Supporting Information

Geometrically diverse lariat peptide scaffolds reveal an untapped chemical space of high membrane permeability

Colin N. Kelly[1], Chad E. Townsend[1], Ajay N. Jain[2], Matthew R. Naylor[1], Cameron R. Pye[3], Joshua Schwochert[3], R. Scott Lokey[1]*

[1] *Department of Chemistry and Biochemistry, University of California, Santa Cruz, USA*

[2] *Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, USA*

[3] *Unnatural Products Inc. 250 Natural Bridges Drive, Santa Cruz, California, USA*

[*] Author to whom correspondence should be addressed; email: slokey@ucsc.edu

Table of Contents

Abbreviations ***S4***

General Materials and Methods ***S5***

Synthetic Procedures ***S5***

Loading of 2-chlorotrityl resin **S5**

Library synthesis **S6**

Synthesis of pure lariat peptides **S7**

Manual solid-phase peptide synthesis **S7**

Automated peptide synthesis (Prelude X, Protein Technologies) **S7**

Ester formation using DIC **S7**

Cleavage from 2-chlorotrityl resin **S8**

Cyclization **S8**

N-Terminal acetylation **S8**

Scheme S1. **Split-pool scheme for libraries 2-4.** **S9**

Purification **S9**

Fmoc protection of free amino acids **S9**

N-Methylation of Fmoc amino acids **S9**

ANALYTICAL PROCEDURES ***S10***

Purity assessment **S10**

Library analysis by UPLC-MS^2 ..... **S10**

***Table S1*. LCMS conditions for library analysis** **S10**

PAMPA procedure ***S11***

Library Data Processing ***S13***

MDCK Assay ***S14***

***Figure S1*. 1NMe3 used as permeability benchmark** **S16**

Library 1: Representation and permeability statistics ***S17***

***Table S2*. Key characteristics of Library 1 (table) (MW, AlogP, # H-bond donors)** **S17**

***Figure S2*. Recovery of data vs mass redundancy in Library 1** **S17**

***Table S3*: Representation of each molecular weight in Library 1** **S18**

***Table S4*: Number of compounds in Library 1 dataset with each degree of methylation** **S19**

***Table S5*: Permeability of 1NMe3 standard in each Library 1 sub-library** **S19**

***Table S6*: Basic permeability statistics for Library 1** **S20**

***Table S7*: Number of identified compounds in Library 1 dataset with each methylation pattern** **S20**

***Table S8*: Number of identified compounds in Library 1 dataset from each sub-library** **S21**

***Table S9*: Number of identified compounds in the Library 1 dataset with L and D stereochemistry at each stereocenter** **S21**

***Table S10*: Number of identified compounds in the Library 1 dataset with homochiral and heterochiral diastereochemistry for each adjacent residue pair** **S22**

***Table S11:* Effect of stereochemistry for each degree of N-methylation on permeability** **S23**

***Table S12:* Relative number of identified compounds in the library with L vs. D stereochemistry at each stereocenter** **S23**

***Table S13*: Effect of heterochirality for each degree of N-methylation on permeability** **S24**

***Table S14:* Relative number of identified compounds in the library with homochirality vs. heterochirality between each pair of adjacent residues** **S24**

***Table S15*: Representation of adjacent stereochemical configurations** **S25**

***Table S16*: Average permeability of adjacent stereochemical configurations** **S25**

***Table S17*: Representation of each diastereomer for Leu^6/Ala^7, Ala^7/Leu^8, and Leu^8/Pro^9 at each degree of N-methylation** **S26**

Table S18: Permeability deviation for each diastereomer of Leu^6/Ala^7, Ala^7/Leu^8, and Leu^8/Pro^9 at each degree of N-methylation.....S27

Figure S3. Number of compounds in data set vs. average logP_{app} from each (a) methylation pattern and (b) sub-library S28

Figure S4. Correlation between permeability of compounds with heterochiral stereochemistry and number of compounds with heterochiral stereochemistry.....S29

Table S19: Average retention time and retention time variance by degree of N-methylation.....S30

Figure S5. Permeabilities of compounds in each sub-library of Library 1. .....S31

Figure S6. Effect of heterochirality on permeability by degree of N-methylation.....S32

Figure S7. Effect of stereochemistry at each position on permeability of compounds in Library 1. S33

Figure S8. Effect of relative stereochemistry between Leu8 and adjacent residues on permeability S33

Figure S9. Effect of stereochemistry at residues 7-9. .....S34

Table S20. Mann-Whitney U-tests evaluating significance of effects of relative stereochemistry between Leu^8 and adjacent residues on permeability. S35

NMR Solution Structure Generation for Compound 2 .....S36

Table S21. Peak assignments for Compound 2 .....S37

Table S22. ROE-distance restraints of Compound 2 in $CDCl_3$. Distances are presented as ranges ..S39

Table S23: Amide NH temperature shift coefficients for Compound 2 .....S39

Table S24: ^3J vicinal coupling constants between Hα-HN protons and derived ϕ dihedral restraints for Compound 2.....S39

Mass spectra of the sub-libraries S40

Table S25. Data on individual compounds.....S60

Analytical data of individual compounds (UV, TIC, selected ion, MS, H1 NMR).....S61

References.....S99

Abbreviations

Physicochemical

MW: molecular weight

HBD: hydrogen-bond donor

IMHB: intramolecular hydrogen bond

Analytical

HPLC: high-performance liquid chromatography

UPLC: ultra-high performance liquid chromatography

MS: mass spectrometry

PAMPA: parallel artificial membrane permeability assay

NMR: nuclear magnetic resonance

$CDCl_3$: deuterated chloroform

Synthesis

DMSO: dimethylsulfoxide

DMF: N,N-dimethylformamide

ACN: acetonitrile

H_2O: water

PBS: phosphate buffer solution (pH 7.4)

MeOH: methanol

TFA: trifluoroacetic acid

FA: formic acid

DCM: dichloromethane

HFIP: 1,1,1,3,3,3-hexafluoroisopropanol

HATU: 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate

SPE: Solid phase extraction

COMU: (1-Cyano-2-ethoxy-2-oxoethylideneaminooxy)dimethylamino-morpholino-carbenium hexafluorophosphate

DIPEA: diisopropylethylamine

Fmoc: 9-fluorenylmethyloxycarbonyl

DIC: diisopropylcarbodiimide

DMAP: N,N-dimethylaminopyridine

DBU: 1,8-diazabicyclo[5.4.0]undec-7-ene

General Materials and Methods

Solvents and reagents, including dry solvents, were purchased from Fisher Scientific. Fmoc-protected amino acids and coupling agents were purchased from Combi-Blocks, Oakwood, or Chem-Impex. L-Leucine-d_3 and L-alanine-d_3 were purchased from Cambridge Isotope Laboratories. NMR solvents were purchased from Cambridge Isotope Laboratories. HPLC solvents were Optima grade, purchased from Fisher

Scientific. 2-Chlorotrityl chloride polystyrene resin and L-Leu-2-chlorotrityl polystyrene resin were purchased from Rapp-Polymere. Dodecane and soy lecithin used in PAMPA were purchased from Alfa Aesar.

Synthetic Procedures

Loading of 2-chlorotrityl resin

The desired amount of 2-chlorotrityl chloride resin was swelled with DCM in a SPE tube for 1 h, at which time a solution of Fmoc-D-leucine (2 equiv) and DIPEA (3 equiv) in DCM was added. The tube was capped and inverted, and the stopcock opened to allow gas evolution. Once gas evolution subsided, the tube was shaken for 3 h. The resin was capped with a solution of 2:1:17 MeOH:DIPEA:DCM (2 x 30 min). The resin was washed with DMF (3x) followed by DCM (3x). The loading value was calculated by quantifying UV absorbance of the dibenzofulvene byproduct (300 nm) after Fmoc removal.

Library synthesis

The libraries were synthesized starting with 2-chlorotrityl resin loaded with residue 8 (Scheme S1). The procedures for manual amide coupling and Fmoc deprotection were used to install residues 1-7 without protection of the threonine side chain. At steps requiring transfer of resin, transfers were carried out prior to Fmoc

deprotection. Residue 9 was added as Fmoc-proline by ester coupling. After Fmoc deprotection and cleavage from resin, cyclization (residue 9 – residue 8) was carried out in solution. The lariat peptides were purified by reversed-phase chromatography.

Library 1 was synthesized as 16 sub-libraries at a scale of 0.025 mmol/sub-library using a split-pool strategy (Scheme S2). Sub-libraries **1-8** were prepared from L-Leu-2CT and sub-libraries **9-16** from D-Leu-2CT. Couplings were performed manually. For the installation of residues 7, 6, 5, and 4, the resin was separated into four tubes for the addition of different amino acids and recombined and mixed prior to Fmoc deprotection.

Libraries **2-4** were synthesized as single mixtures at a scale of 0.025 mmol starting from L-Leu-2CT resin and installing amino acids in the appropriate order. For Library 4, the N-terminal acetyl group was installed using the procedure for N-terminal acetylation.

Library **5** was synthesized as a single mixture at a scale of 0.2 mmol starting from L-Leu-2CT resin using a split-pool strategy. For the installation of residues 6, 5, and 4, the resin was separated into eight tubes for the addition of different amino acids and recombined and mixed prior to Fmoc deprotection.

Synthesis of pure lariat peptides

Lariat peptides were synthesized starting with 2-chlorotrityl resin loaded with L-Leu or D-Leu. Further residues were added using the procedure for automated peptide synthesis yielding a linear octapeptide bound to resin with an unprotected Thr sidechain. The free hydroxyl group was then acylated with Fmoc-L-Pro or Fmoc-D-Pro by the procedure for ester coupling. After final Fmoc removal, the peptide was removed from the resin using the procedure for cleavage from 2-chlorotrityl resin. The peptide was cyclized using the procedure for cyclization and purified using the procedure for purification.

Manual solid-phase peptide synthesis

To a solution of the Fmoc amino acid (2 eq, 0.5 M in DMF) was added HATU (1.9 eq, 0.5 M in DMF) followed by DIPEA (2.5eq). The resultant solution was swirled and allowed to stand for 5 minutes, then added to the drained resin. The resin was heated to 50°C for 2 h. After coupling, the resin was washed with DMF (3x) and DCM (3x). For Fmoc deprotection, the resin was treated with a solution containing 2% 1,8-diazabicycloundec-7-ene (DBU) and 2% piperidine in DMF for 15 min at room temperature. The resin was then washed with DMF (3x) and DCM (3x).

Automated peptide synthesis (Prelude X, Protein Technologies)

Synthesis was performed on 0.1 mmol scale using loaded 2-chlorotrityl resins. Fmoc deprotection was accomplished with 2% DBU and 2% piperidine in DMF for 1 min at 90°C. Couplings were carried out with Fmoc-protected amino acids (5 eq), COMU (4 eq), and DIPEA (6 eq) in DMF for 10 min at 90°C. Each coupling and deprotection step was followed by a wash with DMF (4x) and DCM (2x).

Ester formation using DIC

Fmoc-proline (10 eq) was dissolved in DMF/DCM (1:9, roughly 3mL/g Fmoc-proline). DMAP was added (0.25 eq) followed by DIC (10 eq). The solution was swirled rapidly until a precipitate formed. The mixture was added to the resin and the SPE tube capped. The reaction was shaken at room temperature for 3 h, then drained. Another portion of reactants (10 eq Fmoc-Pro-OH, 0.25 eq DMAP, 10 eq DIC) was immediately added without washing the resin and the resin was allowed to react overnight at room temperature. The resin was drained and washed with DMF until the thick precipitate which formed during the reaction was removed, then washed with DCM (3x).

Cleavage from 2-chlorotrityl resin

Branched linear peptides were cleaved from the resin with 25% HFIP in DCM (2 x 30 min). The resin was rinsed with DCM between treatments. Solvent was

evaporated under a stream of nitrogen. DCM was added and evaporation repeated. The residue was stored overnight in a vacuum desiccator prior to cyclization.

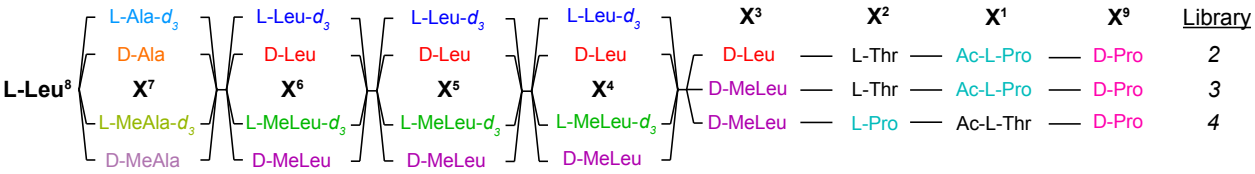
Cyclization

The solvent volumes in this procedure are for 0.1 mmol of peptide. The concentration during cyclization was approximately 0.001 M.

COMU (3 eq) was placed in a round-bottom flask with a stir bar, followed by THF (90 mL) and DIPEA (3eq). In a separate vessel, the branched linear peptide was dissolved in ACN (10mL) and DIPEA (3 eq). The peptide was added dropwise and in portions to the round-bottom flask with rapid stirring during 30 minutes. Stirring was continued for 16 h. The solution was concentrated under reduced pressure.

N-Terminal acetylation

A mixture containing acetic anhydride (6eq), DIPEA (7.5eq), and DMF (0.6mL) was added to the drained resin. The resin was shaken for 2 h at room temperature, drained, and washed with DMF (3x) and DCM (3x).



Scheme S1. **Split-pool scheme for libraries 2-4.**

Purification

Crude cyclic peptides were purified on a Biotage Isolera Prime automated chromatography system equipped with a SNAP Bio C18 25g column eluting with water/acetonitrile containing 0.1% TFA. Libraries were purified with the following gradient (%ACN in water): 30% (50mL), 30-100% (25mL), 100% (25mL). Individual compounds were purified with the following gradient (%ACN in water): 20% (50mL), 20-80% (450mL), 100% (75mL).

Fmoc protection of free amino acids

Fmoc-protected L-Ala-d_3 and L-Leu-d_3 were prepared by Fmoc protection of the commercially available deuterated amino acids according to literature procedure.[1]

N-Methylation of Fmoc amino acids

Fmoc-N-methyl amino acids were prepared from Fmoc amino acids by reduction of the formaldehyde oxazolidinone as previously described.[2]

ANALYTICAL PROCEDURES

Purity assessment

Compounds were analyzed using the Orbitrap with the same gradient as the library analysis and purity was quantified by UV absorbance at 200nm. For compounds occurring in the library, retention time agreement corroborated proper sequencing of the compound.

Library analysis by UPLC-MS^2

Acquisition of MS^2 data for sequencing was carried out as reported previously, with some modifications.[3] Source ionization was used, as we suspected this may facilitate fragmentation at the ester bond. Source ionization was optimized to produce the strongest signal from the M+Na mass and minimize signal from the M+H mass, which

could interfere with sequencing. For libraries 1-4, the top 7 most intense non-isotopic peaks were selected for MS^2 acquisition at each MS^1 acquisition. For Library 5, the top 10 were selected.

***Table S1*. LCMS conditions for library analysis**

Liquid chromatography	Thermo Scientific Dionex UltiMate system (RS pump, RS autosampler, RS column compartment)
Mass spectrometer	Thermo Scientific Velos Pro mass spectrometer
Eluent	H_2O/ACN containing 0.1% FA (all solvents were Optima® grade)
Column	Thermo Scientific Acclaim™ RSLC 120 C18 (2.2µm 120Å, 2.1 X 250 mm) (product# 074812, serial# 001101)
Flow rate	0.4mL/min
Temperature	50°C
Gradient for Libraries 1-4 (%ACN)	0-45min: 52-67, 45-55min: 100, 55-60min: 52
Gradient for Library 5 (%ACN)	0-45min: 50-80, 45-55min: 100, 55-60min: 50
Source ionization voltage	80V
MS1	FTMS
MS2	ITMS
MS^n activation	Collision-induced dissociation
Normalized collision energy	35.0
Isolation width	2.0 m/z
Activation Q.	0.250
Activation time	10.0 ms

PAMPA procedure

The PAMPA assay was carried out and the peak volumes interpreted using a procedure utilized previously in our lab.[4]

The analyte concentration in the PAMPA assay was 250 μM for sub-libraries (total concentration, roughly 1 μM per compound) and 1 μM for compounds assayed individually. Internal standards were included in the assay at a concentration of 1 μM. For sub-libraries, 1NMe3, synthesized according to published procedures,[5] was used as the internal standard and for compounds assayed individually, carbamazepine was used as the standard. The assay was run for 18 h.

A 96-well donor plate with 0.45 μm hydrophobic Immobilon-P membrane supports (Millipore MAIPNTR10) and a 96-well Teflon acceptor plate (Millipore MSSACCEPTOR) were used in the PAMPA permeability test. The acceptor plate was prepared by adding 300 μL of 5% DMSO in PBS to each well. Donor well solutions were prepared by diluting 50 μL DMSO stock solutions prepared above to a final volume of 1000 μL with PBS and mixed thoroughly. The frits were infused with 5 μL of dodecane containing 1% (w/v) soy lecithin (90%, Alfa Aesar). The membranes were allowed to equilibrate for 5 minutes before adding the donor well solution and placing on top of the acceptor well solution to begin the assay.

Samples were prepared for LC-MS analysis by diluting with an equal volume of ACN. The donor wells were further diluted tenfold with 1:1 ACN/H_2O for approximately even analyte concentration in the donor and acceptor wells.

Sink PAMPA conditions

The assay was run as described above, except that the donor well contained 0.2% (v/v) polysorbate 80 and the acceptor well contained 0.2% (w/v) TPGS-750M. Donor well samples were prepared for LC-MS analysis by diluting with an equal volume of 9:1 ACN / 2% (w/v) TPGS-750M in water. Acceptor well samples were prepared for LC-MS analysis by diluting with an equal volume of 9:1 ACN / 2% (v/v) polysorbate 80 in water.

Library Data Processing

The PAMPA LC-MS data was processed using AUTOPAMPA, CycLS, RTMerge. These programs and instructions for installation and use are available on GitHub at https://github.com/LokeyLab/PAMPA-Analysis-Support-Tools.

Library chromatograms and the PAMPA data spreadsheet are available in the supporting information as separate files.

Although the parent mass provides the degree of N-methylation and number of residues of each stereochemistry, the parent mass does not provide the order of residues. Previously we reported an algorithm for deconvoluting cyclic peptide libraries based on matching their MS^2 fragment ions to a virtual library derived from the theoretical compounds present in each sub-library.[3] After processing the raw MS^2 spectra using CycLS, we removed all data with sequencing confidence scores below 0.01 (the confidence score refers to the difference between the sequencing scores of the highest scoring sequence and the second highest scoring sequence divided by the highest score).[3] Duplicate sequences were resolved by removing the entry with lower confidence score. For data with assigned permeability values of 0, we integrated the LC-MS peaks manually wherever possible. When the peaks could not be integrated, the data were discarded.

MDCK Assay

The MDCK assay was carried out according to the procedure utilized by Furukawa et al.[6]

Transcellular transport of the test compounds from apical to basal direction using MDCK II cells was investigated. For the transcellular transport assay, the culture medium on the apical side and the basal side was replaced with buffer (pH 6.5) and buffer (pH 7.4) containing BSA, respectively. The buffer on the apical side was replaced with the buffer containing 10 μM of test compounds to start the incubation. After the incubation, aliquots of the solutions were sampled from both the apical side and basal side, and the concentrations of each compound were determined by LC-MS/MS.

Apparent permeability coefficient (Papp) was calculated by the following equations.

$P_{app} = (Cb \times V) / (Ca \times t \times A)$

P_{app}: apparent permeability (10^{-6} cm/sec)

Ca: test compound concentration added to the apical side (μM)

Cb: test compound concentration on the basal side (μM)

A: surface area of cell monolayer (cm^2)

V: volume of buffer on the basal side (cm^3)

t: incubation time (sec)

Natural Products Atlas analysis

To explore the prevalence and nature of lariat architectures in macrocyclic peptide natural products we computationally interrogated the Natural Products Atlas (van Santen, et al., *ACS Cent Sci* **2019**, *5* (11), 1824-1833) encompassing bacterial, fungal and cyanobacterial compounds. Natural products were considered macrocyclic peptides if they had a ring of greater than 13 atoms which contained a combined 3 amides or esters in this largest ring. To identify lariats, macrocyclic peptides were fragmented about all amide and ester bonds, and the subsequent "monomer fragments" were interrogated for the presence of functional groups that indicate that they are themselves the product of multiple monomer additions (therefore constituting a lariat "tail"). Monomer fragments were considered a "tail" if they contained some combination of at least two: amides, esters, thiazol(e/ine)s, or oxazol(e/ines). Finally, the branch-point was assigned (through visual inspection) as either an amide, ester, thioester, or aromatic ring. For example, an isoacyl-threonine; whose sidechain oxygen is in the macrocyclic backbone as an ester, and whose backbone nitrogen is elaborated as a "tail", would be defined as an ester branchpoint.

The Natural Products Atlas (accessed April 2020) contains 25,523 natural products, of which 1928 are macrocyclic peptides. A substantial portion of the cyclic peptides, 573 or ~30%, were identified as lariats per the definition above. Lariats which have multiple

macrocycles were removed from subsequent analysis as they may contain multiple branch-points or other complex architectures that defy a simple delineation between macrocyclic backbone and lariat tail. A large majority of the remaining lariats, 312 or 64%, contained an ester branchpoint. While not explicitly quantified in our analysis, a vast majority of these ester branched lariats are iso-acyl threonines or serines, highlighting the relative importance of this architecture. Amide branch-points make up a majority of those remaining, primarily through monomers such as lysine and diaminobutyric acid. Finally, complex natural products such as thiocillin, whose tails are linked through aromatic rings such as pyridine contribute a further 59 structures.

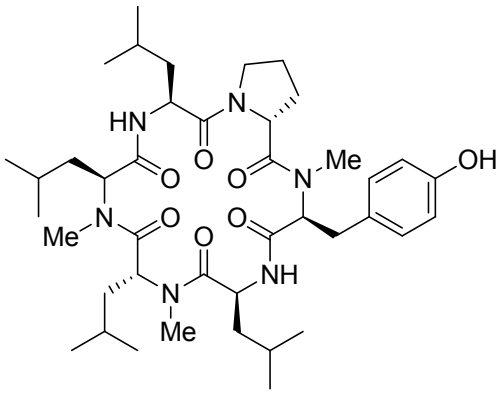


***Figure S1*. 1NMe3 used as permeability benchmark**

Library 1: Representation and permeability statistics

Table S2. **Key characteristics of Library 1 (table) (MW, AlogP, # H-bond donors)**

Degree of Methylation (#R = Me)	MW (non-deuterated)	H-bond donors	Number of compounds (expected)	AlogP
0	987.64	6	256	2.59
1	1001.65	5	1024	2.94
2	1015.67	4	1536	3.28
3	1029.67	3	1024	3.62
4	1043.70	2	256	3.96

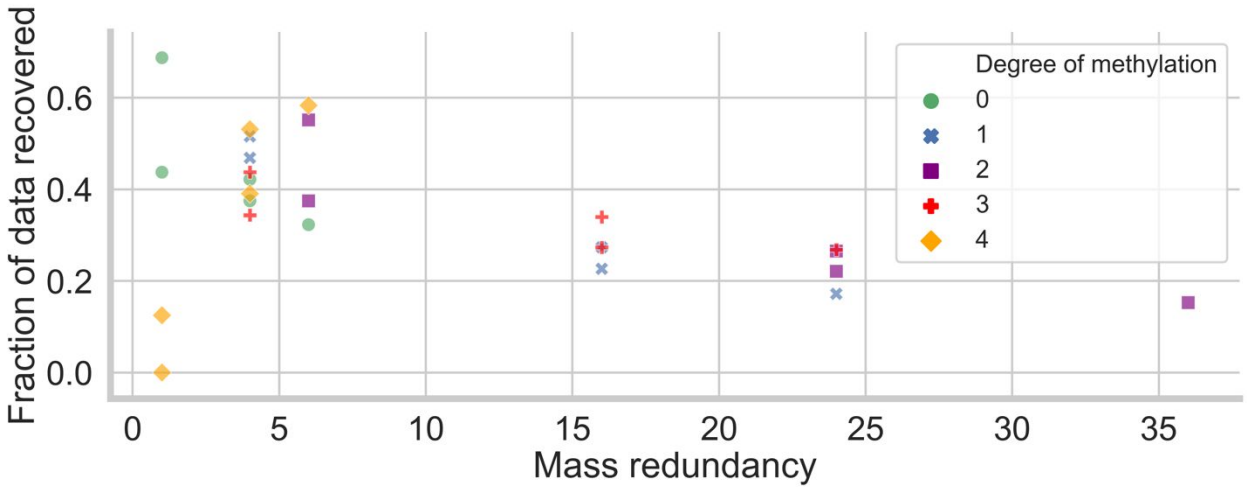


***Figure S2*. Recovery of data vs mass redundancy in Library 1.**
Each point represents a particular mass. Fraction of data recovered = (number of identified compounds in Library 1 dataset with a given mass) / (theoretical number of compounds with that mass). The list of masses may be found in Table S3.

***Table S3*: Representation of each molecular weight in Library 1**

Degree of methylation	MW	Theoretical number	Total number in data set	Efficiency of data recovery (% of theoretical)
0	987.64	16	7	44
0	990.66	64	27	42
0	993.67	96	30	31
0	996.69	64	24	38
0	999.71	16	11	69
1	1001.65	64	33	52
1	1004.67	256	54	21
1	1007.69	384	63	16
1	1010.71	256	62	24
1	1013.73	64	28	44
2	1015.67	96	30	31
2	1018.69	384	97	25
2	1021.71	576	85	15
2	1024.72	384	83	22
2	1027.74	96	51	53
3	1029.68	64	21	33
3	1032.70	256	66	26
3	1035.72	384	98	26
3	1038.74	256	85	33
3	1041.76	64	28	44
4	1043.70	16	2	13
4	1046.72	64	25	39
4	1049.74	96	56	58
4	1052.76	64	33	52
4	1055.78	16	0	0

***Table S4*: Number of identified compounds in Library 1 dataset with each degree of methylation**

Degree of methylation	Theoretical number	Number in dataset	Data recovery (% of theoretical)
0	256	99	39
1	1024	240	23
2	1536	346	23
3	1024	298	29
4	256	116	45

***Table S5*: Permeability of 1NMe3 standard in each Library 1 sub-library**

Library	P_{app} (10^{-6} cm/s) ± SD	Pro[1]	MeLeu[2]	Leu[8]	Pro[9]	1/2	8/9
1.1	10.27	L	L	L	L	Hom.	Hom.
1.2	10.17	L	L	L	D	Hom.	Het.
1.3	7.74	D	L	L	L	Het.	Hom.
1.4	8.93	D	L	L	D	Het.	Het.
1.5	7.42	L	D	L	L	Het.	Hom.
1.6	6.31	L	D	L	D	Het.	Het.
1.7	6.63	D	D	L	L	Hom.	Hom.
1.8	7.44	D	D	L	D	Hom.	Het.
1.9	11.24	L	L	D	L	Hom.	Het.
1.1	6.75	L	L	D	D	Hom.	Hom.
1.11	6.79	D	L	D	L	Het.	Het.
1.12	6.61	D	L	D	D	Het.	Hom.
1.13	7.94	L	D	D	L	Het.	Het.
1.14	8.09	L	D	D	D	Het.	Hom.
1.15	7.59	D	D	D	L	Hom.	Het.
1.16	10.10	D	D	D	D	Hom.	Hom.
Average	**8.13 ± 1.55**						
Avg L		8.52	8.56	8.11	8.20		
Avg D		7.73	7.69	8.14	8.05		
Avg hom.						8.77	7.95
Avg het.						7.48	8.30
Ratio		**1.10**	**1.11**	**1.00**	**1.02**	**1.17**	**0.96**

Hom. = homochiral, Het. = heterochiral

***Table S6*: Basic permeability statistics for Library 1**

	# of P_{app} > 5 × 10^{-6} cm/s	Mean $logP_{app}$	Median $logP_{app}$	Variance of $logP_{app}$
Full library	181 (16%)	-5.95	-5.83	0.47
DoM = 0	0 (0%)	-7.02	-6.86	0.45
DoM = 1	13 (5%)	-6.26	-6.16	0.37
DoM = 2	54 (15%)	-5.90	-5.83	0.32
DoM = 3	72 (24%)	-5.61	-5.53	0.19
DoM = 4	42 (37%)	-5.44	-5.37	0.09

DoM = degree of methylation

***Table S7*: Number of identified compounds in Library 1 dataset with each methylation pattern**

Methylation pattern (Leu4, Leu5, Leu6, Ala7)	Number in dataset
NNNN	99
NNNY	82
NNYN	70
NYNN	41
YNNN	47
NNYY	115
NYNY	86
NYYN	55
YNNY	42
YNYN	24
YYNN	24
NYYY	92
YNYY	103
YYNY	82
YYYN	21
YYYY	116

Theoretical number is 256 in all cases.

***Table S8*: Number of identified compounds in Library 1 dataset from each sub-library**

Sub-library	Number of compounds in dataset
1	95
2	53
3	81
4	81
5	70
6	53
7	59
8	37
9	76
10	80
11	85
12	47
13	61
14	63
15	92
16	66

Theoretical number is 256 in all cases.

***Table S9*: Number of identified compounds in the Library 1 dataset with L and D stereochemistry at each stereocenter**

Stereocenter	L	D
Pro[1]	551	548
MeLeu[2]	598	501
Thr[3]	1099	0
Leu[4]	534	565
Leu[5]	544	555
Leu[6]	520	579
Ala[7]	668	431
Leu[8]	529	570
Pro[9]	619	480

***Table S10*: Number of identified compounds in the Library 1 dataset with homochiral and heterochiral diastereochemistry for each adjacent residue pair**

Stereocenters	Homochiral	Heterochiral
$Pro^{1}/MeLeu^{2}$	558	541
$MeLeu^{2}/Thr^{3}$	598	501
Thr^{3}/Leu^{4}	534	565
Leu^{4}/Leu^{5}	595	504
Leu^{5}/Leu^{6}	601	498
Leu^{6}/Ala^{7}	659	440
Ala^{7}/Leu^{8}	552	547
Leu^{8}/Pro^{9}	561	538

Table S11: **Effect of stereochemistry for each degree of N-methylation on permeability**

Degree of methylation	Pro[1]	MeLeu[2]	Thr[3]	Leu[4]	Leu[5]	Leu[6]	Ala[7]	Leu[8]	Pro[9]
0	0.134	0.019	N/A	-0.003	0.013	0.083	-0.021	0.274	0.507
1	0.112	0.104	N/A	-0.440	0.138	-0.021	-0.006	0.319	0.078
2	0.122	0.021	N/A	-0.396	0.033	-0.096	-0.116	0.166	-0.047
3	-0.079	-0.062	N/A	-0.236	-0.035	-0.010	-0.053	0.207	-0.088
4	0.056	-0.067	N/A	-0.158	-0.038	0.061	-0.078	-0.027	0.029
All	0.077	0.048	N/A	-0.282	0.039	0.080	-0.161	0.224	0.015

The numbers correspond to the differences between median log(P_{app}) for L stereochemistry vs. D stereochemistry, with positive numbers indicating higher log(P_{app}) for L stereochemistry.

Degree of methylation	Pro[1]	MeLeu[2]	Thr[3]	Leu[4]	Leu[5]	Leu[6]	Ala[7]	Leu[8]	Pro[9]
0	0.980	1.020	N/A	1.250	0.707	0.547	2.667	0.833	1.250
1	0.935	1.051	N/A	0.739	1.087	0.655	1.857	0.805	1.500
2	1.000	1.190	N/A	0.870	1.097	0.912	1.602	1.023	1.247
3	1.000	1.328	N/A	0.935	0.987	1.328	1.275	0.961	1.129
4	1.231	1.367	N/A	1.636	0.731	0.902	1.071	0.933	1.522
All	1.005	1.194	N/A	0.945	0.980	0.898	1.550	0.928	1.290

Table S12: **Relative number of identified compounds in the library with L vs. D stereochemistry at each stereocenter**

The numbers correspond to the ratio of the number of compounds with L stereochemistry to the number of compounds with D stereochemistry. Larger numbers indicate a preponderance of compounds with L stereochemistry at that stereocenter for library members with the specified number of N-methyl groups in the macrocycle.

***Table S13*: Effect of heterochirality for each degree of N-methylation on permeability**

Degree of methylation	Pro[1]/ MeLeu[2]	MeLeu[2]/ Thr[3]	Thr[3]/ Leu[4]	Leu[4]/ Leu[5]	Leu[5]/ Leu[6]	Leu[6]/ Ala[7]	Ala[7]/ Leu[8]	Leu[8]/ Pro[9]
0	-0.539	0.019	-0.003	-0.208	-0.138	0.099	0.199	-0.369
1	-0.244	0.104	-0.440	-0.202	-0.220	-0.109	-0.099	-0.254
2	-0.277	0.021	-0.396	-0.227	-0.185	-0.125	-0.101	-0.183
3	-0.203	-0.062	-0.236	-0.003	-0.148	0.066	-0.009	0.045
4	-0.092	-0.067	-0.158	-0.048	-0.056	0.082	-0.016	-0.001
All	-0.233	0.048	-0.282	-0.195	-0.265	-0.074	-0.051	-0.128

The numbers correspond to the differences between median log(P_{app}) for homochiral vs. heterochiral, with positive numbers indicating higher log(P_{app}) for homochiral stereochemistry.

***Table S14:* Relative number of identified compounds in the library with homochirality vs. heterochirality between each pair of adjacent residues**

Degree of methylation	Pro[1]/ MeLeu[2]	MeLeu[2]/ Thr[3]	Thr[3]/ Leu[4]	Leu[4]/ Leu[5]	Leu[5]/ Leu[6]	Leu[6]/ Ala[7]	Ala[7]/ Leu[8]	Leu[8]/ Pro[9]
0	0.941	1.020	1.250	1.152	1.605	1.676	0.941	1.063
1	0.920	1.051	0.739	1.424	1.553	1.637	1.202	0.983
2	1.097	1.190	0.870	1.471	1.471	1.746	0.989	1.023
3	1.113	1.328	0.935	1.099	1.014	1.346	0.874	1.159
4	0.949	1.396	1.614	0.513	0.474	0.949	1.170	0.917
All	1.031	1.194	0.945	1.181	1.207	1.498	1.009	1.043

The numbers correspond to the ratio of the number of compounds with homochiral stereochemistry between two residues to the number of compounds with heterochiral stereochemistry between the two residues. Larger numbers indicate a preponderance

of compounds with homochiral stereochemistry between those stereocenters for library members with the specified number of N-methyl groups in the macrocycle.

***Table S15*: Representation of adjacent stereochemical configurations**

	L,L	L,D	D,L	D,D
Pro[1]/MeLeu[2]	1.11	0.90	1.07	0.92
MeLeu[2]/Thr[3]	1.09		0.91	
Thr[3]/Leu[4]	0.97	1.03		
Leu[4]/Leu[5]	1.04	0.90	0.94	1.12
Leu[5]/Leu[6]	1.03	0.95	0.86	1.16
Leu[6]/Ala[7]	1.36	0.53	1.07	1.04
Ala[7]/Leu[8]	1.18	1.25	0.74	0.83
Leu[8]/Pro[9]	1.11	0.82	1.14	0.93

The values are calculated as (# library members found / # library members expected). In cases involving Thr[3], the expected number of library members was equal to half the total number of library members in Library 1. Otherwise, the expected number of library members was one quarter of the total library members in Library 1.

***Table S16*: Average permeability of adjacent stereochemical configurations**

	L,L	L,D	D,L	D,D
Pro[1]/MeLeu[2]	-6.00	-5.81	-5.86	-6.15
MeLeu[2]/Thr[3]	-5.93		-5.98	
Thr[3]/Leu[4]	-6.10	-5.82		
Leu[4]/Leu[5]	-6.16	-6.02	-5.68	-5.93
Leu[5]/Leu[6]	-6.02	-5.84	-5.78	-6.12
Leu[6]/Ala[7]	-6.01	-5.67	-6.03	-5.95
Ala[7]/Leu[8]	-5.95	-6.08	-5.66	-6.03
Leu[8]/Pro[9]	-5.90	-5.75	-5.99	-6.15

The values are mean log(P_{app}) for each configuration.

***Table S17*: Representation of each diastereomer for Leu^6/Ala^7, Ala^7/Leu^8, and Leu^8/Pro^9 at each degree of N-methylation**

Leu^6/Ala^7	L,L	L,D	D,L	D,D
All	374	146	294	285
0 Me	1.04	0.00	1.40	1.05
1 Me	0.98	0.47	1.18	1.11
2 Me	1.07	0.85	0.94	1.05
3 Me	1.04	1.64	0.78	0.85
4 Me	0.71	1.75	1.03	0.96

Ala^7/Leu^8	L,L	L,D	D,L	D,D
All	325	343	204	227
0 Me	1.13	1.26	0.65	0.73
1 Me	1.08	1.05	0.67	1.09
2 Me	1.05	0.98	1.06	0.91
3 Me	0.87	0.97	1.25	1.01
4 Me	0.90	0.80	1.16	1.29

Leu^8/Pro^9	L,L	L,D	D,L	D,D
All	305	224	314	256
0 Me	0.95	0.94	1.03	1.08
1 Me	0.98	0.86	1.15	0.97
2 Me	1.02	1.09	0.95	0.96
3 Me	1.00	1.04	0.88	1.11
4 Me	1.03	0.97	1.12	0.85

Each field contains the number of library members with the indicated stereochemistry at each degree of N-methylation. The color corresponds to the number of library members found relative to the number of library members expected at each degree of N-methylation based on the total number of library members with the indicated stereochemistry. Green indicates higher than expected abundance while red indicates lower.

***Table S18*: Permeability deviation for each diastereomer of Leu^6/Ala^7, Ala^7/Leu^8, and Leu^8/Pro^9 at each degree of N-methylation**

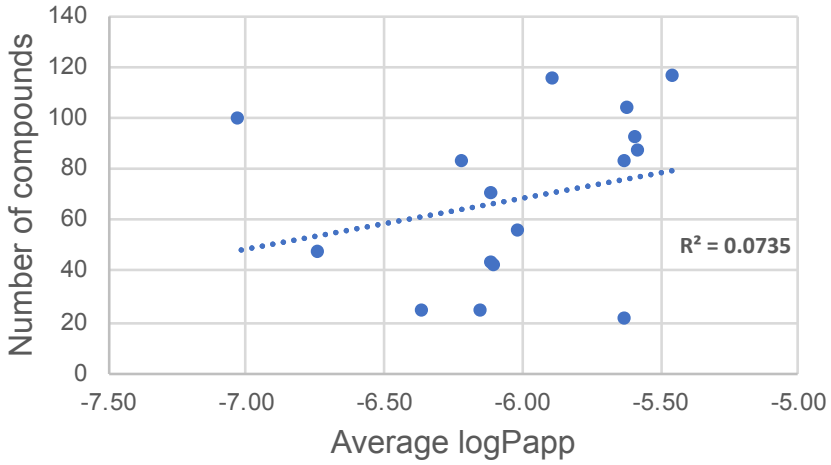
Leu^6/Ala^7	L,L	L,D	D,L	D,D
All	-0.05	0.29	-0.08	0.00
0 Me	0.05		-0.06	0.02
1 Me	-0.05	0.18	0.05	-0.03
2 Me	-0.11	0.14	0.05	0.04
3 Me	0.00	-0.01	-0.06	0.07
4 Me	0.03	0.03	-0.10	0.05

Ala^7/Leu^8	L,L	L,D	D,L	D,D
All	0.01	-0.13	0.29	-0.07
0 Me	0.17	-0.15	0.09	-0.05
1 Me	0.08	-0.08	0.42	-0.23
2 Me	-0.02	-0.07	0.24	-0.10
3 Me	0.07	-0.10	0.14	-0.10
4 Me	-0.06	-0.02	0.04	0.04

Leu^8/Pro^9	L,L	L,D	D,L	D,D
All	0.05	0.21	-0.04	-0.20
0 Me	0.19	0.09	0.25	-0.56
1 Me	0.06	0.35	0.01	-0.36
2 Me	-0.03	0.22	-0.01	-0.17
3 Me	0.07	0.15	-0.17	-0.04
4 Me	0.00	-0.03	0.02	0.00

Each field contains the difference between the average permeability ($\log P_{app}$) of library members with the indicated stereochemistry and the average permeability of all stereoisomers at each degree of N-methylation.

a



b

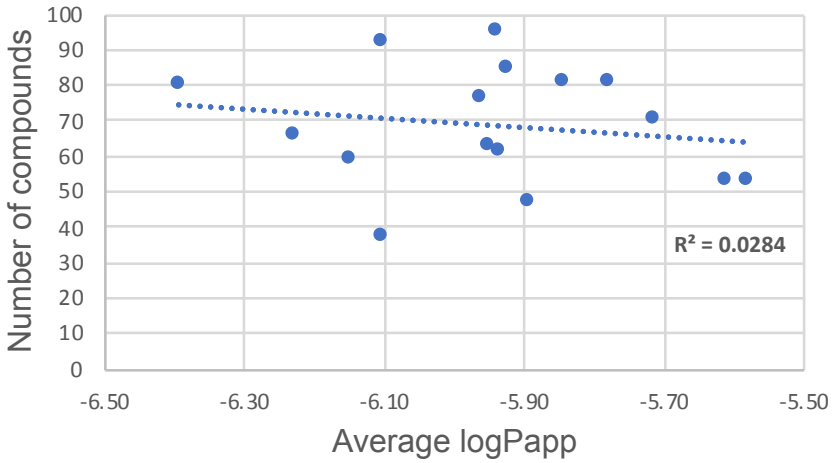


Figure S3. **Number of identified compounds in data set vs. average $logP_{app}$ from each (a) methylation pattern and (b) sub-library**

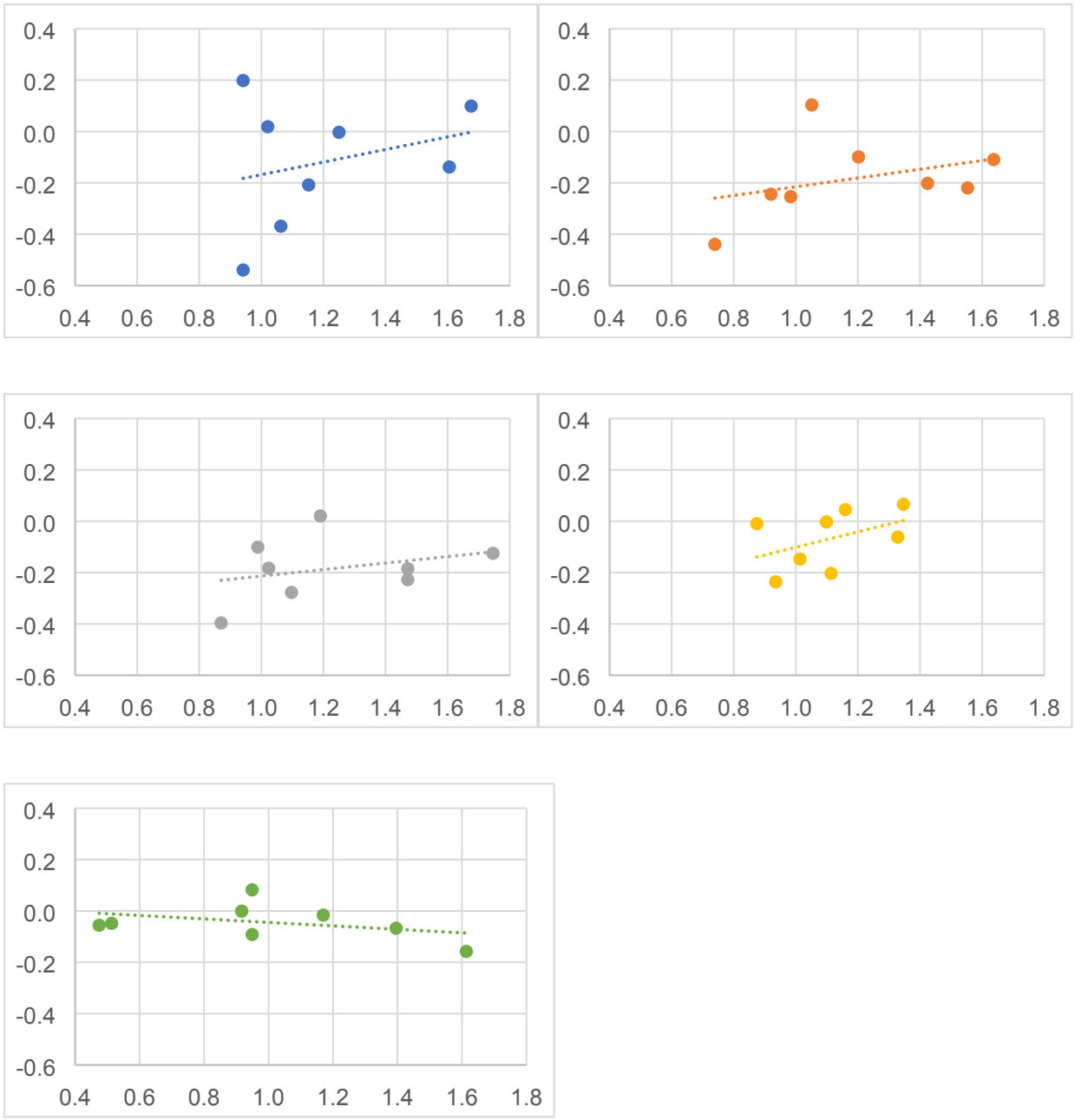


***Figure S4*. Correlation between permeability of compounds with heterochiral stereochemistry and number of compounds with heterochiral stereochemistry**

Each point represents an adjacent stereochemical pair for a given number of N-methyl groups. The y-axis values correspond to the values in Table S13 (permeability effect). The x-axis values correspond to the values in Table S14 (representation disparity). Blue: 0 N-methyl groups in the macrocycle, orange: 1 N-methyl group in the macrocycle, gray: 2 N-methyl groups in the macrocycle, yellow: 3 N-methyl groups in the macrocycle, green: 4 N-methyl groups in the macrocycle.

***Table S19*: Average retention time and retention time variance by degree of N-methylation**

Degree of N-methylation	Average retention time (min)	Variance (min)
0	12.95	9.24
1	18.85	41.82
2	24.22	64.57
3	28.71	55.03
4	32.54	26.37

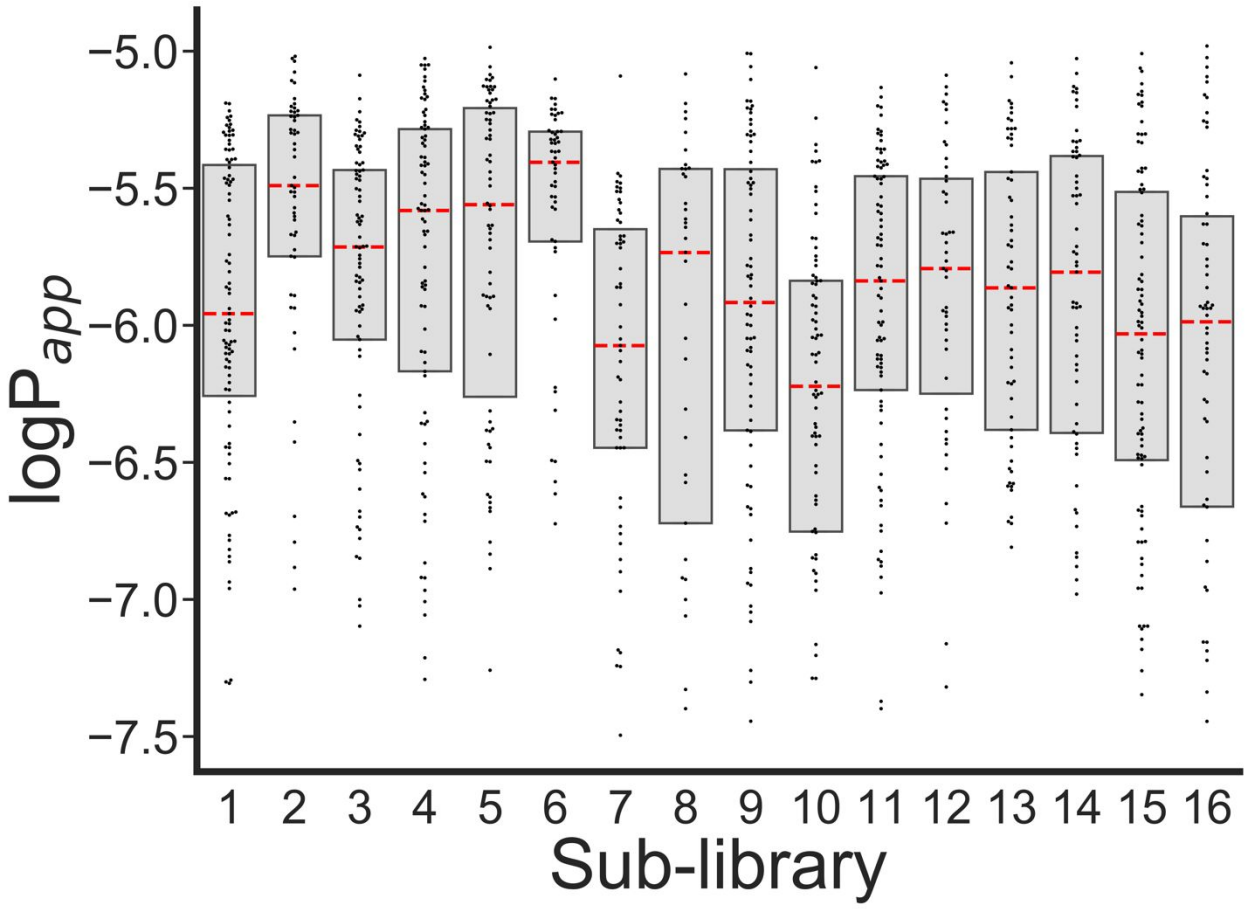


***Figure S5*. Permeabilities of compounds in each sub-library of Library 1.**
The red dashed lines represent medians and the boxes represent quartiles. The red dashed lines represent medians and the boxes represent quartiles. $LogP_{app}$ values below -7.5 were not included in the swarm plot but were used to calculate median and quartiles.

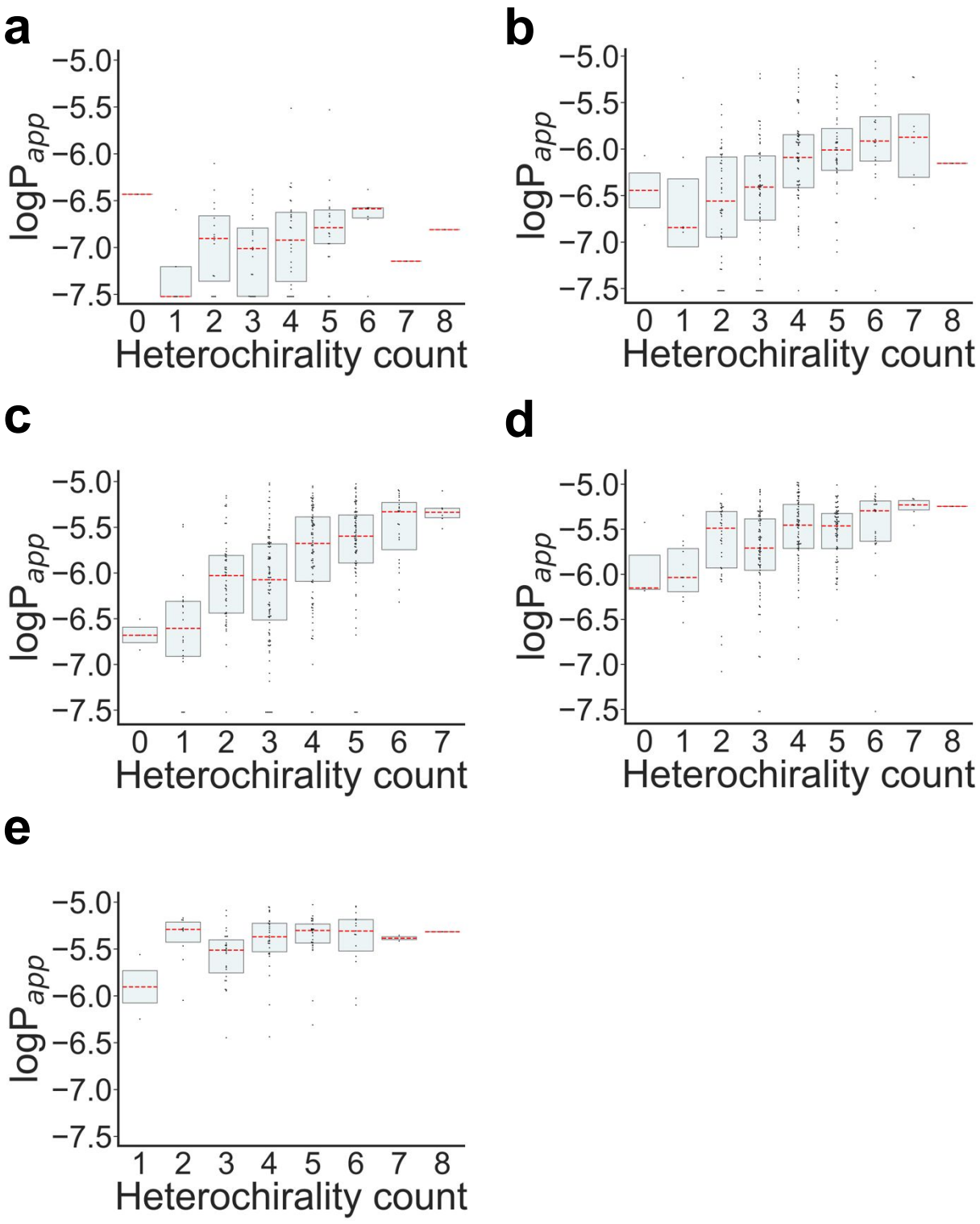


Figure S6. **Effect of heterochirality on permeability by degree of N-methylation** (a) 0 N-methyl groups in the macrocycle, (b) 1 N-methyl group in the macrocycle, (c) 2 N-methyl groups in the macrocycle, (d) 3 N-methyl groups in the macrocycle, (e) 4 N-methyl groups in the macrocycle. The red dashed lines represent medians and the boxes represent quartiles. LogP_{app} values below -7.5 were not included in the swarm plot but were used to calculate median and quartiles.

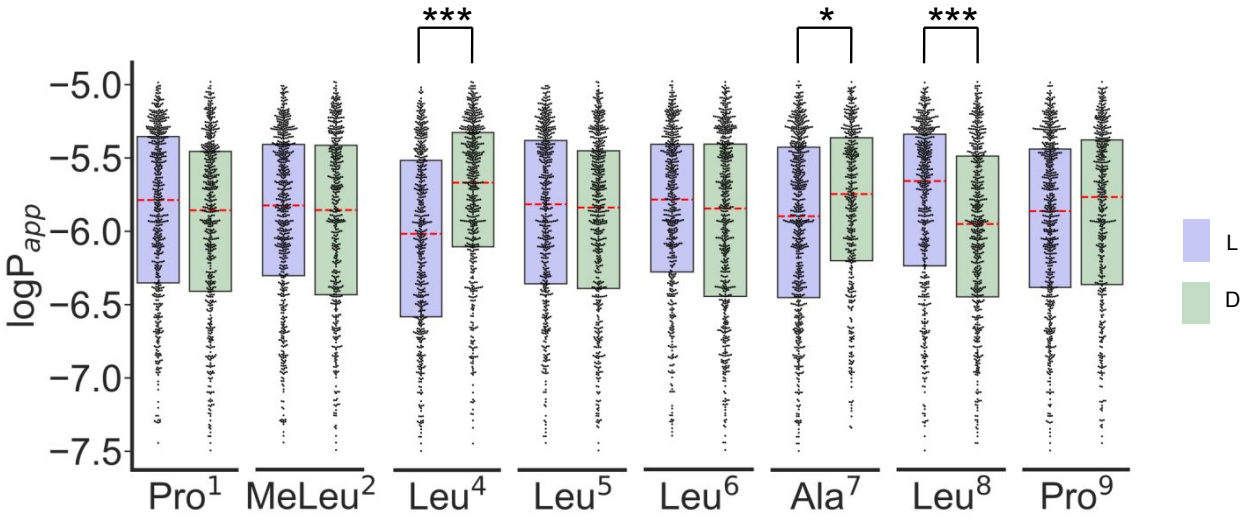


Figure S7. **Effect of stereochemistry at each position on permeability of compounds in Library 1.** The red dashed lines represent medians and the boxes represent quartiles. Statistics are as follows: Mann-Whitney U test; ***$P < 0.0001$, **$P < 0.001$, *$P < 0.01$. LogP$_{app}$ values below -7.5 were not included in the swarm plot but were used to calculate median and quartiles.

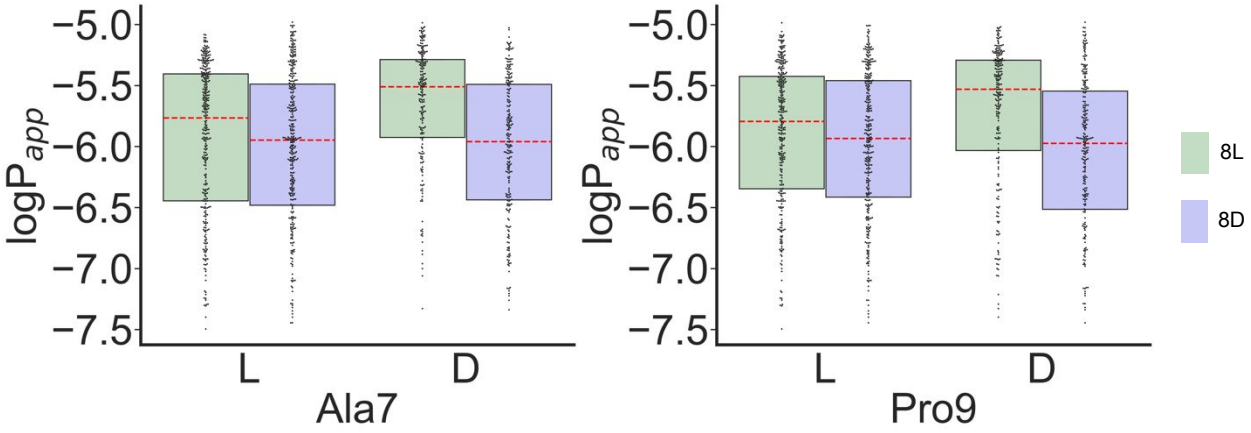


Figure S8. **Effect of relative stereochemistry between Leu8 and adjacent residues on permeability** The red dashed lines represent medians and the boxes represent quartiles. The statistical significance of the effect is listed in Table S20. LogP$_{app}$ values below -7.5 were not included in the swarm plot but were used to calculate median and quartiles.

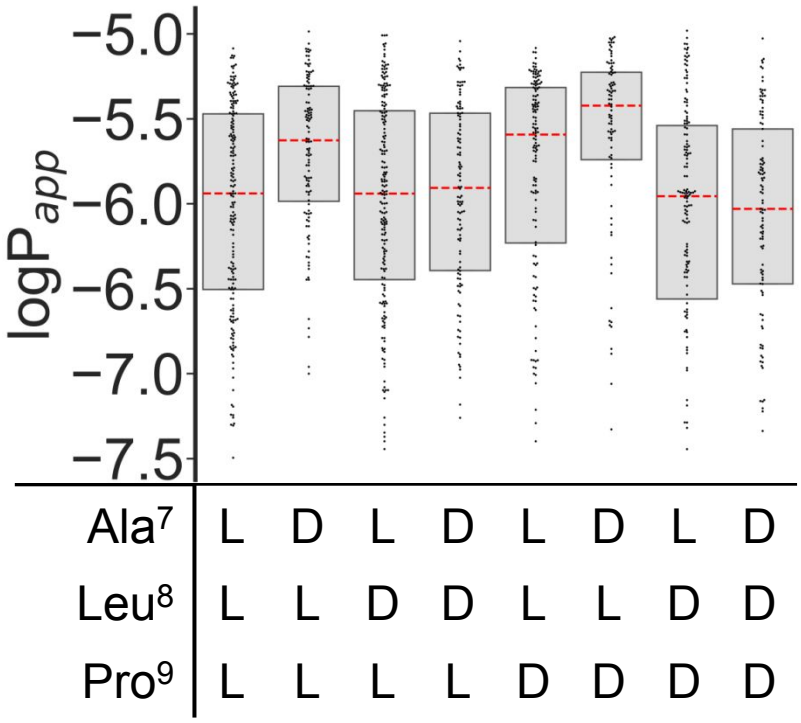


Figure S9. **Effect of stereochemistry at residues 7-9.** Considering the stereochemistry of residues 7-9, three of eight configurations have higher permeability than the others. All three of these have Leu8=L and at least one instance of heterochirality with an adjacent residue. The red dashed lines represent medians and the boxes represent quartiles. LogP_{app} values below -7.5 were not included in the swarm plot but were used to calculate median and quartiles.

***Table S20*. Mann-Whitney U-tests evaluating significance of effects of relative stereochemistry between Leu[8] and adjacent residues on permeability.**

All pairwise comparisons were performed. Tests yielding p-values < 0.001 were considered significant (more permeable diastereomer in green).

Diastereomer		p-value (Mann-Whitney U-test)
7L, 8L	7L, **8D**	0.029
7D 8L	7D **8D**	2.01×10^{-11}
7L 8L	**7D** 8L	1.23×10^{-7}
7L 8D	**7D** 8D	0.48
7L 8L	**7D 8D**	0.034
7L 8D	**7D 8L**	9.59×10^{-12}
8L 9L	**8D** 9L	0.053
8L 9D	**8D** 9D	1.09×10^{-10}
8L 9L	8L **9D**	4.04×10^{-5}
8D 9L	8D **9D**	0.042
8L 9L	**8D 9D**	4.95×10^{-4}
8D 9L	**8L 9D**	1.55×10^{-7}

NMR Solution Structure Generation for Compound 2

A 2D ROESY spectrum of compound **2** was obtained at 277K in chloroform-*d* with a mixing time of 300ms, which was in the linear range in cross-relaxation ROESY buildup curve as determined by performing separate 1D ROESY experiments with mixing times of 100, 200, 300, 400, 500, and 600 ms. Cross peaks were classified as strong, medium, or weak by visual inspection. The following distance restraints were applied:

Strong	1.7 – 2.5 Å
Medium	2.5 – 3.5 Å
Weak	3.5 – 4.5 Å

We avoided ROESY crosspeaks involving leucine sidechain atoms when selecting distance restraints due to the high level of peak overlap in the upfield region of the spectrum. In total, thirteen distance restraints were applied to the agnostic conformer pool (Table S18).

The amide resonances were sufficiently sharp to obtain HN-Hα *J*-coupling values for calculating dihedral restraints. The Karplus relationship was used to obtain estimates of the ϕ dihedral angles. Theses ϕ angle values are listed in Table S13.

To determine the solution structure of **2**, we used the ForceGen approach [7]. Beginning with an initial set of NMR constraints, some of which might be degenerate and some of which might be incorrect, an initial conformer pool is produced without the use of NMR data. That pool is profiled against the full set of NMR constraints, allowing for selection of a subset of

non-degenerate constraints that are consistent with some of the conformers in the pool. In an iterative process, new conformer pools are produced using a set of NMR constraints that explores the space of additions to the prior set of constraints where the new constraints are shown to be feasible from the prior conformer pool. In cases where multiple possibilities exist for a particular constraint, all of which are feasible, they are all explored. The process ends when all choices from among degenerate constraints are made and when no non-degenerate constraints are feasible to add. In this case, all non-degenerate constraints were selected for the final set, and a single choice was made for each degenerate constraint.

***Table S21*. Peak assignments for Compound 2**

Amino acid	Atom group	Proton (^{1}H) δ (ppm)	Carbon (^{13}C) δ (ppm)
Pro1	Carbonyl-C-Acetyl		169.27
	Acetyl-Methyl	2.01	22.37
	α	4.66	56.13
	β	2.17	28.89
	β	1.95	28.83
	γ	2.19	25.53
	γ	1.96	25.53
	δ	3.65	48.11
	δ	3.56	48.14
	Carbonyl-C		173.49
MeLeu2	N-methyl	3.04	31.32
	α	5.35	54.90
	β	1.87	35.97
	β	1.55	35.96
	γ	1.38	24.88
	δ	0.90	?
	δ	0.87	?
	Carbonyl-C		170.57
Thr3	NH	7.55	
	α	4.60	55.49
	β	4.99	70.91
	γ	1.32	14.79
	Carbonyl-C		168.65
Leu4	NH	7.94	
	α	4.56	52.01
	β	1.71	40.78
	β	1.31	40.78
	γ	1.75	24.70
	δ	0.92	?

	δ	0.88	?
	Carbonyl-C		173.92
Leu5	NH	6.91	
	α	4.46	50.15
	β	1.83	37.92
	β	1.54	37.79
	γ	1.59	24.68
	δ	0.90	?
	δ	0.88	?
	Carbonyl-C		173.18
Leu6	NH	7.42	
	α	4.99	45.79
	β	1.72	40.78
	β	1.58	41.13
	γ	1.48	24.76
	δ	0.92	?
	δ	0.88	?
	Carbonyl-C		171.12
MeAla7	N-methyl	2.67	28.09
	α	5.10	54.44
	β	1.31	15.28
	Carbonyl-C		169.18
Leu8	NH	7.82	
	α	4.69	48.73
	β	1.70	40.55
	β	1.52	40.44
	γ	1.77	24.70
	δ	1.01	23.30
	δ	0.96	21.87
	Carbonyl-C		172.03
Pro9	α	4.11	59.41
	β	2.25	28.93
	β	2.03	28.89
	γ	2.25	25.45
	γ	2.03	25.35
	δ	4.08	47.48
	δ	3.68	47.55
	Carbonyl-C		171.53

***Table S22*. ROE-distance restraints of Compound 2 in $CDCl_3$. Distances are presented as ranges**

Atom 1	Atom 2	Experimental Atomic Distance (Å)	Predicted Atomic Distance (Å)	Violation
Leu4.NH	Leu5.NH	2.5-3.5	2.4192	0.127
Leu4.NH	Thr3.HA	2.5-3.5	2.4172	0.0207
Leu4.NH	Pro9.HA	3.5-4.5	3.1653	0.1485
Thr3.NH	Leu2.HA	2.5-3.5	2.7316	0.0945
Leu8.NH	Ala7.HA	1.7-2.5	2.2205	0
Leu6.NH	Leu5.HA	3.5-4.5	2.5213	0.7418
Ala7.HA	Leu6.HA	1.7-2.5	1.9413	0
Leu5.NH	Thr3.HA	3.5-4.5	3.4473	0.0888
Leu2.HA	Leu2.QNMe	3.5-4.5	3.4802	0.5238
Pro1.HA	Leu2.QNMe	2.5-3.5	2.8647	0.3992
Leu8.HA	Pro9.HD	1.7-2.5	2.1925	0
Pro1.HD	Ac.CH3	2.5-3.5	3.5565	0.0184
Ala7.QB	Ala7.QNMe	1.7-2.5	2.8425	0.0925

***Table S23*: Amide NH temperature shift coefficients for Compound 2**

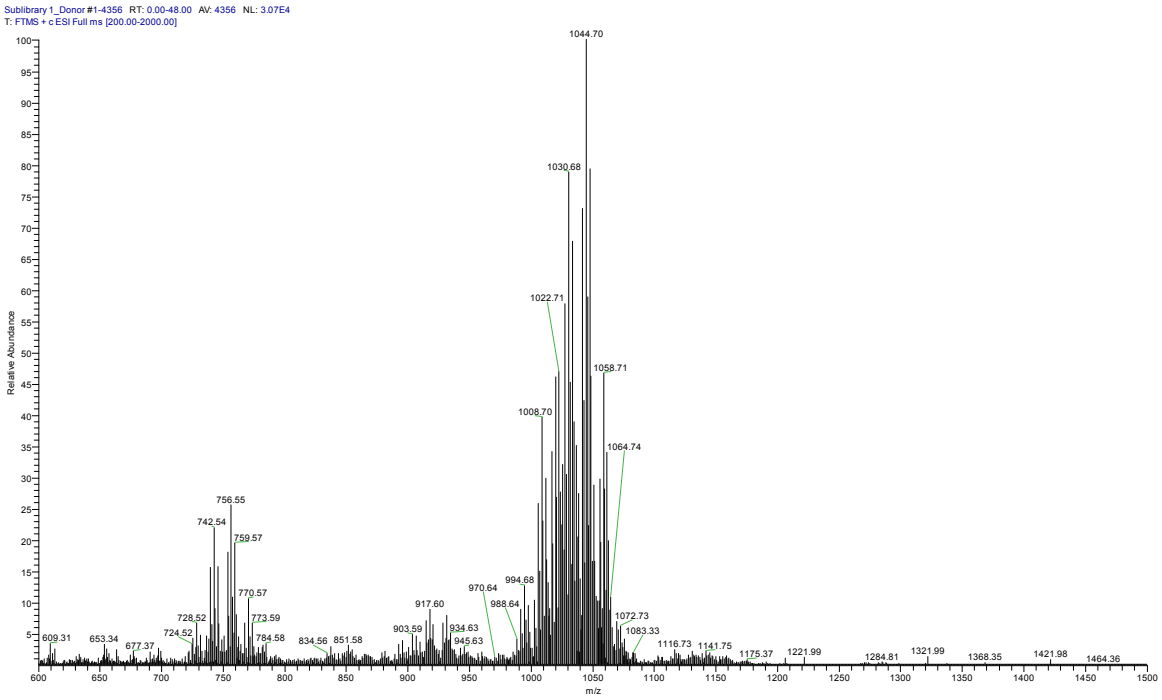
	Thr^3	Leu^4	Leu^5	Leu^6	Leu^8
Δδ/ΔT (ppb/K)	-3.2	-0.8	-1.6	-3.2	-2.4

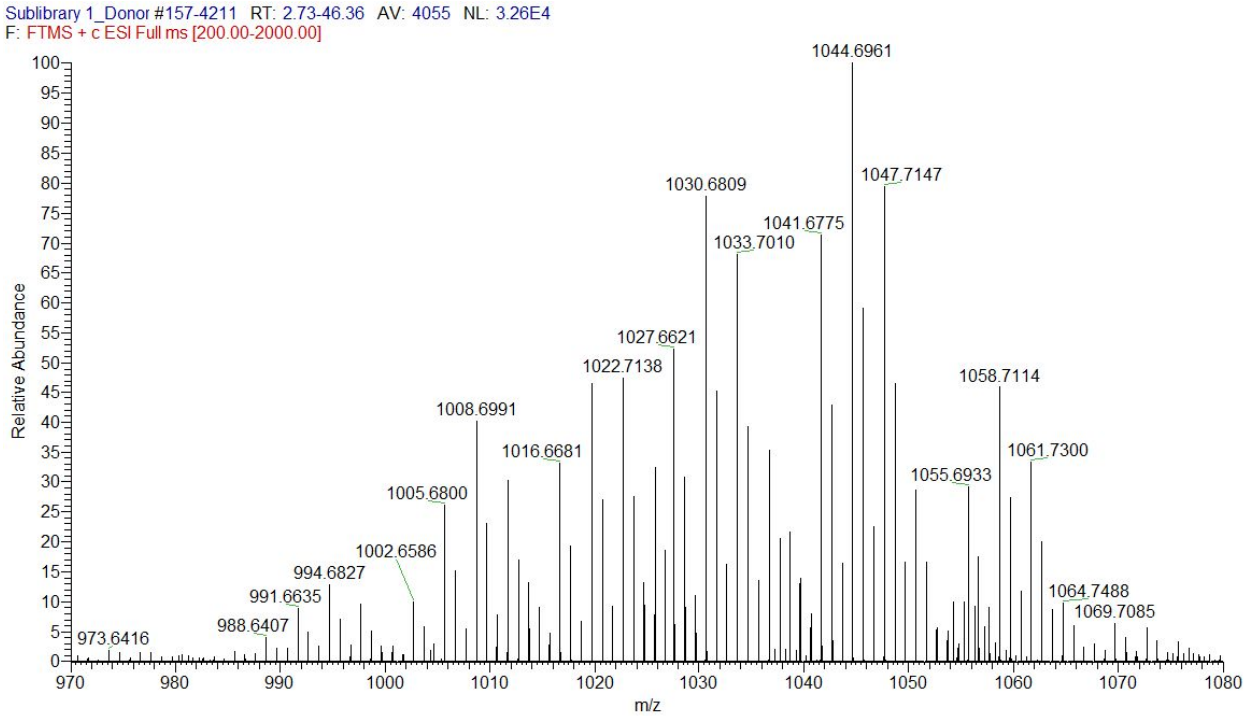
***Table S24*: ^3J vicinal coupling constants between Hα-HN protons and derived ϕ dihedral restraints for Compound 2**

	^{3}J coupling constant (Hz)	Estimated allowed ϕ angles	Predicted ϕ angles
Thr^3	7.14	-158, -82, 50, 70	-92
Leu^4	8.64	94, 146	101
Leu^5	8.22	-149, -90	-101
Leu^6	9.24	101, 139	110
Leu^8	7.92	-152, -88	-95

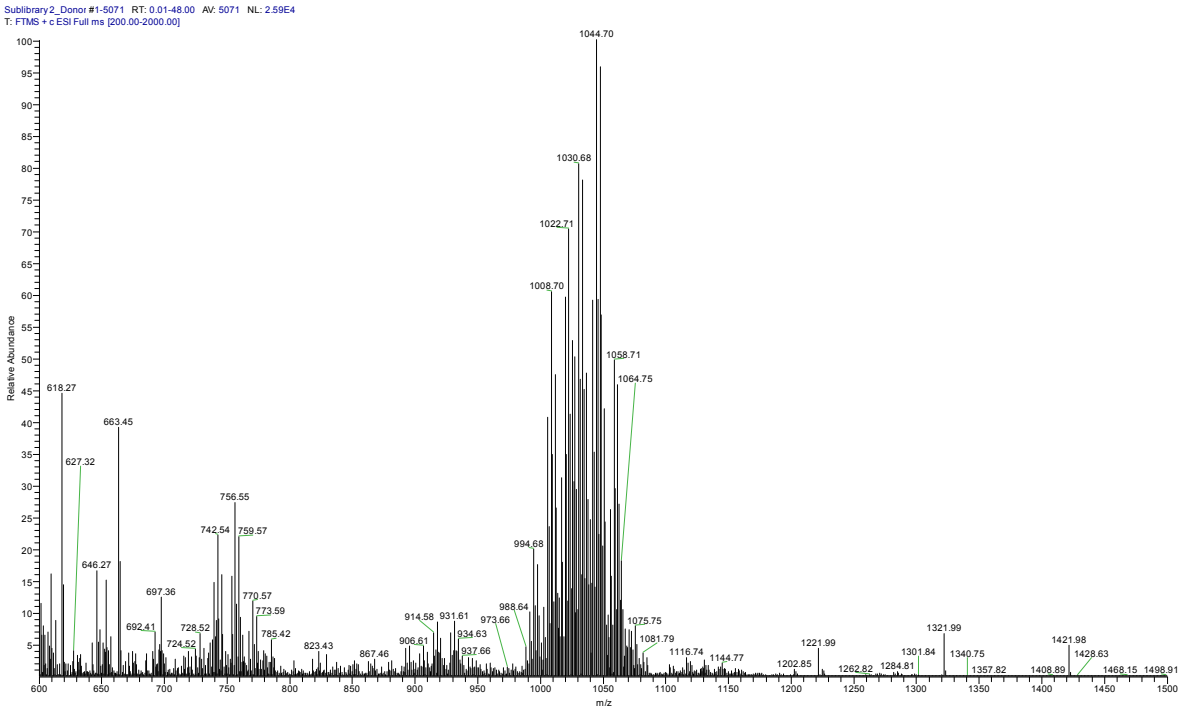
Mass spectra of the sub-libraries

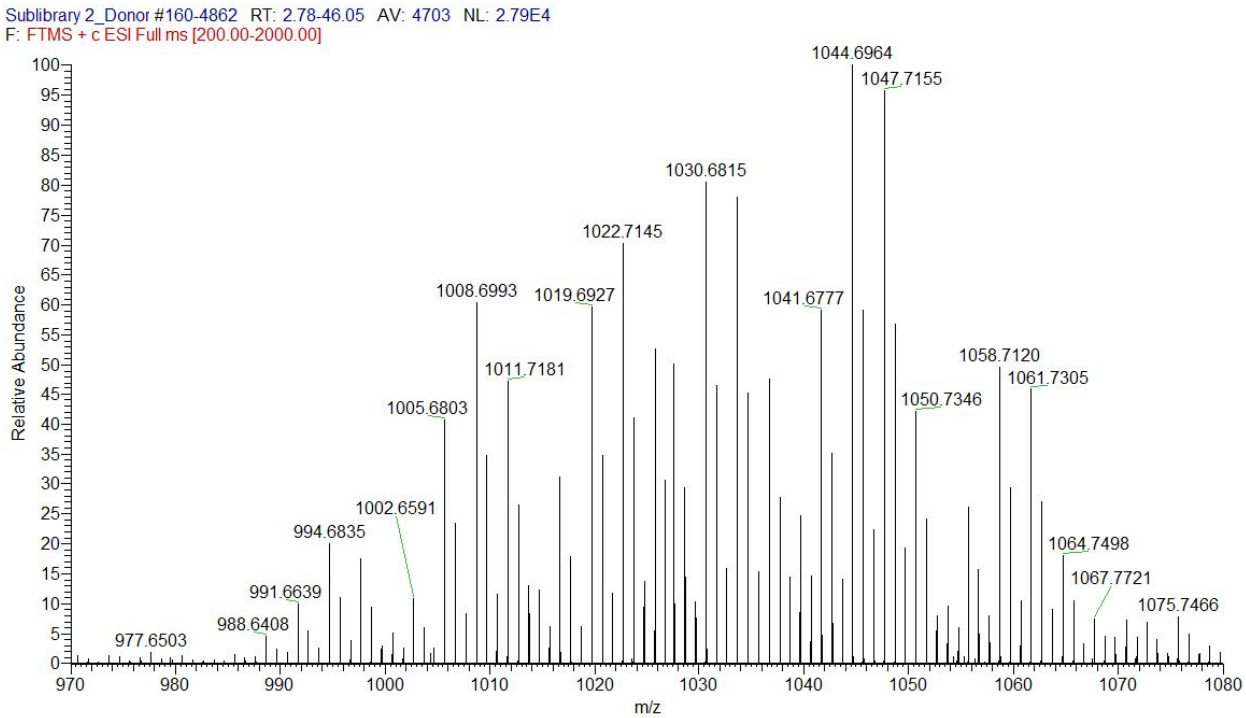
Mass spectrum of sub-library 1



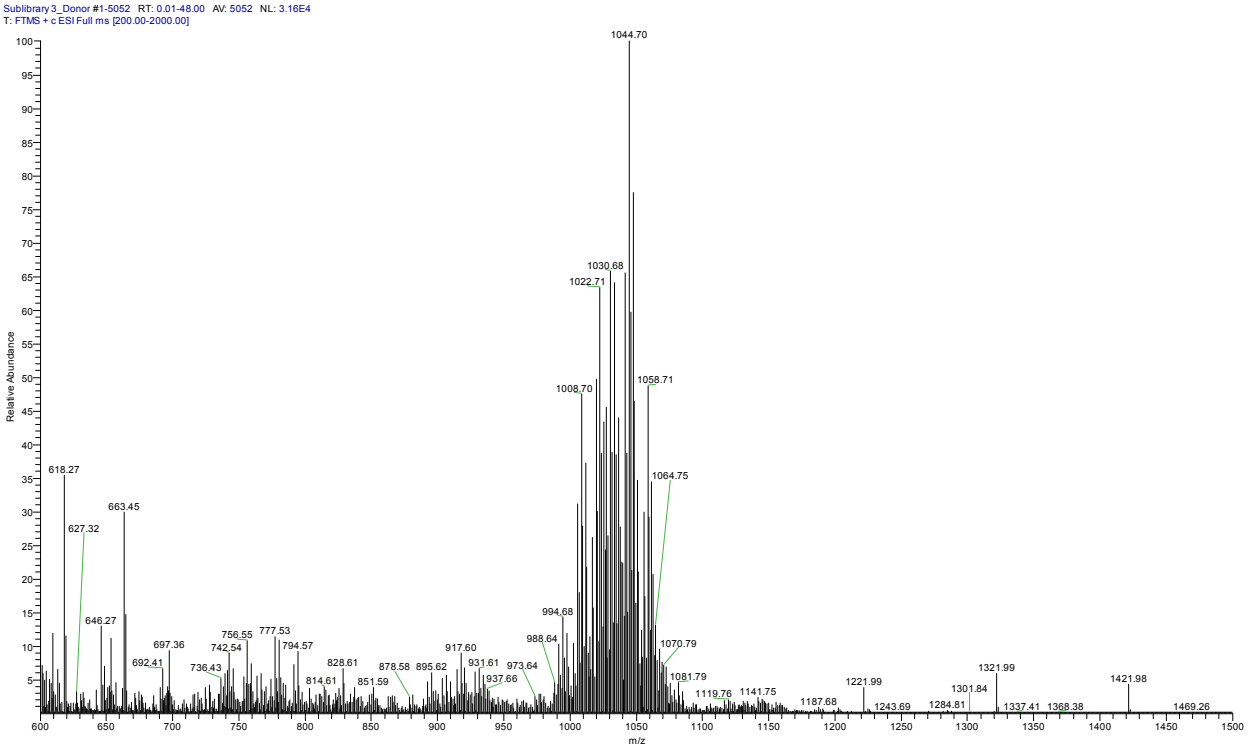


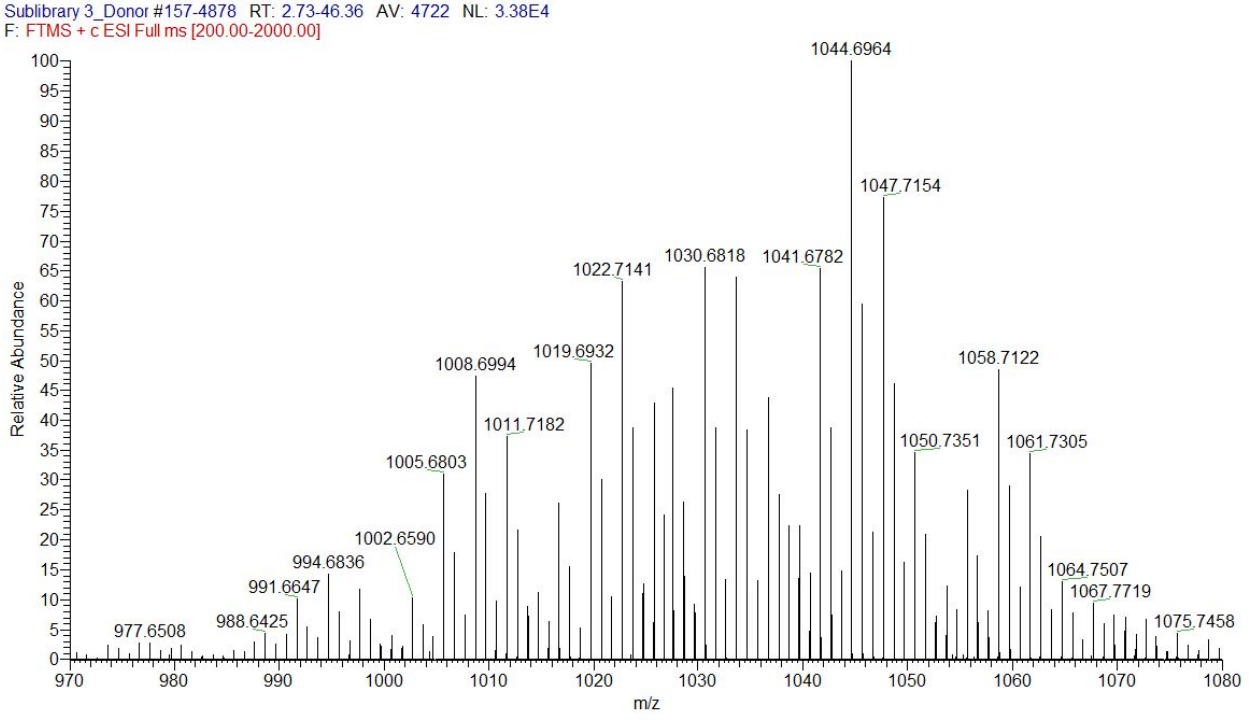
Mass spectrum of sub-library 2



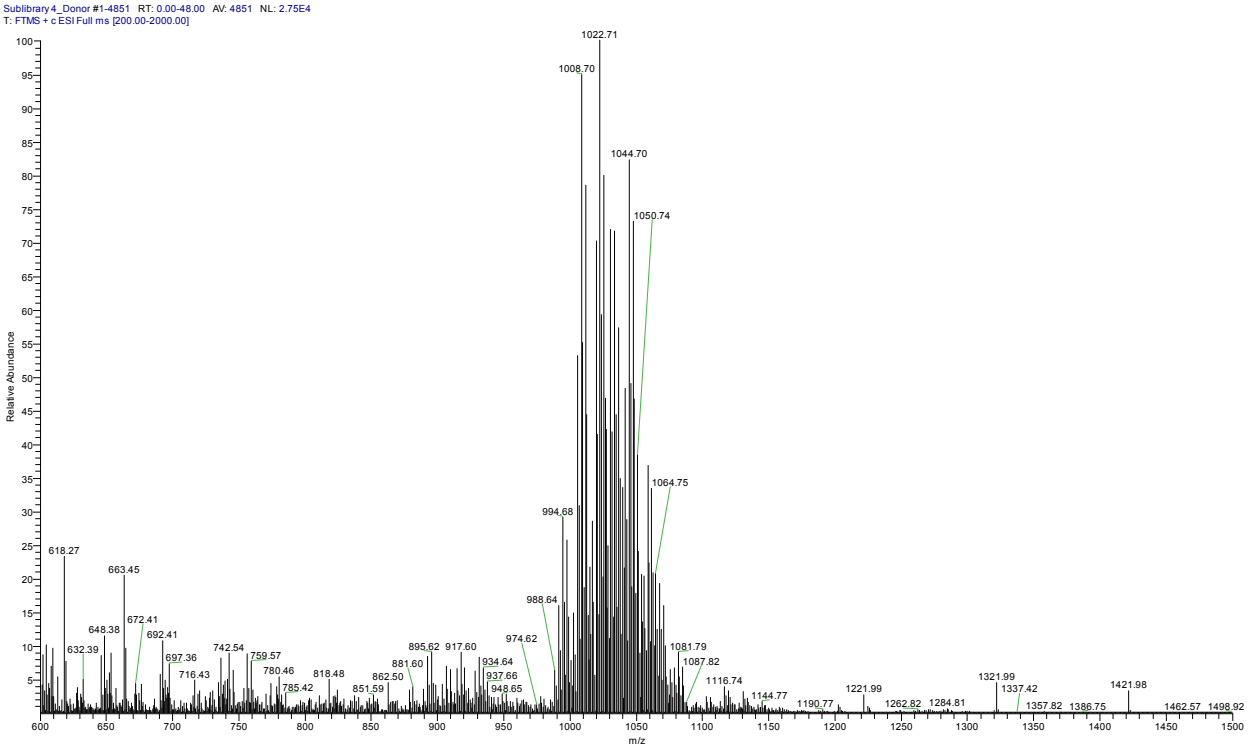


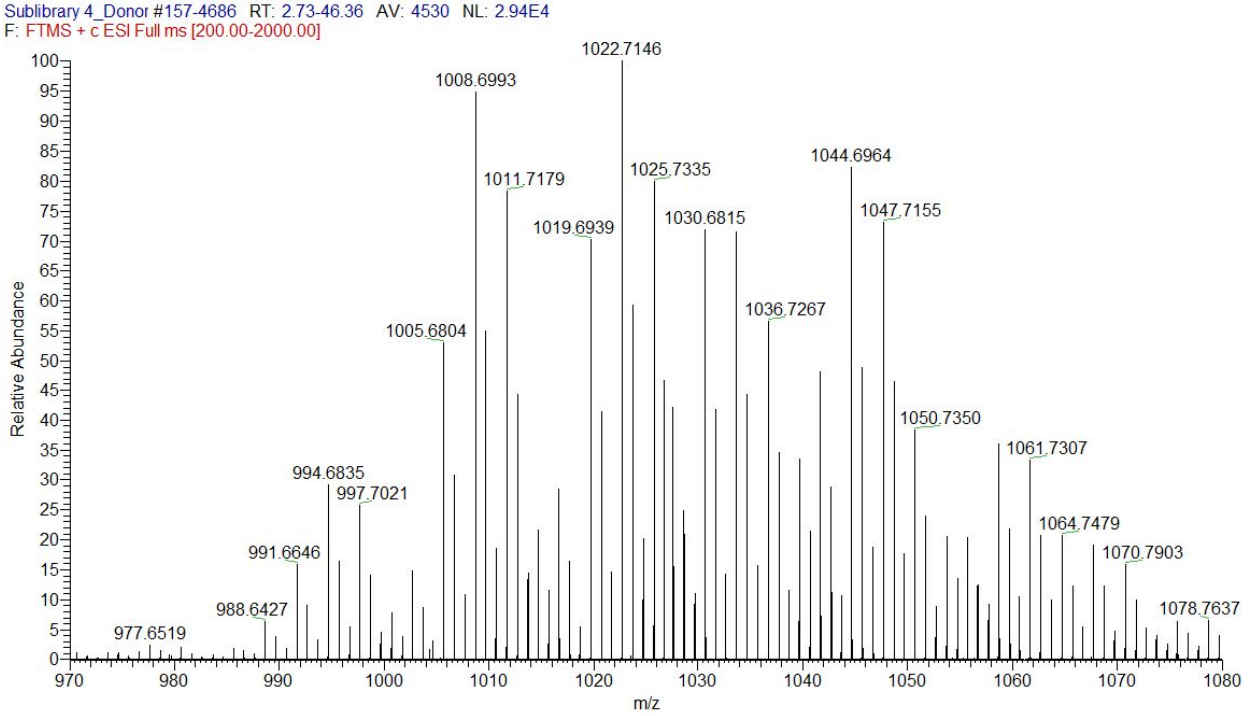
Mass spectrum of sub-library 3



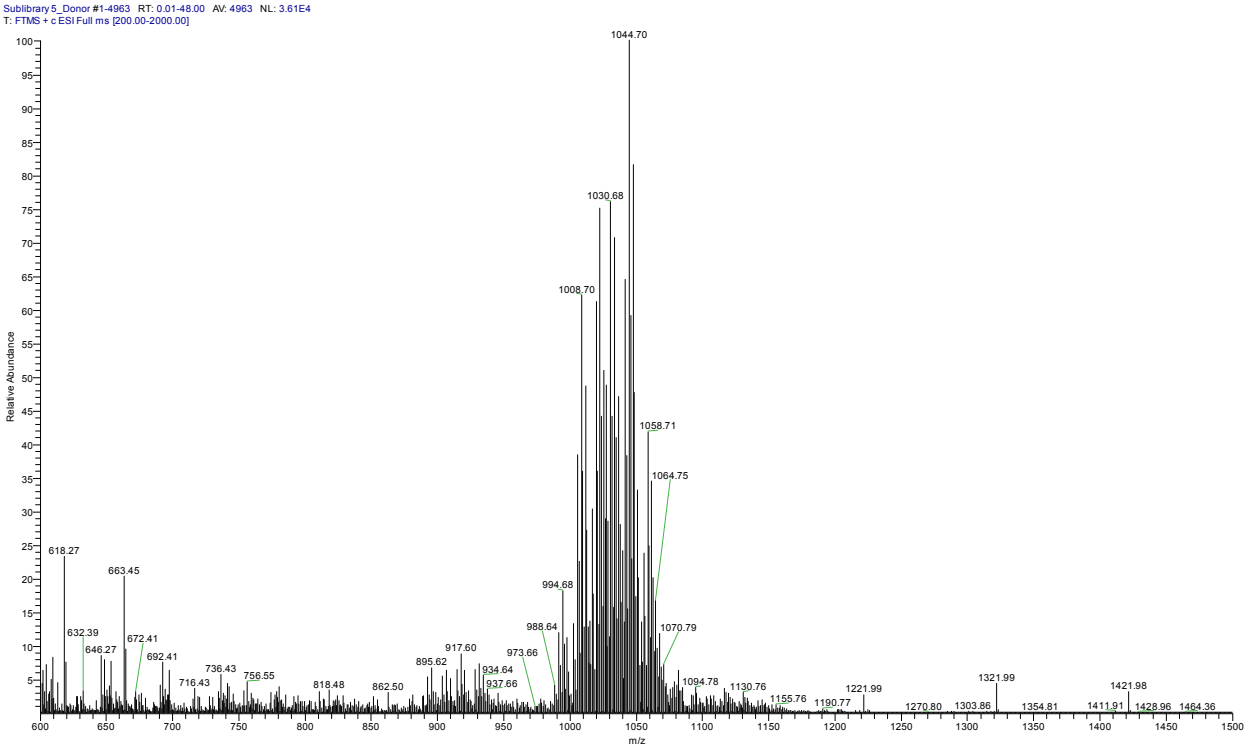


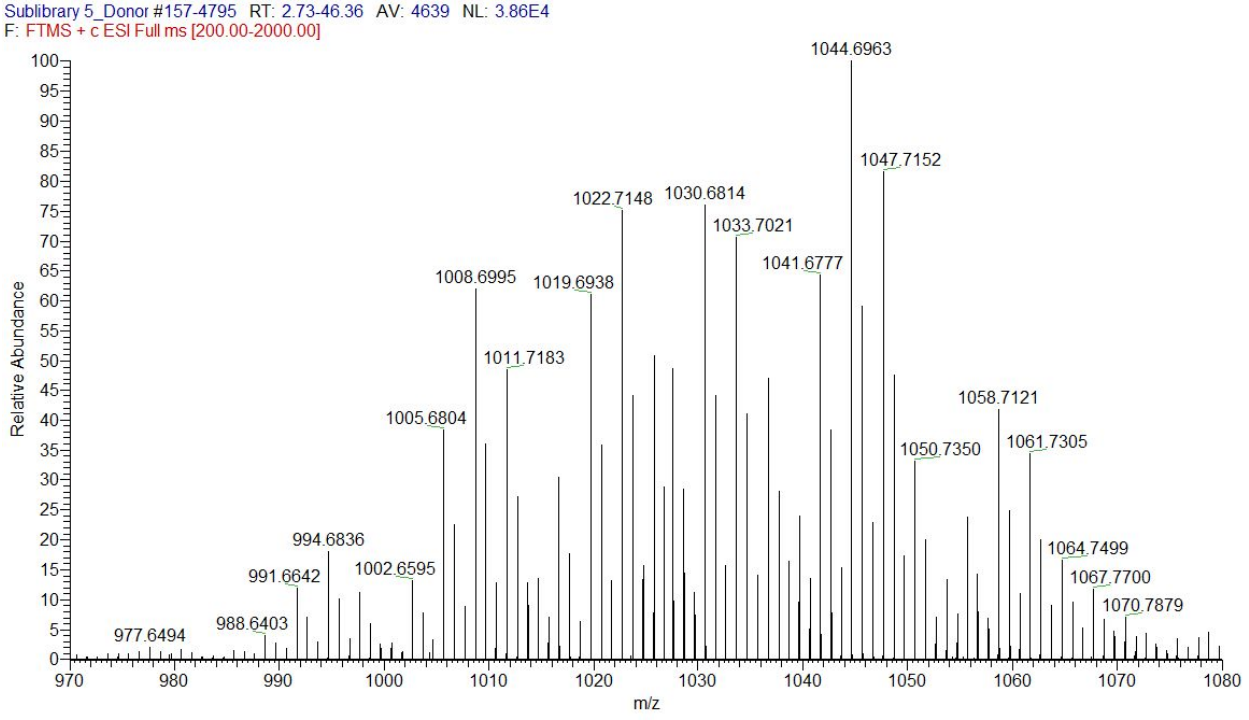
Mass spectrum of sub-library 4



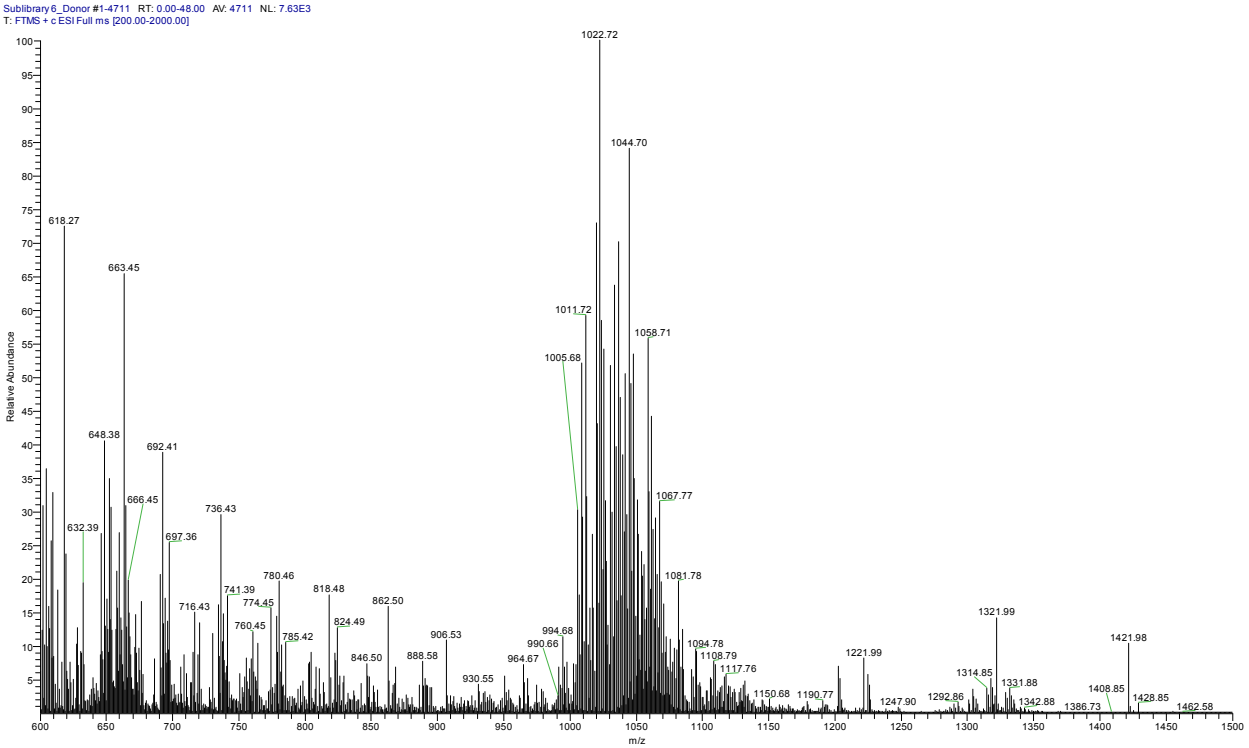


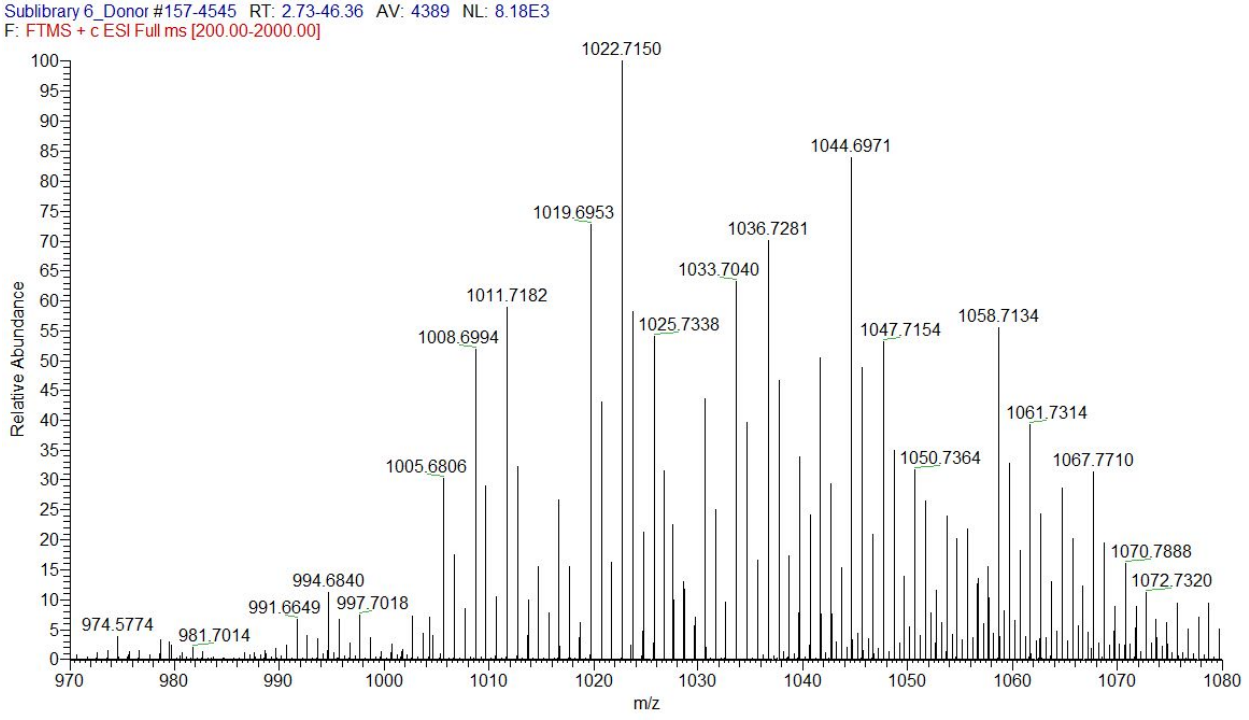
Mass spectrum of sub-library 5



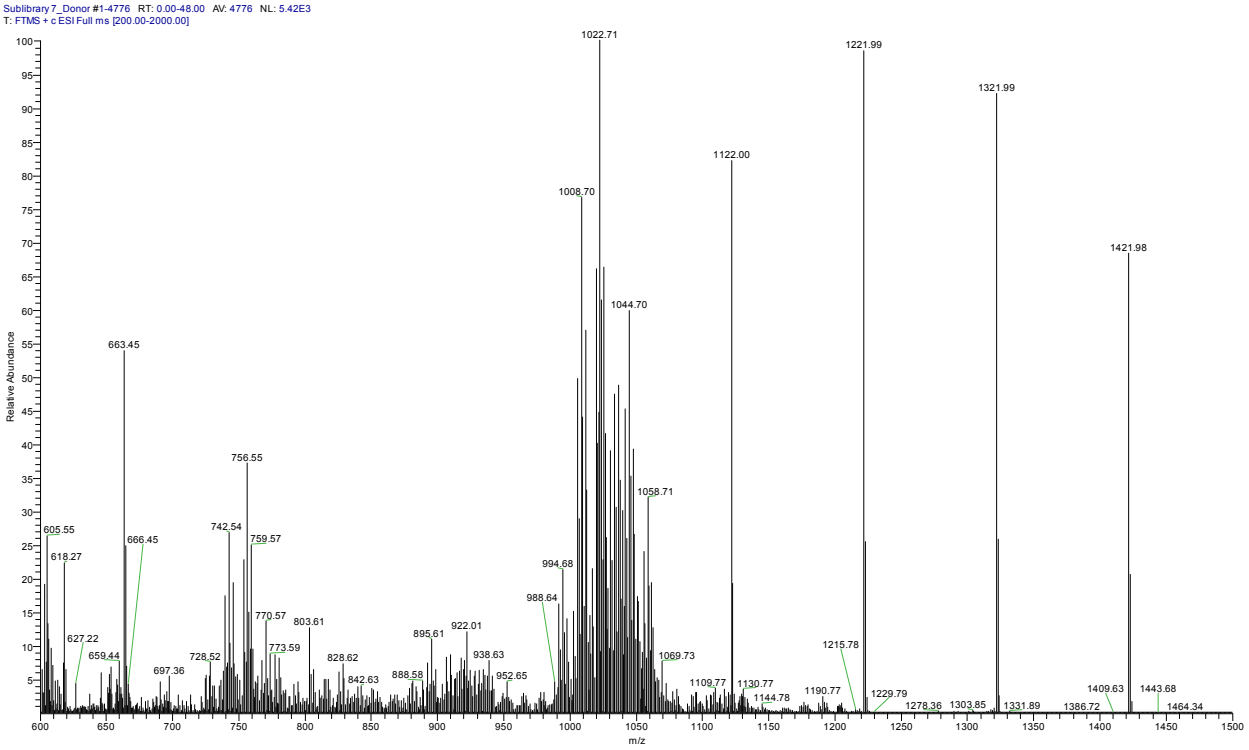


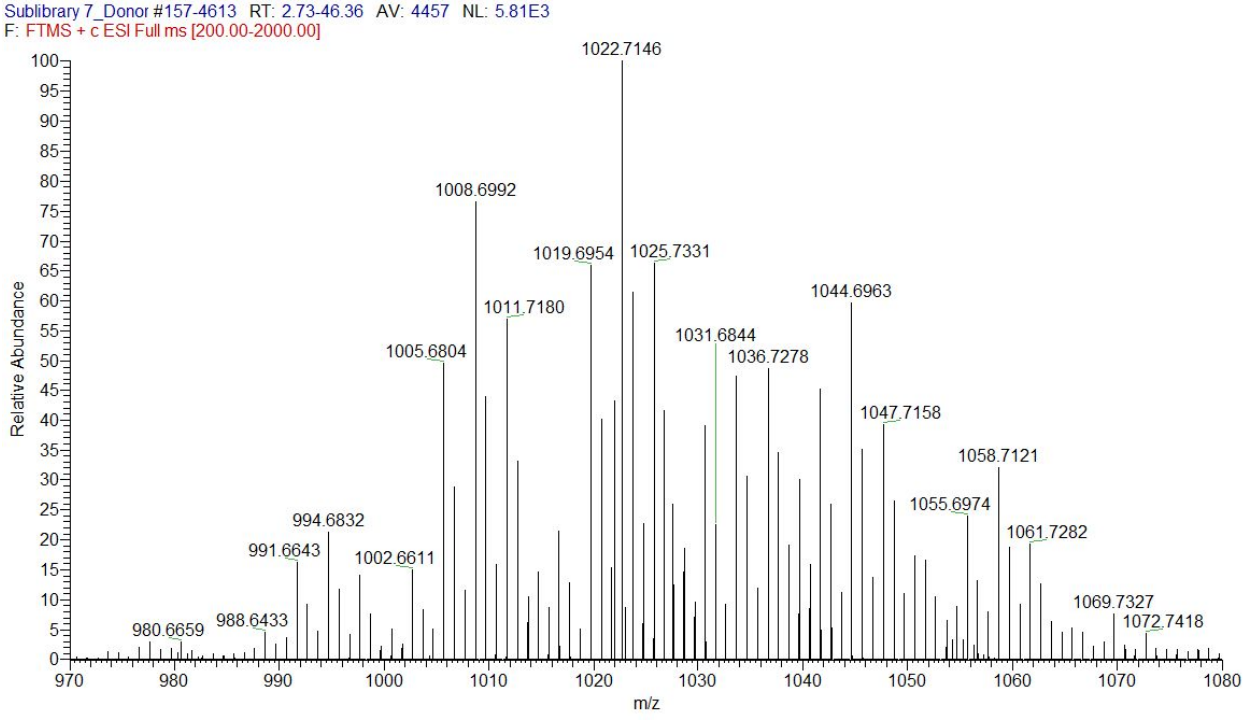
Mass spectrum of sub-library 6



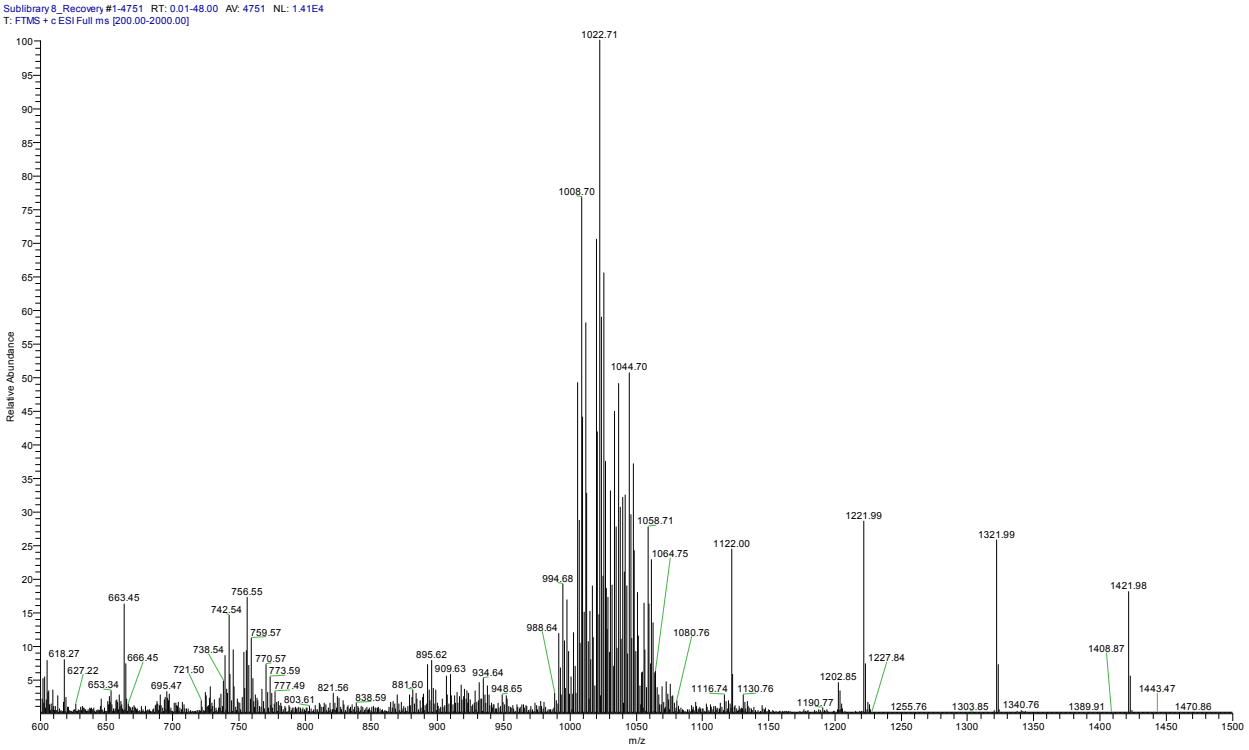


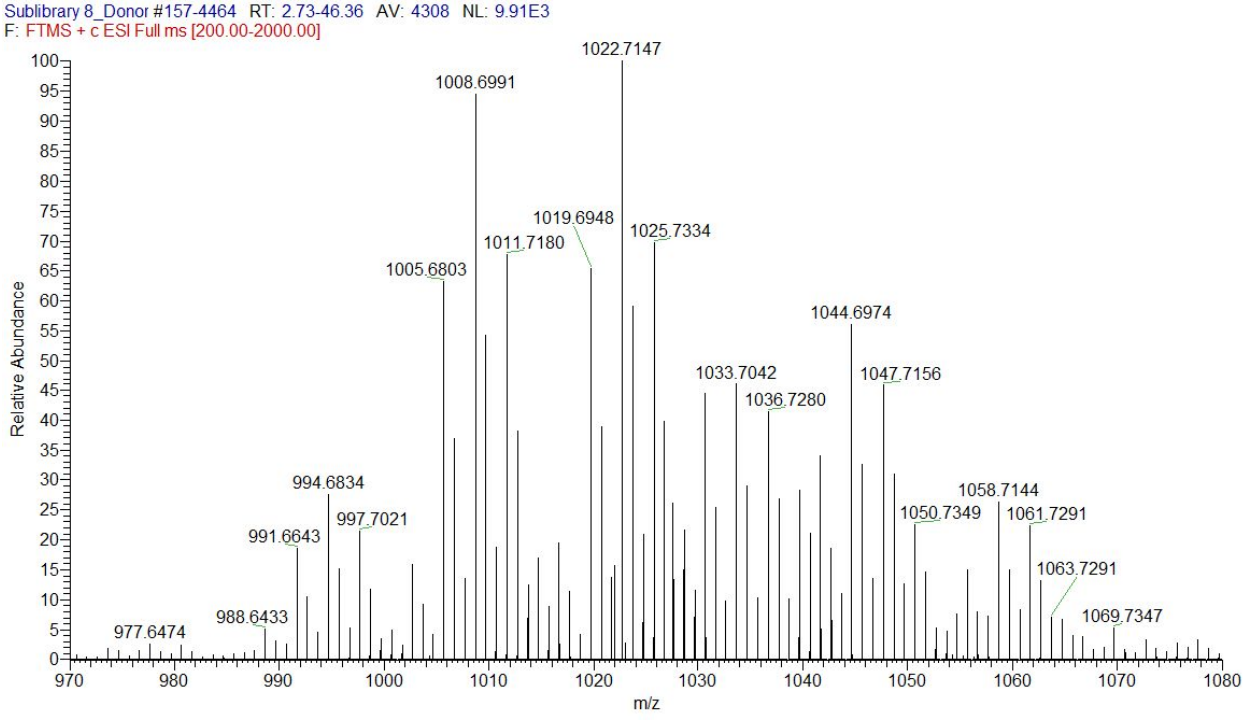
Mass spectrum of sub-library 7



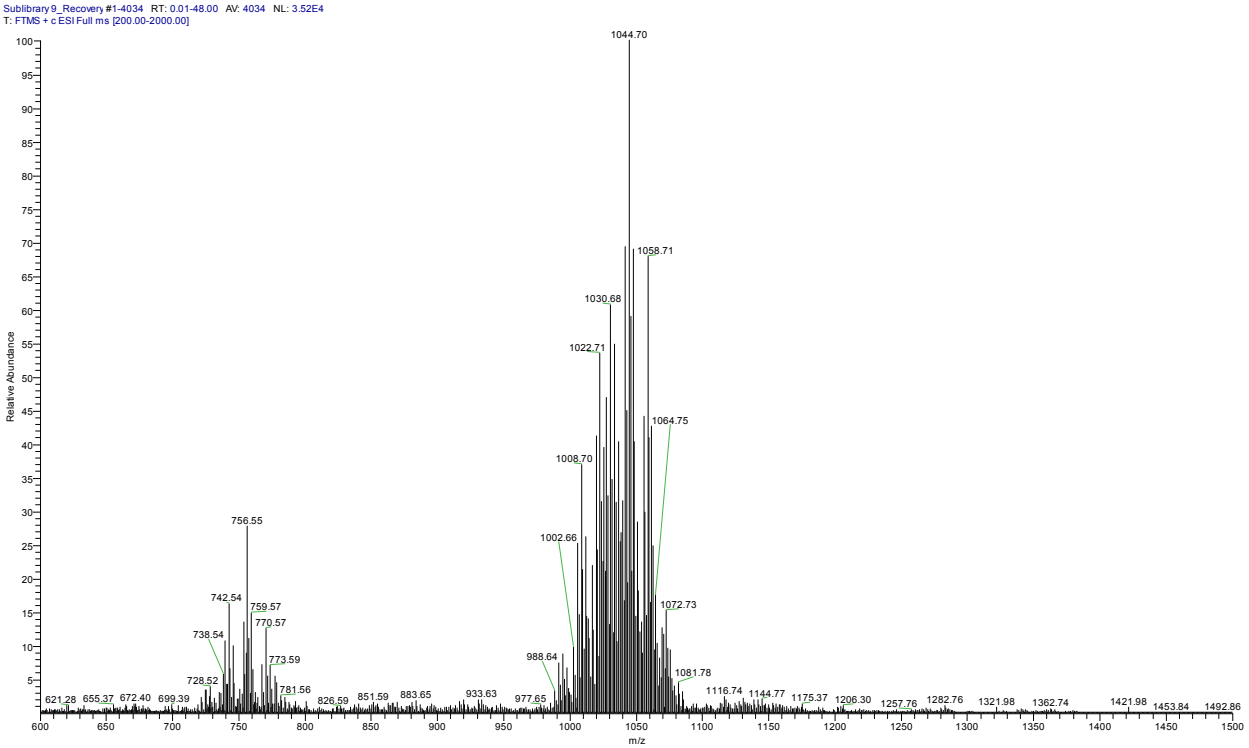


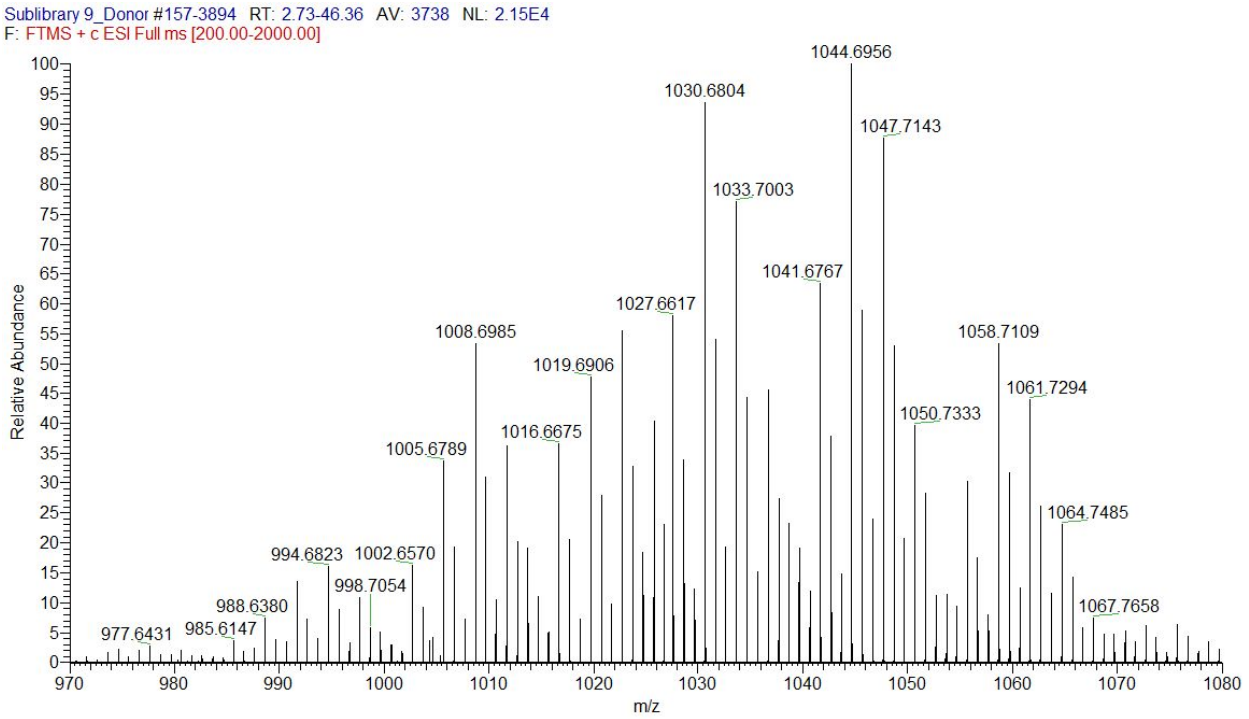
Mass spectrum of sub-library 8



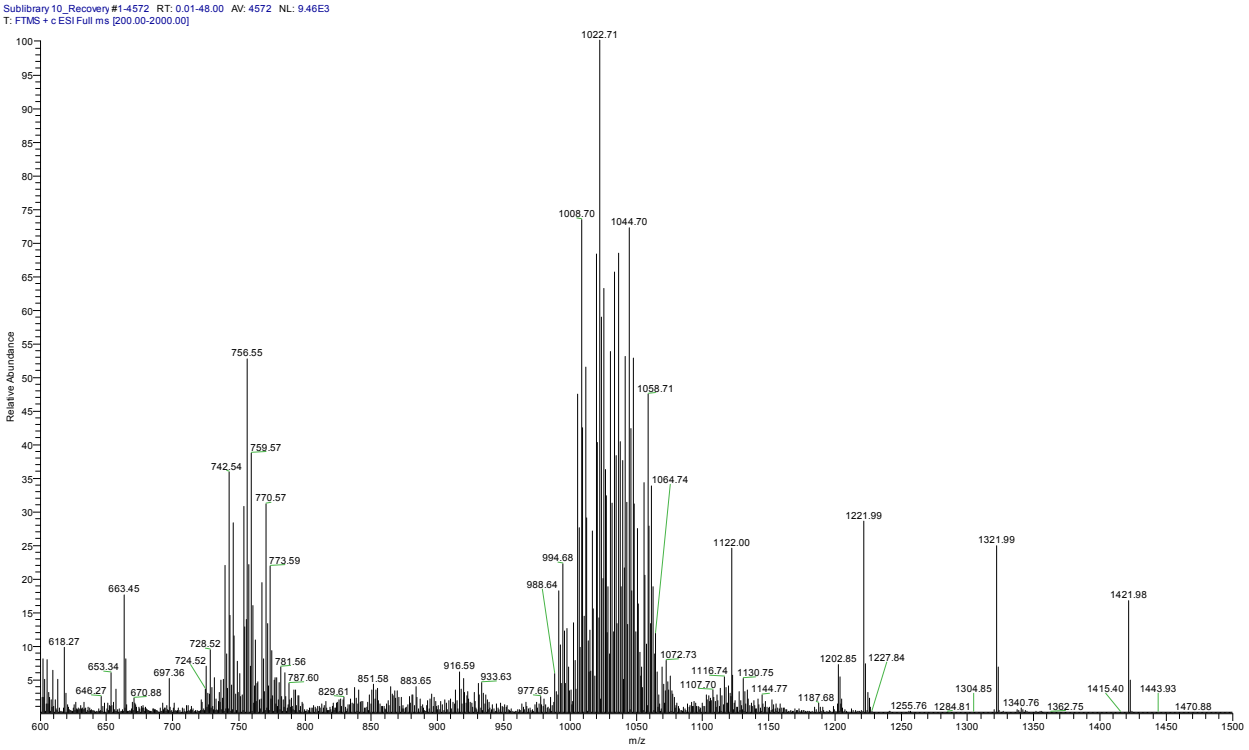


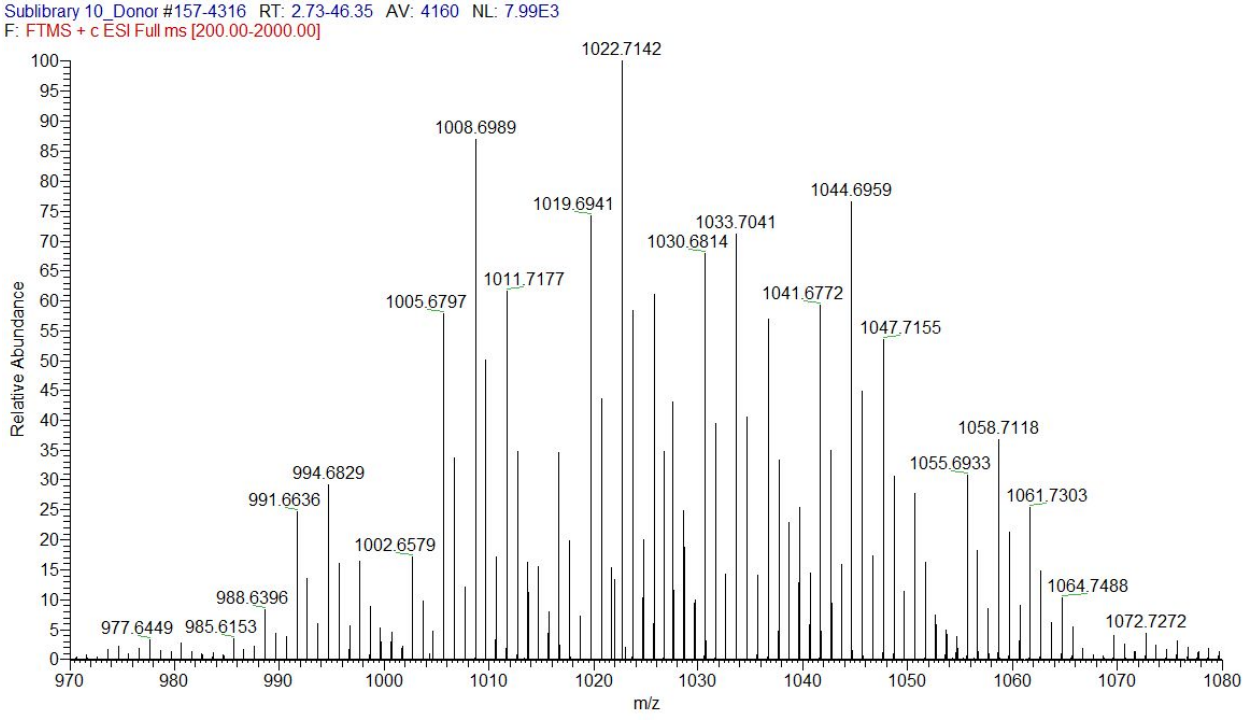
Mass spectrum of sub-library 9



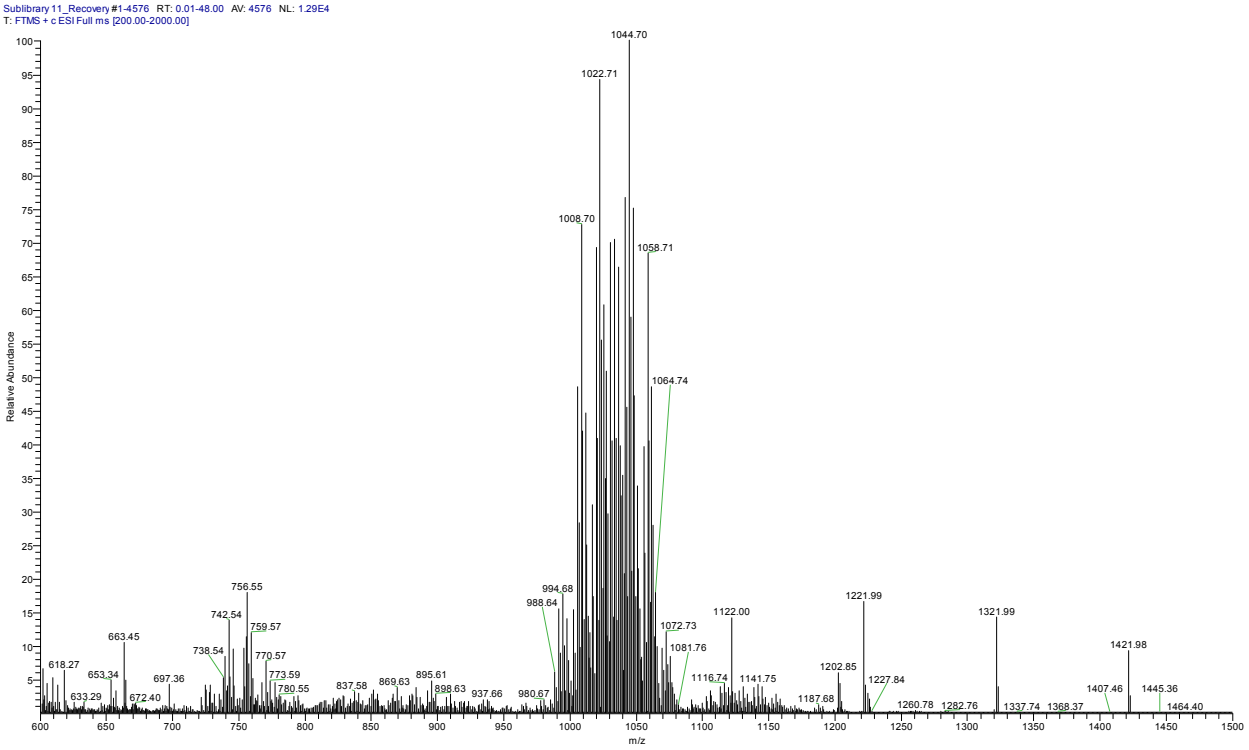


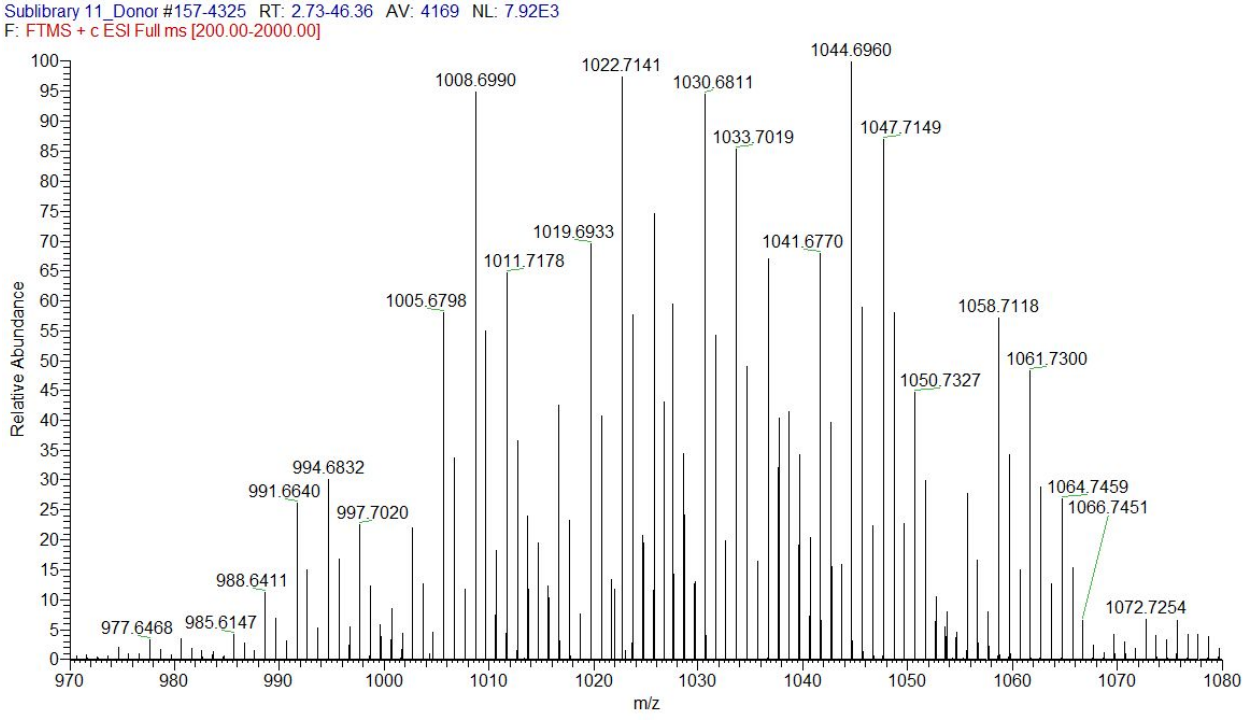
Mass spectrum of sub-library 10



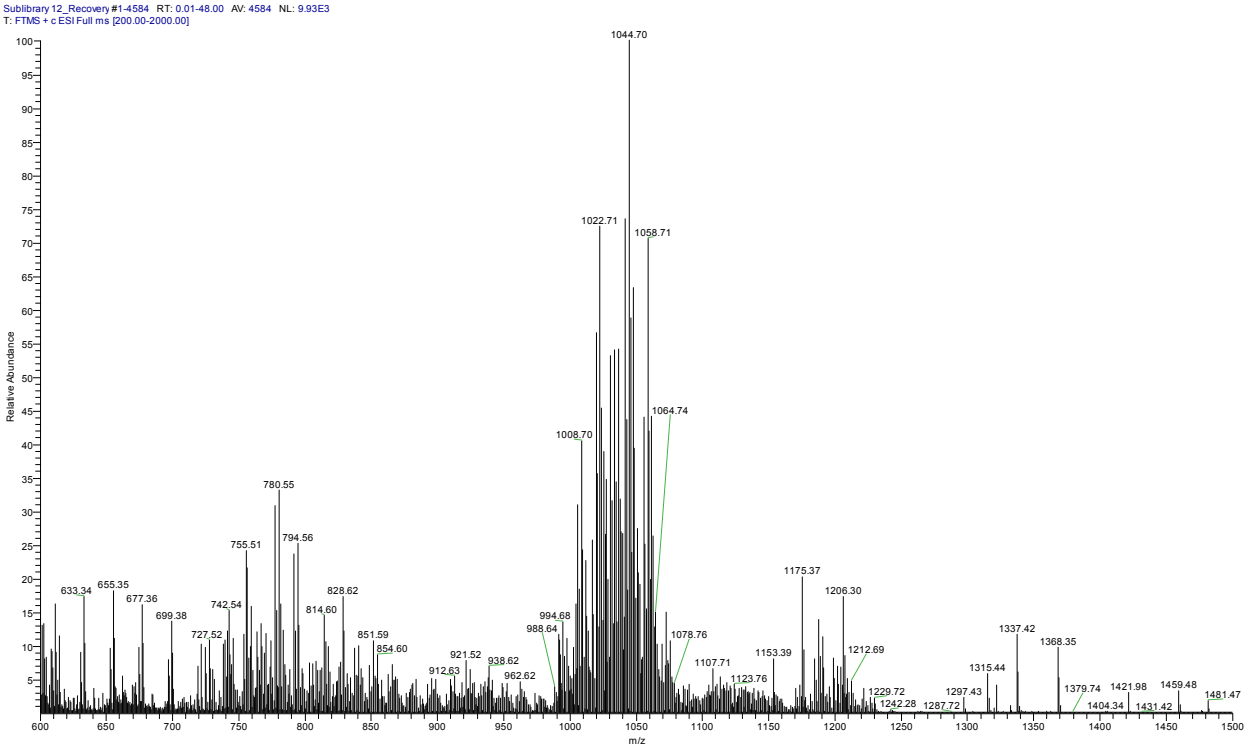


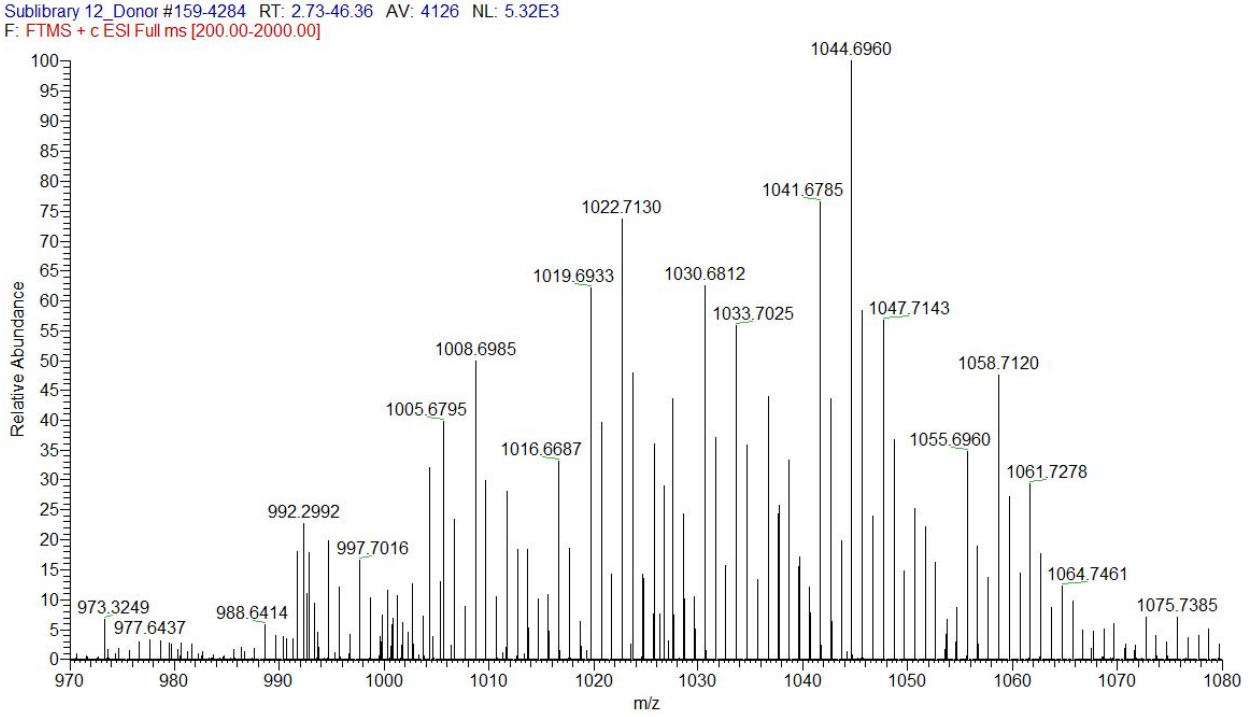
Mass spectrum of sub-library 11



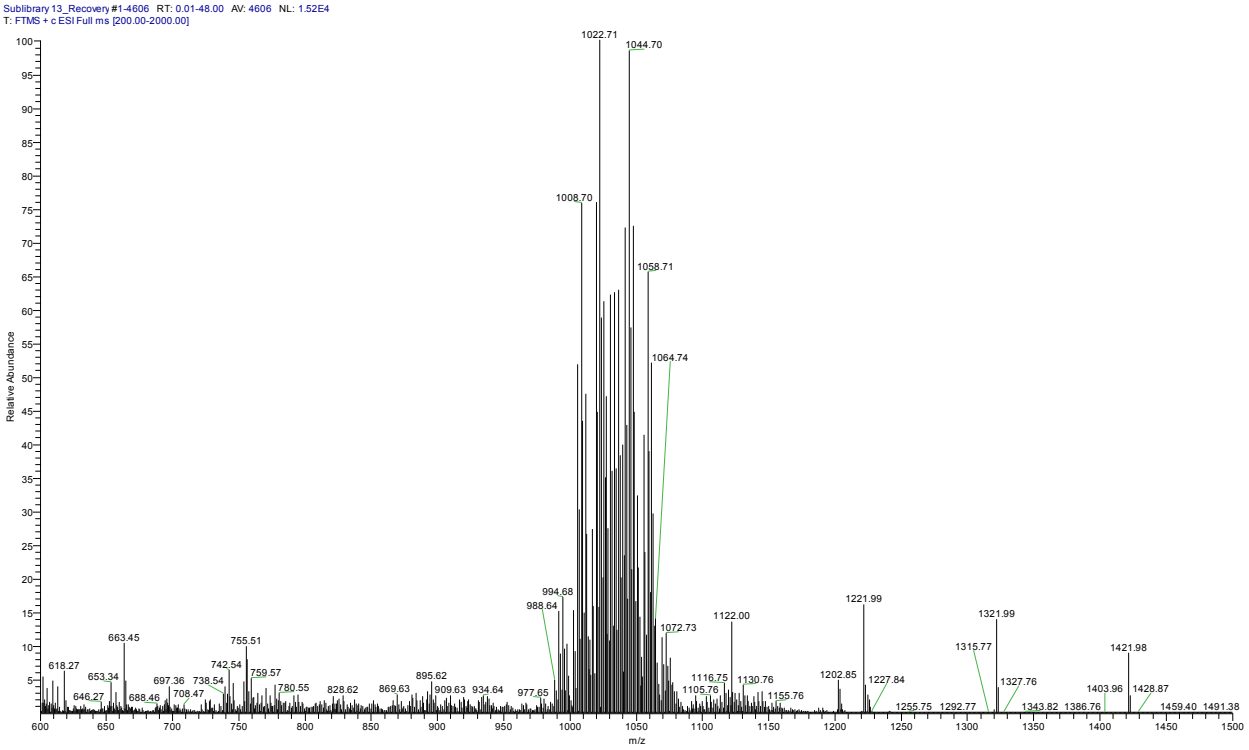


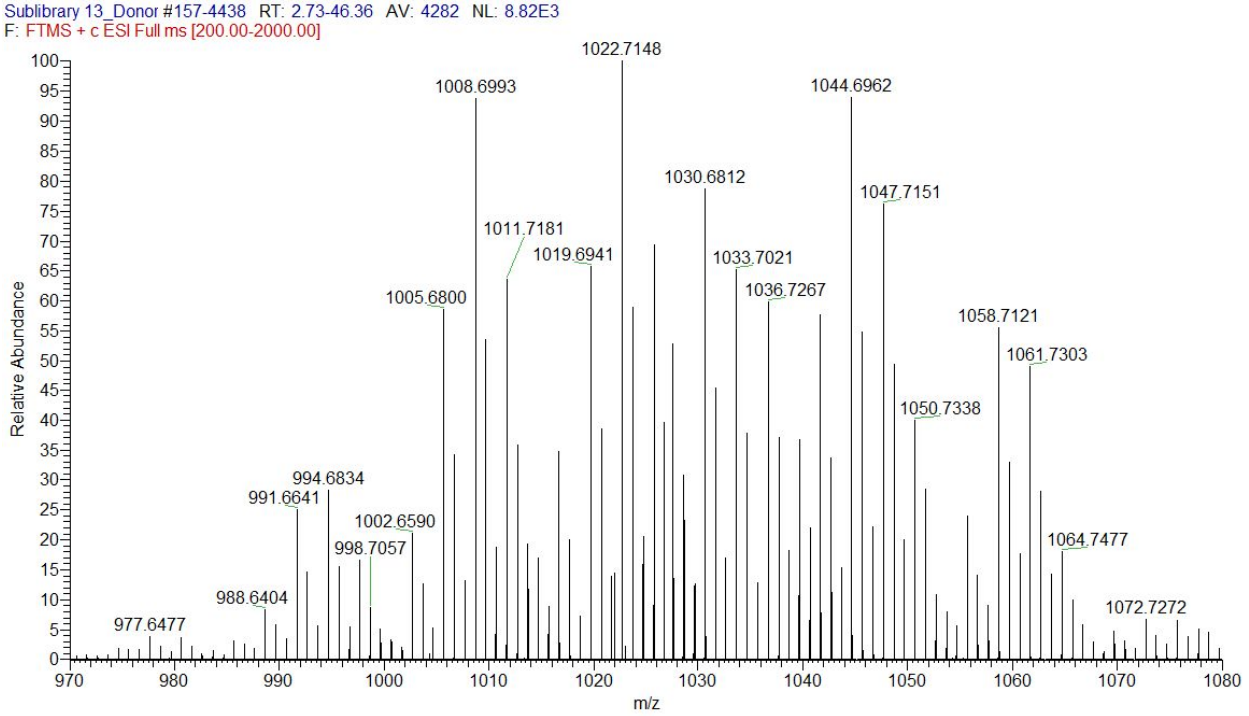
Mass spectrum of sub-library 12



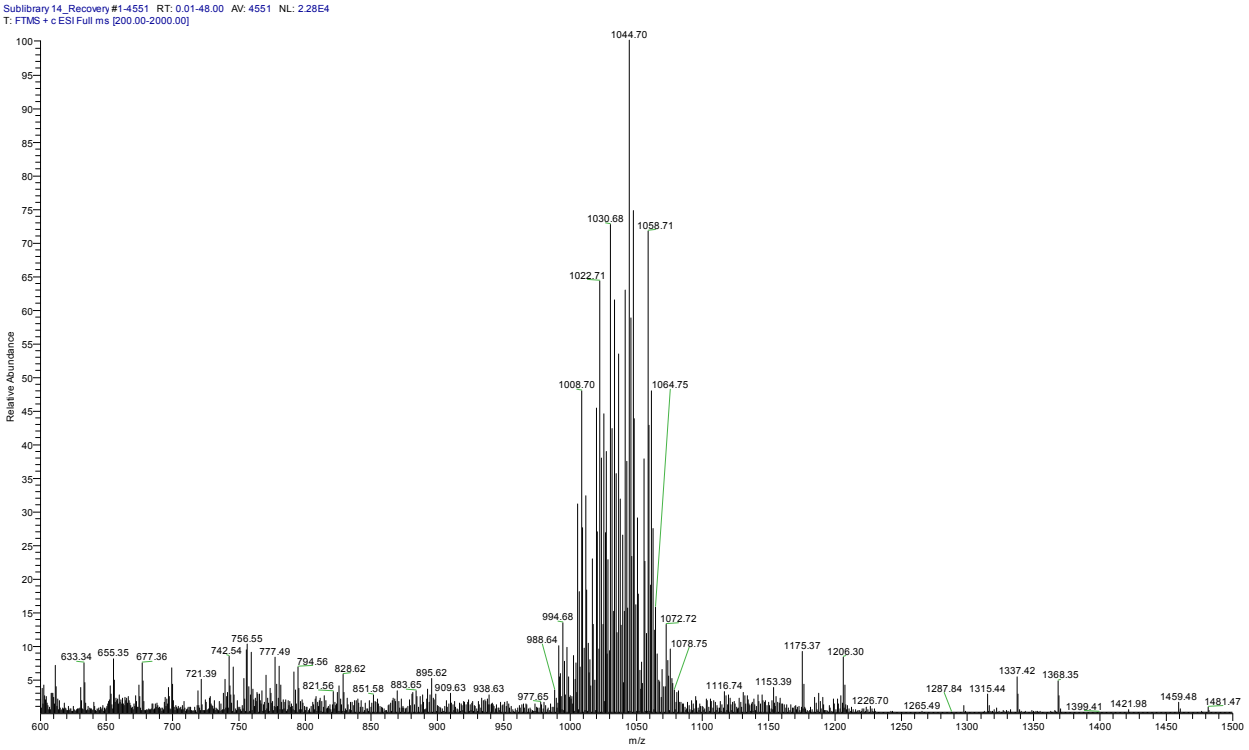


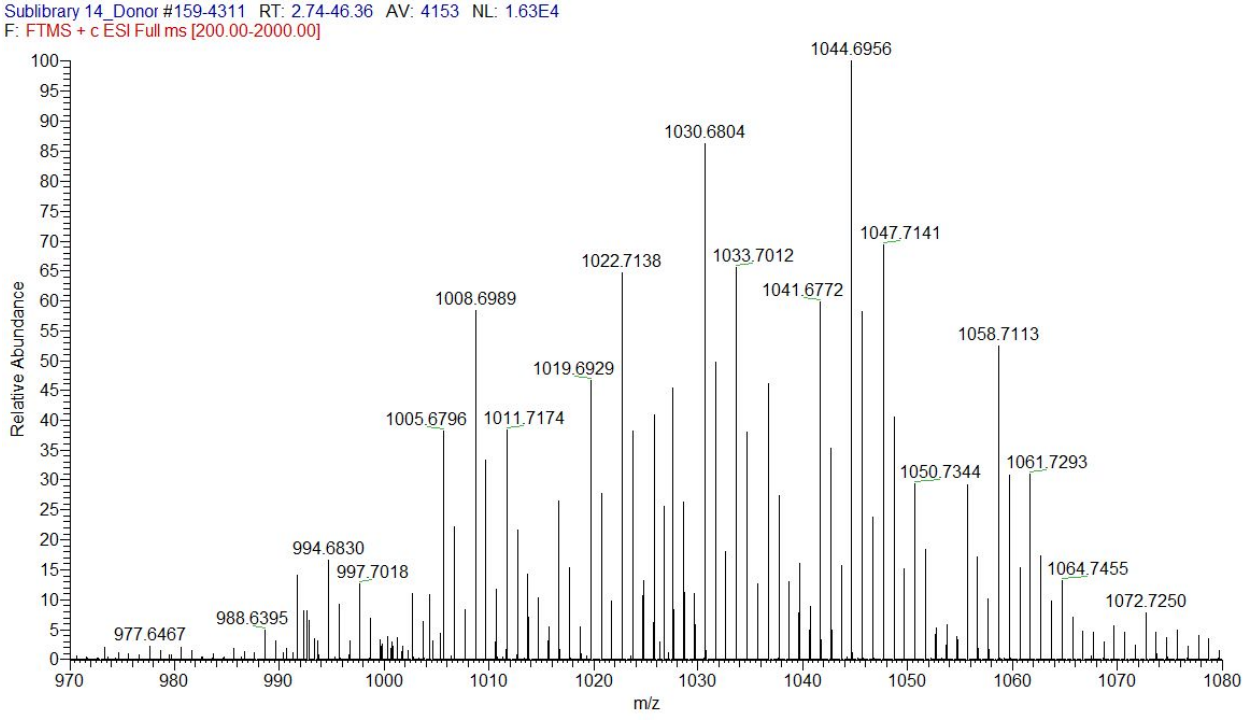
Mass spectrum of sub-library 13



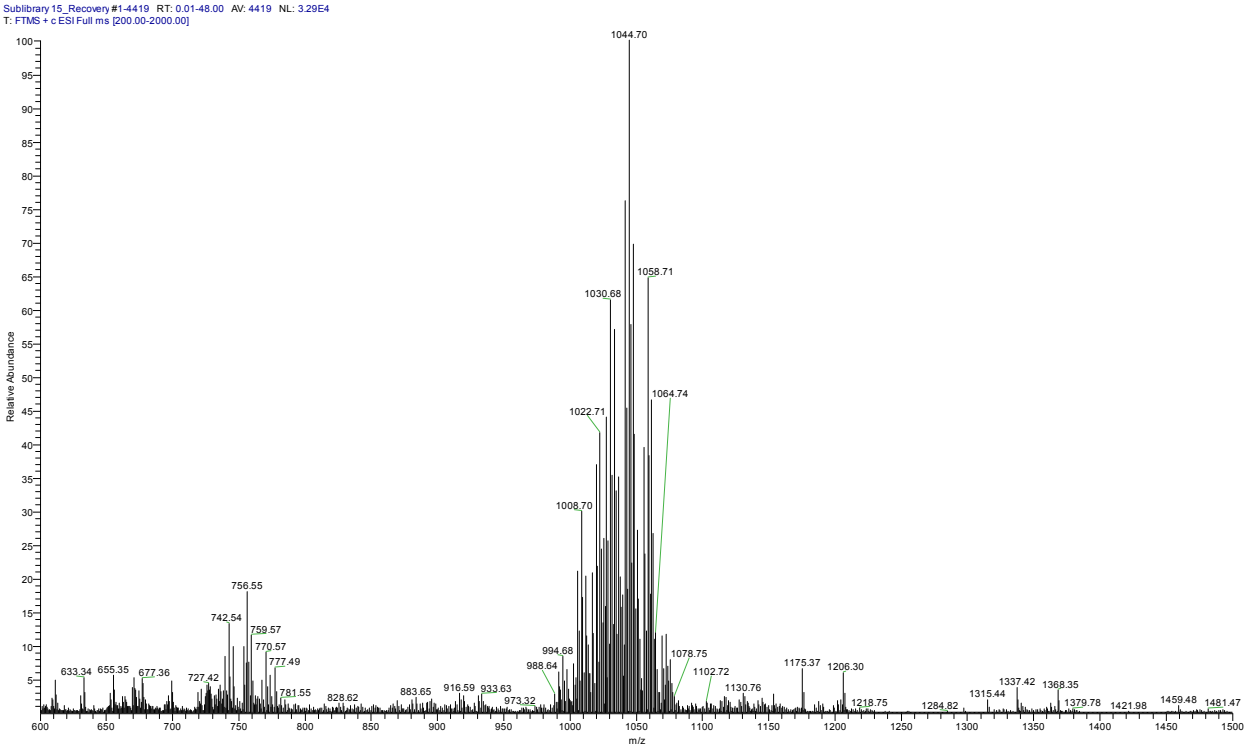


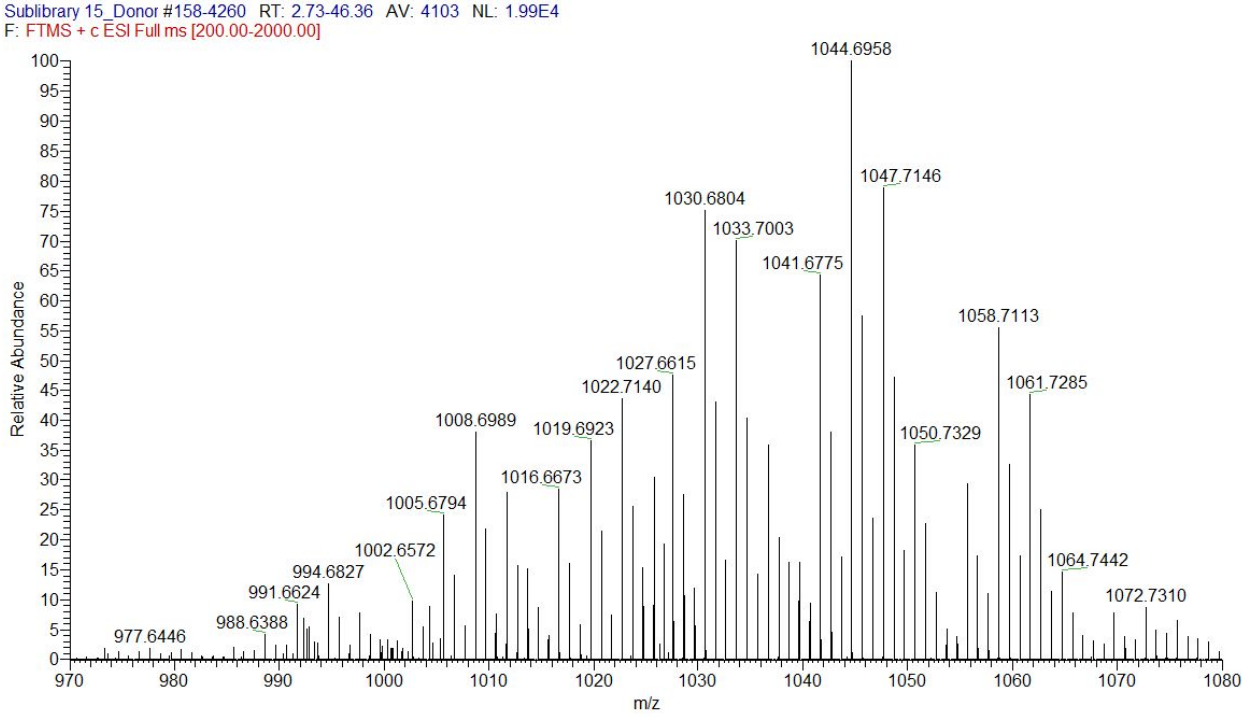
Mass spectrum of sub-library 14



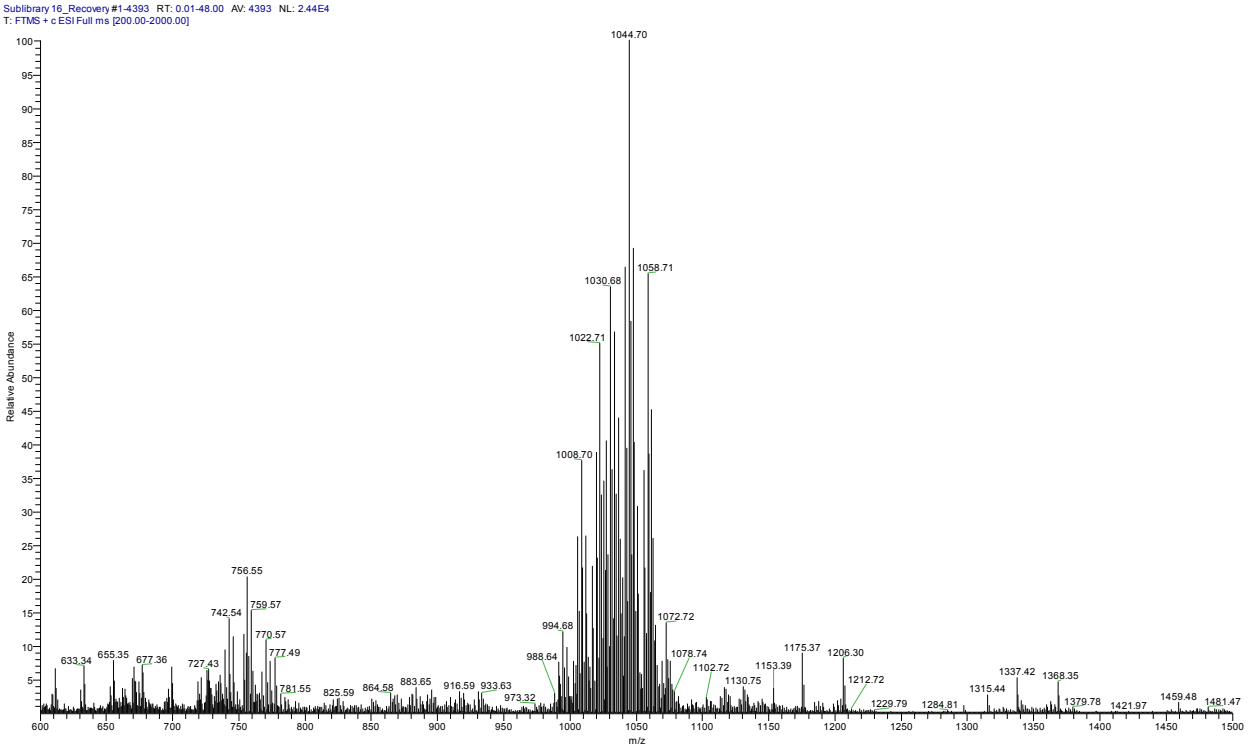


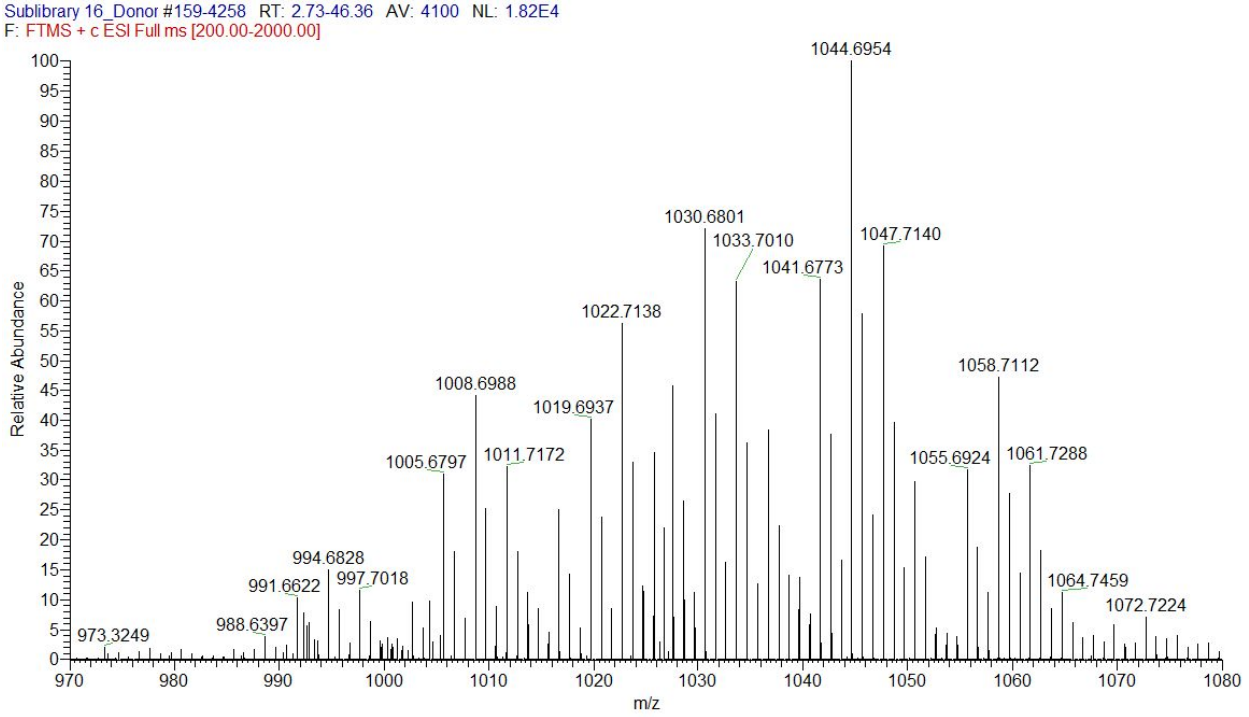
Mass spectrum of sub-library 15



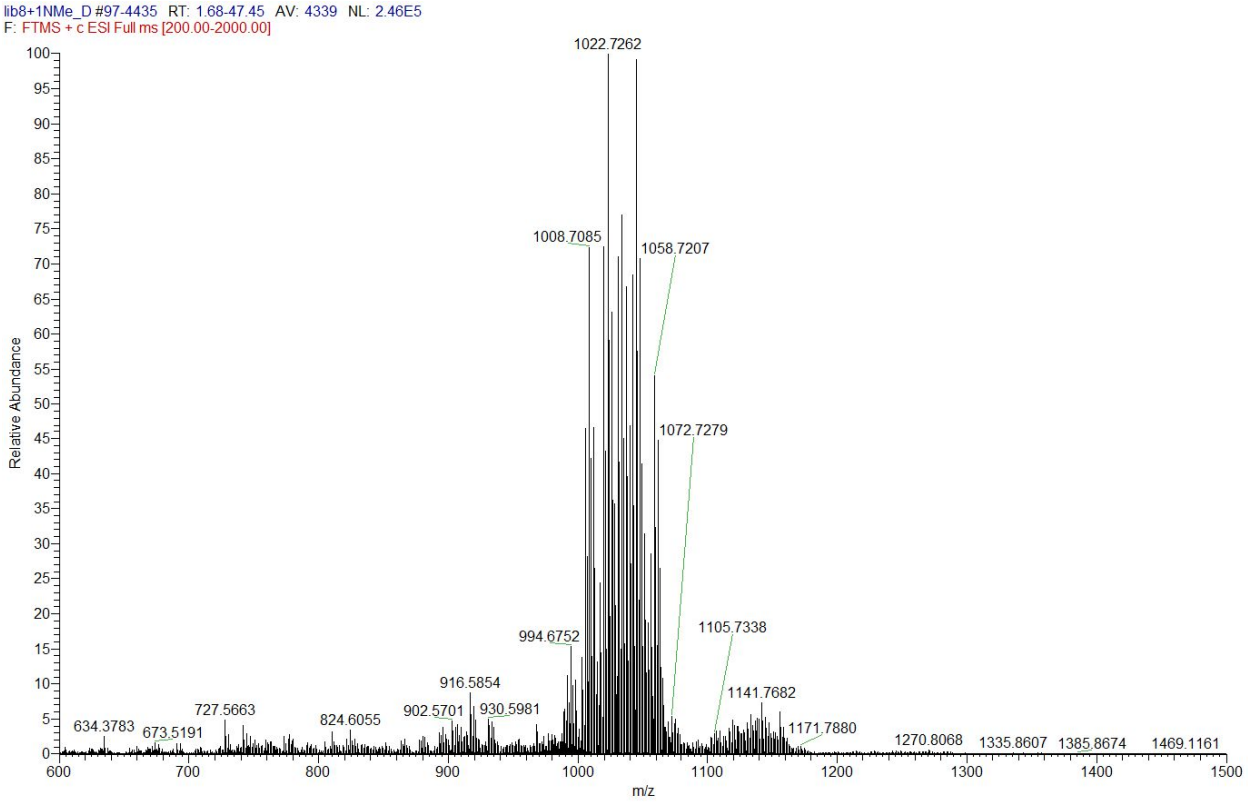


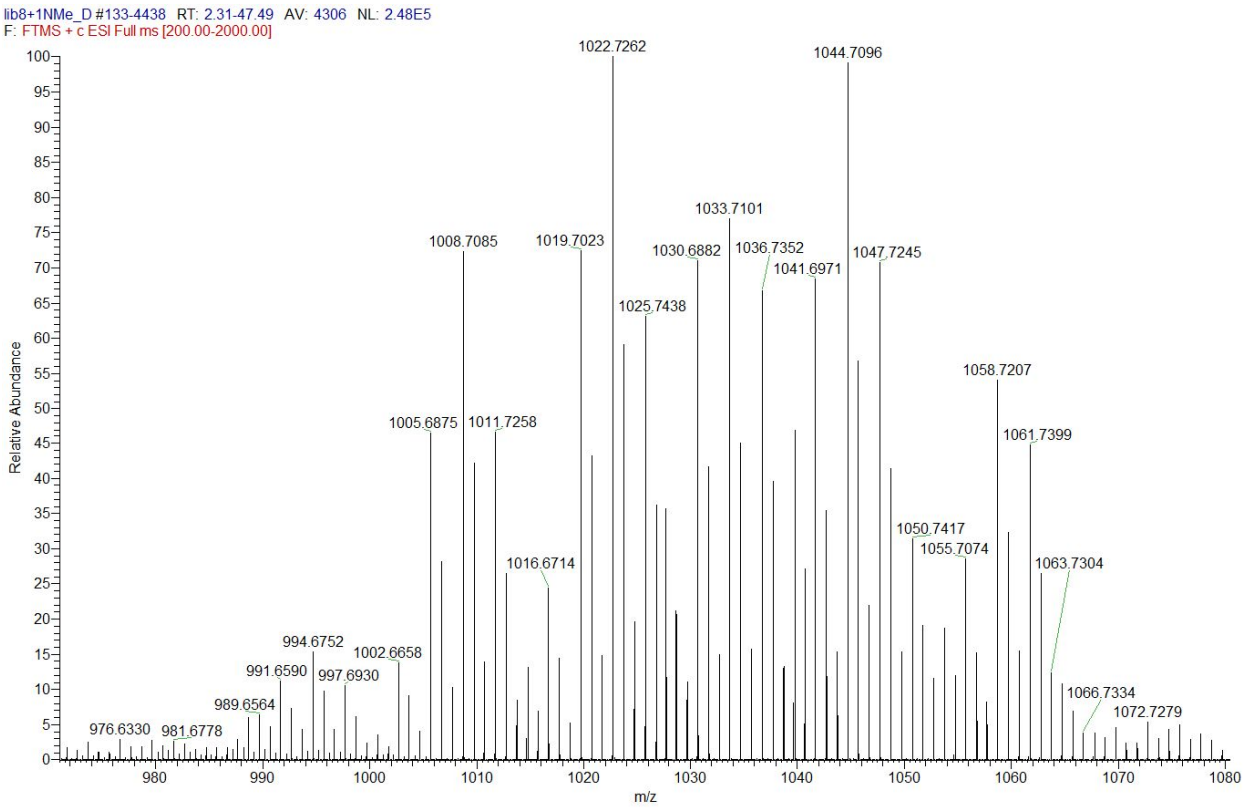
Mass spectrum of sub-library 16



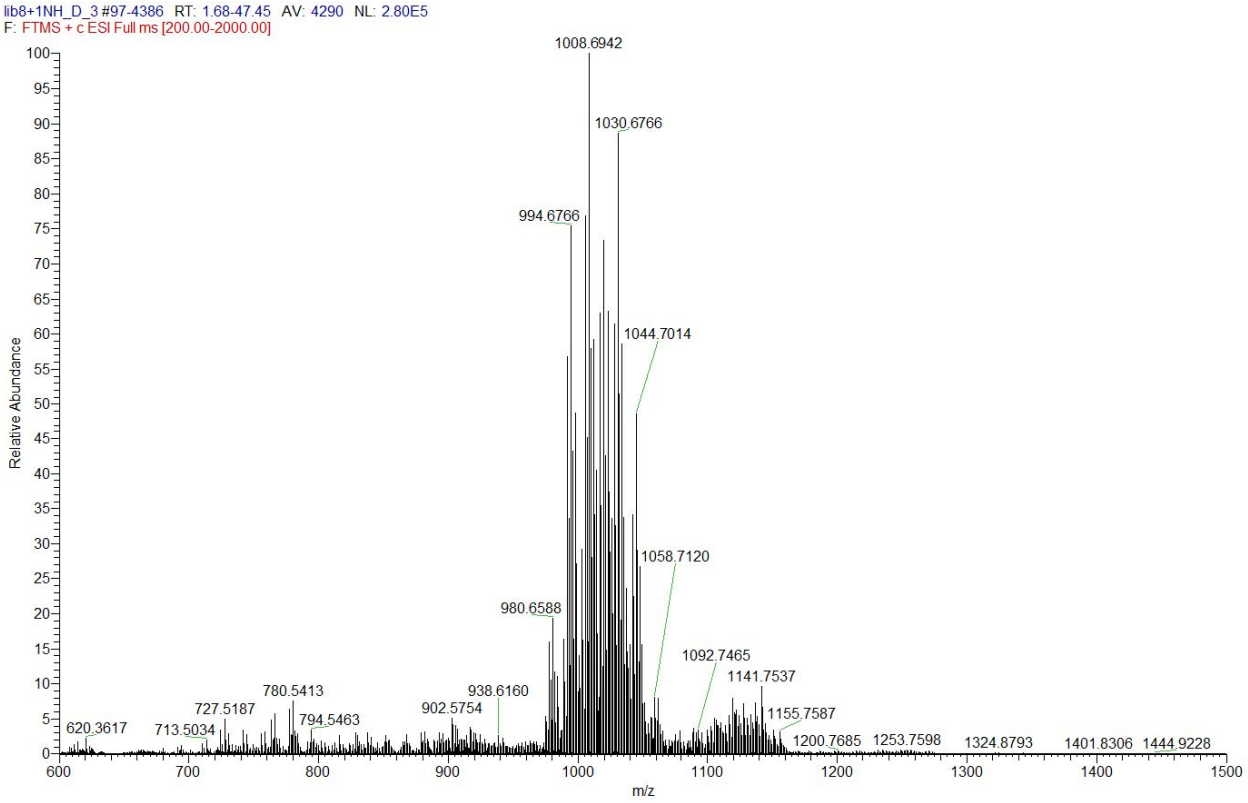


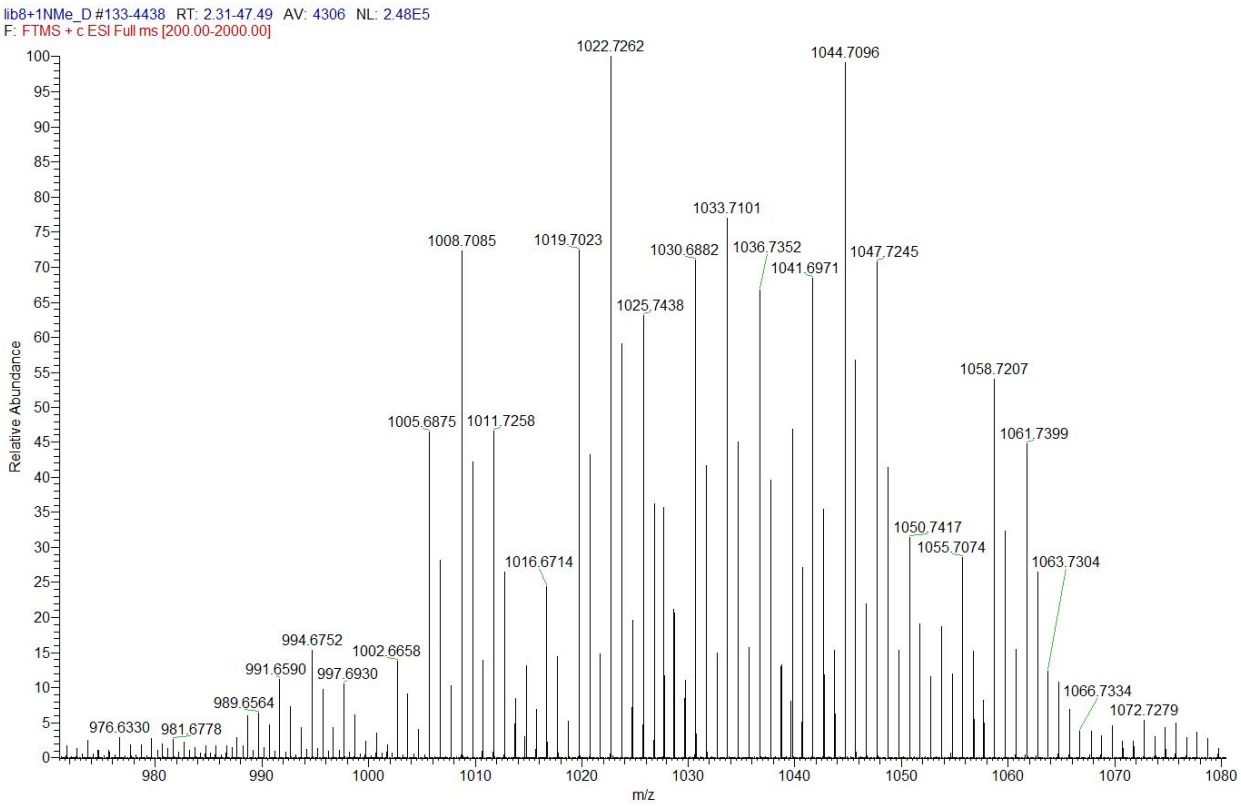
Mass spectrum of Library 2



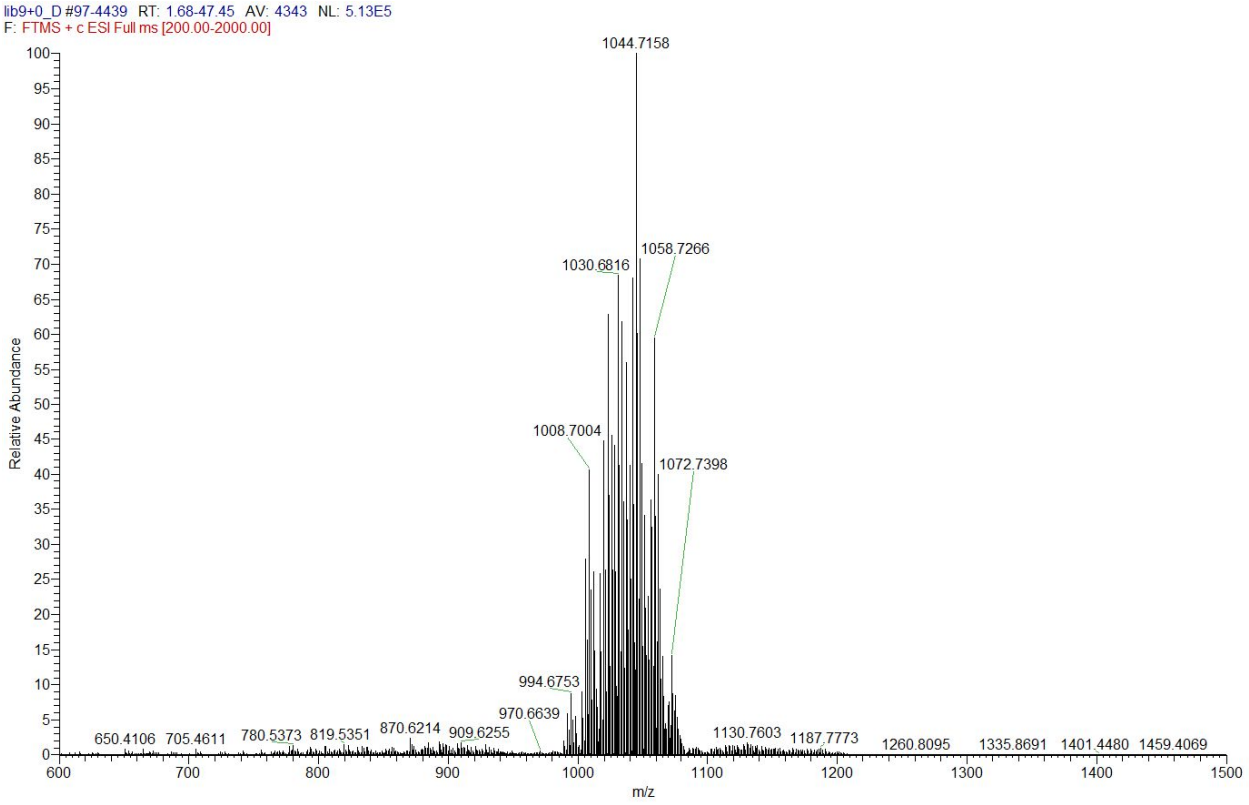

lib8+1NMe_D #133-4438 RT: 2.31-47.49 AV: 4306 NL: 2.48E5
F: FTMS + c ESI Full ms [200.00-2000.00]
Relative Abundance
m/z
1022.7262
1044.7096
1033.7101
1008.7085
1019.7023
1030.6882
1036.7352
1041.6971
1047.7245
1025.7438
1058.7207
1005.6875
1011.7258
1061.7399
1050.7417
1055.7074
1016.6714
1063.7304
994.6752
1002.6658
991.6590
997.6930
1066.7334
989.6564
1072.7279
976.6330
981.6778

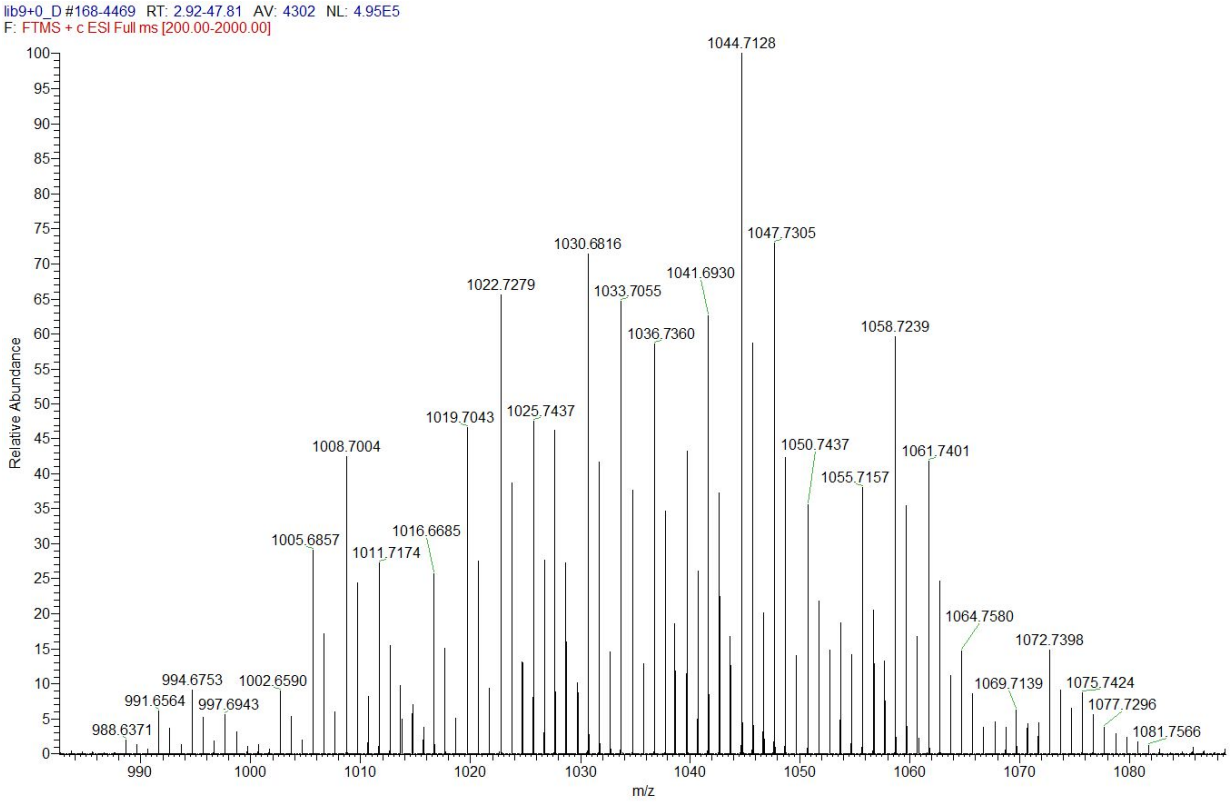
Mass spectrum of Library 3



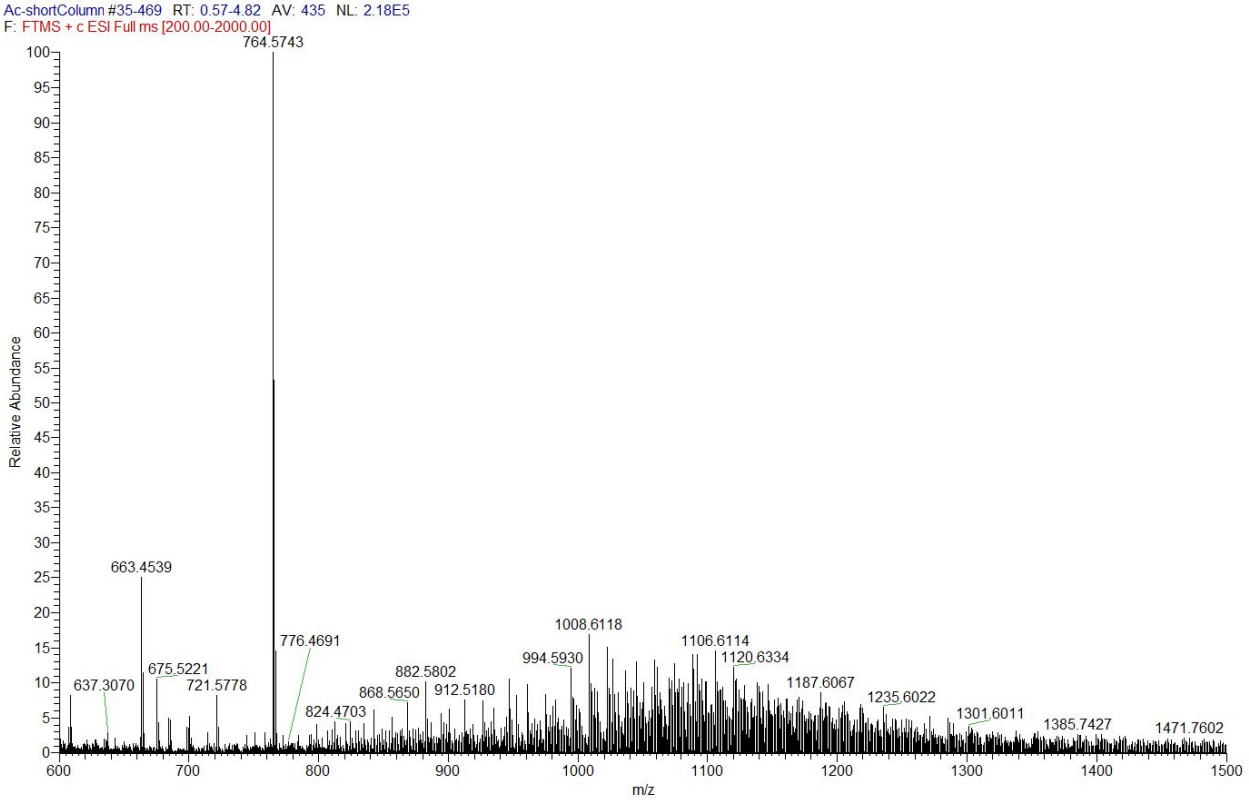

lib8+1NMe_D #133-4438 RT: 2.31-47.49 AV: 4306 NL: 2.48E5
F: FTMS + c ESI Full ms [200.00-2000.00]
Relative Abundance
m/z
1022.7262
1044.7096
1033.7101
1008.7085
1019.7023
1030.6882
1036.7352
1041.6971
1047.7245
1025.7438
1058.7207
1005.6875
1011.7258
1061.7399
1050.7417
1055.7074
1016.6714
1063.7304
994.6752
1002.6658
991.6590
997.6930
1066.7334
989.6564
1072.7279
976.6330
981.6778

Mass spectrum of Library 4


lib9+0_D #97-4439 RT: 1.68-47.45 AV: 4343 NL: 5.13E5
F: FTMS + c ESI Full ms [200.00-2000.00]
1044.7158
1058.7266
1030.6816
1008.7004
1072.7398
994.6753
970.6639
650.4106
705.4611
780.5373
819.5351
870.6214
909.6255
1130.7603
1187.7773
1260.8095
1335.8691
1401.4480
1459.4069
Relative Abundance
m/z


lib9+0_D #168-4469 RT: 2.92-47.81 AV: 4302 NL: 4.95E5
F: FTMS + c ESI Full ms [200.00-2000.00]
1044.7128
1047.7305
1030.6816
1041.6930
1022.7279
1033.7055
1036.7360
1058.7239
1019.7043
1025.7437
1008.7004
1050.7437
1061.7401
1055.7157
1005.6857
1016.6685
1011.7174
1064.7580
1072.7398
994.6753
1002.6590
1069.7139
1075.7424
991.6564
997.6943
1077.7296
988.6371
1081.7566
Relative Abundance
m/z

Mass spectrum of Library 5



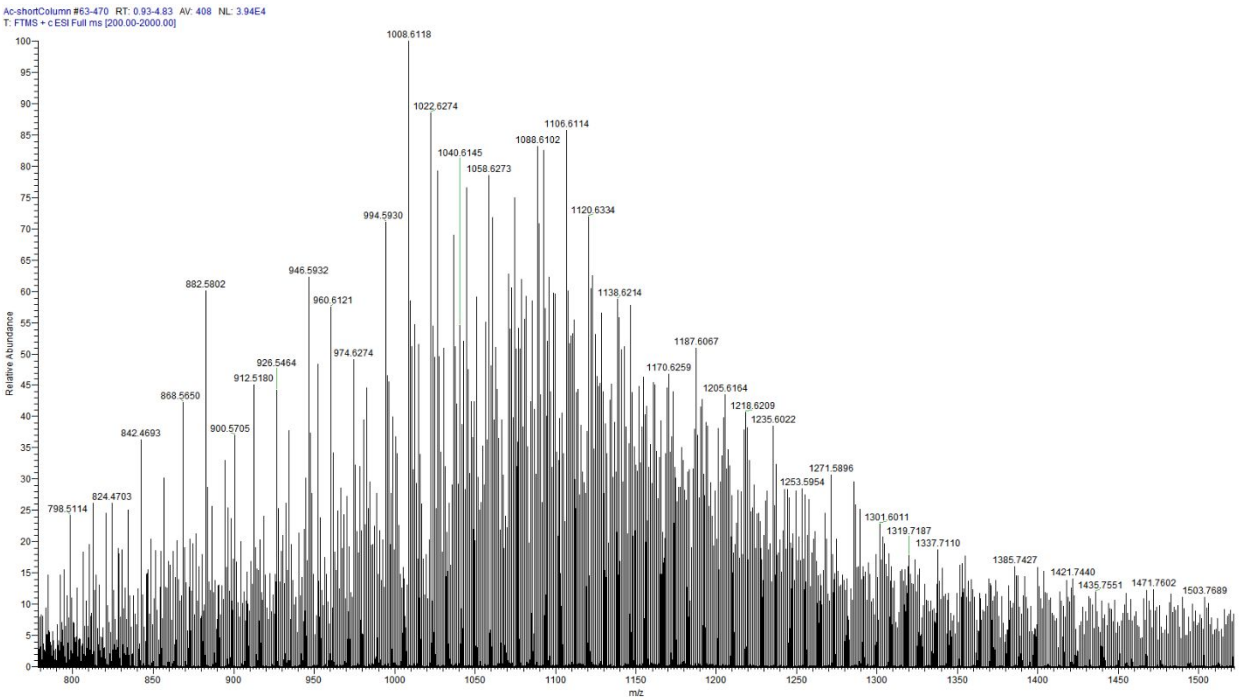


***Table S25*. Data on individual compounds**

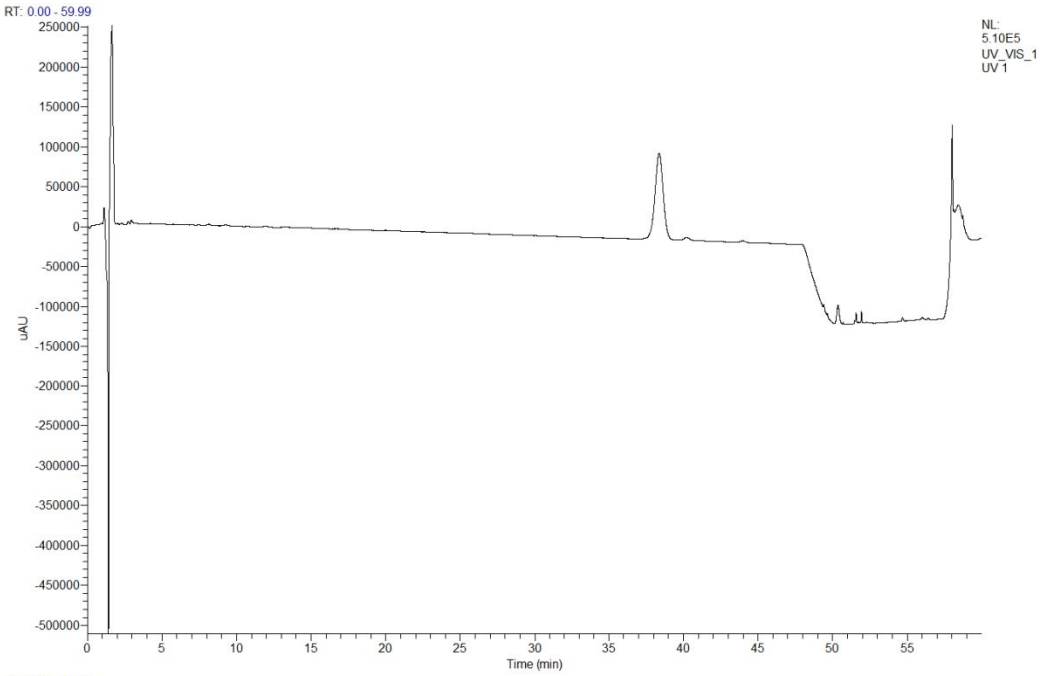
Number	MW	AlogP	Rt (library)	Rt (resynthesized)	P_{app} (library)	P_{app} (resynthesized)	P_{app} (MDCK)	Purity (UV 200nm)
1	1029.68	3.62	38.62	38.51	9.81	2.26	7.6	95%
2	1002.31	2.94	32.45	32.11	8.77	2.29	13.8	95%
3	1029.68	3.62	42.37	42.23	8.72	5.02	26.6	96%
4	1016.34	3.28	37.17	36.94	7.92	2.28	10.8	96% (94%*)
5	1029.68	3.62	29.52	29.53	4.67	2.19	4.8	91%
6	1002.31	2.94	19.92	19.8	1.18	0.39	0.9	99%
7	1016.34	3.28	13.62	13.83	0.11	0.05	0.5	99%
8	988.28	2.59	12.51	12.57	0.06	0.06	0.5	87%
9	988.28	2.59	10.57	10.58	0.01	0.02	0.4	97%
10	1016.34	3.28	**17.33**	**22.38**	NA	0.06	NA	85%*
11	988.28	2.59	**20.72**	**13.37**	NA	0.03	NA	93%*

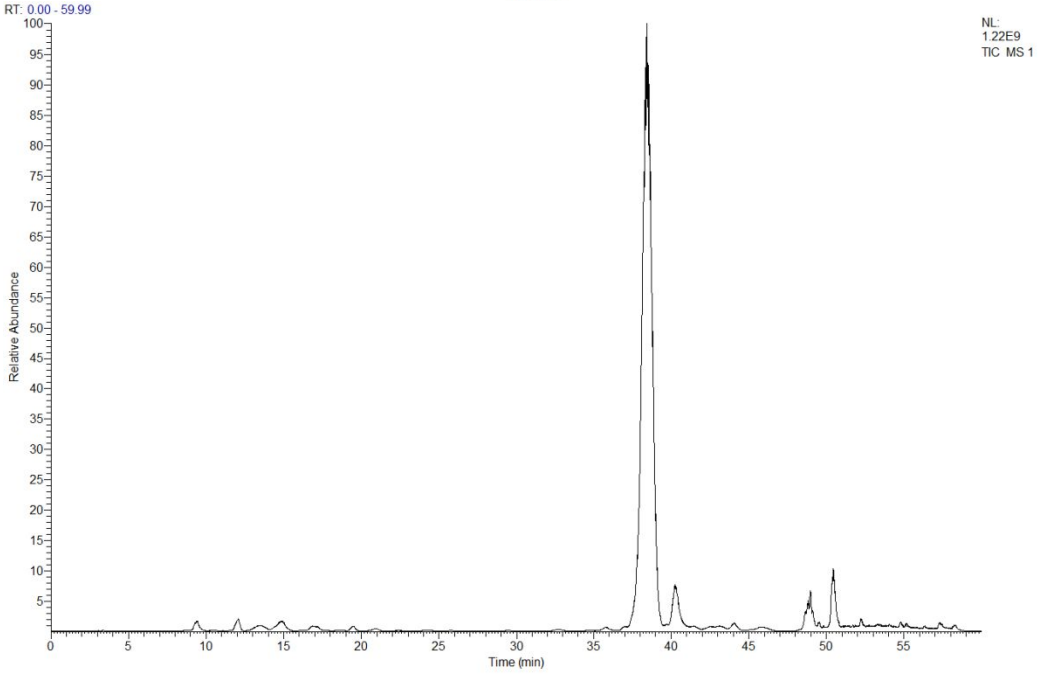
P_{app} values are in units of 10^{-6} cm/s. Retention times differing between library and resynthesized are indicated in red.

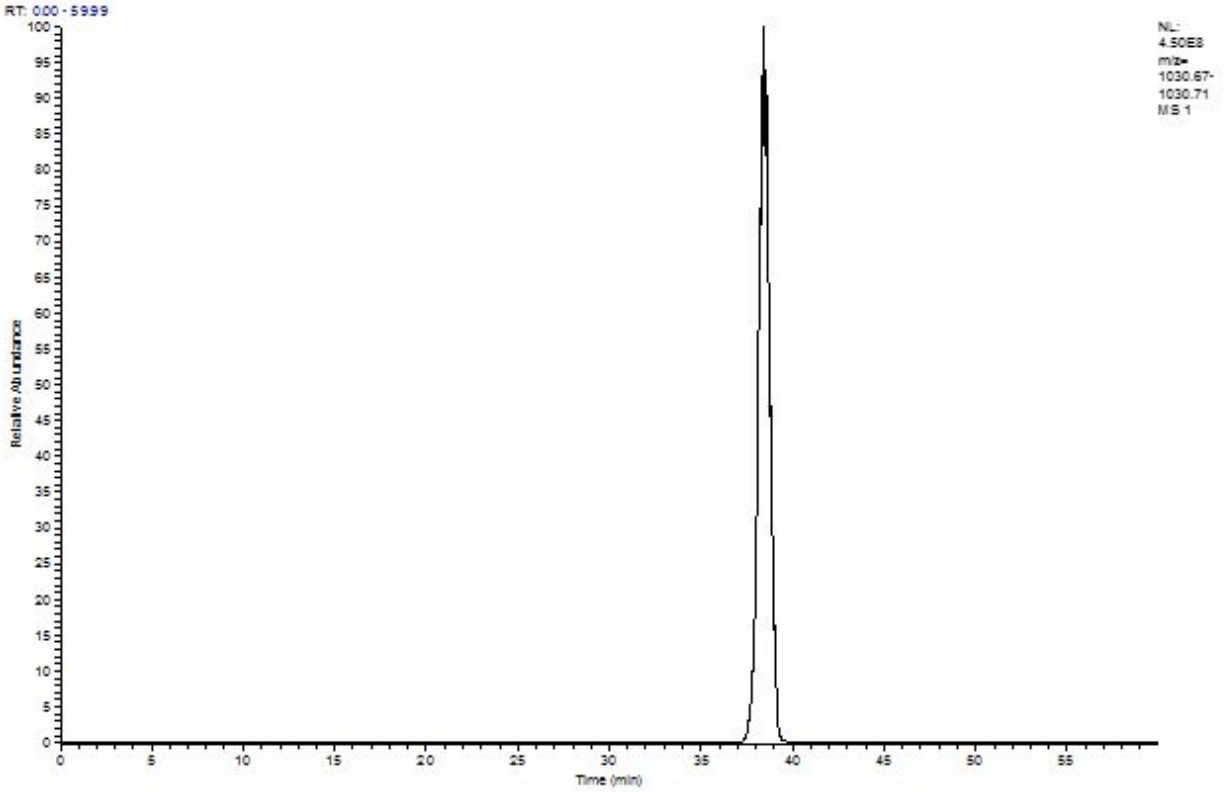
*Purities based on the total ion chromatogram. For compounds **10** and **11**, UV data was not collected.

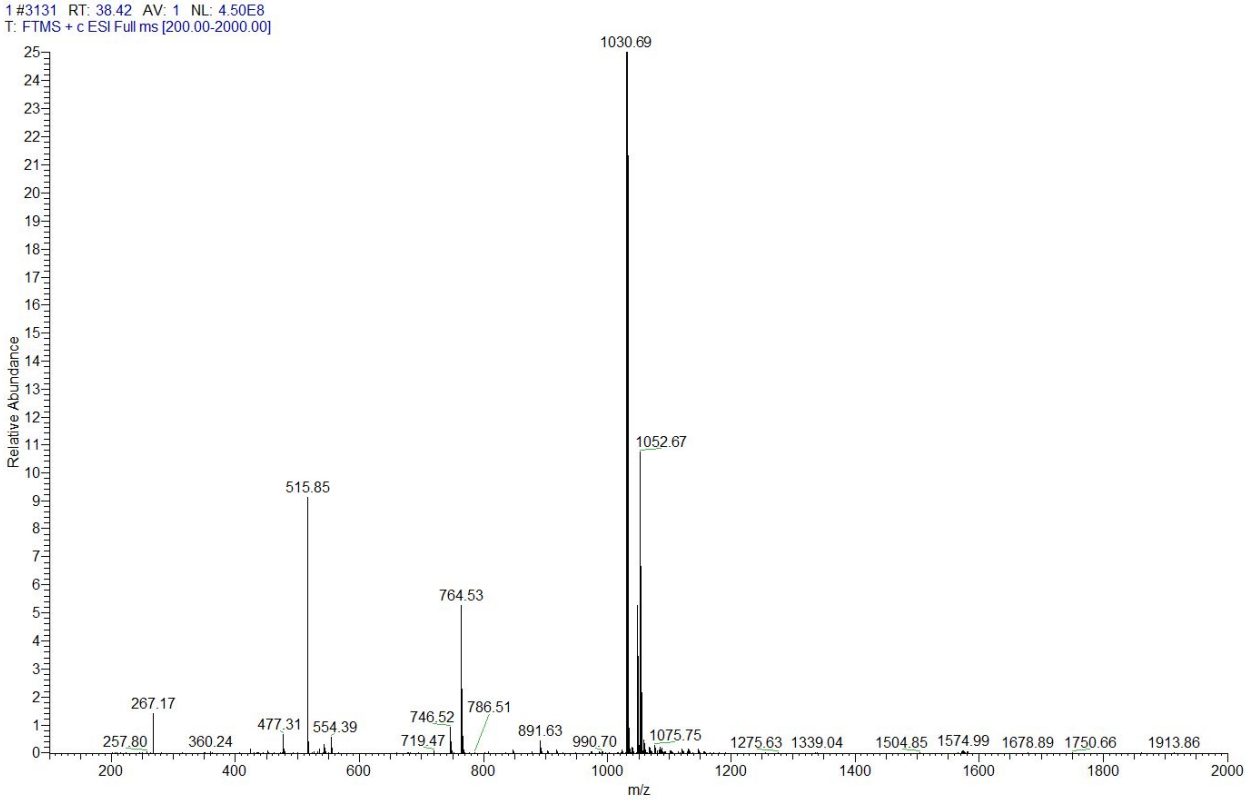
Analytical data of individual compounds (UV, TIC, selected ion, MS, H1 NMR)

LCMS data for **1**

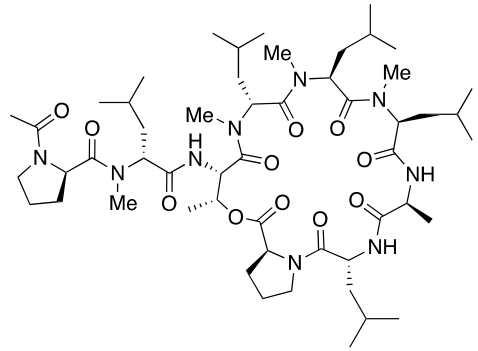


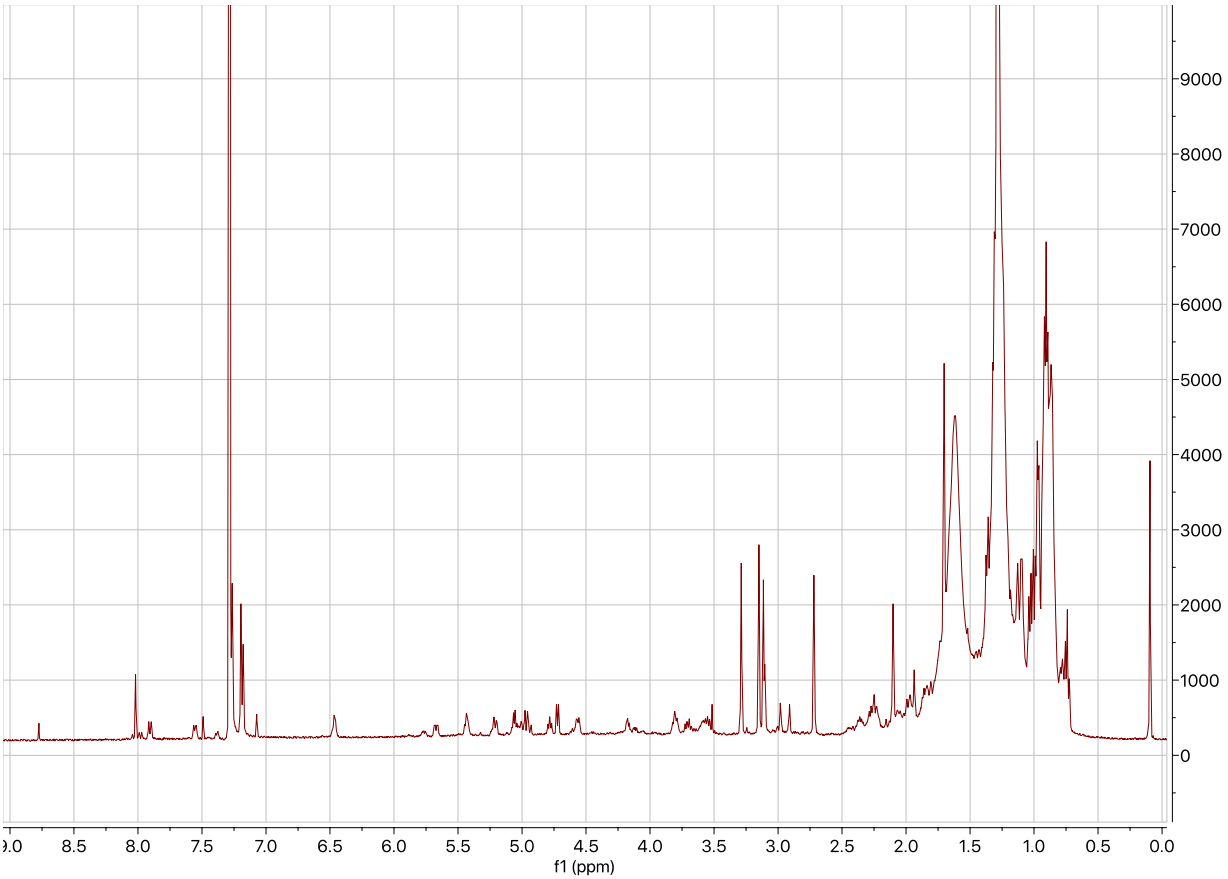



RT: 0.00 - 59.99
NL:
4.50E8
m/z=
1030.67-
1030.71
MS 1
Relative Abundance
Time (min)

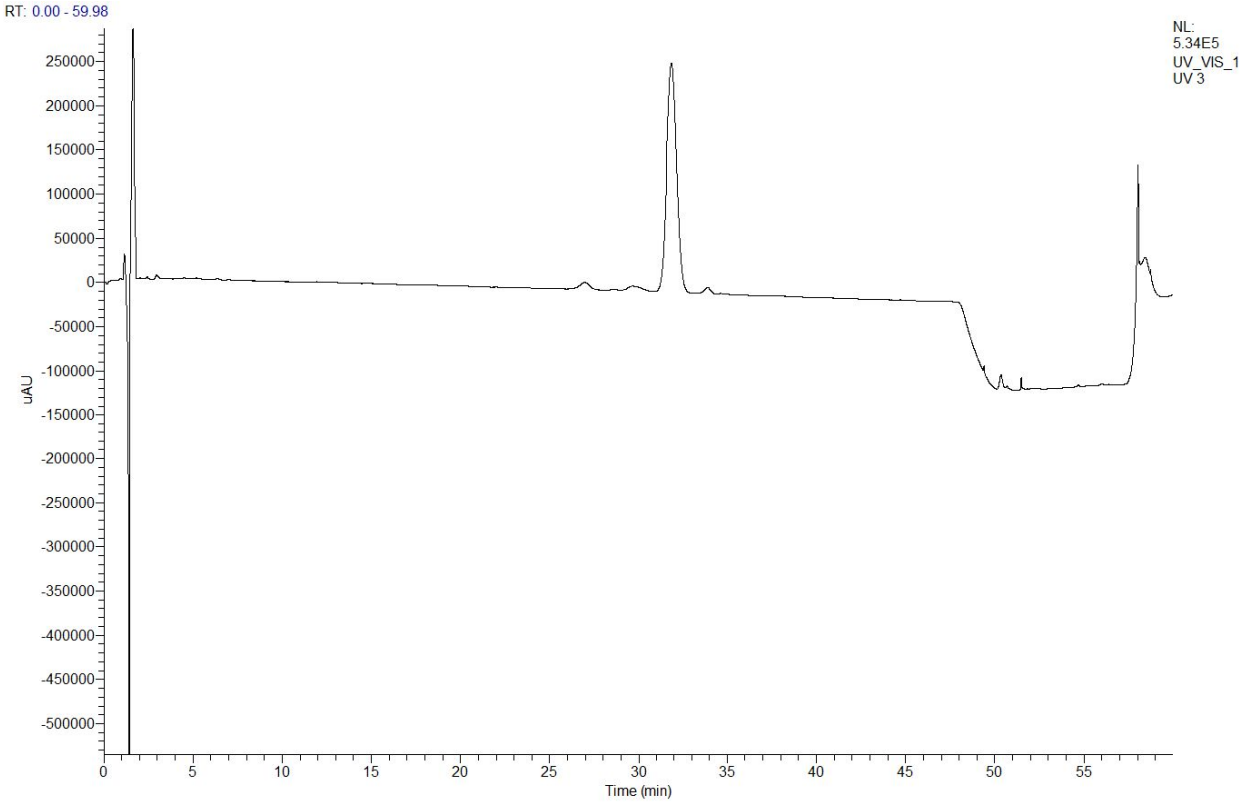

1 #3131 RT: 38.42 AV: 1 NL: 4.50E8
T: FTMS + c ESI Full ms [200.00-2000.00]
1030.69
1052.67
515.85
764.53
267.17
257.80
360.24
477.31
554.39
719.47
746.52
786.51
891.63
990.70
1075.75
1275.63
1339.04
1504.85
1574.99
1678.89
1750.66
1913.86
Relative Abundance
m/z

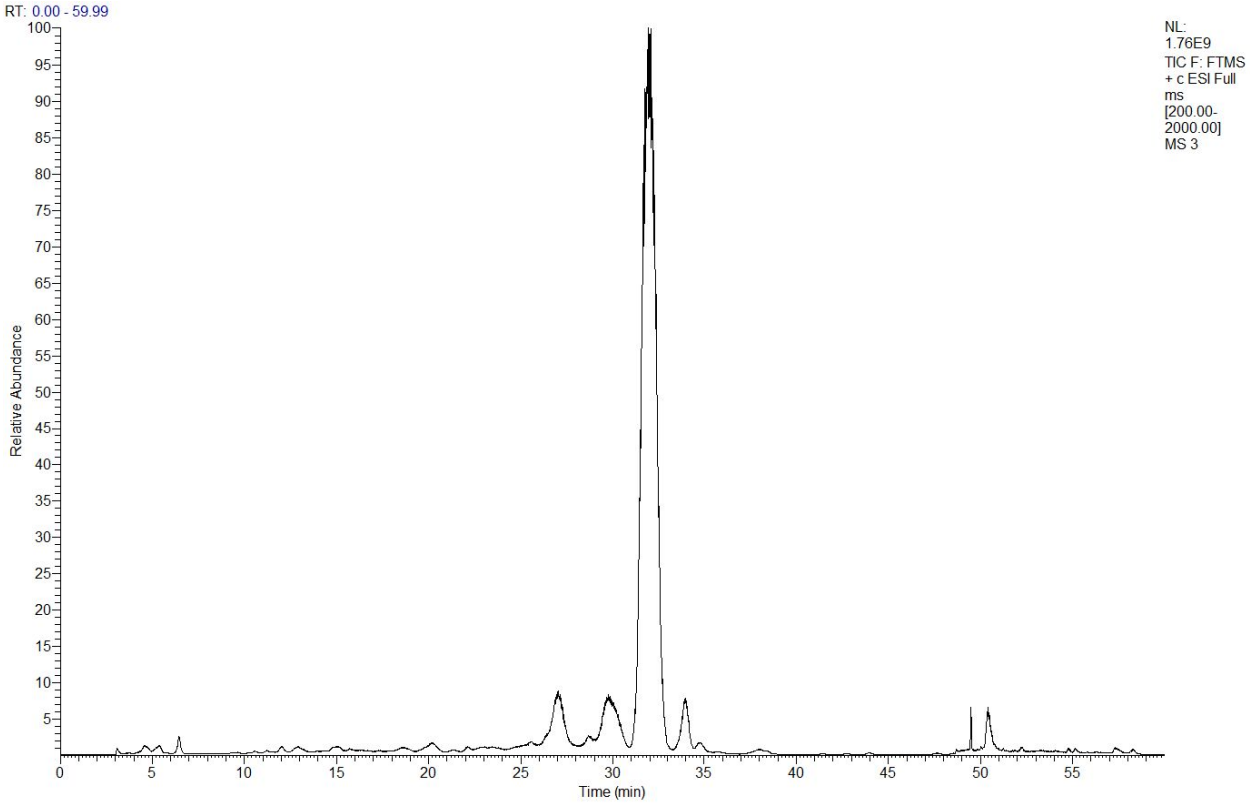
H1 NMR spectrum of **1**

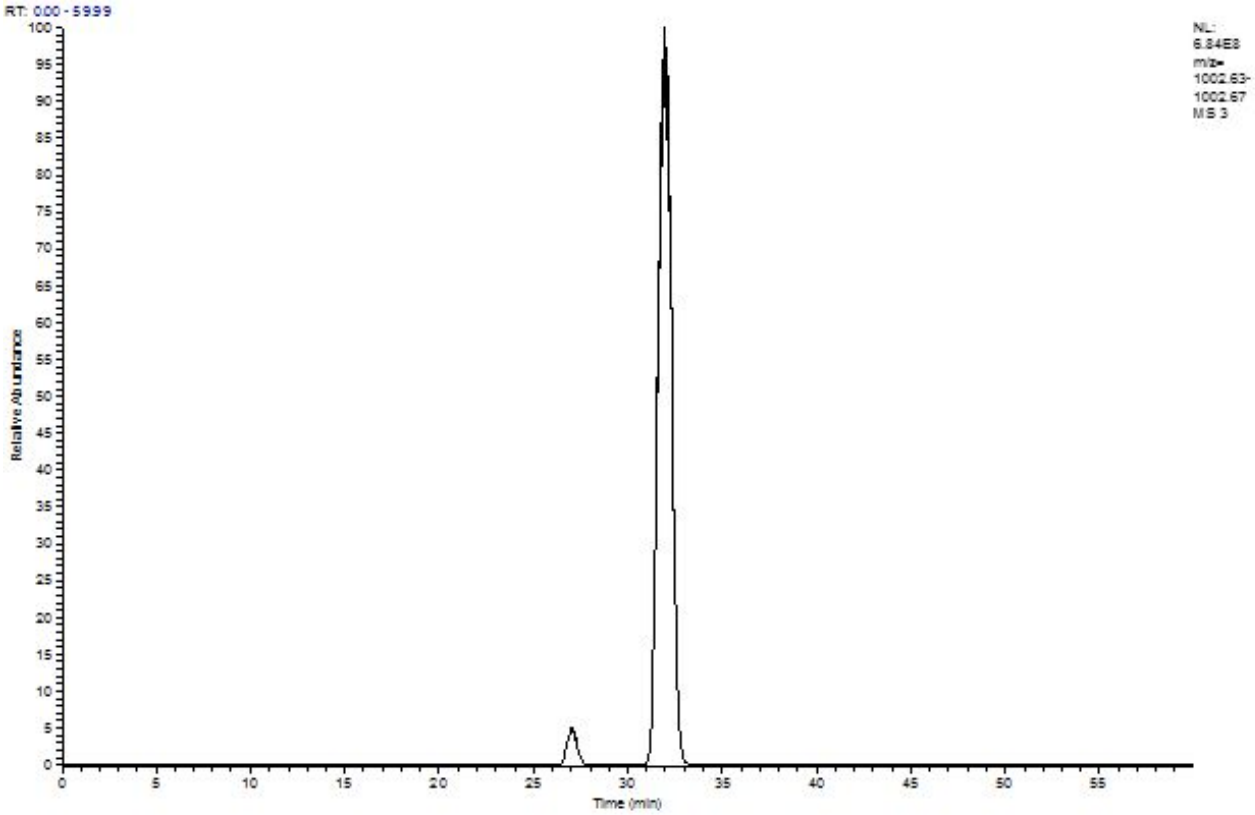


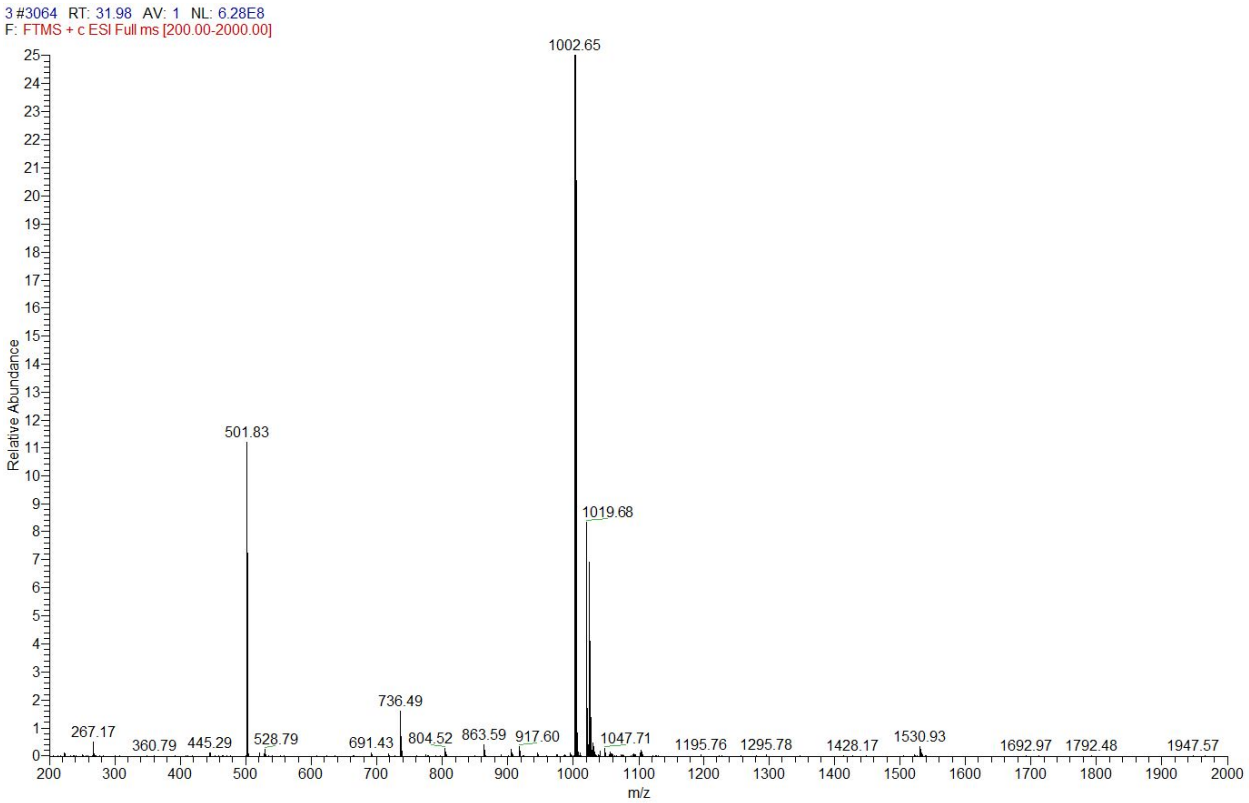


LCMS data for **2**

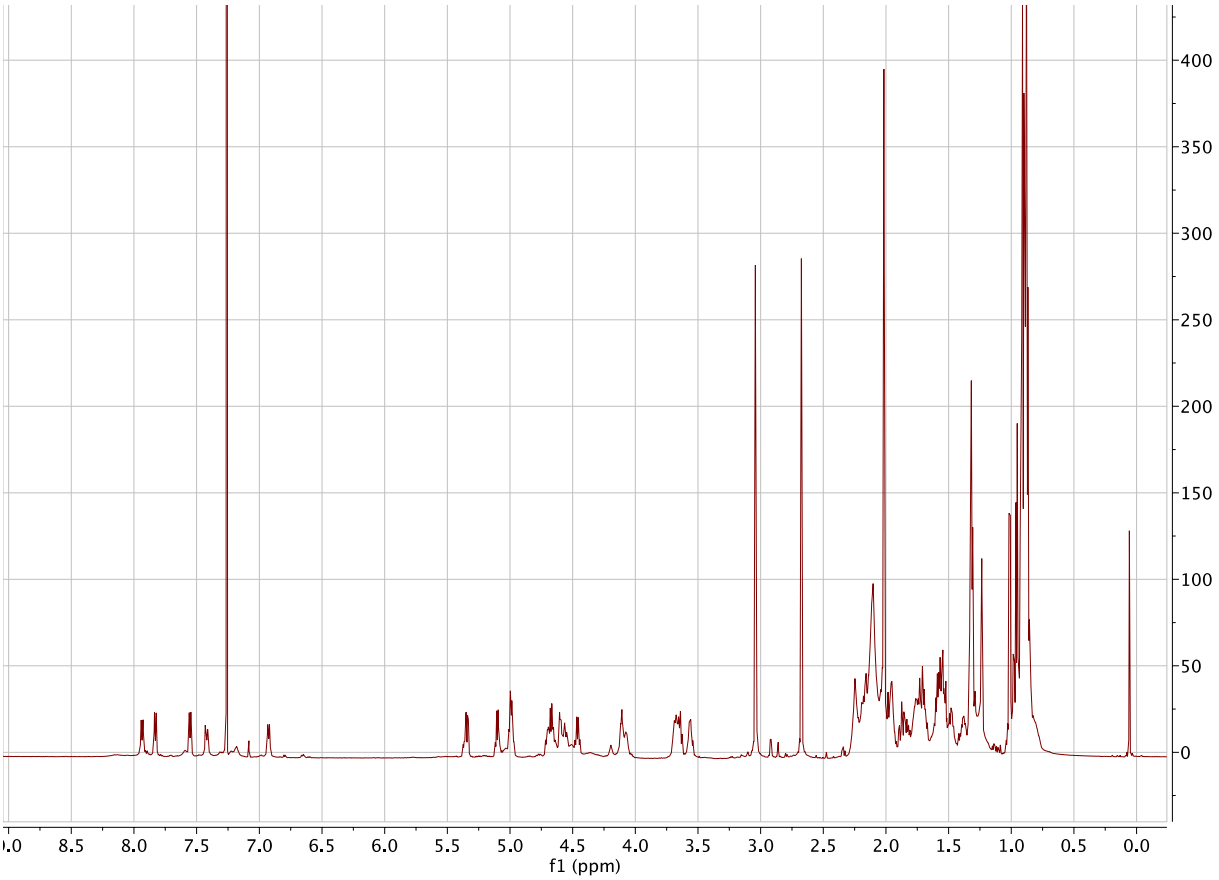


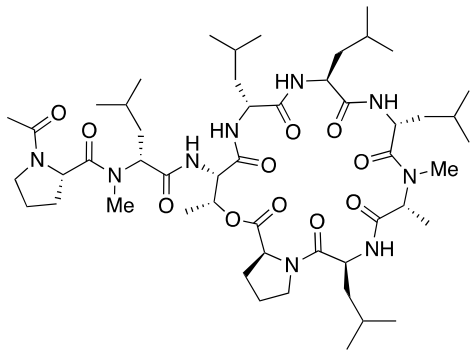



RT: 0.00 - 59.99
NL:
6.84E8
m/z=
1002.63-
1002.67
MS 3
Relative Abundance
Time (min)

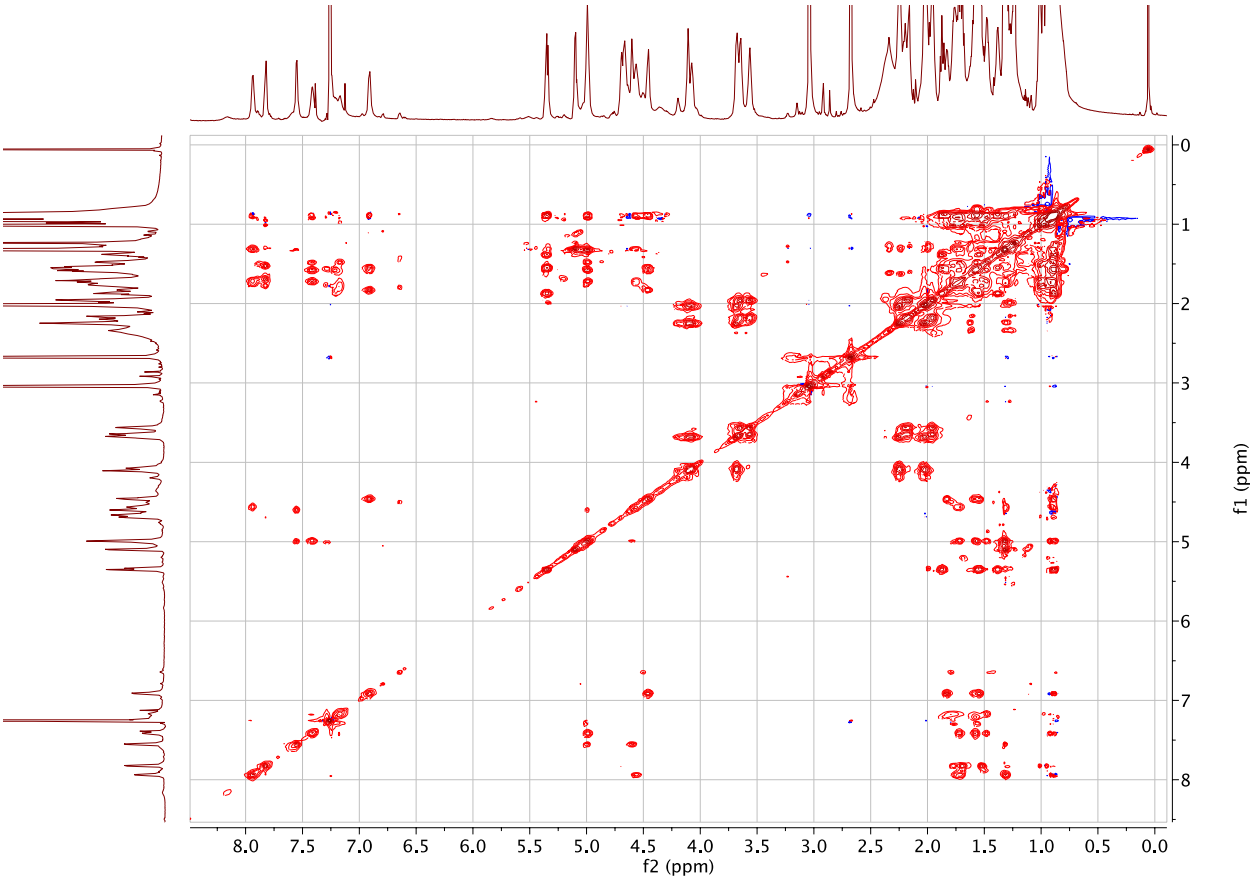

3 #3064 RT: 31.98 AV: 1 NL: 6.28E8
F: FTMS + c ESI Full ms [200.00-2000.00]
1002.65
501.83
1019.68
736.49
267.17
360.79
445.29
528.79
691.43
804.52
863.59
917.60
1047.71
1195.76
1295.78
1428.17
1530.93
1692.97
1792.48
1947.57
Relative Abundance
m/z

H1 NMR spectrum of **2**





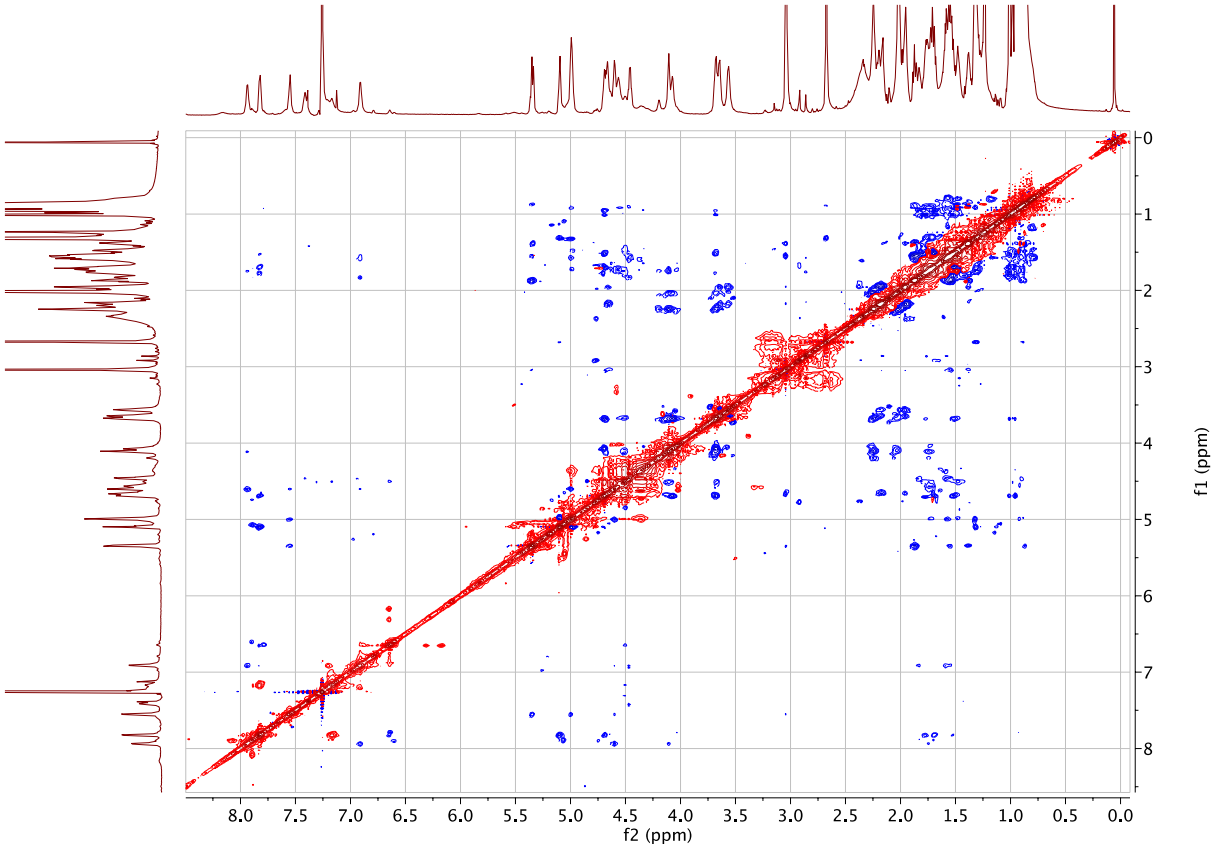
TOCSY spectrum of **2**



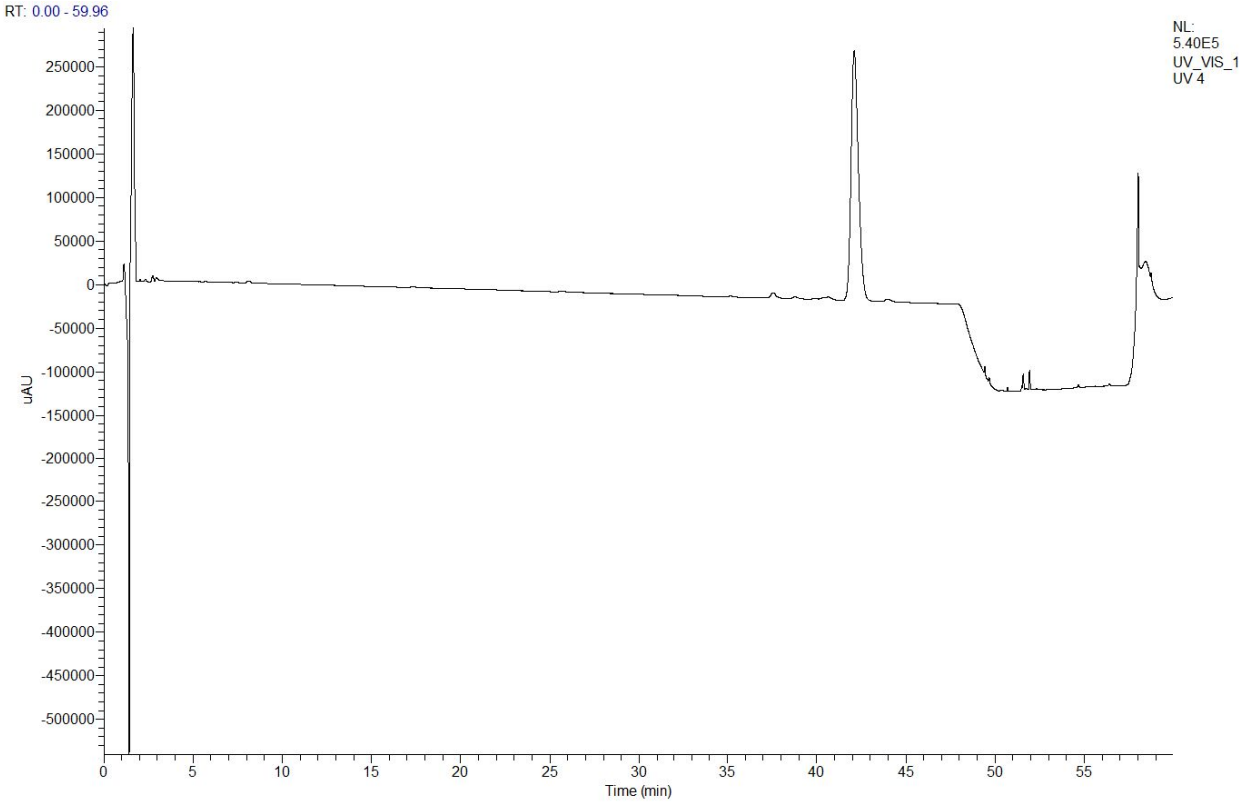
HSQC spectrum of **2**

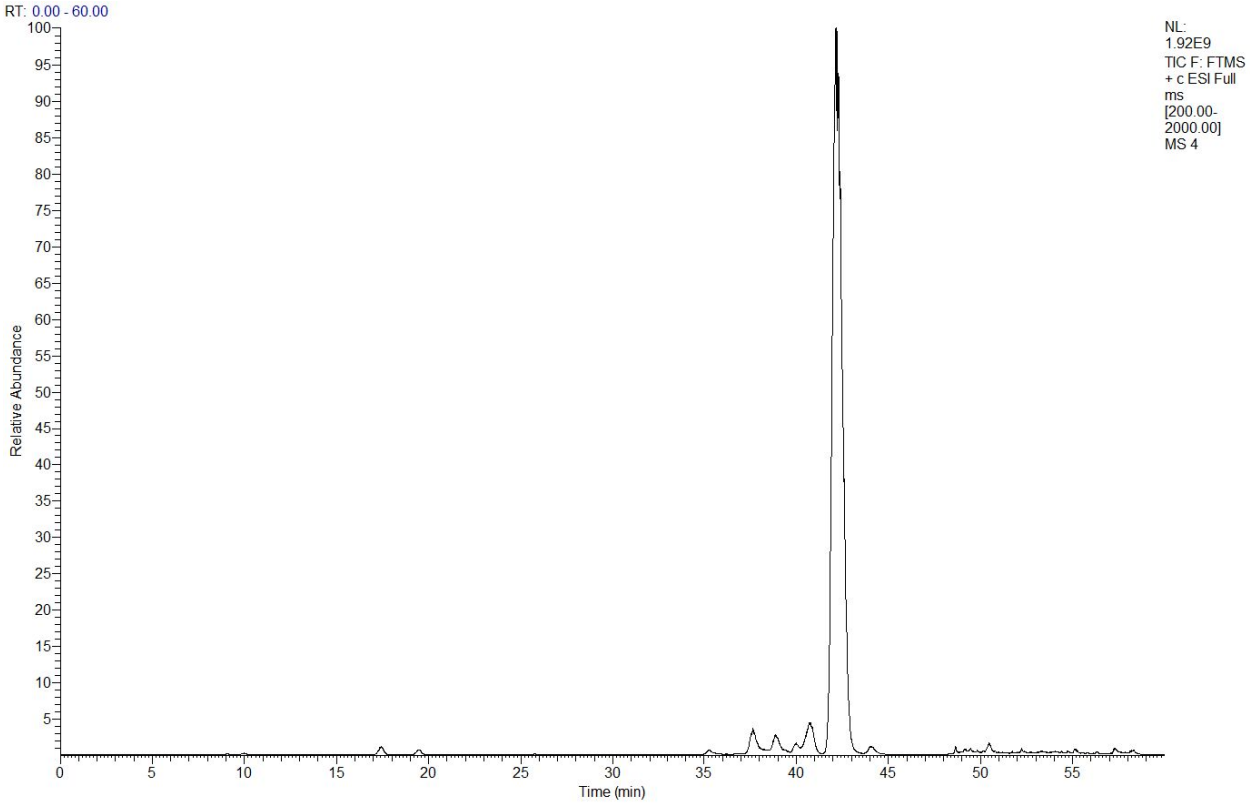
HMBC spectrum of **2**

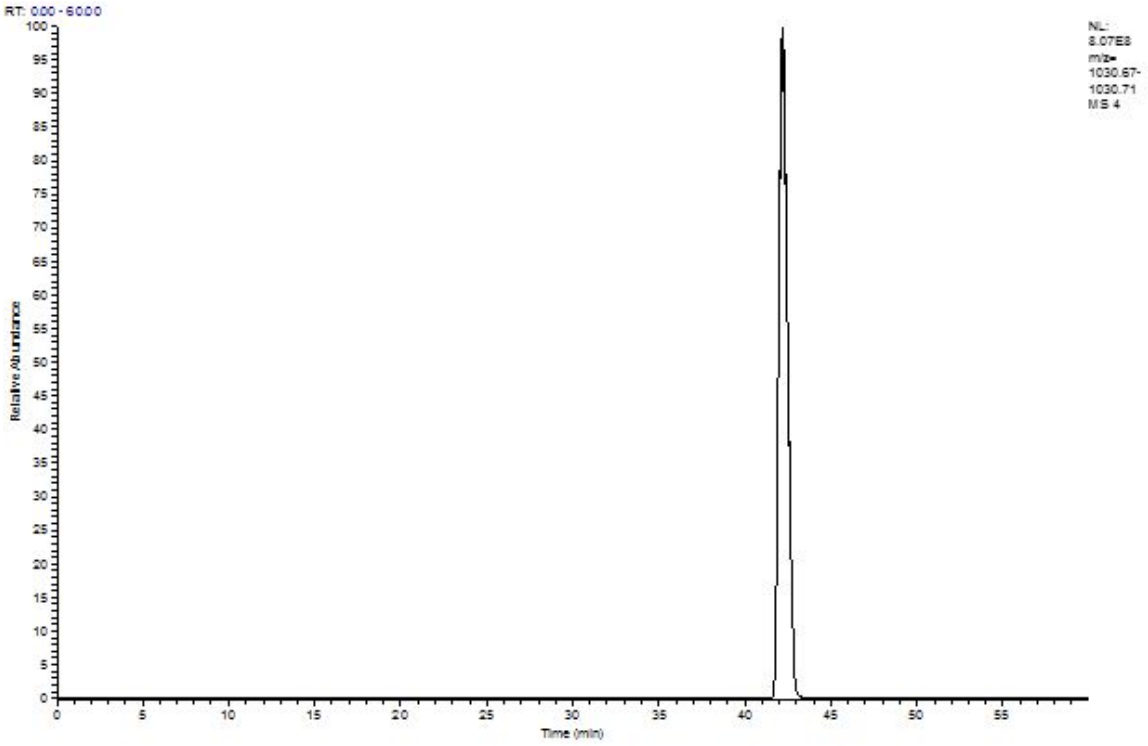
ROESY spectrum of **2**

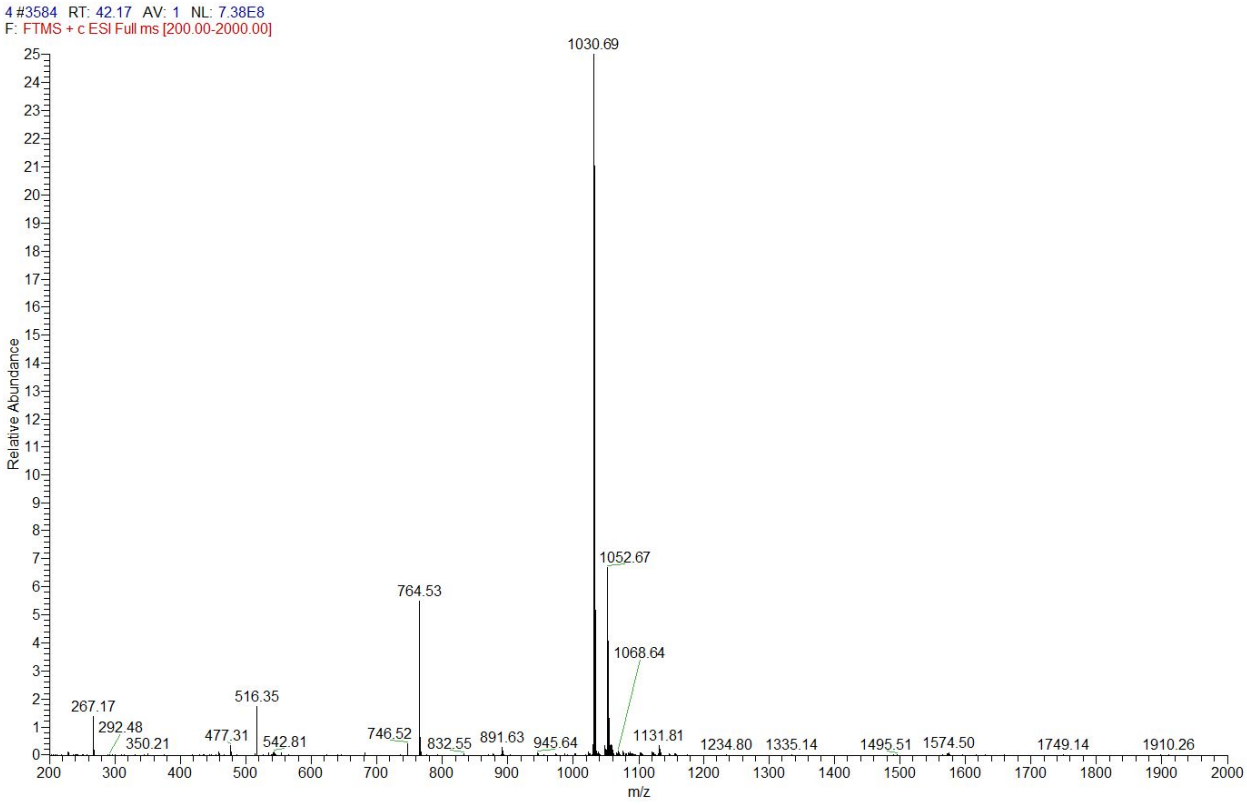


LCMS data for **3**

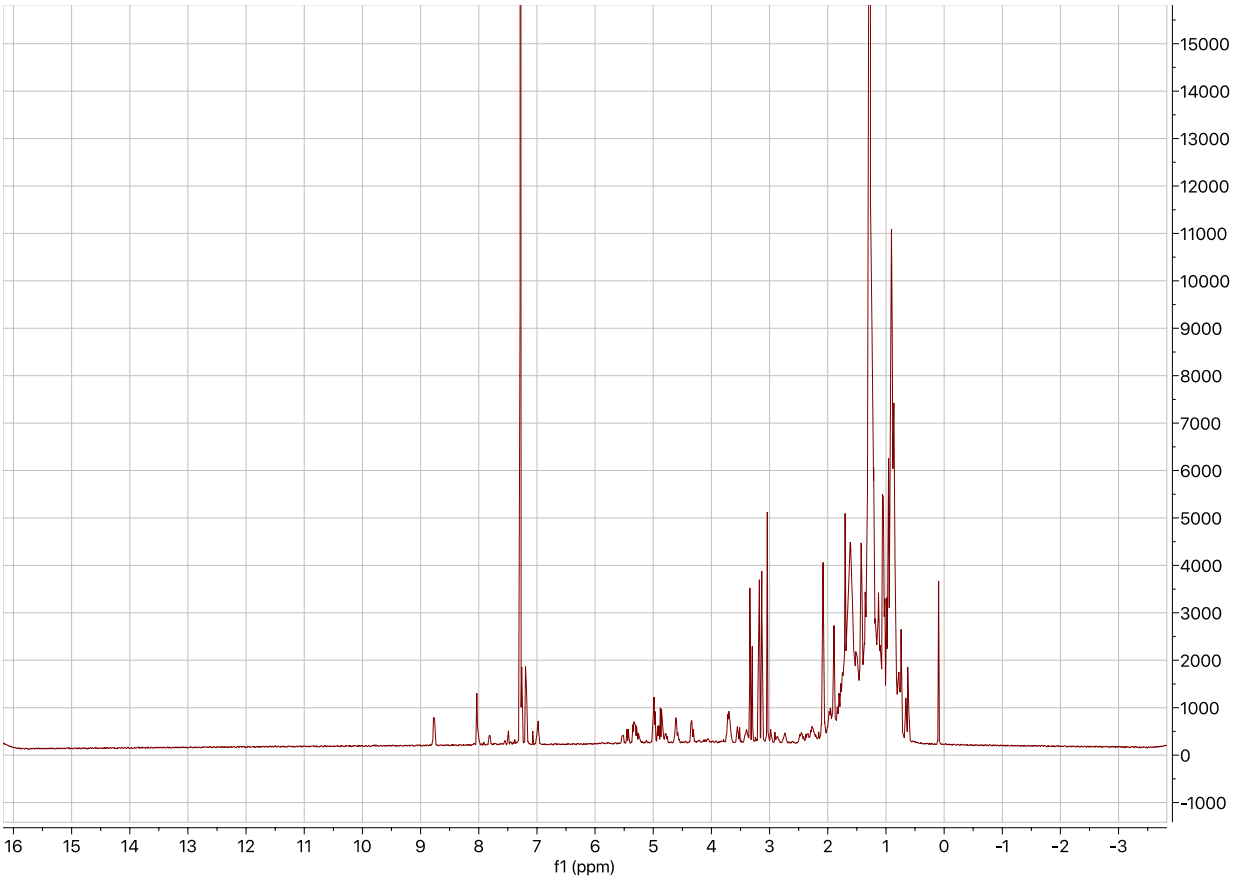


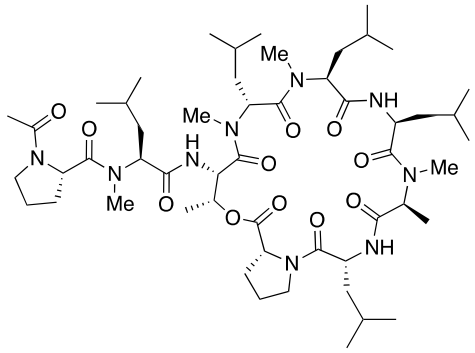



RT: 0.00 - 60.00
NL:
8.07E8
m/z=
1030.67-
1030.71
MS 4
Relative Abundance
Time (min)

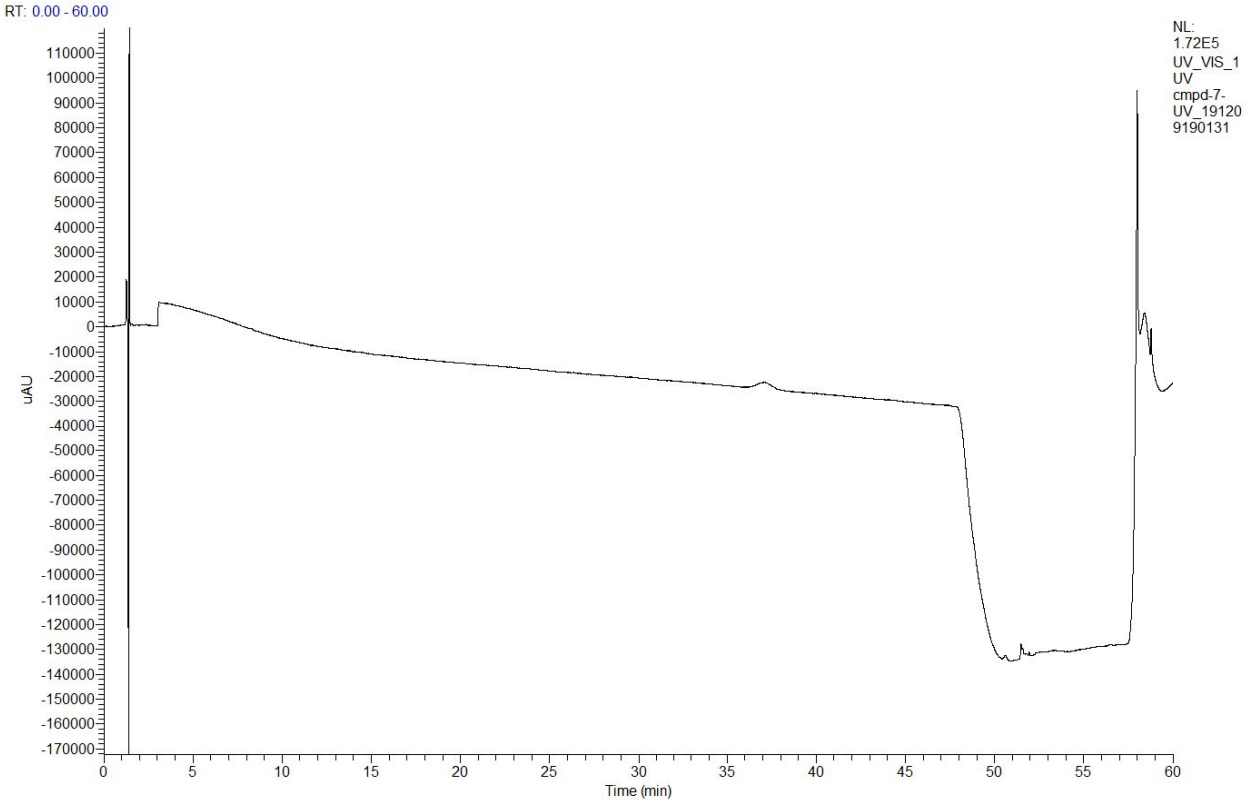

4 #3584 RT: 42.17 AV: 1 NL: 7.38E8
F: FTMS + c ESI Full ms [200.00-2000.00]
1030.69
1052.67
764.53
1068.64
516.35
267.17
292.48
350.21
477.31
542.81
746.52
832.55
891.63
945.64
1131.81
1234.80
1335.14
1495.51
1574.50
1749.14
1910.26
Relative Abundance
m/z

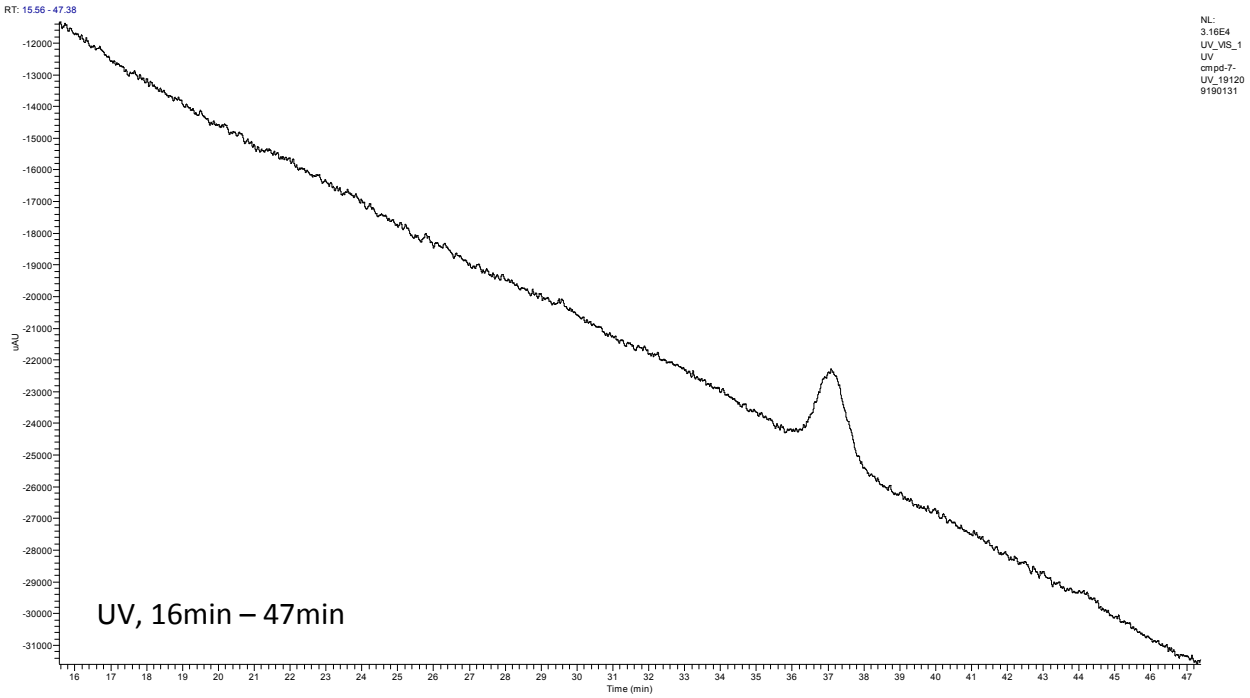
1H NMR spectrum of **3**

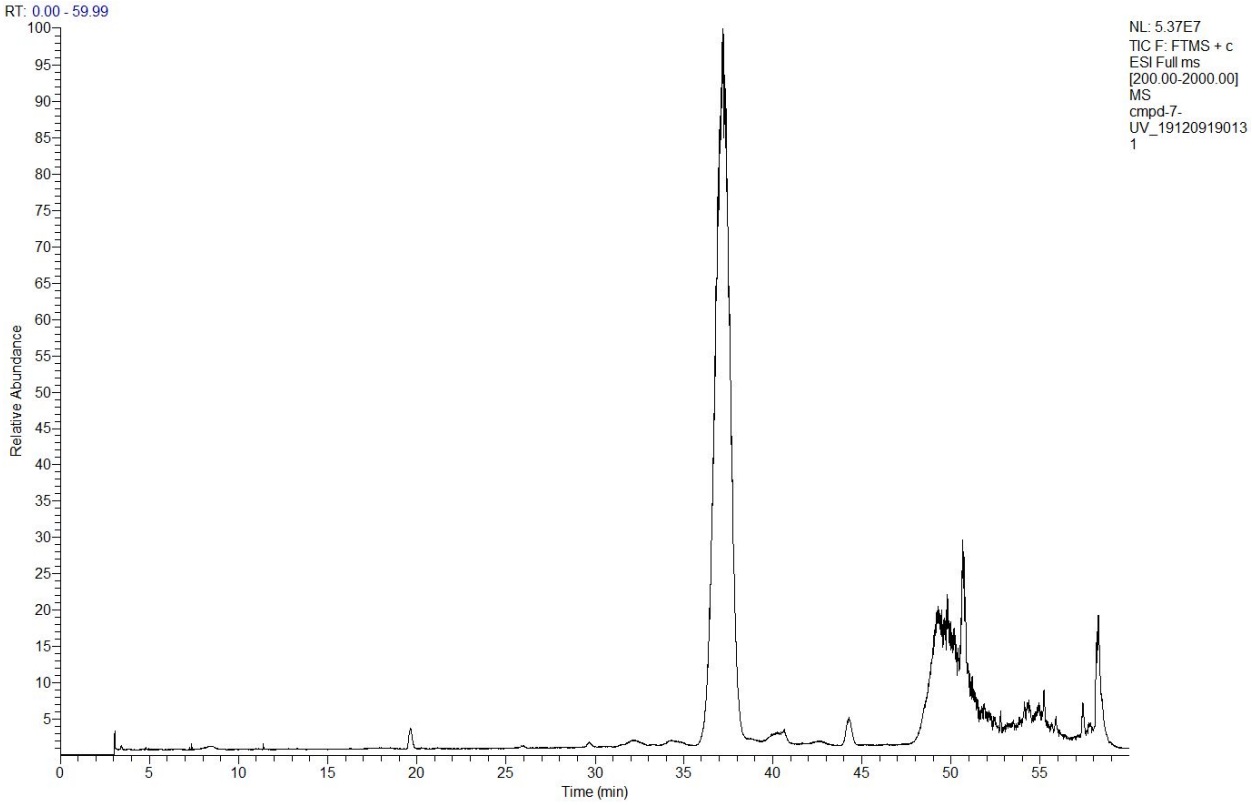


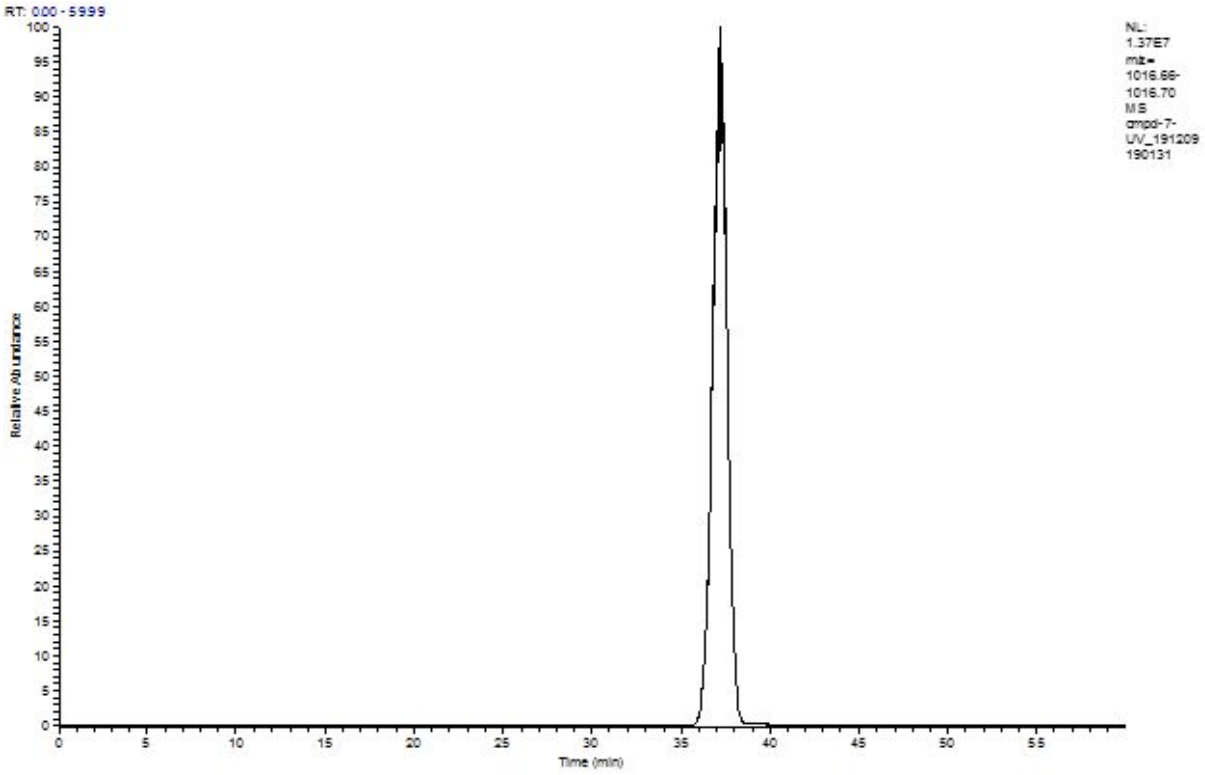


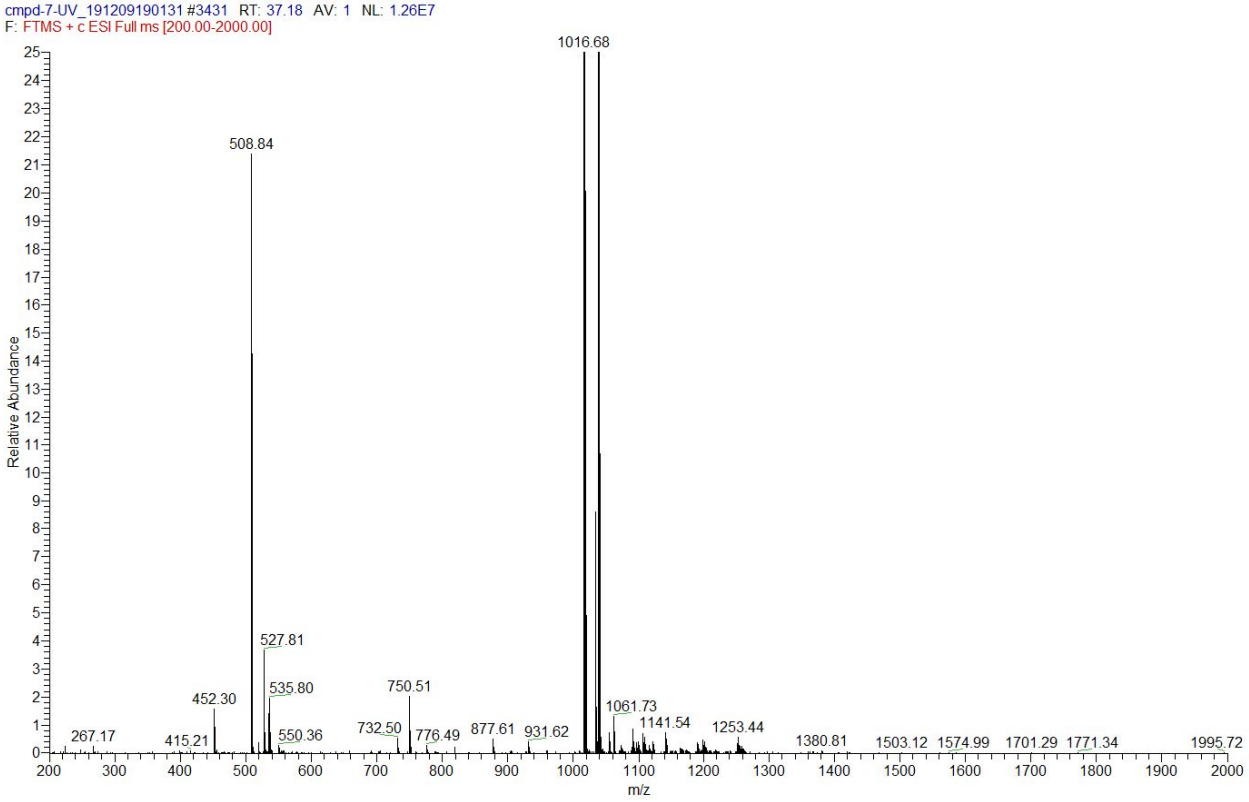
LCMS data for **4**



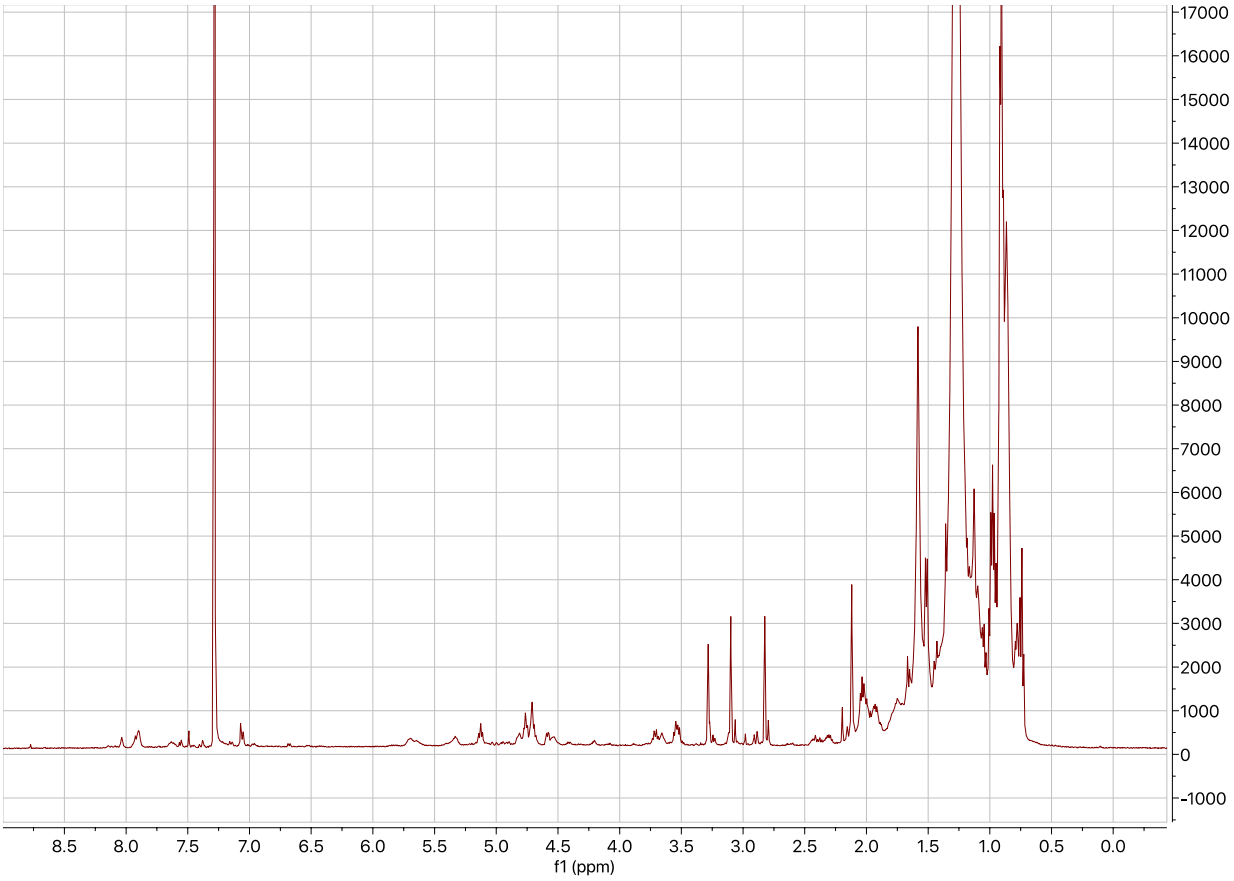


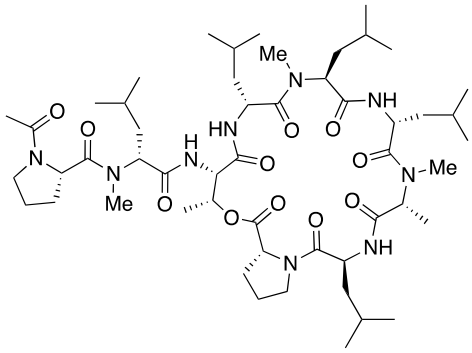

RT: 0.00 - 59.99
NL: 5.37E7
TIC F: FTMS + c
ESI Full ms
[200.00-2000.00]
MS
cmpd-7-
UV_19120919013
1
Relative Abundance
Time (min)


RT: 0.00 - 59.99
NL:
1.37E7
m/z=
1016.66-
1016.70
MS
cmpd-7-
UV_191209
190131
Relative Abundance
Time (min)

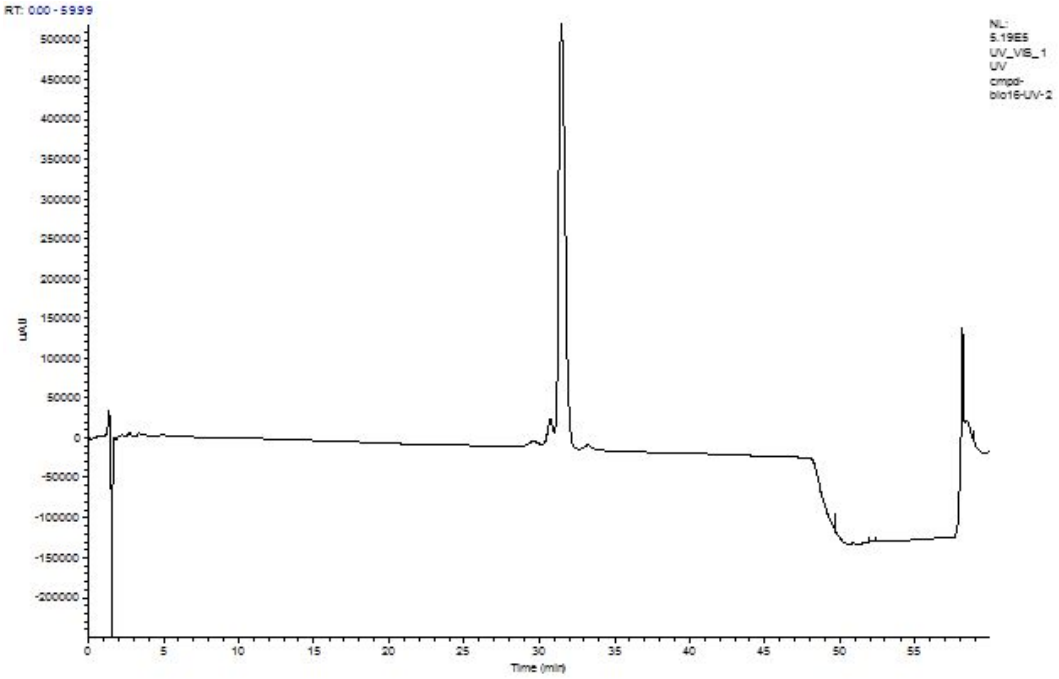

cmpd-7-UV_191209190131 #3431 RT: 37.18 AV: 1 NL: 1.26E7
F: FTMS + c ESI Full ms [200.00-2000.00]
1016.68
508.84
527.81
535.80
452.30
750.51
1061.73
1141.54
1253.44
267.17
415.21
550.36
732.50
776.49
877.61
931.62
1380.81
1503.12
1574.99
1701.29
1771.34
1995.72
Relative Abundance
m/z

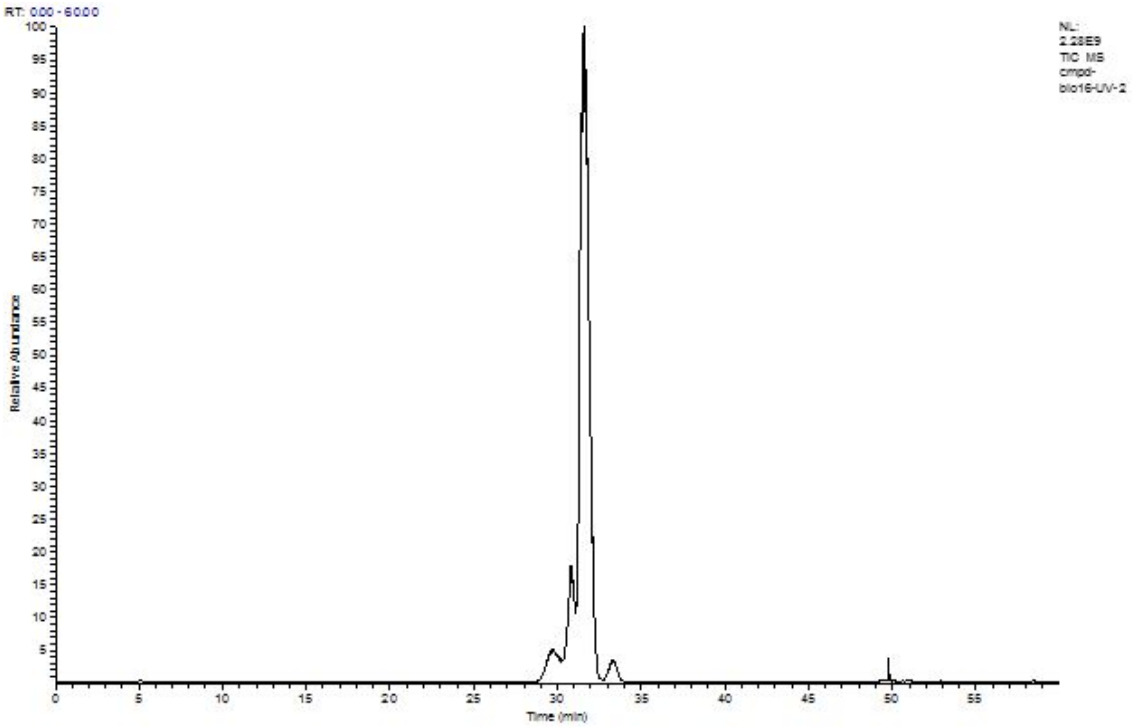
1H NMR spectrum of **4**

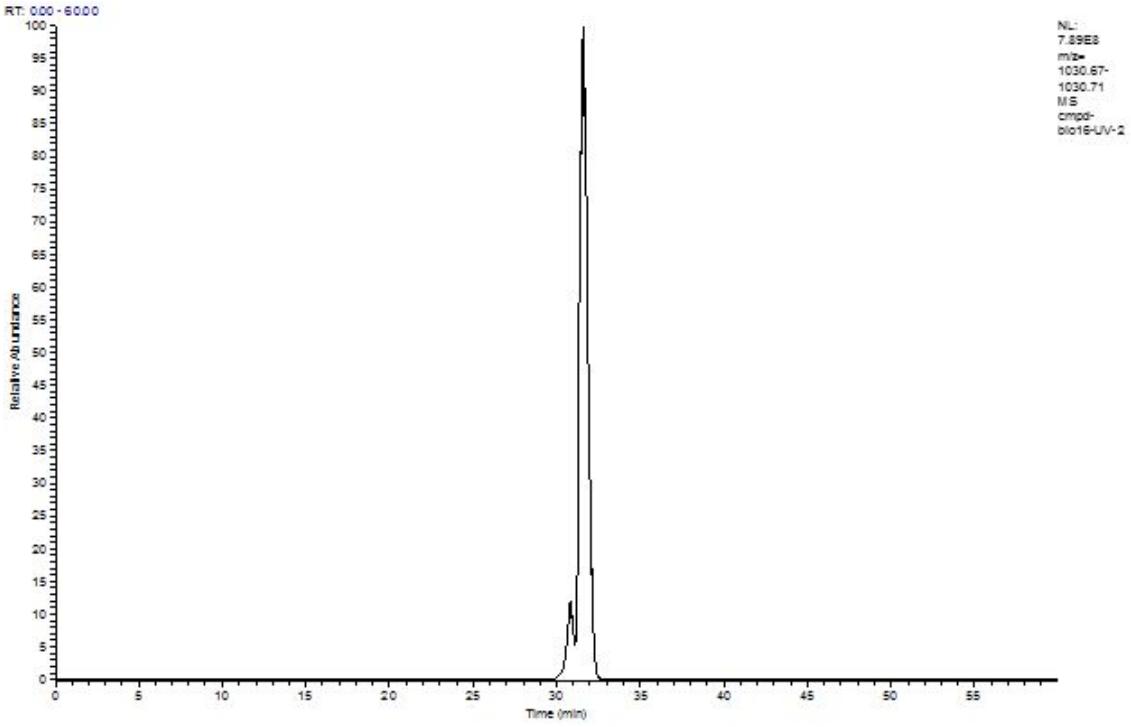


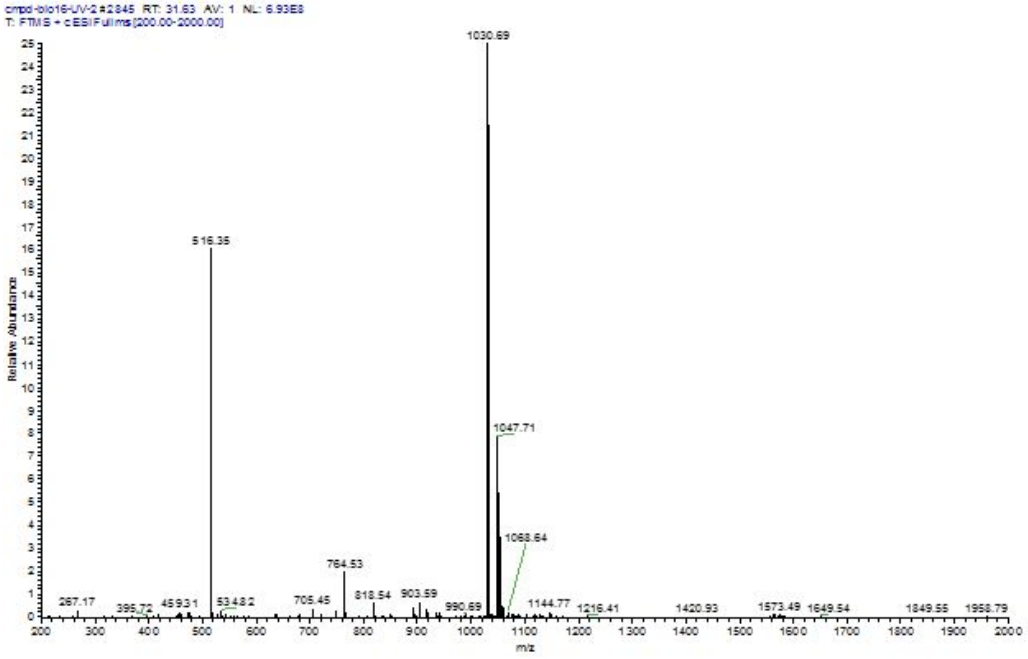


LCMS data for **5**

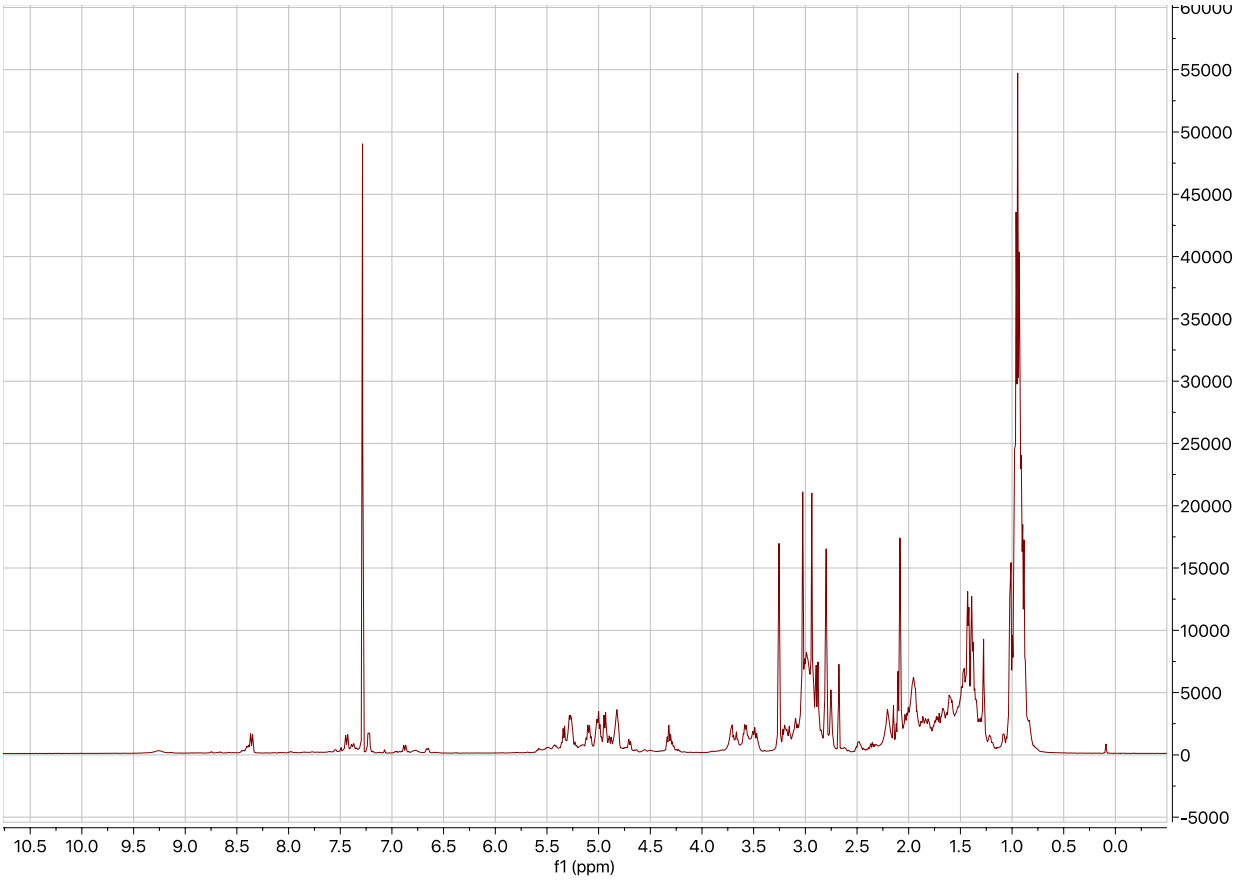


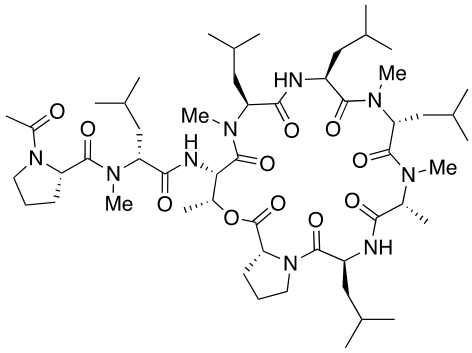



RT: 0.00 - 60.00
NL:
7.89E8
m/z=
1030.67-
1030.71
MS
cmpd-
blo16-UV-2
Relative Abundance
Time (min)

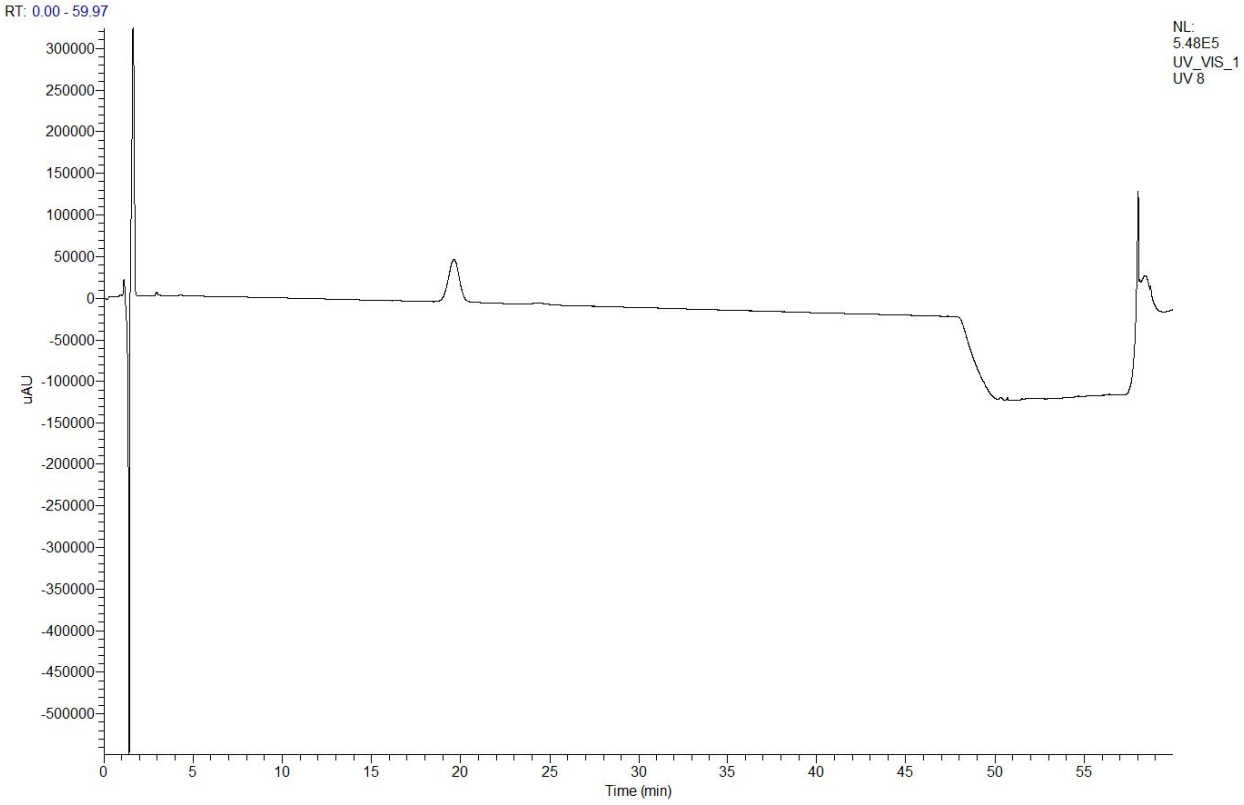

cmpd-blo16-UV-2 #2845 RT: 31.63 AV: 1 NL: 6.93E8
T: FTMS + cESI Full ms [200.00-2000.00]
1030.69
516.35
1047.71
1068.64
764.53
267.17
395.22
459.31
534.82
705.45
818.54
903.59
990.69
1144.77
1216.41
1420.93
1573.49
1649.54
1849.55
1958.79
Relative Abundance
m/z

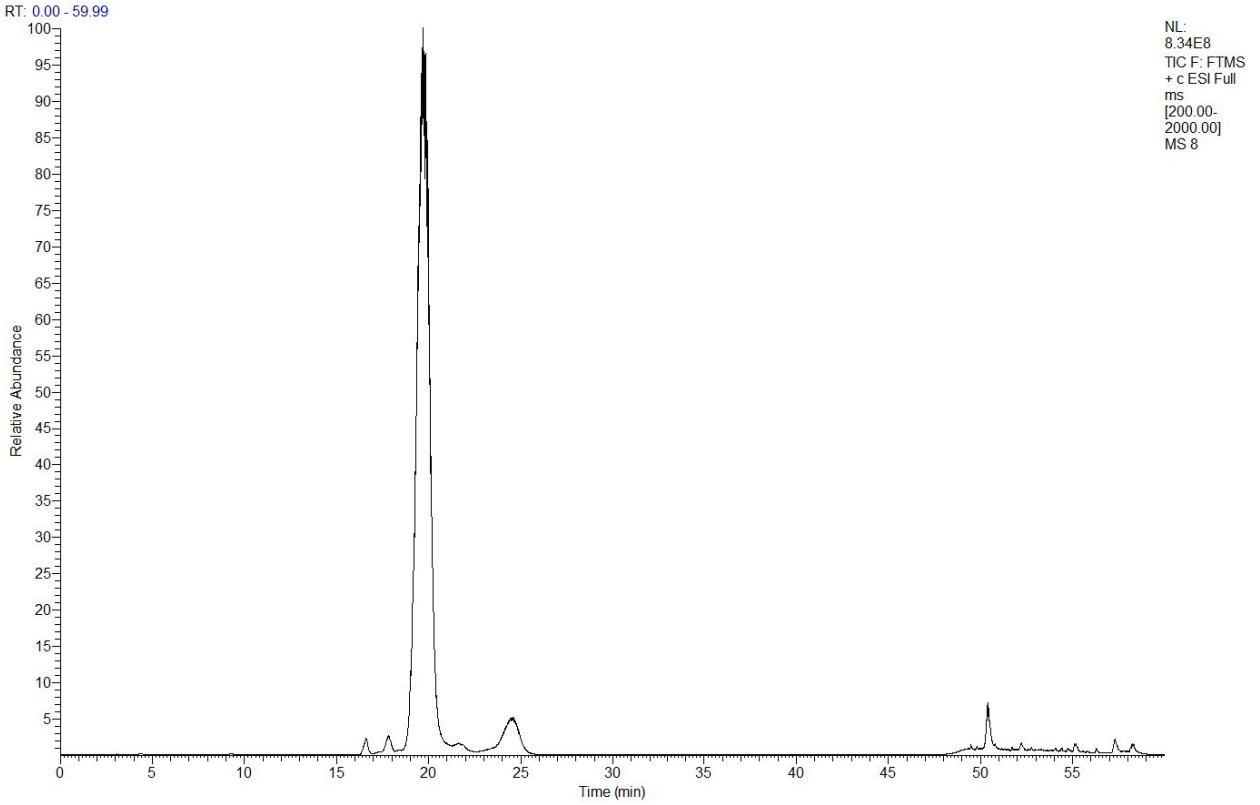
H1 NMR spectrum of **5**

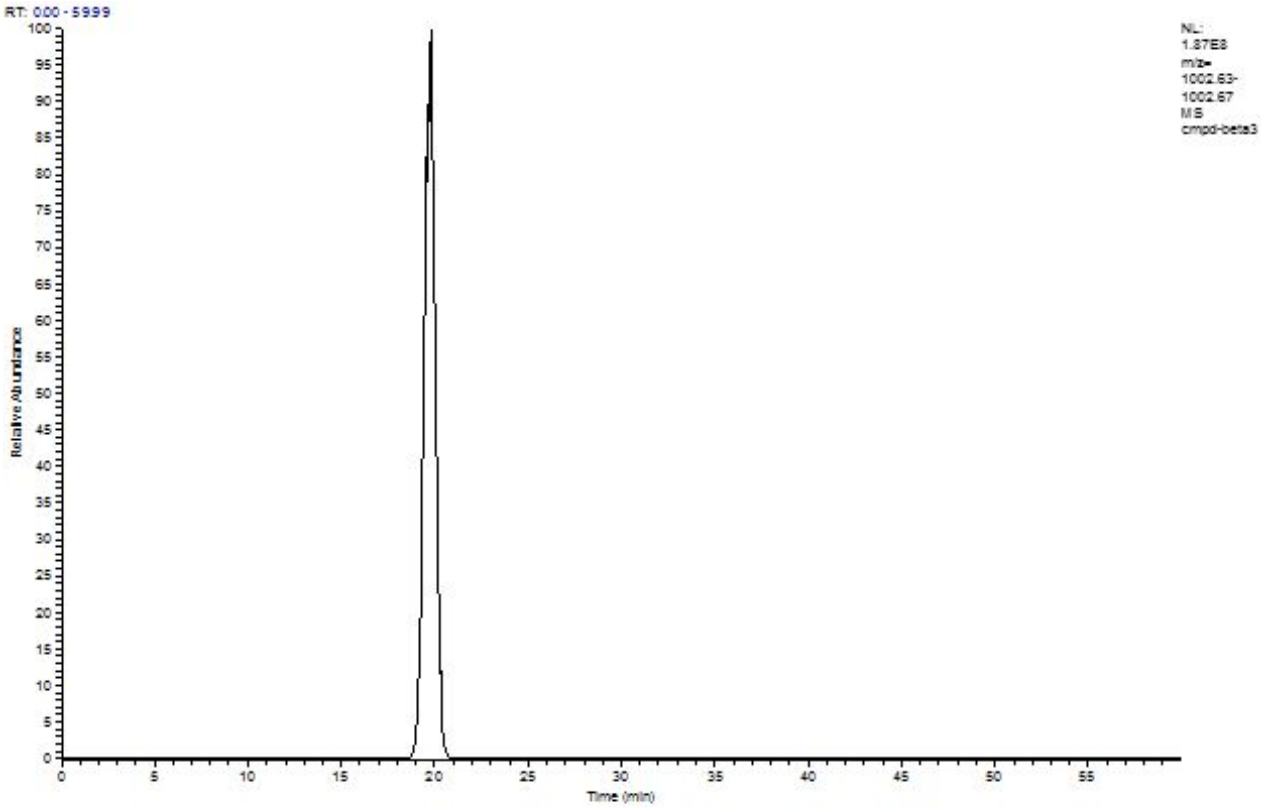


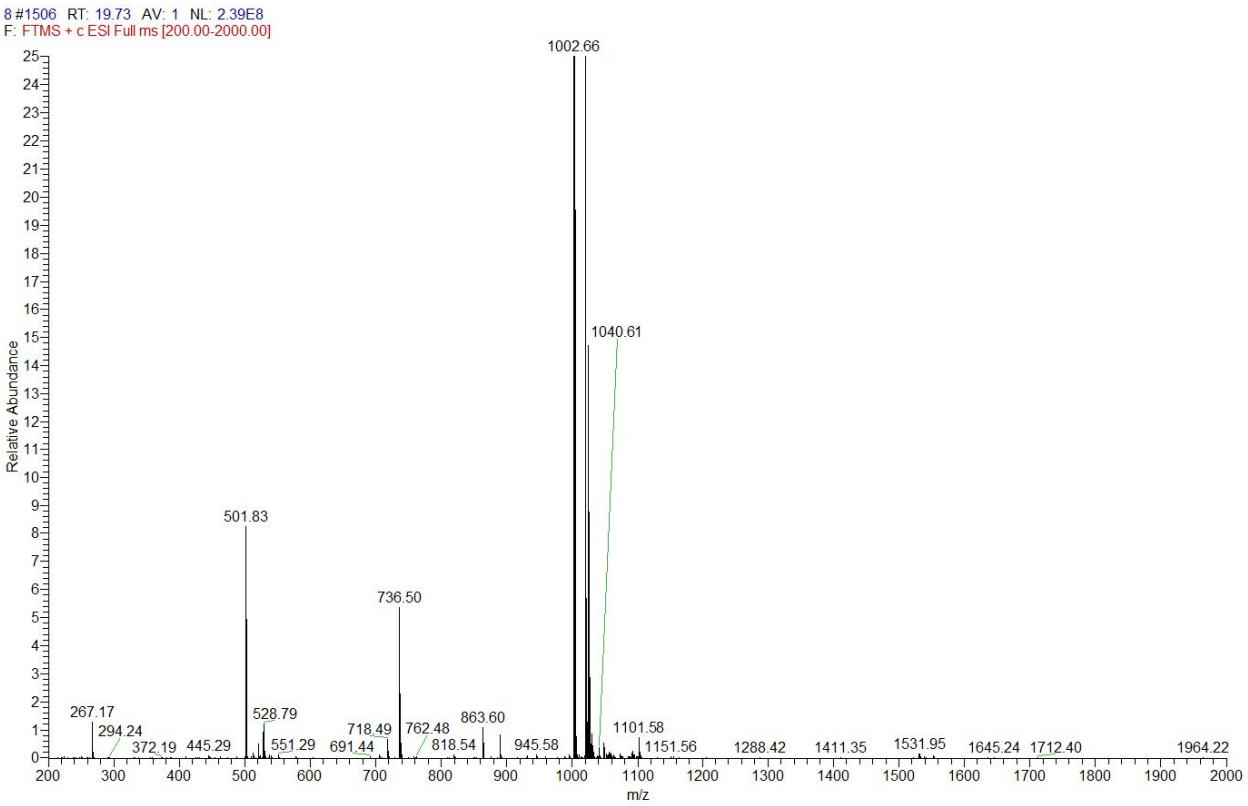


LCMS data for **6**

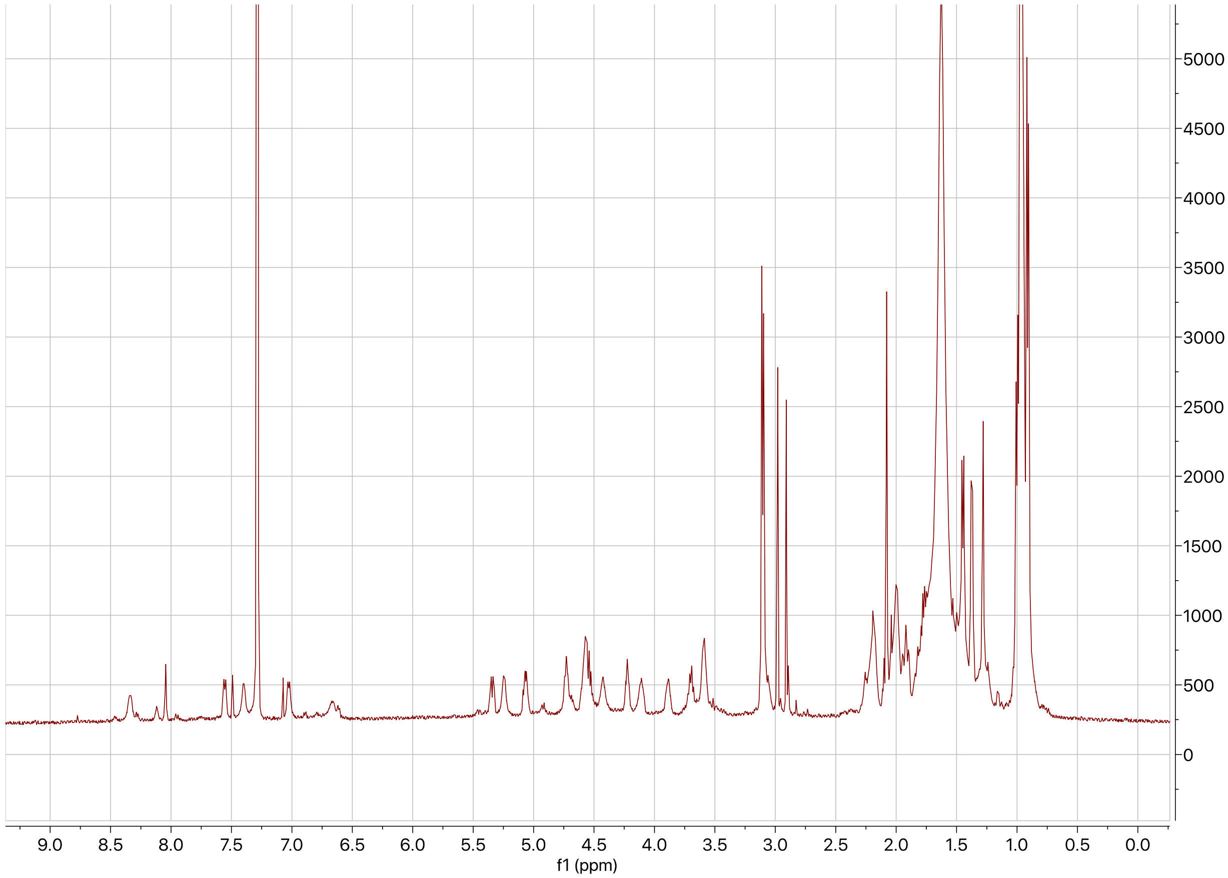


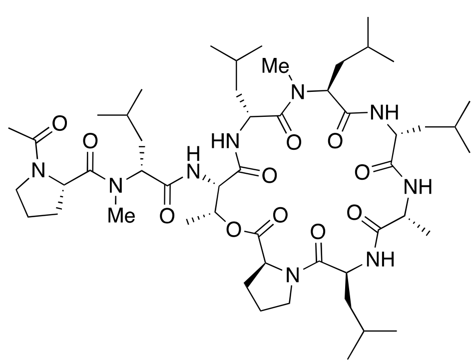

RT: 0.00 - 59.99
NL:
8.34E8
TIC F: FTMS
+ c ESI Full
ms
[200.00-
2000.00]
MS 8
Relative Abundance
Time (min)


RT: 0.00 - 59.99
NL:
1.87E8
m/z=
1002.63-
1002.67
MS
cmpd-beta3
Relative Abundance
Time (min)

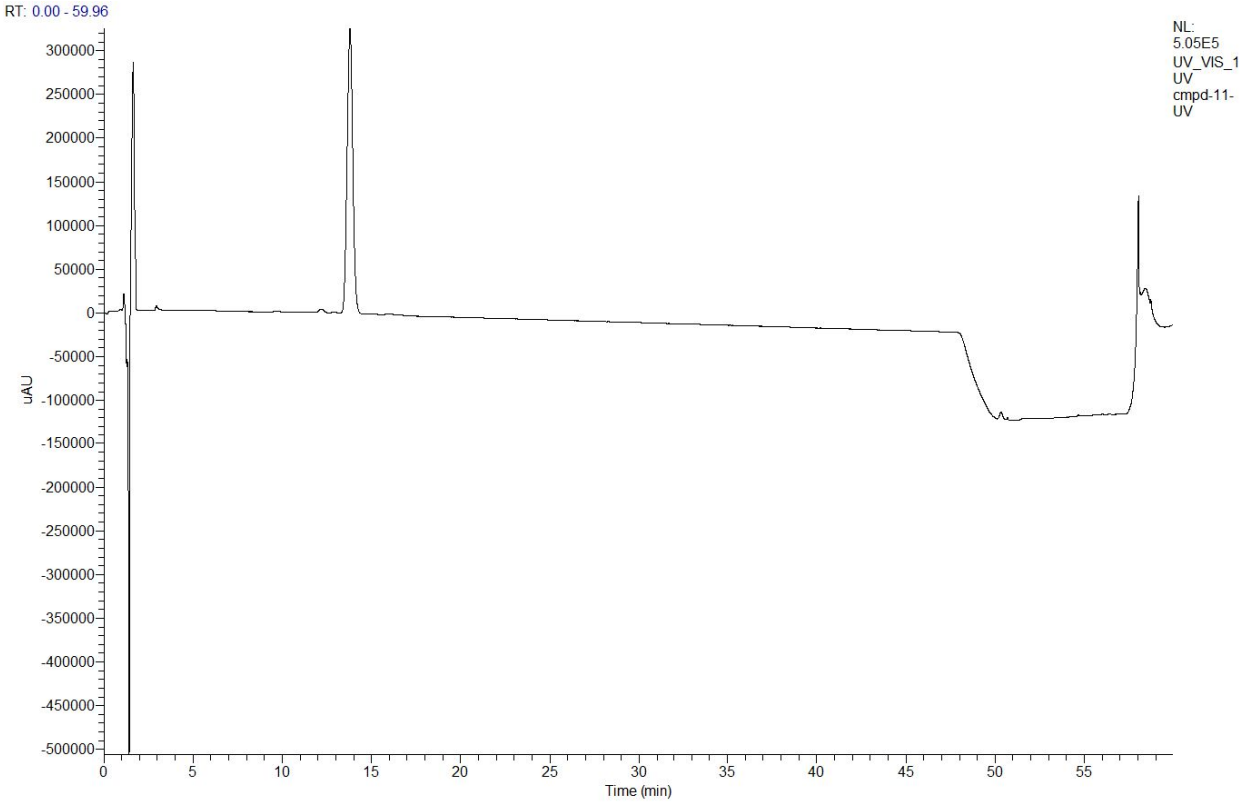

8 #1506 RT: 19.73 AV: 1 NL: 2.39E8
F: FTMS + c ESI Full ms [200.00-2000.00]
Relative Abundance
m/z
1002.66
1040.61
501.83
736.50
267.17
294.24
372.19
445.29
528.79
551.29
691.44
718.49
762.48
818.54
863.60
945.58
1101.58
1151.56
1288.42
1411.35
1531.95
1645.24
1712.40
1964.22

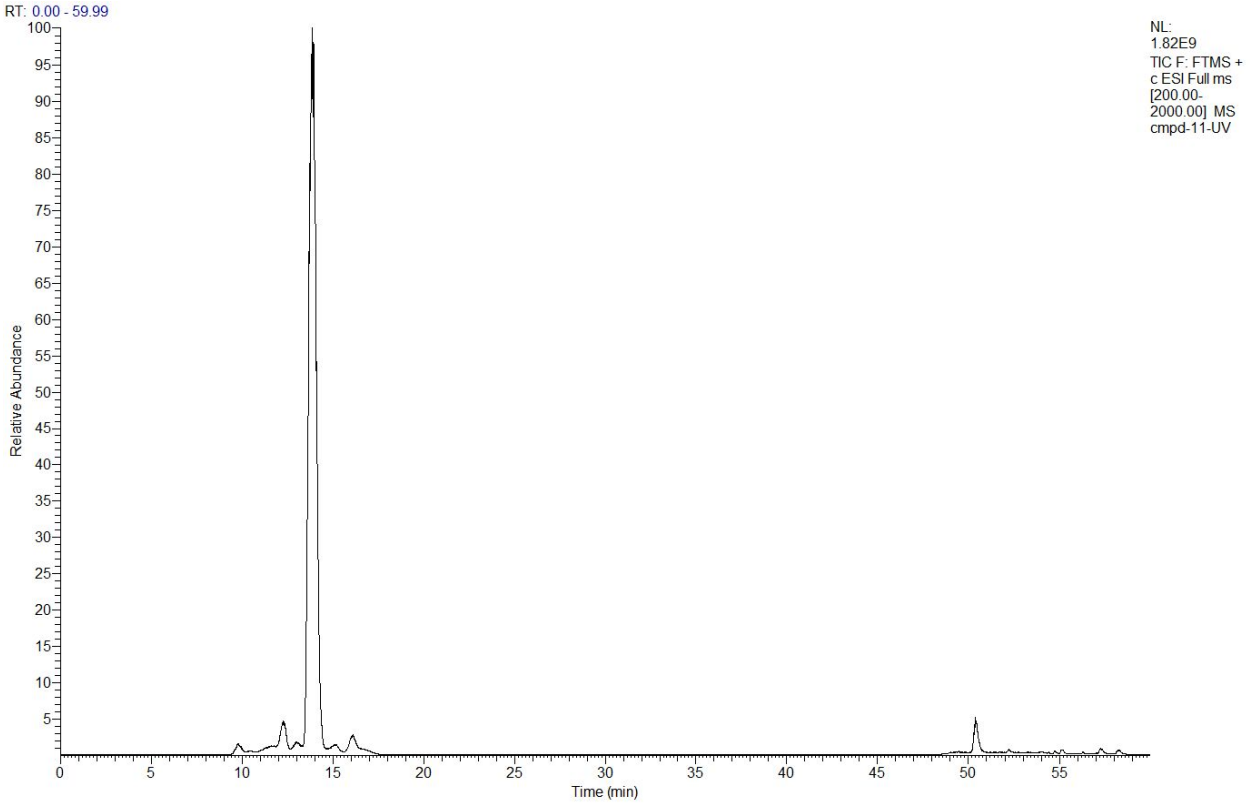
H1 NMR spectrum of **6**

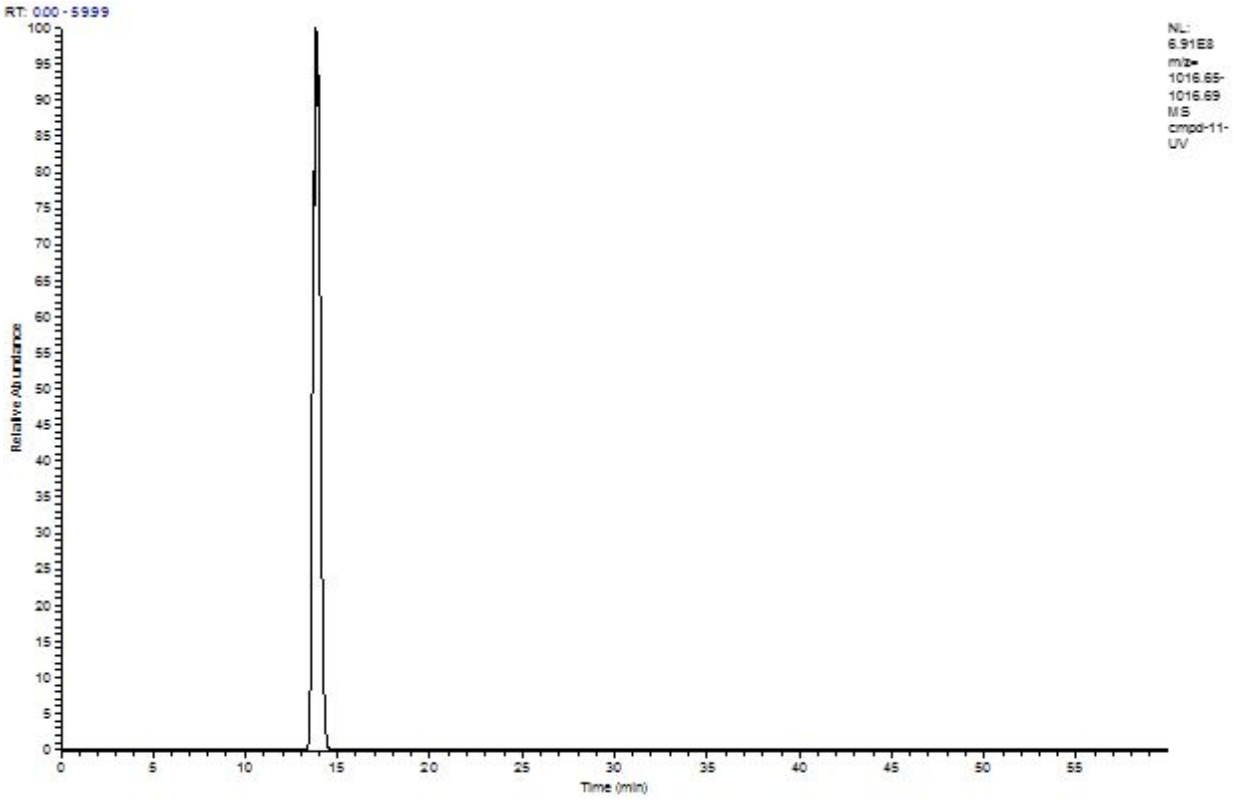


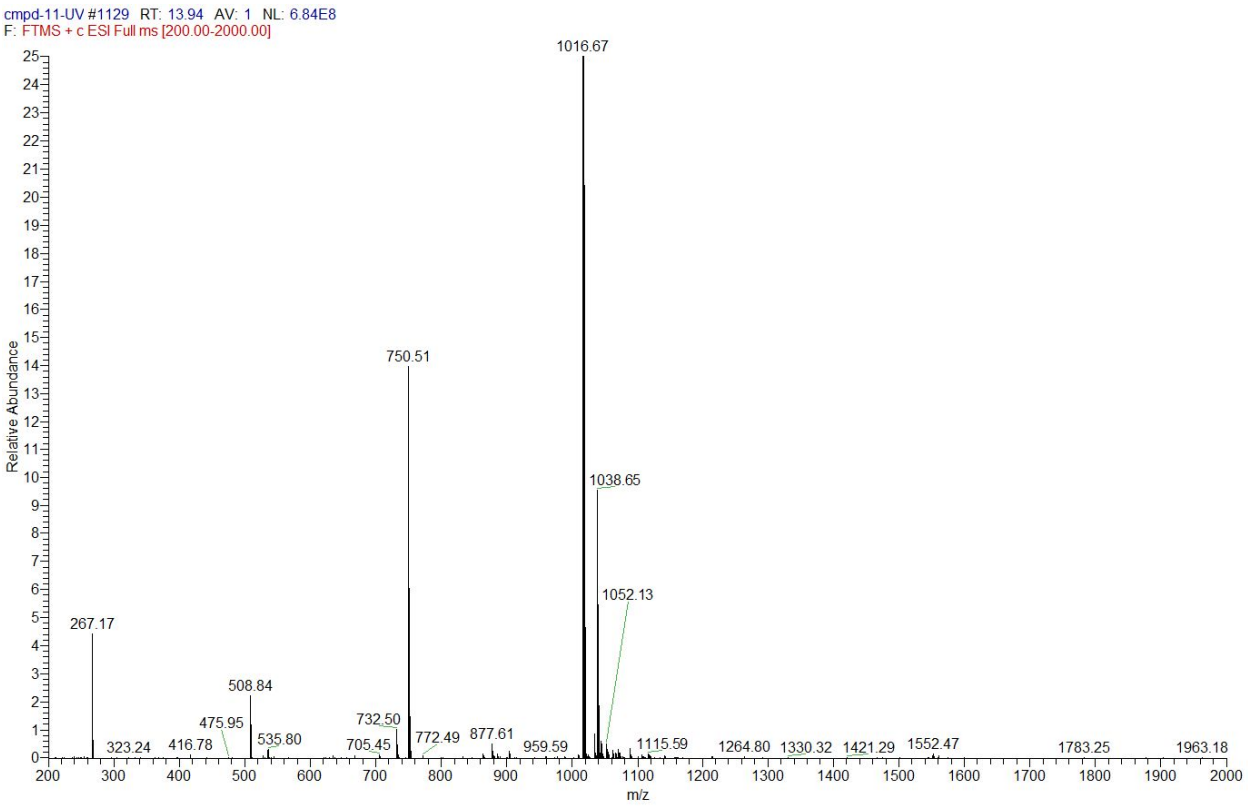


LCMS data for **7**

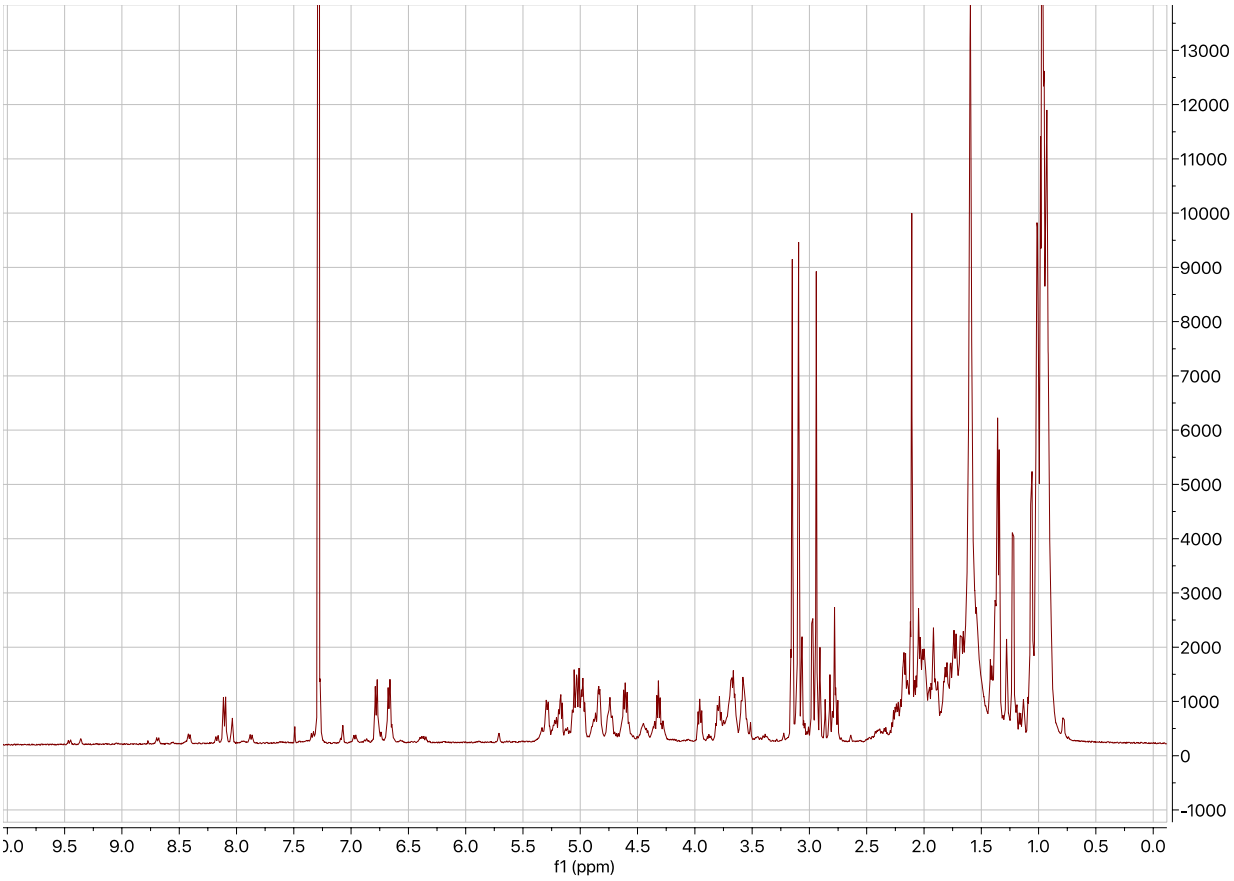


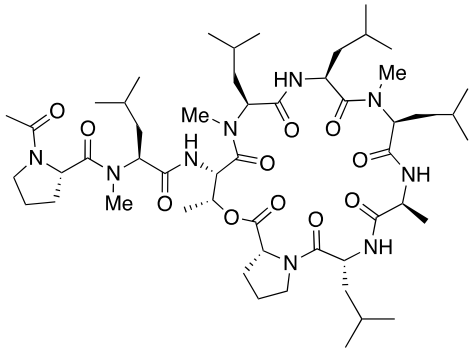

RT: 0.00 - 59.99
NL:
1.82E9
TIC F: FTMS +
c ESI Full ms
[200.00-
2000.00] MS
cmpd-11-UV
Relative Abundance
Time (min)


RT: 0.00 - 59.99
NL:
6.91E8
m/z=
1016.65-
1016.69
MS
cmpd-11-
UV
Relative Abundance
Time (min)

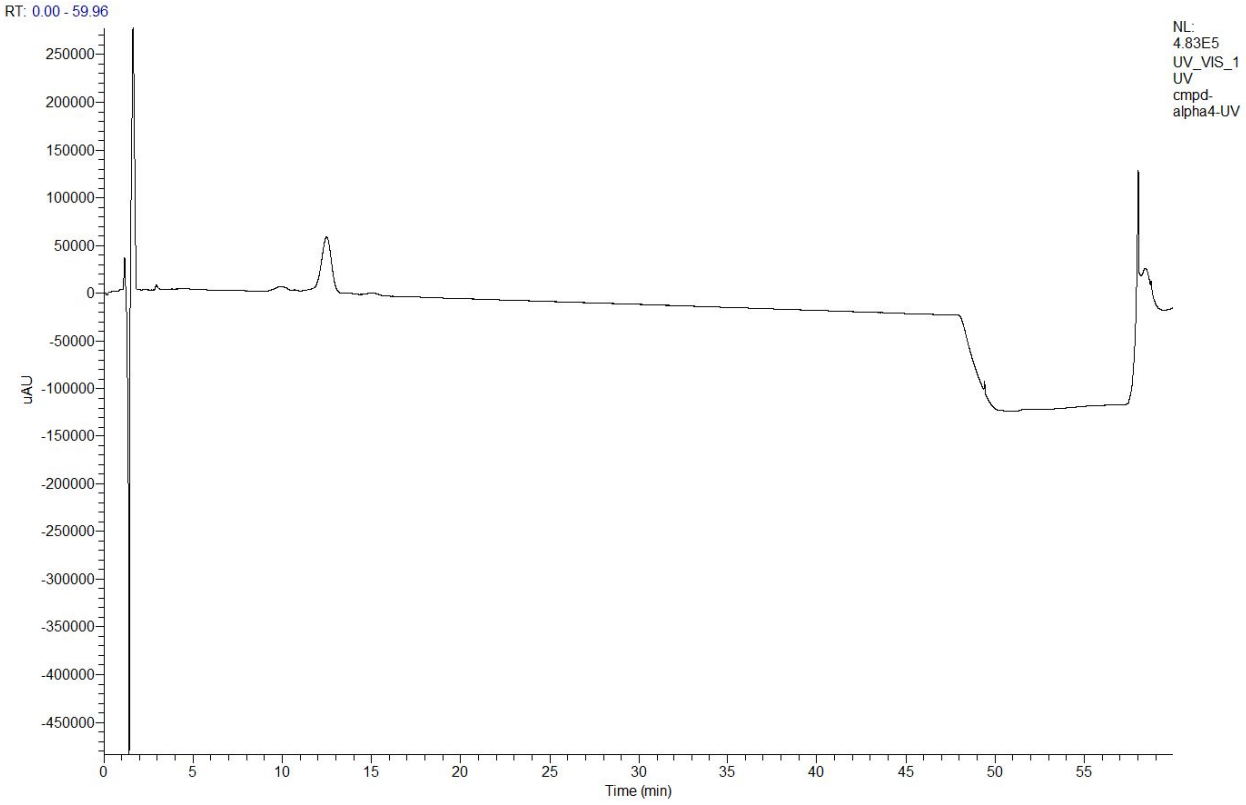

cmpd-11-UV #1129 RT: 13.94 AV: 1 NL: 6.84E8
F: FTMS + c ESI Full ms [200.00-2000.00]
Relative Abundance
m/z
1016.67
750.51
1038.65
1052.13
267.17
508.84
732.50
475.95
323.24
416.78
535.80
705.45
772.49
877.61
959.59
1115.59
1264.80
1330.32
1421.29
1552.47
1783.25
1963.18

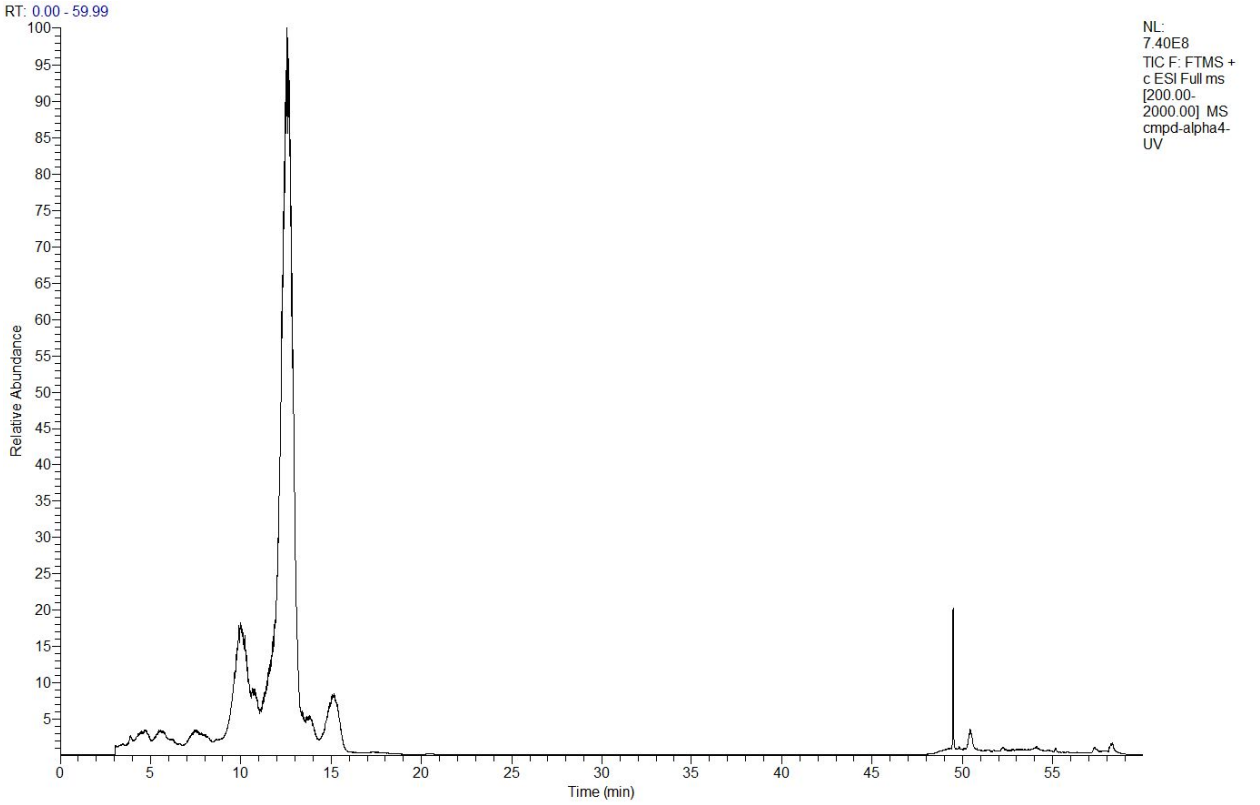
H1 NMR spectrum of **7**

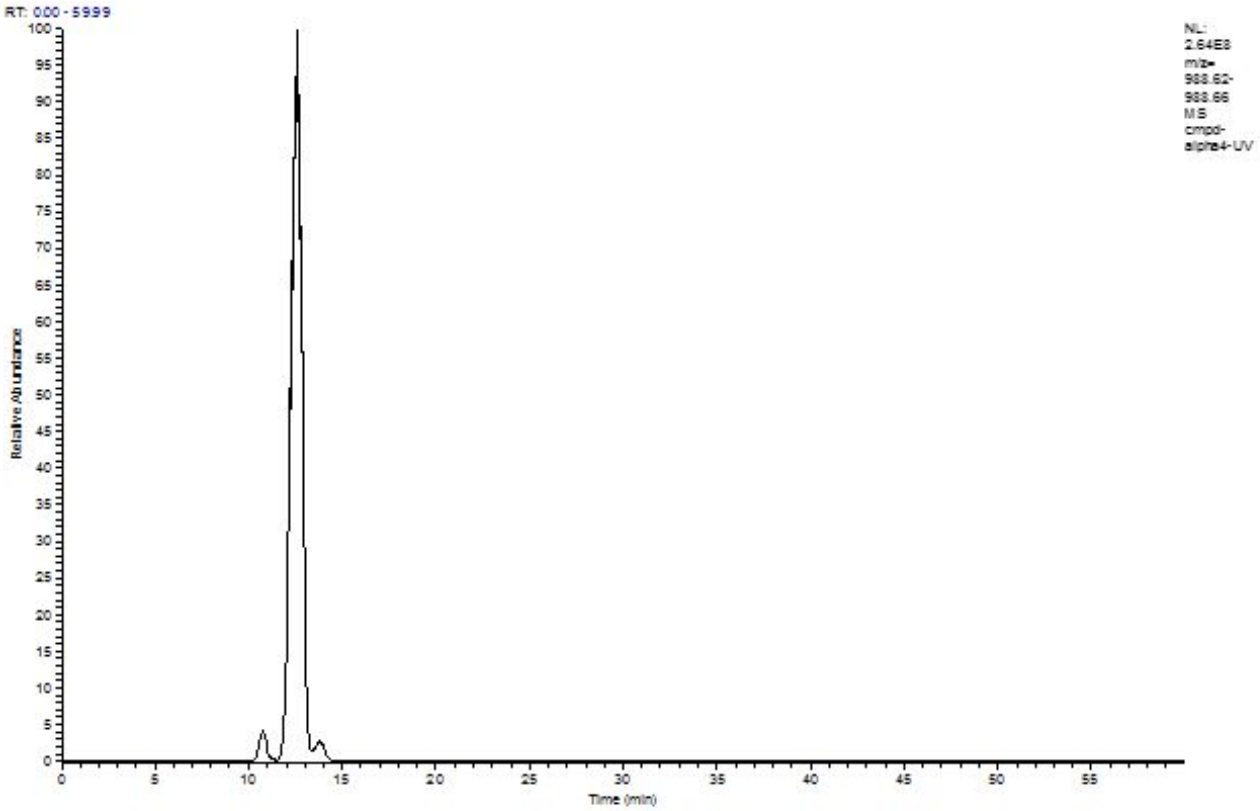


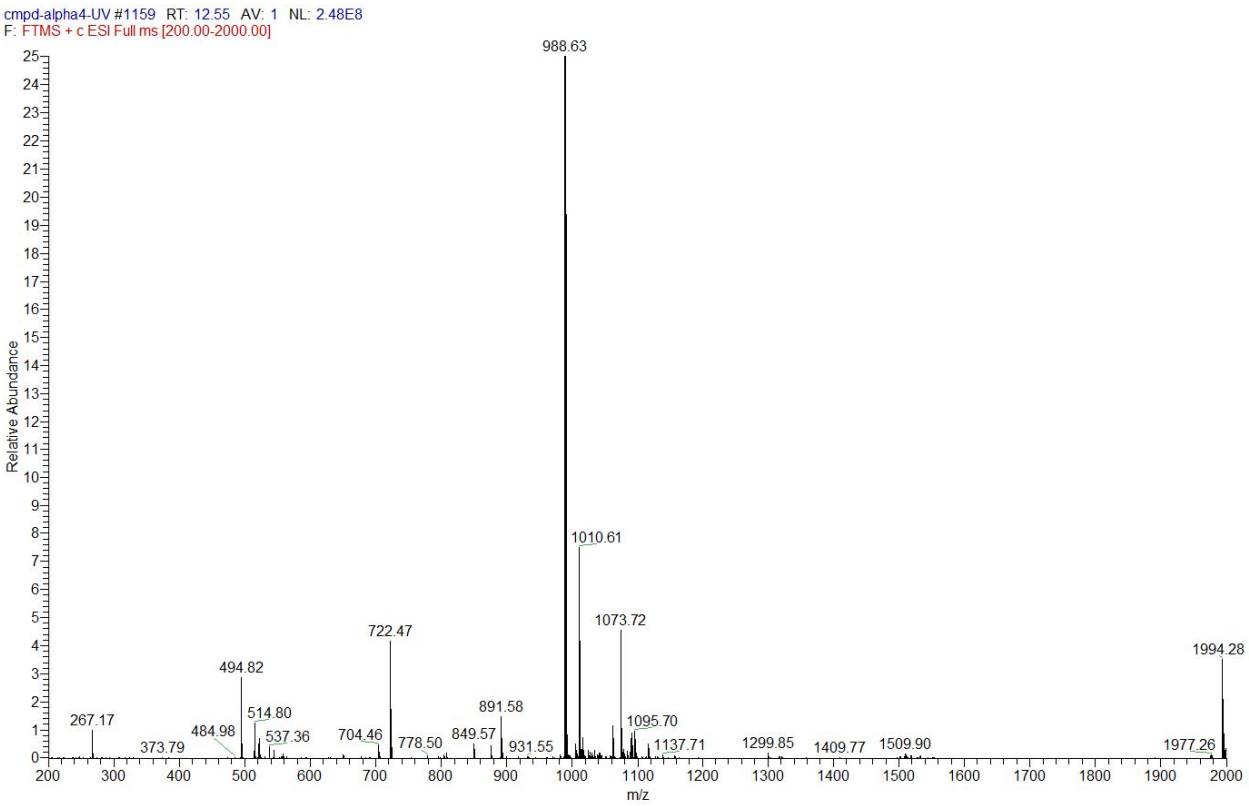


LCMS data for **8**

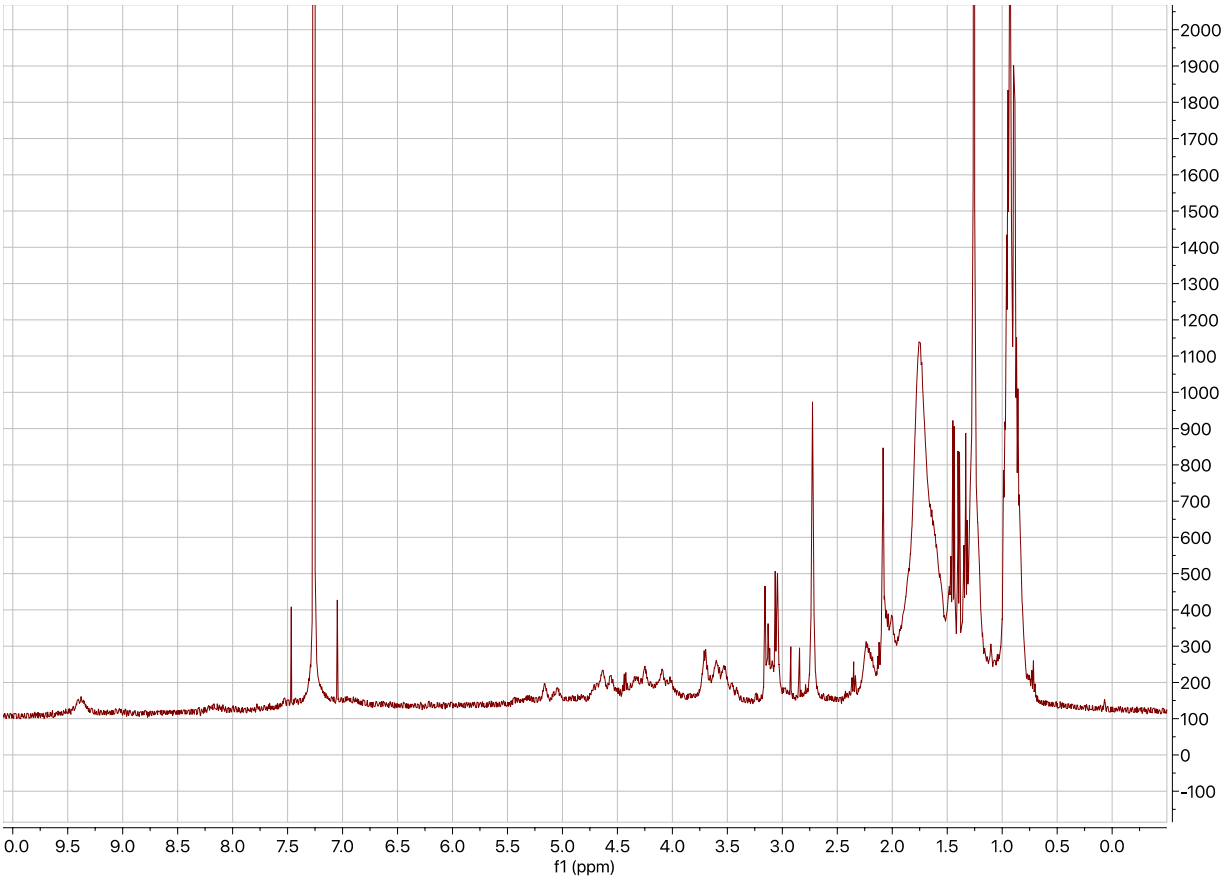


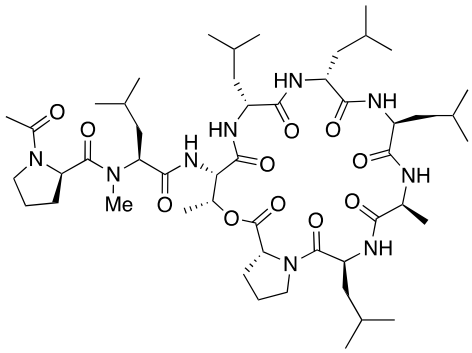

RT: 0.00 - 59.99
NL:
7.40E8
TIC F: FTMS +
c ESI Full ms
[200.00-
2000.00] MS
cmpd-alpha4-
UV
Relative Abundance
Time (min)


RT: 0.00 - 59.99
NL:
2.64E8
m/z=
988.62-
988.66
MS
cmpd-
alpha4-UV
Relative Abundance
Time (min)

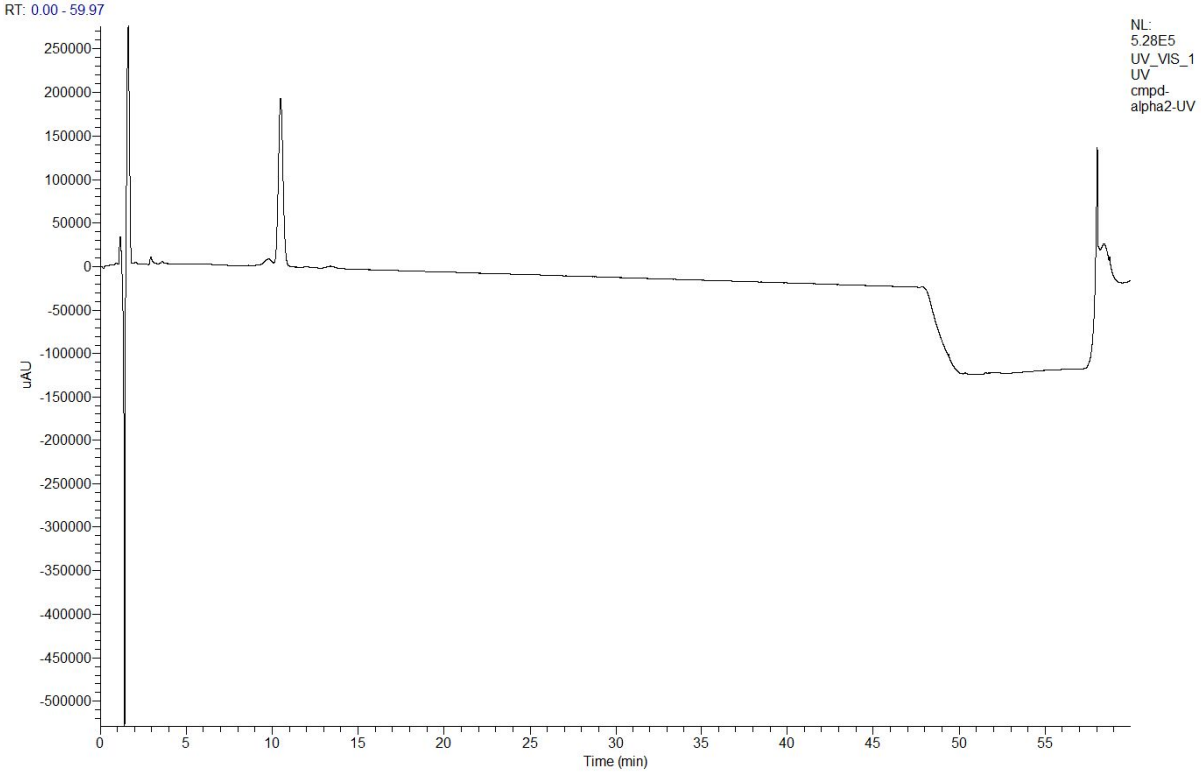

cmpd-alpha4-UV #1159 RT: 12.55 AV: 1 NL: 2.48E8
F: FTMS + c ESI Full ms [200.00-2000.00]
Relative Abundance
m/z
988.63
1010.61
1073.72
722.47
494.82
1994.28
891.58
514.80
267.17
484.98
1095.70
537.36
704.46
778.50
849.57
931.55
1137.71
1299.85
1409.77
1509.90
1977.26
373.79

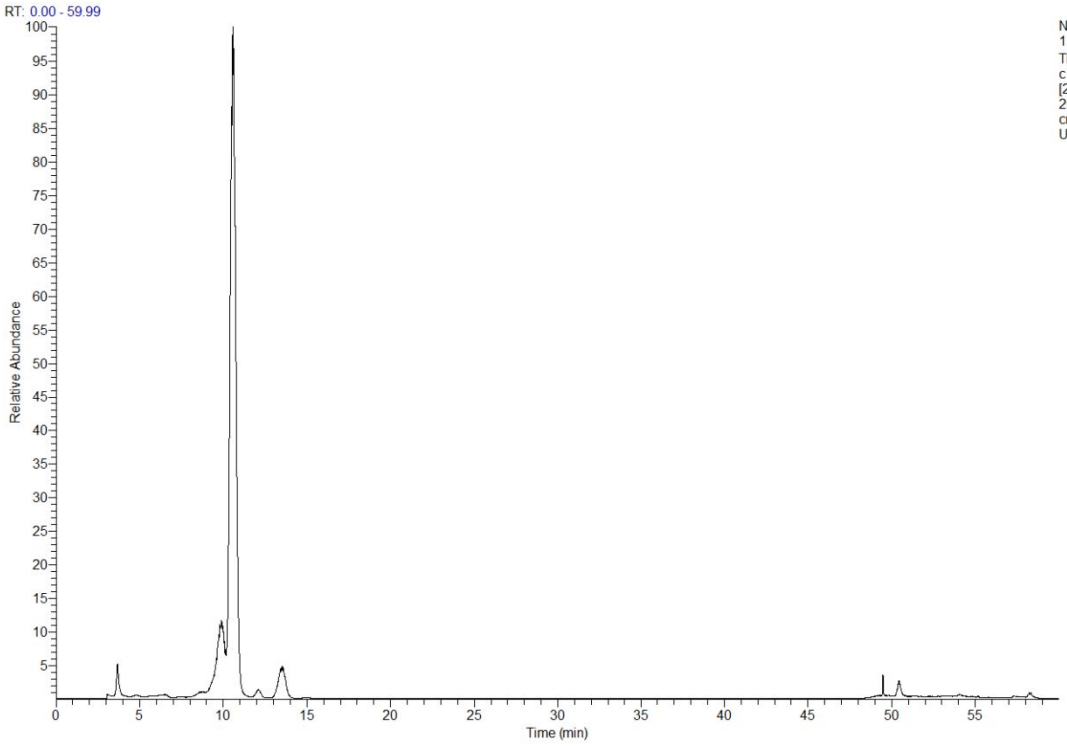
H1 NMR spectrum of **8**

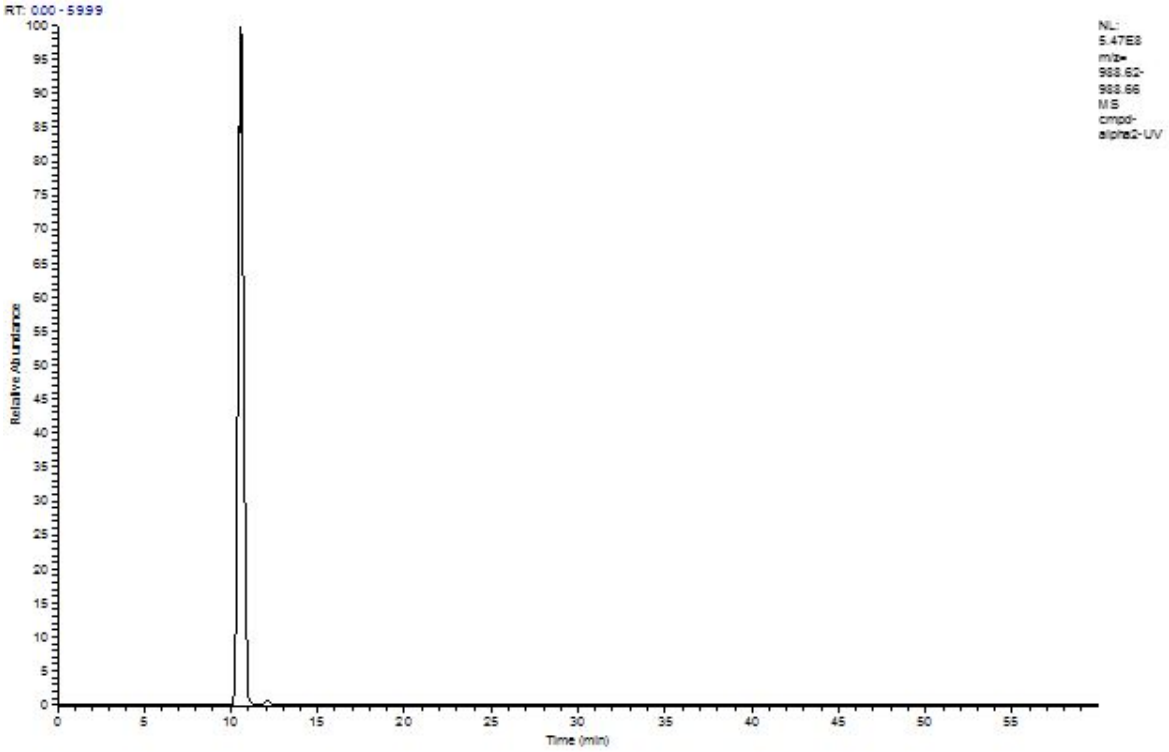


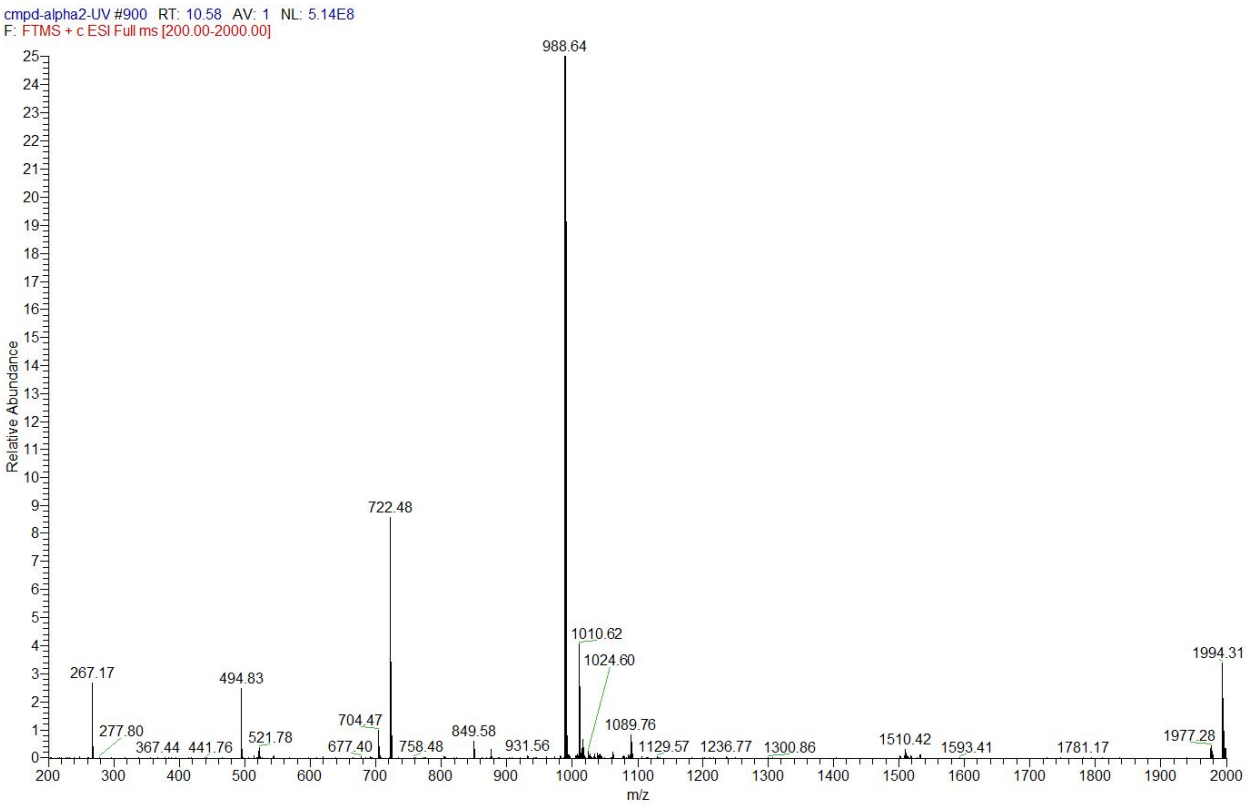


LCMS data for **9**

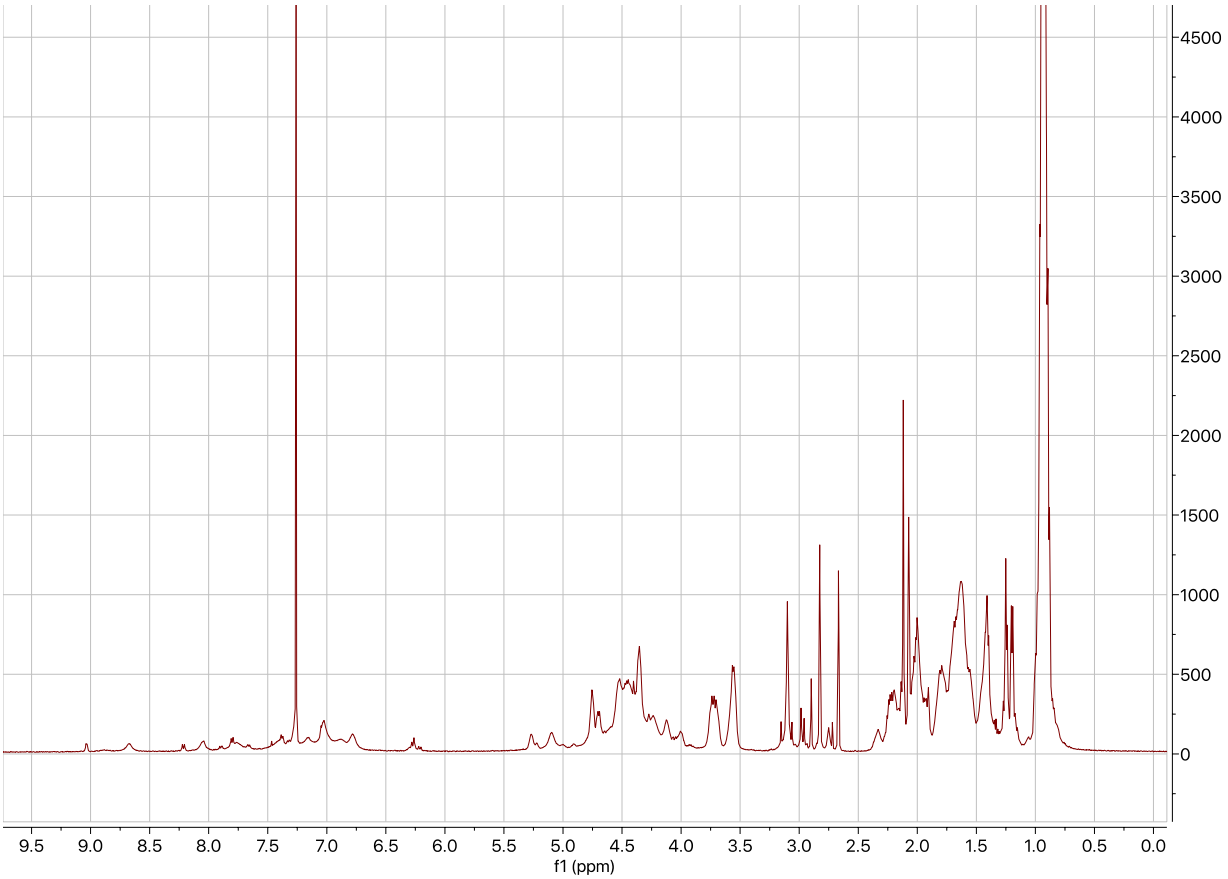


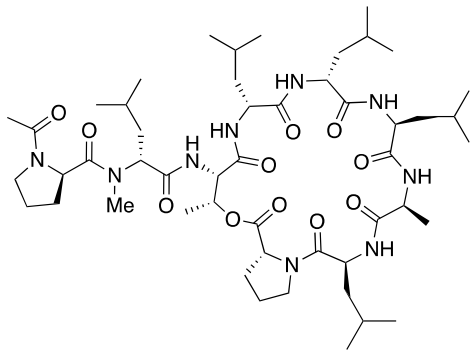

RT: 0.00 - 59.99
NL:
1.31E9
TIC F: FTMS +
c ESI Full ms
[200.00-
2000.00] MS
cmpd-alpha2-
UV
Relative Abundance
Time (min)


RT: 0.00 - 59.99
NL:
5.47E8
m/z=
988.62-
988.66
MS
cmpd-
alpha2-UV
Relative Abundance
Time (min)

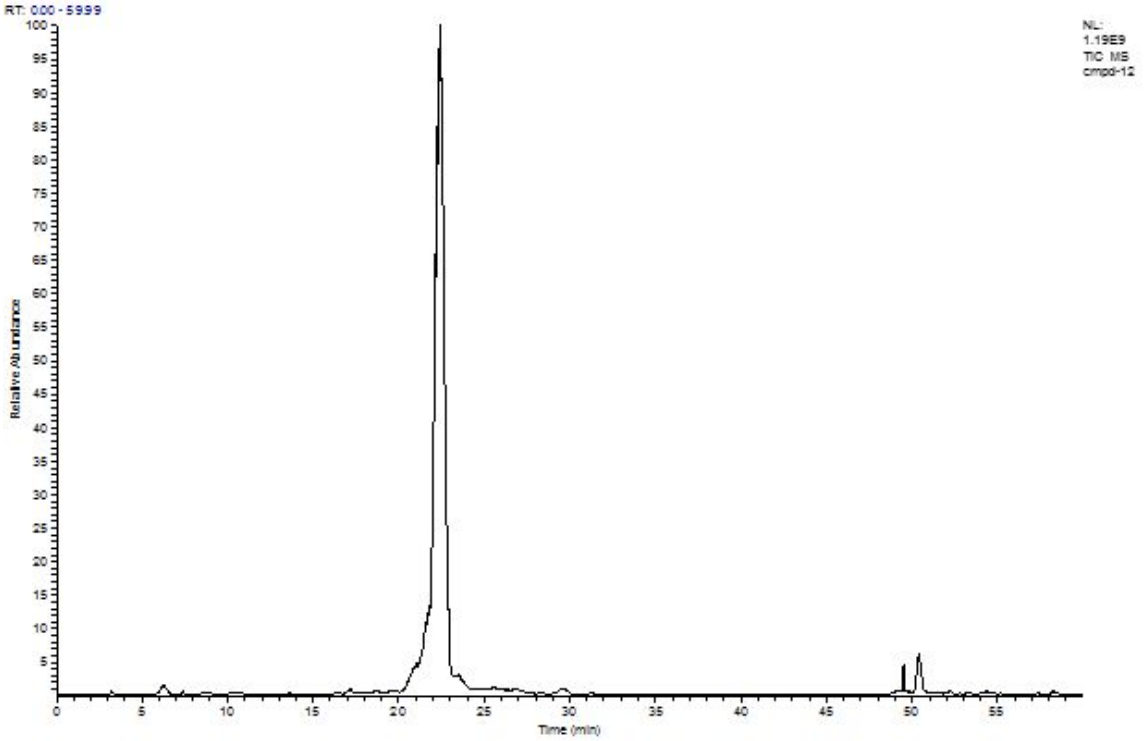

cmpd-alpha2-UV #900 RT: 10.58 AV: 1 NL: 5.14E8
F: FTMS + c ESI Full ms [200.00-2000.00]
Relative Abundance
m/z
988.64
722.48
1010.62
1024.60
1994.31
267.17
494.83
704.47
1089.76
849.58
277.80
367.44
441.76
521.78
677.40
758.48
931.56
1129.57
1236.77
1300.86
1510.42
1593.41
1781.17
1977.28

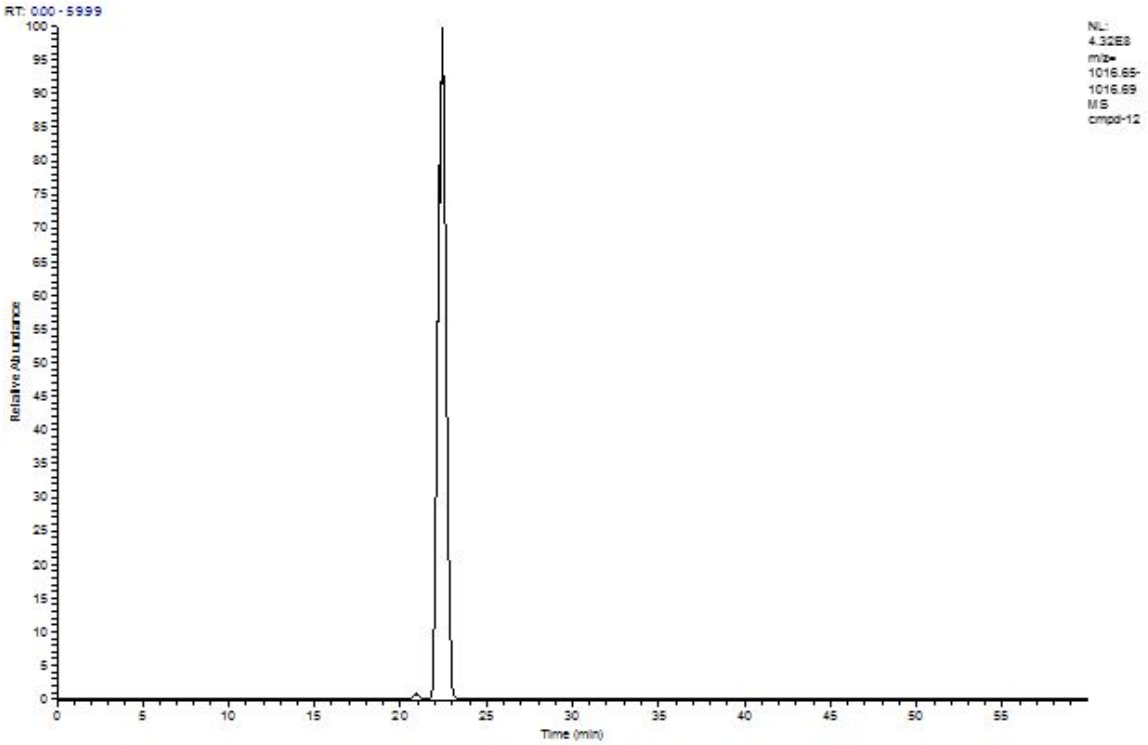
H1 NMR spectrum of **9**

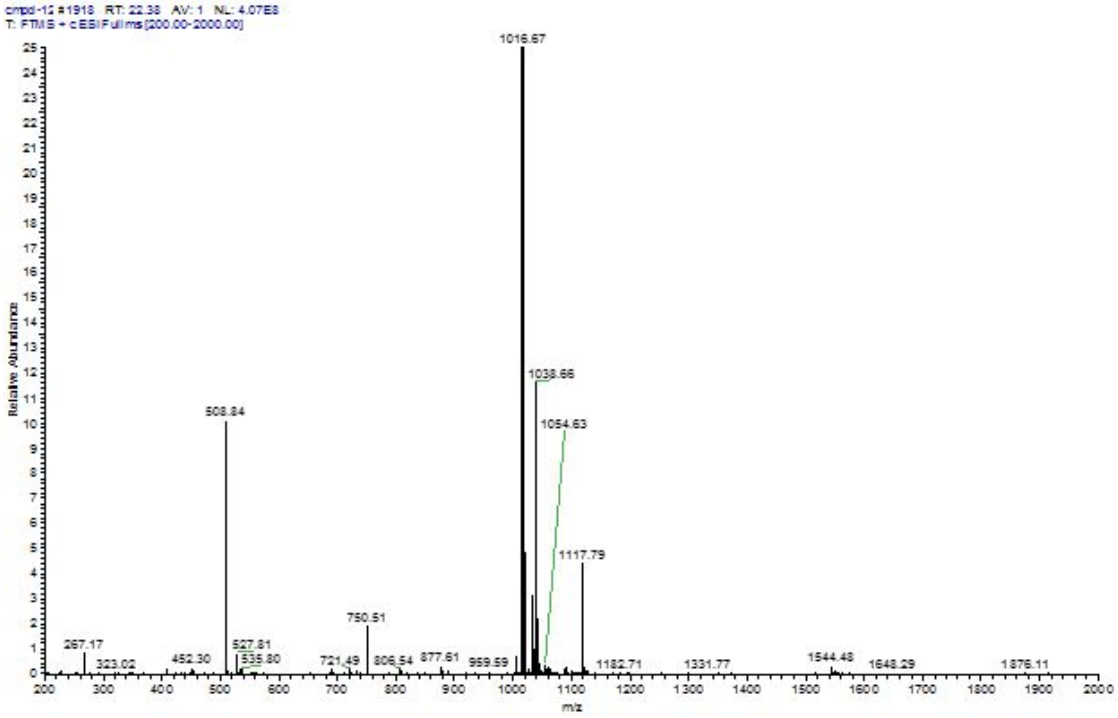




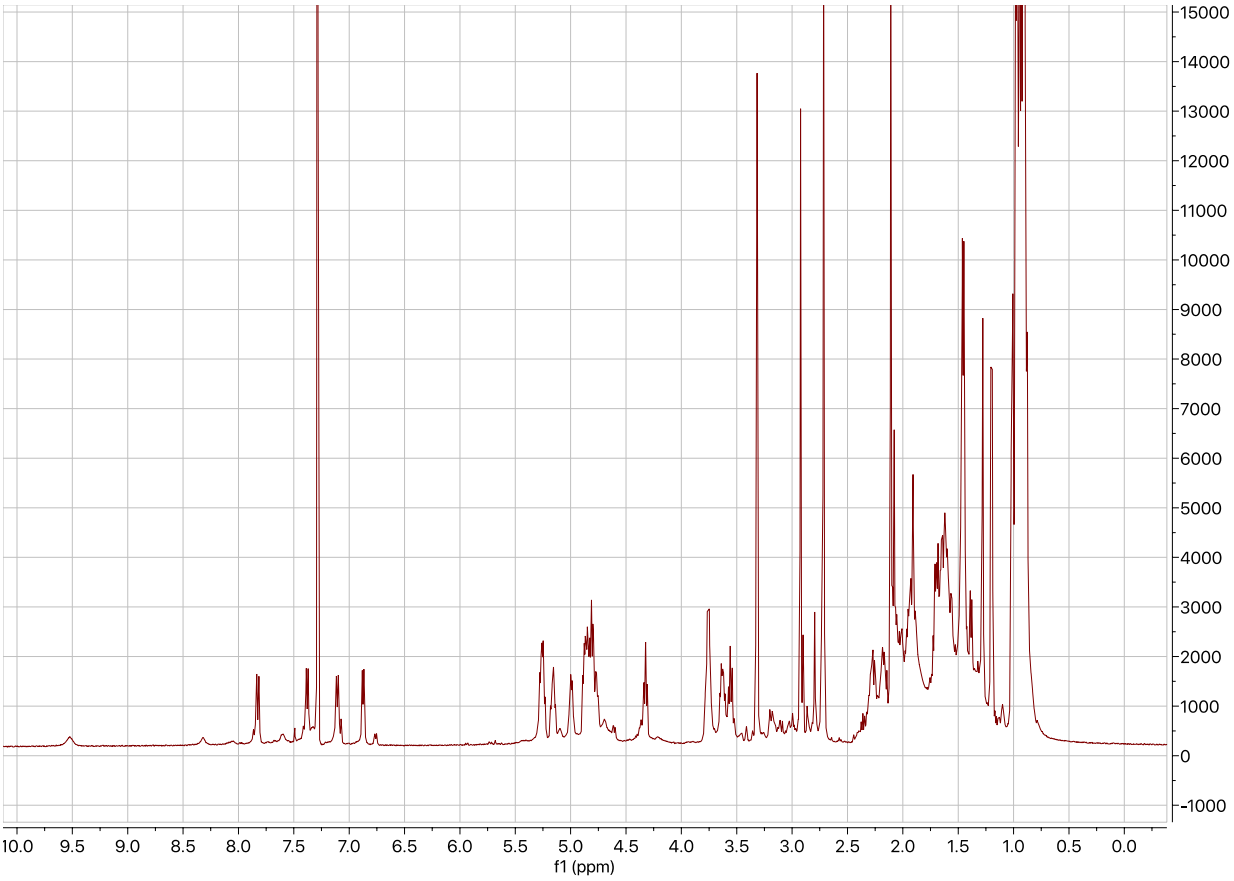
LCMS data for **10**

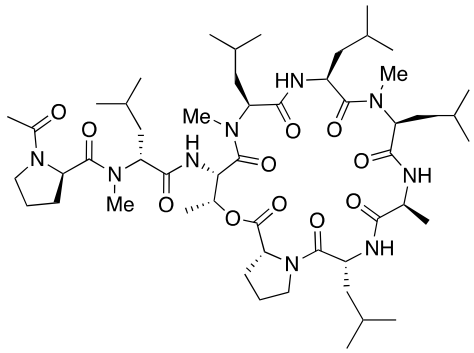




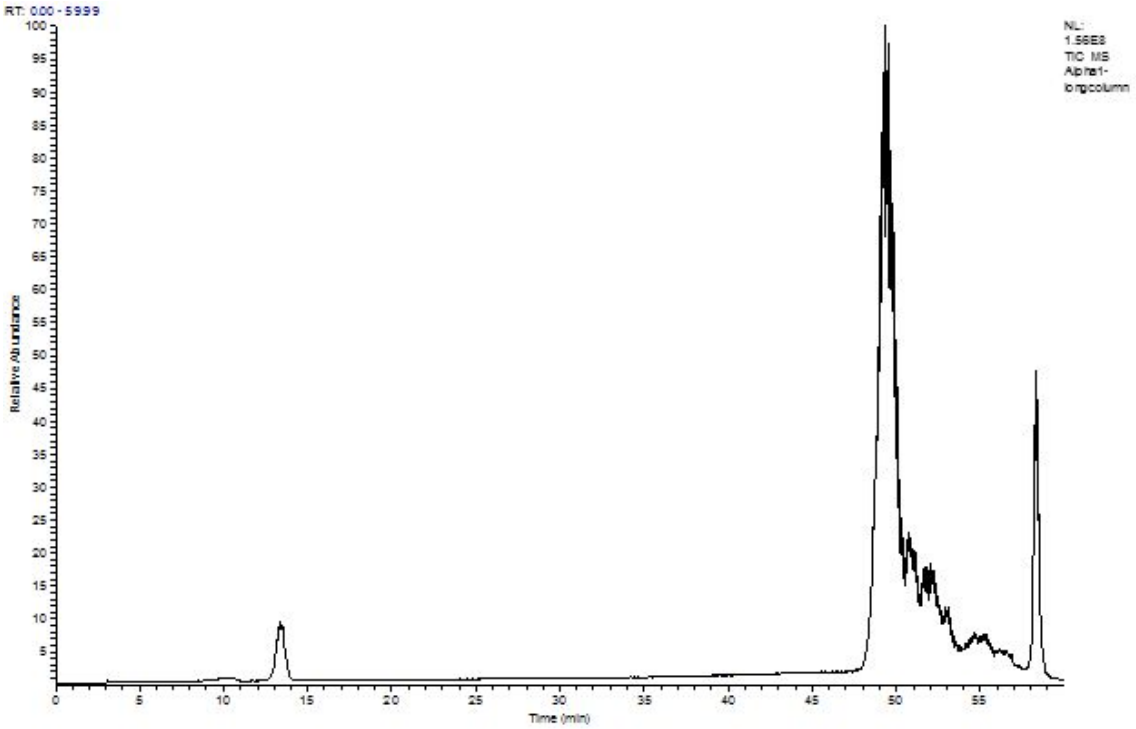

cmpd-12 #1918 RT: 22.38 AV: 1 NL: 4.07E8
T: FTMS + cESI Full ms [200.00-2000.00]
Relative Abundance
m/z
1016.67
1038.66
1054.63
1117.79
508.84
750.51
267.17
323.02
452.30
527.81
535.80
721.49
806.54
877.61
959.59
1182.71
1331.77
1544.48
1648.29
1876.11

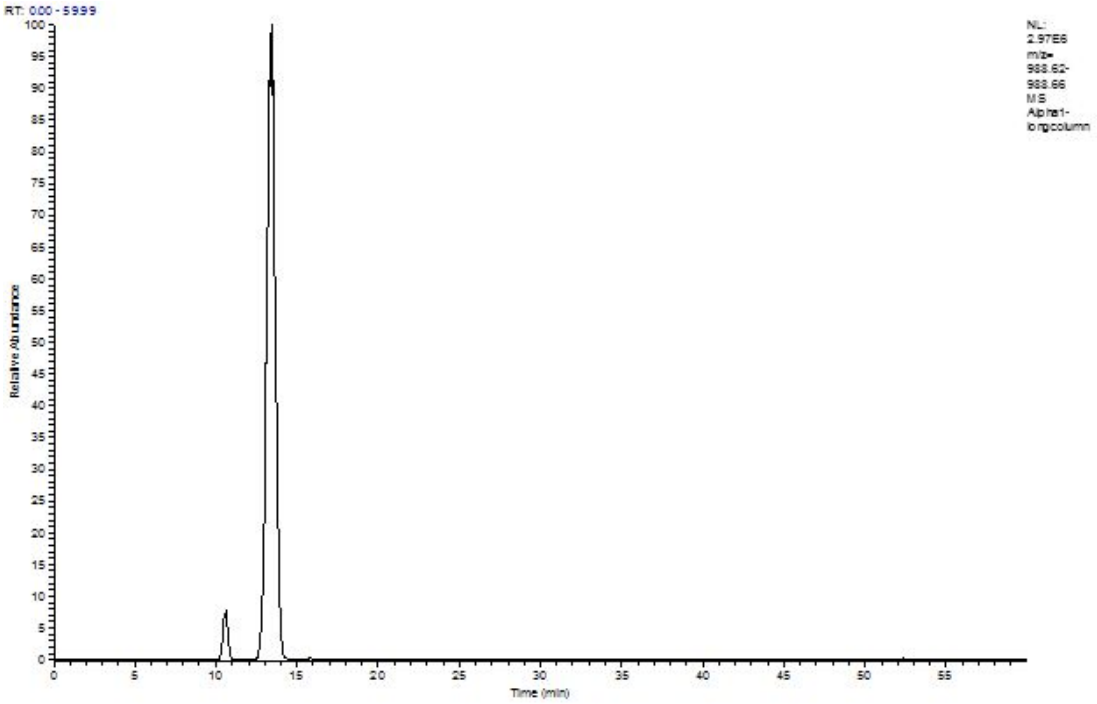
H1 NMR spectrum of **10**

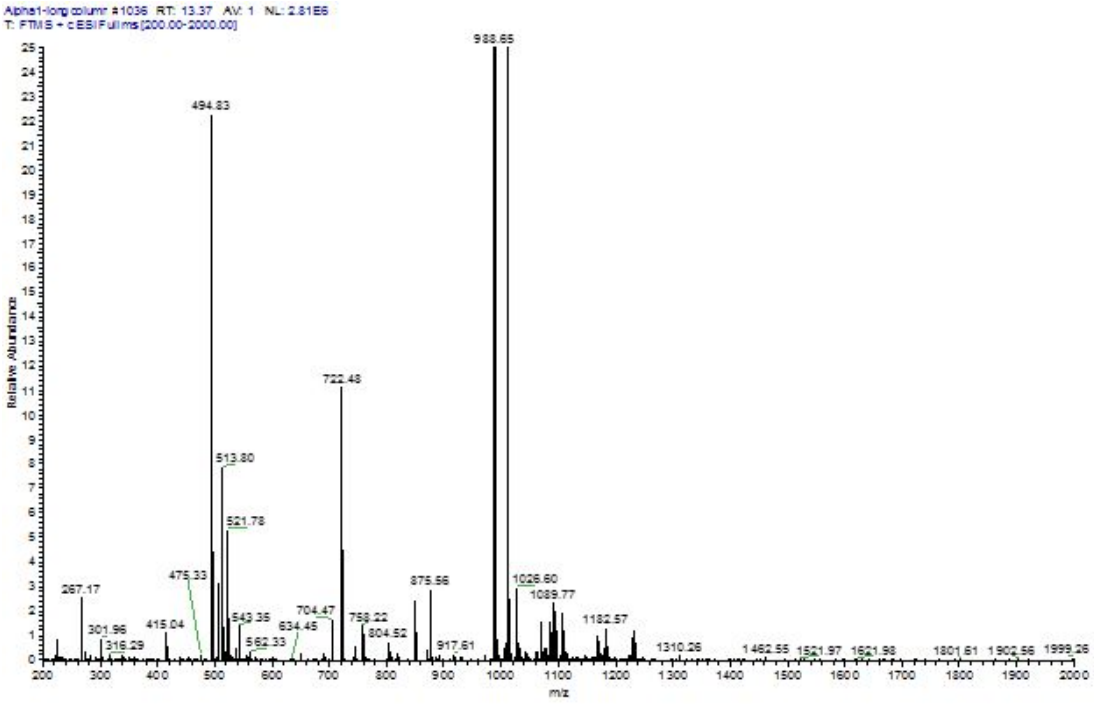




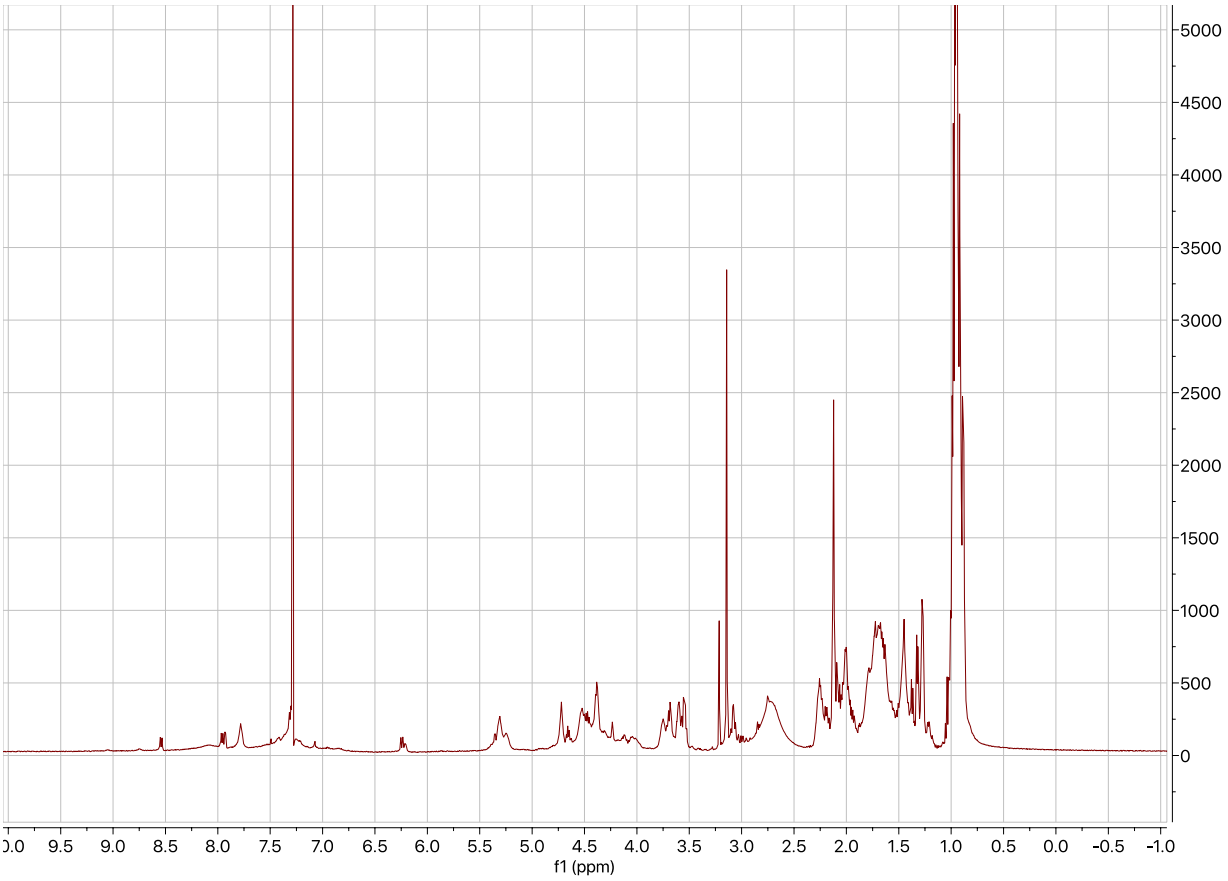
LCMS data for **11**

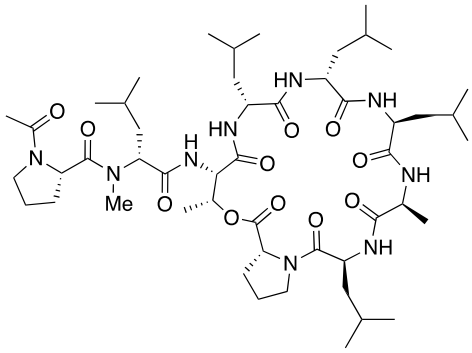





Alpha1-long column #1036 RT: 13.37 AV: 1 NL: 2.81E6
T: FTMS + cESI Full ms [200.00-2000.00]
Relative Abundance
m/z
988.65
494.83
722.48
513.80
521.78
475.33
875.56
1026.60
1089.77
1182.57
267.17
301.96
316.29
415.04
543.35
562.33
634.45
704.47
758.22
804.52
917.61
1310.26
1462.55
1521.97
1621.98
1801.61
1902.56
1999.26

H1 NMR spectrum of **11**





References

1. Sigler, G. F.; Fuller, W. D.; Chaturvedi, C., N., Formation of oligopeptides during the synthesis of 9-fluorenylmethyloxycarbonyl amino acid derivatives. *Biopolymers* **1983,** *22*, 2157-2162.
2. Freidinger, R. M.; Hinkle, J. S.; Perlow, D. S., Synthesis of 9-fluorenylmethyloxycarbonyl-protecte N-alkyl amino acids by reduction of oxazolidinones. *J Org Chem* **1983,** *48*, 77-81.
3. Townsend, C.; Furukawa, A.; Schwochert, J.; Pye, C. R.; Edmondson, Q.; Lokey, R. S., CycLS: Accurate, whole-library sequencing of cyclic peptides using tandem mass spectrometry. *Bioorg Med Chem* **2018,** *26* (6), 1232-1238.
4. Naylor, M. R.; Ly, A. M.; Handford, M. J.; Ramos, D. P.; Pye, C. R.; Furukawa, A.; Klein, V. G.; Noland, R. P.; Edmondson, Q.; Turmon, A. C.; Hewitt, W. M.; Schwochert, J.; Townsend, C. E.; Kelly, C. N.; Blanco, M. J.; Lokey, R. S., Lipophilic Permeability Efficiency Reconciles the Opposing Roles of Lipophilicity in Membrane Permeability and Aqueous Solubility. *J Med Chem* **2018,** *61* (24), 11169-11182.
5. White, T. R.; Renzelman, C. M.; Rand, A. C.; Rezai, T.; McEwen, C. M.; Gelev, V. M.; Turner, R. A.; Linington, R. G.; Leung, S. S.; Kalgutkar, A. S.; Bauman, J. N.; Zhang, Y.; Liras, S.; Price, D. A.; Mathiowetz, A. M.; Jacobson, M. P.; Lokey, R. S., On-resin N-methylation of cyclic peptides for discovery of orally bioavailable scaffolds. *Nat Chem Biol* **2011,** *7* (11), 810-7.
6. Furukawa, A.; Schwochert, J.; Pye, C. R.; Asano, D.; Edmondson, Q. D.; Turmon, A.; Klein, V.; Ono, S.; Okada, O.; Lokey, R. S., Drug-like properties in macrocycles above MW 1000: Backbone rigidity vs. side-chain lipophilicity. *Angew Chem Int Ed* **2020**.
7. Jain, A. N.; Cleves, A. E.; Gao, Q.; Wang, X.; Liu, Y.; Sherer, E. C.; Reibarkh, M. Y., Complex macrocycle exploration: parallel, heuristic, and constraint-based conformer generation using ForceGen. *J Comput Aided Mol Des* **2019,** *33* (6), 531-558.