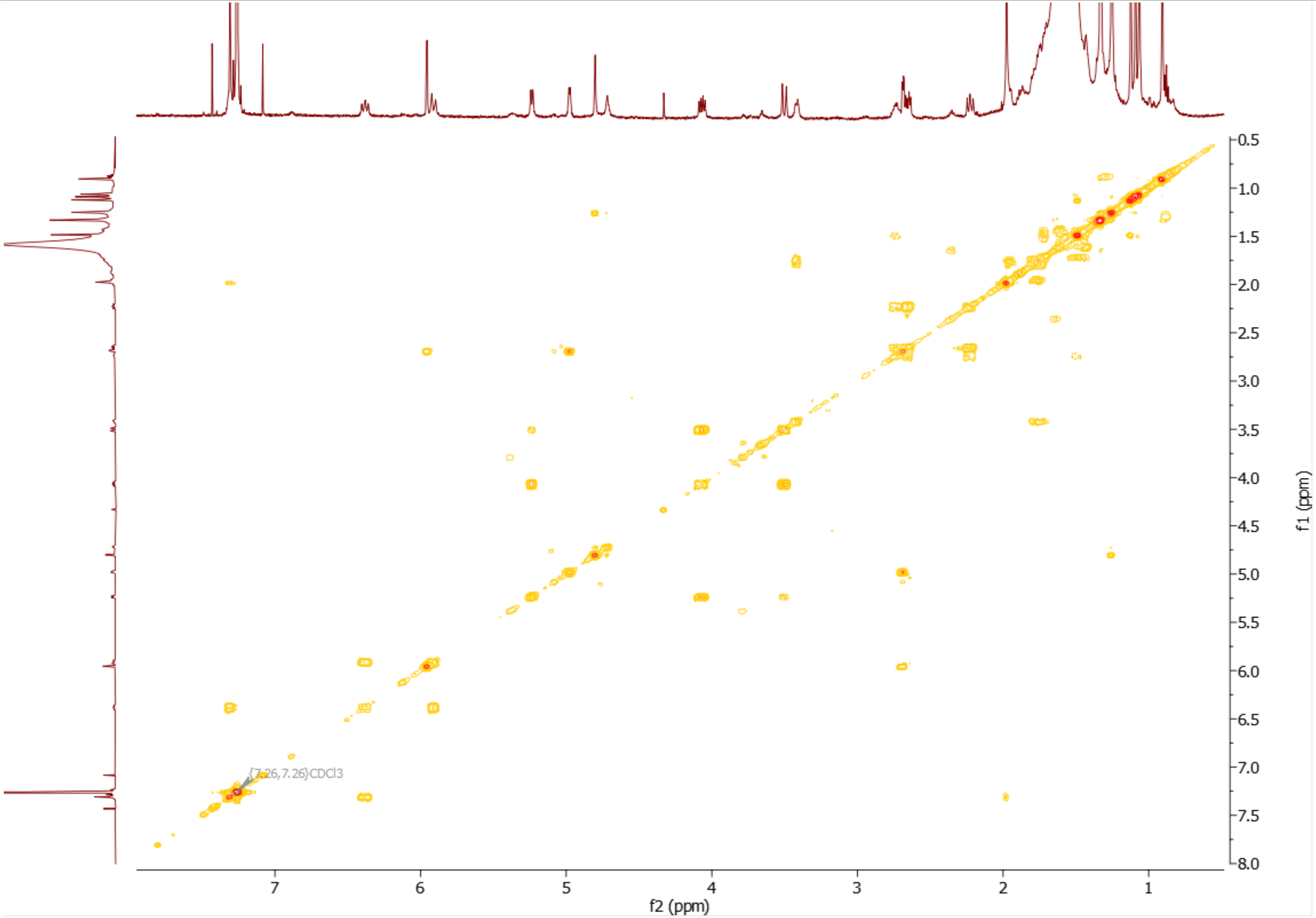


Figure S47. COSY Spectrum of NAB

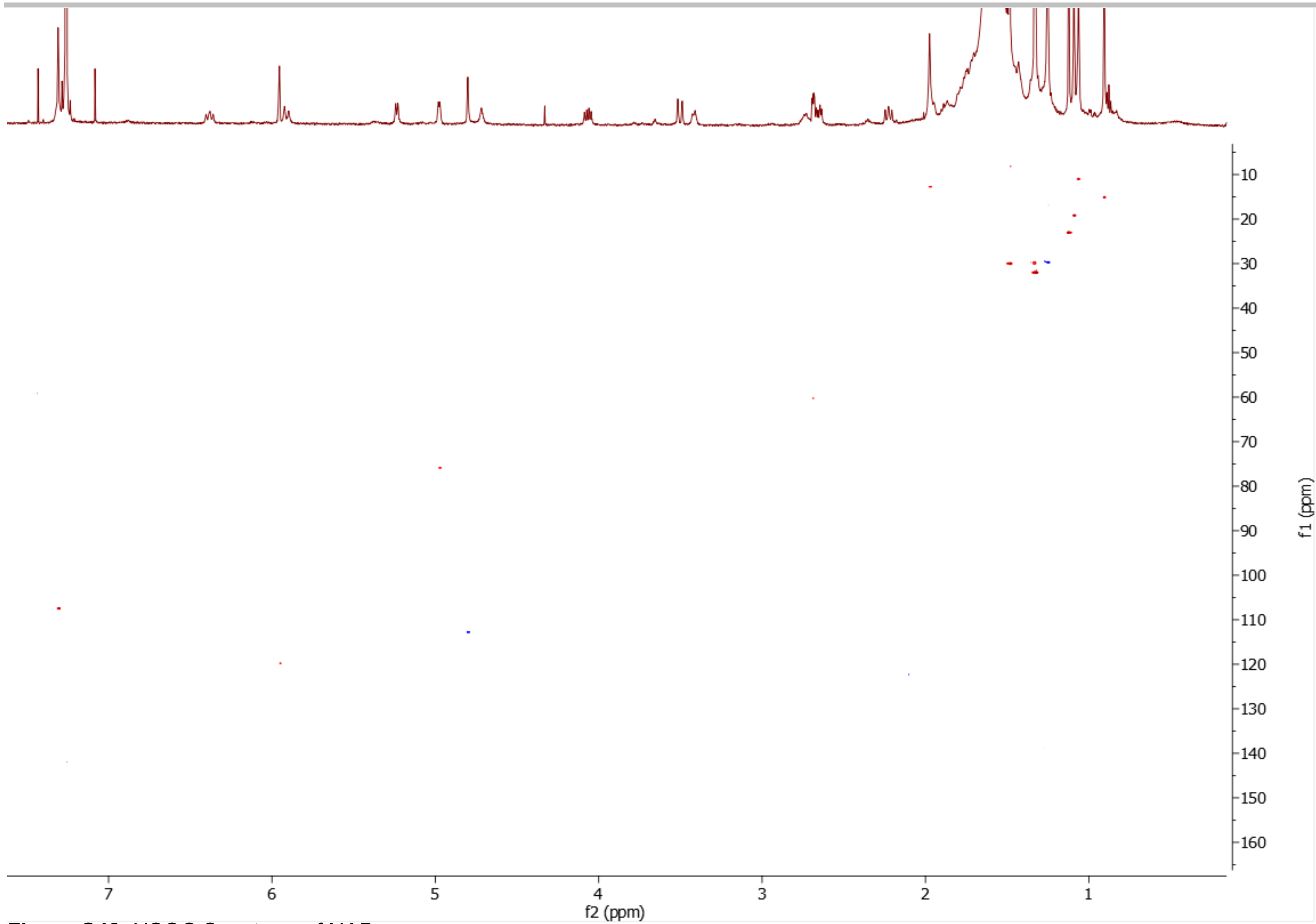


Figure S48. HSQC Spectrum of NAB

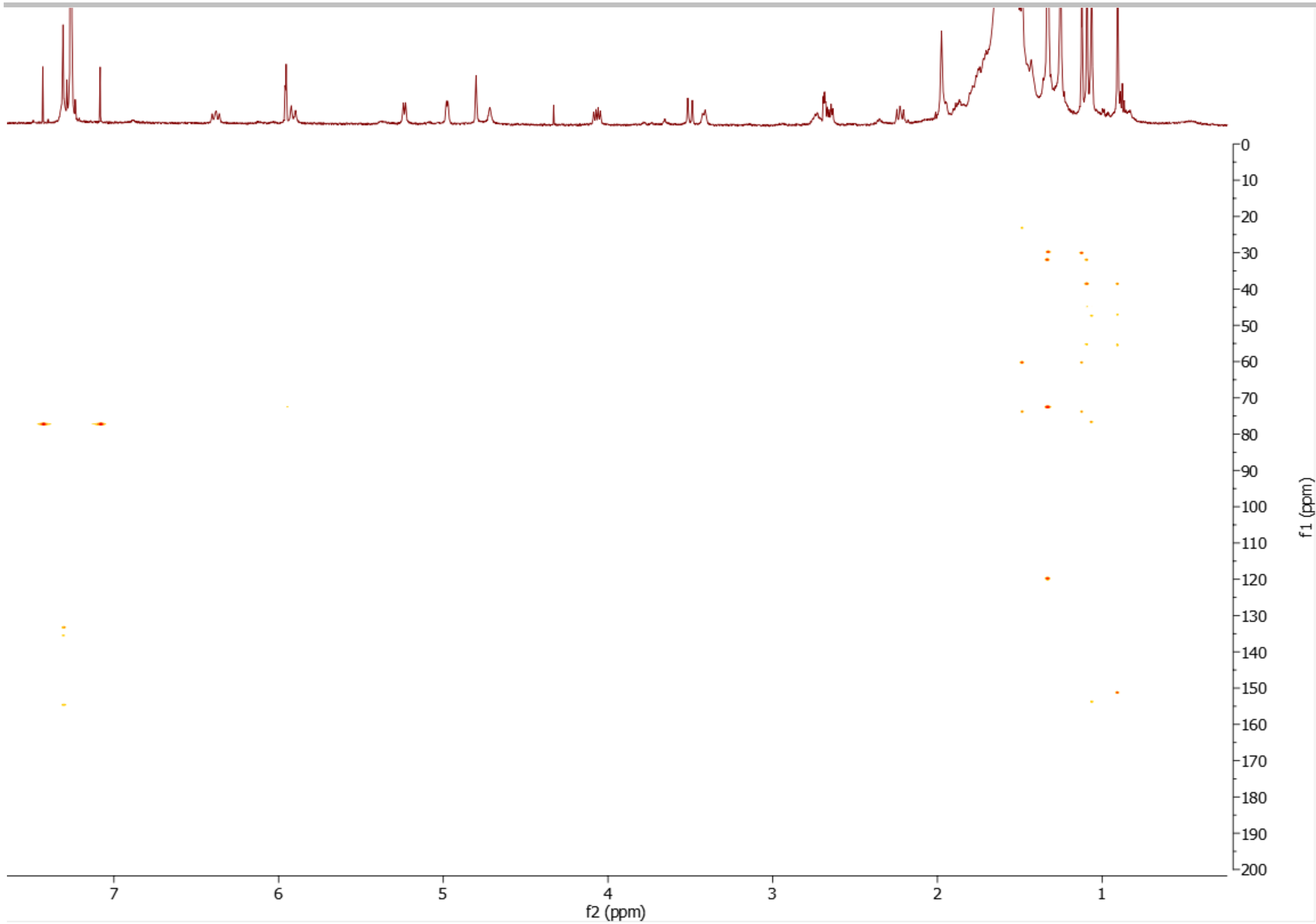


Figure S49. HMBC Spectrum of NAB

NAD6 NMR Spectra

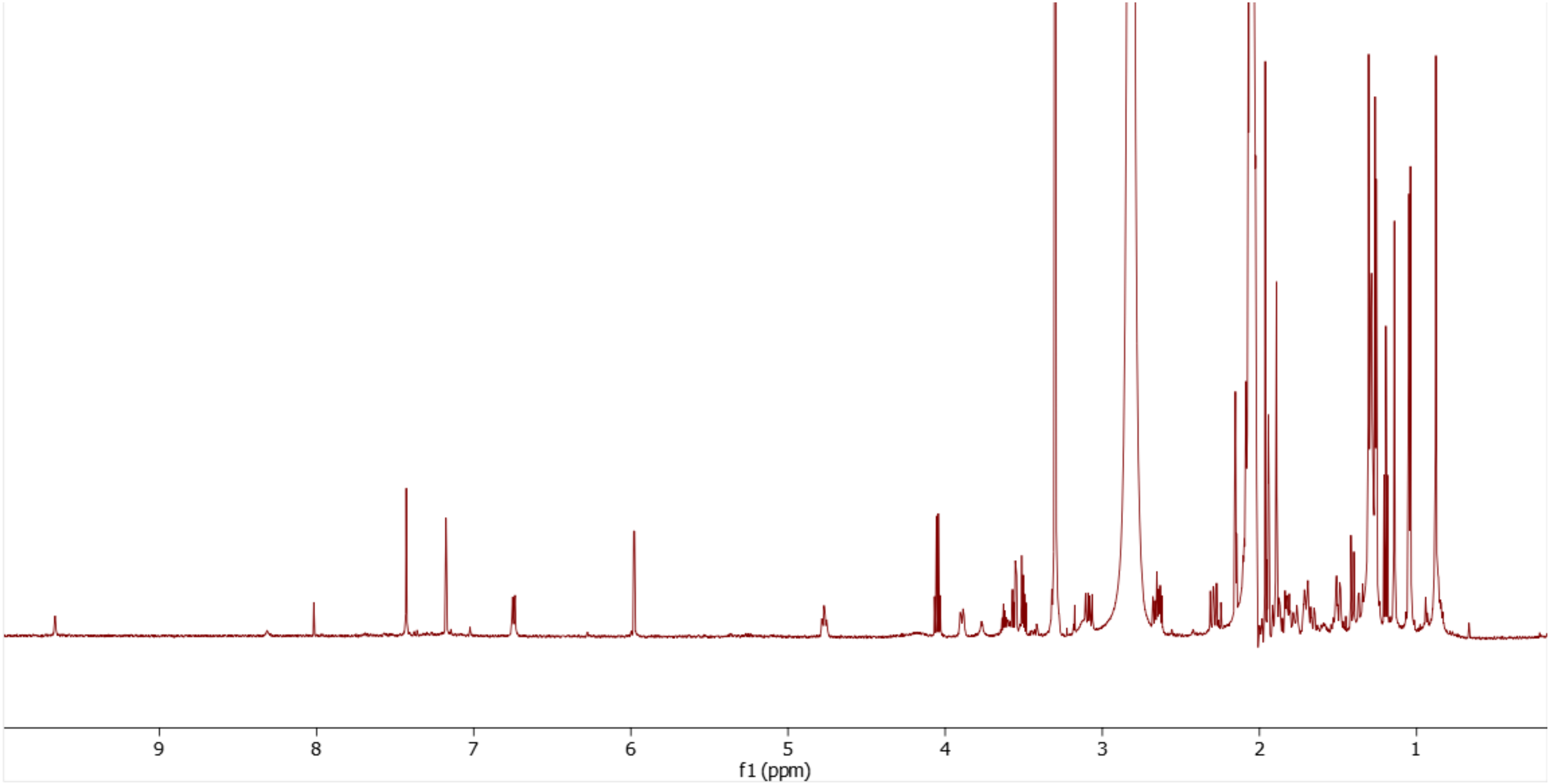


Figure S50. ^{1}H NMR Spectrum of NAD6

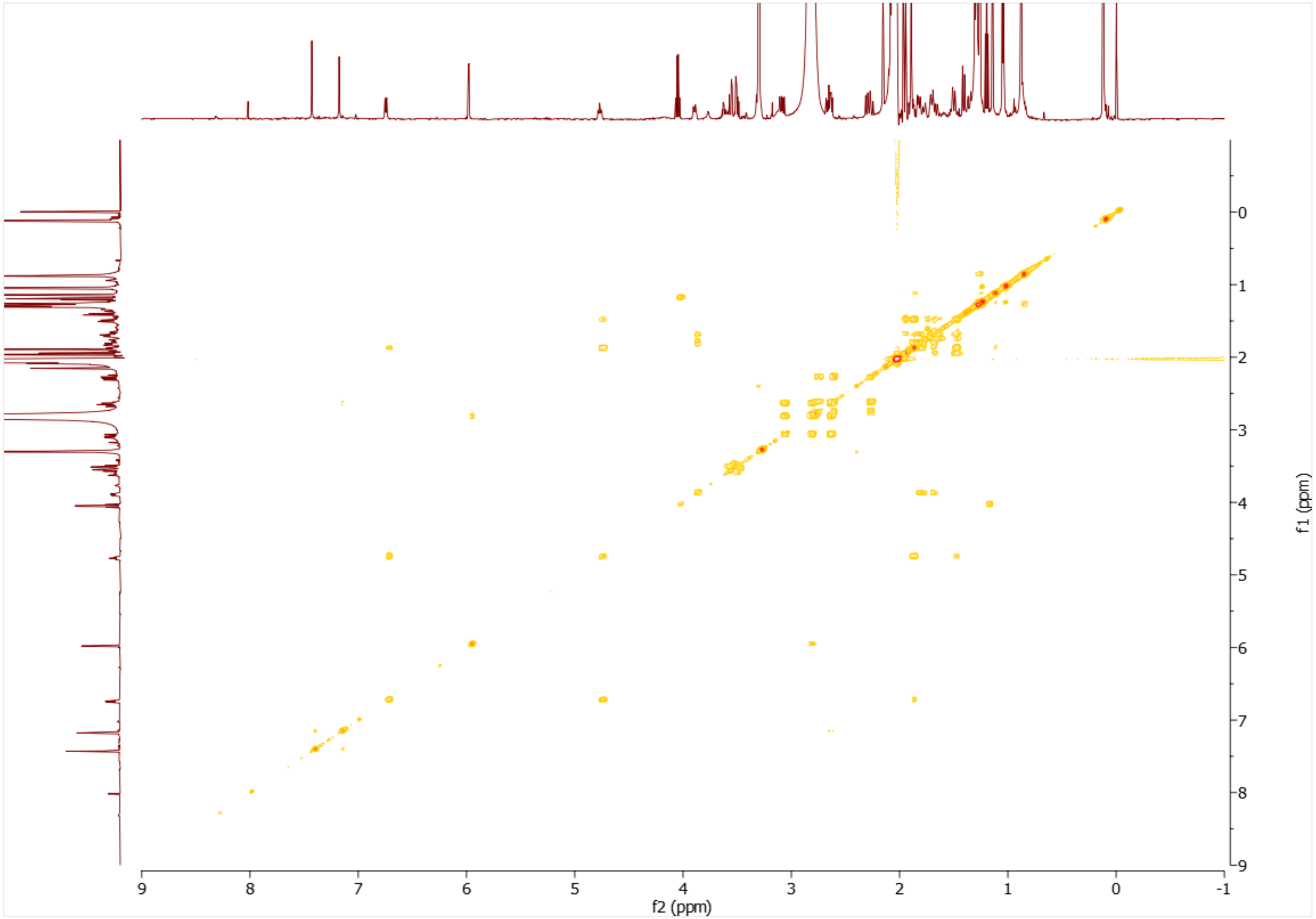


Figure S51. COSY Spectrum of NAD6

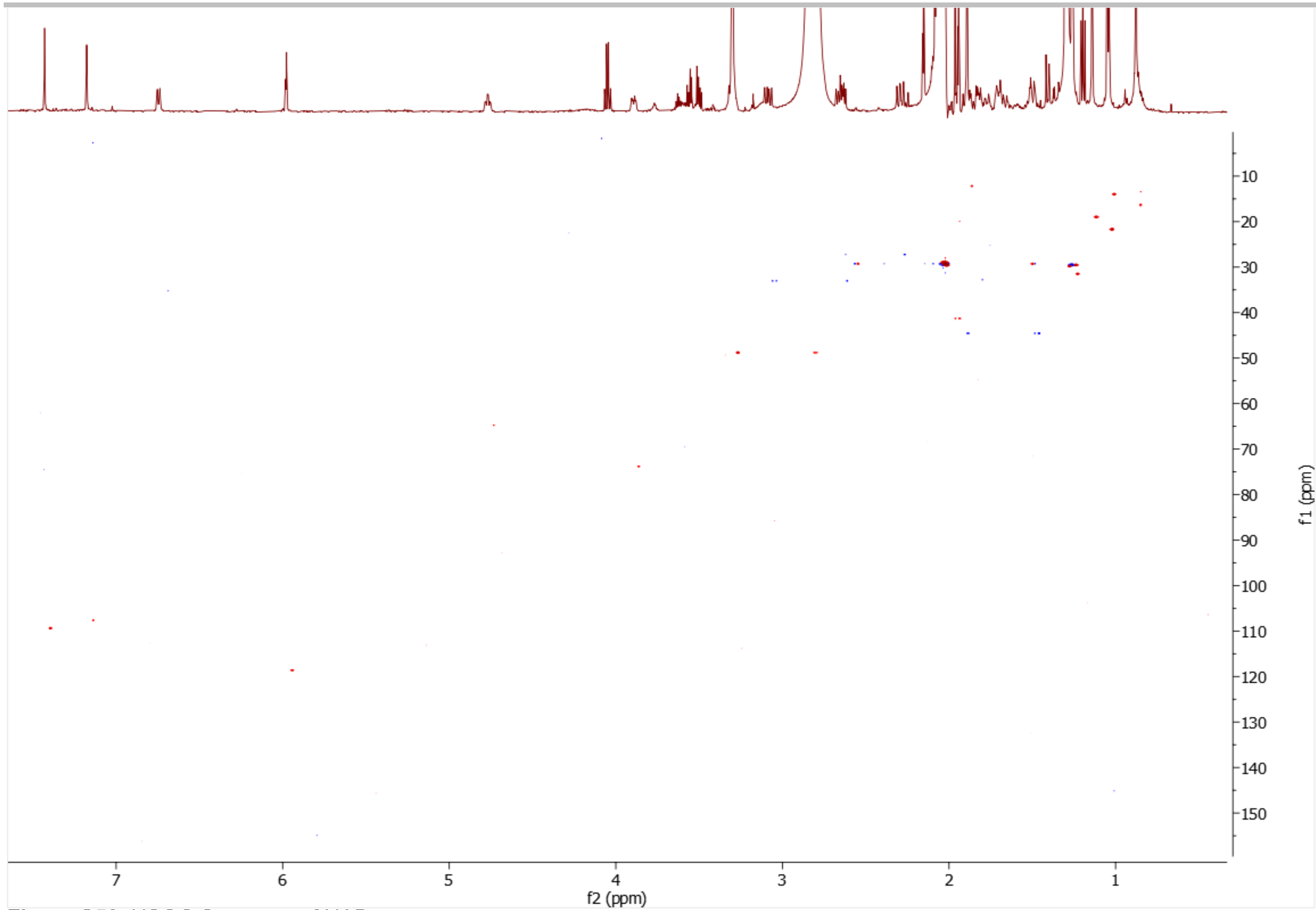


Figure S52. HSQC Spectrum of NAD_6

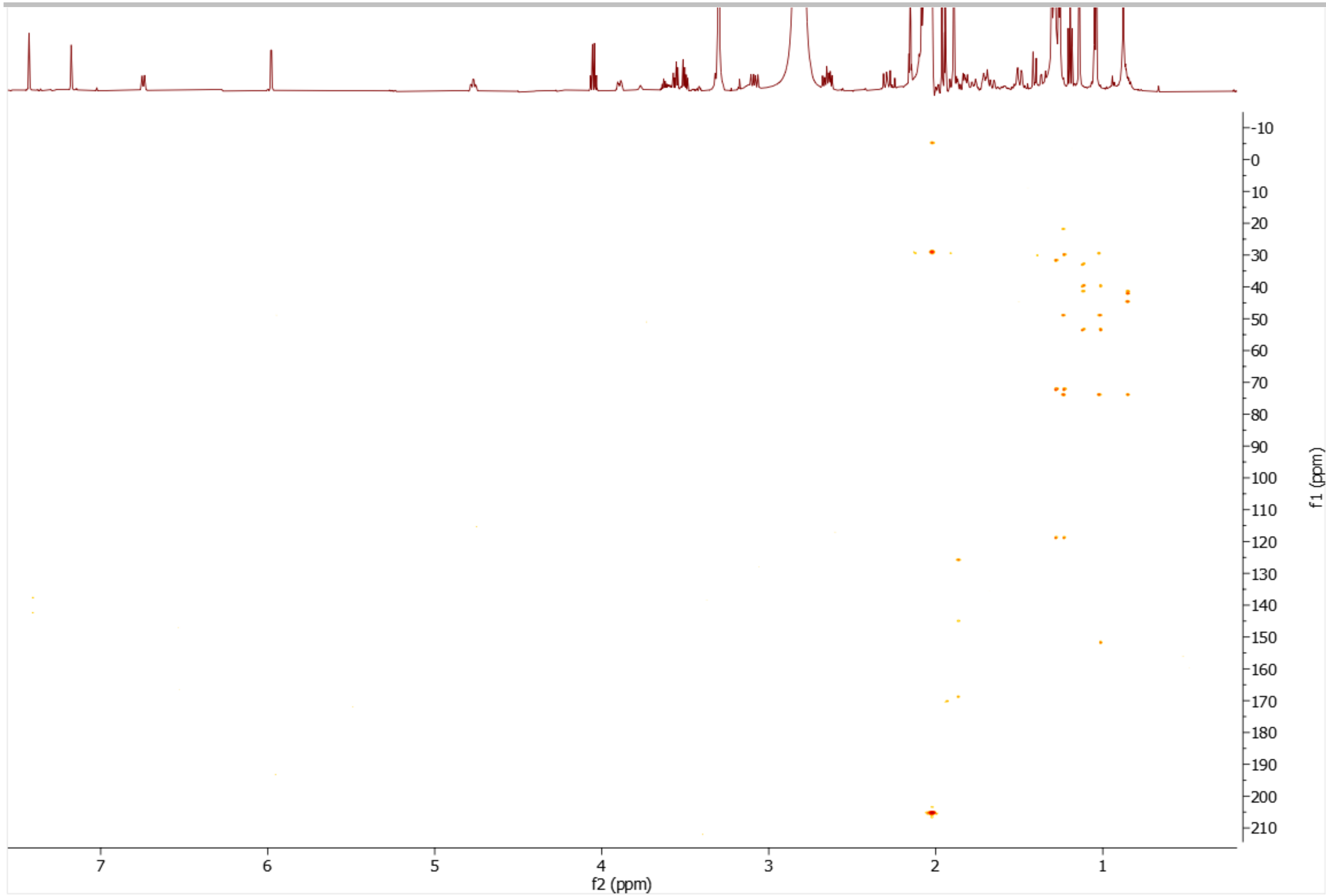


Figure S53. HMBC Spectrum of NAD_6

NAD$_5$ NMR Spectra

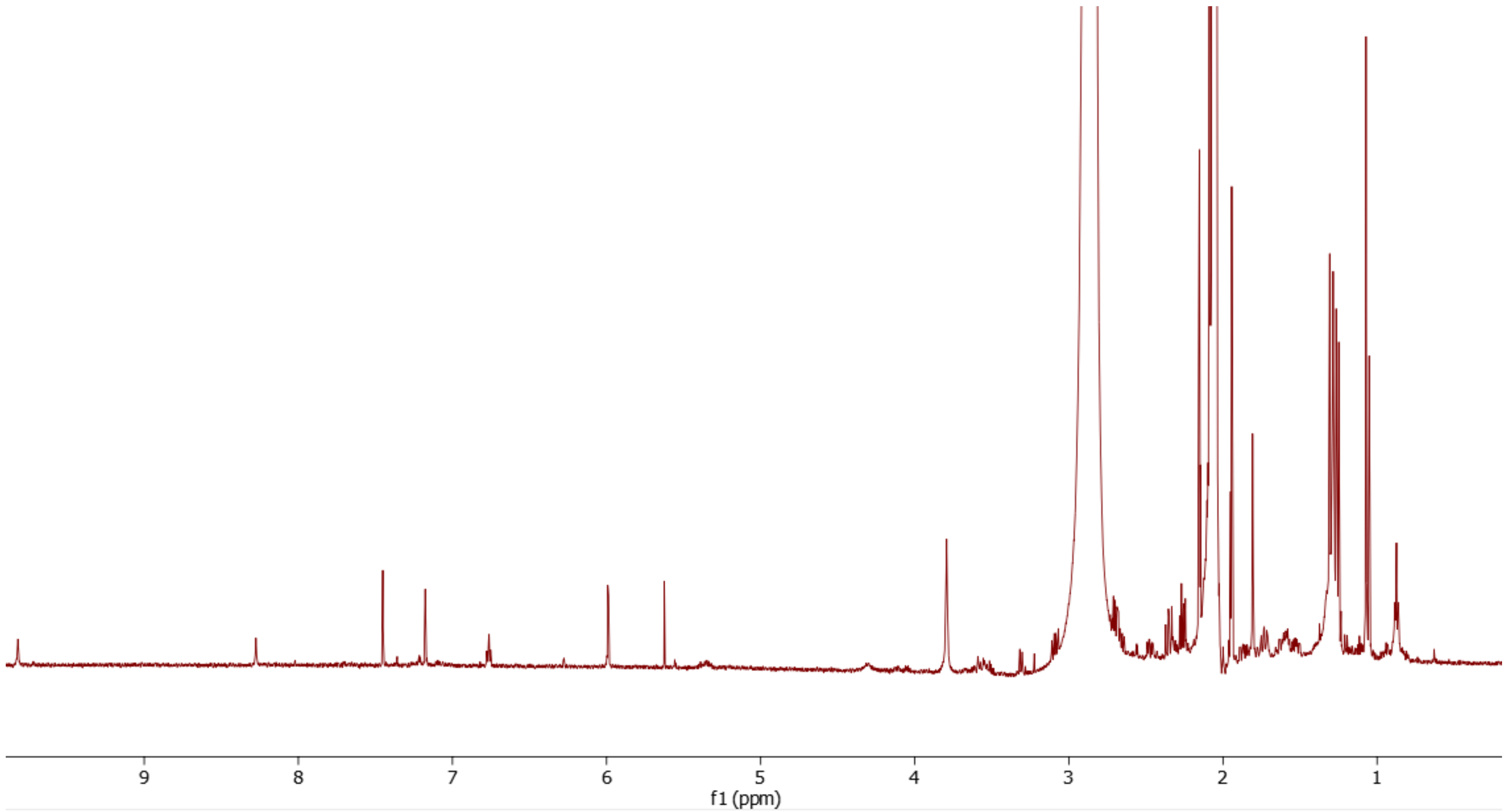


Figure S54. ^{1}H NMR Spectrum of NAD$_5$

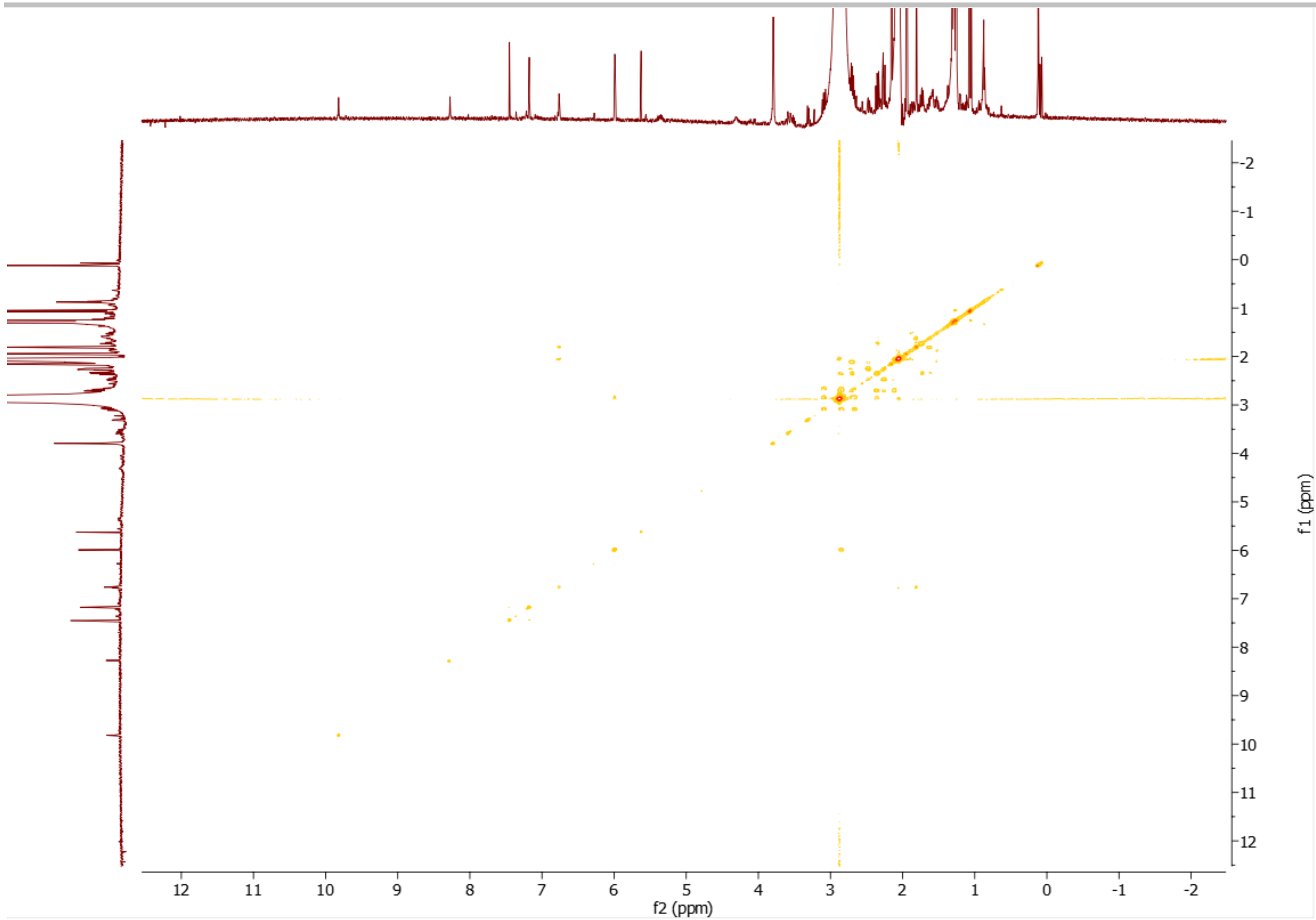


Figure S55. COSY Spectrum of NAD_5

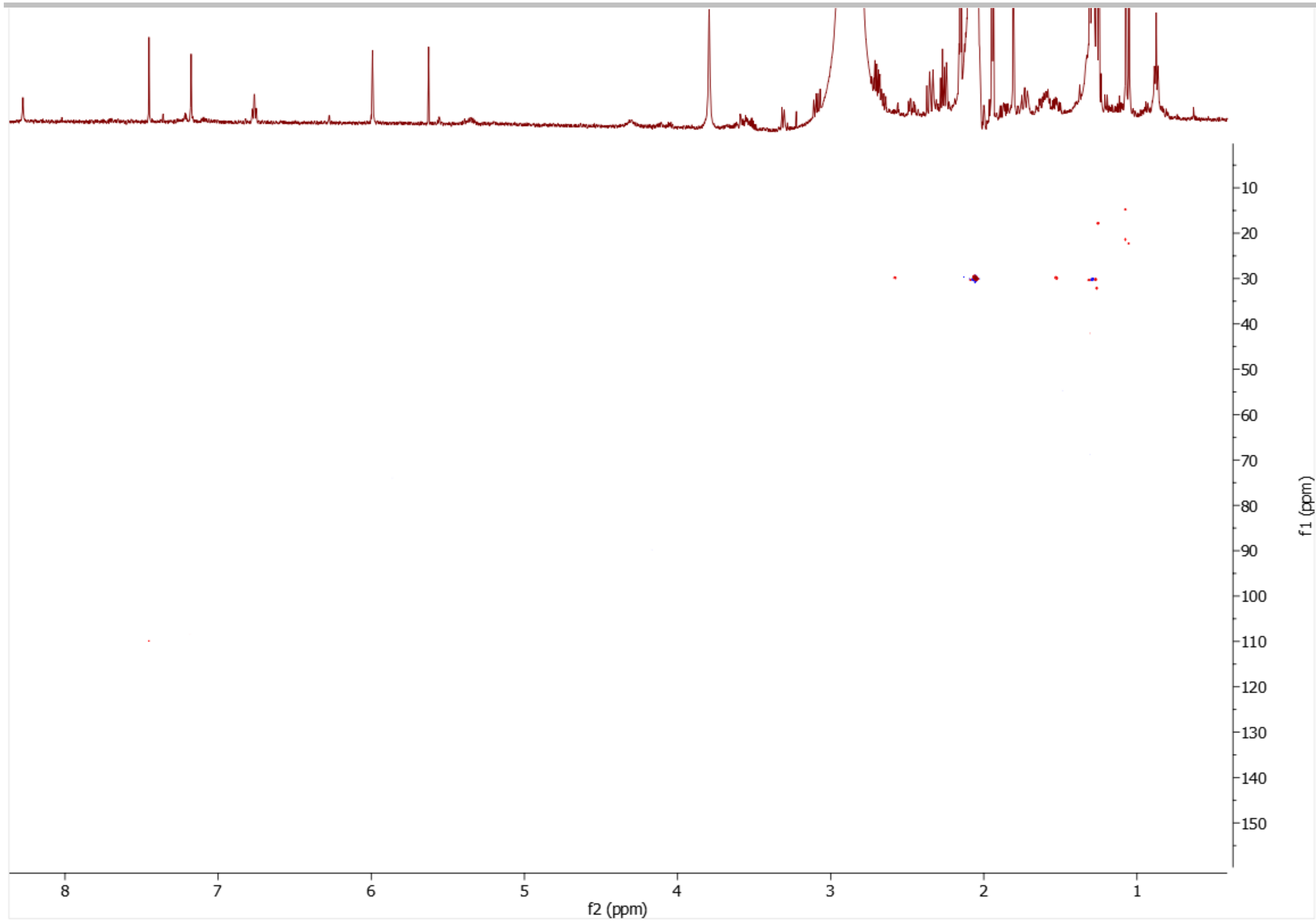


Figure S56. HSQC Spectrum of NAD_5

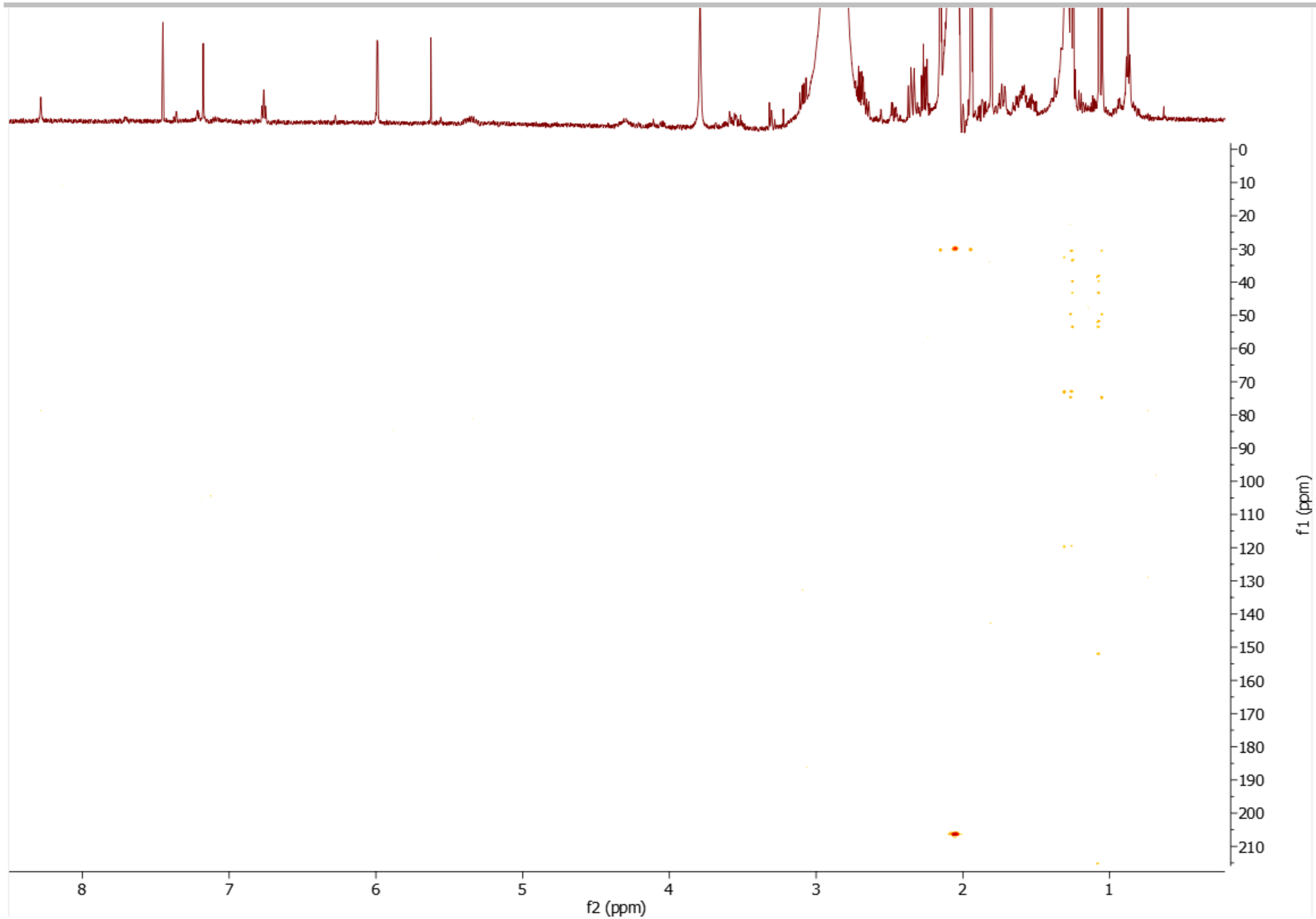


Figure S57. HMBC Spectrum of NAD_5

In Vivo Feeding Studies

To determine whether NAD_5 and NAD_6 were substrates for NodJ, and consequently whether they are precursors for the main and 1-series, *in vivo* feeding studies were undertaken. To conduct these experiments *nodJ* was added to the CY2 strain of *P. paxilli*, in which the entire paxilline biosynthetic gene cluster has been deleted, and the CY2 strain without *nodJ* was used as a control.

Fungi were grown using standard conditions (Pg 5) and nodulisporic acid substrates were provided to the culture on days 0 and 4 as MeOH aliquots (100 µL, 1–2 mg/mL substrate). The substrates included NAD_4, NAD_5, NAD_6 and a MeOH blank as a control (Table S16). Cultures were harvested after 10 days and mycelia were extracted and analysed following standard procedures (Pg 5).

Table S16. NodJ *in vivo* feeding studies

Experiment*	Strain	Substrate provided
A	CY2	MeOH (control)
B	LS29 (CY2 + *nodJ)*	MeOH (control)
C	CY2	NAD_4
D	LS29 (CY2 + *nodJ)*	NAD_4
E	CY2	NAD_6
F	LS29 (CY2 + *nodJ)*	NAD_6
G	CY2	$NAD_{5/1}$**
H	LS29 (CY2 + *nodJ)*	$NAD_{5/1}$**

**All experiments were conducted in triplicate*
***Insufficient NAD_5 was available to conduct the experiment therefore a sample containing both NAD_1 and NAD_5 was used.*

Analysis of experiments A and B did not show the presence of any nodulisporic acids (Figure S58) and culture growth appeared unaffected by addition of MeOH aliquots. Mycelia from experiment C showed the presence of the substrate, NAD_4, but no other nodulisporic acids, confirming CY2 does not contain proteins capable of producing additional nodulisporic acids. Experiment D produced NAD, NAD_1, NAD_5, and NAD_6, which are the expected products of NodJ when provided with NAD_4 as seen in heterologous reconstruction of the biosynthetic pathway (Figure S58). This result confirms both the expression of *nodJ* in this strain and its ability to access substrate provided via spiking the culture.

Singh *et al* predicted that NAD was formed via an enzyme catalysed elimination of 2''-OH from NAD_6, therefore experiments E and F were conducted to test this hypothesis.[3a] However, when NAD_6 was provided as a substrate NAD was not detected (Figure S58), indicating an alternative explanation for the production of main series nodulisporic acids was required. We propose that hydrogen abstraction at C-2'' of NAD_4 gives an allylic radical intermediate which partitions either to NAD through the loss of hydrogen or to NAD_6 through reaction with an iron-bound hydroxyl radical (Scheme 2, Manuscript). Such bifurcation of a radical intermediate has been reported for other P450 oxidases and is consistent with the experimental data obtained through feeding studies.[5] NAD_1 was also not detected in these experiments confirming the inability of NAD_6 to be utilised as a substrate by NodJ to produce any other nodulisporic acids and therefore the 6-series is a shunt pathway.

The final two experiments were conducted to establish whether NAD_5 was a substrate for NodJ and therefore a precursor for NAD_1. Since insufficient NAD_5 could be purified for feeding studies a sample containing both NAD_5 and NAD_1 was used as the substrate. When this substrate was provided to CY2 in experiment G the ratio of NAD_5 to NAD_1 remained unchanged after the 10 days in culture (Figure S59). However, with NodJ present (experiment H) the NAD_5 decreased and NAD_1 increased. This result provides evidence to suggest that NAD_5 can be converted to NAD_1 by NodJ and is therefore an intermediate and not a shunt product. Interestingly no evidence for production of NAD_3, a metabolite characterised by Merck, was observed in this experiment despite the obvious biosynthetic relationship to NAD_5.[3a]

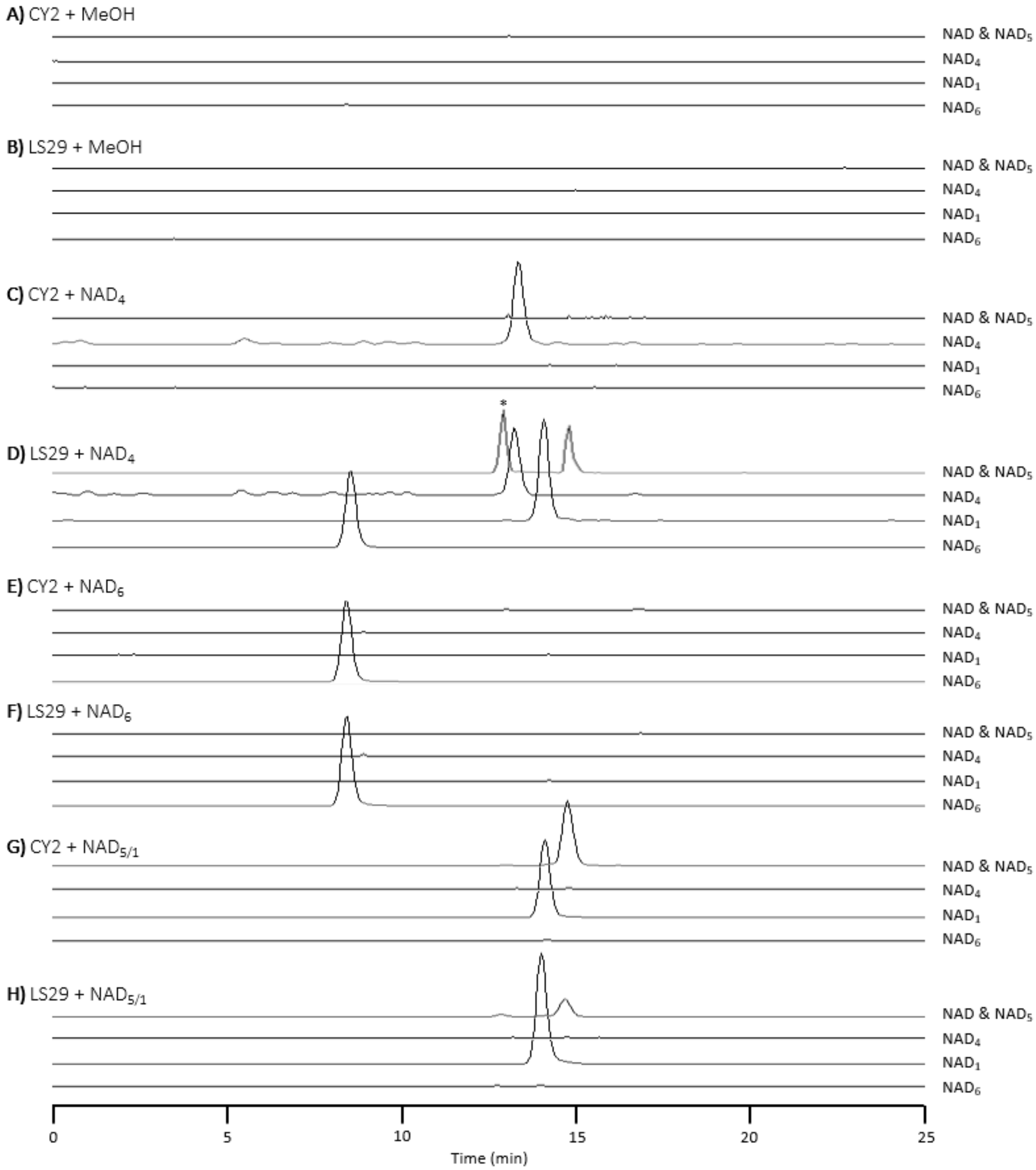


Figure S58. LC-MS analysis of in vivo feeding studies.

LC-MS analysis of in vivo feeding studies leading to the assignment of NAD_6 as a shunt product and NAD_5 as a precursor for NAD_1. Traces are EICs of HR-ESI-MS2 analysis and letter codes refer to experiment details in Table S16. *Since NAD and NAD_5 are isomeric a single trace has been provided and NAD peaks distinguished by an asterix (*).

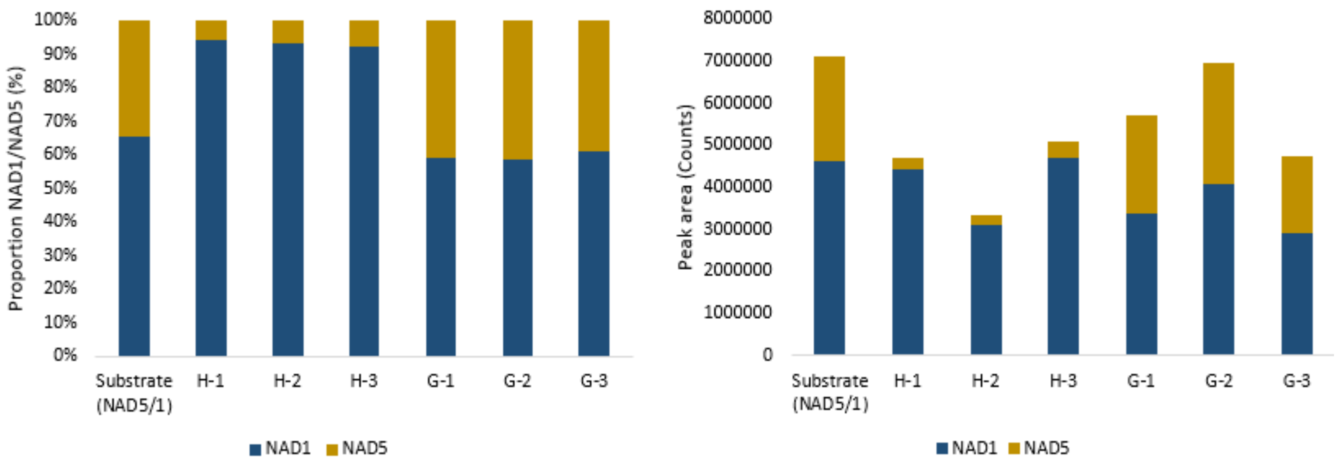


Figure S59. $NAD_{5/1}$ feeding study

Analysis of NAD_1 and NAD_5, and comparison to substrate, in experiments G and H (-1, -2, -3 refer to replicate cultures). Result presented as relative (left) and absolute (right) values.

Non-enzymatic NAD_6 elimination

To determine whether elimination of the 2''-OH of NAD_6 could proceed non-enzymatically a sample of NAD_6 was treated with formic acid and analysed by LC-MS over a period of 12 h. No evidence of NAD production was observed over this time and higher concentrations of acid led to degradation of the metabolite.

Proposed NodJ P450 Monooxygenase Bifurcation Mechanism

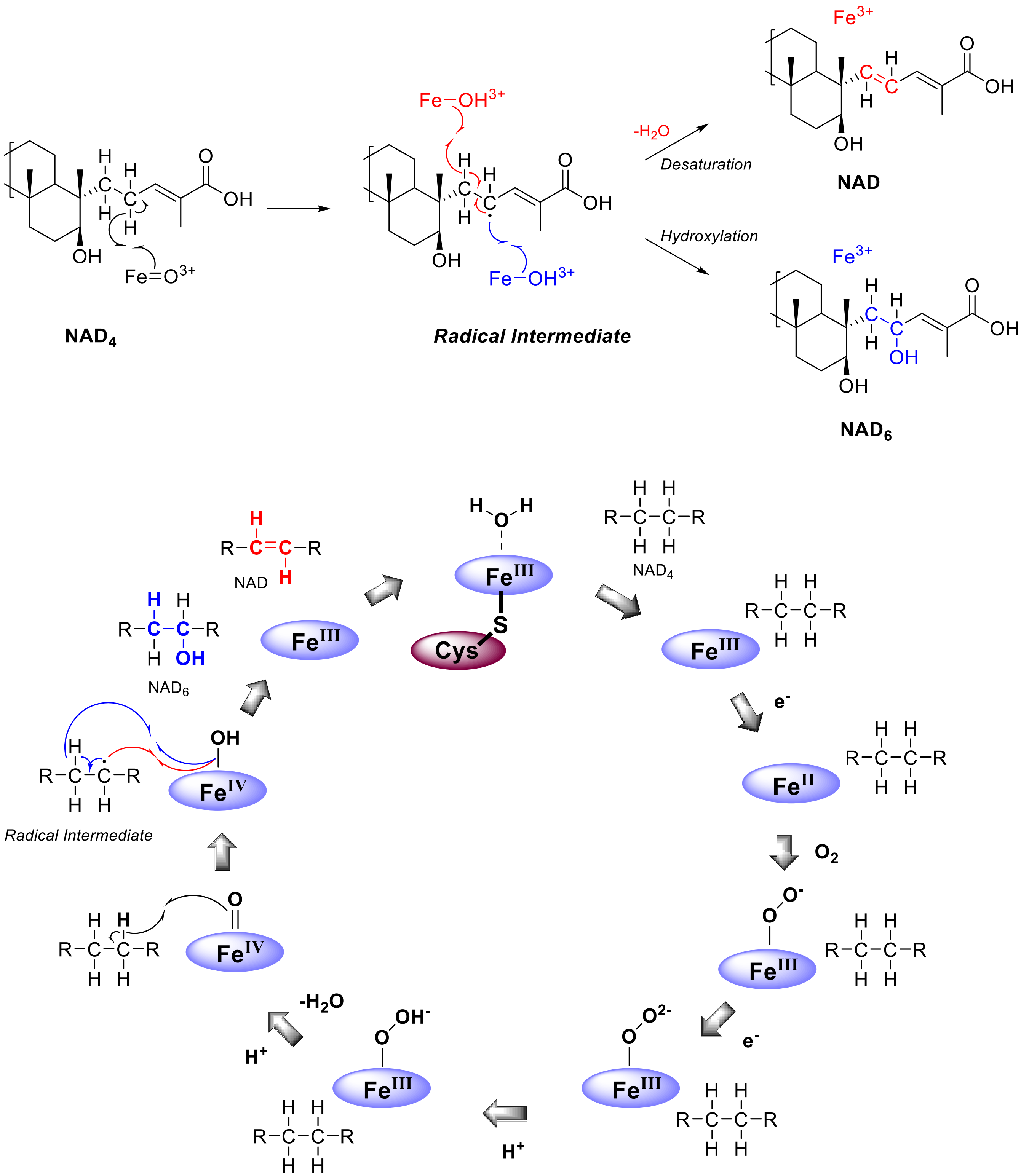


Figure S60. Proposed mechanism for NodJ

Top – Proposed mechanism fro NodJ in the bifurcation of nodulisporic acid biosynthesis to produce both NAD and NAD$_6$ via oxidation at C-2''.
Bottom – NodJ mechanism in the context of the cytochrome P450 catalytic cycle.

Phylogenetic Classification of NodX

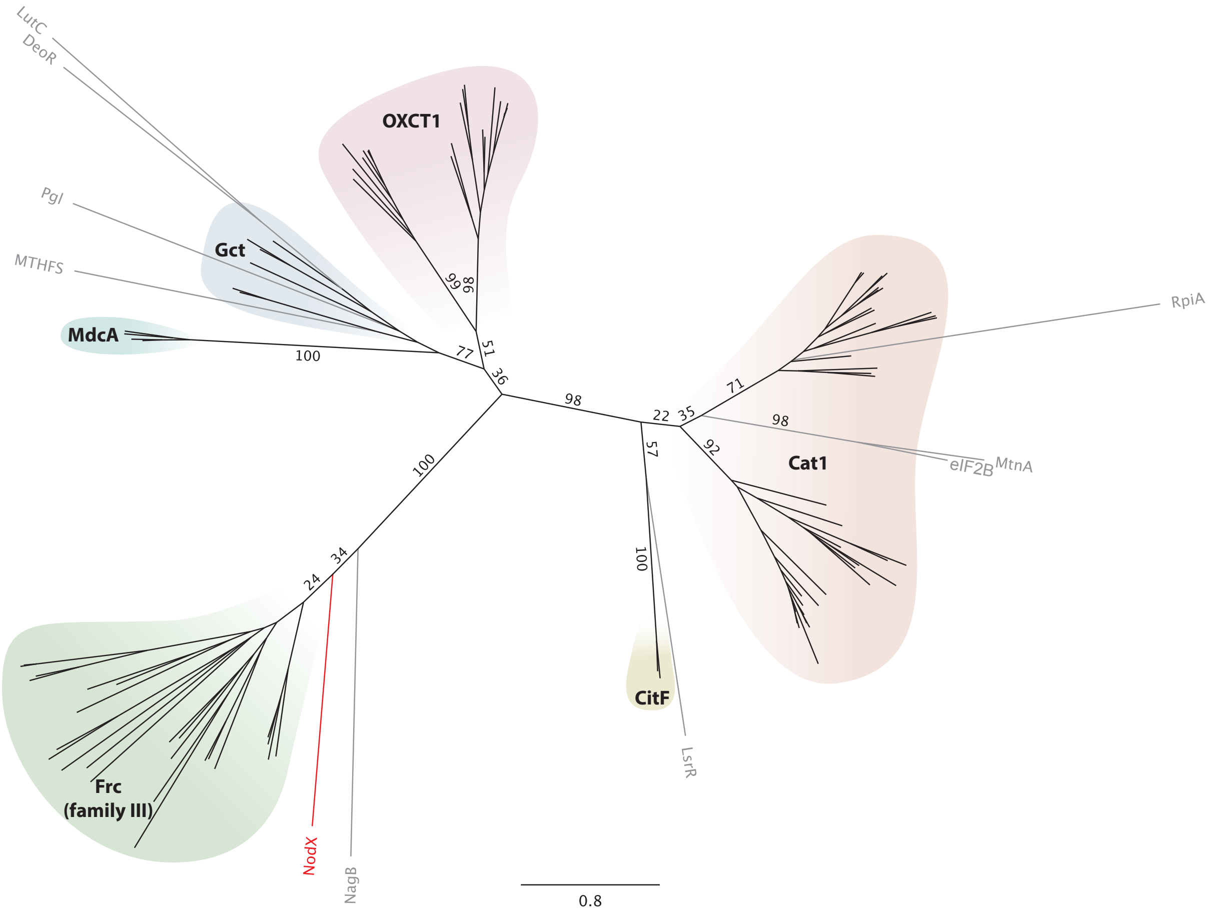


Figure S61. Phylogenetic classification of NodX

Maximum likelihood phylogenetic tree generated from protein sequences of NodX and CoA transferases aggregated by Hackmann, 2022.[6] Tree shows the position of NodX in red relative to the six CoA transferase families proposed by Hackman 2022.[6] Non-CoA transferase proteins identified by Hackmann as being erroneously labelled in online databases as CoA transferases are shown in grey. Protein sequences were aligned using Clustal Omega and this tree was generated using RAxML using the GAMMA BLOSUM62 protein model.[7] Maximum likelihood support values (100 bootstrap replicates) are shown for major branches, space permitting. Importantly, the Frc clade (also known as CoA transferase family III) contains both prokaryotic and eukaryotic CoA transferase sequences.

Nodulisporic Acid Biosynthesis Scheme

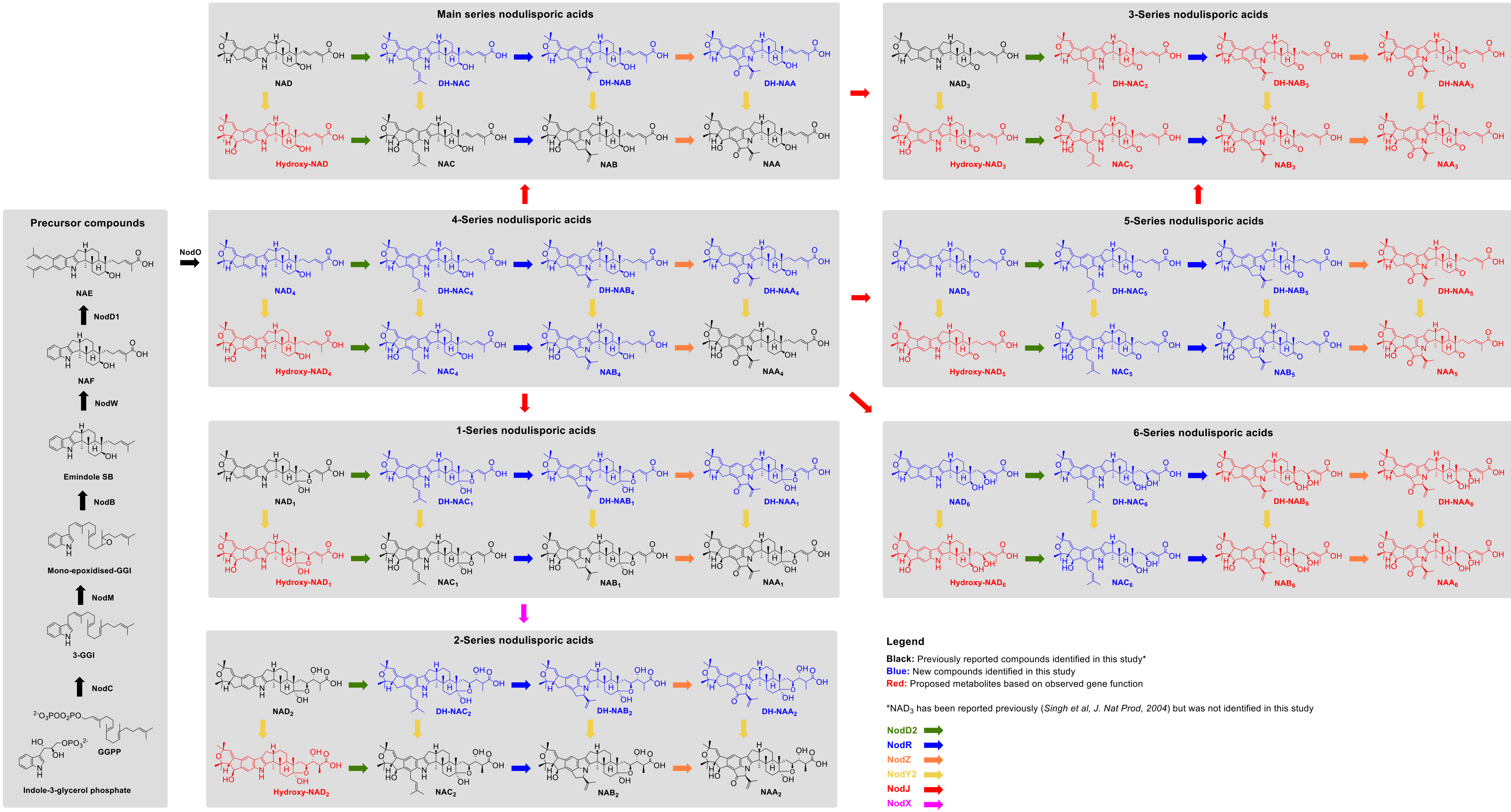


Figure S62. Nodulisporic acid biosynthesis

Author Contributions

Research was conducted in equal parts by lead authors Dr. A. T. Richardson, Dr. R. C. Cameron, and Dr L. J. Stevenson with assistance and contributions from supporting authors Dr A. J. Singh, Dr Y. Lukito, Dr D. Berry and Dr M. J. Nicholson. Prof E. J. Parker and Dr M. J. Nicholson conceived the study and were responsible for funding acquisition, project oversight and administration, and assisted with preparation of the manuscript.

References

[1] C. J. van Dolleweerd, S. A. Kessans, K. C. Van de Bittner, L. Y. Bustamante, R. Bundela, B. Scott, M. J. Nicholson, E. J. Parker, *ACS Synth. Biol.* **2018**, *7*, 1018-1029.
[2] K. C. Van de Bittner, M. J. Nicholson, L. Y. Bustamante, S. A. Kessans, A. Ram, C. J. van Dolleweerd, B. Scott, E. J. Parker, *J. Am. Chem. Soc.* **2018**, *140*, 582-585.
[3] a) S. B. Singh, J. G. Ondeyka, H. Jayasuriya, D. L. Zink, S. N. Ha, A. Dahl-Roshak, J. Greene, J. A. Kim, M. M. Smith, W. Shoop, J. S. Tkacz, *J. Nat. Prod.* **2004**, *67*, 1496-1506; b) J. G. Ondeyka, K. Byrne, D. Vesey, D. L. Zink, W. L. Shoop, M. A. Goetz, S. B. Singh, *J. Nat. Prod.* **2003**, *66*, 121-124; c) J. G. Ondeyka, A. M. Dahl-Roshak, J. S. Tkacz, D. L. Zink, M. Zakson-Aiken, W. L. Shoop, M. A. Goetz, S. B. Singh, *Bioorganic Med. Chem. Lett.* **2002**, *12*, 2941-2944; d) O. D. Hensens, J. G. Ondeyka, A. W. Dombrowski, D. A. Ostlind, D. L. Zink, *Tetrahedron Lett.* **1999**, *40*, 5455-5458; e) J. G. Ondeyka, G. L. Helms, O. D. Hensens, M. A. Goetz, D. L. Zink, A. Tsipouras, W. L. Shoop, L. Slayton, A. W. Dombrowski, J. D. Polishook, D. A. Ostlind, N. N. Tsou, R. G. Ball, S. B. Singh, *J. Am. Chem. Soc.* **1997**, *119*, 8809-8816.
[4] Y. Zou, X. Li, Y. Yang, S. Berritt, J. Melvin, S. Gonzales, M. Spafford, A. B. Smith, *J. Am. Chem. Soc.* **2018**, *140*, 9502-9511.
[5] a) F. P. Guengerich, *ACS Catal.* **2018**, *8*, 10964-10976; b) F. P. Guengerich, *Chem. Res. Toxicol.* **2001**, *14*, 611-650; c) A. E. Rettie, A. W. Rettenmeier, W. N. Howald, T. A. Baillie, *Science* **1987**, *235*, 890-893.
[6] T. Hackmann, *Prot. Sci.* **2022**, *31*, 864-881.
[7] a) A. Stamatakis, *Bionformatics* **2014**, *30*, 1312-1313; b) F. Sievers, A. Wilm, D. Dineen, *Mol. Syst. Biol.* **2011**, *7*, 539.