High resolution mass spectrometry and clouds of points

Réunion Astrostatistique - 8-9 décembre 2011 - Grenoble

R. Thissen

V. Vuitton, F.R. Orthous Daunay, J.Y. Bonnet,
G. Danger, D. Voisin, P. Pernot,

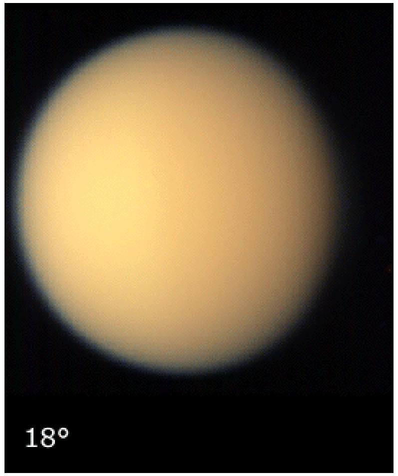
Q. Mérigot, H. Pajot

introduction

Mass spectrometry ⇔ Astrophysics

??

Titan, complex chemistry

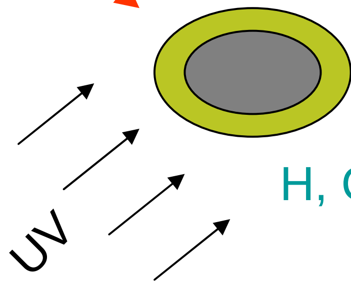


H, C, N

meteorites

H, C, N, O, S, Mg, Fe, ….

Primitive ice
Irradiation, comets

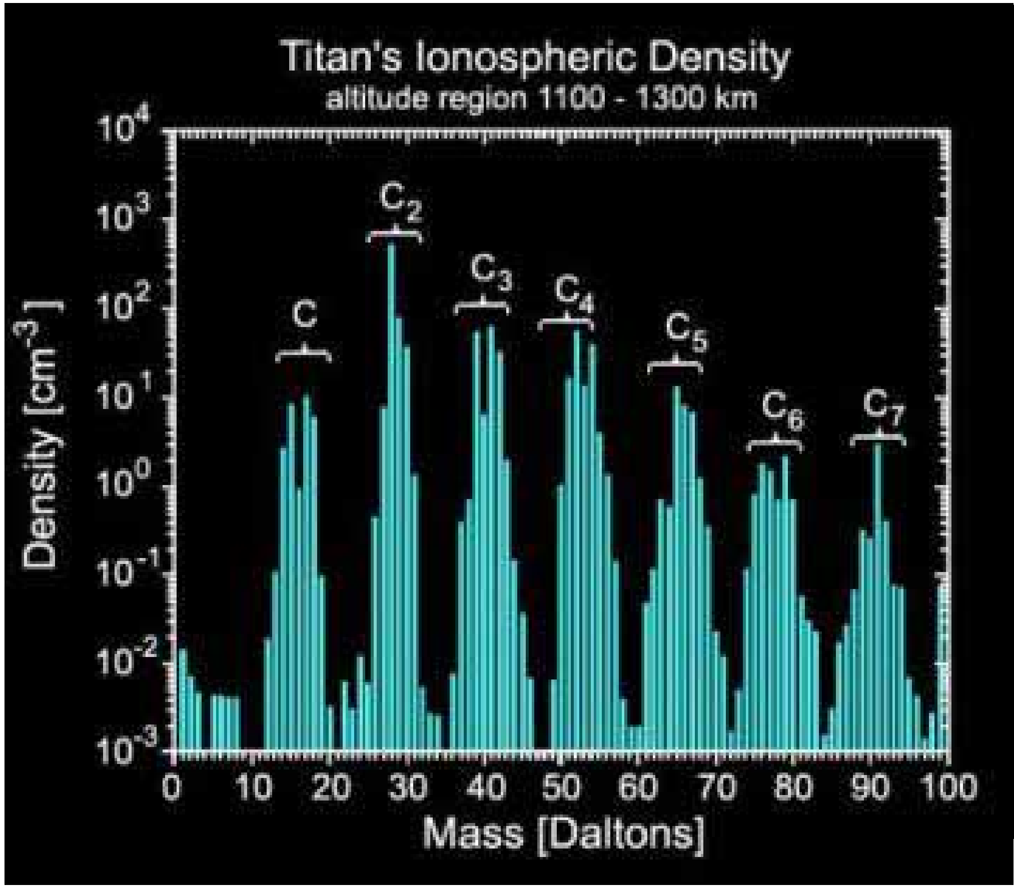


H, C, N, O

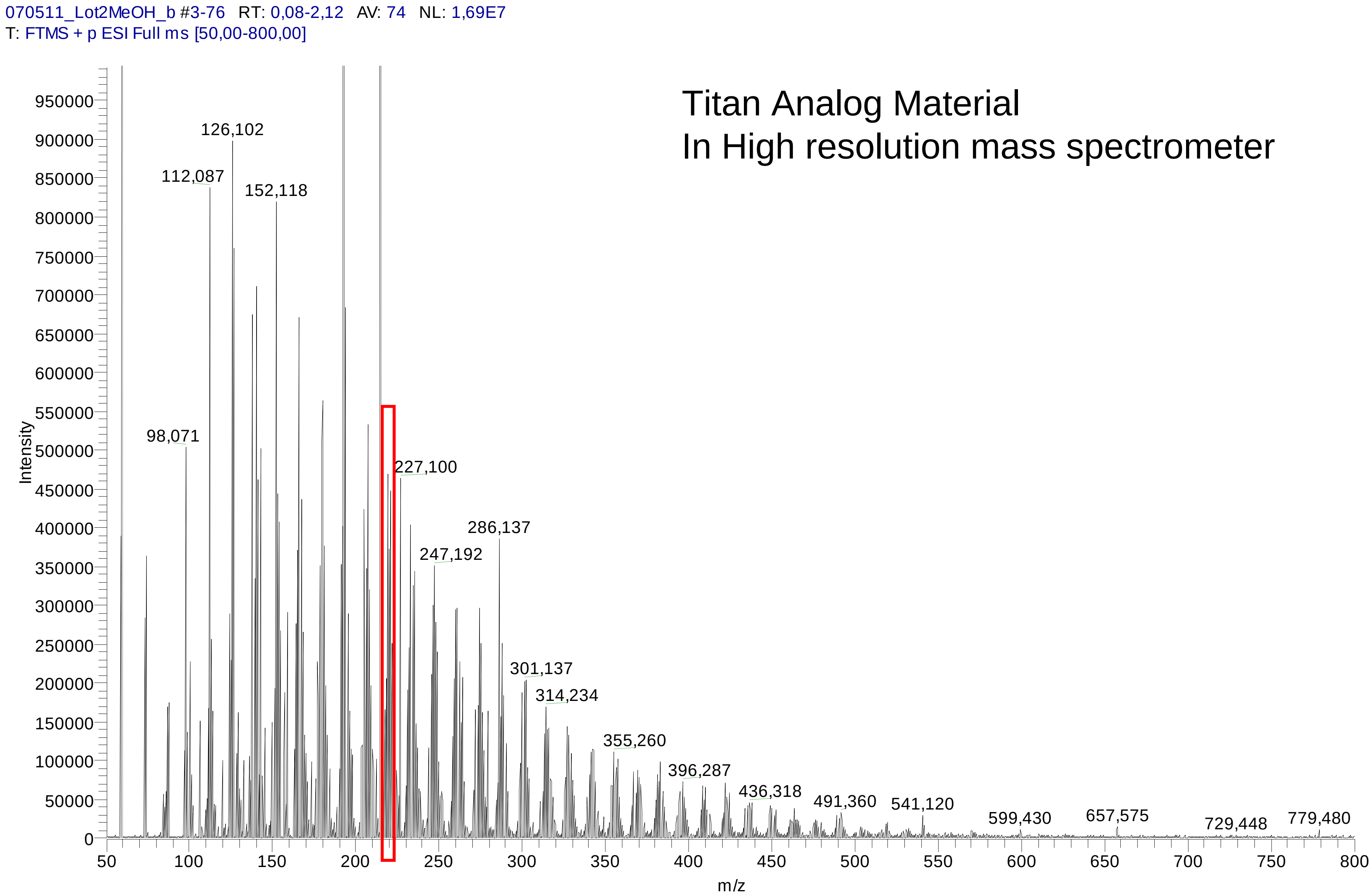
Examples

- Titan analog material
- meteorites

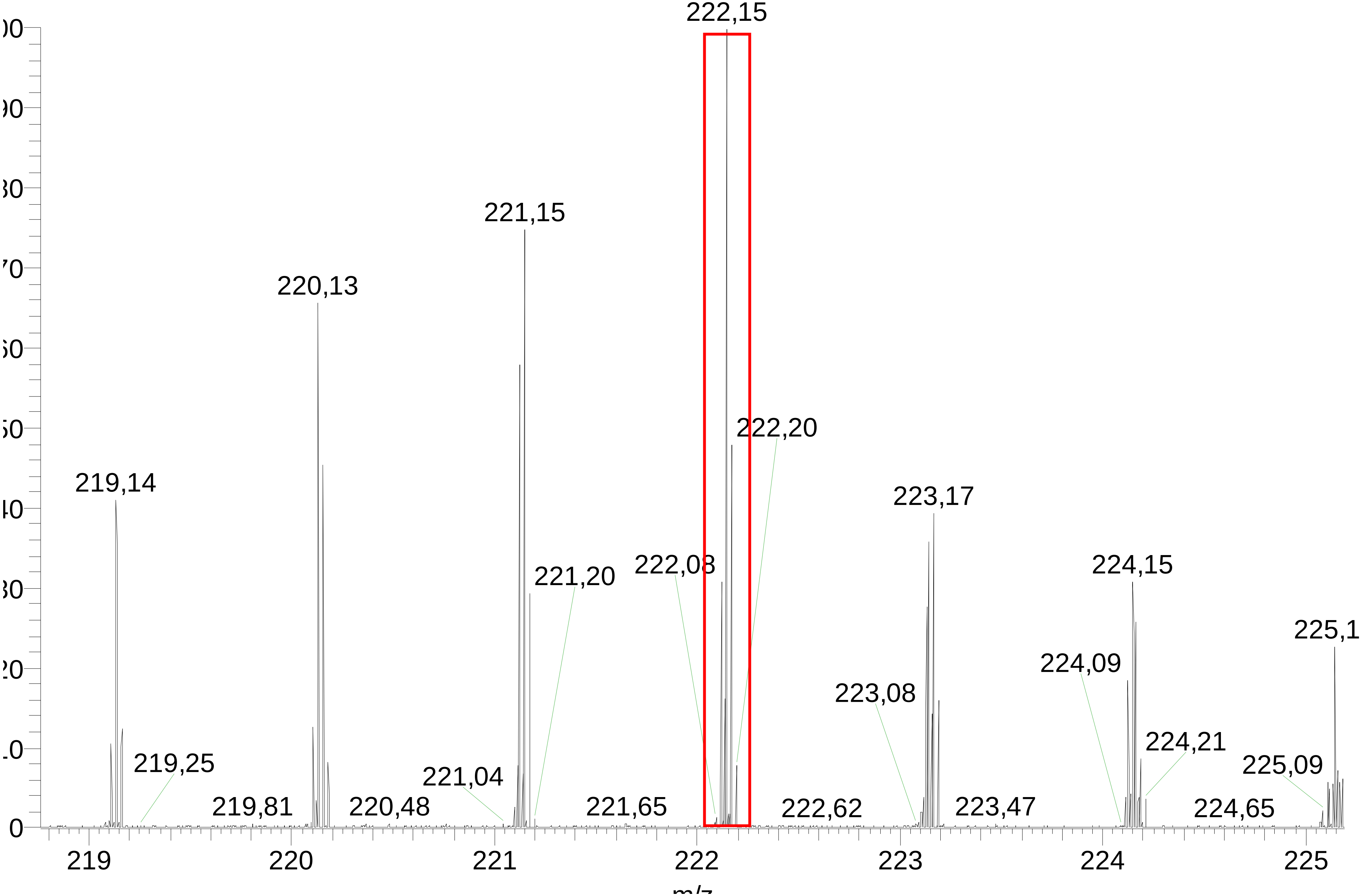
Cassini (INMS)
ionic densities in Titan ionosphere

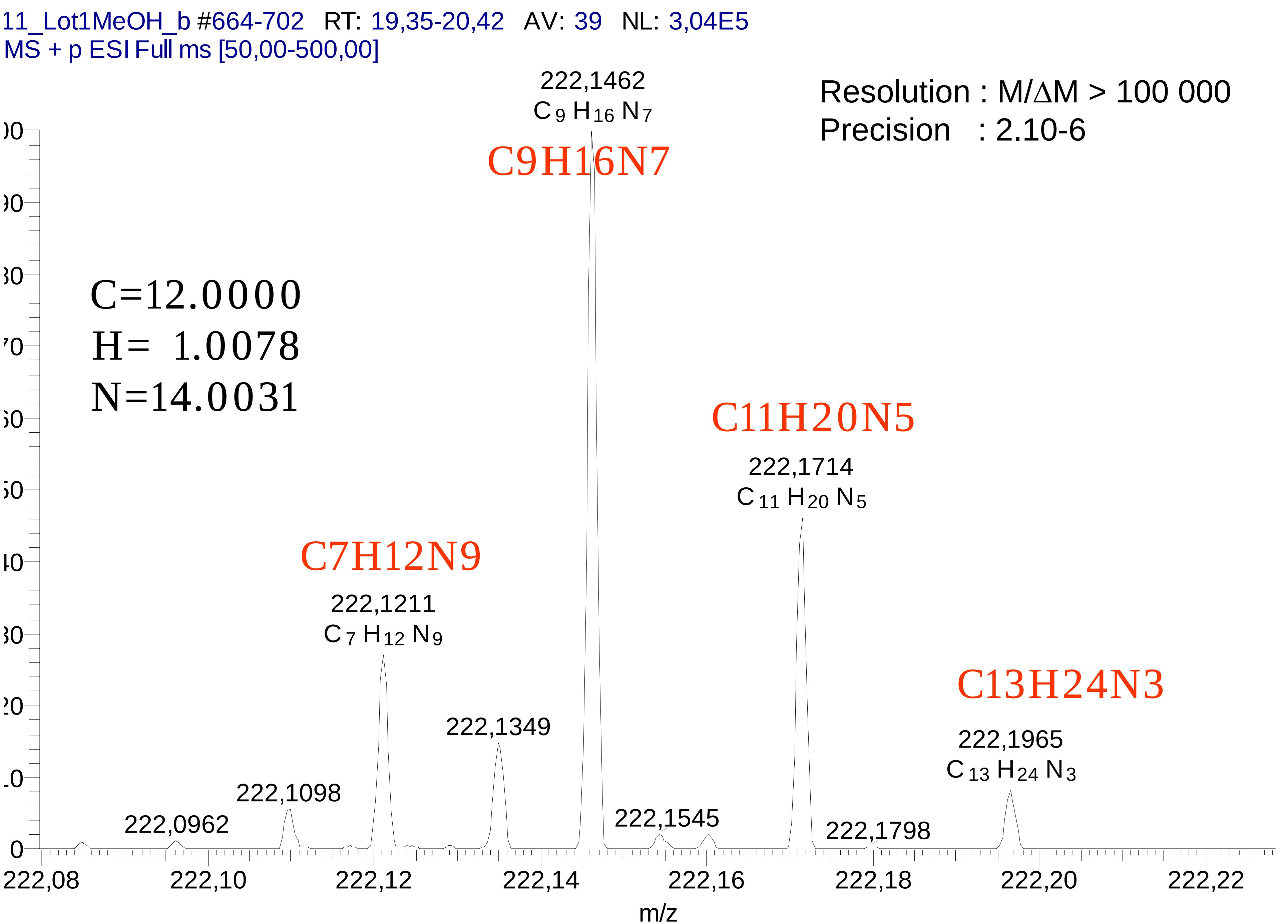


C_7NH^+	99.0109	resolution
$C_8H_3^+$	99.0235	7857
$C_4N_3H_9^+$	99.0797	1761
$C_5N_2H_{11}^+$	99.0923	7857
$C_6NH_{13}^+$	99.1049	7853
$C_7H_{15}^+$	99.1174	7920


070511_Lot2MeOH_b #3-76 RT: 0,08-2,12 AV: 74 NL: 1,69E7
T: FTMS + p ESI Full ms [50,00-800,00]
Titan Analog Material
In High resolution mass spectrometer
Intensity
m/z
126,102
112,087
152,118
98,071
227,100
286,137
247,192
301,137
314,234
355,260
396,287
436,318
491,360
541,120
599,430
657,575
729,448
779,480

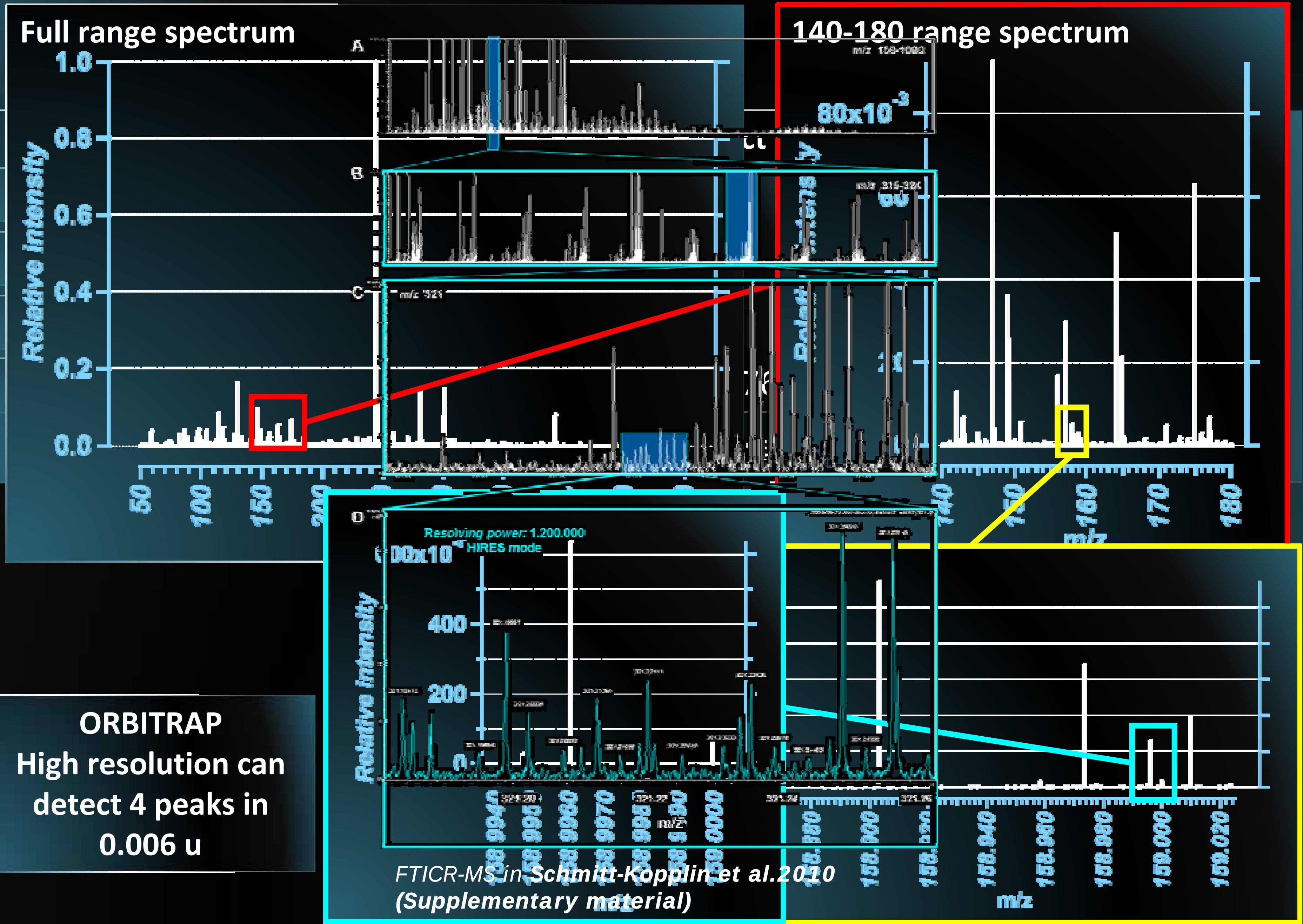
11_Lot1MeOH_b #661-717 RT: 19,27-20,83 AV: 57 NL: 3,02E5
MS + p ESI Full ms [50,00-500,00]





11_Lot1MeOH_b #664-702 RT: 19,35-20,42 AV: 39 NL: 3,04E5
MS + p ESI Full ms [50,00-500,00]
Resolution : M/ΔM > 100 000
Precision : 2.10-6
C=12.0000
H= 1.0078
N=14.0031
222,1462
C 9 H 16 N 7
C9H16N7
C11H20N5
222,1714
C 11 H 20 N 5
C7H12N9
222,1211
C 7 H 12 N 9
222,1349
C13H24N3
222,1965
C 13 H 24 N 3
222,0962
222,1098
222,1545
222,1798
222,08
222,10
222,12
222,14
222,16
222,18
222,20
222,22
m/z

FT-MS required to probe FOM diversity



ORBITRAP
High resolution can detect 4 peaks in 0.006 u

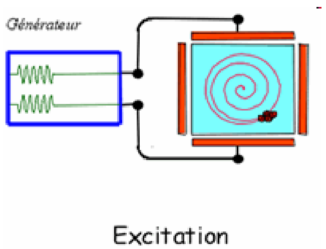
*FTICR-MS in **Schmitt-Kopplin et al.2010** (Supplementary material)*

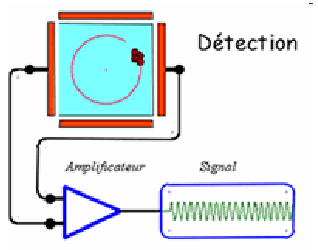
Instruments

- FT ICR
- FT Orbitrap

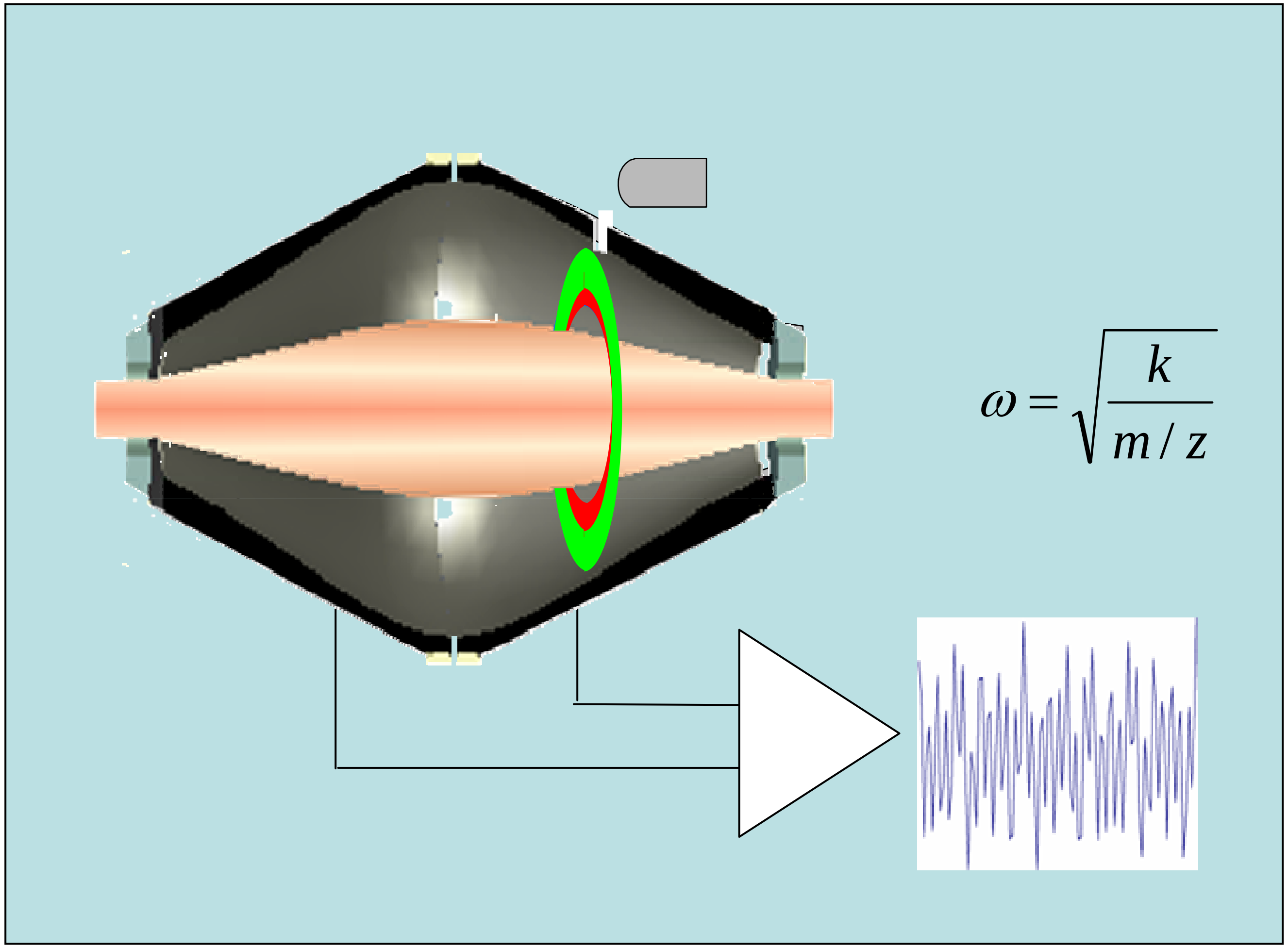
2 methods to record periodic movements of ions

- FTICR
  - intense magnetic fields (~10 tesla)
  - Trapping in 2 dimensions by B
  - Trapping in 3rd dimension by E
  - RF Excitation on 2 plates
  - Non destructive Detection of ω by induced curent on 2 plates
  - □ ω=eB z/m





- FT Orbitrap
  - □ $\omega=k\,(z/m)^{1/2}$


$\omega = \sqrt{\frac{k}{m/z}}$

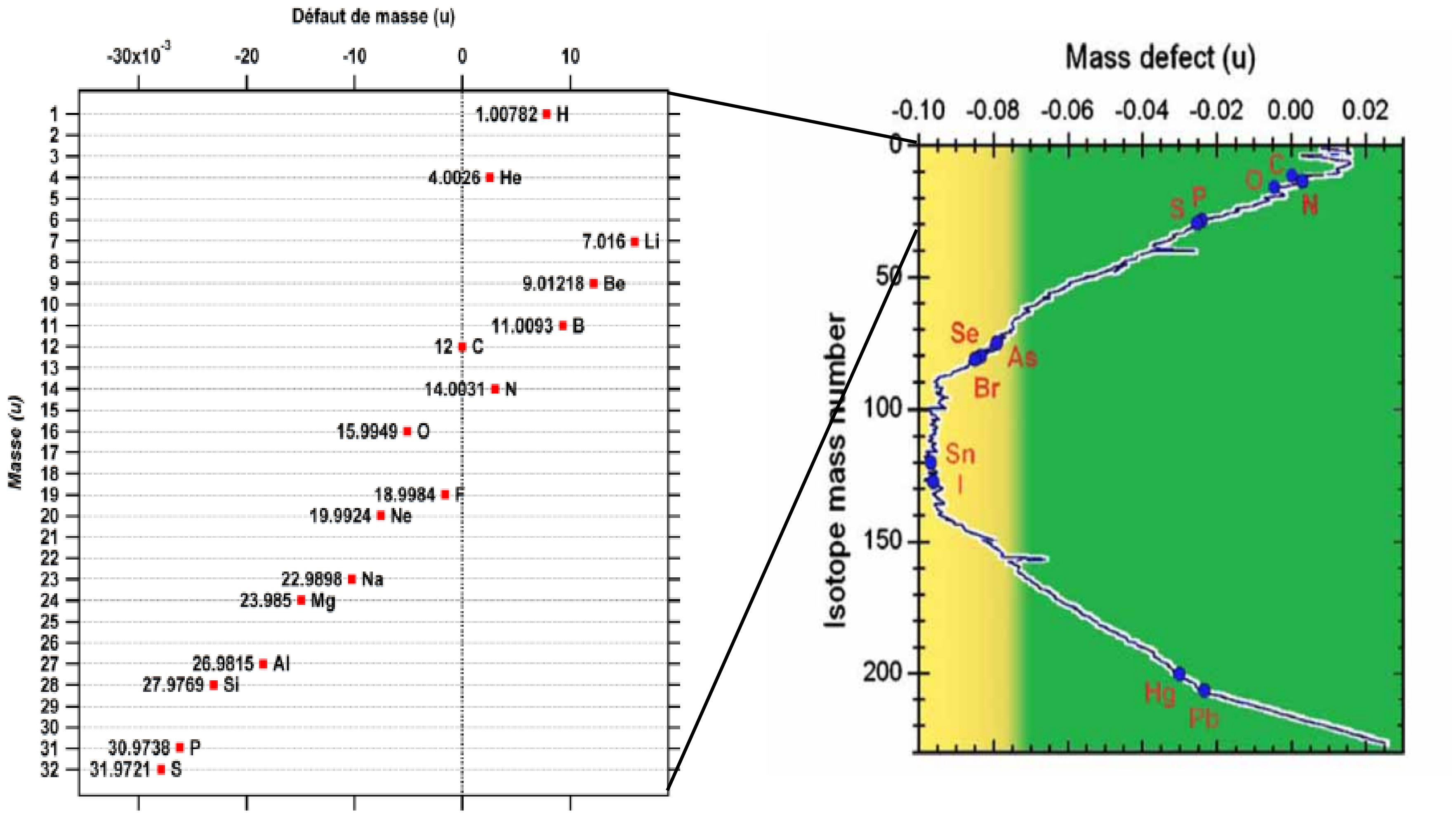
Information used to retrieve the chemical composition

- Mass defect
- Isotopic pattern

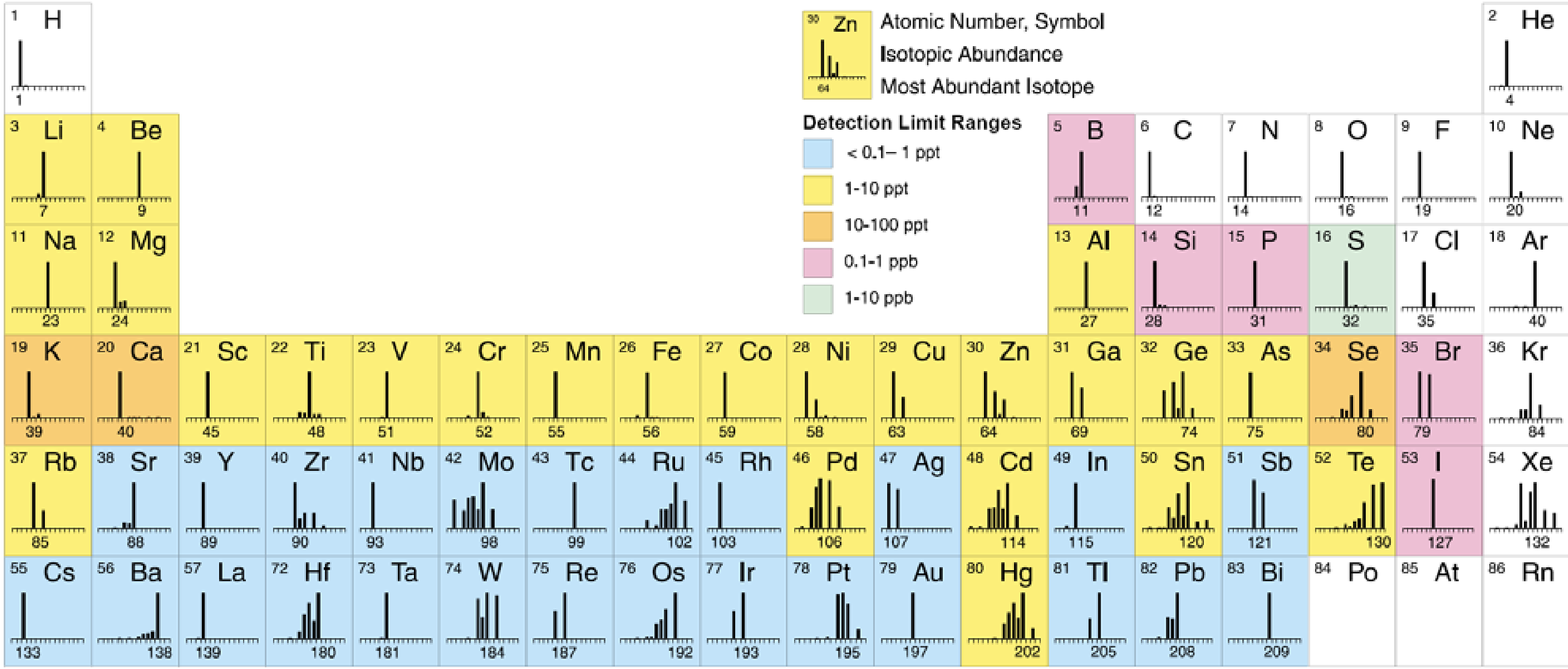
exact masses !

Element	masse	abundance(%)	m+1	abundance(%)	m+2	abundance(%)
H	1.00782	99.99	2.0141	0.015		
C	12	98.9	13.0034	1.1		
N	14.0031	99.63	15.0001	0.37		
O	15.9949	99.76	16.9991	0.038	17.9992	0.2

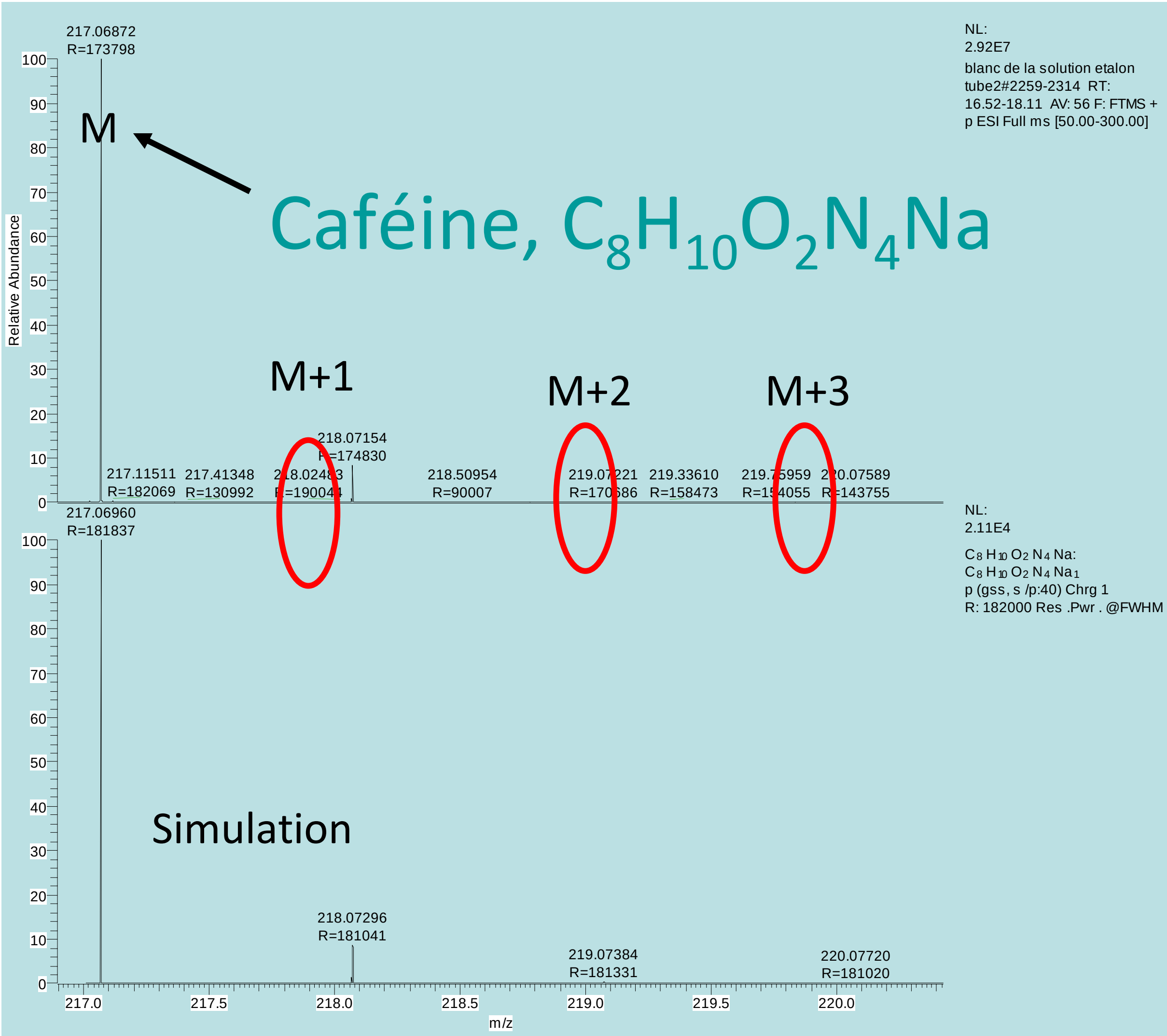
Mass Defect ⇔ Nuclear Physics
=
Chemical Information



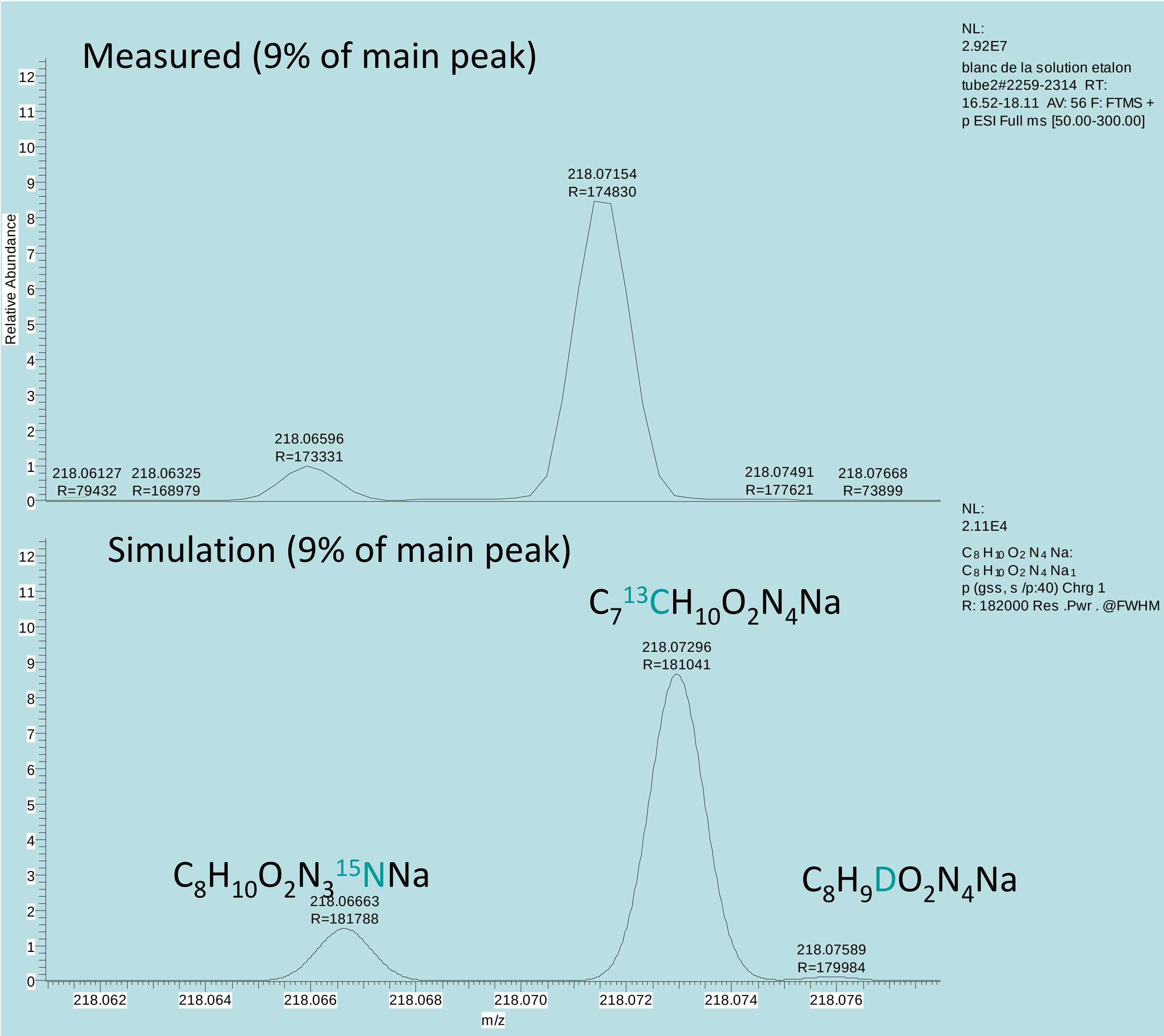
Isotopical abundancies, isotopical fingerprint…



Example on cafeine

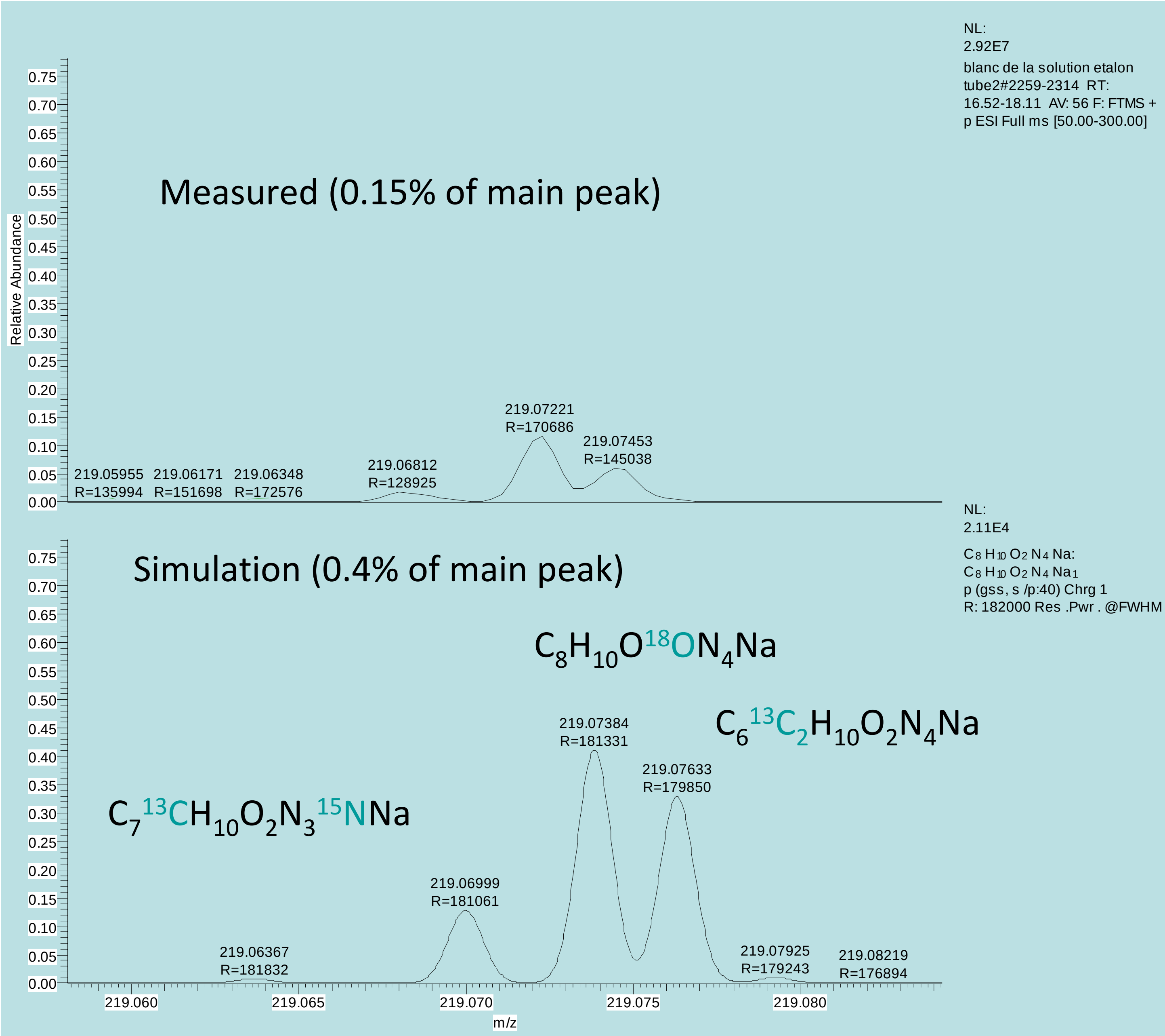


M+1

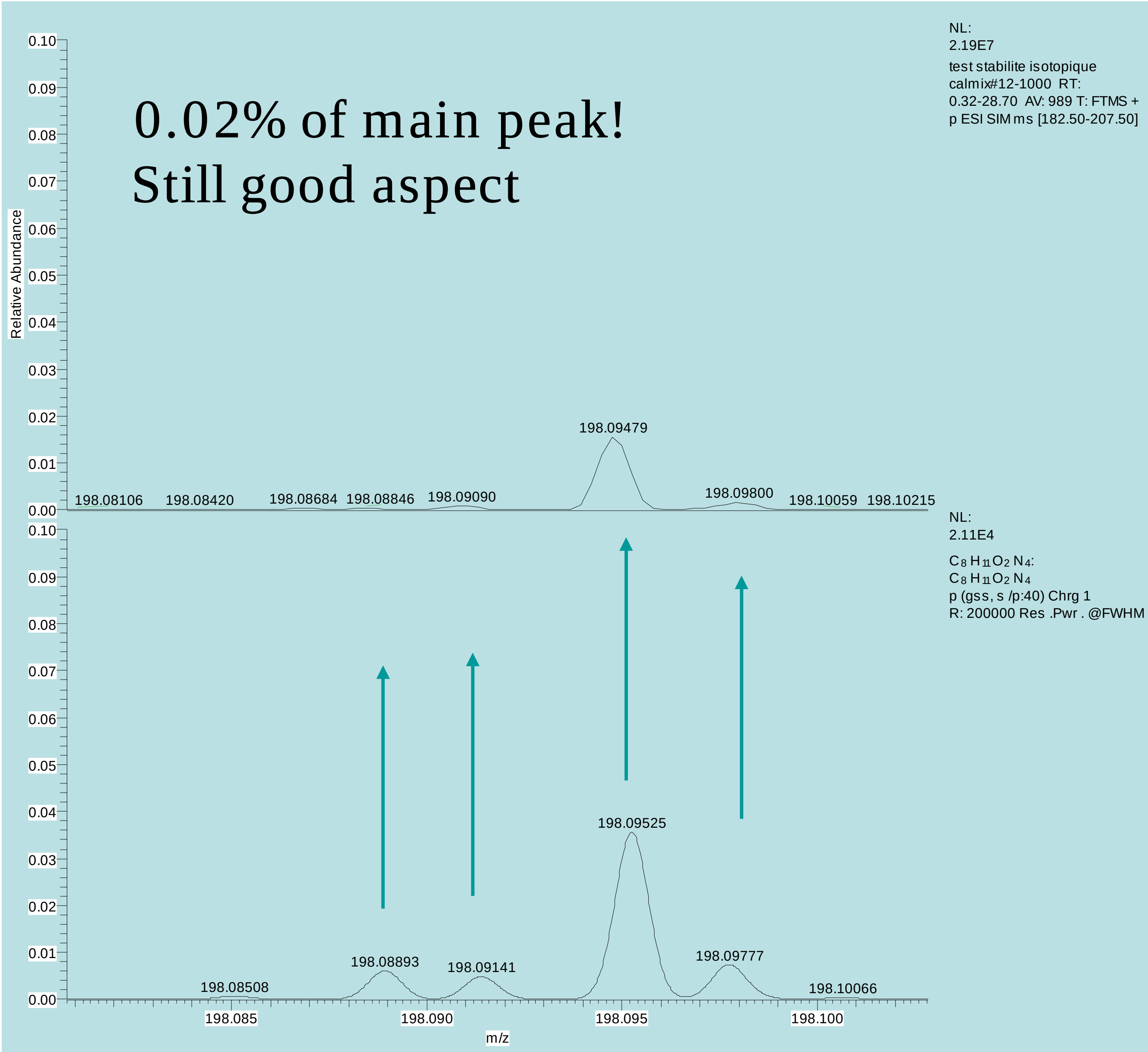


Measured (9% of main peak)
NL:
2.92E7
blanc de la solution etalon tube2#2259-2314 RT: 16.52-18.11 AV: 56 F: FTMS + p ESI Full ms [50.00-300.00]
Relative Abundance
218.07154
R=174830
218.06596
R=173331
218.06127
R=79432
218.06325
R=168979
218.07491
R=177621
218.07668
R=73899
Simulation (9% of main peak)
NL:
2.11E4
C8 H10 O2 N4 Na:
C8 H10 O2 N4 Na1
p (gss, s /p:40) Chrg 1
R: 182000 Res .Pwr . @FWHM
$C_7{}^{13}CH_{10}O_2N_4Na$
218.07296
R=181041
$C_8H_{10}O_2N_3{}^{15}NNa$
218.06663
R=181788
$C_8H_9DO_2N_4Na$
218.07589
R=179984
218.062
218.064
218.066
218.068
218.070
218.072
218.074
218.076
m/z

M+2


NL:
2.92E7
blanc de la solution etalon
tube2#2259-2314 RT:
16.52-18.11 AV: 56 F: FTMS +
p ESI Full ms [50.00-300.00]
Measured (0.15% of main peak)
Relative Abundance
219.05955 R=135994
219.06171 R=151698
219.06348 R=172576
219.06812 R=128925
219.07221 R=170686
219.07453 R=145038
NL:
2.11E4
C8 H10 O2 N4 Na:
C8 H10 O2 N4 Na1
p (gss, s /p:40) Chrg 1
R: 182000 Res .Pwr . @FWHM
Simulation (0.4% of main peak)
$C_8H_{10}O^{18}ON_4Na$
$C_6{}^{13}C_2H_{10}O_2N_4Na$
$C_7{}^{13}CH_{10}O_2N_3{}^{15}NNa$
219.06367 R=181832
219.06999 R=181061
219.07384 R=181331
219.07633 R=179850
219.07925 R=179243
219.08219 R=176894
m/z

M+3

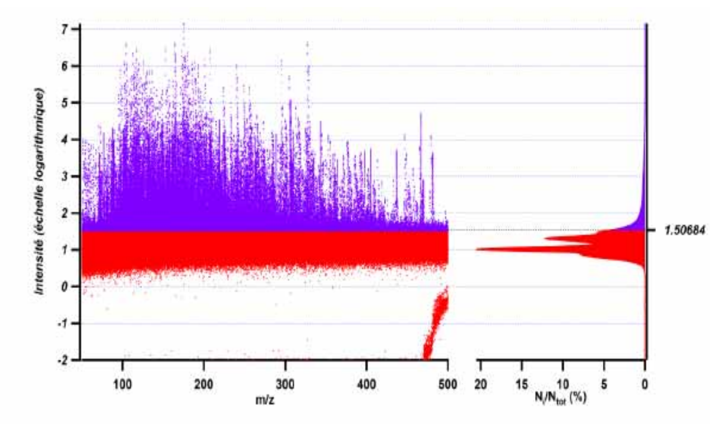


0.02% of main peak!
Still good aspect
NL:
2.19E7
test stabilite isotopique
calmix#12-1000 RT:
0.32-28.70 AV: 989 T: FTMS +
p ESI SIM ms [182.50-207.50]
NL:
2.11E4
$C_8 H_{11} O_2 N_4$:
$C_8 H_{11} O_2 N_4$
p (gss, s /p:40) Chrg 1
R: 200000 Res .Pwr . @FWHM
Relative Abundance
m/z
198.09479
198.08106
198.08420
198.08684
198.08846
198.09090
198.09800
198.10059
198.10215
198.09525
198.08508
198.08893
198.09141
198.09777
198.10066
198.085
198.090
198.095
198.100

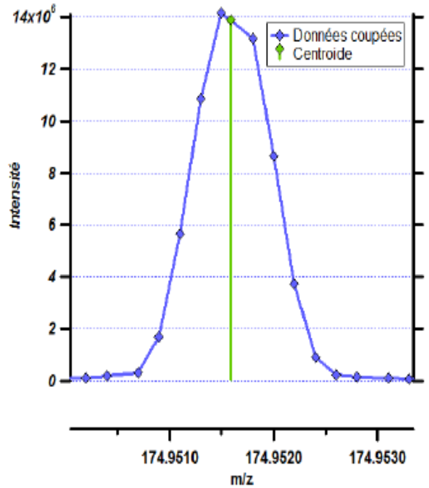
methods to retrieve the chemical composition

- Data reduction
- data representation
- Best candidate formula search
- Molecular Data test and interpretation

Procedure of data treatment



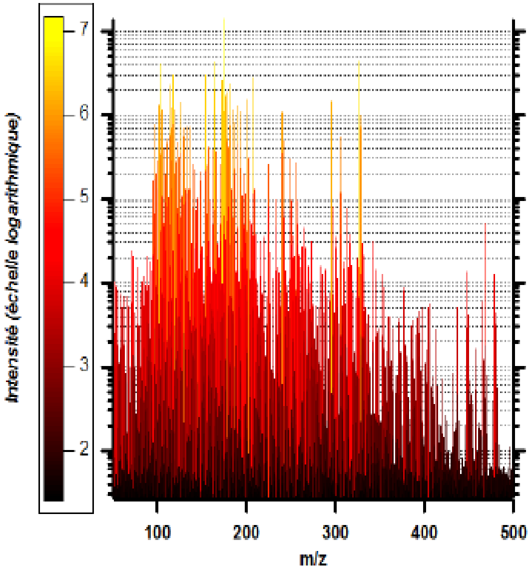
+



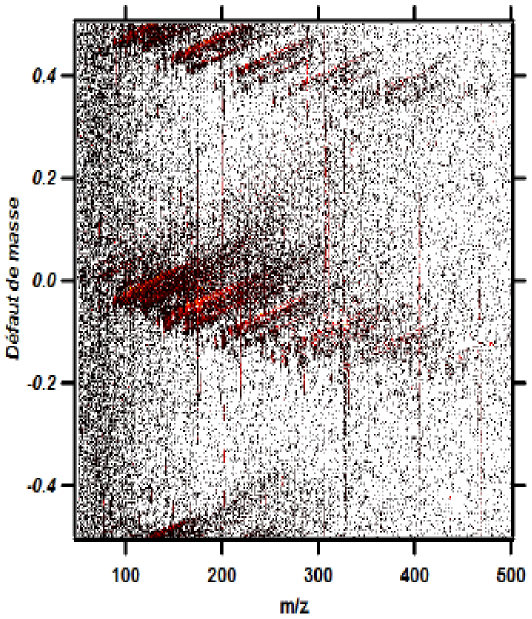
Remove « noise »

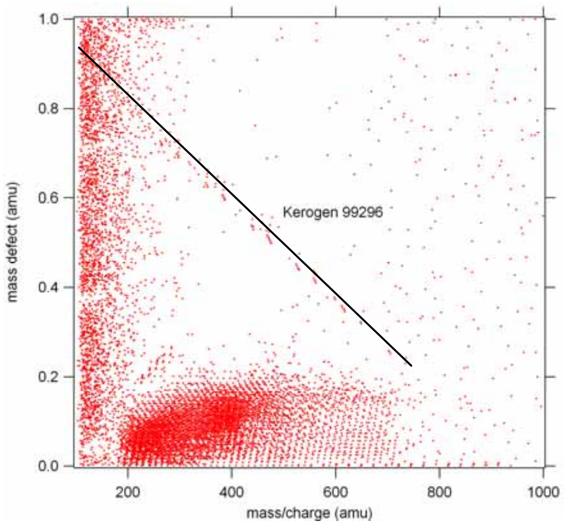
Reduce to centroids

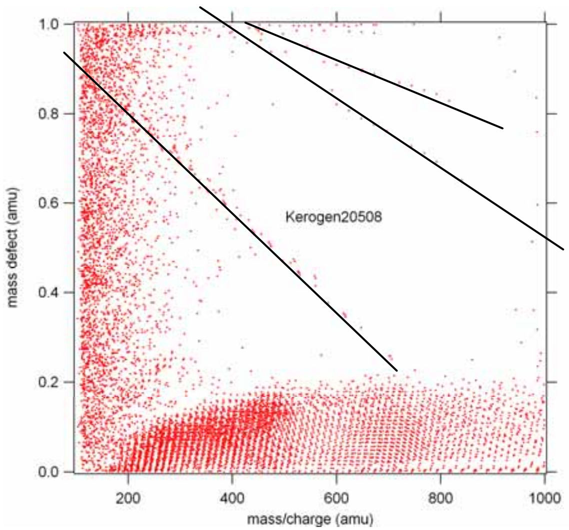
Question, where to cut between noise and signal?
Question, what about peak overlaps?



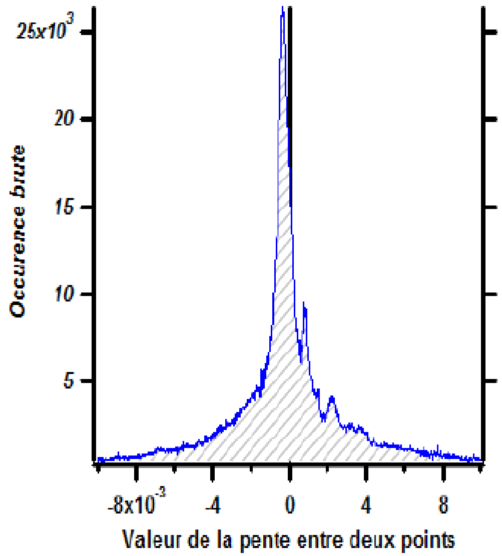
Intensity vs Mass
⇔
mass defect vs mass spectra

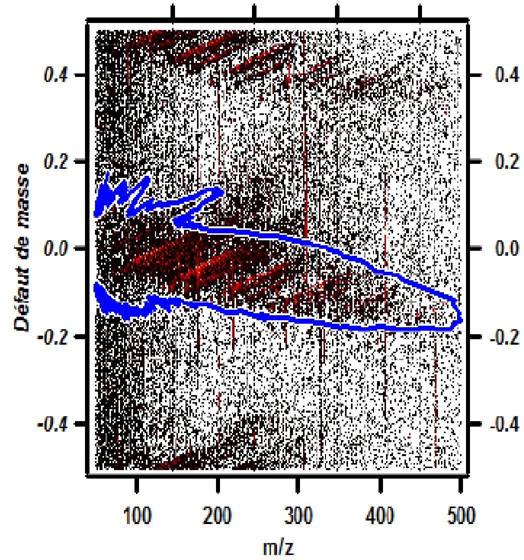


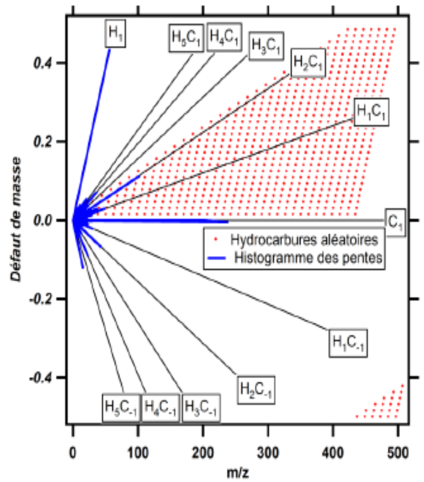




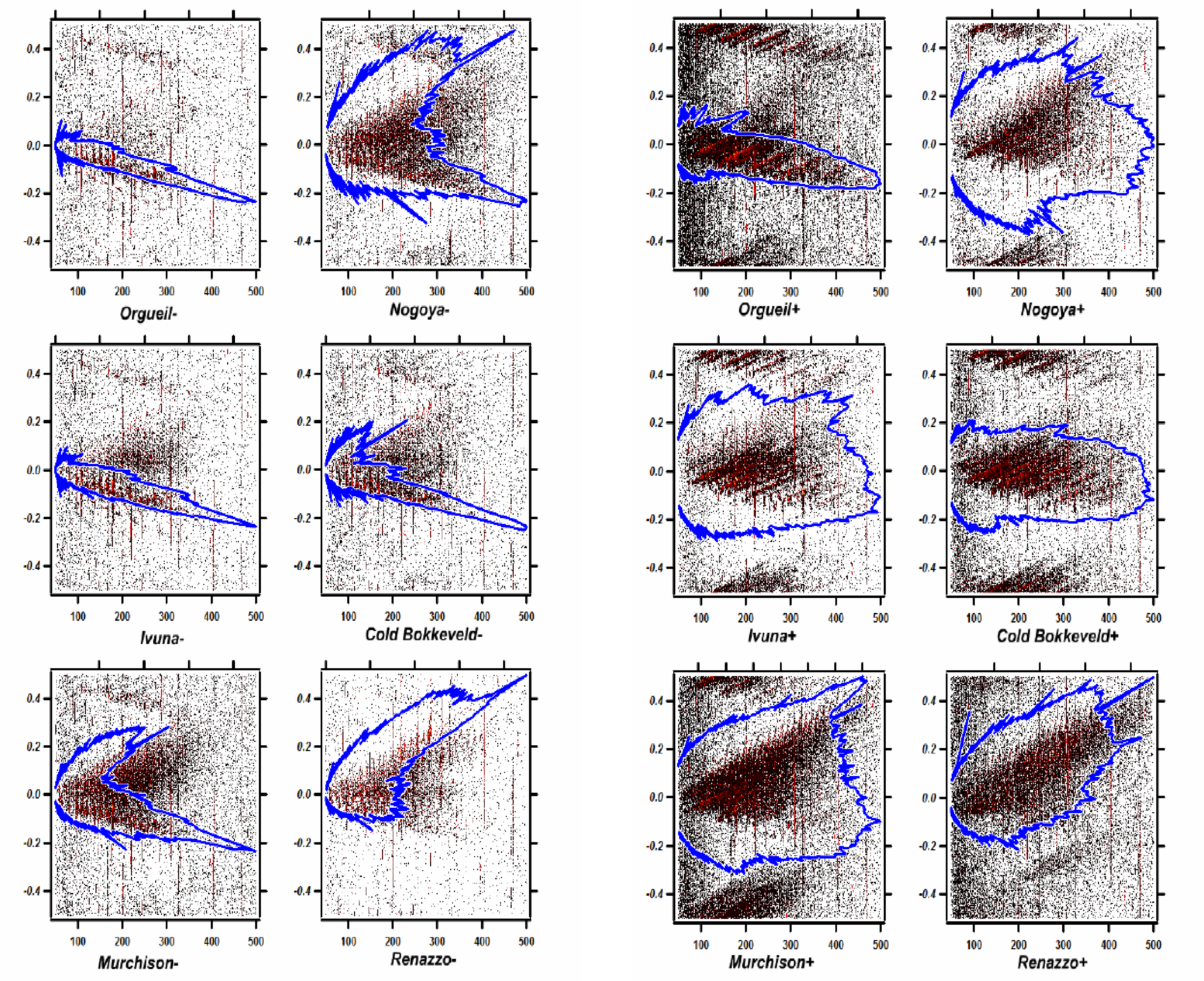
Slope statistics







Question : graph theory?


Orgueil-
Nogoya-
Orgueil+
Nogoya+
Ivuna-
Cold Bokkeveld-
Ivuna+
Cold Bokkeveld+
Murchison-
Renazzo-
Murchison+
Renazzo+
0.4
0.2
0.0
-0.2
-0.4
100
200
300
400
500

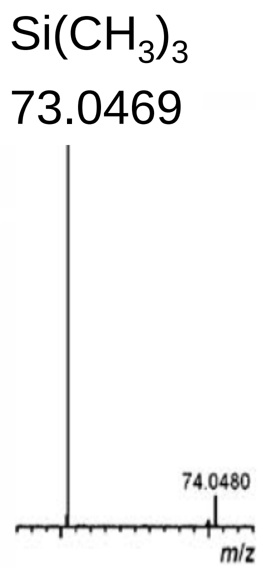
Search for ?best? molecular formula

Input variables ⇒ $\sum_{X} \sum_{i}^{n} n_i X_i = 73$ ⇒ solutions

Diophantine equation

Si_n	(n = 1–2)
O_n	(n = 0–4)
C_n	(n = 0–6)
$(OH)_n$	(n = 0–4)
$(CH_3)_n$	(n = 0–4)
$(CH_2)_n$	(n = 0–5)
$(CH)_n$	(n = 0–5)

	exact mass
Si_2OH	72.9566
$SiOC(OH)$	72.9746
$SiO_2(CH)$	72.9746
$Si(OH)(CH_2)_2$	73.0110
$Si(OH)(CH)CH_3$	73.0110
$SiO(CH_2)(CH_3)$	73.0110
$Si(CH_3)_3$	<u>73.0473</u>



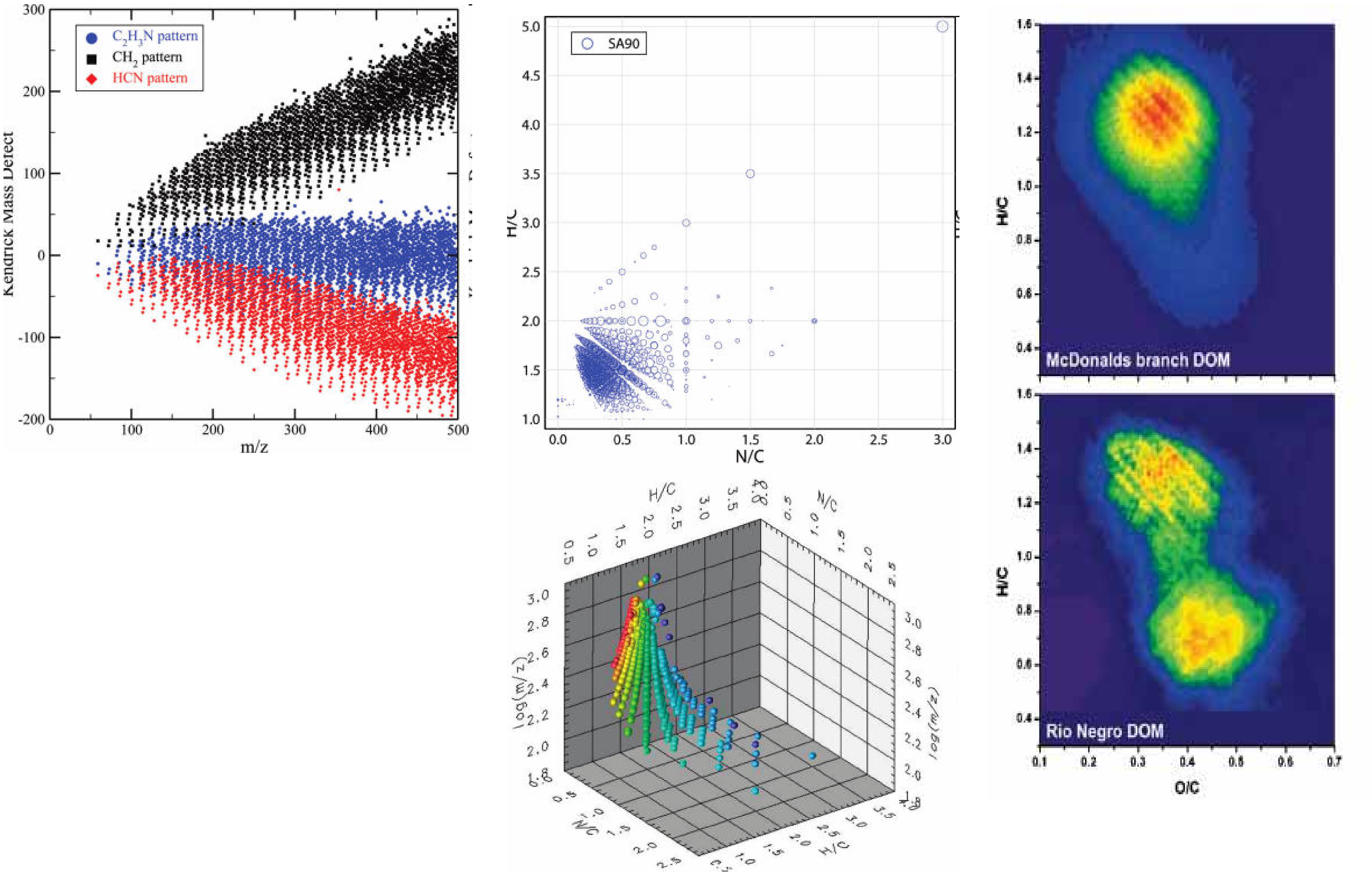
Problem, the method explodes with molecular size and number of species

Question, autres méthodes plus efficaces pour trouver la meilleure solution

Result : Mass Intensity atoms

mass	measured	formula	C	H	N	O	intensity
83.0364	83.0363	C2 H3 N4	2	3	4	0	49192176
92.0255	92.0254	C4 H2 N3	4	2	3	0	2483336
93.0207	93.0207	C3 H N4	3	1	4	0	8880951
95.0363	95.0363	C3 H3 N4	3	3	4	0	44190828
97.052	97.052	C3 H5 N4	3	5	4	0	31413156
105.0458	105.0458	C6 H5 N2	6	5	2	0	2722807
106.0411	106.0411	C5 H4 N3	5	4	3	0	12958549
107.0363	107.0363	C4 H3 N4	4	3	4	0	1.15E+08
108.0316	108.0316	C3 H2 N5	3	2	5	0	2.47E+08
109.052	109.052	C4 H5 N4	4	5	4	0	32081540
110.0472	110.0472	C3 H4 N5	3	4	5	0	3877392
110.0724	110.0724	C5 H8 N3	5	8	3	0	4217568
111.0676	111.0676	C4 H7 N4	4	7	4	0	19245324
112.0414	112.0404	C5 H6 N O2	5	6	1	2	3951274
112.0629	112.0629	C3 H6 N5	3	6	5	0	8516299
116.0254	116.0254	C6 H2 N3	6	2	3	0	20673458
117.0207	117.0207	C5 H N4	5	1	4	0	41064336
118.0159	118.0159	C4 N5	4	0	5	0	44674168
118.0411	118.0411	C6 H4 N3	6	4	3	0	18135442
119.0363	119.0363	C5 H3 N4	5	3	4	0	98598008
119.0614	119.0615	C7 H7 N2	7	7	2	0	4025135
120.0316	120.0316	C4 H2 N5	4	2	5	0	1.54E+08

Data representation : Kendrick and Van Krevelen diagram



Problem, qualitative representation of dense data sets

Conclusions

- Astromaterial or astro analogs are very complex, polymer like, mixtures
- High Resolution Mass Spectrometry is a precious tool to explore those materials
- Recorded spectra (I,M) are very informative and contain strongly correlated information
- Need for improved tools (concept and algo) to perform
  - Data reduction
  - Qualitative analysis
  - Data analysis
  - Data representation

THANK YOU