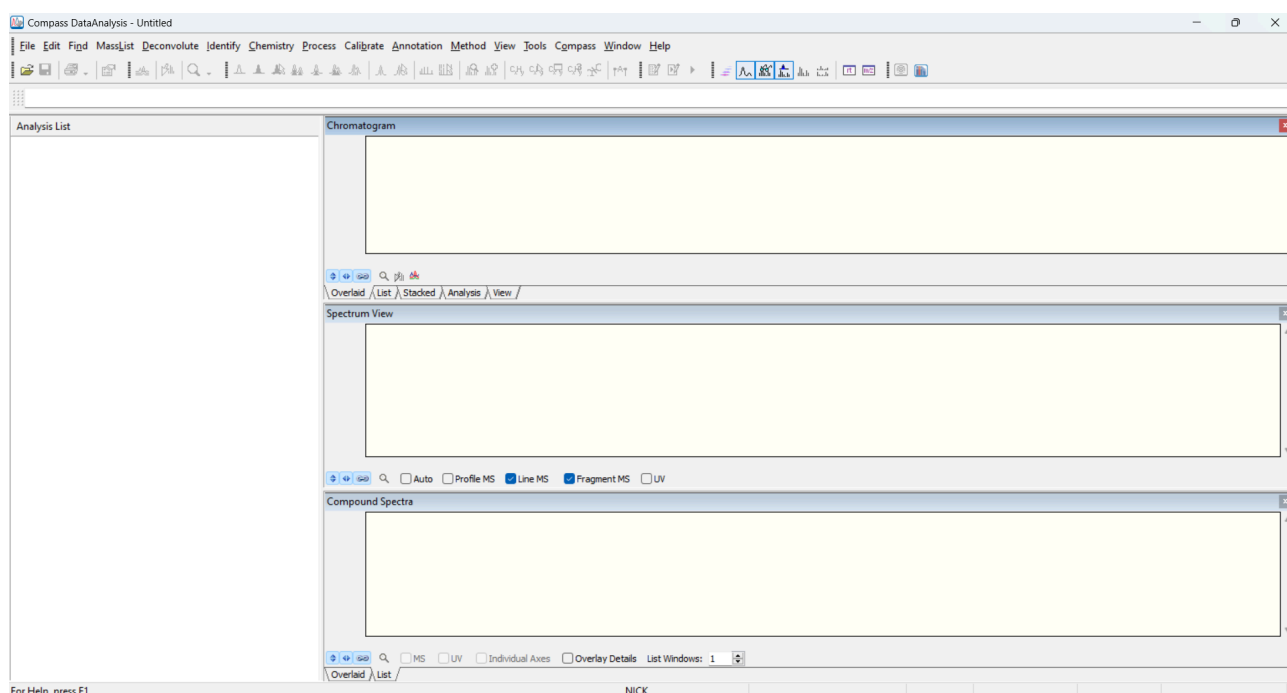


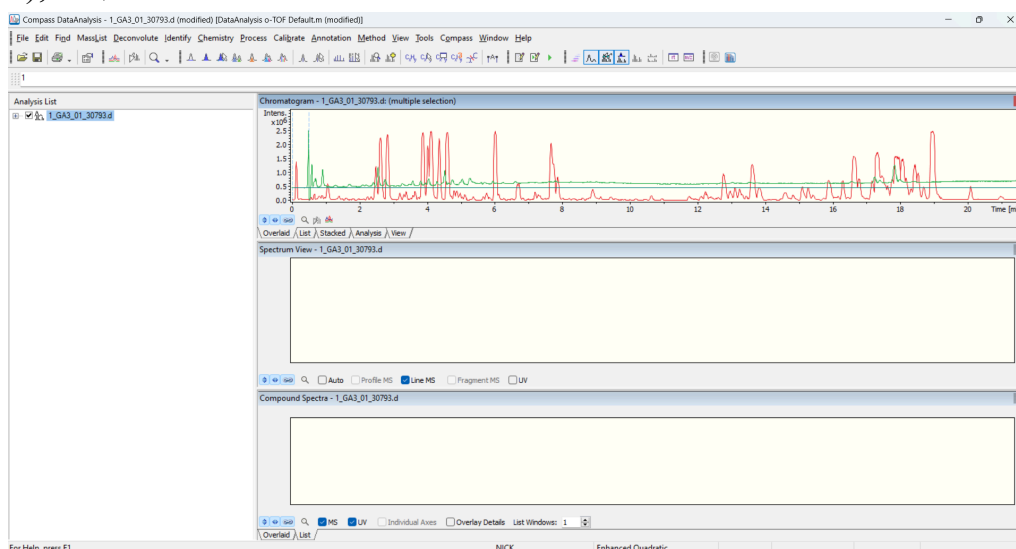
Основи методики ручної дереплікації ESIMS хромато-мас спектрів

Дереплікація навіть електроспрей-іонізаційних спектрів — це ручна праця, що вимагає досить великих затрат часу і людських ресурсів. Втім, вона є необхідною і дозволяє отримати велику кількість даних з одного мас-спектру. Для її виконання потрібне специфічне програмне забезпечення, та, як найголовніше, безпосередньо мас-спектрометричні дані.

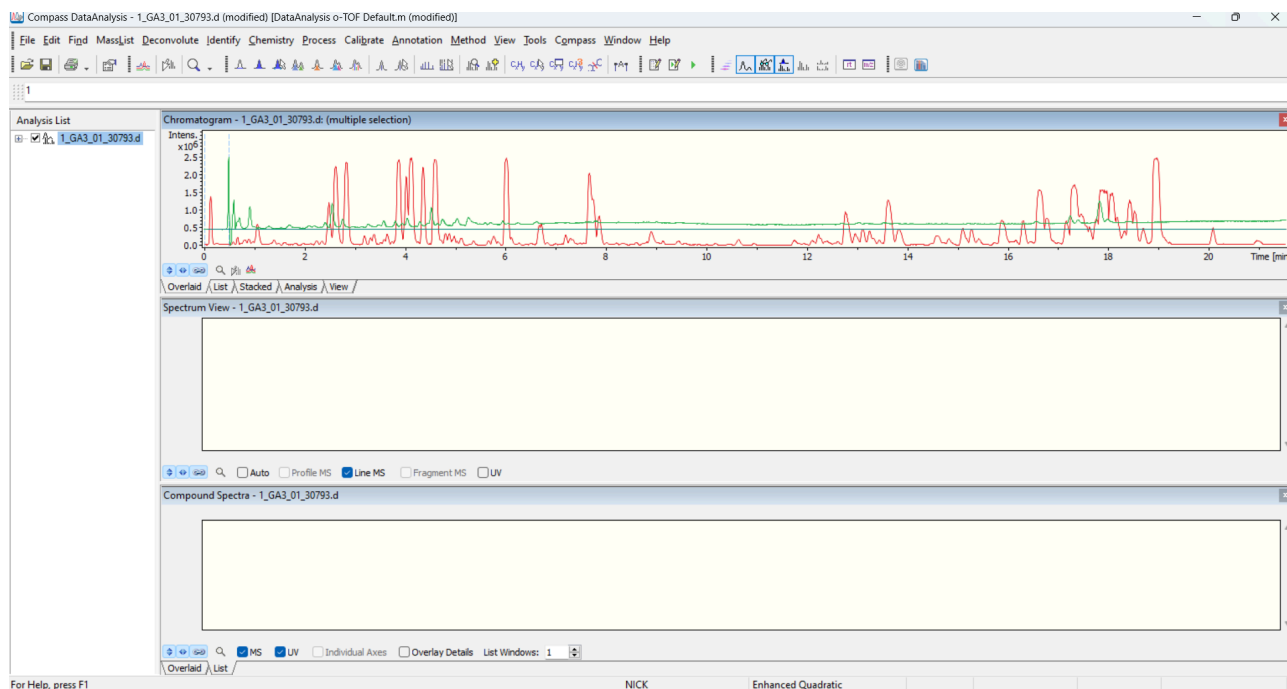
Нижче буде продемонстровано процес дереплікації на прикладі реальних наукових даних.



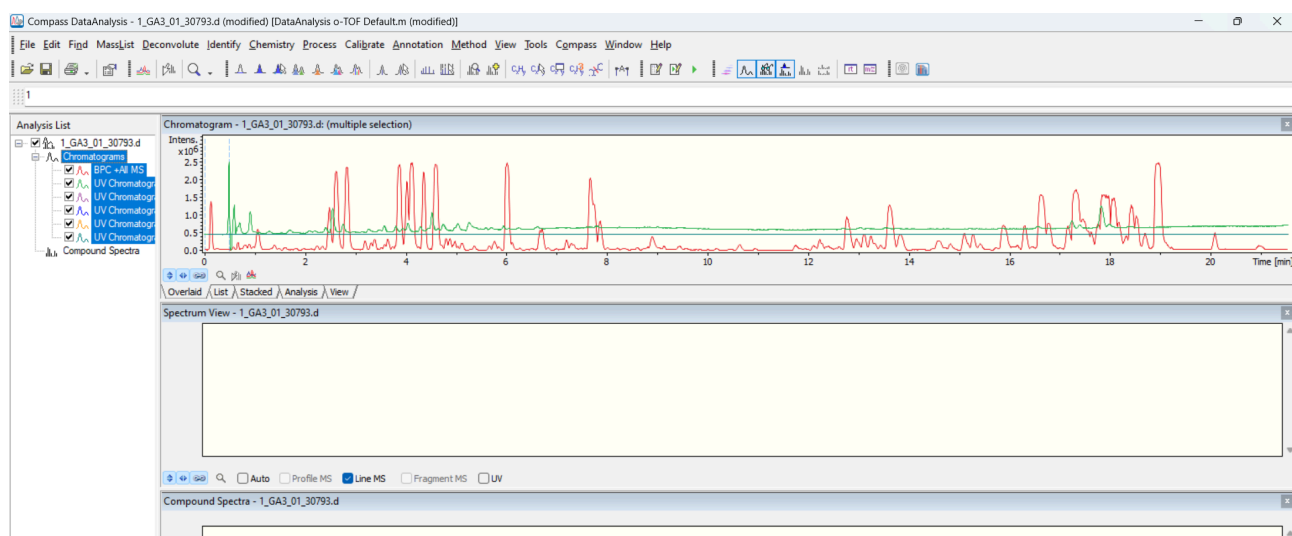
Вигляд вікна програми одразу після запуску. Віконце Compound Spectra можна відразу закрити (натиснувши в ньому на хрестик угорі праворуч). Потім можна відкрити файл спектра (через меню File вгорі ліворуч). Вибираємо спектр з папки (зауважимо, що операційна система сприймає наш спектр як звичайну папку з файлами і єдиним цілим він є лише для нашої програми), і що бачимо:



Знову закриваємо Compound Spectra (вибачте, я його забув закрити і на кількох наступних скріншотах він ще буде). Можна схопити мишкою кордон робочої області в центрі і перетягнути вліво, щоб не муляв очі білий простір.



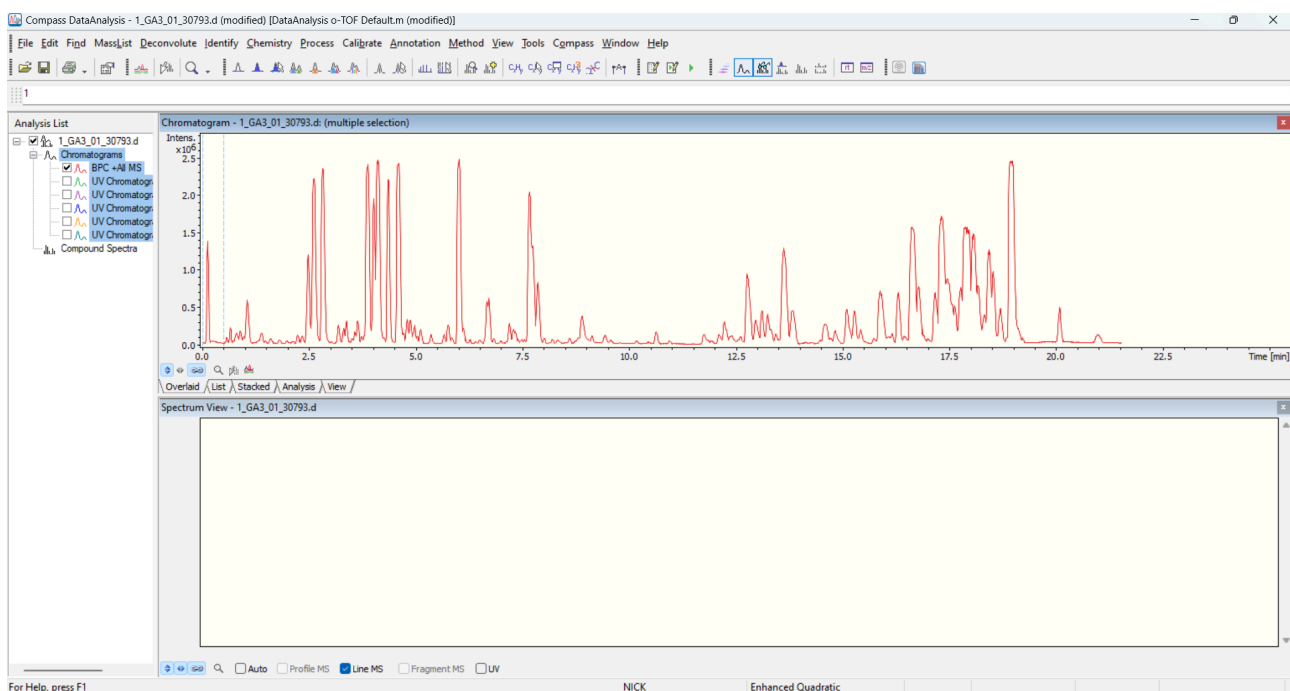
Червоний графік - це інтенсивність мас-спектрометричного сигналу (по суті, щільність потоку іонів, яка приблизно еквівалентна кількості речовини, що виходить із колонки хроматографа в кожен момент). Кольорові графіки - поглинання в різних діапазонах світла. Їх можна приховати, оскільки для роботи вони не критичні і на ранньому етапі можуть скоріше заважати. Клікаємо на хрестик зліва, потім ще на один, що випаде...



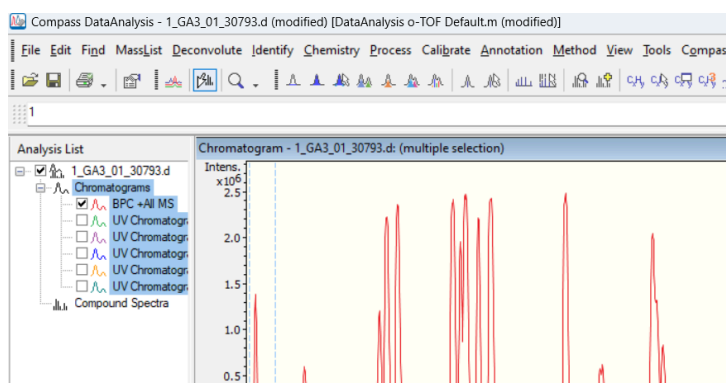
Прибираємо галочки на всьому, крім червоного графіка. Отримуємо красу. Під червоним графіком видно шкалу - це час виходу піків у хвилинах.

Для роботи зі спектром потрібно активувати курсор спектра (стрілочка вліво вниз на панелі інструментів).

В вікна Chromatogram за допомогою вибраної команди Chromatogram Compaunds, програма обере піки які вважає головними. Програма видає найбільш доречні піки, які ми надалі розглядали та досліджували. Але було прийнято рішення спиратись не тільки на вибір програми, але й перевіряти не обрані програмою піки. Запис піків вели починаючи з другої хвилини і закінчували двадцятою хвилиною. В вікні Chromatogram натискали на вершину піку, а в вікні Spectrum View висвічувались маси піків. Обирали масу сполуки, маса якої було найбільше (її «під» максимальний). Маси головних піків (M+N) виписувались в табличку програми Excel, в колонку M+N; шляхом арифметичного віднімання від цього значення відносної атомної маси протону (1,007276) обраховували значення молекулярної маси сполуки в колонці M.



Тепер клікніть на іконку зі стрілочкою на панелі інструментів. Після цього вона виділяється блакитним.



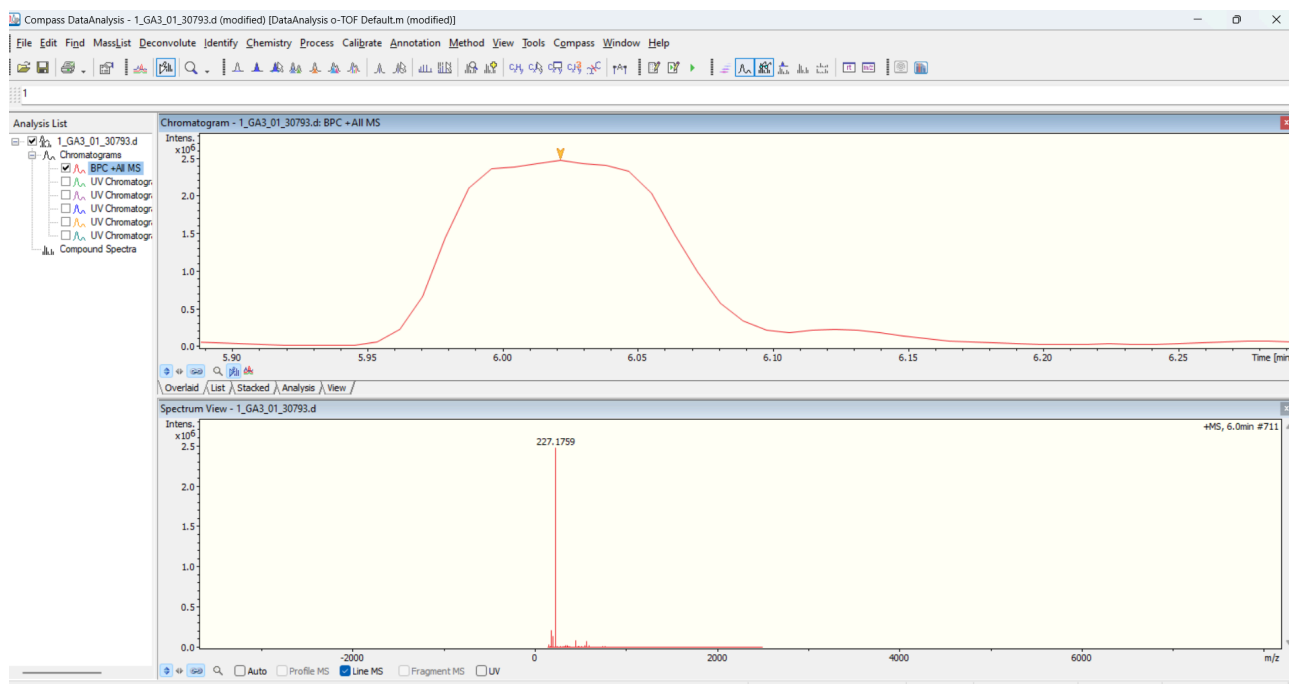
Тепер
піки на графіку.

можна вибрати
Наведіть мишку

на графік і клікніть. У вікні Chromatogram на графіку з'явиться жовта стрілка - ваш курсор, встановлений у певній точці. А в Spectrum View - мас-спектр цього піка.

До речі, ви можете змінювати масштаб обох графіків, підводячи курсор миші ПІД шкалу і крутячи коліщатко.

Ось тут я зменшив масштаб обраного піка на верхньому графіку і збільшив на нижньому. Це звичайна тактика - щоб повністю бачити мас-спектр у цей момент і знати, в якій точці хроматограми він узятий.



Тут у нас домінує пік з m/z 227.1759. У нього час виходу (Retention Time), як видно зверху, 6.02 (6 хвилин і 2 сотих хвилини). Записуємо!

Без назви 1 — LibreOffice Calc

	A	B	C	D	E	F
1	RT		M+N	Measured mass (M — H+)		
2	6,02		227,1759	226,168624		
3						
4						
5						
6						
7						
8						

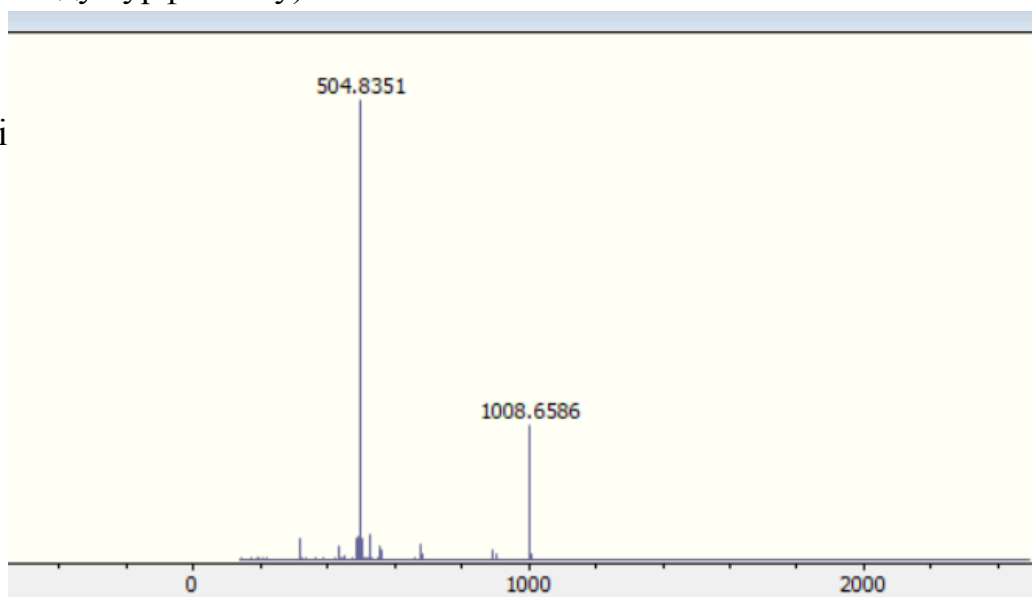
Обирали масу сполуки, маса якої було найбільше (її «пік»

максимальний). Маса головних піків (M+N) виписувались в табличку програми Excel, в колонку M+N; шляхом Арифметичного віднімання від цього значення відносної атомної маси протону (1,007276) обраховували значення молекулярної маси сполуки в колонці M. На думку автора, зручніше всього спочатку переписати всі базові піки, а потім переходити до наступного етапу — пошуку адуктів. Важливо, що ми враховуємо лише піки в діапазоні RT від 2 до 20 хвилин — це пов'язано з особливостями процесу хроматографії.

Варто враховувати, що в спектрах пептидів M+N практично ніколи не буває домінуючим. В бацилярних ліпопептидів, як правило, базовим є пік M+2H, а M+N слідує за ним по числу каунтів. Відповідно, спостерігається

характерна картина, як на ілюстрації нижче (ESIMS-спектр аналогу гептапептиду сурфактину):

в вікні



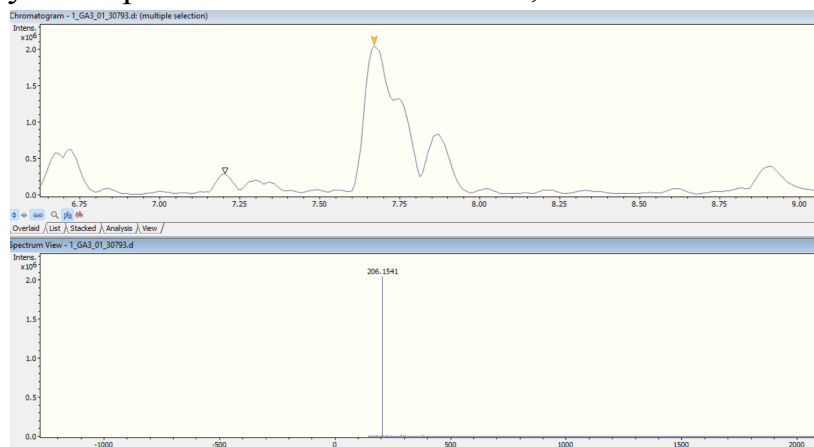
Також

Spectrum View додатково перевірялись ion mode.

В positive mode перевірялися адукти, в порядку зниження пріоритету: $M+Na$, $M+NH_4$, $M+K$, $M+ACN+K$, $2M+H$, $M+ACN+Na$, $2M+NH_4$, $2M+Na$, $2M+3H_2O+2H$ та $2M+K$.

Ці дані теж вносились до таблиці. Враховували те, що відхилення від обрахованого значення m/z не має перевищувати $\pm 2,5$ ppm.

В нашому спектрі виявлено йон з m/z 206,1541.



перевіряти

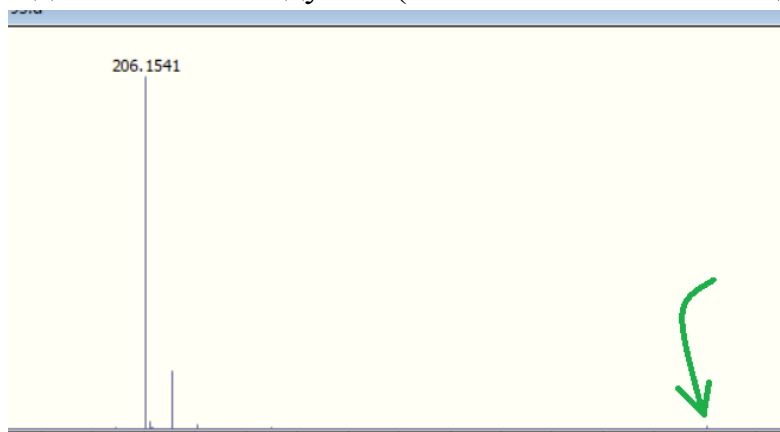
Починаємо

гіпотезу про

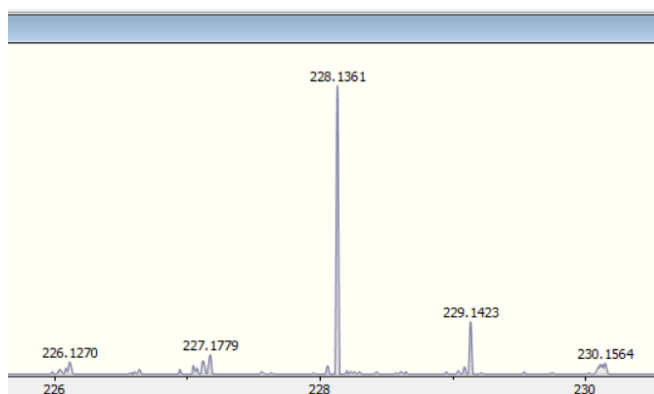
те, що перед нами — $M+H$. Обраховуємо гіпотетичну молекулярну масу сполуки: $206,1541 - 1,0073 = 205,1468$. Обраховуємо m/z найбільш частотного адукту (завдячуємо натрію в лабораторному посуді...) за допомогою таблицки Huang et al, 1999:

Table 1. Monoisotopic exact masses of molecular ion adducts often observed in ESI mass spectra					Your M here:	Your M+X or M-X
					205,146824	413,2659
Ion name	Ion mass	Charge	Mult	Mass	Result:	Reverse:
1. Positive ion mode						
M+3H	M/3 + 1.007276	3+	0,33	1,007276	69,389551	136,748024
M+2H+Na	M/3 + 8.334590	3+	0,33	8,334590	76,716865	129,420710
M+H+2Na	M/3 + 15.7661904	3+	0,33	15,766190	84,148465	121,989110
M+3Na	M/3 + 22.989218	3+	0,33	22,989218	91,371493	114,766082
M+2H	M/2 + 1.007276	2+	0,5	1,007276	103,580688	205,625674
M+H+NH ₄	M/2 + 9.520550	2+	0,5	9,520550	112,093962	197,112400
M+H+Na	M/2 + 11.998247	2+	0,5	11,998247	114,571659	194,634703
M+H+K	M/2 + 19.985217	2+	0,5	19,985217	122,558629	186,647733
M+ACN+2H	M/2 + 21.520550	2+	0,5	21,520550	124,093962	185,112400
M+2Na	M/2 + 22.989218	2+	0,5	22,989218	125,562630	183,643732
M+2ACN+2H	M/2 + 42.033823	2+	0,5	42,033823	144,607235	164,599127
M+3ACN+2H	M/2 + 62.547097	2+	0,5	62,547097	165,120509	144,085853
M+H	M + 1.007276	1+	1	1,007276	206,154100	412,258624
M+NH ₄	M + 18.033823	1+	1	18,033823	223,180647	395,232077
M+Na	M + 22.989218	1+	1	22,989218	228,136042	390,276682
M+CH ₃ OH+H	M + 33.033489	1+	1	33,033489	238,180316	380,232411
M+K	M + 38.963158	1+	1	38,963158	244,109982	374,302742
M+ACN+H	M + 42.033823	1+	1	42,033823	247,180647	371,232077
M+2Na-H	M + 44.971160	1+	1	44,971160	250,117984	368,294740
M+IsoProp+H	M + 61.06534	1+	1	61,065340	266,212164	352,200560
M+ACN+Na	M + 64.015765	1+	1	64,015765	269,162589	349,250135
M+2K+H	M + 76.919040	1+	1	76,919040	282,065864	336,346860

Шукаємо даний адукт в тому самому локальному спектрі. І знаходимо! Його видно навіть у великому масштабі відображення спектру. Втім, таке характерне лише для частотних адуктів (з часткою більше 1% від бази).



Зменшуємо масштаб, аби визначити точне значення m/z даного йону і порівняти його з розрахованим вище.



Бачимо, що 228,1361 відрізняється від еталонного 228,1360 лише на +0,25 ppm. Адукт M+Na знайдено; гіпотезу про те, що базовий пік є M+H, підтверджено. Подальший пошук адуктів сенсу не має. Якщо б це адукт не був знайдений, то перевірка йшла б у зазначеній вище послідовності.

В negative mode перевірялись M+FA-H, M-3H, M-2H, M-H₂O-H, M+Na-H, M+Cl, M+K-2H, M+Hac-H, 2M-H, 2M+FA-H, 2M+Hac-H, 3M-H та найірідкісніший M+Br (адуктанти походять або з компонентів елюенту, або вилуговуються зі посуду, як лужні метали; бром у таких умовах майже нізвідки взятись). В негативній формі, на жаль, в кожному з головних піків не кожен додатковий пік було знайдено-виписано до таблиці через його відсутність.

Наступний крок виконували з використанням програми *Dictionary of Natural Products*. Використовували як програму 2019 року так і інтернет версію, де база даних є більш розширеною. З допомогою цієї форми знаходились речовини причетні до кожного з виписаного піку досліджуваного спектру. Пошук в програмі по кожному піку брали діапазон з відхиленням в 3 ppm. Не кожен пік був знайдений через відсутність його в базі даних або ж через деградацію відповідної сполуки в процесі іонізації.

Задля демонстрації процесу ідентифікації сполуки візьмемо для прикладу йон, адукт для якого було виявлено вище. Розрахуємо діапазон мас для пошуку: $X \pm 3 \text{ ppm}$, де X — це розрахована відносна молекулярна маса нашої сполуки.

Нижня межа: $205,1468 - ((205,1468/1000000)*3) = 205,1462$

Верхня межа: $205,1468 + ((205,1468/1000000)*3) = 205,1474$

Тепер ці дані готові для передачі до бази даних, в даному випадку - *Dictionary of Natural Products*. Дана база підтримує різноманітні варіанти пошуку, нас цікавить дуже конкретний — Accurate Mass, Range.

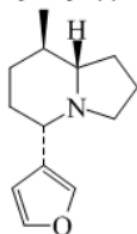
В поля вносимо значення верхньої та нижньої межі і натискаємо Search. Увага! DNP НЕ підтримує десяткові частки, розділені комою. Отриманий список хітів вручну фільтруємо на предмет наявності бациллярних сполук.

Hit List		History
		Hit 1 of 8
☐ ...	Chemical Name	Molecular Formula
☐ 1	5-(3-Furanyl)octahydro-8-methylir	C ₁₃ H ₁₉ NO
☐ 2	5-(3-Furanyl)octahydro-8-methylir (5S,8R,8aS)-form	C ₁₃ H ₁₉ NO
☐ 3	2-(2-Hydroxy-2-phenylethyl)-1-me	C ₁₃ H ₁₉ NO
☐ 4	2-(2-Hydroxy-2-phenylethyl)-1-me (2R*,2'R*)-form	C ₁₃ H ₁₉ NO
☐ 5	2-(2-Hydroxy-2-phenylethyl)-1-me (2R,2'R)-form	C ₁₃ H ₁₉ NO
☐ 6	2-(2-Hydroxy-2-phenylethyl)-1-me (2R,2'S)-form	C ₁₃ H ₁₉ NO
☐ 7	2-(2-Hydroxy-2-phenylethyl)-1-me (2R*,2'S*)-form	C ₁₃ H ₁₉ NO

Entry Name: 5-(3-Furanyl)octahydro-8-methylindolizine, CAS



Synonym(s): 5-(3-Furanyl)-8-methylindolizidine.



CRC Number: GPM70-P

Molecular Formula: C₁₃H₁₉NO

Molecular Weight: 205.301

Accurate Mass: 205.146664

Percentage Composition: C 76.06%; H 9.33%; N 6.82%; O 7.79%

General Statement: Closely related to the *Nuphar* alkaloids

InChi Key: FYEDVVFLFWFKNZ-UHFFFAOYSA-N

InChi: InChI=1S/C13H19NO/c1-10-4-5-13(11-6-8-15-9-11)14-7-2-3-12(10)14/h6,8-10,12-13H,2-5,7H2,1H3

Smiles: CC1CCC(N2CCCC12)c1ccoc1

В таблицю вписувались хвилини (RT), назви ідентифікованих сполук, значення Accurate Mass, Measured mass ([гіпотетичне M+H] – маса протону)б похибка в ppm Measured mass відносно Accurate, характер біологічної активності сполуки (антибіотик, інгібування певних ферментів тощо), організм/и продуцент/и і джерело інформації про сполуку (DNP чи посилання на статтю). За допомогою знання про те де було знайдено речовину робили висновок - вносити чи ні речовину до таблиці, могла чи не могла дана сполука бути в досліджуваному штамі.

В результаті були отримані наступні таблиці по кожному з досліджуваних спектрів:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	RT	Name	Accurate mass	M+H	Measured mass	Error, ppm	Props	Source					
2		20 (4-Hydroxy-3-tetraprenylphenyl)acetic acid	424,297745	425,3049	424,297624	-0,285177	Constit. of	-	DNP				
3		15,53 14-Norplakortide Q	342,27701	343,2838	342,276524	-1,419905	Cytotoxic?	Isol. from the sponge Plakobolus	DNP				
4		10,12 1-Heneicosene-4,6,8,10,12,14,16-heptol	406,293055	407,3002	406,292924	-0,322427	https://archimer.ifremer.fr/doc/00314/	-	DNP				
5		11,93 1-Heptacosene-4,6,8,10,12,14,16,18,24-nonol	522,376785	523,3839	522,376624	-0,808207	-	-	DNP				
6		7,67 1,5-Dihydro-5-methyl-1-phenazinone	210,079313	211,0864	210,079124	-0,899661	Prod. by B	Active against Gram-positiv	DNP				
7		3,35 2-Allyloxyphenol.	150,06808	151,0754	150,068124	0,2932002	found in	Streptomyces sundarbansensis	https://link.springer.com/article/10.1007/s00253-009-23				
8		14,96 3,6,9,12-Tetraoxatetracosan-1-ol Monododecyl ether	362,30	363,3104	362,303124	-0,278772							
9		3,124 4-Formylquinoline	157,052764	158,06	157,052724	-0,254692		Alkaloid prod. by Archangium	DNP				
10		3,7 5,7-Dihydroxypyrimido[5,4-e]-1,2,4-triazine	165,028675	166,0356	165,028324	-2,126908	Numerous	-	DNP				
11		4,52 5-Deoxyenterocin	428,110732	429,1177	428,110424	-0,719441	Deoxyente	algiculous strain Streptomyces	https://www.mdpi.com/1660-3397/15/12/389				
12		9,9 5-Prenyltryptophol/6-Prenyltryptophol	229,146664	230,1537	229,146424	-1,047365	Cytotoxic	Prod. by a marine Streptomyces	DNP				
13		5,01 6-Epimer, 8-deoxyenterocin	428,110735	429,1175	428,110224	-1,193618	Enterocin	Prod. by a marine Streptomyces	DNP				
14		4,28 6-N-Cyclopentyl, N9-Me adenine	217,132745	218,1397	217,132424	-1,478361							
15		6,49 Alchorneine	221,152812	222,1596	221,152324	-2,206624	Effective a	Alkaloid from Alchornea floribunda	DNP				
16		3,2 Cyclo(alanyl isoleucyl)	184,121178	185,1283	184,121024	-0,836406	Shows viru	L-L Prod. by Aspergillus phoenicis	DNP				
17		3,4 Cyclo(alanyl leucyl)	184,121178	185,1284	184,121124	-0,293285	natural product found in	Streptomyces	https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/anie.201602000				
18		3,57 Cyclo(alanyl phenylalanyl)	218,105528	219,1126	218,105324	-0,935328		L-L from Isol. from Entoloma	DNP				
19		4 Cyclo(isoleucyl prolyl)	210,136828	211,1443	210,137024	0,9327247		Isol. from the sponges Calyptopora	DNP				
20		3,64 Cyclo(isoleucyl tyrosyl)	276,147393	277,1546	276,147324	-0,249867	Inhibits bio	Constit. of Panax notoginseng	DNP				
21		3,9 Cyclo(leucyl prolyl)	210,136828	211,1444	210,137124	1,408604	Marine Ascidian-Derived	Streptomyces	https://www.mdpi.com/1660-3397/16/6/189				
22		6,53 Cyclo(phenylalanyl phenylalanyl)	294,136828	295,144	294,136724	-0,353577		L-L from Prod. by Aspergillus	DNP				
23		4,34 Cyclo(phenylalanyl prolyl)	244,121178	245,1289	244,121624	1,8269582	D-D form	P	DNP				
24		4,55 Cyclo(phenylalanyl prolyl)	244,121178	245,1289	244,121624	1,8269582							
25		4,69 Cyclo(phenylalanyl tyrosyl)	310,131742	311,1387	310,131424	-1,025372	Bacillus sp	-	https://link.springer.com/article/10.1007/s10600-009-927				
26		2,66 Cyclo(prolyl tyrosyl)	260,116093	261,1239	260,116624	2,041392	Prod. by Streptomyces sp.	GW33/1593	DNP				
27		2,5 Cyclo(prolyl tyrosyl)	260,116093	261,1236	260,116324	0,8880642	D-D from	P	DNP				
28		4,81 Cyclo(tryptophanyl dehydrobutyryne)	283,132077	284,1388	283,131524	-1,953156	Streptomyces	TDD shows weak activity against	https://www.istage.ist.go.jp/article/antibiotics1968/27/10				
29		4,46 Cyclo(tryptophanyl prolyl)	283,132077	284,1392	283,131924	-0,540384			DNP				