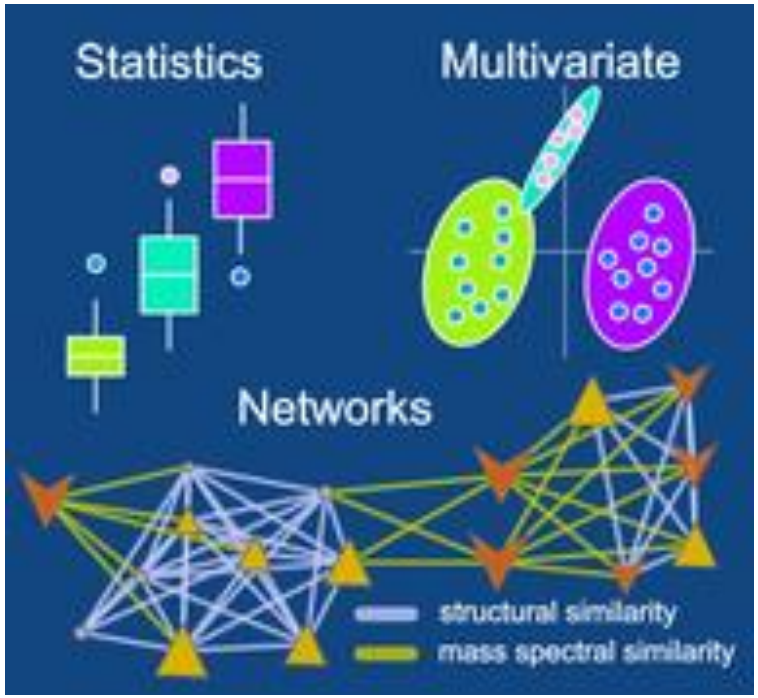
Quelques outils et idées en spectrométrie de masse assistée par vos neurones et la bioinformatique



David TOUBOUL

CNRS
ICSN
Gif-sur-Yvette

David.Touboul@cnrs.fr

La problématique des « big data »

L'étude des systèmes biologiques conduit à la nécessité de gérer des données de plus en plus conséquentes de données afin d'évaluer un plus grand nombre de variables (pertinentes ou non).

Deux approches sont possibles et complémentaires:

- Approches ciblées (connaissance *a priori* de l'objet à étudier ou hypothèses initiales plus ou moins crédibles)
- Approches non ciblées (génération de « big data »)

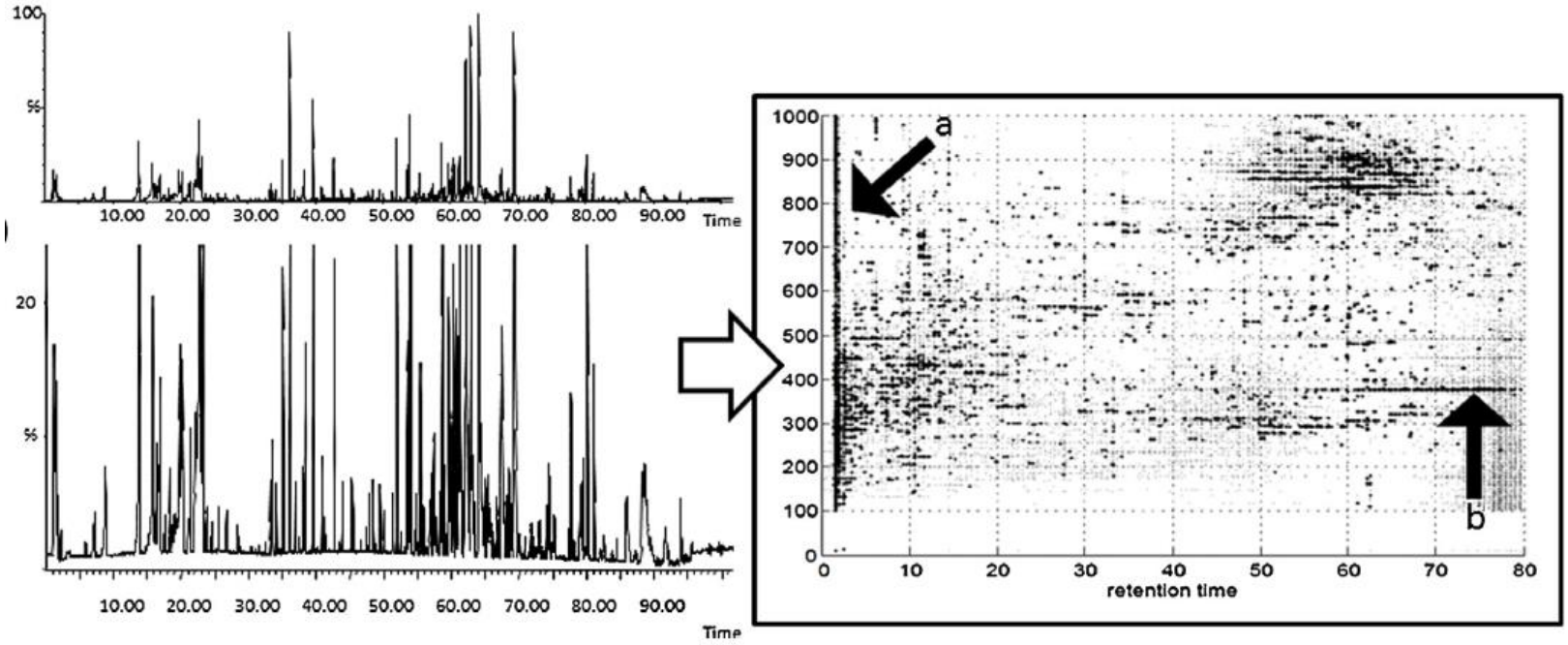
- Comment les générer ?
- Comment les stocker ?
- Comment les analyser ?

Un cas simple: la protéomique

Simple car:

- 20 acides aminés principaux
- Des modifications connus et facile à implémenter
- Des bases de données riches (mais avec des erreurs d'annotation aussi)

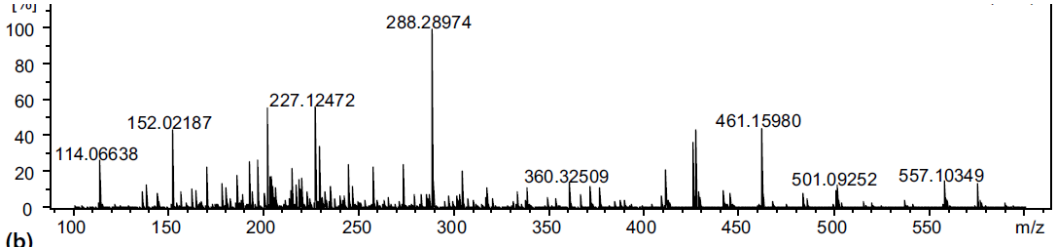
Un cas compliqué: la métabolomique



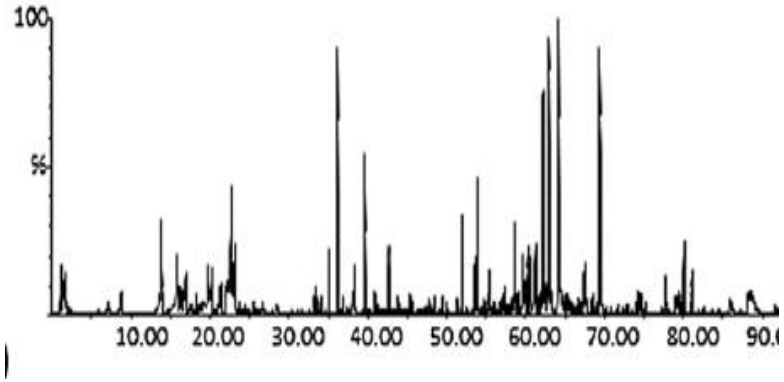
Pas plus de complexité qu'en protéomique en terme de nombre de molécules détectées
Mais plus compliqué d'interprétation car:

- un grand nombre de fonctions chimiques et de squelettes carbonés
- base de données incomplètes

Première étape: choisir sa stratégie !



Shotgun:
Rapide
MS haute résolution nécessaire
Suppression ionique
Difficile pour la quantification



Chromatographie-MS:
Long
Pas de méthode universelle pour toutes les classes de molécules
MS haute résolution nécessaire
Moins de suppression ionique
Plus facile pour la quantification

Première étape: choisir sa stratégie !

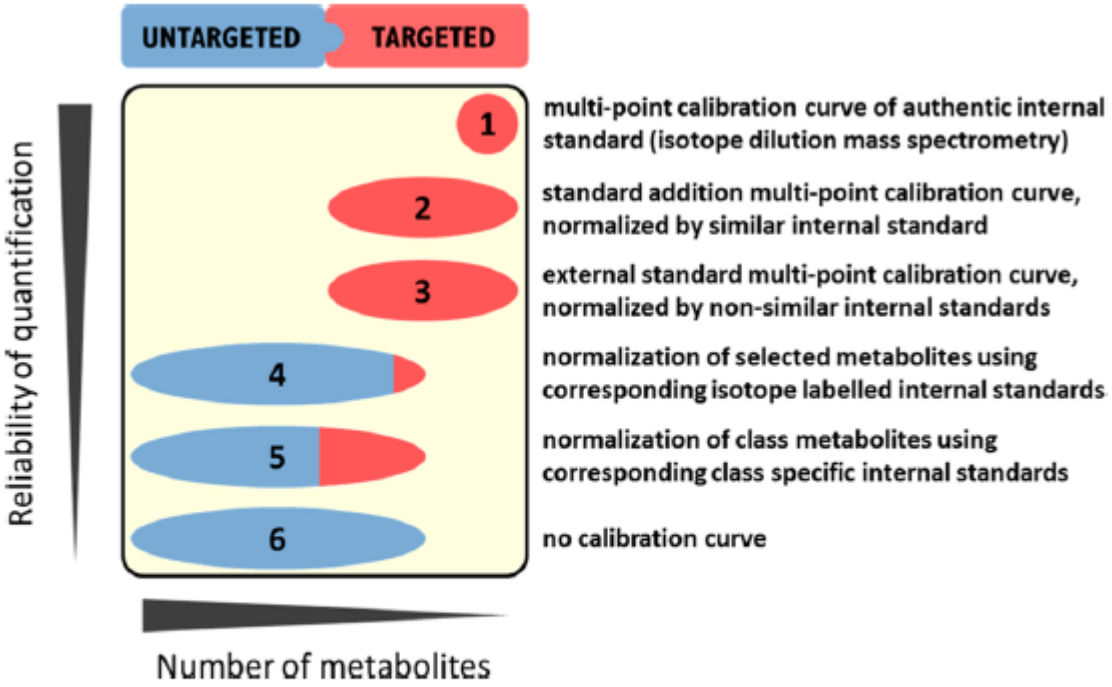


Figure 1. Untargeted and targeted metabolomics/lipidomics methods in relation to the number of detected metabolites and reliability of quantitative results.

DOI: 10.1021/acs.analchem.5b04491
Anal. Chem. 2016, 88, 524−545

La physico-chimie au cœur du problème

MTBE: Méthyl tert-butyl éther

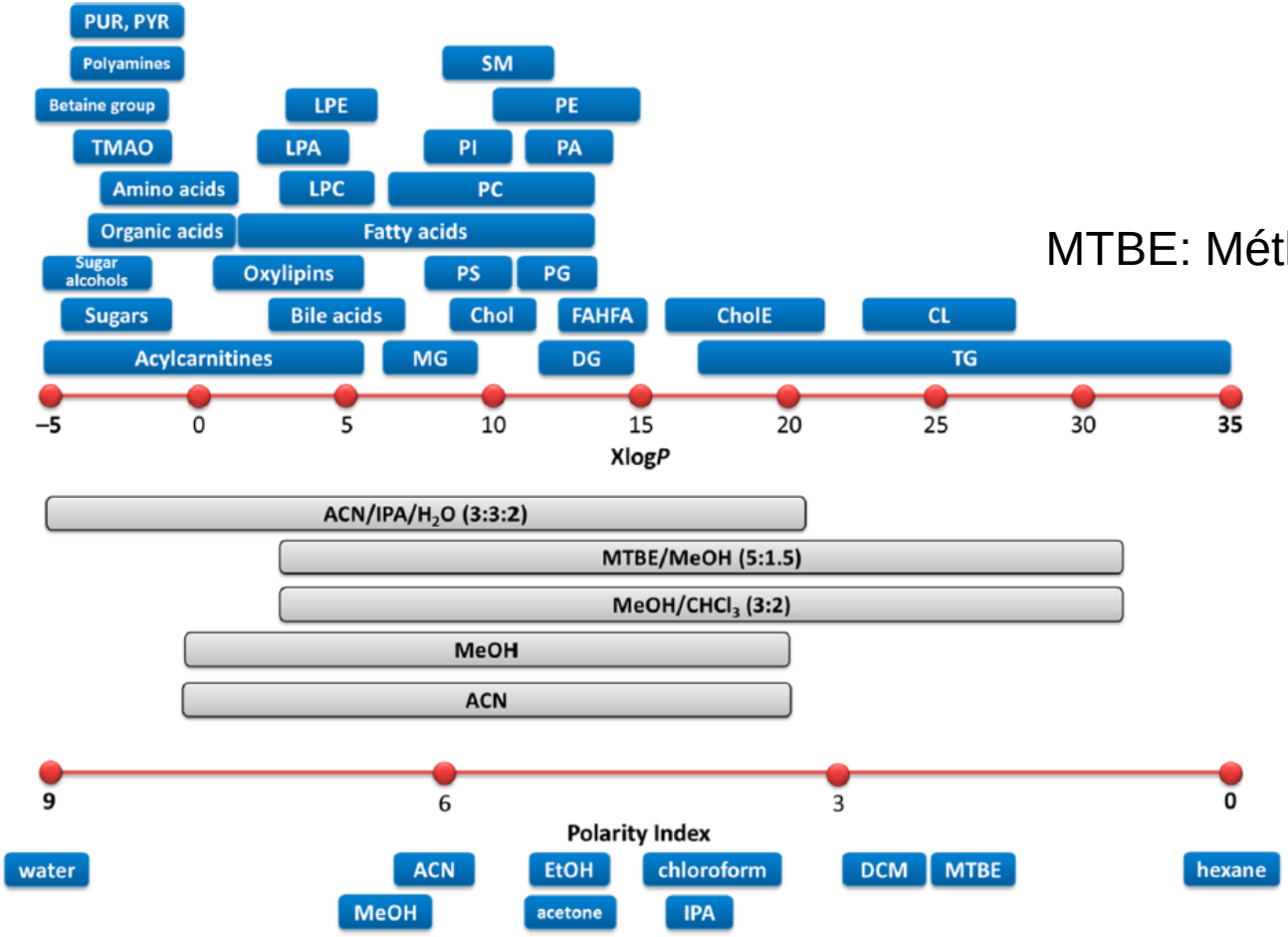


Figure 2. Predicted octanol/water partition coefficient (*X* log *P*) range of common metabolites (data for representative metabolites taken from ref 25) in blood plasma and polarity index of solvents[26] used for sample extraction. Typical solvents or solvent mixtures (in gray color) used in metabolomics and lipidomics indicate the polarity range of isolated metabolites with high recovery. Extraction ranges increase if ternary mixtures are used that include water, or if water is added in simultaneous extraction/fractionation procedures similar to the classic Bligh–Dyer[42] or Matyash–Schwudke[43] protocols. Legend: Cer, ceramides; Chol, cholesterol; CholE, cholesteryl esters; CL, cardiolipins; DG, diacylglycerols; FAHFA, fatty acid esters of hydroxyl fatty acids; LPA, lysophosphatidic acids; LPC, lysophosphatidylcholines; LPE, lysophosphatidylethanolamines; MG, monoacylglycerols; PA, phosphatidic acids; PC, phosphatidylcholines; PE, phosphatidylethanolamines; PG, phosphatidylglycerols; PI, phosphatidylinositols; PS, phosphatidylserines; PUR, purines; PYR, pyrimidines; SM, sphingomyelins; TG, triacylglycerols; TMAO, trimethylamine *N*-oxide.

Shotgun: tout est dans la préparation de l'échantillon

Table 1. Combined Extraction Methods for the Analysis of Hydrophilic, Amphiphilic, and Lipophilic Metabolites

matrix	extraction method	phase collected	resuspension solvent(s)	LC–MS	ref
plant tissue	MeOH/MTBE/H_2O	MTBE/MeOH fraction	ACN/IPA (7:3, v/v)	C8 column ESI(+)	50
		H_2O/MeOH fraction	water	C18 column ESI(+)	
liver and muscle tissue	MeOH/MTBE/H_2O	MTBE/MeOH fraction	chloroform/MeOH mixture (2:1, v/v) diluted in ACN/IPA/H_2O (65:30:5, v/v/v)	C8 column ESI(+)	51
		MTBE/MeOH fraction and H_2O/MeOH fraction (2:1 ratio)	20% MeOH	C18 column ESI(+)	
plasma	MeOH/MTBE/H_2O	MTBE/MeOH fraction	MeOH	C18 column ESI(+), ESI(−)	52
		H_2O/MeOH fraction	5% ACN	HILIC column ESI(+), ESI(−)	
plasma	MeOH/MTBE/H_2O	MTBE/MeOH fraction	Direct injection	C8 column ESI(+), ESI(−)	11,12
		H_2O/MeOH fraction	direct injection	C18 column ESI(+), ESI(−)	
plasma	MeOH/MTBE/H_2O	MTBE/MeOH fraction	IPA	C18 column ESI(+)	53
		H_2O/MeOH fraction	water (0.1% formic acid)	C18-PFP column ESI(+)	
plasma	SPE 96-well plates Step 1: ACN Step 2: chloroform/MeOH (2:1, v/v)	ACN fraction	MeOH/H_2O (1:1, v/v)	C18 column ESI(+)	37
		chloroform/MeOH fraction	chloroform/MeOH (2:1, v/v)	C18 column ESI(+)	
serum spot	MeOH/MTBE/H_2O	MTBE/MeOH fraction	direct injection	C18 column ESI(+)	54
		H_2O/MeOH fraction	direct injection	C18 column ESI(+)	

DOI: 10.1021/acs.analchem.5b04491
Anal. Chem. 2016, 88, 524−545

Chromatographie: un vaste choix …

Table 3. Overview of LC–MS Platforms Typically Used in Metabolomics and Lipidomics

chromatography mode	metabolite scope	typical mobile phase	typical stationary phase	limitations
RPLC (metabolomics)	• Polar and medium polar metabolites	• H_2O → ACN or MeOH	• C18	• Not for matrixes that include TG/CholE because ACN/MeOH are too weak solvents to elute these lipid classes, leading to ghost peaks and column deterioration
RPLC (lipidomics)	• Separation of lipids based on lipophilicity, which is governed by the carbon-chain length and the number of double bonds	• H_2O/ACN (~1:1) → high % IPA (with ACN)	• C18, C8	• When using IPA, high back pressures are observed • High temperatures may be used to improve separations (may be problematic for some column types) • Different mobile-phase modifiers needed for positive and negative ion mode to increase lipidome coverage • Class specific internal standard not eluted at close proximity of the rest lipids of the same lipid class • Phosphatidic acids (PA) and phosphatidylserines (PS) tend to elute as extensively broad peaks • Separation between different classes is decreased as compared to HILIC/NPLC
IP-RPLC (metabolomics)	• Very polar metabolites	• H_2O → ACN or MeOH with an ion-pair reagent	• C18	• Permanent contamination of LC system and ion source with ion-pair reagents possible • Cleanup system needed after several injections
HILIC (metabolomics)	• Very polar metabolites	• ACN → H_2O	• Amide, silica	• Longer column equilibration time compared to RPLC • Scope of the method strongly depends on pH of the mobile phase • When using basic pH (~9) shorter column lifetime can be expected
HILIC (lipidomics)	• Separation of lipids according to headgroup classes (from nonpolar to polar)	• ACN → H_2O	• Amide, silica	• Longer column equilibration time compared to RPLC • Much narrower spread of peaks within each class of lipids as compared to RPLC
NPLC (lipidomics)	• Separation of lipids according to headgroup classes (from nonpolar to polar)	• Heptane, chloroform, hexane → MeOH, ACN	• Silica	• Less robust compared to HILIC • Use of chlorinated solvents may raise operational costs and environmental concerns • Low ionization capacity of solvents used
SFC (lipidomics)	• Separation of lipids according to headgroup classes (from nonpolar to polar)	• CO_2 → MeOH	• Silica	• Possible retention time shift due to stationary phase degradation (regeneration of stationary phase recommended

Ne pas négliger la GC-MS !

DOI: 10.1021/acs.analchem.5b04491
Anal. Chem. 2016, 88, 524−545

Chromatographie: un vaste choix …

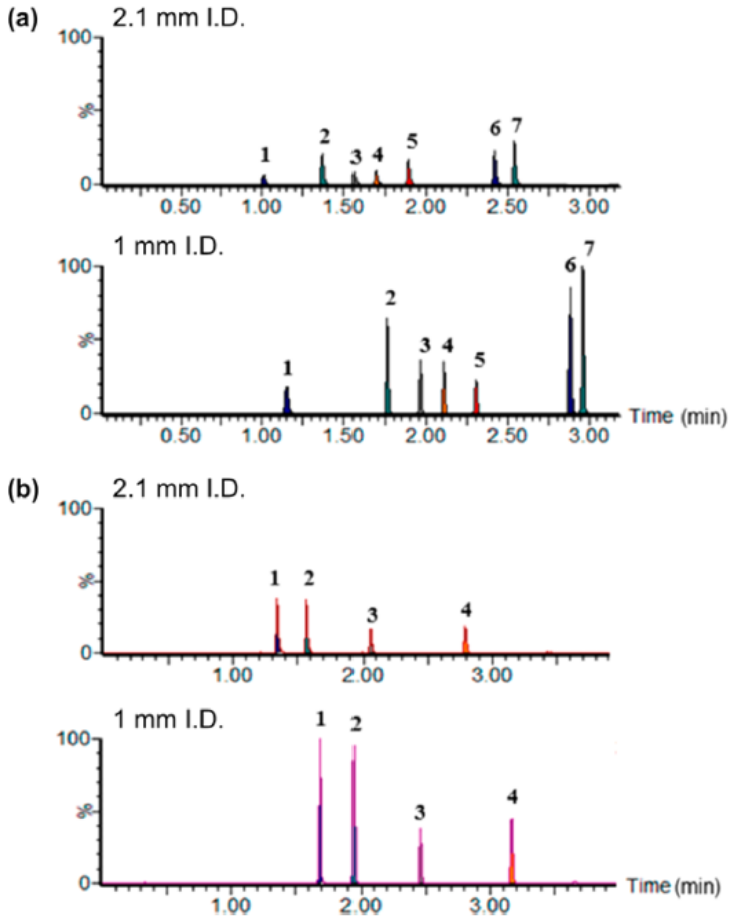
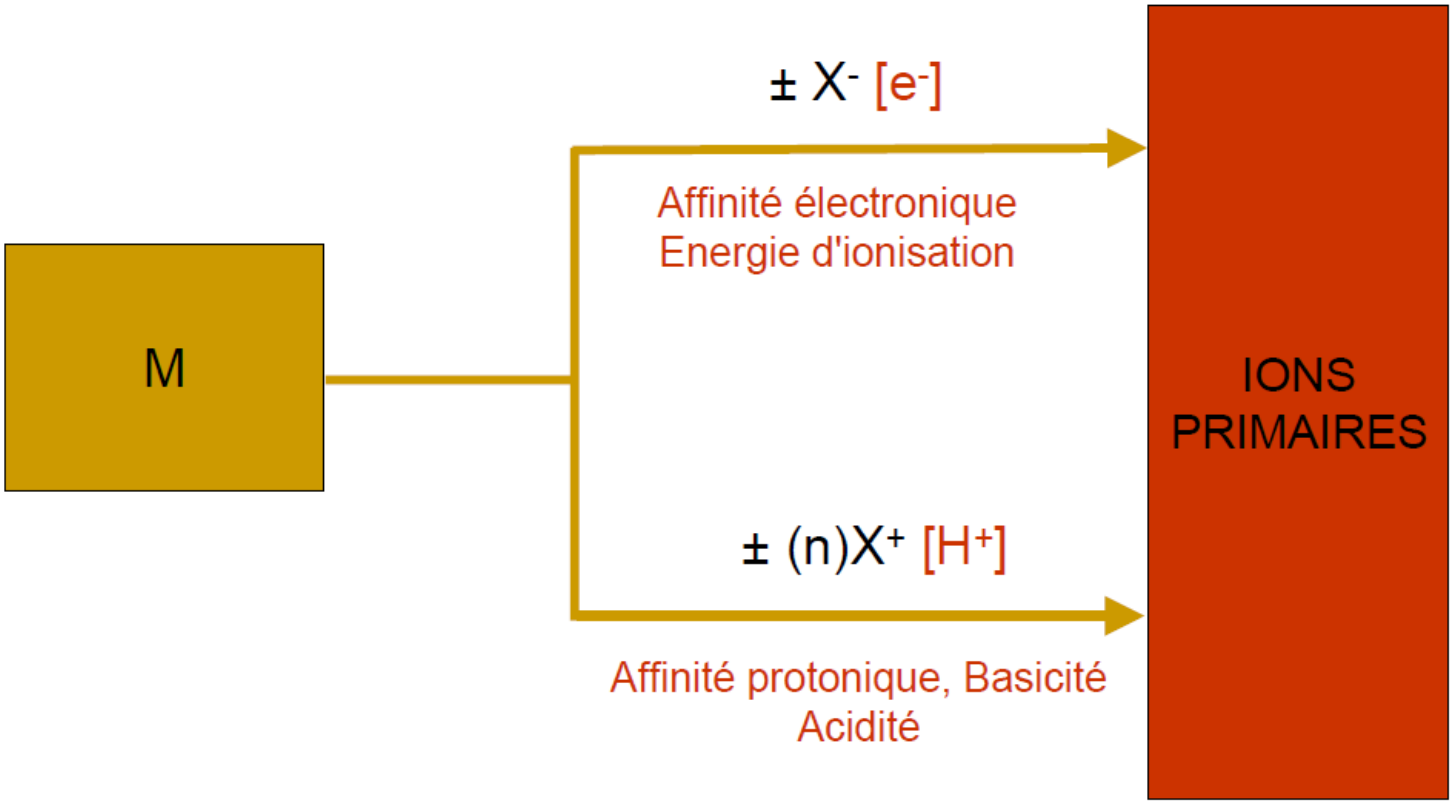


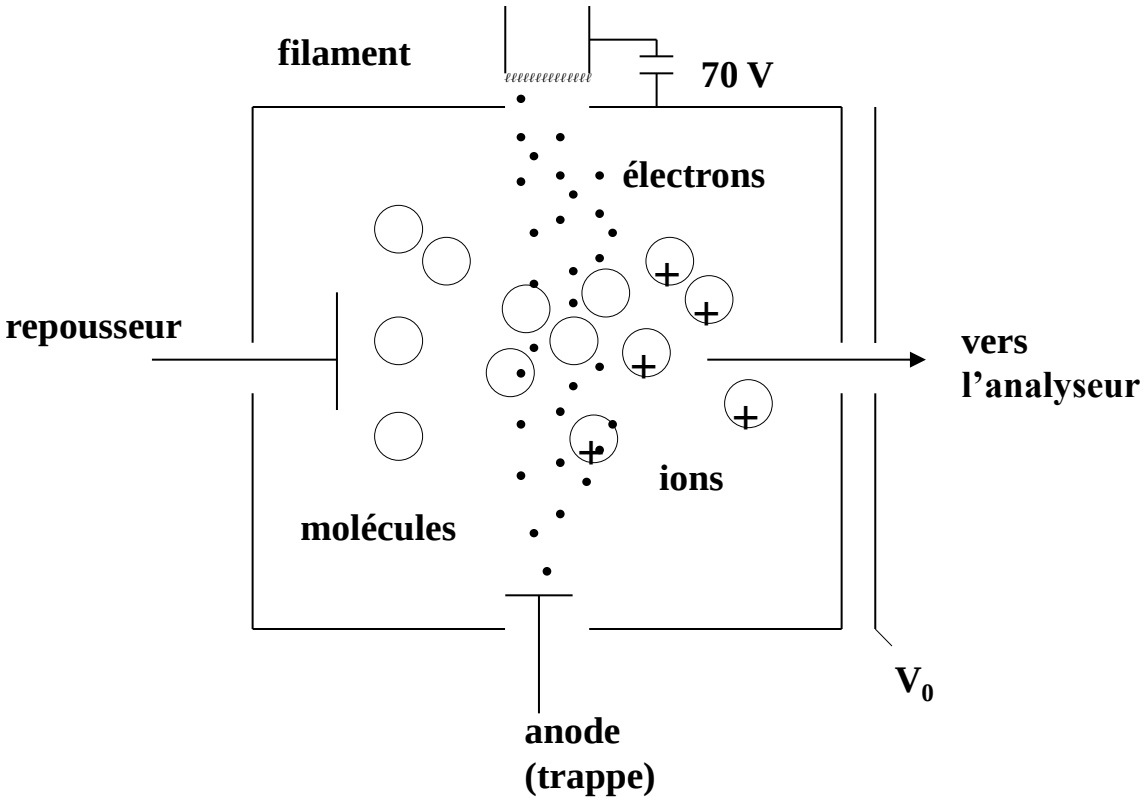
Figure 5. Comparison of 2.1 and 1 mm i.d. scale separations using the optimized system in (a) ESI(+): sulfaguanidine (1), acetaminophen (2), caffeine (3), hippurate (4), leucine encephalin (5), sulfadimethoxine (6), and verapamil (7). (b) ESI(−): hippurate (1), leucine enkephalin (2), sulfadimethoxine (3), and cholic acid (4). The same gradient scaled from 0.5 to 0.11 mL/min and the same injection volume used for experiments. Reproduced from Gray, N.; Lewis, M.R.; Plumb, R.S., Wilson, I.D.; Nicholson, J.K., *J. Proteome Res.* **2015**, *14*, 2714−2721 (ref 91). Copyright 2015 American Chemical Society.

Source d'ionisation et polarité

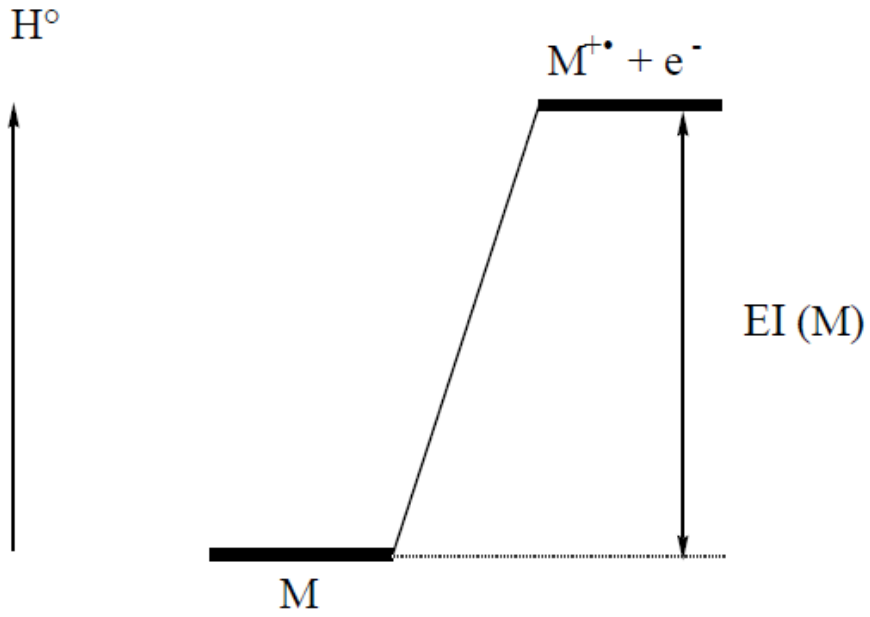


Impact électronique

- $M + e^- \rightarrow M^{+\cdot} + 2\ e^-$



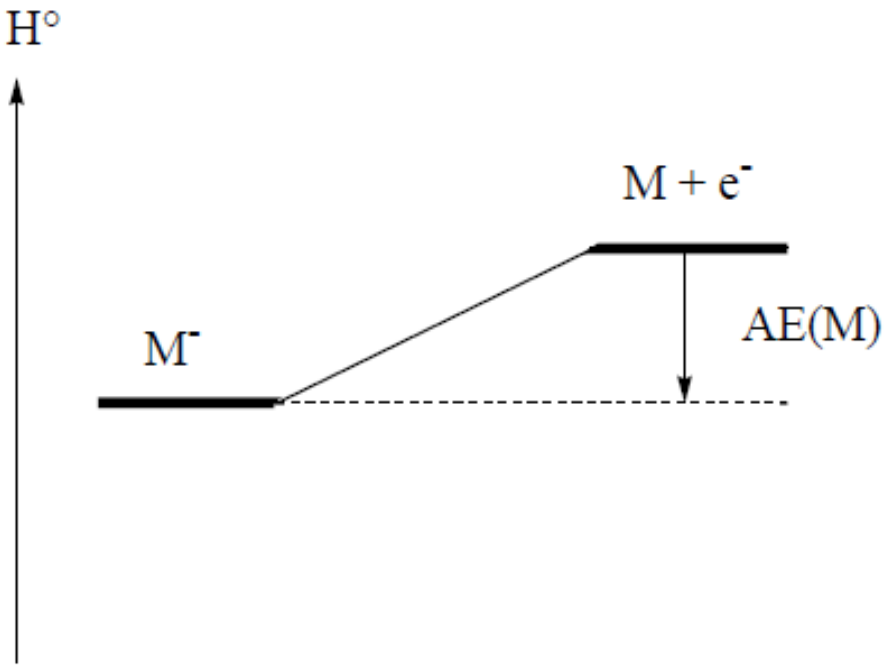
Impact électronique



EI ~ 7 à 12 eV
(700-1200 kJ/mol)

- L'énergie d'ionisation dépend des électrons de la HOMO:
  - Type d'orbitale: σ, π, n
  - Électronégativité de l'atome porteur
- Et de la stabilité de l'ion formé

Impact électronique



AE ~ -1 à +3 eV
(-100 à +300 kJ/mol)

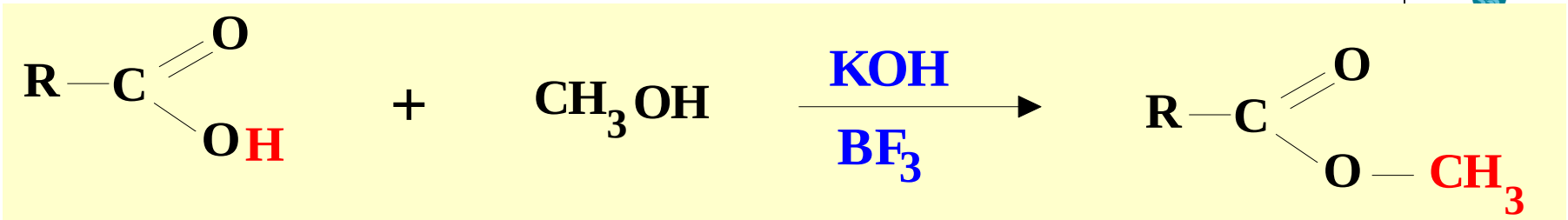
Affinité électronique positive, l'ion M^- est plus stable que $M + e^-$

AE(M) augmente lorsque M présente:

- des substituants attracteurs d'électrons
- un système d'e π largement délocalisés

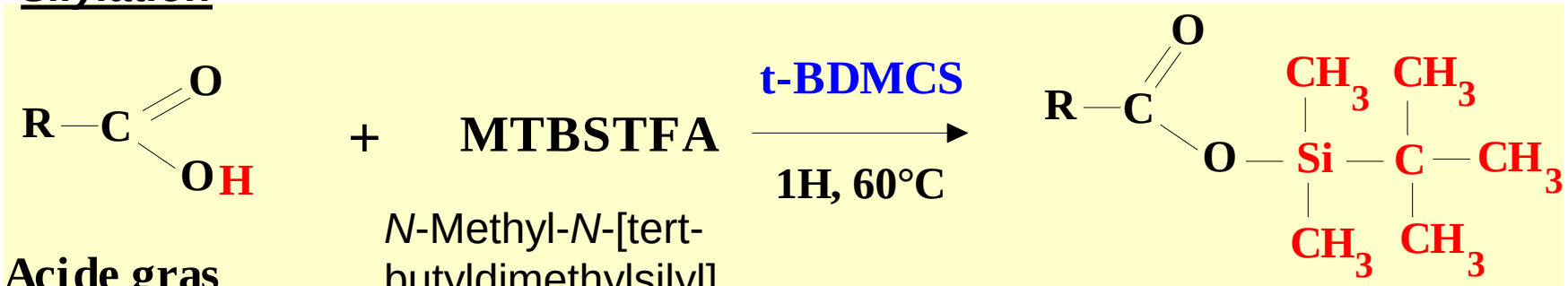
Impact électronique

Esterification



Acide gras **Ester méthylique**

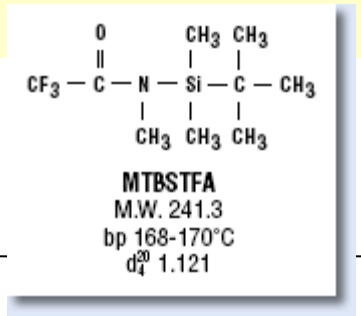
Silylation



Acide gras

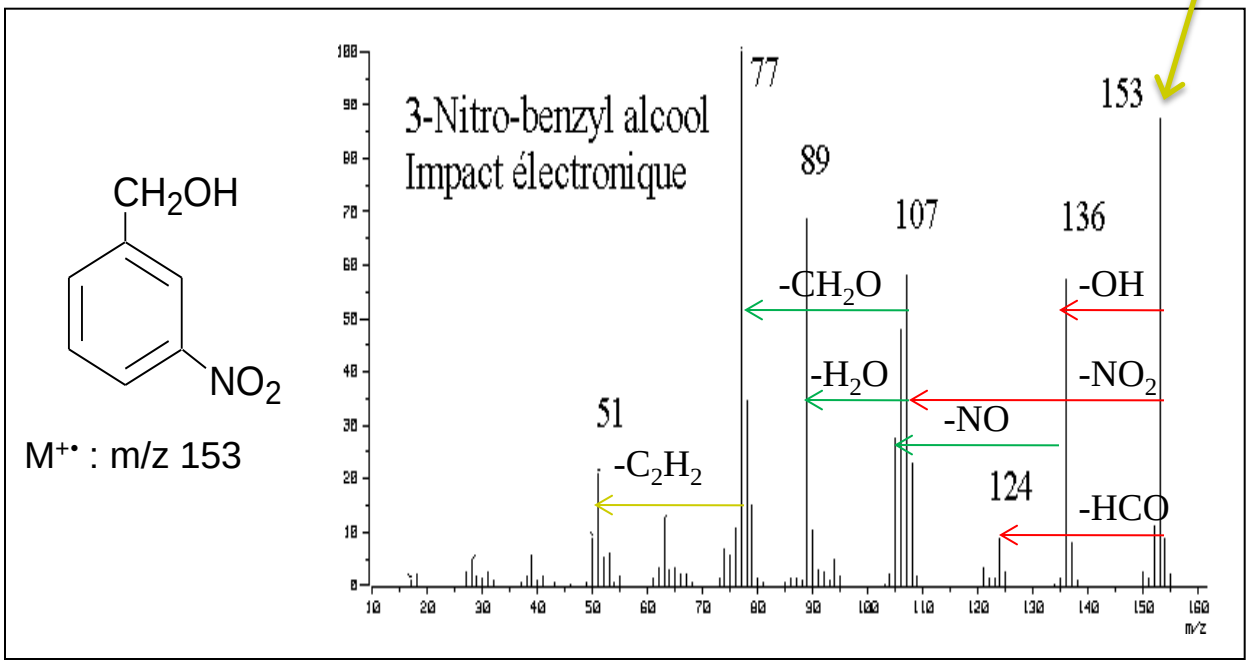
N-Methyl-*N*-[tert-butyldimethylsilyl] trifluoroacetimide

Dérivé t-BDMS



Impact électronique

Un spectre impact électronique 70 eV



Banque de données en IE

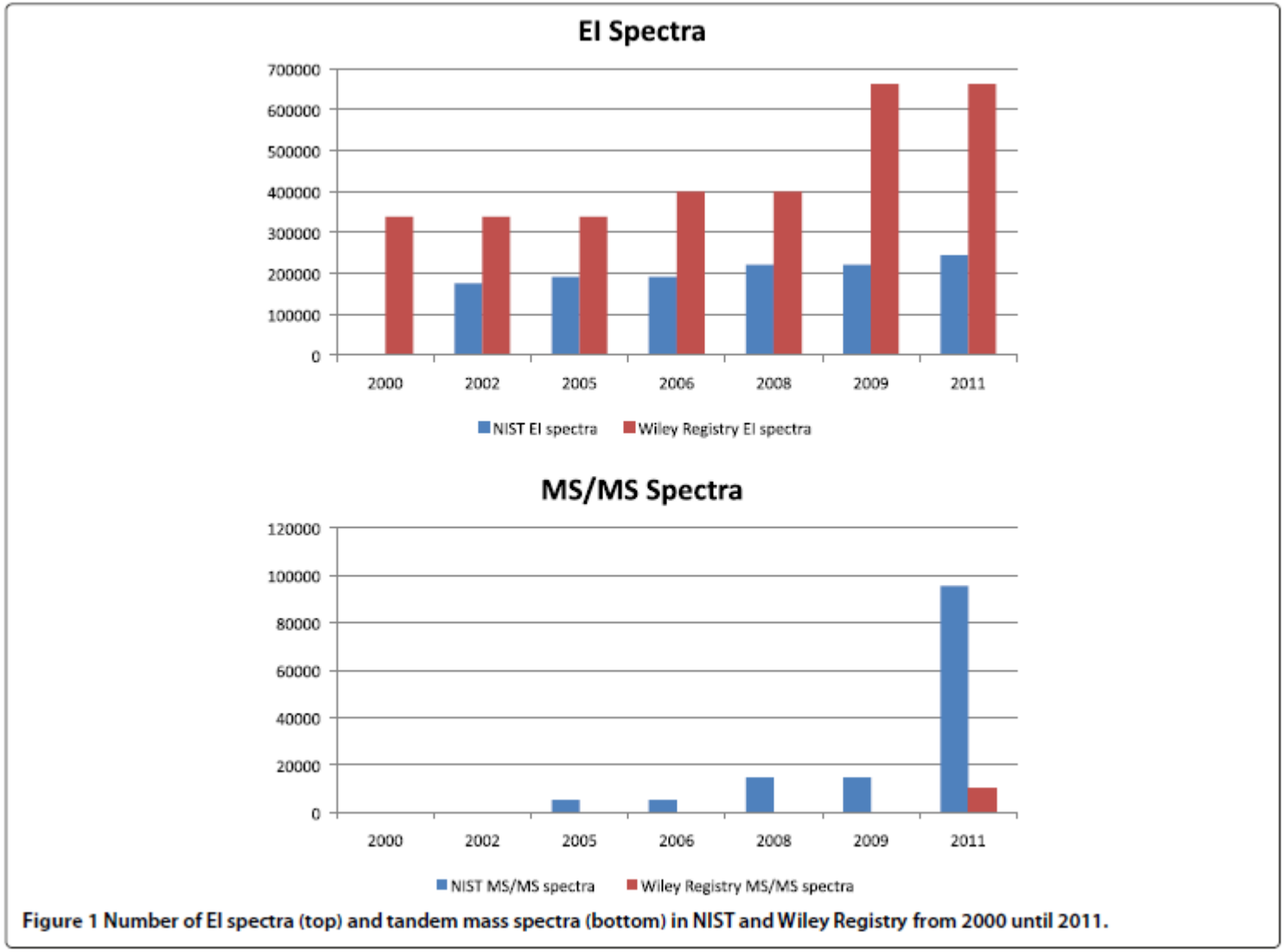


Figure 1 Number of EI spectra (top) and tandem mass spectra (bottom) in NIST and Wiley Registry from 2000 until 2011.

Scheubert *et al. Journal of Cheminformatics* 2013, **5**:12

GC-IE-MS

Un outil extrêmement robuste et fiable pour la métabolomique mais un accès restreint au métabolome …

GC-IE-HRMS

Performance Characteristics	
Resolving Power:	100,000 @ *m/z* 272
Mass Range:	50 to 3,000 *m/z*
Scan Rate:*	Up to 18 Hz at resolution setting of 12,500 @ *m/z* 272
Mass Accuracy:**	Internal: <1 ppm RMS External: <3 ppm RMS
Quantitative Dynamic Range*:	$>10^6$
In-Spectrum Dynamic Range:	>5000:1
Multiplexity:	Up to 10 precursors/scan

* *Under defined conditions*

** *Under conditions defined in 1 µL, 100 fg/µL octafluoronaphthalene EI Full MS installation specification*

Resolution	17,500	35,000	70,000	140,000
Mass Analysis Time (ms)	80	150	300	700
Scan Rate (Hz)	12	7	3	1.3

Ionisation chimique

- *Première* étape: Ionisation d'un gaz réactif (exemple: l'ammoniac):

  $NH_3 + e^- \rightarrow NH_3^{+\cdot} + 2\ e^-$

  $NH_3^{+\cdot} + NH_3 \rightarrow NH_4^+ + NH_2^{\cdot}$

- *Deuxième étape*: protonation de la molécule M par l'ion ammonium

  $M + NH_4^+ \rightarrow (M \ldots\ldots H \ldots\ldots NH_3)^+ \rightarrow MH^+ + NH_3$

  ***m/z* 18** ***m/z* M+18** ***m/z* M+1**

- Lorsque l'on utilise un gaz réactif GH+, la réaction qui se produit avec la molécule M peut s'écrire de la façon suivante:

  $M + GH^+ \rightarrow MH^+ + G$

  $GH^+ \rightarrow G + H^+$ $\quad \Delta H°(G) = AP(G)$

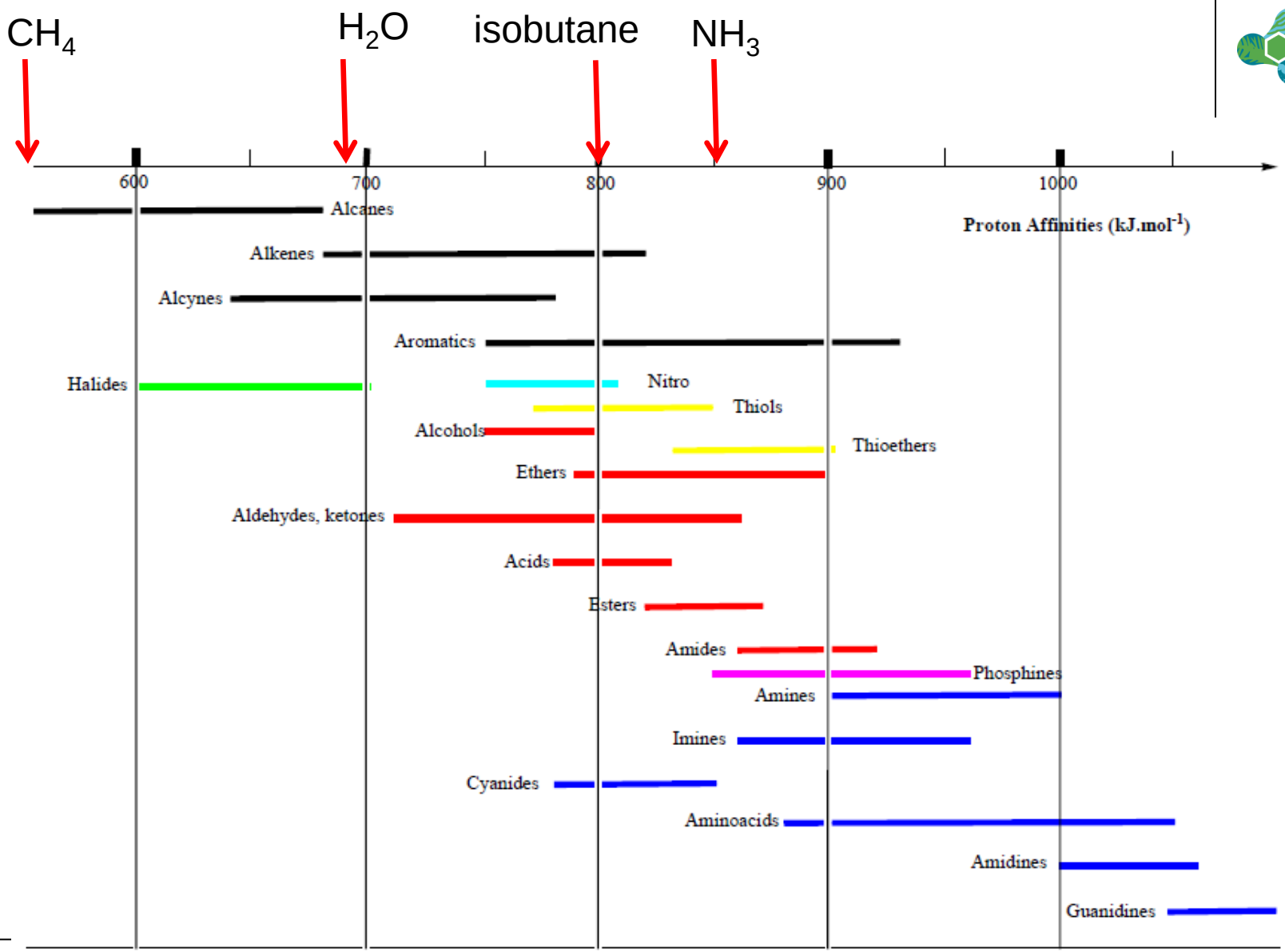
  $MH^+ \rightarrow M + H^+$ $\quad \Delta H°(M) = AP(M)$

- Les enthalpies de ces deux réactions s'appellent "affinités protoniques" (respectivement du gaz et de la molécule à analyser); elles correspondent au terme enthalpique de la basicité en phase gazeuse.

- La réaction de protonation a lieu lorsque AP(G)< AP(M)

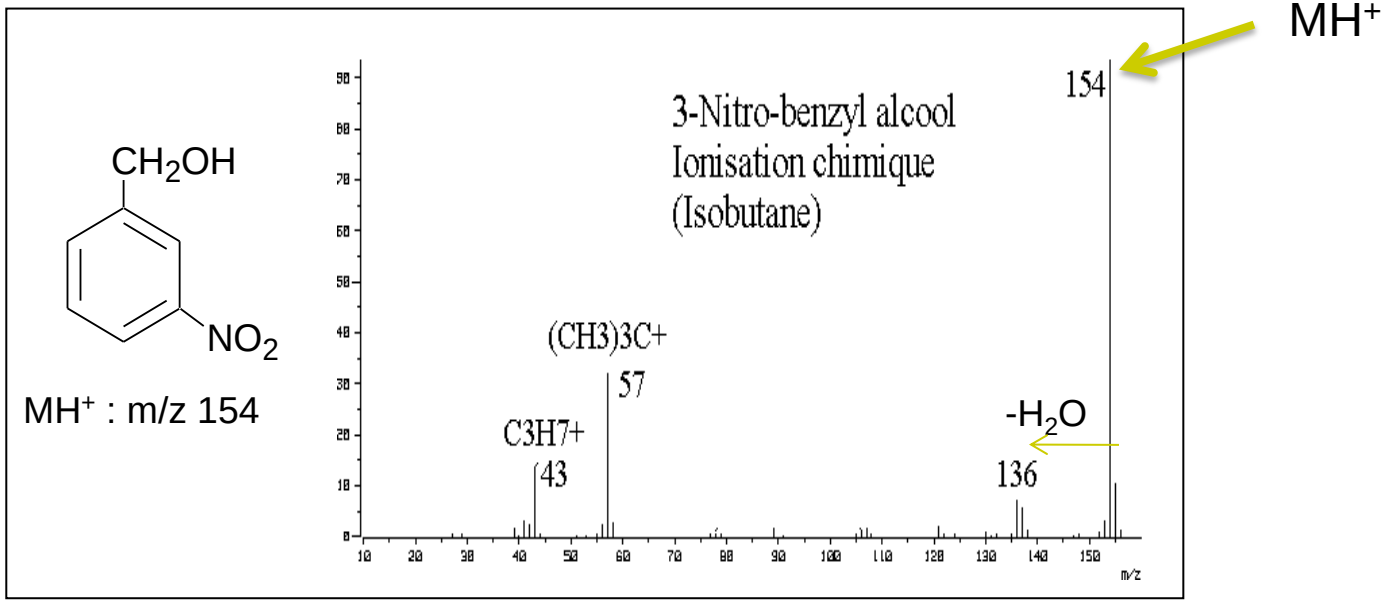
- Concrètement, on choisira un gaz réactif adapté à la molécule étudiée (de plus faible AP), en général dans le but de limiter les phénomènes de fragmentation.

Ionisation chimique

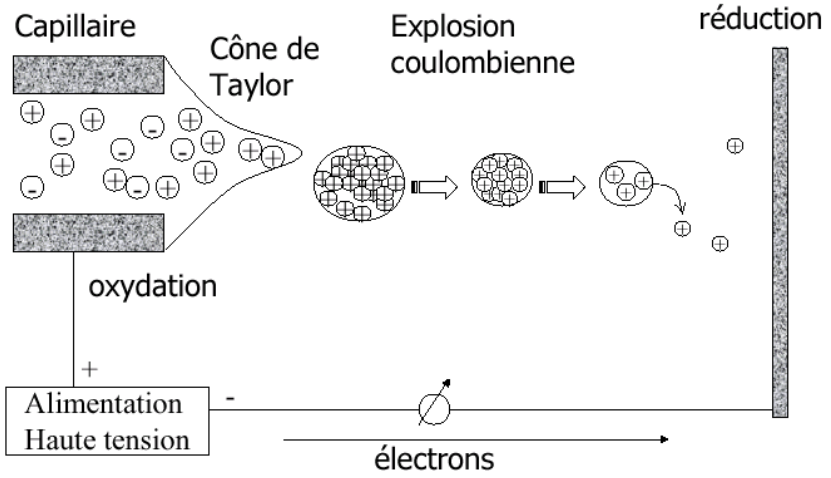


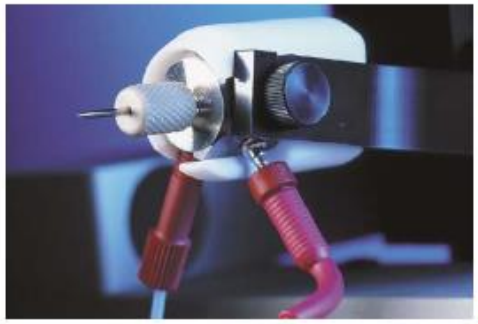
Ionisation chimique

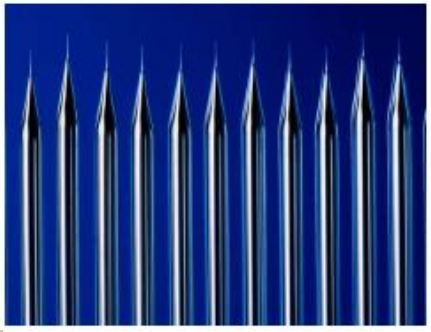
Un spectre IC positive

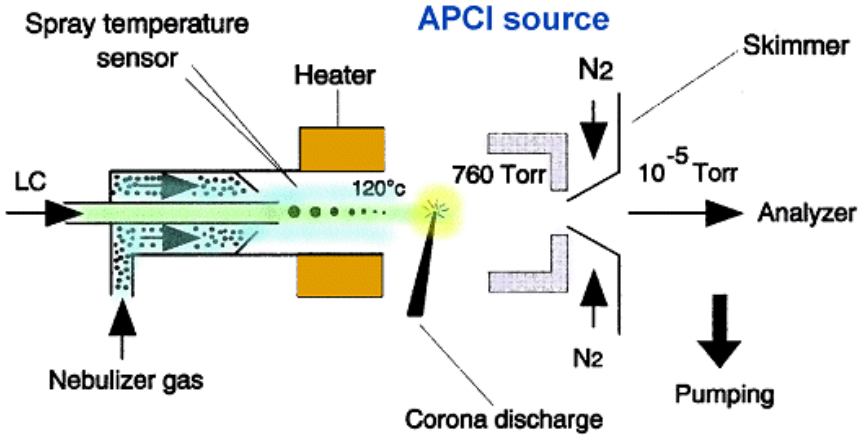


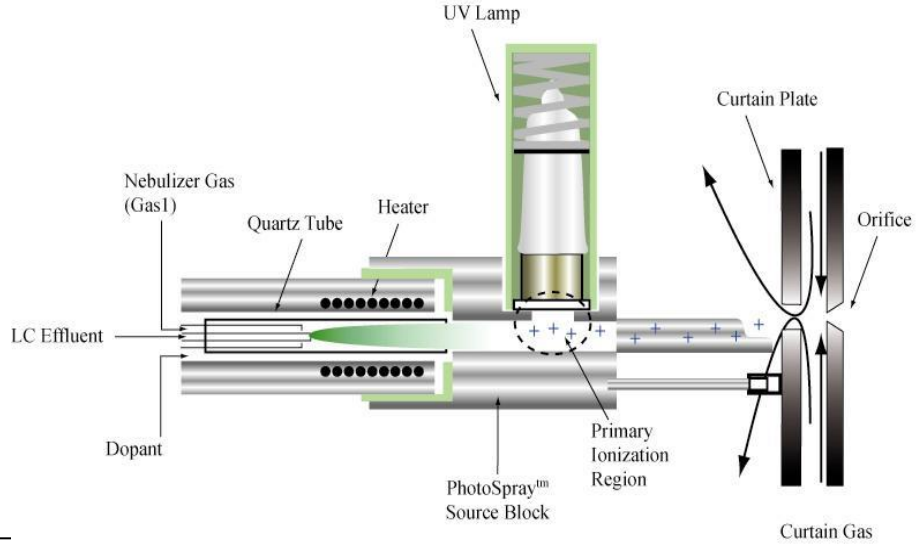
Source à pression atmosphérique











Une vue non globale du métabolome

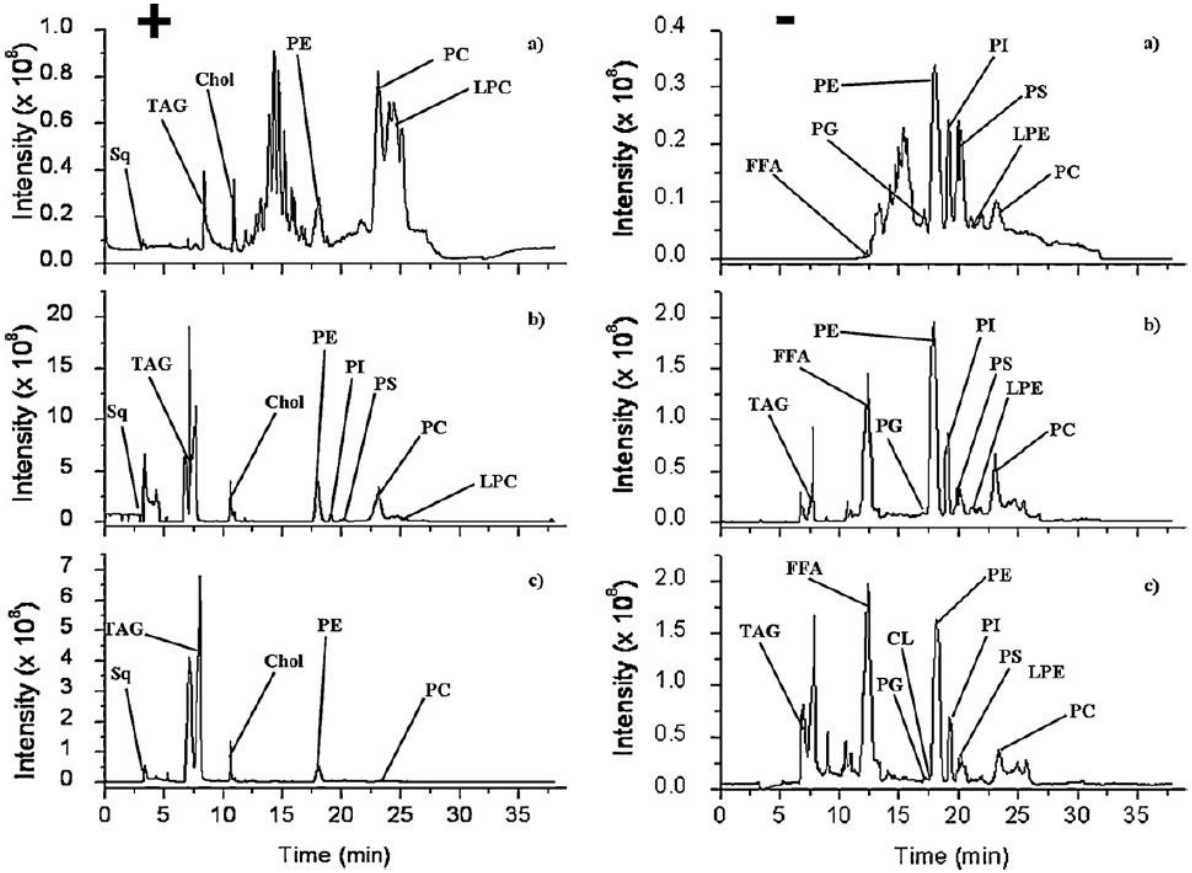


Fig. 4. Chromatograms obtained with ESI (a), APCI (b) and APPI (c) in positive (left) and negative (right) ion mode on a *Leishmania donovani* HePC-RT sample.

Phase A: n-heptane/ethyl acetate (99.8:0.2, v/v),
Phase B : acetone/ethyl acetate (2:1; v/v) +0.02% acetic acid,
Phase C : propan-2-ol/water (85:15, v/v) + 0.05% acetic acid and ethanolamine.

Imbert L, Gaudin M, Libong D, Touboul D, Abreu S, Loiseau PM, Laprévote O, Chaminade P.
J Chromatogr A. 2012 Jun 15;1242:75-83.

Une vue non globale du métabolome

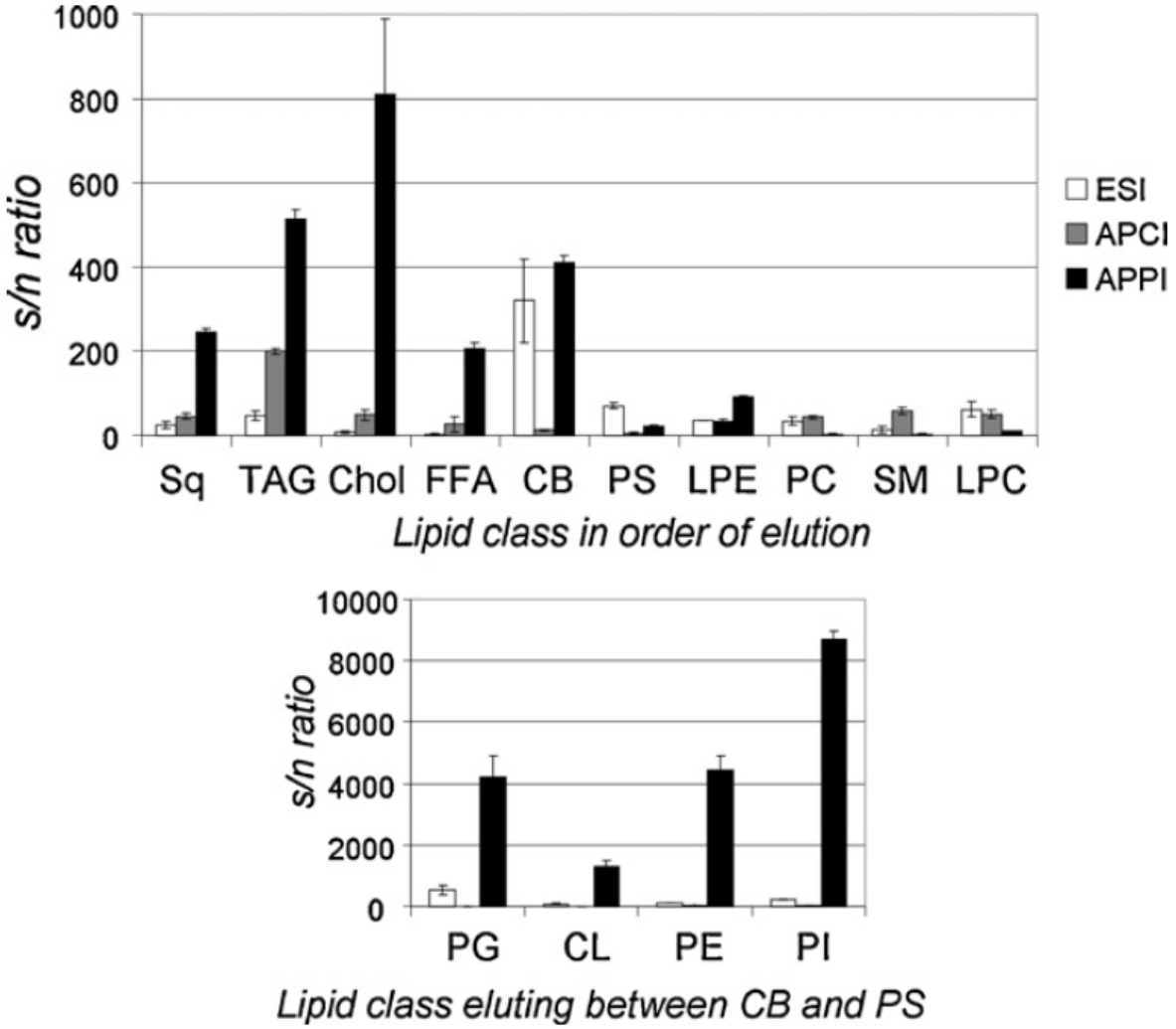


Fig. 2. Comparison of the signal-to-noise ratio of lipid ion peaks from ESI, APCI and APPI analyses of a 100 µg mL^{-1} mixture of lipids standards ($n = 3$).

Imbert L, Gaudin M, Libong D, Touboul D, Abreu S, Loiseau PM, Laprévote O, Chaminade P. J Chromatogr A. 2012 Jun 15;1242:75-83.

Une vue non globale du métabolome

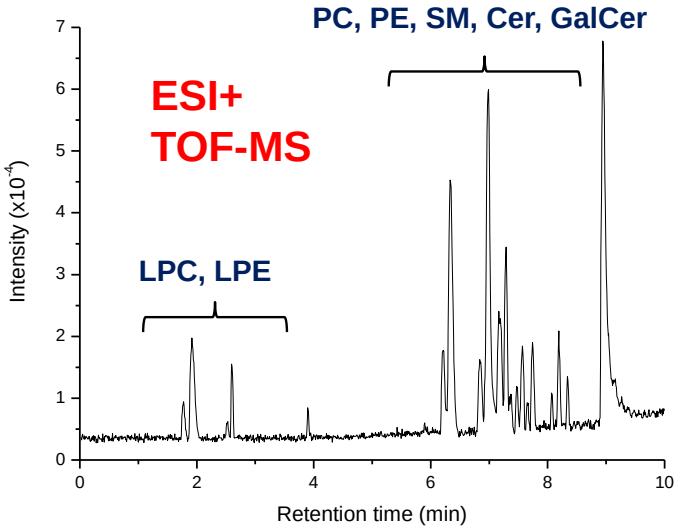


Fig 1 : Positive ESI UPLC-MS lipid profile of a control brain sample

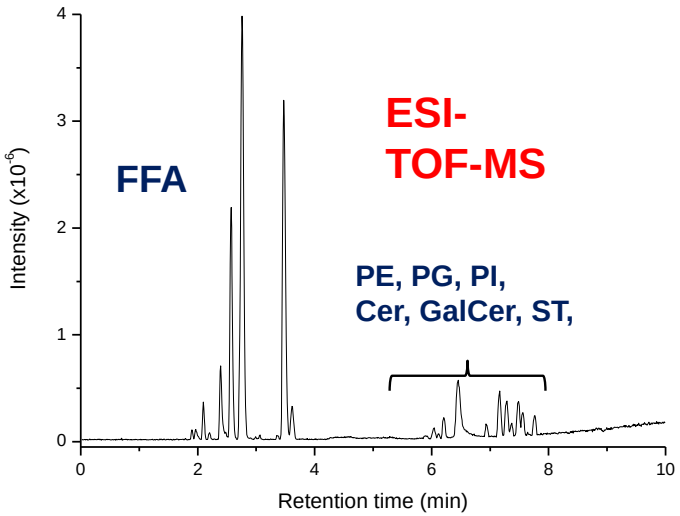


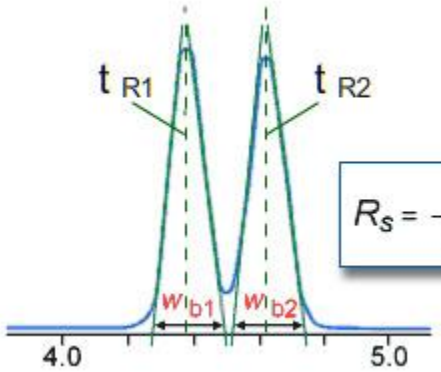
Fig 2 : Negative ESI UPLC-MS lipid profile of a control brain sample

Solvent A : H_2O/ACN 60/40 + 10 mM $AcONH_4$
Solvent B : ACN/iPrOH 50/50 + 10 mM $AcONH_4$

Le spectromètre de masse

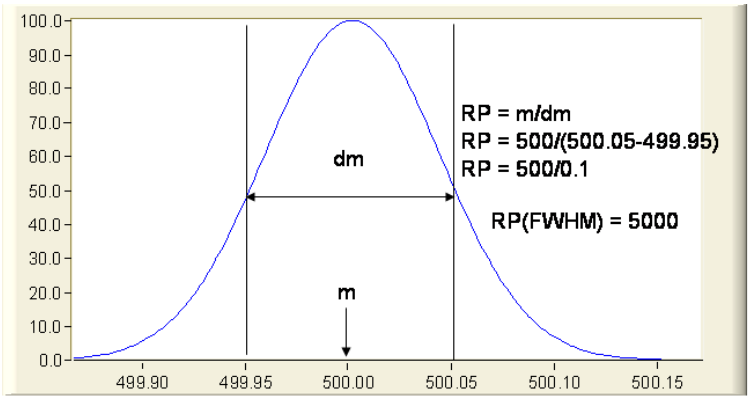
- Résolution spectrale $R = \frac{\lambda}{\Delta\lambda}$

- Résolution chromatographique

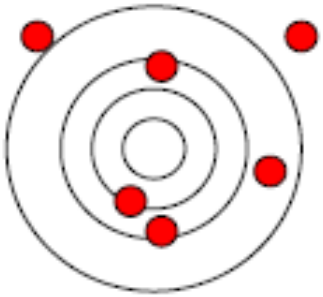


$$R_s = \frac{(t_{R2} - t_{R1})}{(w_{b1} + w_{b2})/2} = \frac{2(t_{R2} - t_{R1})}{w_{b1} + w_{b2}}$$

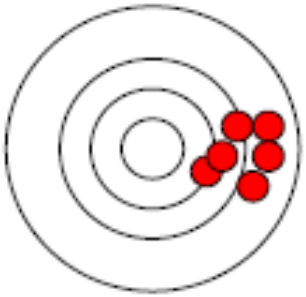
- Résolution en masse



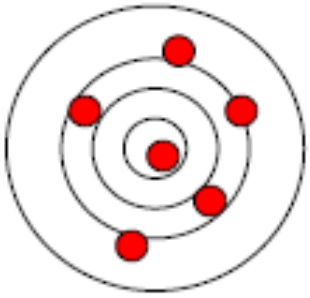
Le spectromètre de masse



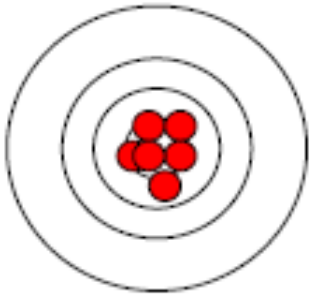
Ni juste, ni fidèle



Pas juste mais fidèle



Juste mais pas fidèle



Juste et fidèle

Il est donc impératif de bien étalonner un spectromètre de masse pour obtenir des mesures en haute précision en masse (HRMS)!!!

Le spectromètre de masse

Élément	Isotope	Masse isotopique	Rapport isotopique
H	1	1,00782503207	0,999885
	2	2,0141017778	0,000115
C	12	12,0000000000	0,9893
	13	13,0033548378	0,0107
N	14	14,0030740048	0,99636
	15	15,0001088982	0,00364
O	16	15,99491461956	0,99757
	17	16,99913170	0,00038
	18	17,9991610	0,00205

Défaut de masse!

Le spectromètre de masse

Élément	Isotope	Masse isotopique	Rapport isotopique
Cl	35	34.96885268	0.7576
	37	36.96590259	0.2424
Br	79	78.9183371	0.5069
	81	80.9162906	0.4931
Pt	190	189.959932	0.00014
	192	191.9610380	0.00782
	194	193.9626803	0.32967
	195	194.9647911	0.33832
	196	195.9649515	0.25242
	198	197.967893	0.07163

Le spectromètre de masse

- Basse résolution: quadripôle / trappe ionique (2D ou 3D)
- Haute résolution: TOF, Orbitrap, FT-ICR
- Et toutes les combinaisons de ces techniques !

Un seul impératif: il faut que le spectromètre de masse ait une vitesse de scan compatible avec le système chromatographique (identification *versus* quantification)

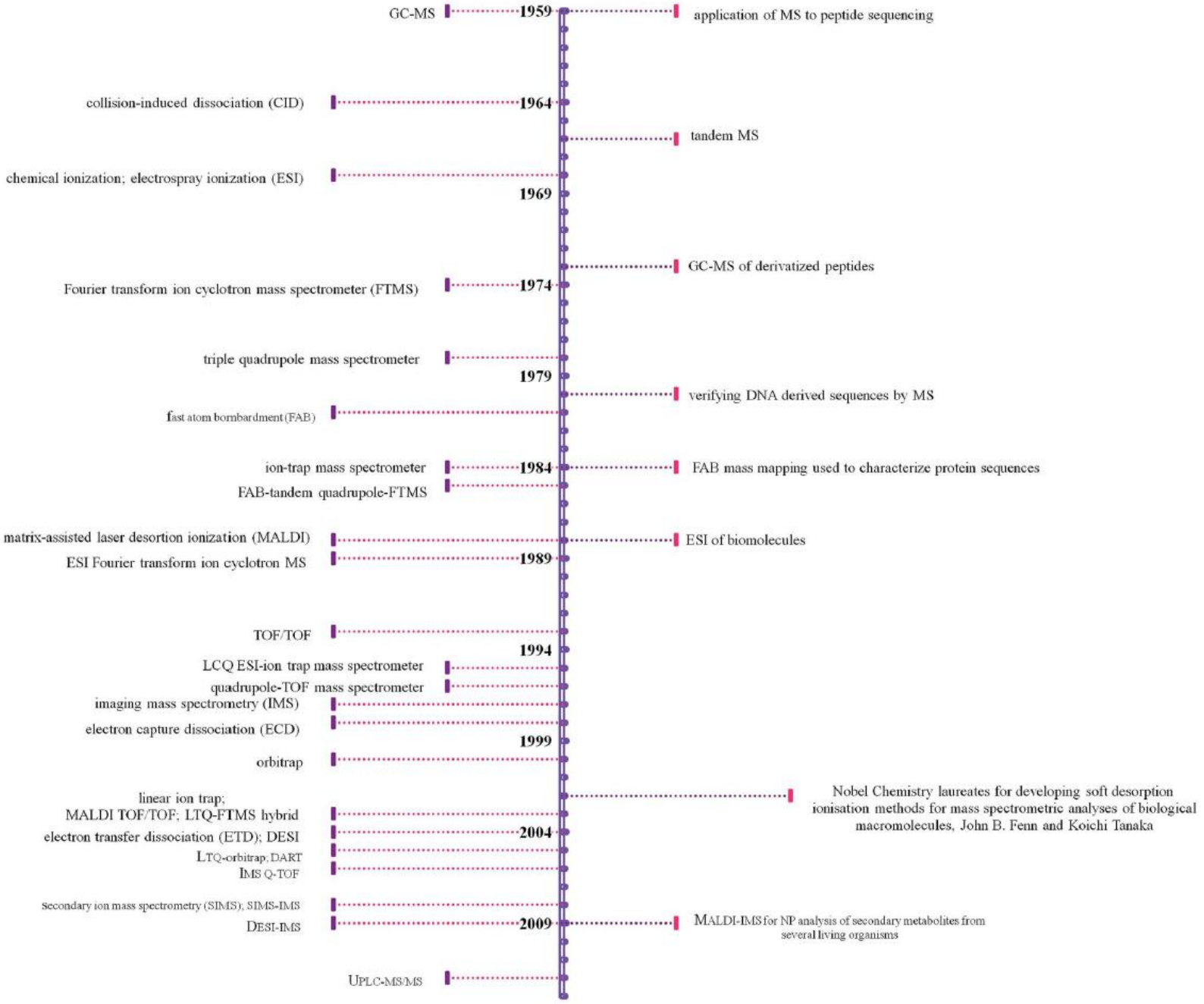


Fig. 7 Timeline illustrating the major advances in MS hardware and methodologies, period 1959–2012.

Workflow en métabolomique

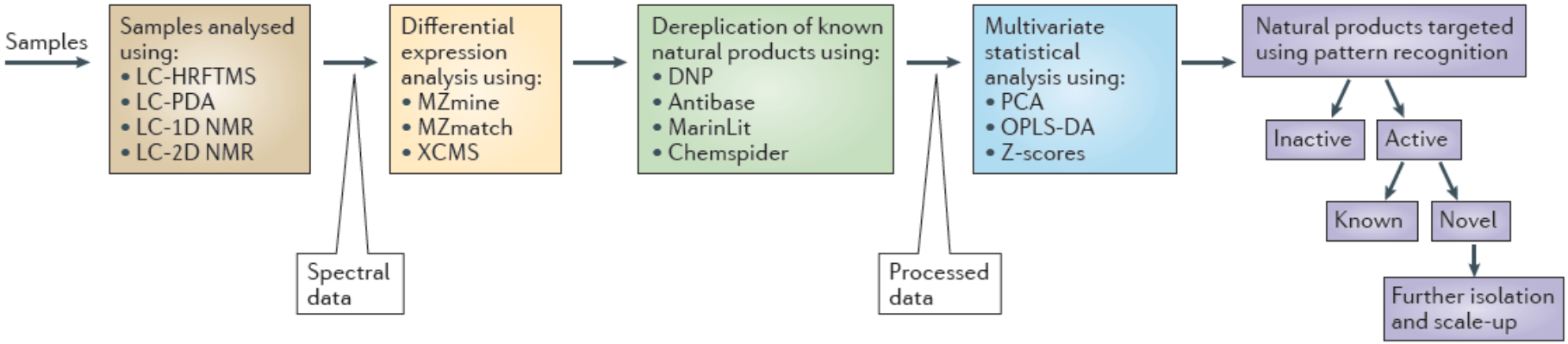


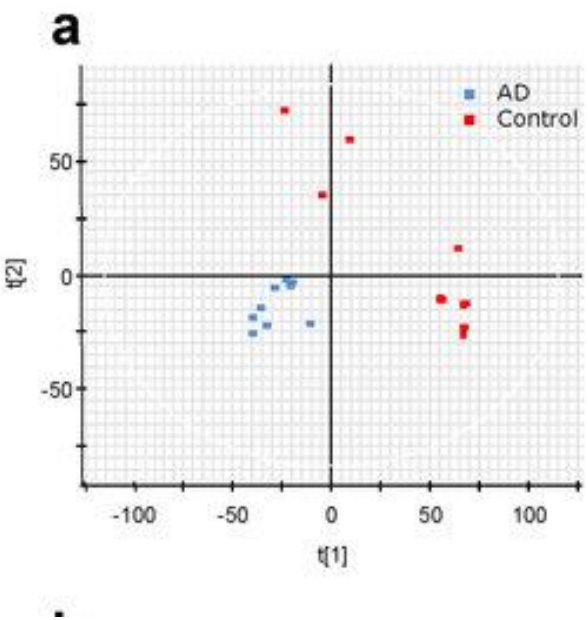
Figure 2 | **Metabolomics data workflow in natural product research.** Samples are submitted to LC–HRFTMS (liquid chromatography–high-resolution Fourier-transform mass spectrometry), LC-PDA (liquid chromatography–photodiode array) and LC-1D/2D NMR (liquid chromatography-one dimensional/two dimensional NMR spectroscopy) analysis. The mass spectrometry data are further processed using differential expression analysis software such as MZmine, MZmatch and XCMS. This software is coupled to databases such as the Dictionary of Natural Products (DNP), AntiBase, or MarinLit to dereplicate known natural products against the novel secondary metabolites. Pre-collected LC-PDA and LC-1D/2D NMR data confirm the dereplication results. The processed data are subjected to multivariate analysis using both PCA (principal component analysis) and/or OPLS-DA (orthogonal partial least squares discriminant analysis). The results are then plotted in S-plots and heat maps. Through pattern recognition, inactive versus active and known versus novel natural products are sorted to define the natural products that will be targeted for further isolation and scale-up work.

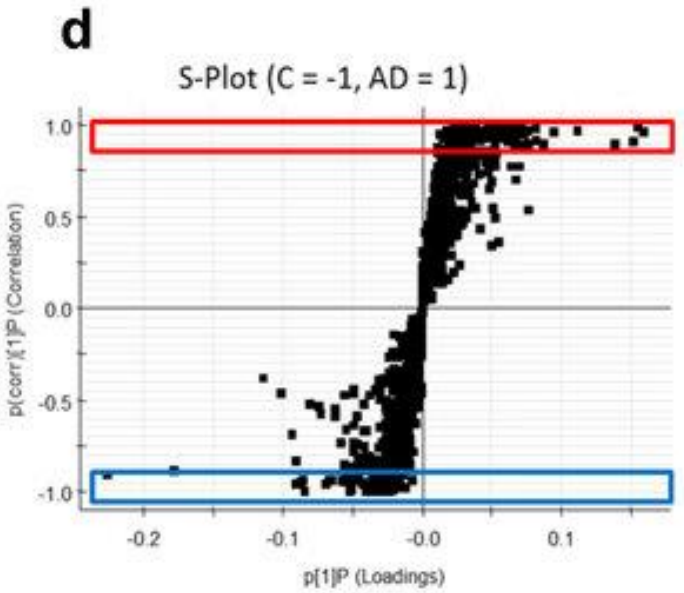
doi:10.1038/nrd4510

Comparaison de deux populations: analyse statistique univarié et multivarié

- Univariée: une seule variable comme observable

- Multivariée: prise en compte de tous les paramètres





Annotation de pics

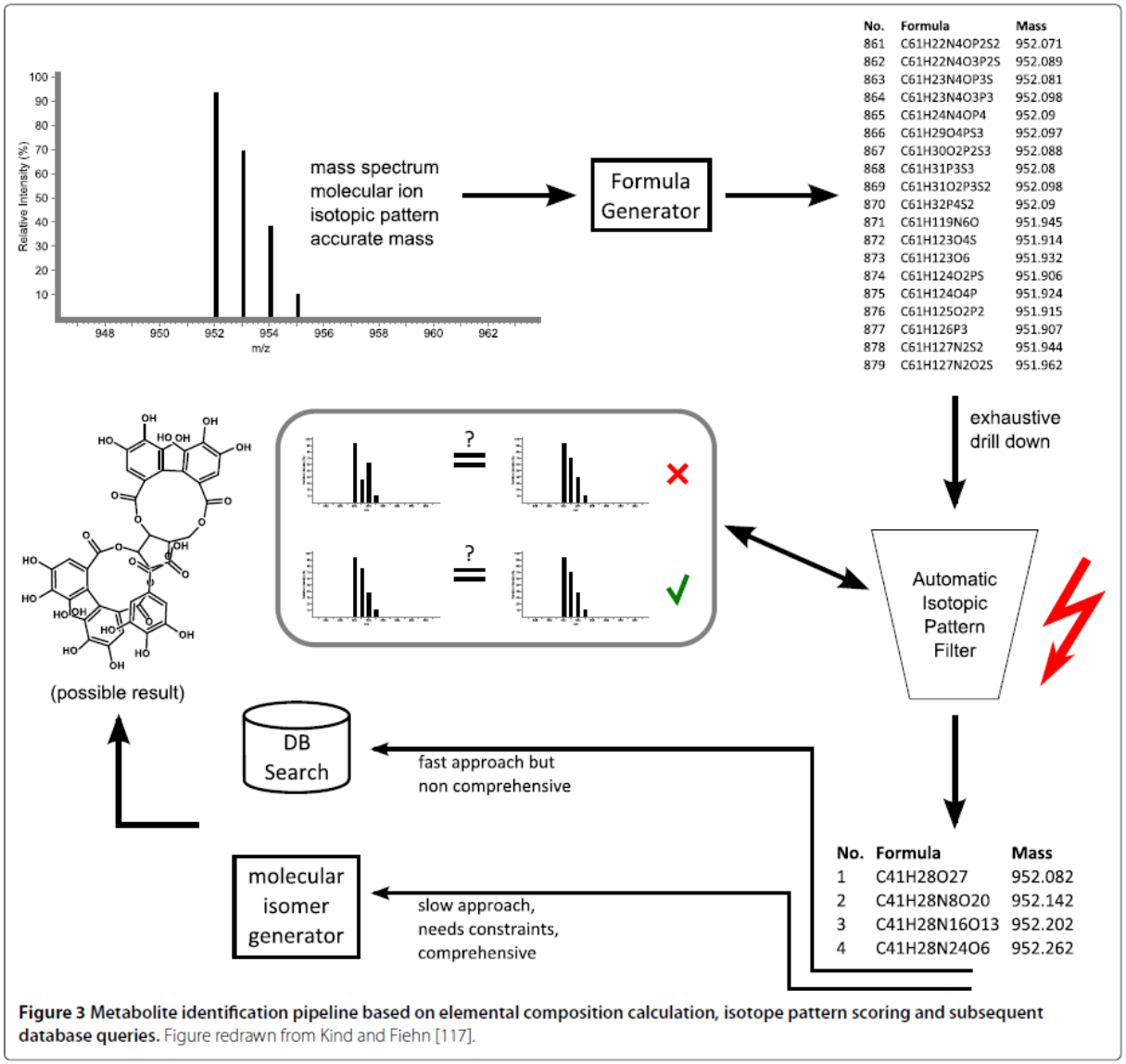


Figure 3 Metabolite identification pipeline based on elemental composition calculation, isotope pattern scoring and subsequent database queries. Figure redrawn from Kind and Fiehn [117].

Annotation de pics

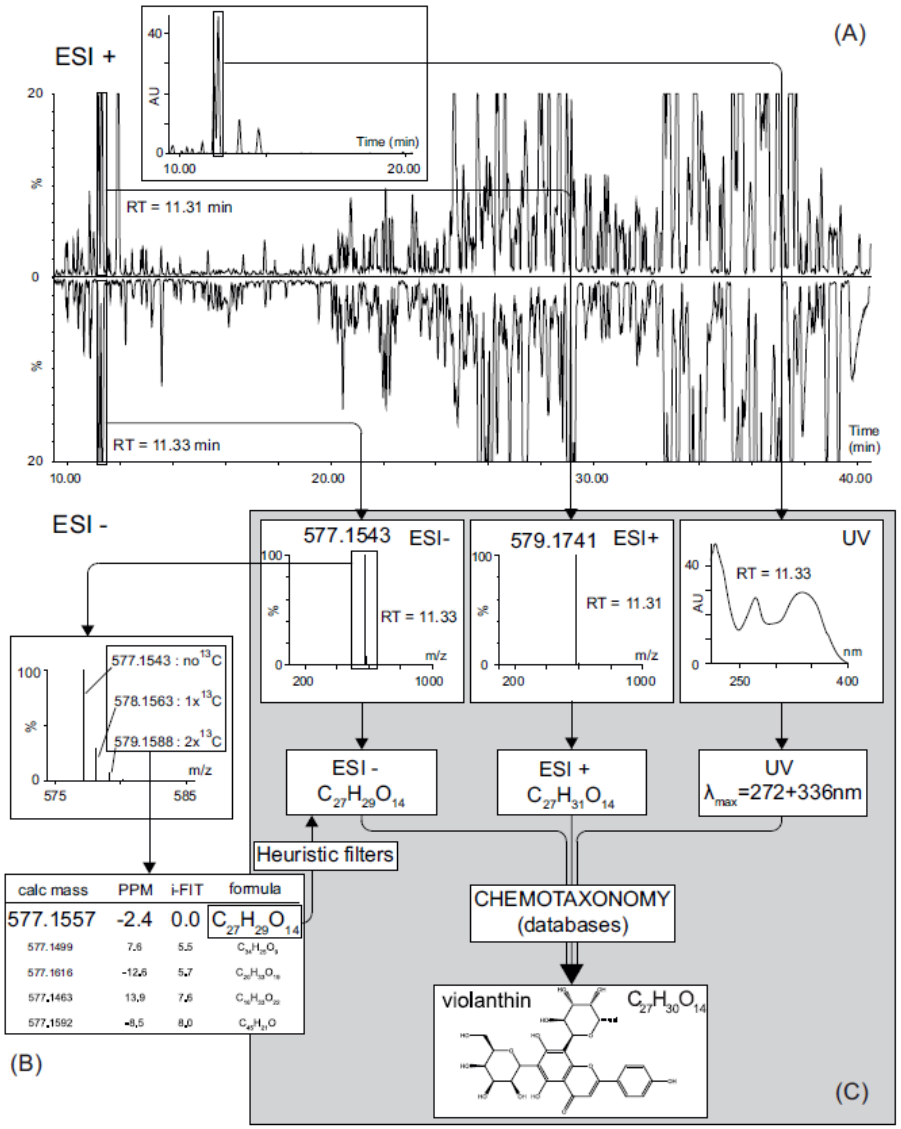
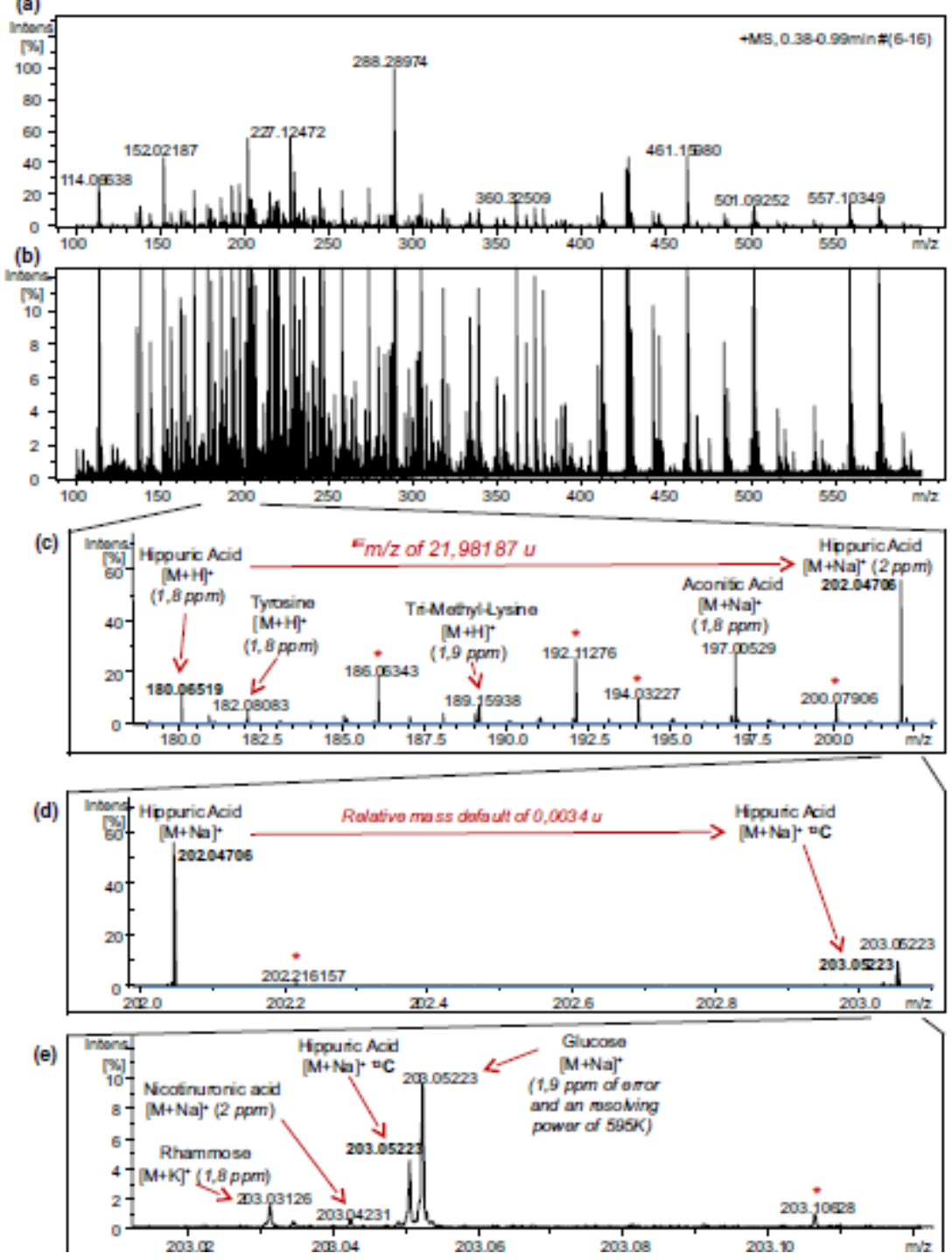


Fig. 6. Example of LC–MS peak annotation based on a high resolution *Viola tricolor* profiling on a C_{18} UHPLC column (150 mm × 2.1 mm; 1.7 µm) obtained with a slow gradient (5–95% ACN in 50 min). (A) PI (upper trace) and NI (lower trace) ESI-TOFMS BPI chromatograms and a UV trace (366 nm) is displayed in the inset. (B) Putative molecular formulas assignment based on the 15 ppm precision and isotopic pattern (iFIT) obtained from the NI ESI-TOFMS spectrum of the LC peak at RT 11.33 min. Application of heuristic filtering enable to ascertain the molecular formula assignment [237] (C) The LC peak at RT 11.33 min is annotated based on PI and NI molecular formula assignment and the UV PDA spectrum. Final structural assignment is based on a cross search with chemotaxonomic information which considerably reduces the number of possibilities. Such an approach for metabolite identification is still ambiguous (level of ID 2 according to MSI [4]) and assume that the metabolite has been previously characterised. Adapted from [41] with permission of The Royal Society of Chemistry.

ation de pics



Annotation de pics

Table 1 Software for the three basic steps of molecular formula identification using isotope patterns

Decomposing monoisotopic peaks	
Decomp [100,101]	for arbitrary alphabets of elements
	requires only little memory
	swift in practice
SIRIUS [102,103]*	implementing *Decomp* approach for MS
	decomposing real-valued masses
"Seven Golden Rules" [104]	to filter molecular formulas
Simulating isotope patterns	
IsoPro [105]	multinomial expansion to predict "center masses"
	memory- and time-consuming
Mercury [106]	pruning by probability thresholds and/or mass range
	reduced memory and time consumption
	reduced accuracy of the predictions
Emass [107]* & *SIRIUS* [102]*	iterative (stepwise) computation of isotope pattern
	probability-weighted center masses
	probabilities and masses are updated as atoms are added
IsoDalton [108]	models the folding procedure as a Markov process
BRAIN [109]*	Newton-Girard theorem and Vietes formulae to calculate intensities and masses
Fourier [110]*	2D Fast Fourier Transform that splits up the calculation in a coarse and a fine structure
	running time improvement for large compounds
Scoring candidate compounds	
SigmaFit	commercial software by Bruker Daltonics
SIRIUS [102]*	Bayesian statistics for scoring intensities and masses of the isotope pattern
MZmine [111]	simple scoring based only on intensities

*Recommended tools.

Scheubert *et al. Journal of Cheminformatics* 2013, **5**:12

Banque de données

Table 1 Essential features of selected databases for NPs dereplication

Database	Compounds[a]		Period	MW	MF	UV[b]	NMR[c]	MS[d]	Bioactivity	Taxonomy	SSS[e]
	Total	NPs									
CAS/SciFinder	8.9×10^7	>283 000	Current	+	+	–	–	–	+	+	+
CSLS	4.6×10^7	Extracts	~2010	+	+	–	–	–	+	–	+
ChemSpider	3.2×10^7	>7800	Current	+	+	+	–	–	+	–	+
PubChem	5.1×10^7	>438 00	Current	+	+	–	–	–	+	–	+
ZINC	3.4×10^7	>19 000	Current	+	–	–	–	–	+	–	+
NAPROC-13		>6000	~2007	+	+	–	+[c1,c2,c3]	–	–	–	+
NMRShiftDB	42 000	?[f]	Current	+	+	–	+[c1,c2,c3]	–	–	–	+
Massbank	13 000	>2500	Current	+	+	–	–	+[d1,d2,d3]	–	–	+
ReSpect		>3595	Current	+	+	–	–	+[d1,d2,d3]	–	–	+
Metlin		64 000	Current	+	+	–	–	+[d1,d3]	–	–	+
GNPS	1.6×10^5	$>1.4 \times 10^5$	Current	+	+	–	–	+[d1,d3]	–	–	+
NaprAlert		>150 000 extracts	~2003[g]	+	+	+[h]	–	–	+	+	–
Dictionary NP		>260 000	Current	+	+	+	–	–	+	+	+
Dictionary MNP		25 000	Current	+	+	+	–	–	+	+	+
MarinLit		23 500	Current	+	+	+	+[c1,c2,c3]	–	+	+	+[h]
AntiBase		42 950	Current	+	+	+[h]	+[c1,h]	–	+	+	+
AntiMarin		53 000	2013[i]	+	+	+[h]	+[c1,c2,c3,h]	–	+	+	+[h]

[a] When possible an estimate number of NPs in the database is given. [b] λ UV data values. [c] Three NMR data options have been used: [c1] δ values (experimental or calculated), [c2] spectra or [c3] ^{1}H NMR structural features (^{1}H-SF). [d] Three MS data options have been used: [d1] positive, negative, and neutral MSn m/z-value, [d2] spectra or [d3] fragment ion (m/z). [e] Sub-structure searching. [f] NPs reported in the database, without numbers. [g] Only includes *ca.* 15% of the literature from 2004 to present time. [h] Partial data only. [i] Is the result of a merger between AntiBase (a database of all terrestrial and marine microbial natural products) and MarinLit (a database of marine natural products) that finished in 2013.

Banque de données

Table 2 | **Natural-product databases that can be used for virtual screening campaigns**

Database	Number of entries	Additional information	Refs
Super Natural II	355,000	2D structures; vendor information for over 215,000 compounds	*
Universal Natural Product Database	197,201	3D structures assembled from other available Chinese databases	289
Chinese Natural Product Database	53,000	Has been used in a virtual screen for PPAR-γ agonists	290
Drug Discovery Portal	40,000	Not all natural products, but all based on available samples	49
iSMART	20,000	Based on components from traditional Chinese medicines	291, 292
Database from historical medicinal plants, DIOS	6,702	Successfully used in several virtual screening campaigns	293
AfroDb	1,000	Compounds from African medicinal plants	294
NuBBE	640	Compounds from Brazilian sources	295‡

2D, two-dimensional; 3D, three-dimensional; iSMART, integrated systems biology-associated research with traditional Chinese medicine; PPAR-γ, peroxisome proliferator-activated receptor-γ. *See the Super Natural II database. ‡See the NuBBE database.

doi:10.1038/nrd4510

Comment identifier ?

- Spectrométrie de masse tandem (MS/MS, principalement CID ou HCD pour les métabolites)

- Spectroscopie optique non destructive (UV-Vis, fluorescence…)

- Isolement et RMN

Quid des énergies de collision et de la transposition des spectres MS/MS par collision ?

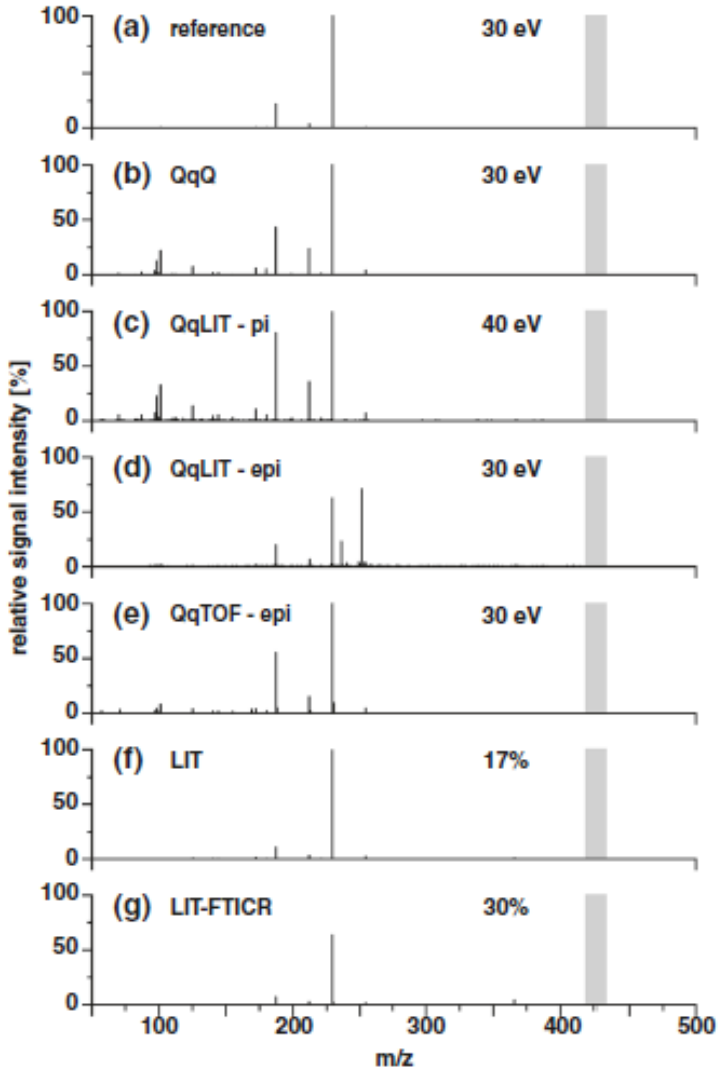
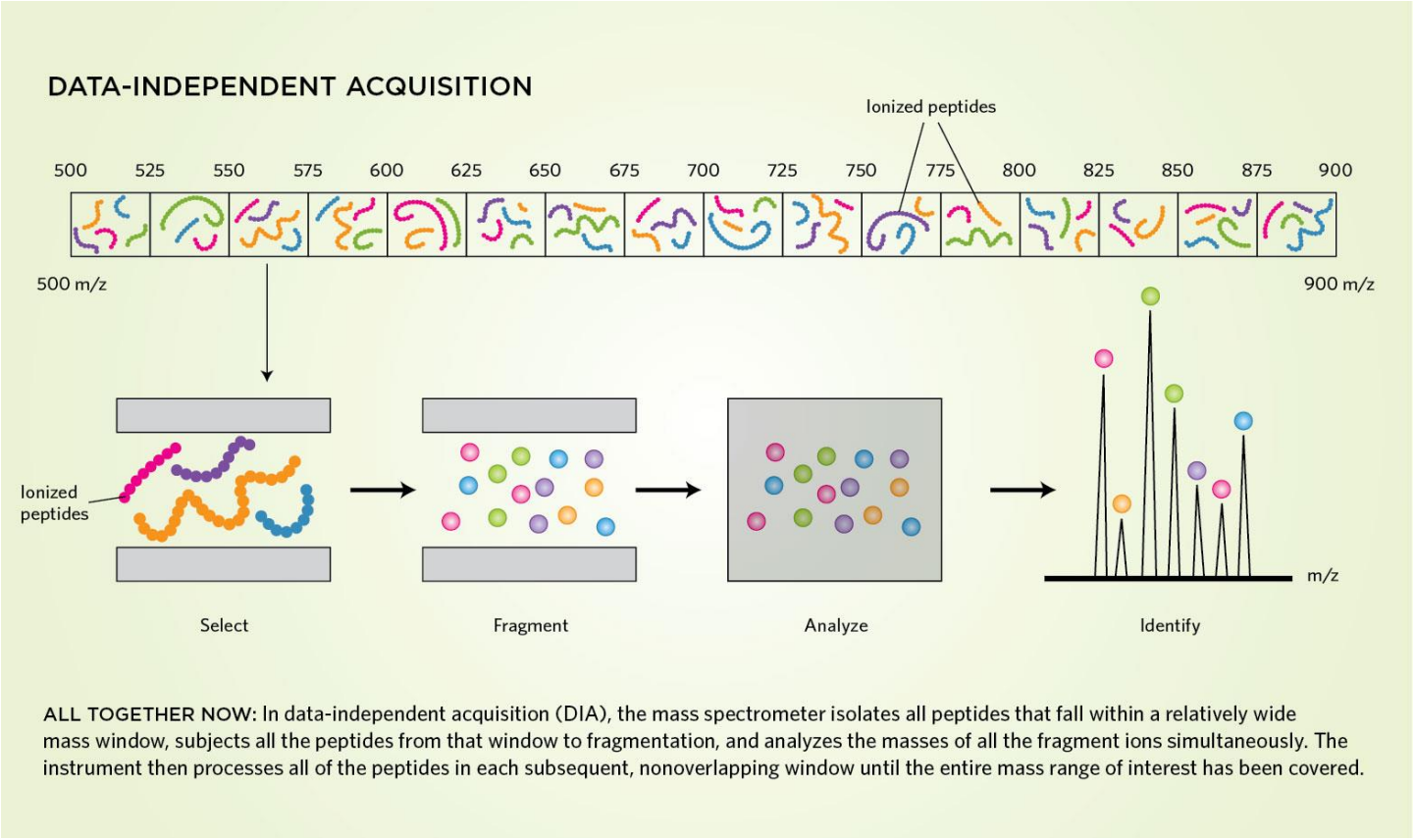


Figure 2 Inter-instrument comparability of dixyrazine-specific tandem mass spectra collected on different instrumental

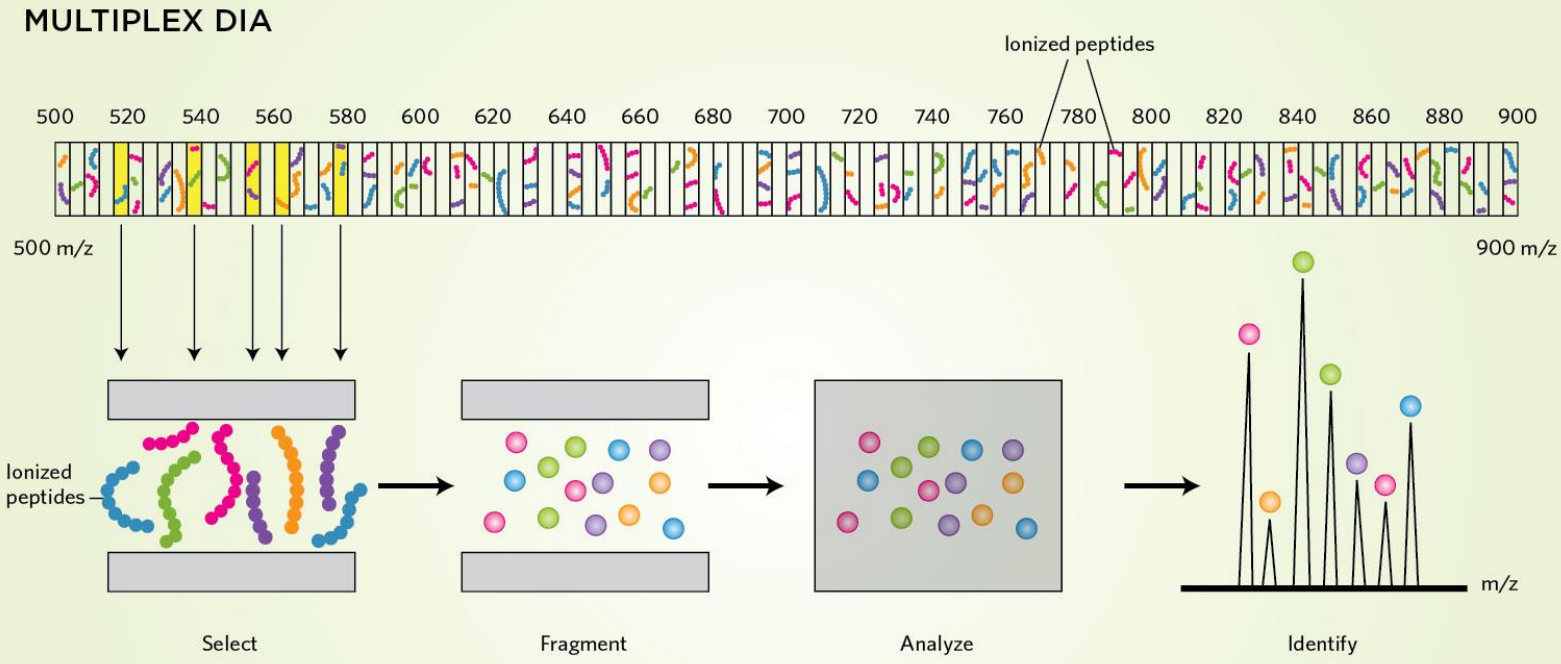
Scheubert *et al.* *Journal of Cheminformatics* 2013, **5**:12

Mode d'acquisition des données



ALL TOGETHER NOW: In data-independent acquisition (DIA), the mass spectrometer isolates all peptides that fall within a relatively wide mass window, subjects all the peptides from that window to fragmentation, and analyzes the masses of all the fragment ions simultaneously. The instrument then processes all of the peptides in each subsequent, nonoverlapping window until the entire mass range of interest has been covered.

Mode d'acquisition des données



REDUCING NOISE: In multiplex DIA, the mass spectrometer isolates all peptides that fall within five, randomly selected, relatively narrow mass windows, subjects all the peptides from each window to fragmentation, and analyzes the masses of all the fragment ions simultaneously. The instrument then processes all of the peptides in subsequent sets of five, randomly chosen, nonoverlapping windows until the entire mass range of interest has been covered.

Méthode SWAT (Sciex)

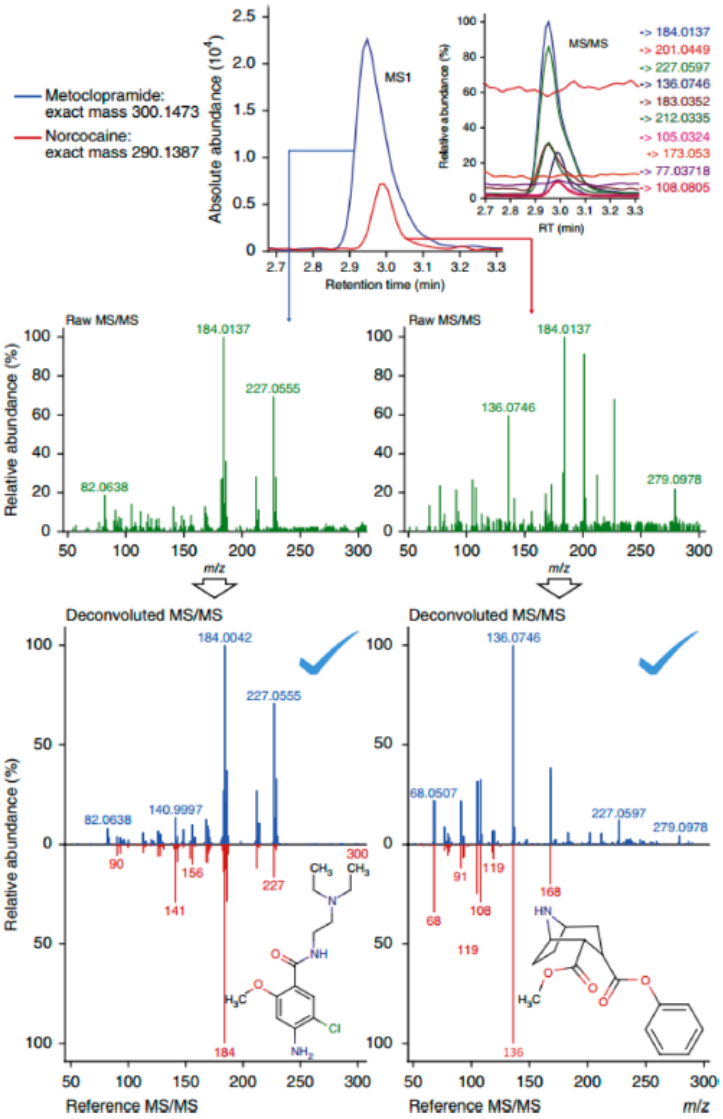
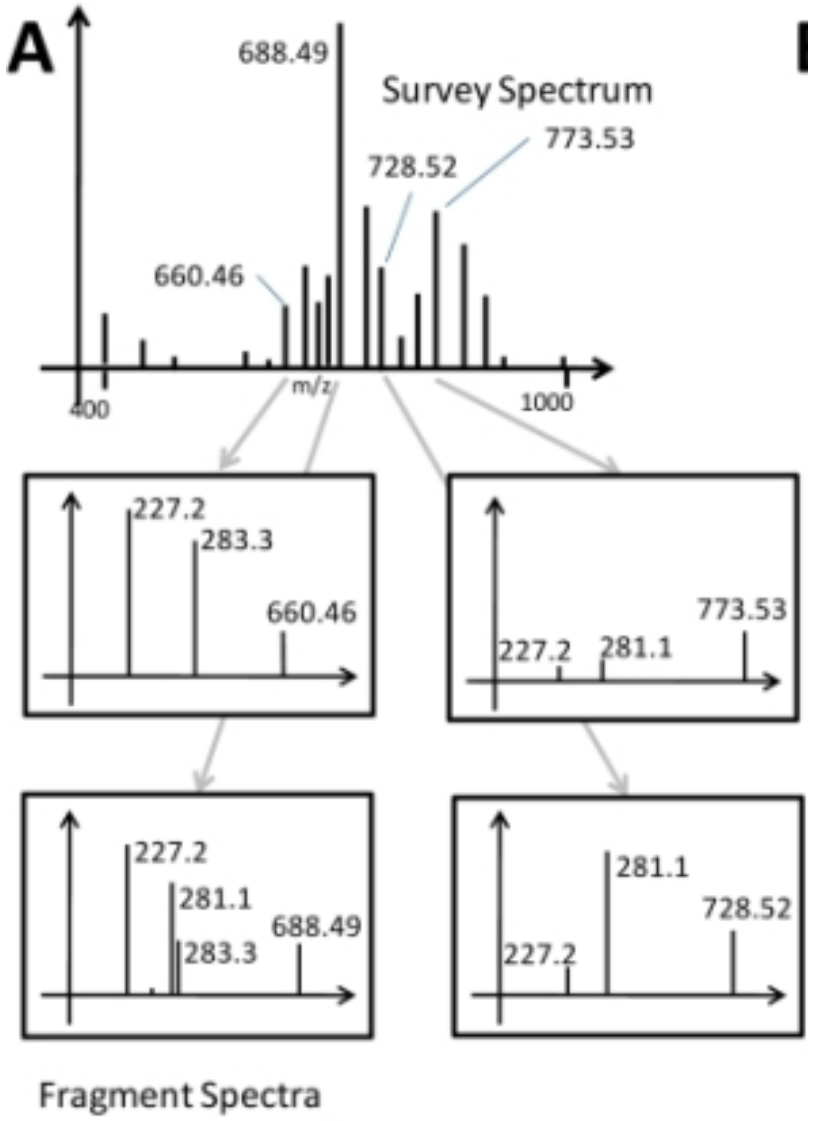


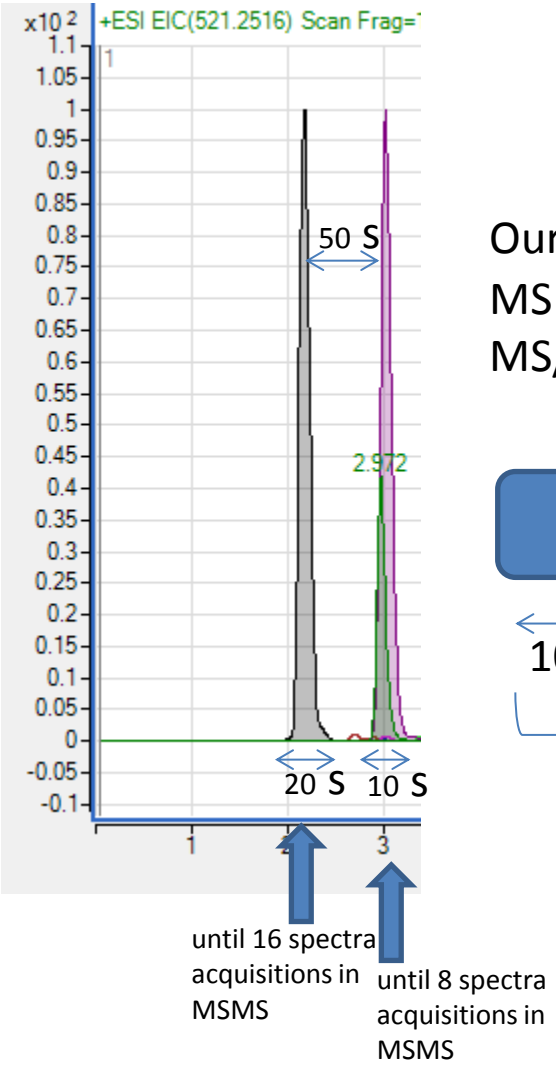
Figure 7. Deconvolution example using SWATH acquisition with HILIC positive ion mode MS. Two pharmaceutical agents, metoclopramide and norcocaine, were detected in untargeted metabolomics screens and coeluted within a 1.8-s peak top difference. The MS/MS ion traces with respect to these two metabolites are also shown in the top right panel of the precursor-ion traces. The middle panels show raw MS/MS spectra of metoclopramide (left) and norcocaine (right), respectively. The spectrum of metoclopramide dominates and masks that of norcocaine, making detection of the latter highly difficult. The bottom panels show the deconvoluted MS/MS spectrum and spectra matching results of metoclopramide (left) and norcocaine (right), yielding dot-product scores of 0.80 and 0.86, respectively. Reprinted with permission from Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. *Nat. Methods* **2015**, *12*, 523–526 (ref 116). Copyright 2015 Nature Publishing Group.

Mode d'acquisition des données

Data-Dependant Acquisition (DDA)



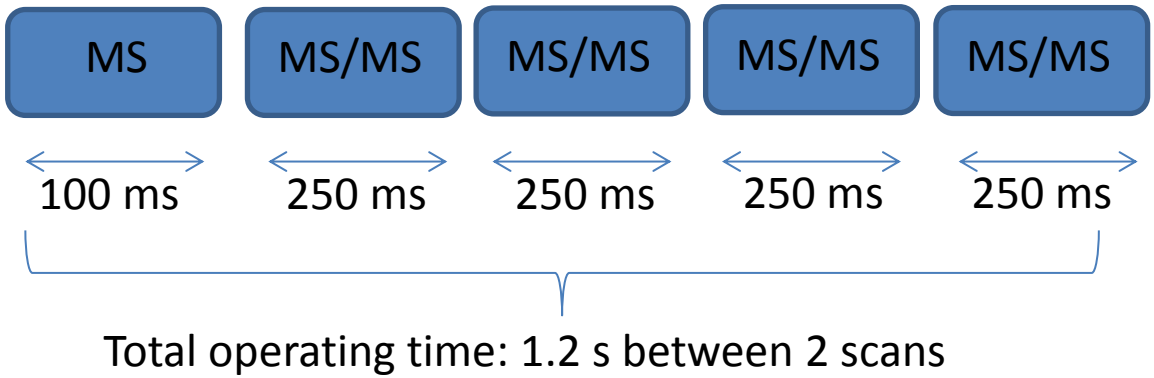
DDA



Our condition in DDA mode:

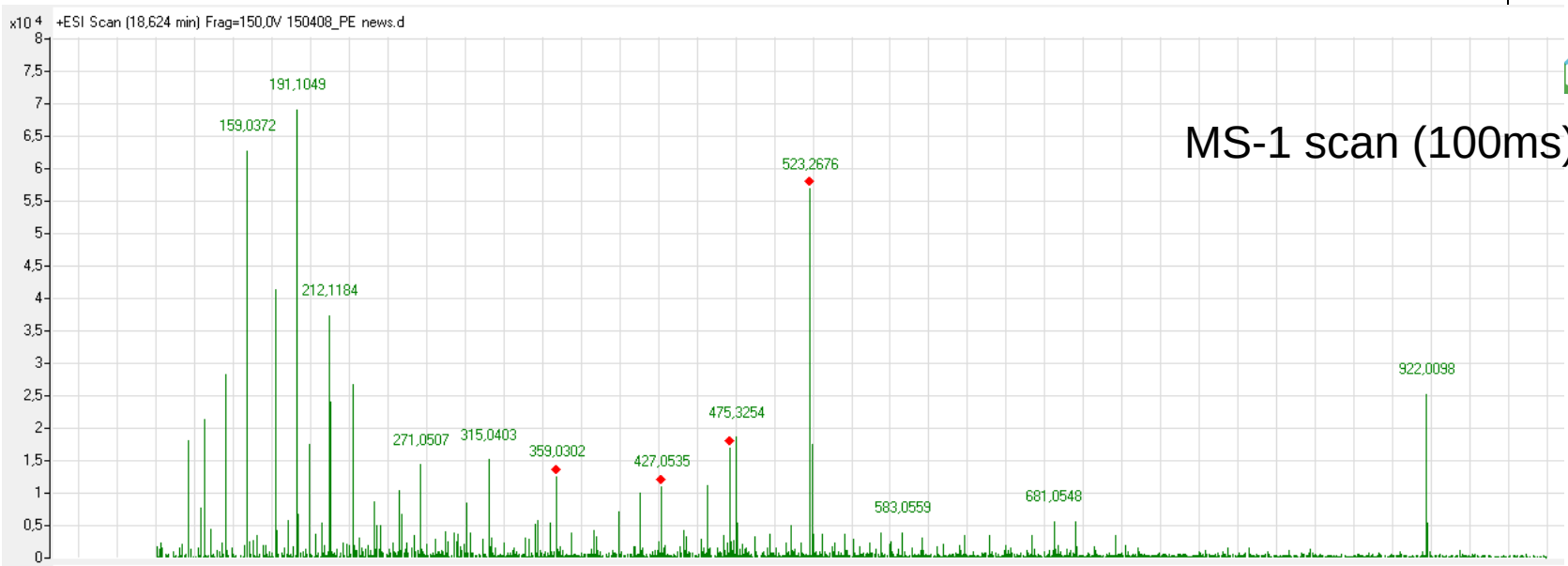
MS 10 spectres/s

MS/MS a maximum of 4 precursor ions per cycle

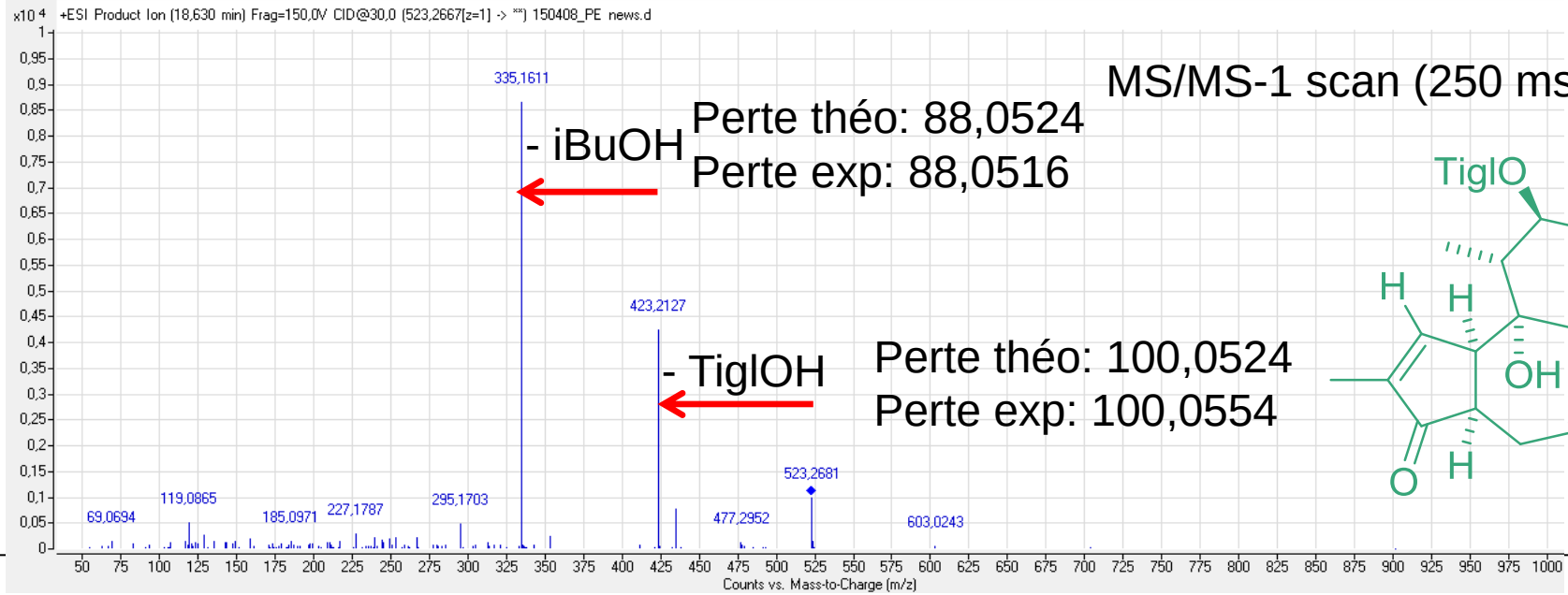


Total operating time: 1.2 s between 2 scans

DDA



MS-1 scan (100ms)



MS/MS-1 scan (250 ms)

- iBuOH

Perte théo: 88,0524
Perte exp: 88,0516

- TiglOH

Perte théo: 100,0524
Perte exp: 100,0554

Arbre décisionnel

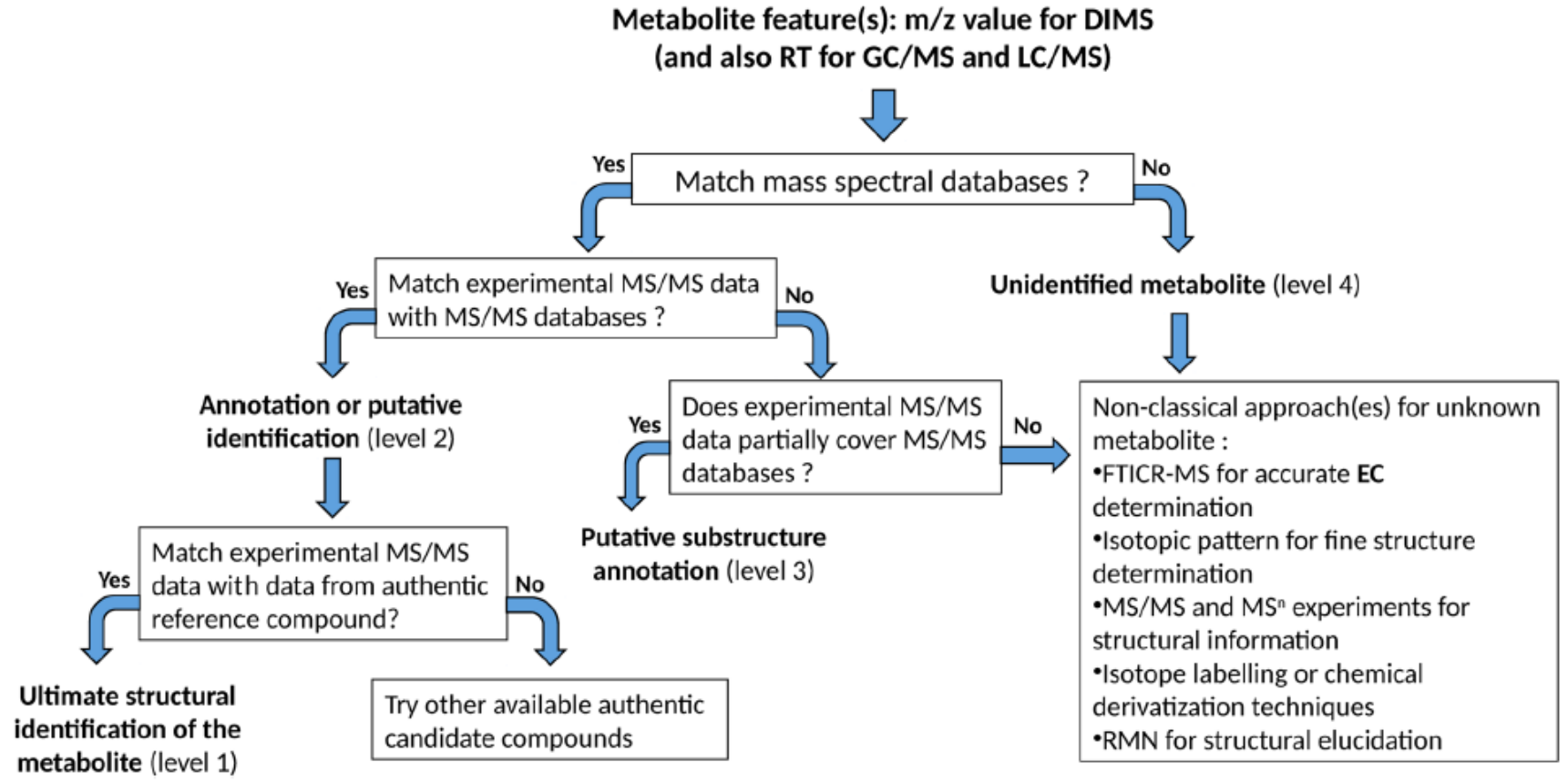


Fig. 3 Strategy for putative metabolite annotation or identification (built from Sumner et al. 2007)

DOI 10.1007/s11306-015-0882-8

Aide à l'identification des spectres MS/MS

Stratégies ???

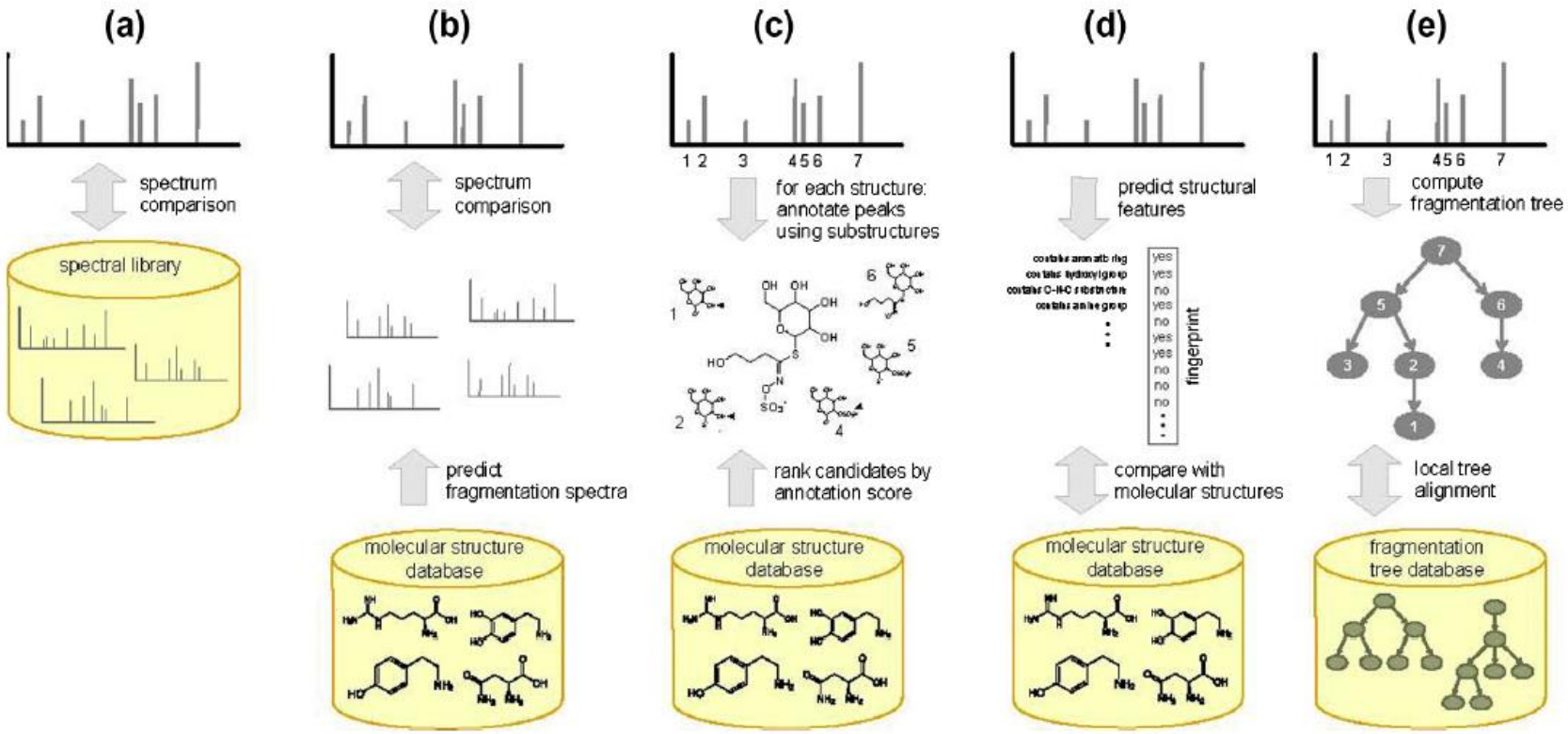


Fig. 1. The five basic approaches of dealing with metabolite fragmentation data: (a) searching spectral libraries; (b) fragmentation spectrum prediction; (c) combinatorial fragmentation; (d) predicting structural features; and, (e) fragmentation trees.

Stratégies ???

Table 2 Approaches for analyzing fragmentation mass spectra of *unknown unknowns* that is, "unexpected" compounds that are not present in spectral libraries [31]

		In silico fragmentation		
Searching for similar compounds	**Mass spectral classifiers**	**Rule-based spectrum prediction**	**Combinatorial fragmentation**	**Fragmentation trees**
searching for similar spectra in a library, assuming that spectral similarity is based on structural similarity	predicting substructures or compound classes by learning spectral classifiers	predicting spectra by applying fragmentation rules to known molecular structures	mapping the fragmentation spectrum to the compound structure to explain the peaks	computing a fragmentation tree that explains the peaks; aligning fragmentation trees to find similar compounds
NIST MS Interpreter [153]	*FingerID* [169]	*Mass Frontier*, *ACD/MS Fragmenter*, *MOLGEN-MS* [196]	*MetFrag* [179]	*SIRIUS* [147,221]

Fragmentation *in silico*

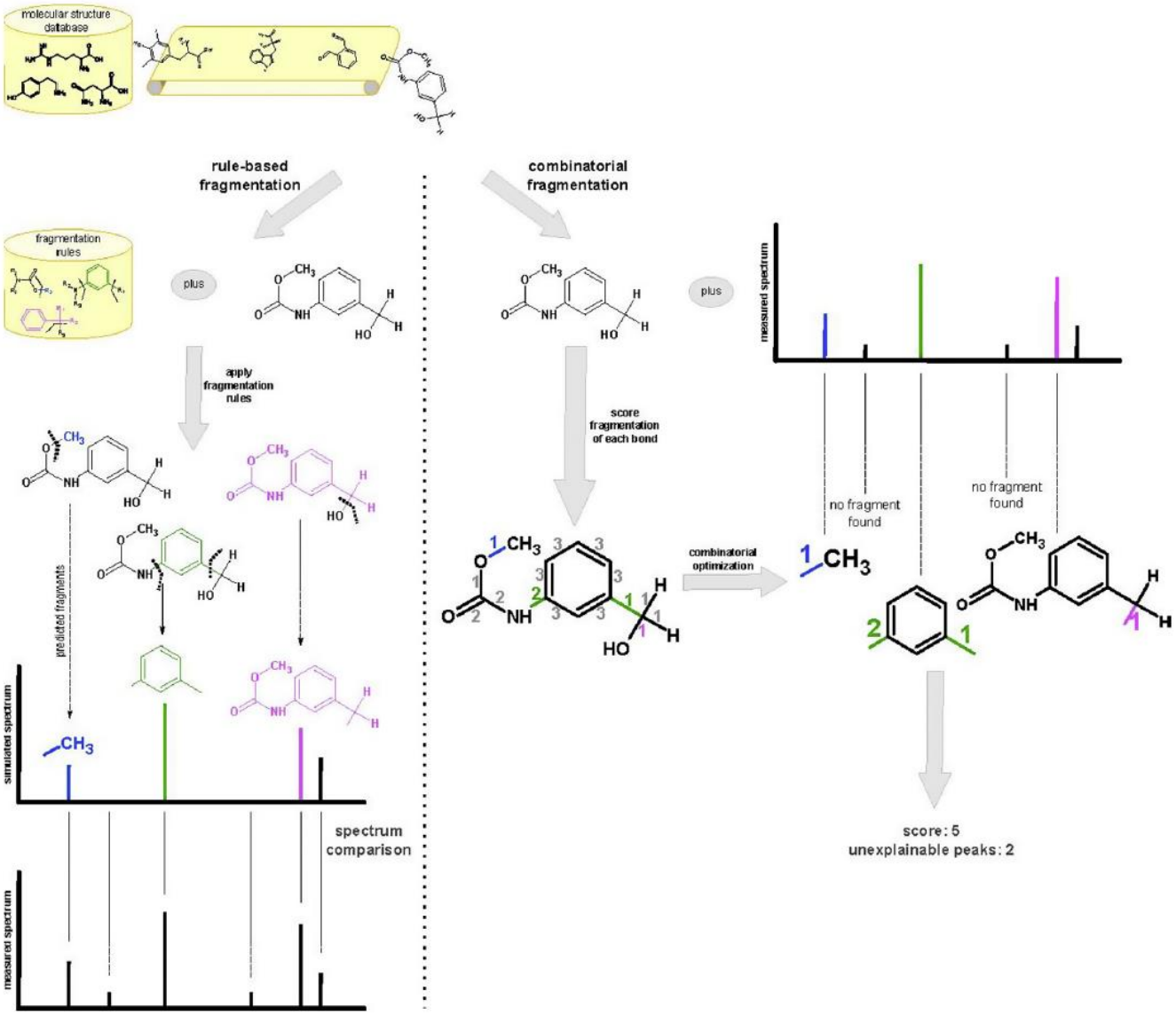


Fig. 2. *In silico* fragmentation. Given a set of known molecular structures, spectra can be predicted by applying fragmentation rules to these structures (left). The simulated spectrum is then compared to the measured spectrum in order to rank candidates. In contrast, combinatorial fragmentation (right) attempts to explain the peaks in the measured spectrum. Costs for cleaving are assigned to all bonds in the structure. Each peak in the spectrum is explained with a substructure of minimal cost.

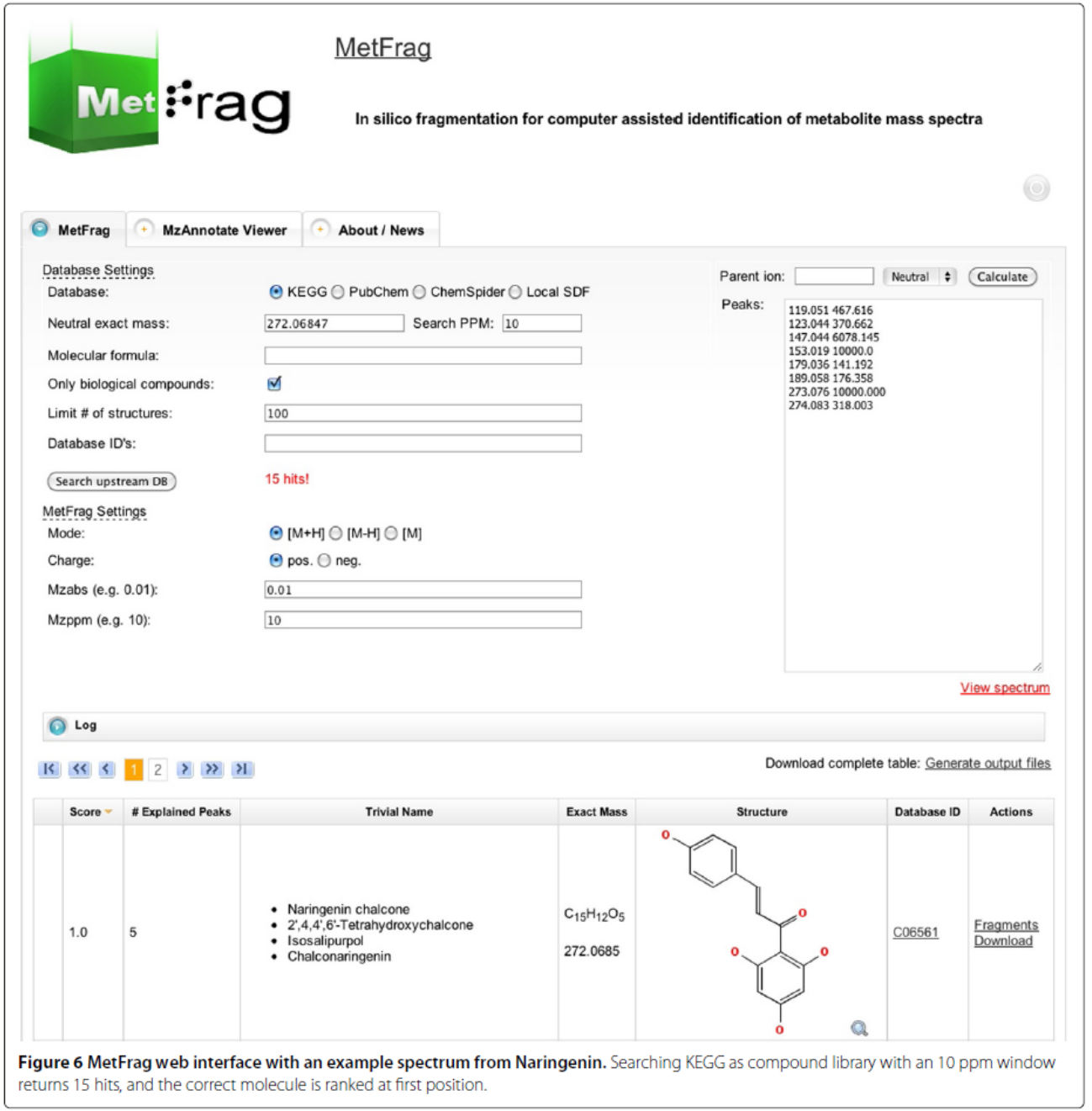


Figure 6 MetFrag web interface with an example spectrum from Naringenin. Searching KEGG as compound library with an 10 ppm window returns 15 hits, and the correct molecule is ranked at first position.

Prédictions de structure

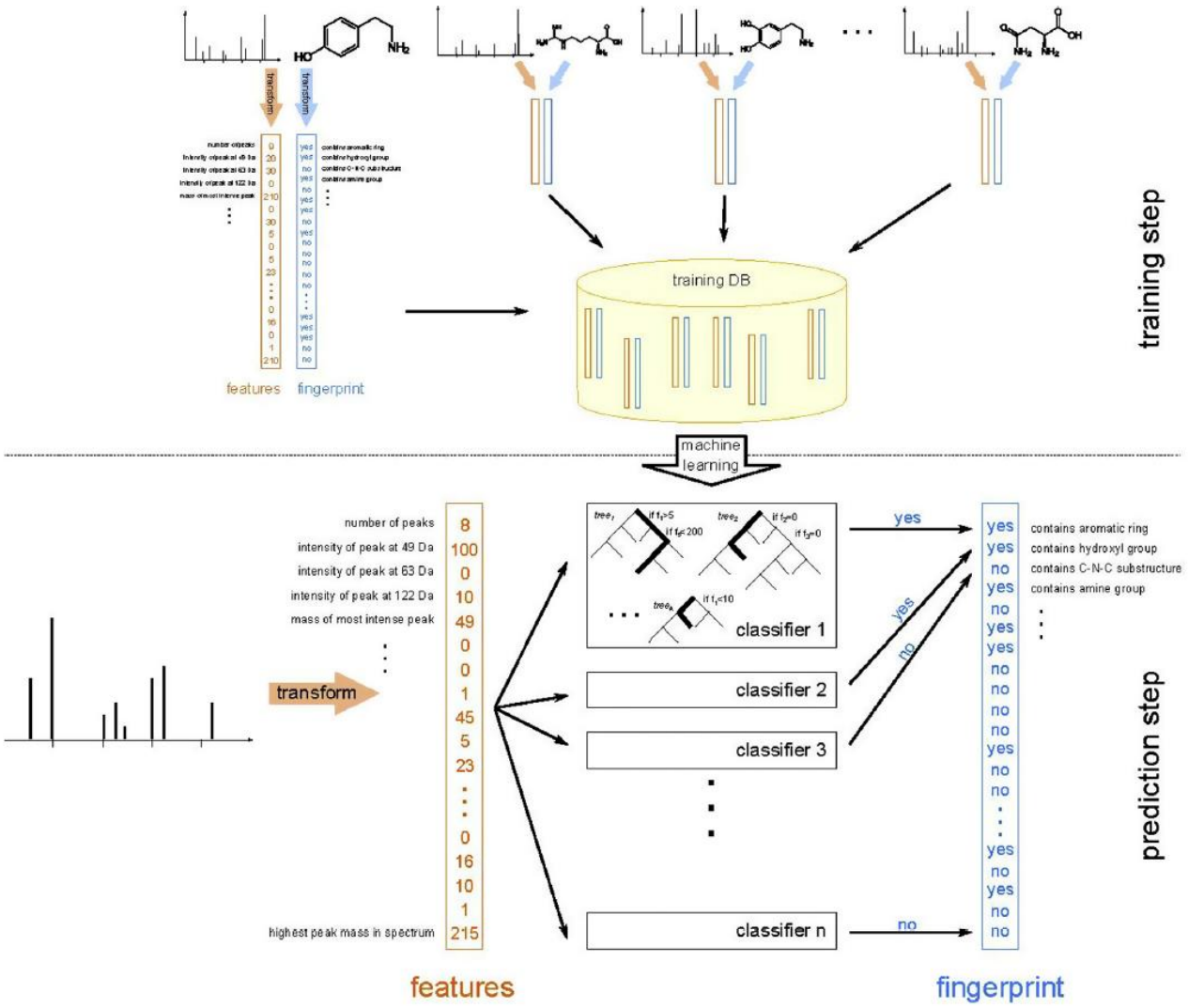


Fig. 3. Feature-based substructure prediction. In the training step (top), pairs of mass spectra and corresponding molecular structures are transformed to feature vectors (describing the spectra) and fingerprints (describing the structures). These are used to train the classifiers using, for example, random forests. In the prediction step (bottom), a query spectrum is transformed to a feature vector using the same transformation. These features are passed to the classifiers, which decide whether a substructure is present or not in the investigated compound.

Arbre de fragmentation

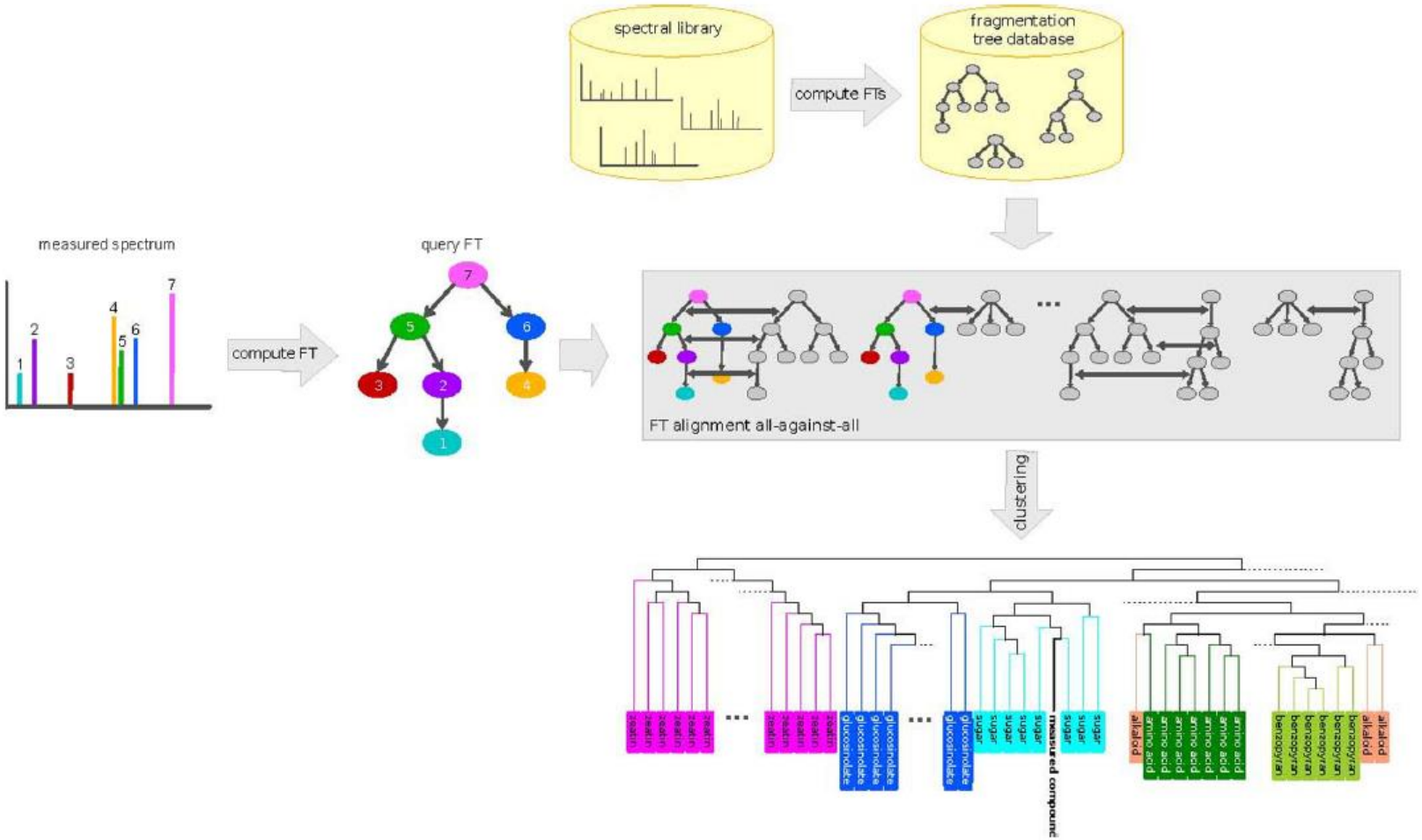


Fig. 4. Fragmentation tree alignment for compound classification. A fragmentation tree is computed from the measured spectrum. The tree is aligned to a database of fragmentation trees in an all-against-all manner. The compounds are clustered based on the resulting similarity scores. Similar compounds (belonging to the same compound class) cluster together. The class of the unknown compound can be concluded from the cluster into which it falls.

Arbre de fragmentation

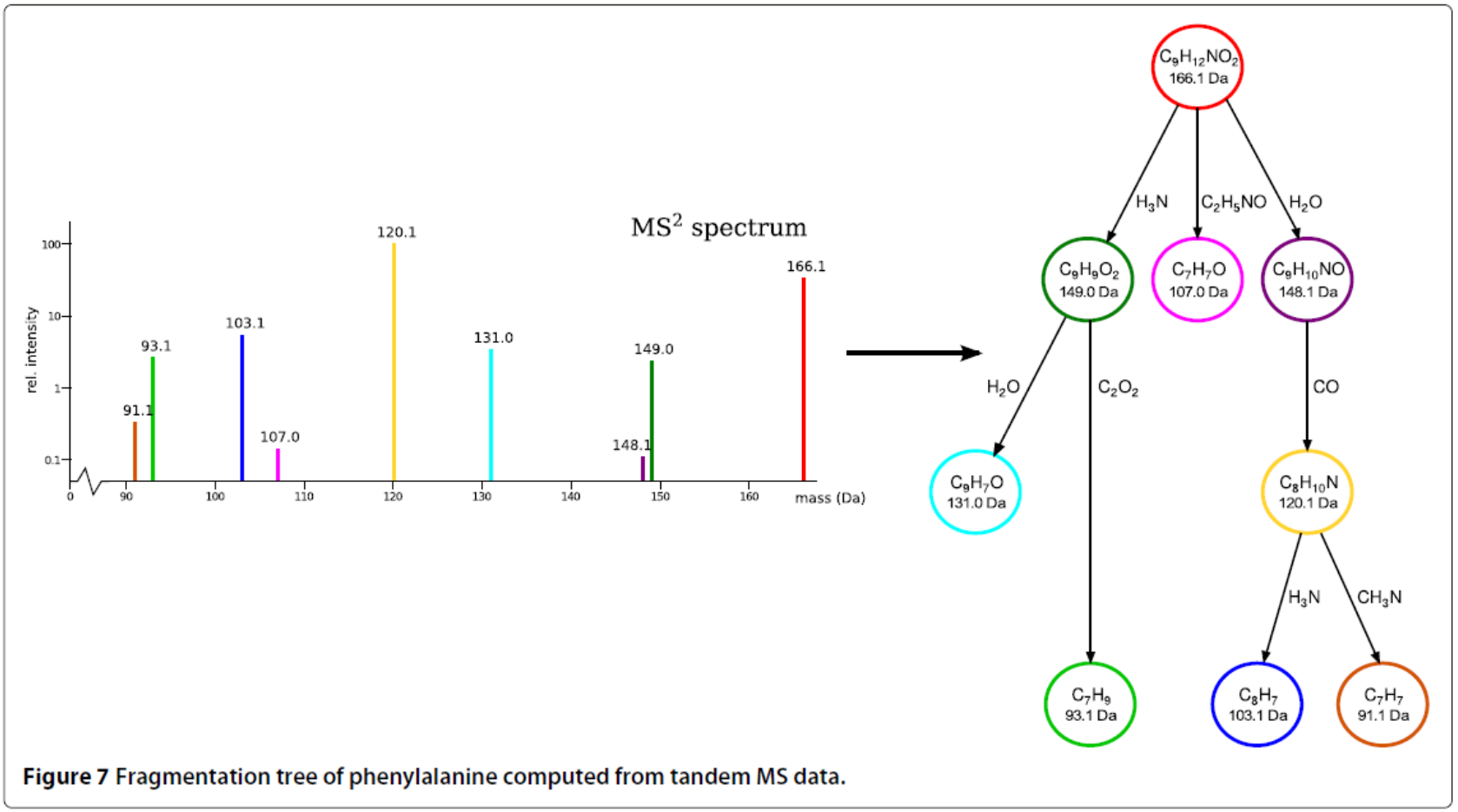
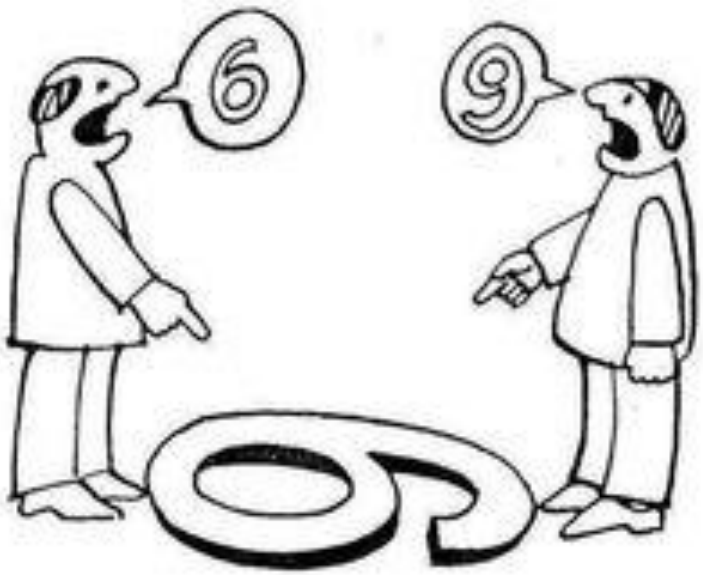


Figure 7 Fragmentation tree of phenylalanine computed from tandem MS data.

Un autre point de vue … la déréplication



seen from Fig. 1. The urge to fill the industrial pipeline and to discover novel lead-like compounds for drug discovery, which can meet the societal challenge of the lack of suitable therapeutic agents for a broad range of diseases, has never been greater. Antibiotic resistance, for instance is a "ticking time bomb", we are currently on "red alert", having a poor drug repertoire, in which commonly treated infections are becoming lethal. One dominant tailback in NPs discovery is dereplication, *i.e.* the discard of known compounds. With the ultimate objective of speeding up and improving drug discovery program efficiency, researchers have been using multifaceted approaches, either merging different areas of knowledge or creating totally innovative ways to advance this field. Consequently, dereplication, which is the object of our review, has become a matter of great interest in recent years. The

Nat. Prod. Rep., 2015, 32, 779–810 |

Domaine en plein essor

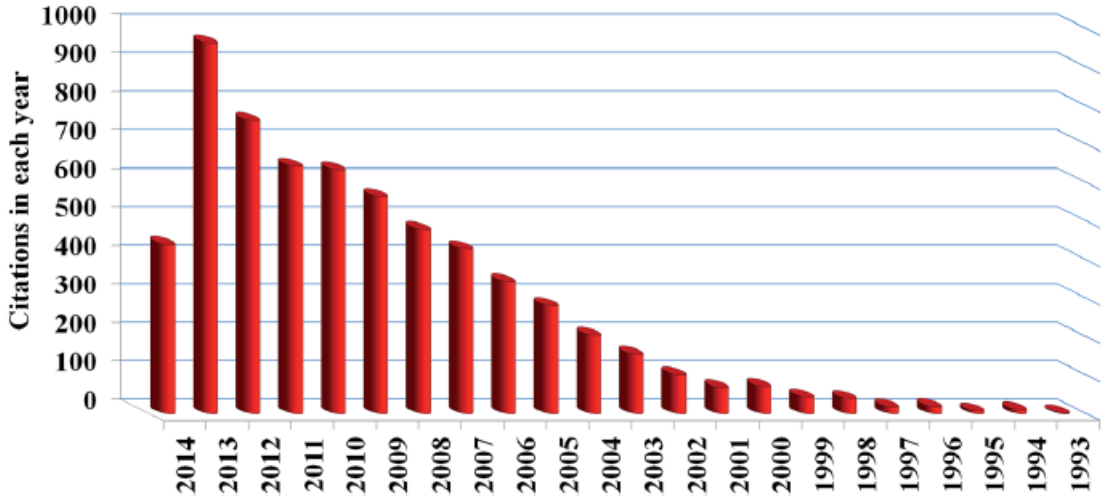
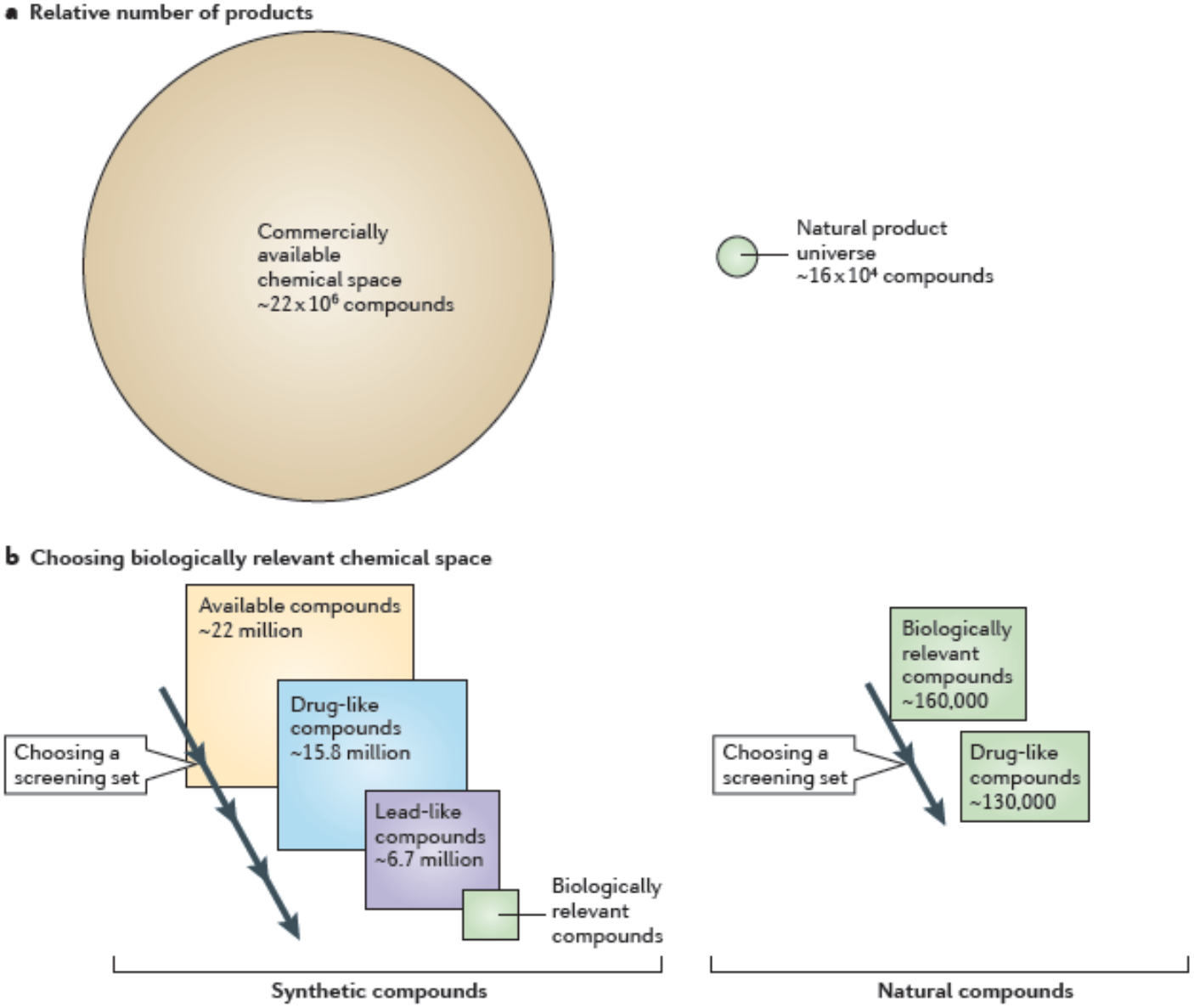


Fig. 3 Number of citations per year covering dereplication topic, period 1993–2014. Data source from Web of Science™ Core Collection.

Complexité et intérêt du vivant



doi:10.1038/nrd4510

Problématique

- Faire le tri entre le connu et l’inconnu
- Classer l’intérêt de l’inconnu
- Rapide et simple

Réseau moléculaire

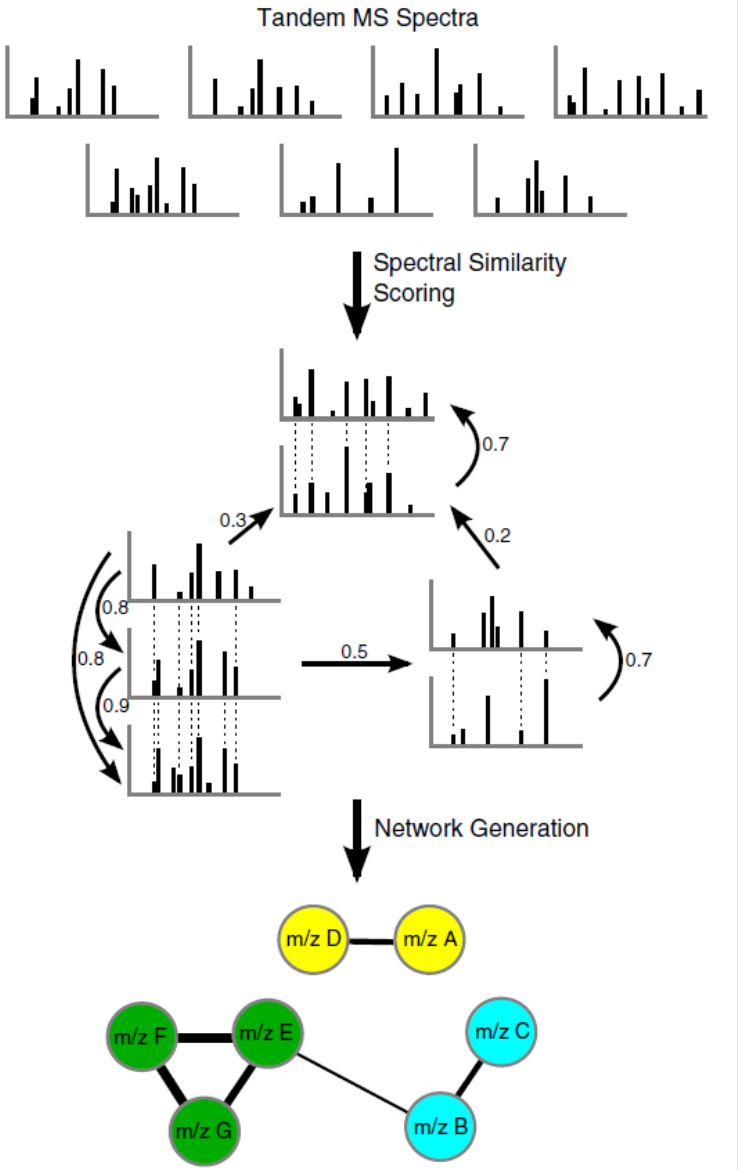


Figure 8 Using spectral alignment of tandem MS data to generate a molecular network. The thickness of the edges indicates the similarity between the spectra. Figure redrawn from Watrous *et al* [230].

Ce n'est, ni plus ni moins, qu'une manière de classer des spectres de masse tandem par homologie, donc assez similaire des arbres de fragmentation sans décrire réellement les voies de fragmentation!

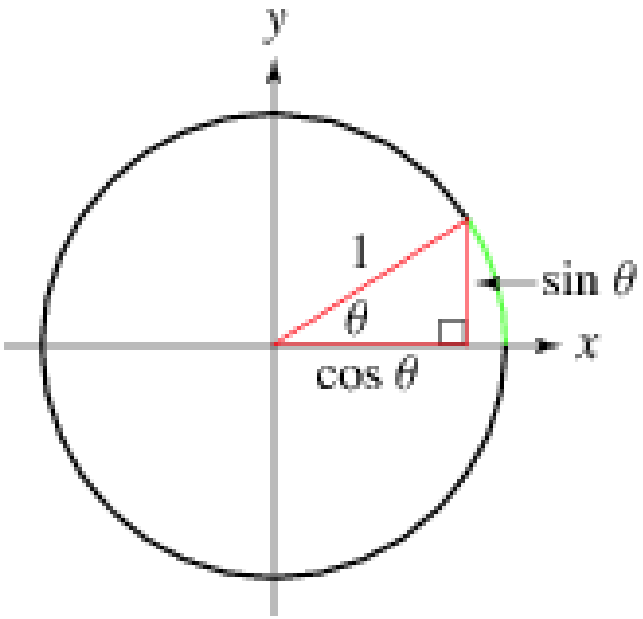
Facteur de similarité

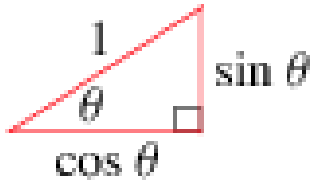
of the spectra. The spectra are treated as vectors $f = (f_1, \ldots, f_M)$ and $g = (g_1, \ldots, g_M)$, and the scalar product $\langle f, g \rangle = \Sigma_m f_m g_m$ is computed. This is particularly applied for unit mass accuracy data, where spectra can be directly mapped to vectors. For data with high mass accuracy, we can treat the spectra as continuous functions f, g with scalar product $\int f(m)g(m)dm$. Often, the raw peak shapes are not used but, instead, peaks are idealized as Gaussian functions. We can also introduce a weight function to weight the terms of the product differently, depending on the mass. Often, it is not the dot product that is reported but the enclosed angle θ or its cosine,

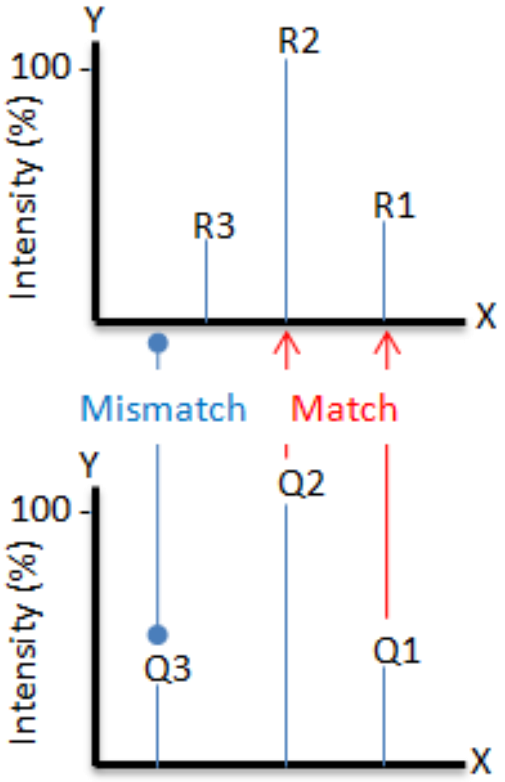
$$\cos\theta = \frac{\langle f, g \rangle}{\sqrt{\langle f, f \rangle}\sqrt{\langle g, g \rangle}}.$$

Similarité cosinus ou mesure cosinus

Facteur de similarité







$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

Initialement …
pour l'analyse de séquences peptidiques

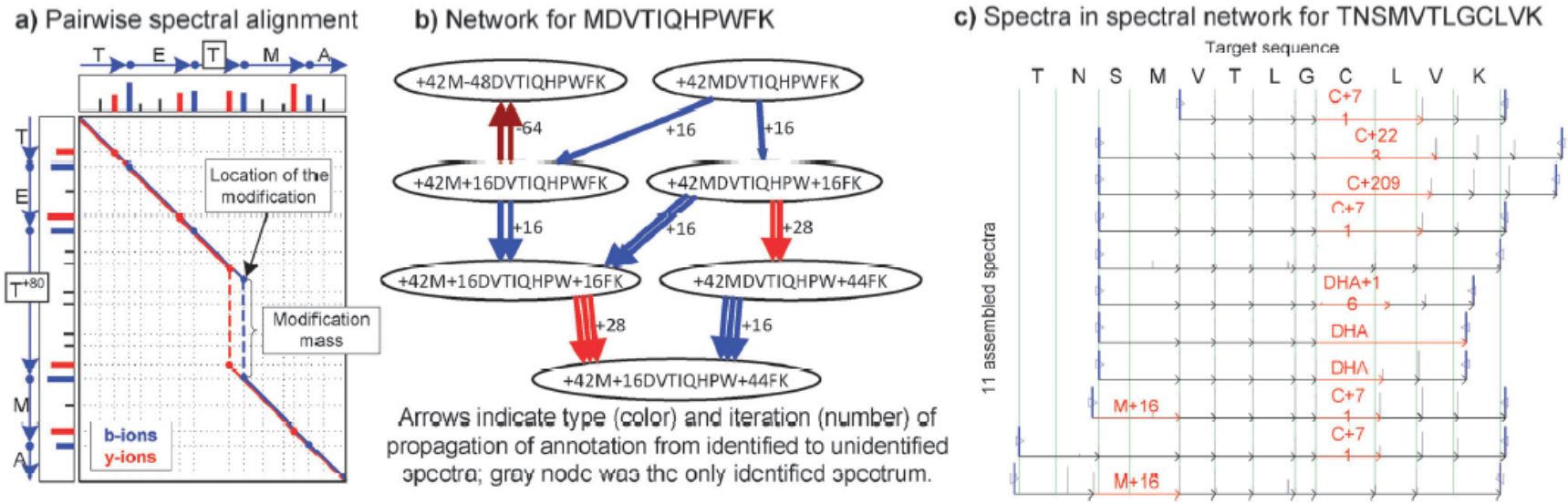


Fig. 1 Discovery and identification of post-translational modifications through spectral networks; (a) Spectral alignment between modified and unmodified variants of the peptide TETMA (*b*-ions shown in blue, *y*-ions in red, blue/red lines track consecutively matched *b*/*y*-ions); (b) Grouped modification states of the peptide MDVTIQHPWFK from a sample of cataractous lenses. Nodes in the spectral network represent individual MS^2 spectra and edges between nodes represent significant spectral alignments such as that shown in part (a); (c) Spectra assembled in the spectral network for TNSMVTLGCLVK with diverse Cysteine modifications on a monoclonal antibody. Each arrow corresponds to the propagation of a sequence and/or PTM from an identified spectrum to an unidentified spectrum (repeated arrows are iterative propagations). Arrow colors correspond to types of modifications transferred.

Mol. BioSyst., 2012, **8**, 2535–2544

Initialement …
pour l'analyse de séquences peptidiques

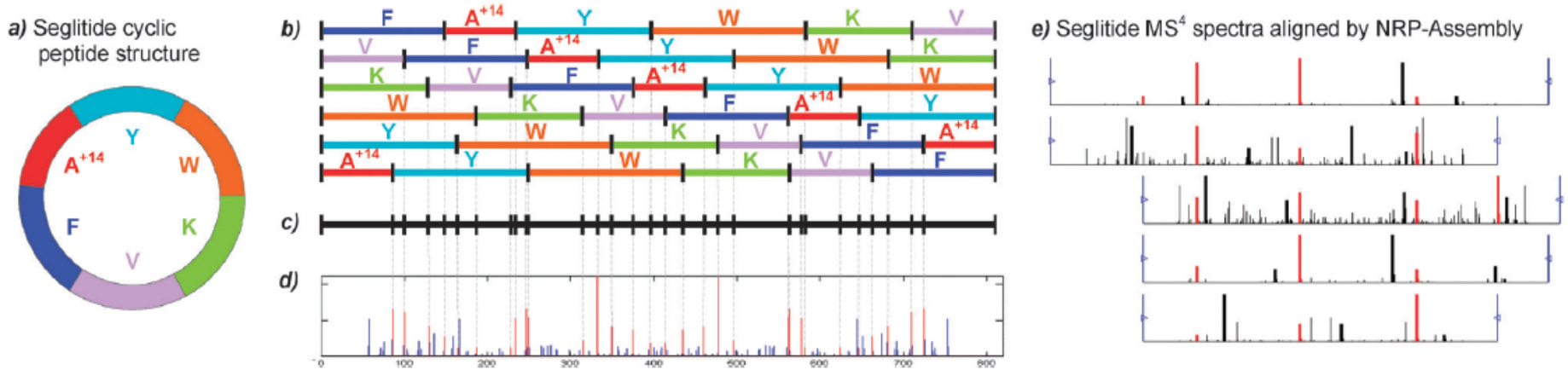


Fig. 3 Analysis of the cyclic peptide Seglitide. (a) The circular structure of Seglitide is schematically illustrated with each residue represented by a different color (slice sizes not scaled to corresponding masses of the residues). A^{+14} denotes a non-standard residue with integer mass 71 + 14 = 85 Da. (b) MS^2 fragmentation of Seglitide generates up to 6 linear peptides representing different rotated variants of the same cyclic peptide. (c) Theoretical spectrum for Seglitide by superposition of the fragment masses of the linearized peptides. (d) Experimental spectrum of Seglitide resulting from a mixture of 6 linear peptides (the peaks corresponding to fragment ions are shown in red). (e) Spectral network from assembled Seglitide MS^n spectra and used for *de novo* sequencing with unknown amino acid masses.

Initialement …
pour l'analyse de séquences peptidiques

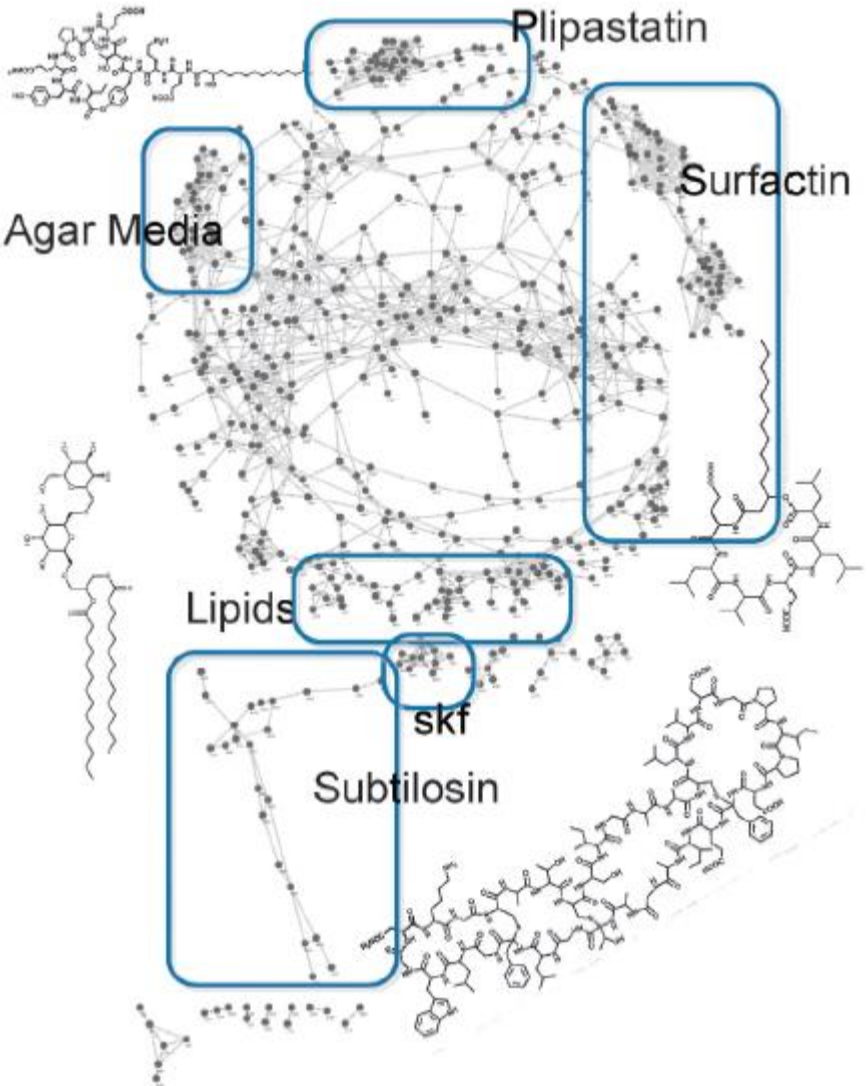
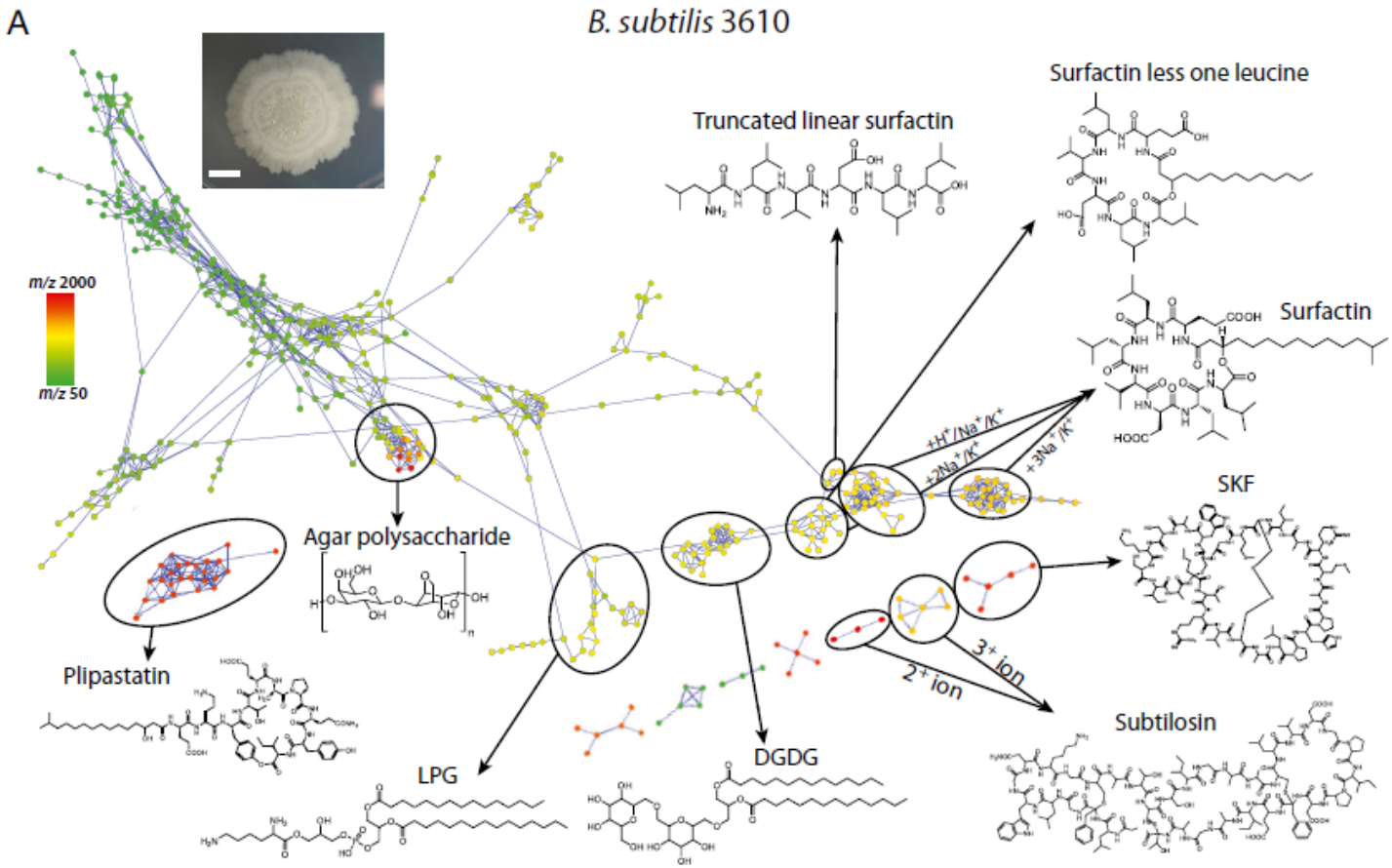


Fig. 4 Molecular spectral network of a partial Bacillus subtilis secretome; nodes indicate MS^2 spectra of initially-unknown compounds of any class of molecules (no peptide-specific assumptions were made), and edges indicate significant similarity between the MS^2 fragmentation patterns of different spectra, mostly between intermediates/variants of the same compounds. Selected molecular structures are shown in black overlaid with the network and next to the correspondingly highlighted network clusters.

Initialement …
pour l'analyse de séquences peptidiques



www.pnas.org/cgi/doi/10.1073/pnas.1203689109

Initialement …
pour l'analyse de séquences peptidiques

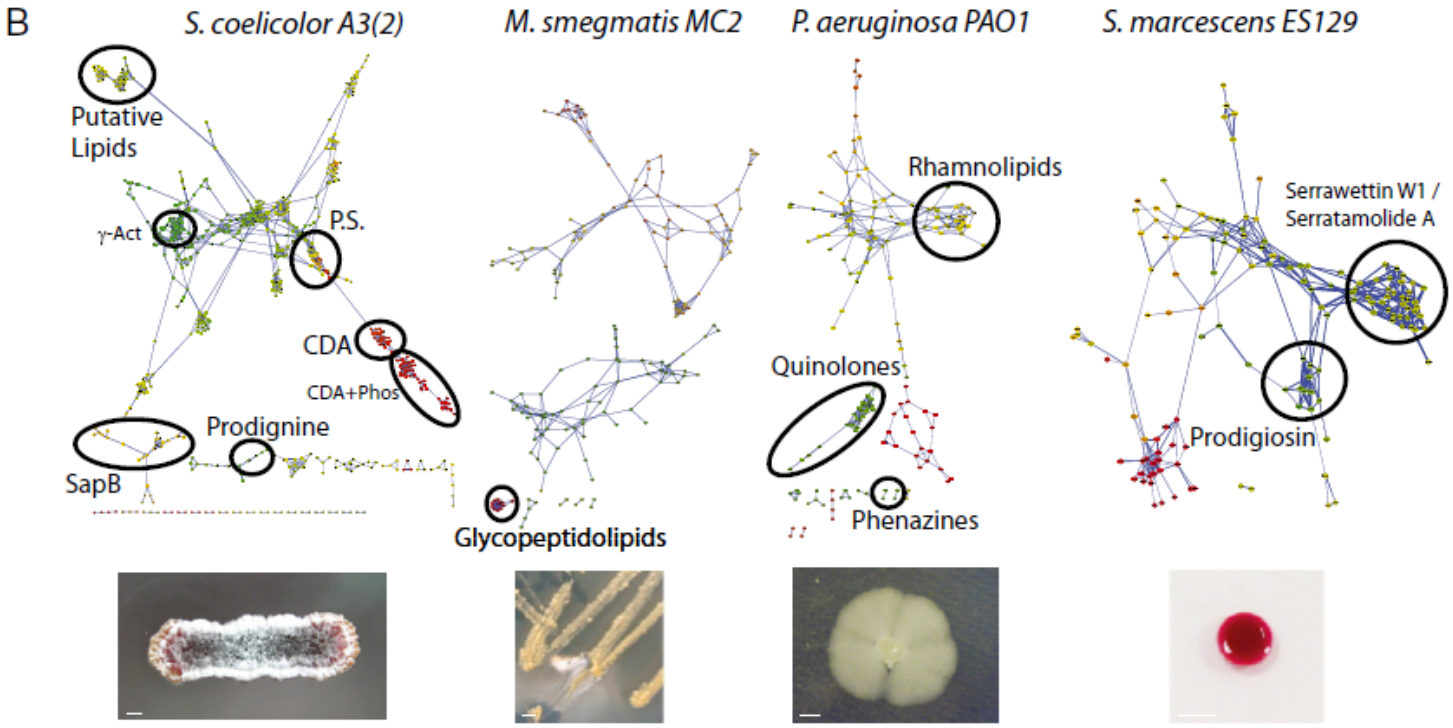


Fig. 3. Molecular networks of nanoDESI fragmentation data obtained from single microbial colonies. (*A*) The annotated molecular network from *B. subtilis* 3610. (*B*) The annotated molecular network of *S. coelicolor* A3(2), *M. smegmatis MC2*, *P. aeruginosa PAO1*, and *S. marcescens* ES129. *Insets*: Images of samples were probed with nanoDESI. The structures of each of the annotated clusters are shown in *SI Appendix*, Figs. S1, S4, and S5. The color scale shows the mass range of the parent ions: green nodes represent the smallest masses; red nodes represent the largest masses fragmented. (Scale bar: 1 mm.)

www.pnas.org/cgi/doi/10.1073/pnas.1203689109

Initialement …
pour l'analyse de séquences peptidiques

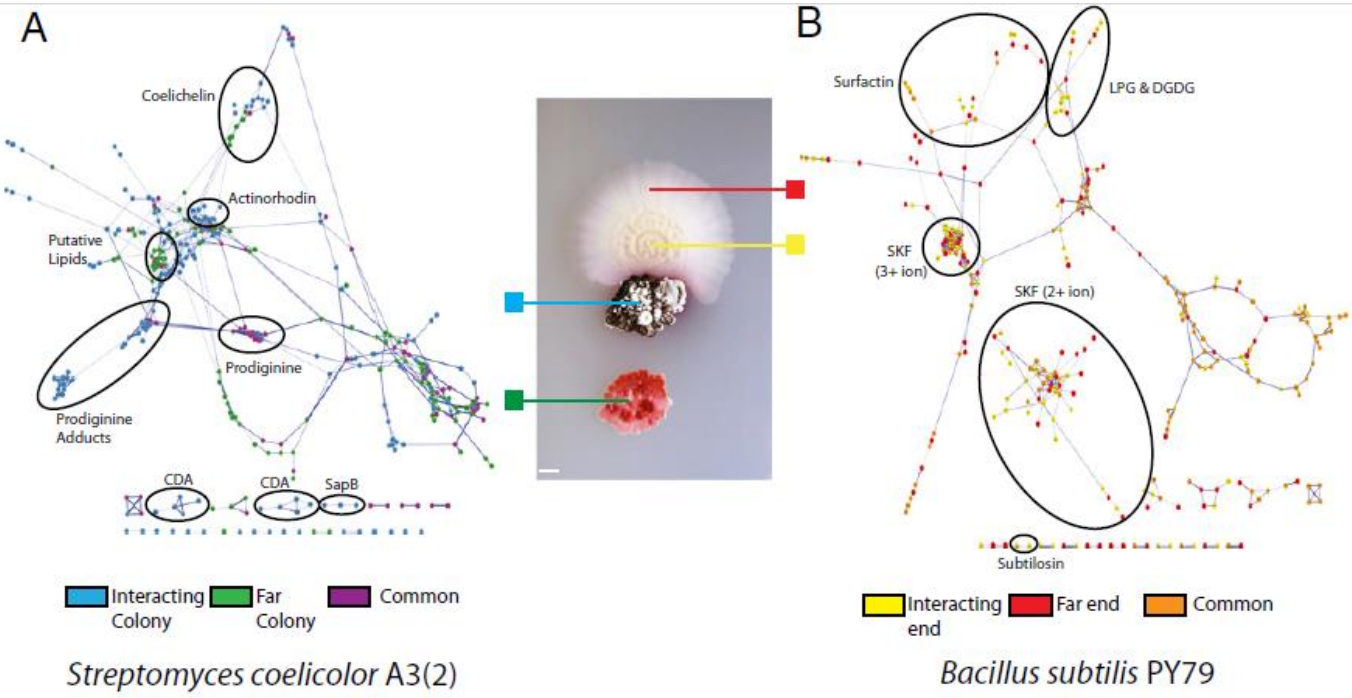
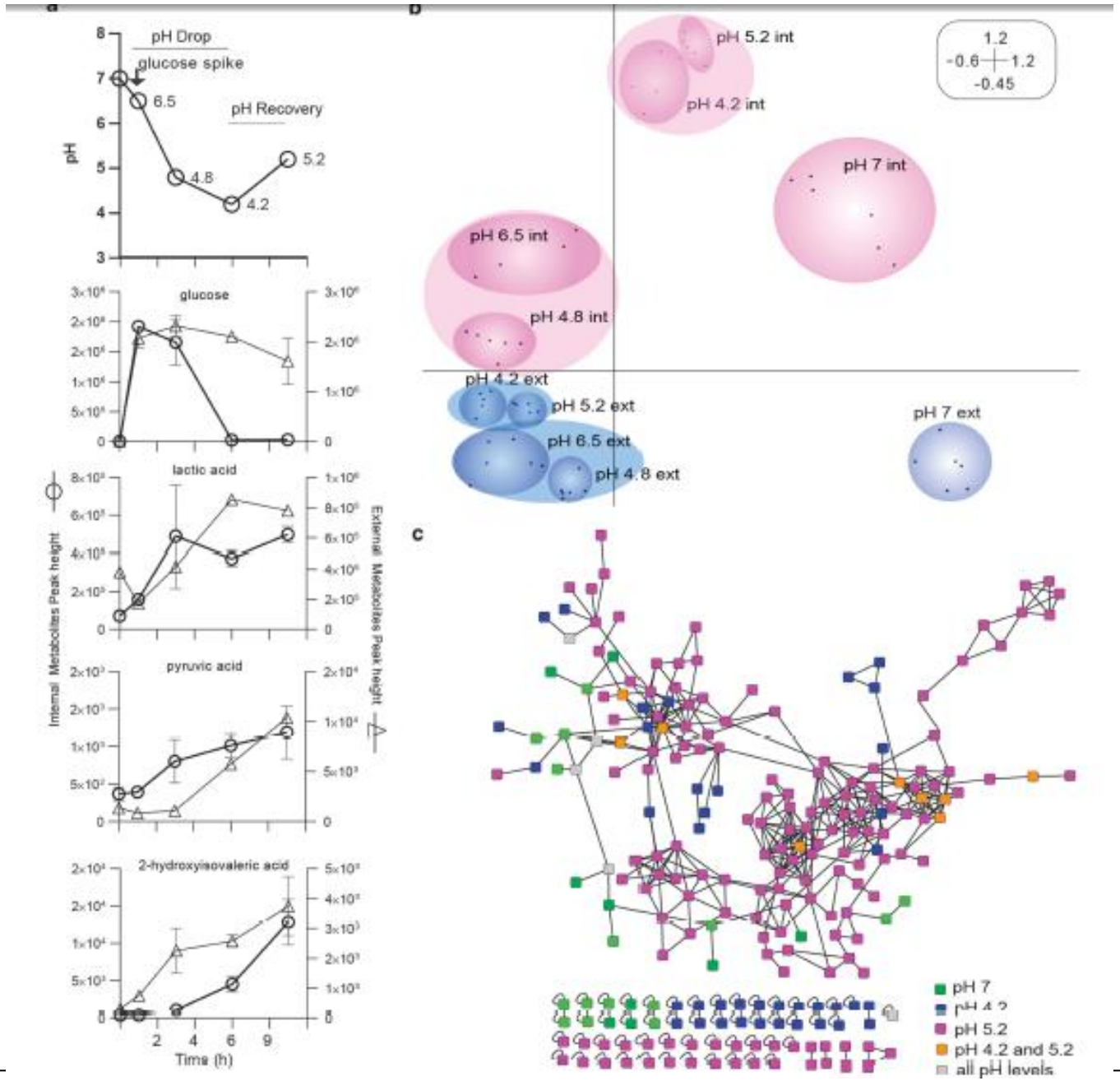


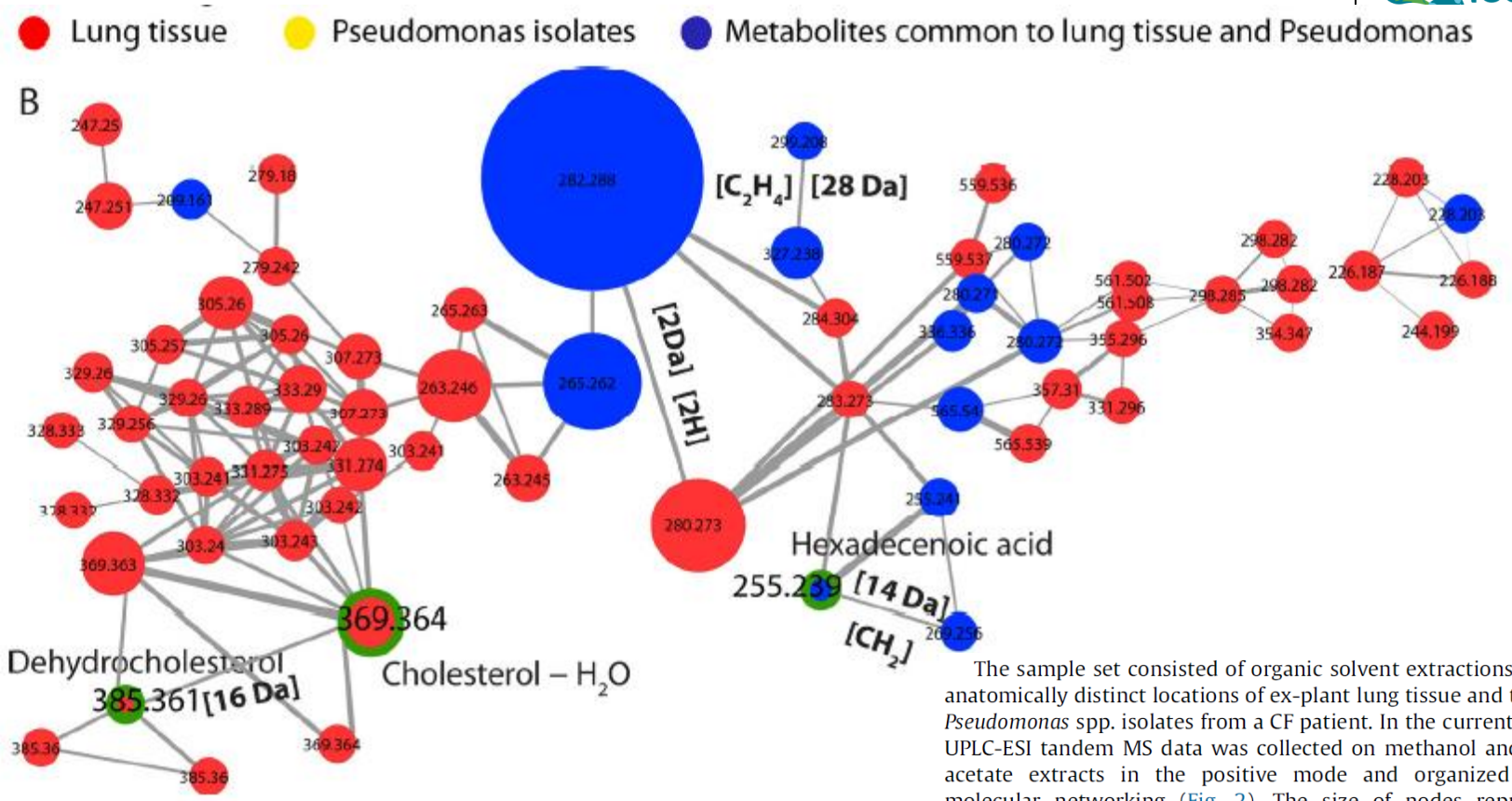
Fig. 5. The molecular network of *S. coelicolor* A3(2) interacting with *B. subtilis* PY79. (*A*) The comparison of the molecular data from the *S. coelicolor* colony adjacent to *B. subtilis* vs. the *S. coelicolor* colony further away. (*B*) The comparison of the molecular data from the interacting and noninteracting sides of the *B. subtilis* PY79 colony. It should be noted that, although PY79 has a frame shift in *sfp*, the phosphopantetheinyl transferase required for surfactin and plipastatin biosynthesis, surfactin is still produced in small amounts (41). This has been observed before by MALDI imaging, as well as imprint desorption electrospray ionization, and can be attributed to promiscuity of another phosphopantetheinyl transferase or a ribosome slippage providing a low amount of in-frame translation of the frame-shifted *sfp* gene (19, 23).

www.pnas.org/cgi/doi/10.1073/pnas.1203689109



The ISME Journal (2015) 9, 2605–2619

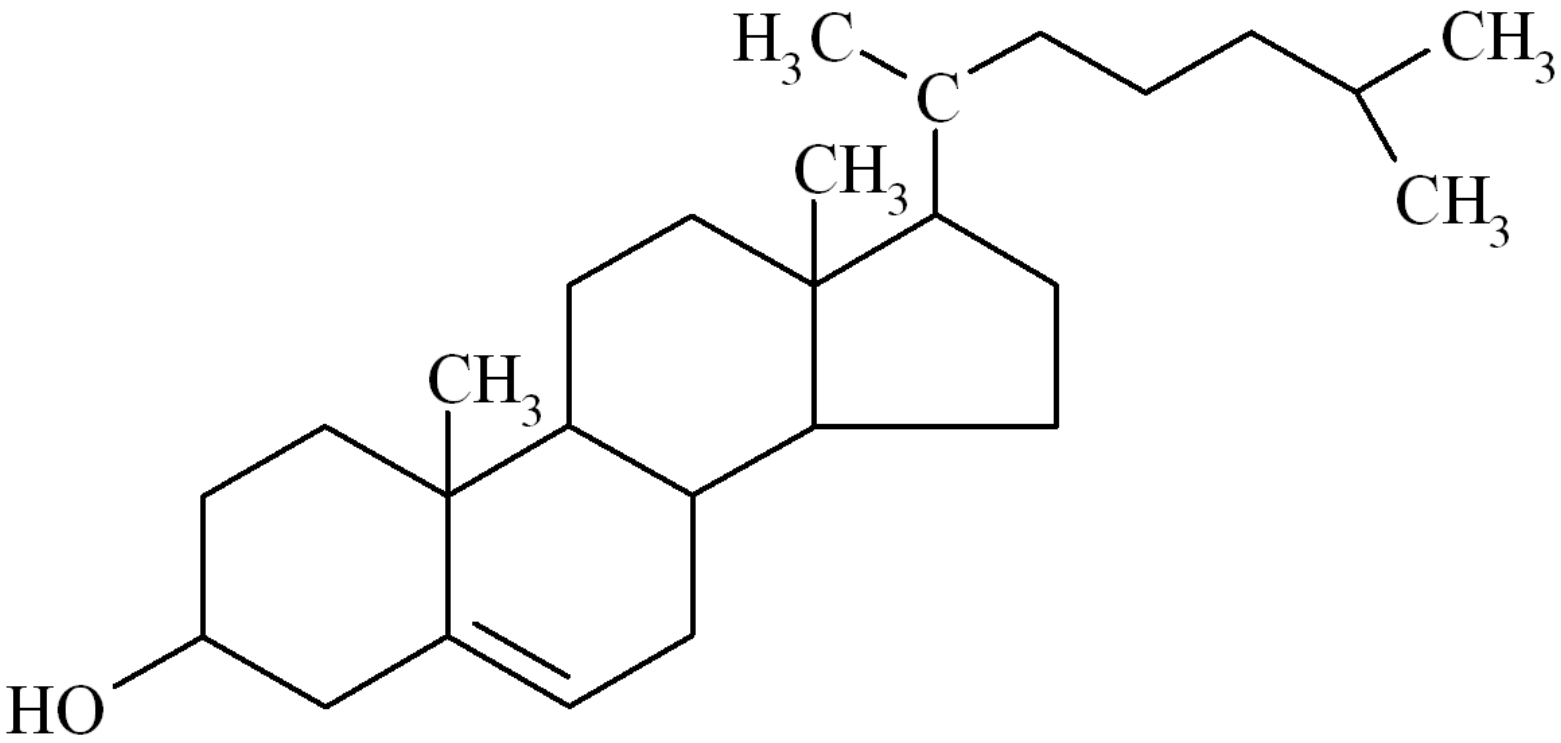
Tous les pièges …



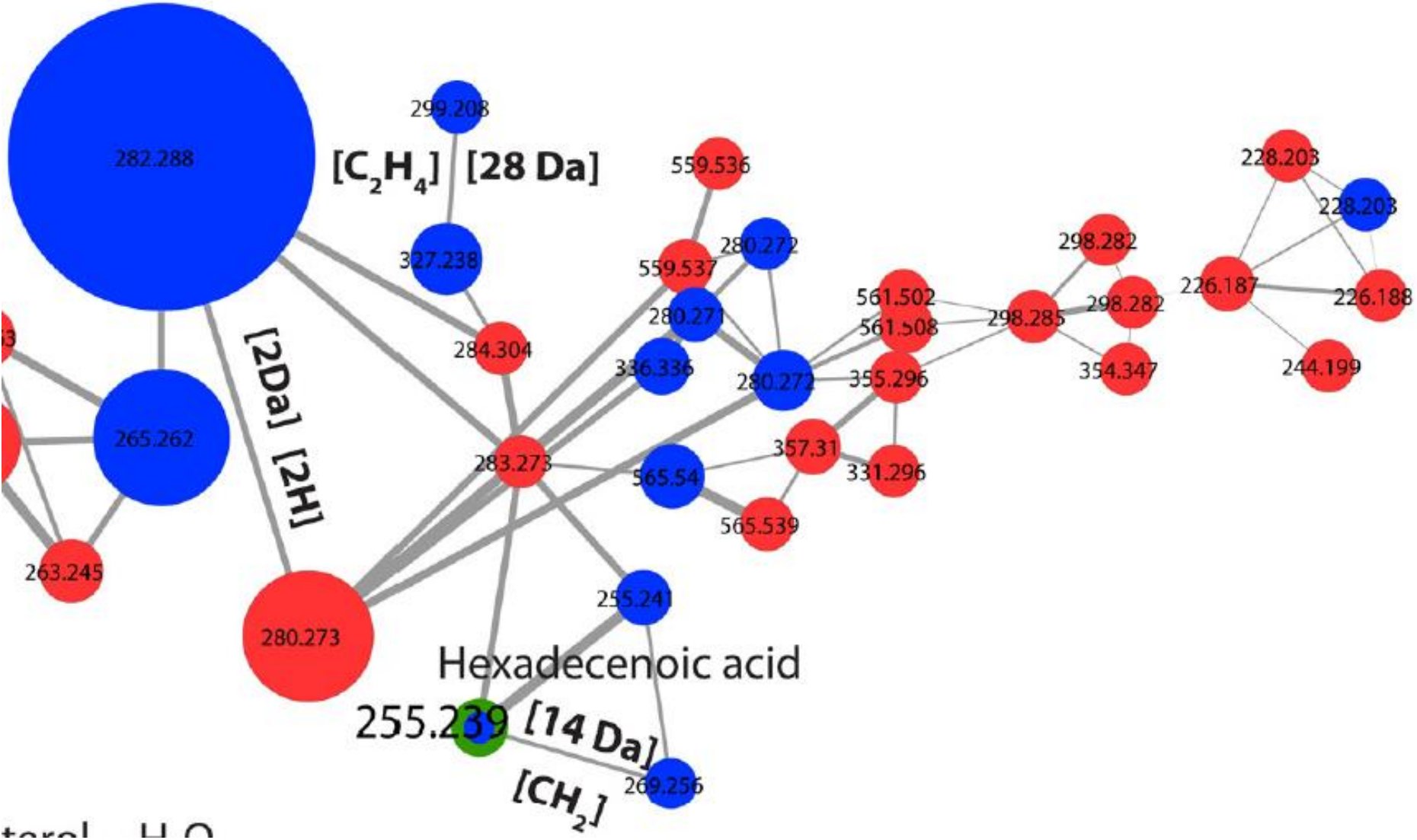
The sample set consisted of organic solvent extractions of ten anatomically distinct locations of ex-plant lung tissue and twenty *Pseudomonas* spp. isolates from a CF patient. In the current study, UPLC-ESI tandem MS data was collected on methanol and ethyl acetate extracts in the positive mode and organized using molecular networking (Fig. 2). The size of nodes represents intensity of the parent ion. As an example, one of the clusters is highlighted in Fig. 2b showing mass shifts of 2 Da, 14 Da and 28 Da between nodes suggesting a molecular family of fatty acids or lipids. The molecular networks form the lungs revealed matches to lipids, fatty acids, sterols, as well as drugs that were –administered

International Journal of Mass Spectrometry 377 (2015) 719–727

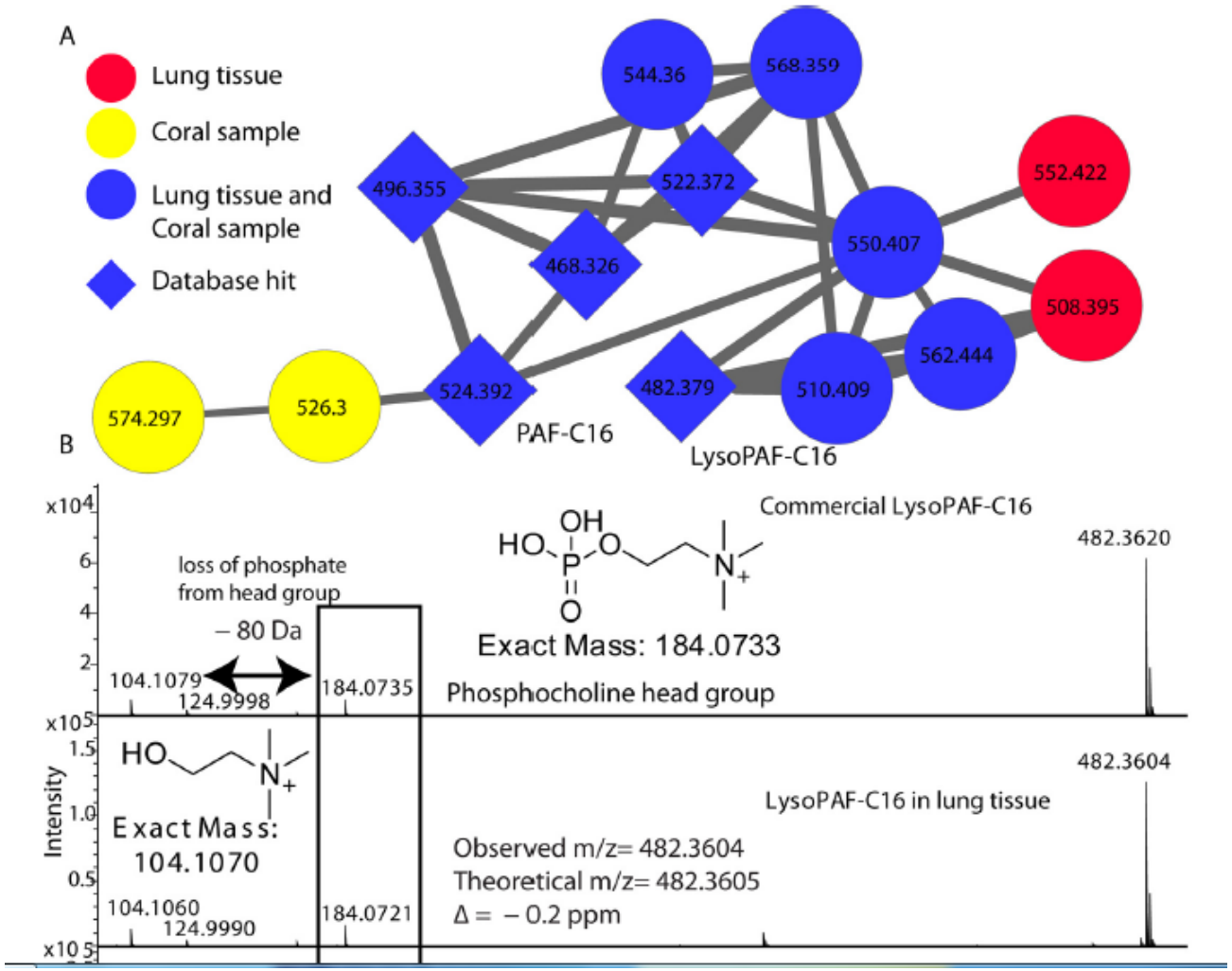
Cholestérol



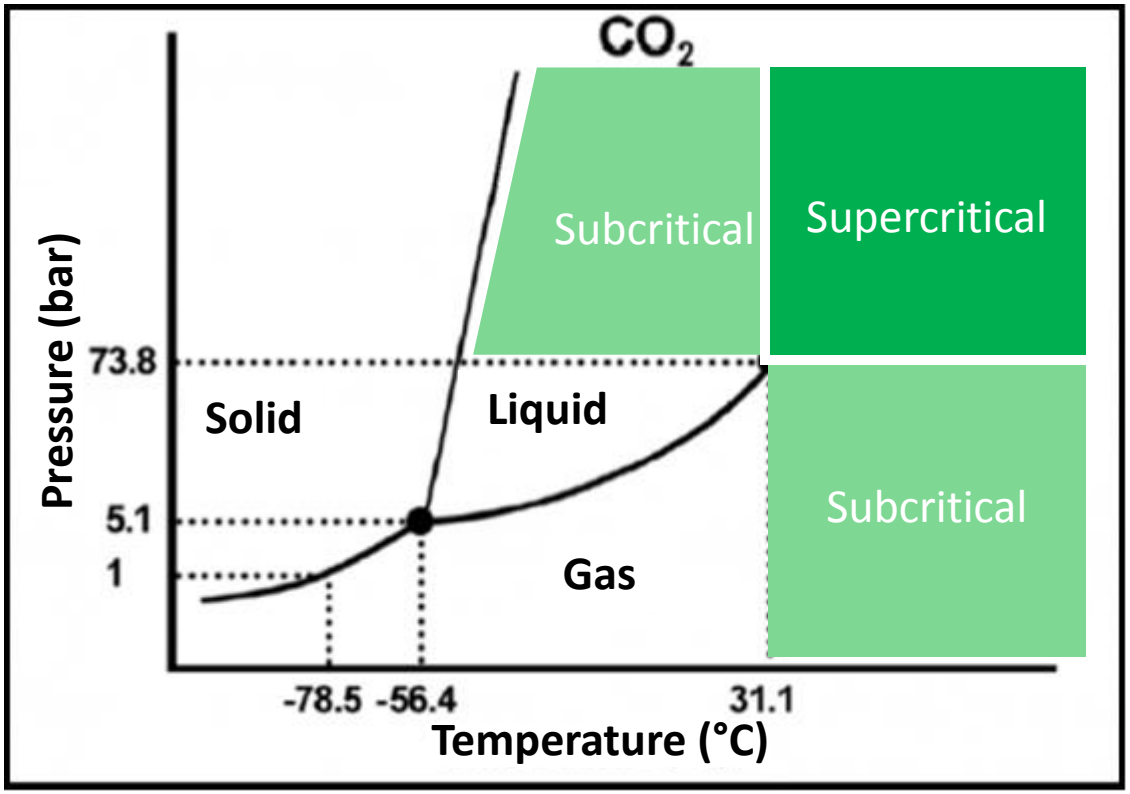
Données redondantes …



Initialement …
pour l'analyse de séquences peptidiques



CO_2


CO2
Subcritical
Supercritical
Pressure (bar)
73.8
Solid
Liquid
5.1
1
Gas
Subcritical
-78.5
-56.4
31.1
Temperature (°C)

Supercritical Fluid Chromatography

SCIENCE, VOL. 222

21 OCTOBER 1983

Dennis R. Gere

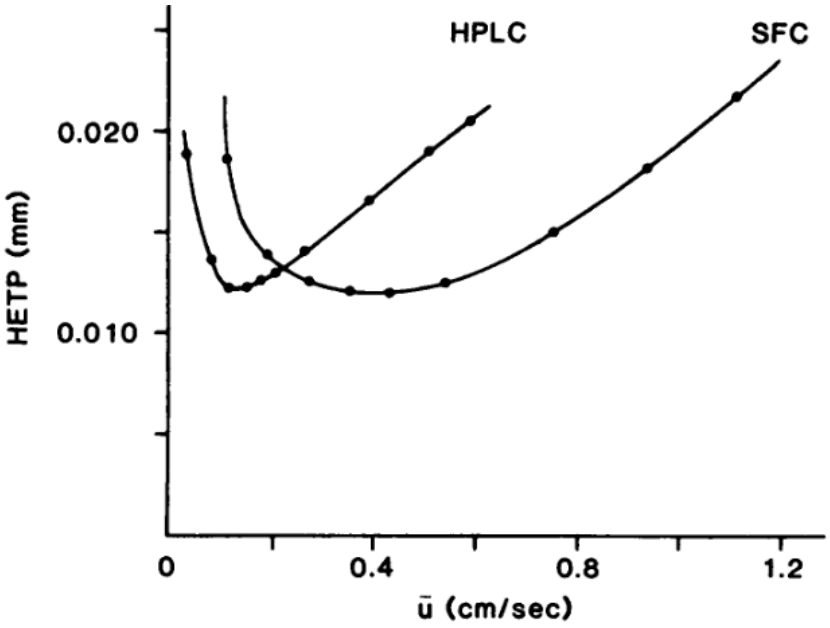


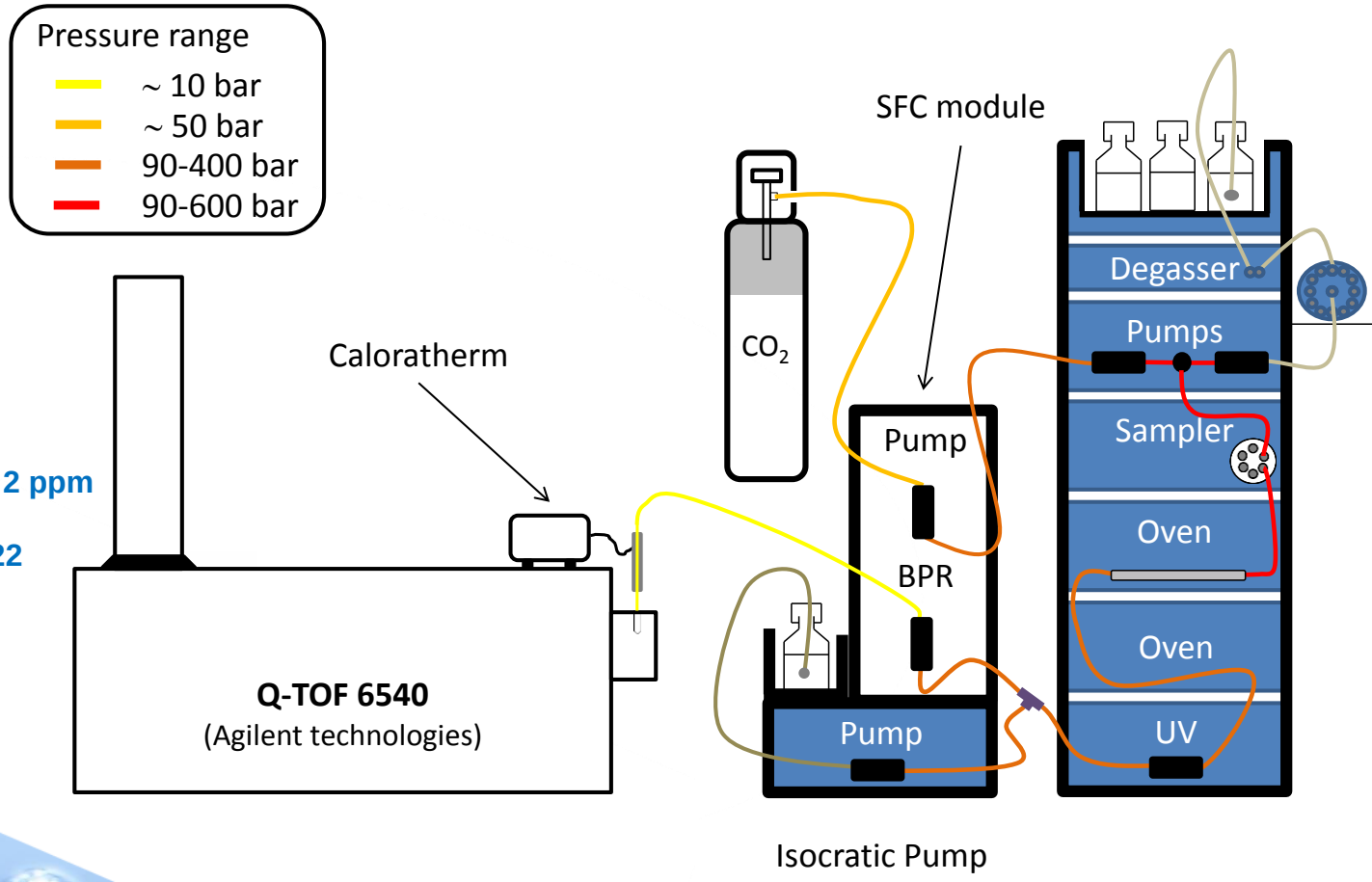
Fig. 1. Van Deemter plots for chromatographic data for HPLC and SFC elution of pyrene (*27*); *HETP* is height equivalent to a theoretical plate.

Table 1. Typical values of parameters important in chromatographic band broadening.

Parameter	GC	SFC	HPLC
Diffusion coefficient, $D_{1,2}$ (cm^2/sec)	10^{-1}	10^{-4}	10^{-5}
Density (g/cm^3)	10^{-3}	0.8	1
Viscosity (g/cm-sec)	10^{-4}	5×10^{-4}	10^{-2}

SFC coupled to a Q-TOF mass spectrometer

- **SFC 1260 Infinity (Agilent Technologies)**
- **Q-TOF 6540 (Agilent Technologies)**



- **Mass accuracy down to 2 ppm**
- **High mass resolution :**
  20 000 - 40 000 at *m/z* 922
- **MS/MS capabilities**

Choice of the column

N°	Polarité de la phase	Colonne	Fournisseur
1	apolaire	Octadécyle silane C18 Eclipse 150 * 2,1 mm et 1,8 µm de diamètre de particules	Agilent
2		Carbone graphite poreux Hypercarb^{R} 100 * 2,1 mm et 3 µm de diamètre de particules et 100 * 2,1 et 5µm de diamètre de particules	Thermo
3		Zirconium ZR carbon 150 * 2,1 mm et 3µm de diamètre de particules	Supelco
4	Hybride	Torus 1-aminoanthracène (1-AA) 150 * 2,1 mm et 1,7 µm de diamètre de particules	Waters
5		Torus 2-picolylamine (2-PIC) 150 * 2,1 mm et 1,7 µm de diamètre de particules	Waters
6		Torus diéthylamine (DEA) 150 * 2,1 mm et 1,7 µm de diamètre de particules	Waters
7	Polaire	Pentafluorophényl (PFP) 150 * 2 mm et 3 µm de diamètre de particules	Agilent
8		Diphényl 250 * 2 mm et 3 µm de diamètre de particules	Agilent
9		Ethyl pyridine (EP) 100 * 2,1 mm et 1,5 µm de diamètre de particules	Agilent
10		Cyano (CN) 100 * 2,1 mm et 1,8 µm de diamètre de particules	Agilent
11		Silanol (Rx sil) 150 * 4,6 mm et 5 µm de diamètre de particules et 100 * 2,1 mm et 1,8 µm de diamètre de particules	Agilent

Etude de l'activité anti-CHIKV d'*Euphorbia* de Corse

Investigation d'*Euphorbia amygdaloides* subsp *semiperfoliata*

Euphorbia amygdaloïdes subsp _semiperfoliata_
Materiel végétal sec (m = 450 g) (LF-015)

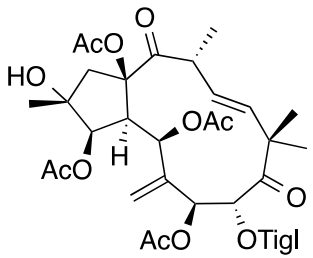
Extrait AcOEt (Zypertex)
m = 30,4 g / σ = 6 %
EC_{90} CHIKV < 0.8 µg / mL

10 Fractions (**A**-**J**)

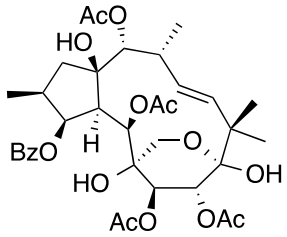
Fraction **I** = 5,1 g
EC_{90} CHIKV < 0.8 µg/mL

Plusieurs étapes chromatographiques (!)

15 molécules isolées

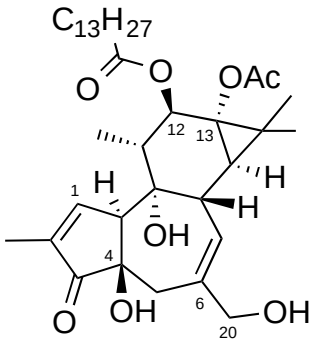


Ester de jatrophane (σ = 0.0006 %)
EC_{50} CHIKV = 0.76 ± 0.14 µM/ SI = 208



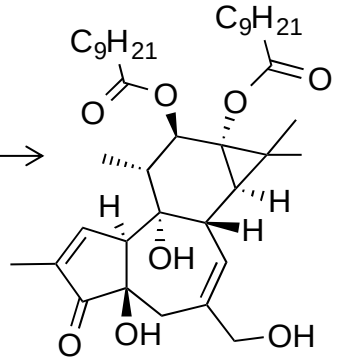
Jatrohemiketal

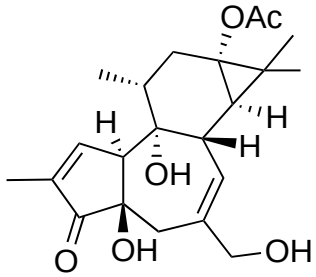
Évaluation de phorboids commerciaux connus pour leurs activité anti-HIV



TPA
EC_{50} CHIKV = 2.9 ± 0.3 nM
SI = 1965

P-12,13-DD
EC_{50} CHIKV = 6.0 ± 0.9 nM
SI = 686





Prostratin
EC_{50} CHIKV = 2.7 ± 1.2 µM
SI = 23

Nothias-Scaglia, L.-F.; Dumontet, V.; Neyts, J.; Roussi, F.; Costa, J.; Leyssen, R.; Litaudon, M.; Paolini, J. *Fitoterapia* **2015**, (in press). Nothias-Scaglia, L.-F.; Retailleau, P.; Paolini, J.; Neyts, J.; Dumontet, V.; Roussi, F.; Leyssen, P.; Costa, J.; Litaudon, -M. *J. Nat. Prod*. **2014**, 77 (6), 1505–1512. Nothias-Scaglia, L.-F.; Gallard J.-F.; Dumontet V.; Roussi F.; Costa J. Bogdan I. I.; Paolini J.; Litaudon M. *J. Nat. Prod*. **2015** (accepted). Bourjot, M.; Delang, L.; Nguyen, V. H.; Neyts, J.; Guéritte, F.; Leyssen, P.; Litaudon, M. *J. Nat. Prod.* **2012**. Nothias-Scaglia, L.-F.; Pannecouque, C.; Renucci, F.; Delang, L.; Neyts, J.; Roussi, F.; Costa, J.; Leyssen, P.; Litaudon, M.; Paolini, J. *J. Nat. Prod.* **2015**.

Etude de l'activité anti-CHIKV d'*Euphorbia* de Corse

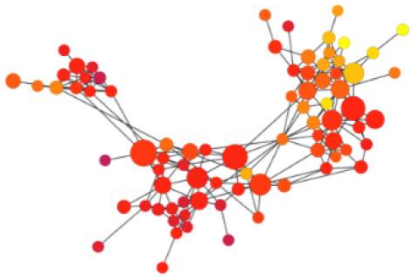
Investigation d'*Euphorbia amygdaloides* subsp *semiperfoliata* : Récolte n°2

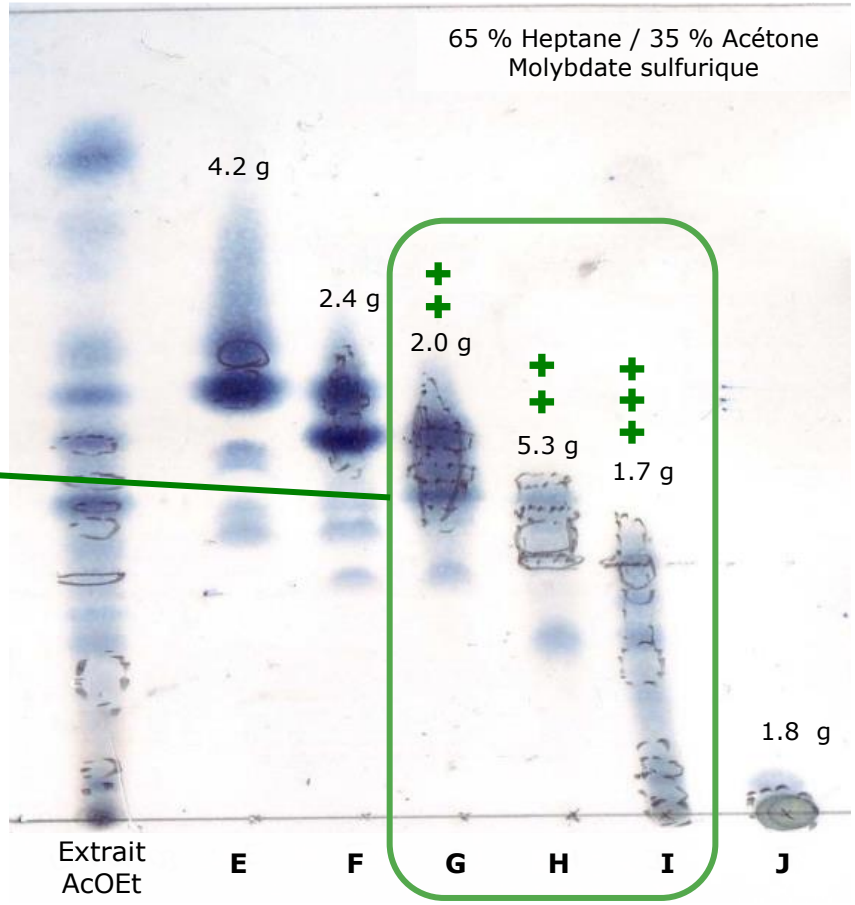
Euphorbia amygdaloïdes subsp *semiperfoliata*
Materiel végétal sec (m = 370 g) (LF-023)

Extrait AcOEt (Zypertex)
m = 24,7 g / σ = 6.7 %
EC_{90} CHIKV < 0.8 µg / mL

10 Fractions (**A-J**)

Analyse par SFC-MS/MS et
génération de réseaux moléculaires.

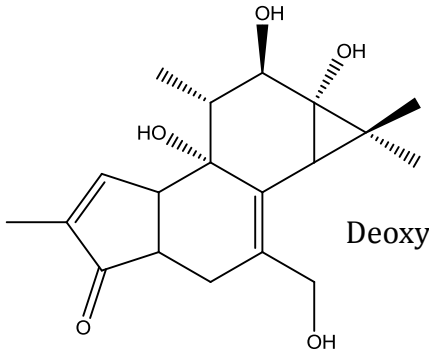




Résultats validés comme **CHIKV** spécifiques par observation au microscope

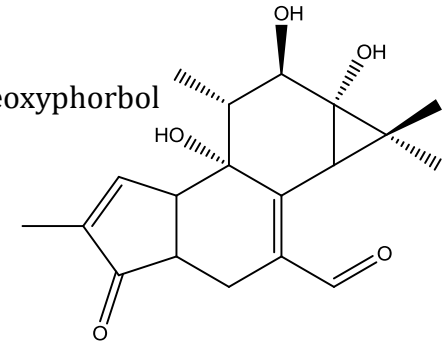
EC_{90} CHIKV en µg/ml	**< 0,8**	> 100	> 100	**1,2**	**3,4**	**< 0,8**	77
KB à 10 µg/ml en % d'inhibition	16 ± 13 %	0 %	0 %	0 %	51 ± 6 %	45 ± 15 %	0 %

Prenols: diterpinoids

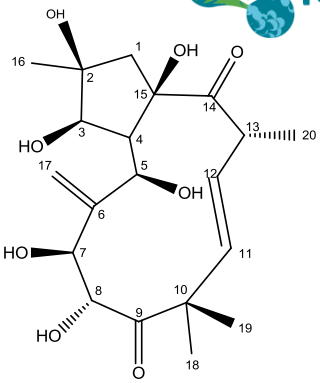


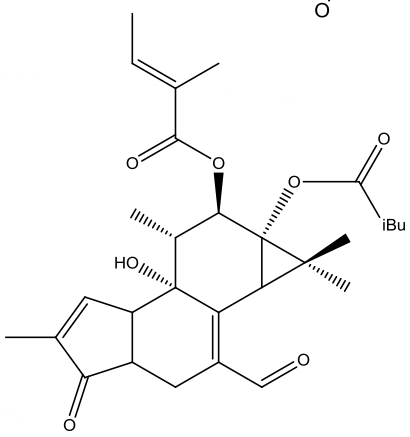
Deoxyphorbol

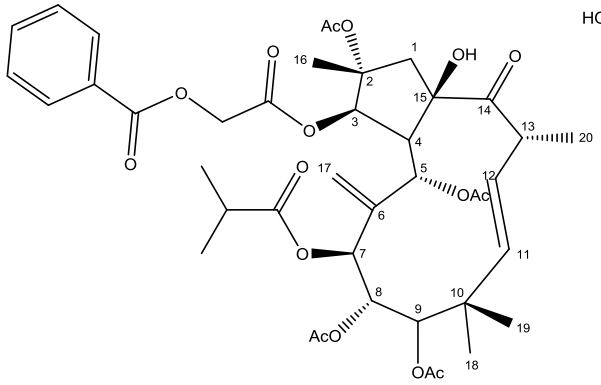
Dideoxyphorbol

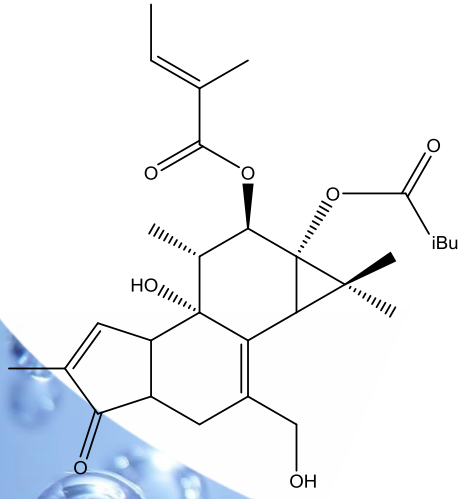


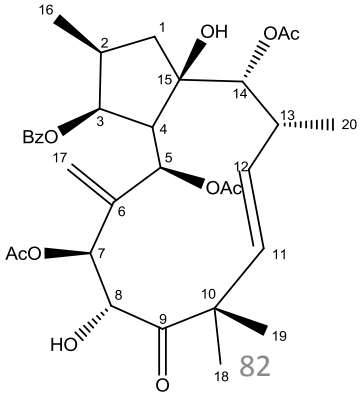
Jatrophane







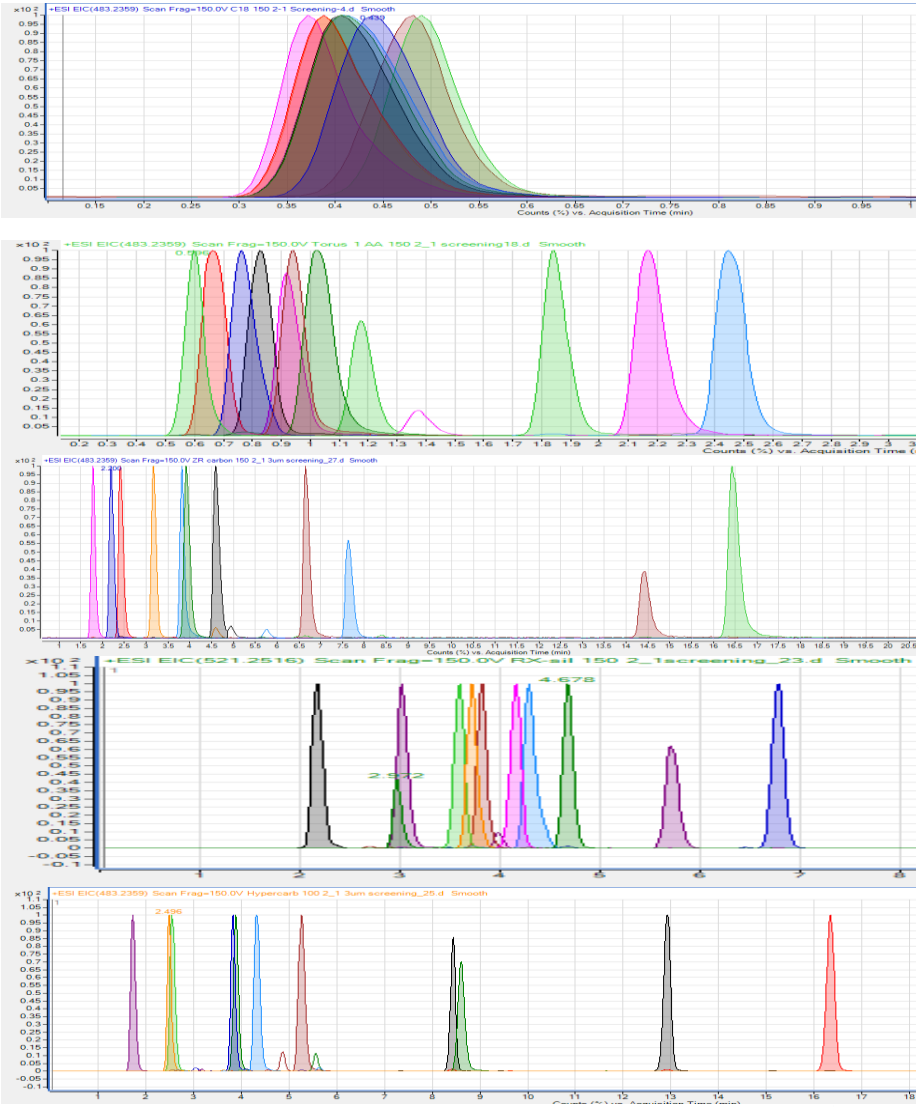




Analyses by LC with strong percentage of organics solvent
30 mL acétonitrile by one LC run

Collaboration with ICSN and University of Corte

Choice of the column



C18: 1 min

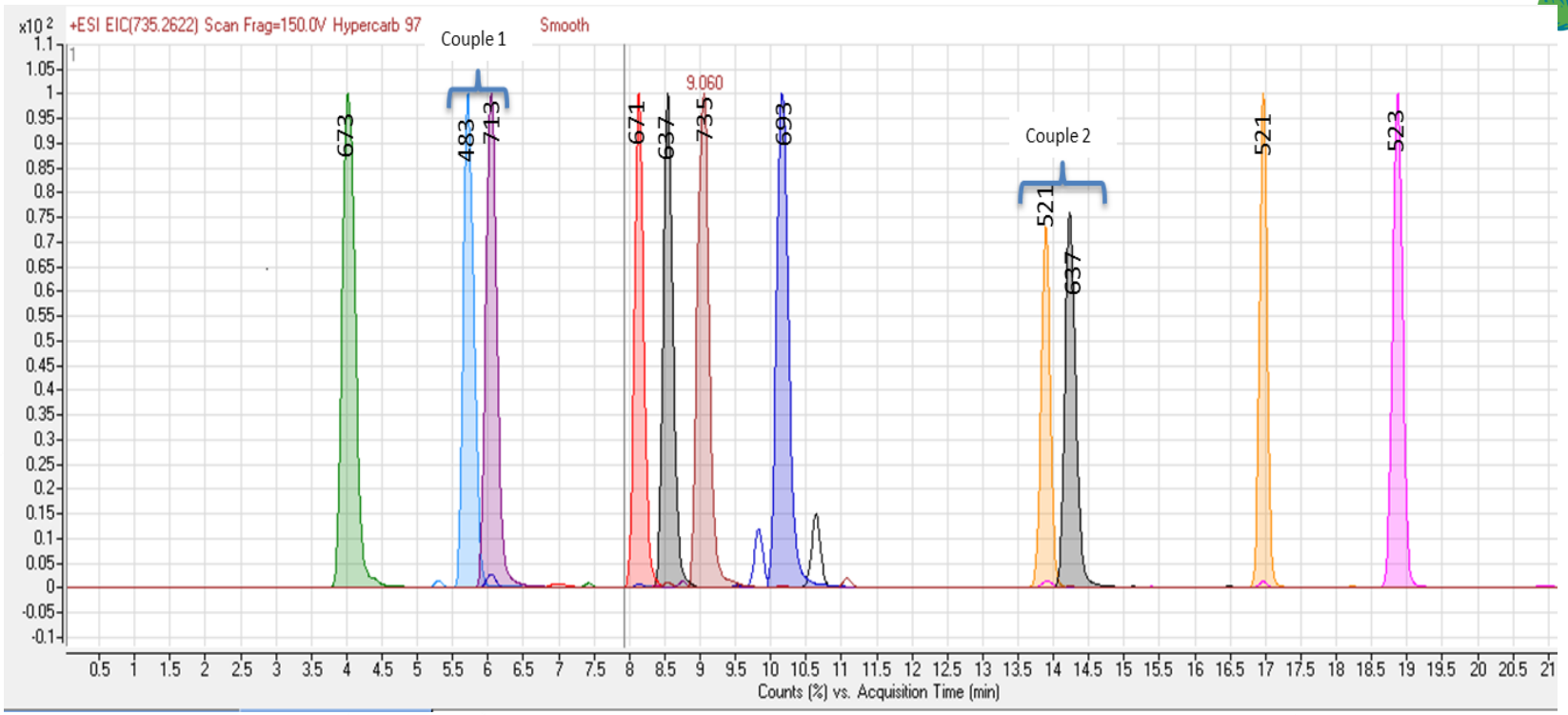
Torus 1-AA: 3 min

ZR carbon: 18 min

Si: 7 min

Hypercarb: 17 min

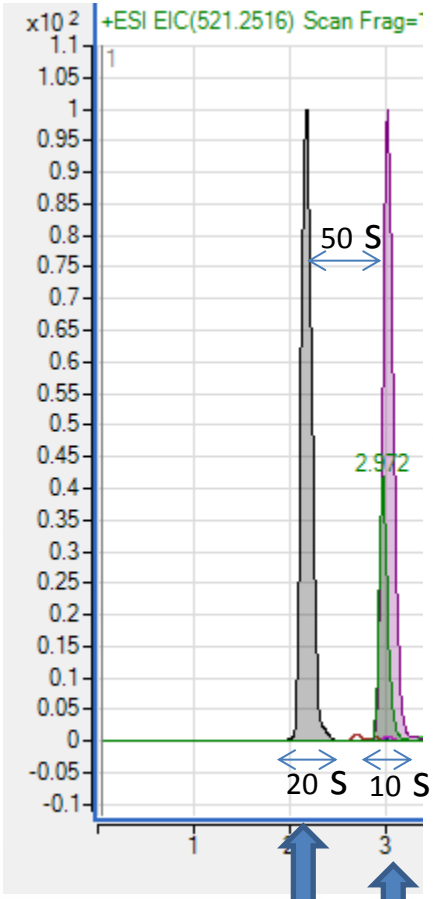
Hypercarb



Time (min)	EtOH (%)
0	3
3	3
13	10
17	20
20	20
21	3
23	3

Flow rate: 1.5 mL/min

DDA analysis



until 16 spectra acquisitions in MSMS

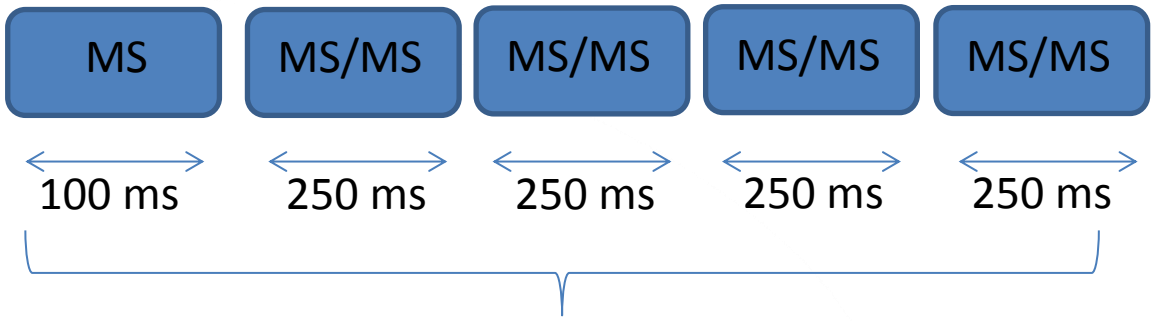
until 8 spectra acquisitions in MSMS

to attribute a peak to a compound 5 at 10 MS/MS spectra are necessary

Our condition in DDA mode:

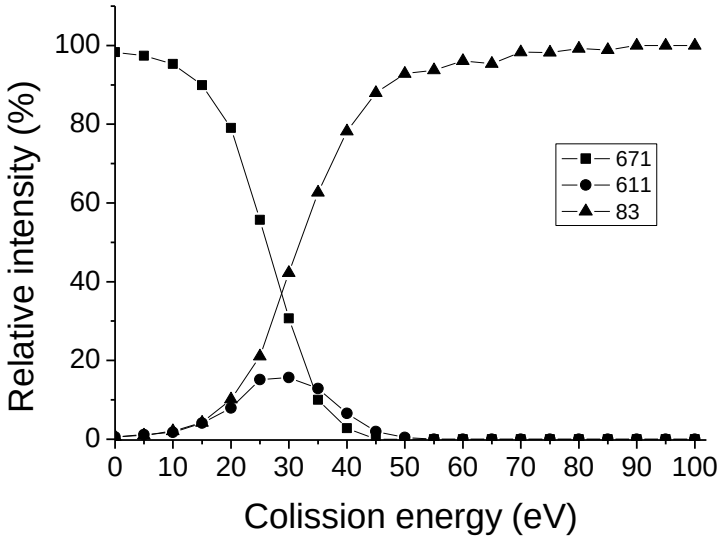
MS 10 spectres/s

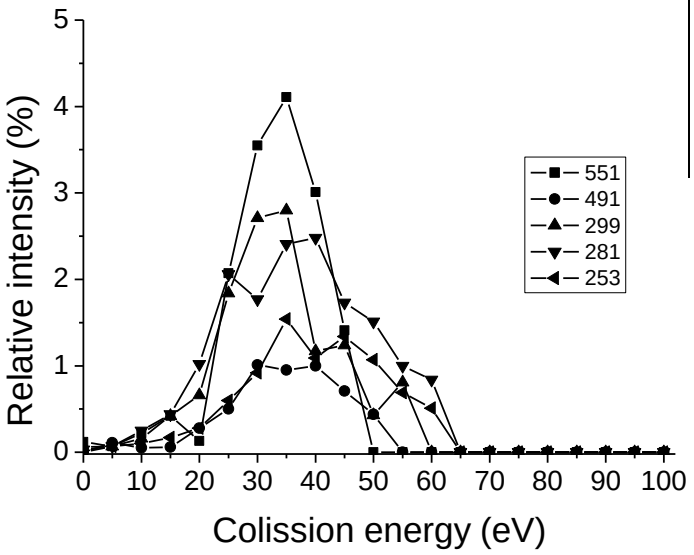
MS/MS a maximum of 4 precursor ions per cycle

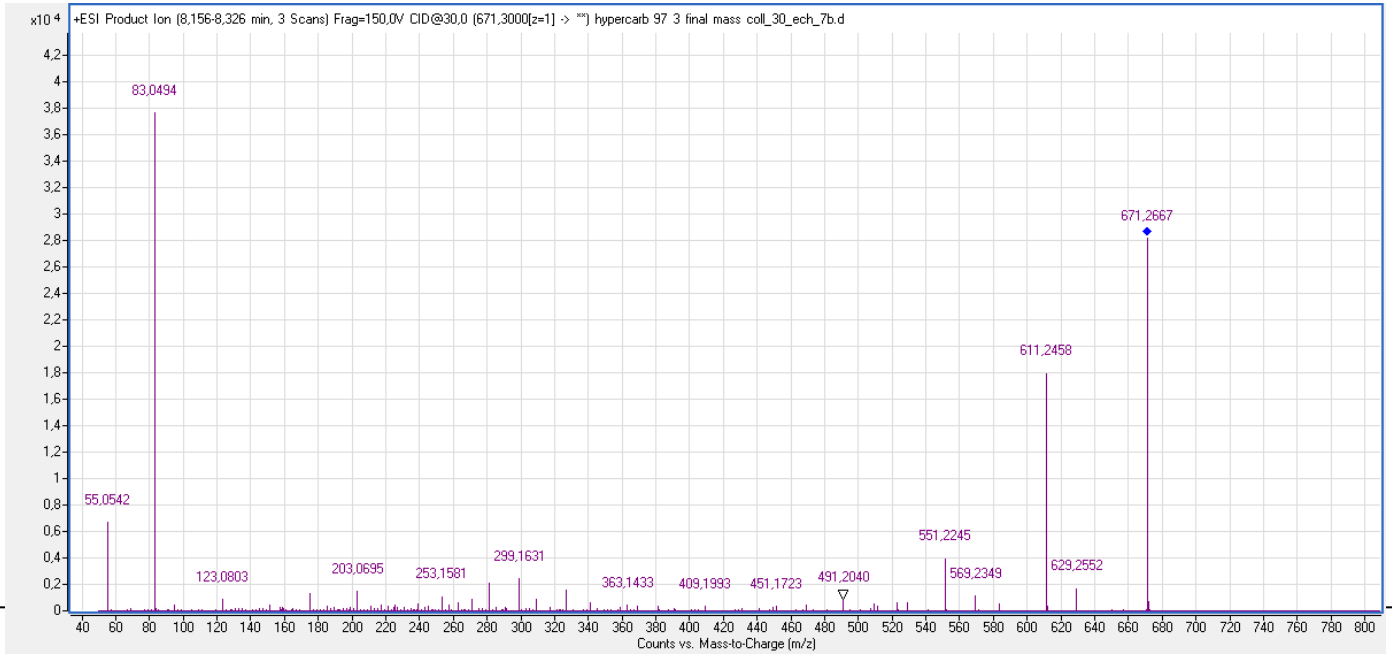


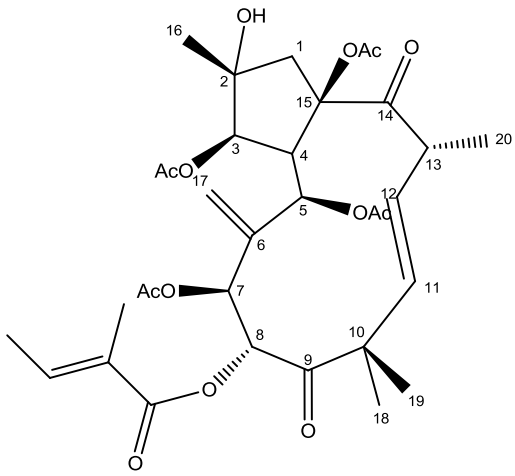
Total operating time: 1.2 s between 2 scans

Energie de collision …

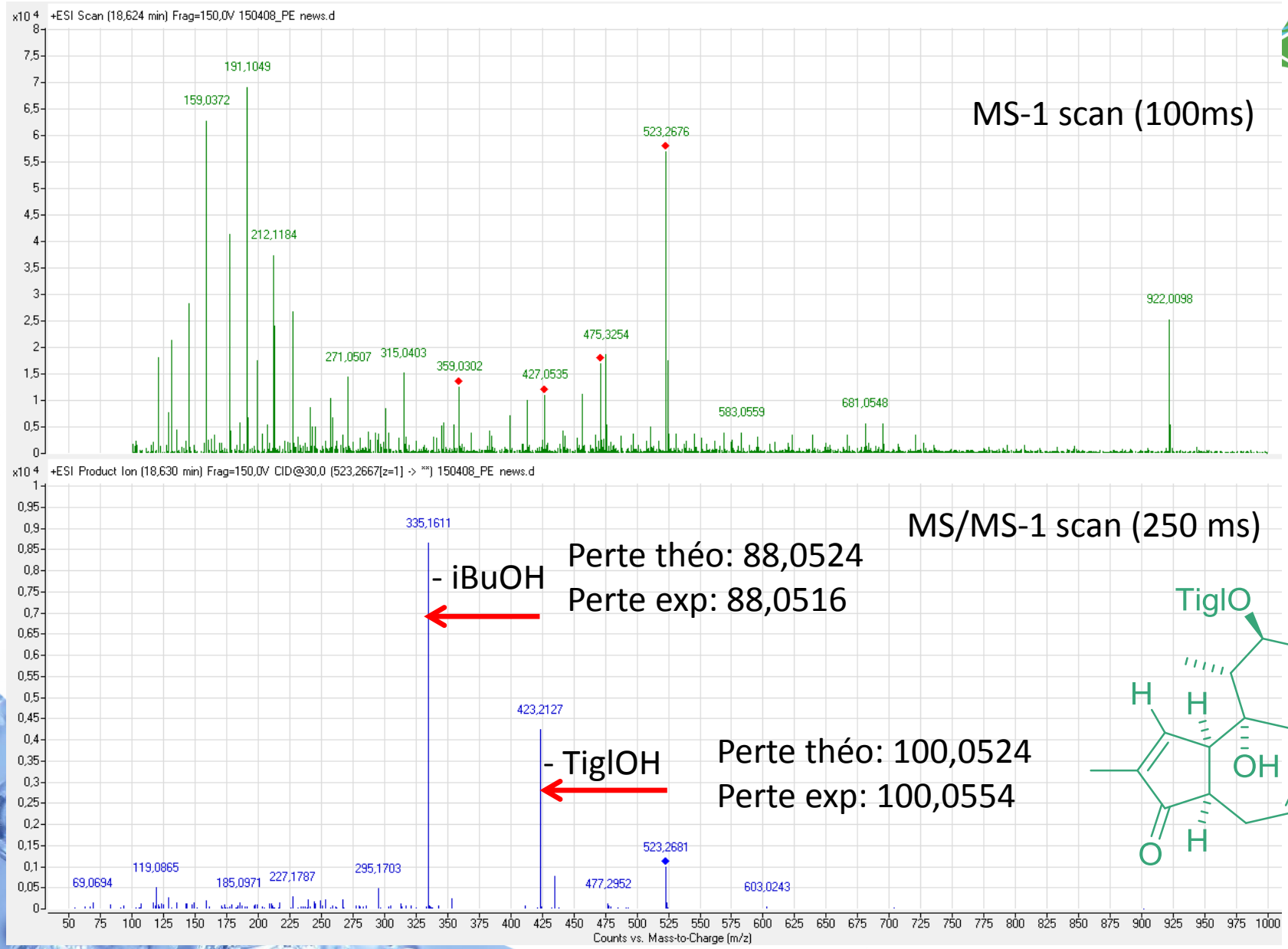


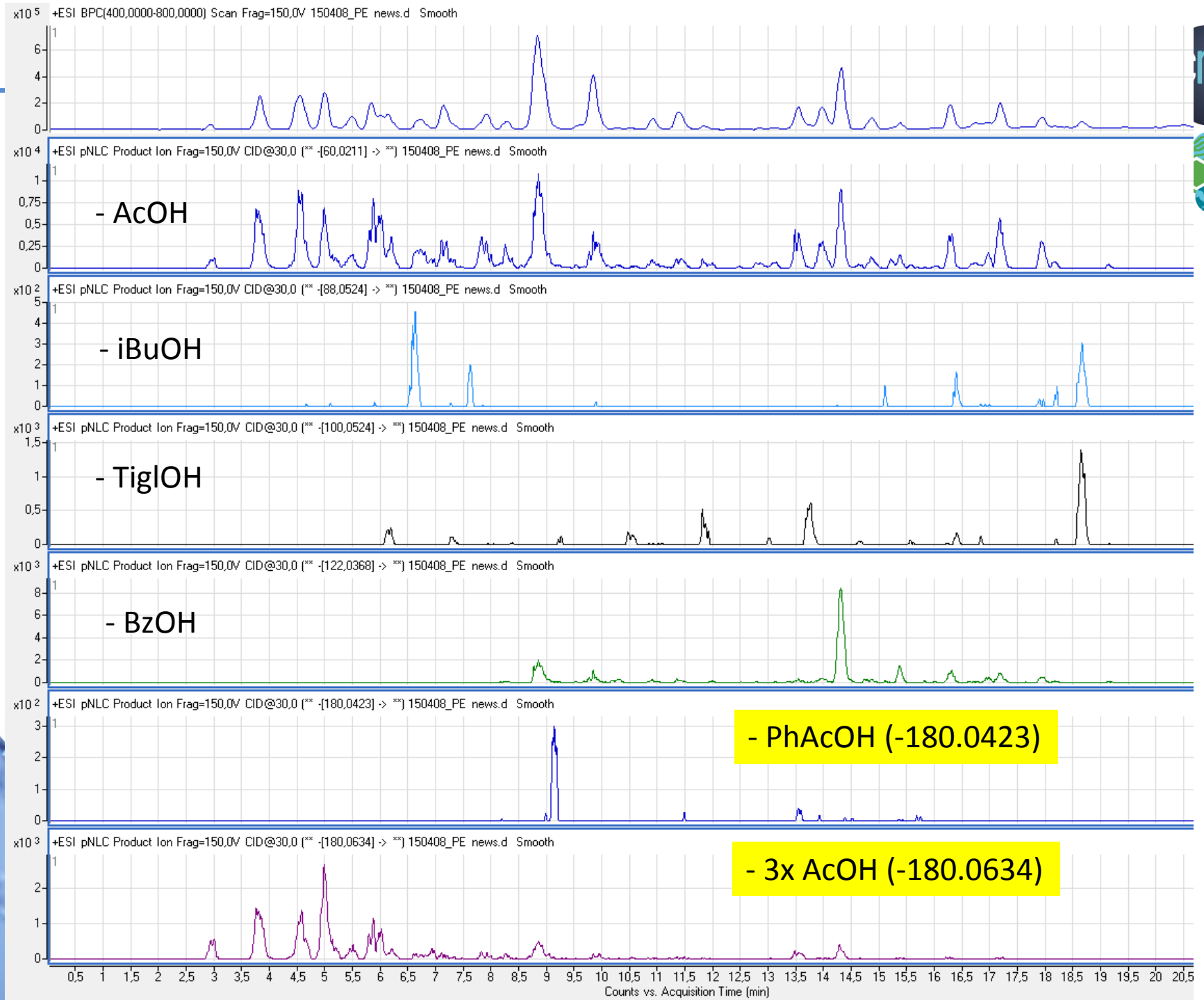






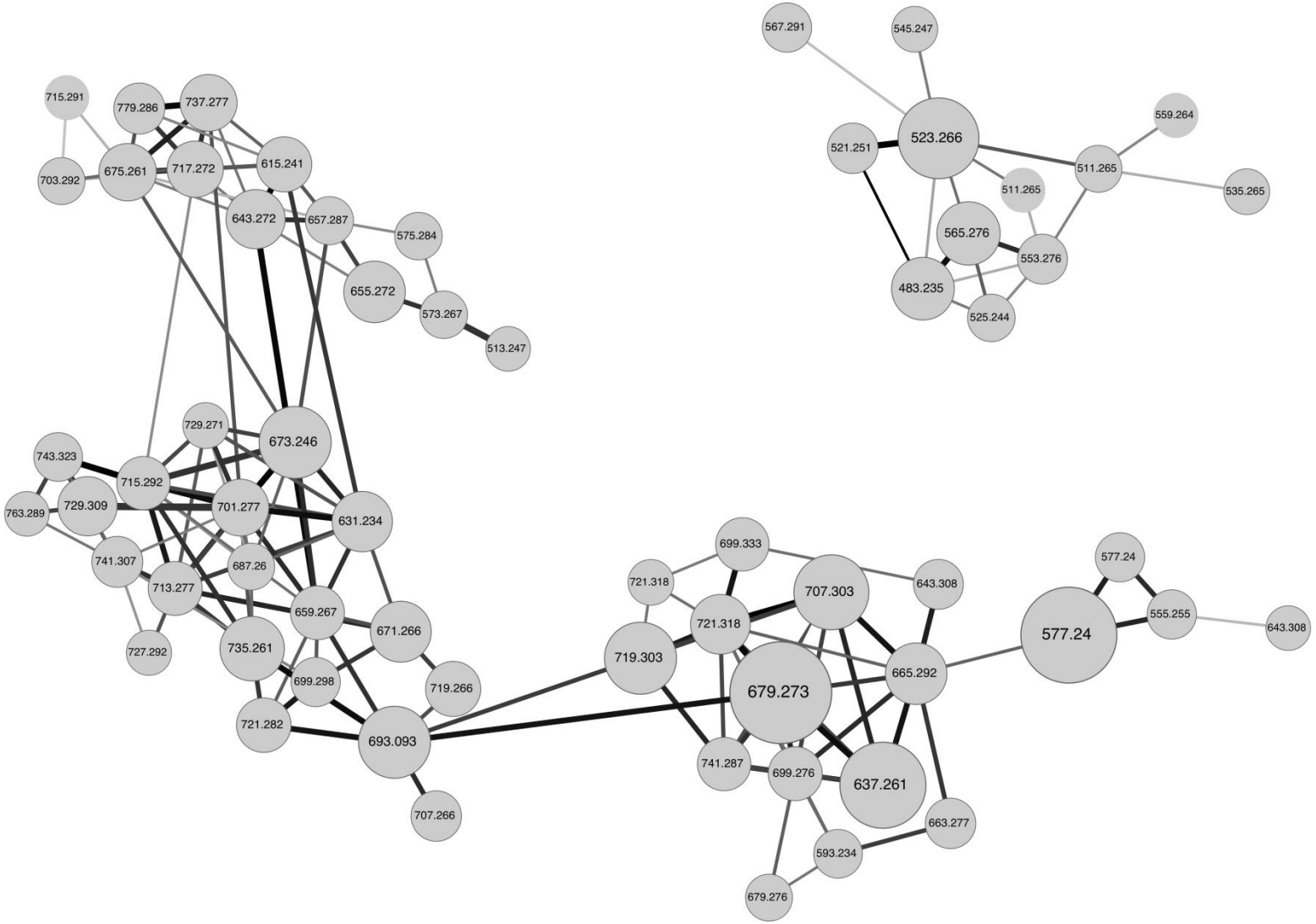
DDA analysis




x10 5 +ESI BPC(400,0000-800,0000) Scan Frag=150,0V 150408_PE news.d Smooth
x10 4 +ESI pNLC Product Ion Frag=150,0V CID@30,0 (** -[60,0211] -> **) 150408_PE news.d Smooth
- AcOH
x10 2 +ESI pNLC Product Ion Frag=150,0V CID@30,0 (** -[88,0524] -> **) 150408_PE news.d Smooth
- iBuOH
x10 3 +ESI pNLC Product Ion Frag=150,0V CID@30,0 (** -[100,0524] -> **) 150408_PE news.d Smooth
- TiglOH
x10 3 +ESI pNLC Product Ion Frag=150,0V CID@30,0 (** -[122,0368] -> **) 150408_PE news.d Smooth
- BzOH
x10 2 +ESI pNLC Product Ion Frag=150,0V CID@30,0 (** -[180,0423] -> **) 150408_PE news.d Smooth
- PhAcOH (-180.0423)
x10 3 +ESI pNLC Product Ion Frag=150,0V CID@30,0 (** -[180,0634] -> **) 150408_PE news.d Smooth
- 3x AcOH (-180.0634)
Counts vs. Acquisition Time (min)

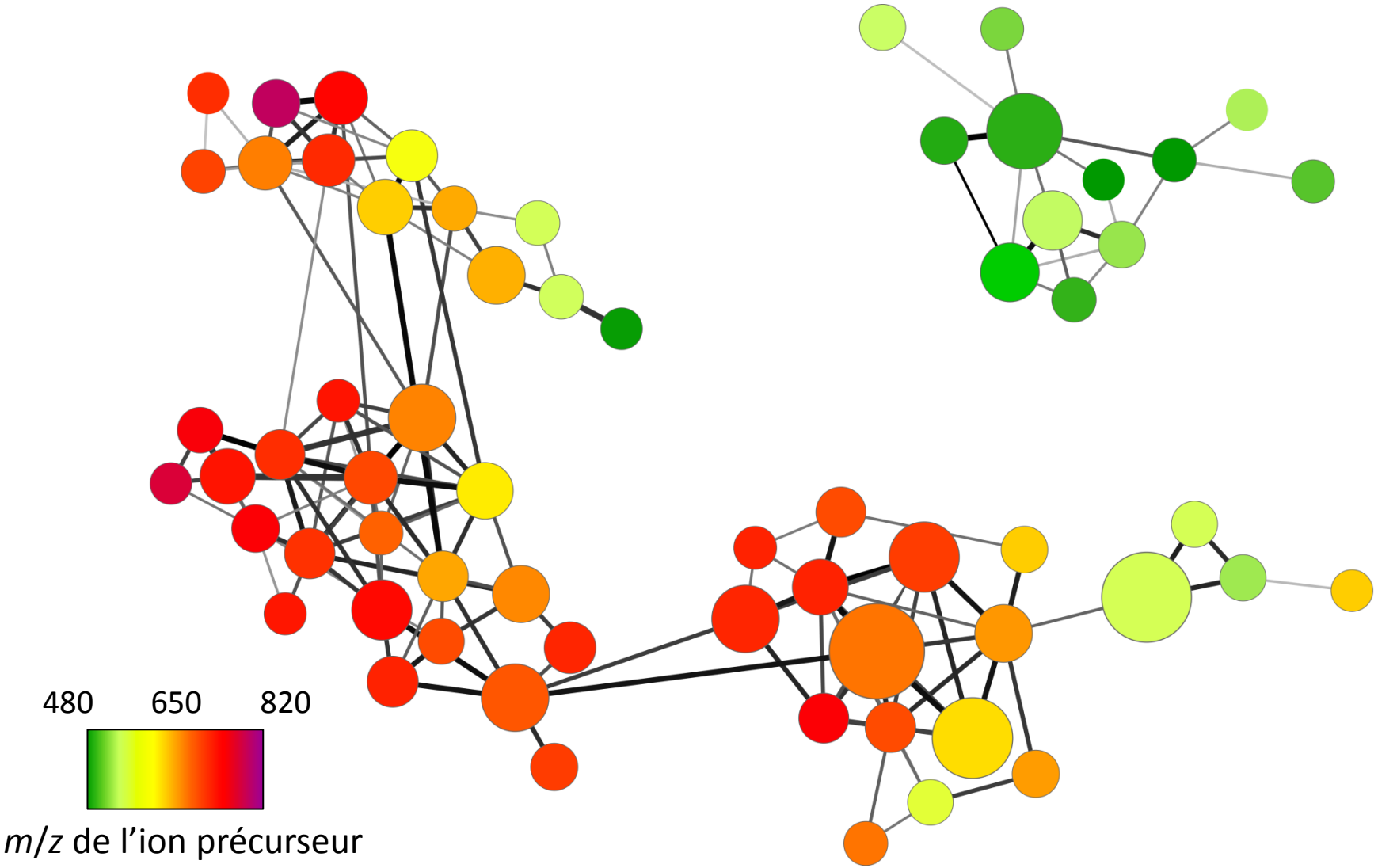
Etude de l'activité anti-CHIKV d'*Euphorbia* de Corse par réseaux moléculaires MS/MS

Représentation des réseaux moléculaires (analyses SFC-qTOF) des fractions anti-CHIKV



Etude de l'activité anti-CHIKV d'*Euphorbia* de Corse par réseaux moléculaires MS/MS

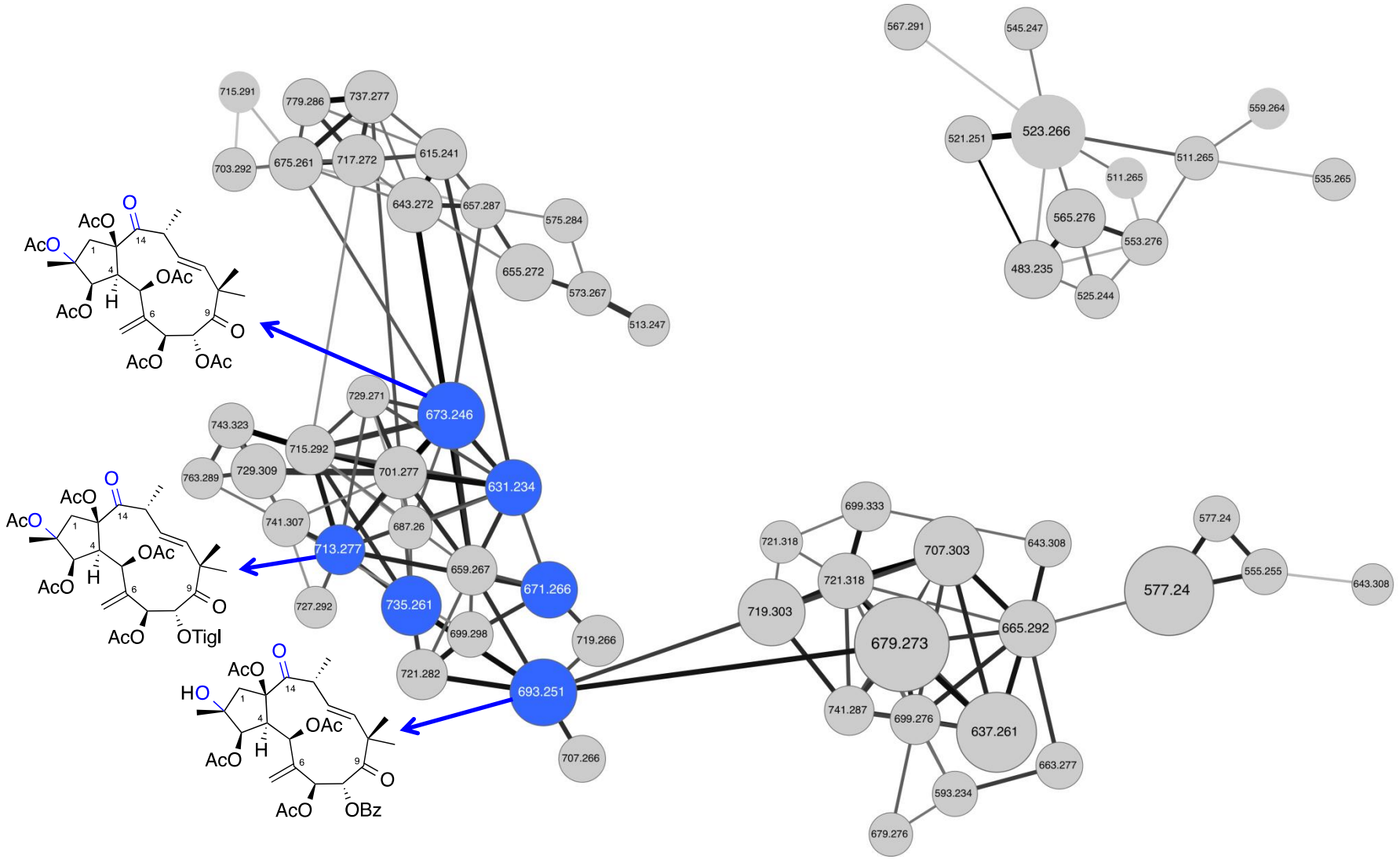
Représentation avec calque *m/z*



Etude de l'activité anti-CHIKV d'*Euphorbia* de Corse par réseaux moléculaires MS/MS

Ajout de spectres MS^2 de références isolées de la même plante:

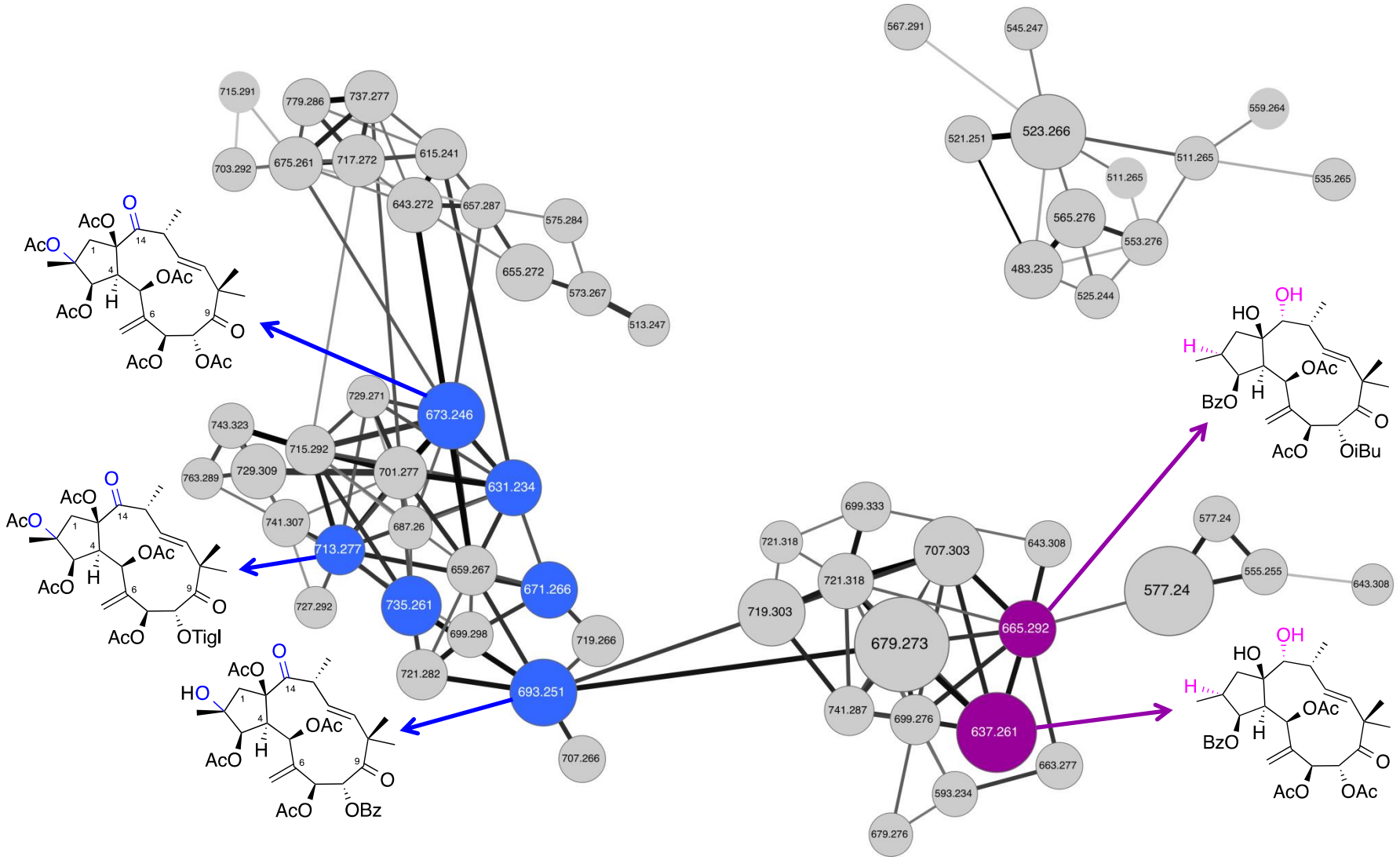
- Esters de jatrophane du groupe A



Etude de l'activité anti-CHIKV d'*Euphorbia* de Corse par réseaux moléculaires MS/MS

Ajout de spectres MS^2 de références isolées de la même plante:

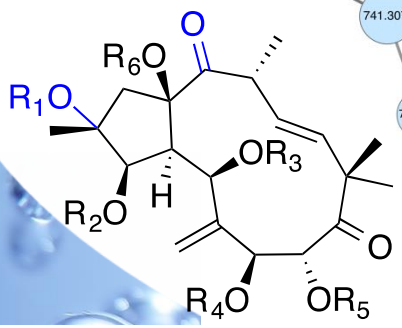
- Esters de jatrophane du groupe B



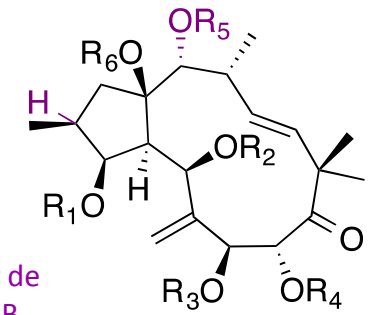
Molecular network

MS² fingerprint
m/z 313, 295, 285, ...
Esters de deoxyphorbol ?
Activité anti-CHIKV ?

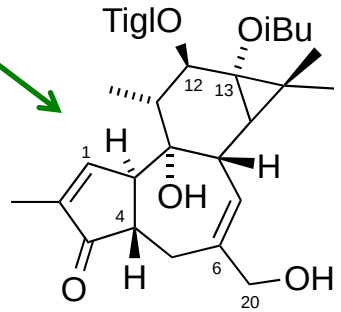
MS² fingerprint
m/z 311, 293, 283, ...
Autre type d'esters de jatrophane ?



MS² fingerprint d'ester de jatrophane du group A
m/z 327, 309, 299, ...

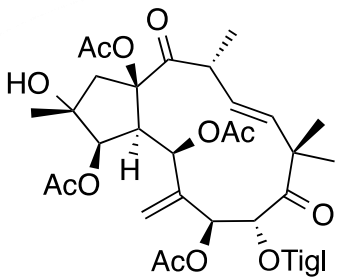


MS² fingerprint d'ester de jatrophane du group B
m/z 313, 295, 285, ...

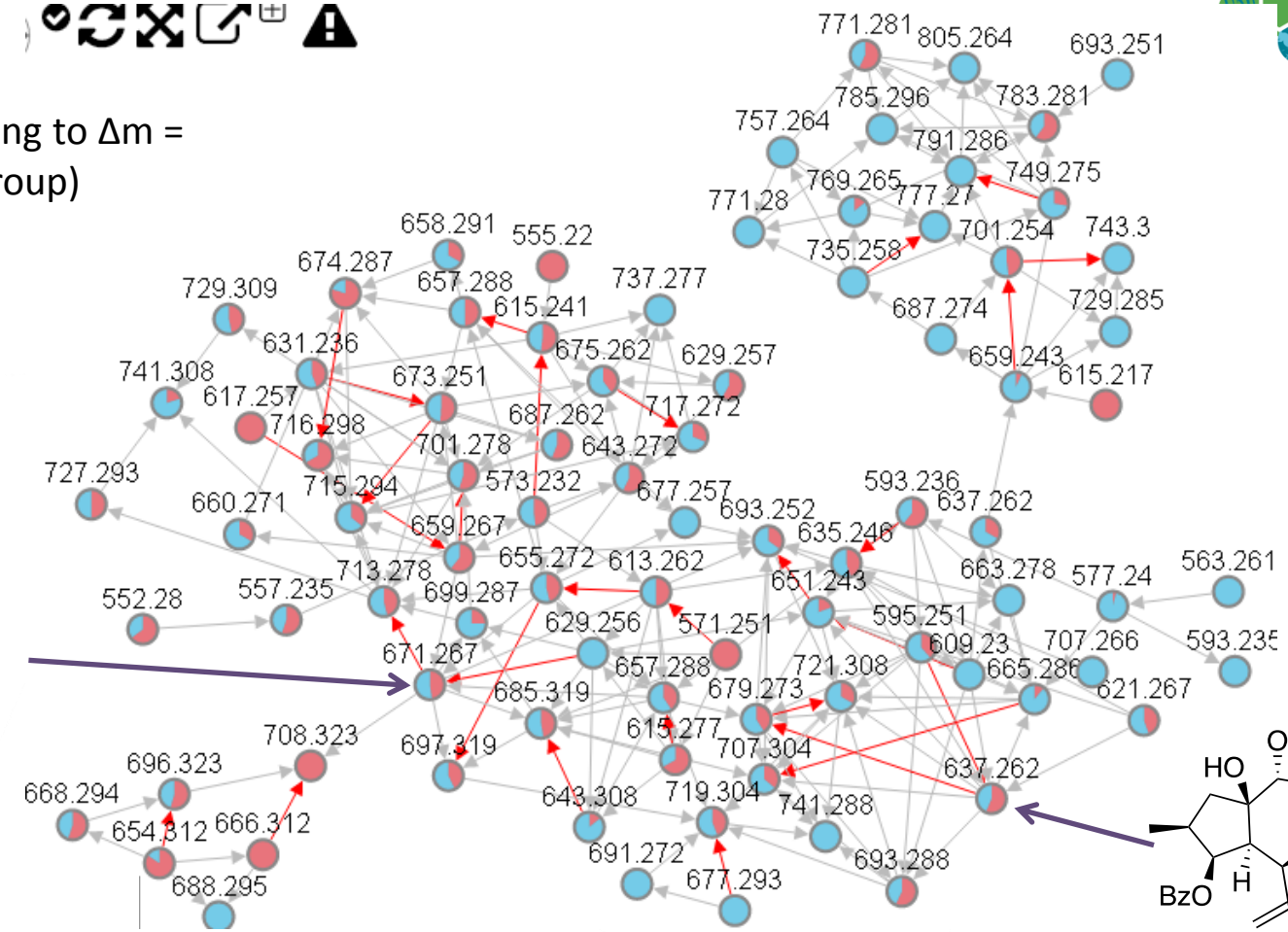


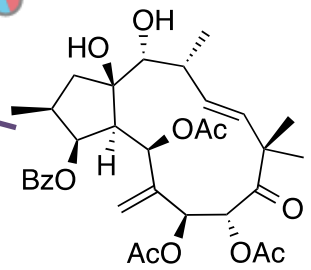
Molecular network

For example:
Red arrow corresponding to Δm = 42.011 uma (acetate group)



jatrophane ester A



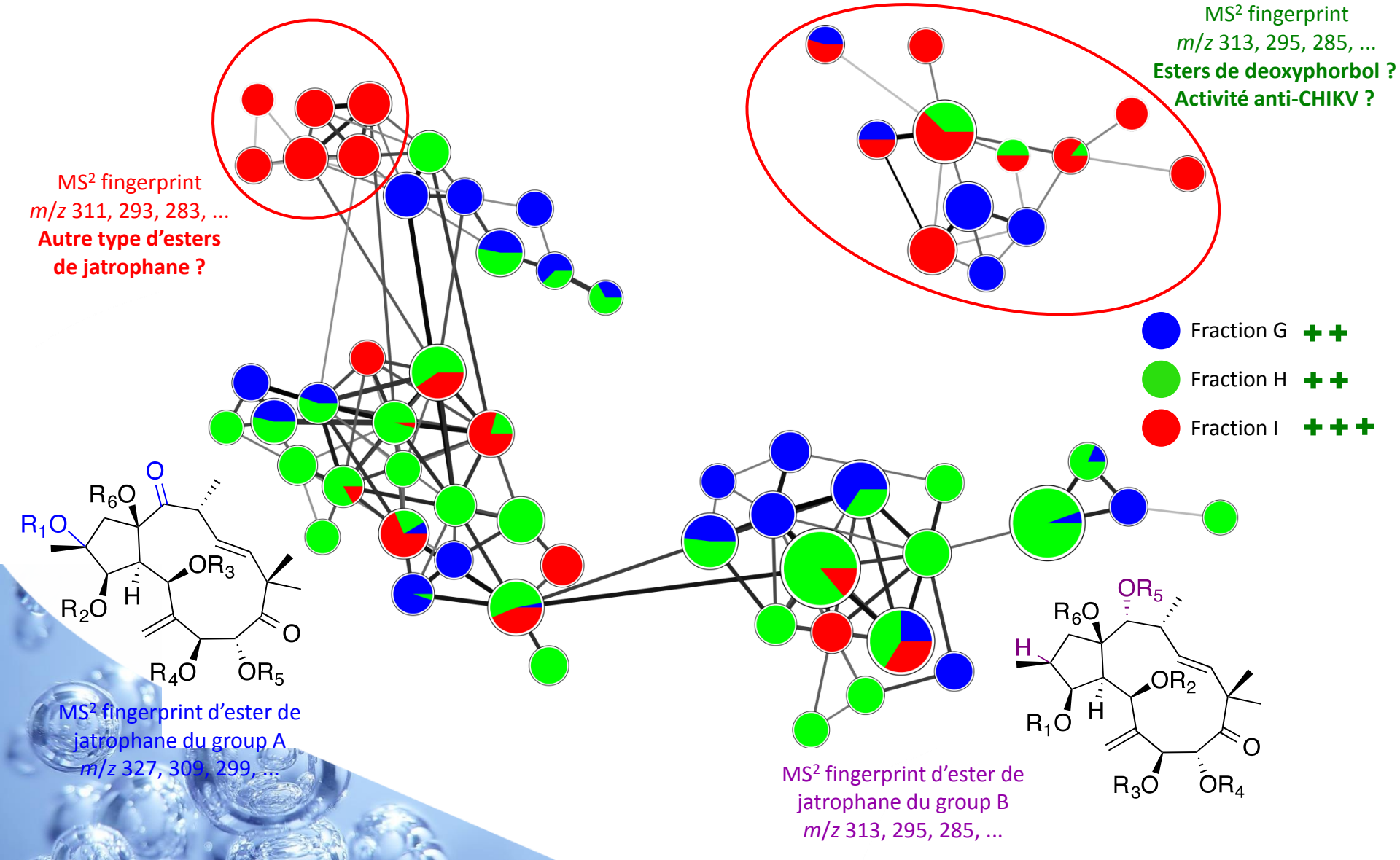


Jatrophane ester B

Latex
Extraction totale

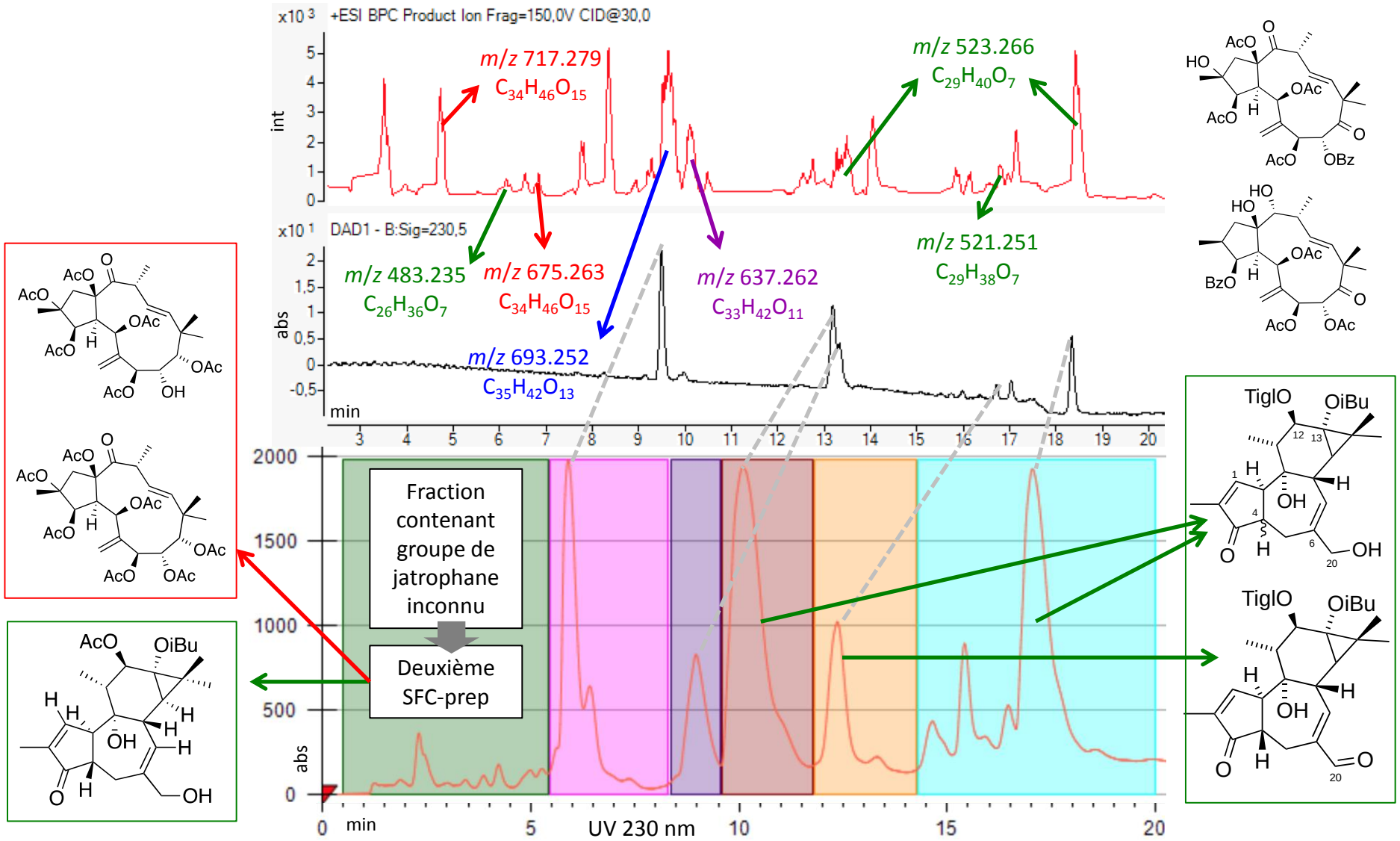
Etude de l'activité anti-CHIKV d'*Euphorbia* de Corse par réseaux moléculaires MS/MS

Représentation de la répartition par fraction



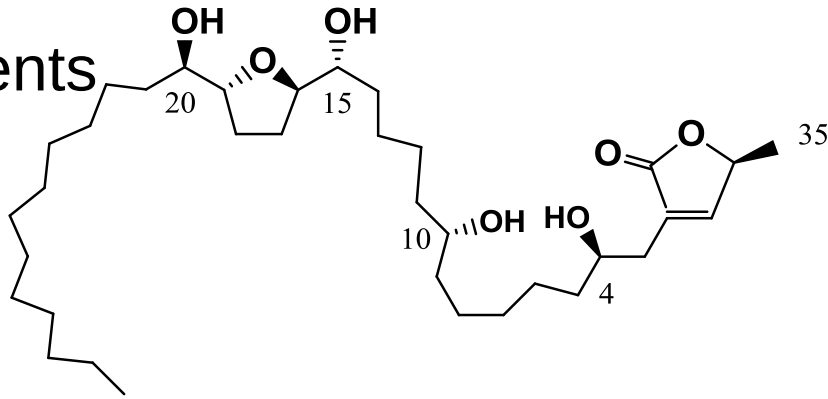
Etude de l'activité anti-CHIKV d'*Euphorbia* de Corse par réseaux moléculaires MS/MS

Purification par SFC-semiprep, guidée par molecular networking MS/MS



Aetiology: Annonaceae

- Environment: link with consumption of Annonaceae
  - Edible fruits (as *Annona muricata*)
  - Traditional medicine
  - High consumption for all patients
- Neurotoxicity of acetogenins
  - Lipophilic polyketides
  - Highly cytotoxic
  - Strong inhibitors of mitochondrial complex I
- Major safety concern for the FDA and ANSES
- What about the football world cup in Brazil ?



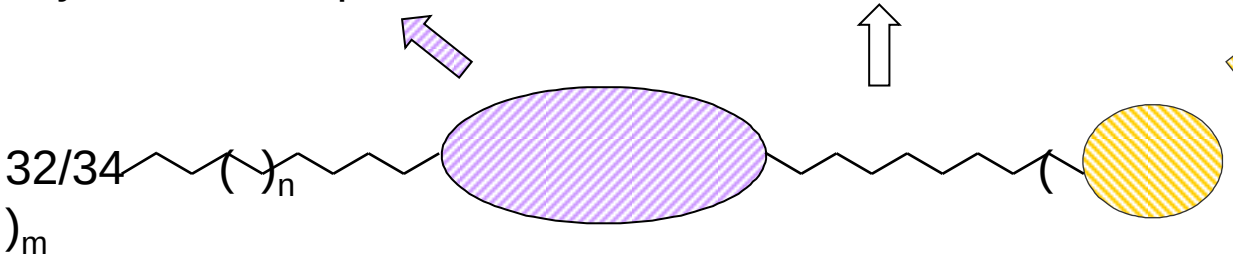
Structural diversity

Classification

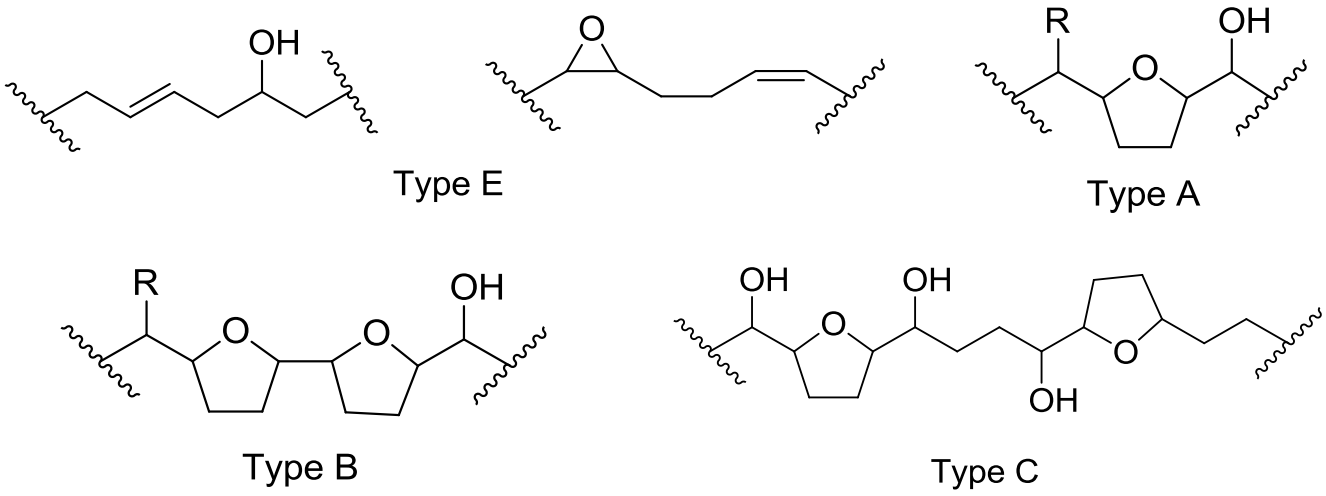
Polyoxygenated domain :
Tetrahydrofurans, epoxides...

+/- substitutions
Hydroxyls, ketones...

γ-methyl-γ-lactone

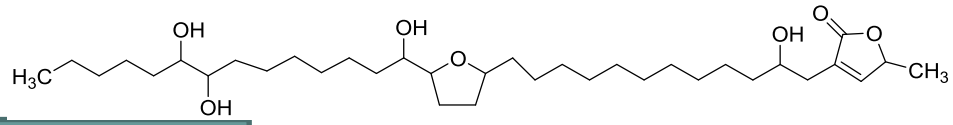


Types



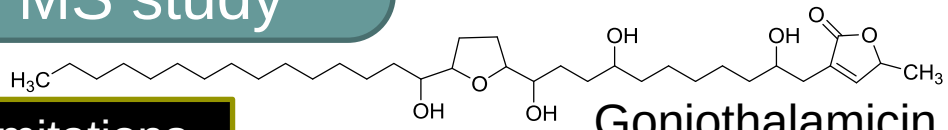
Bermejo, *Nat. Prod. Rep*. 2005

Structural diversity



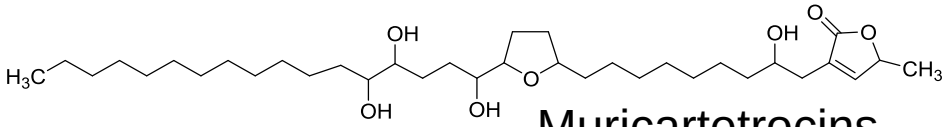
Muricin-A

MS study



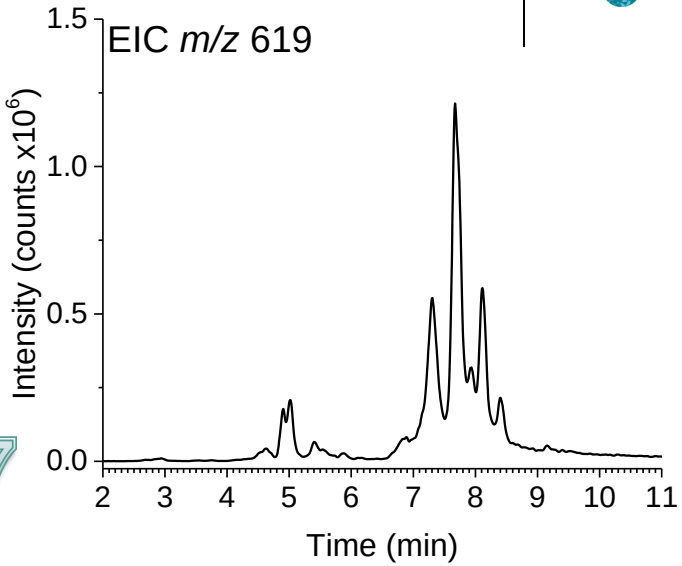
Goniothalamicin

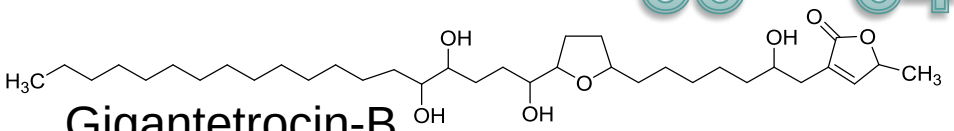
Limitations



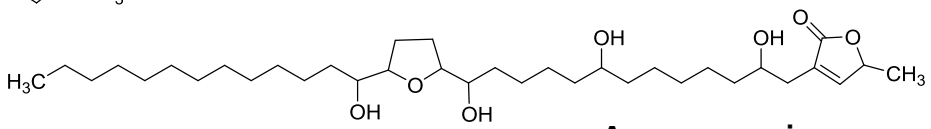
Muricartetrocins

$C_{35}H_{64}O_7$

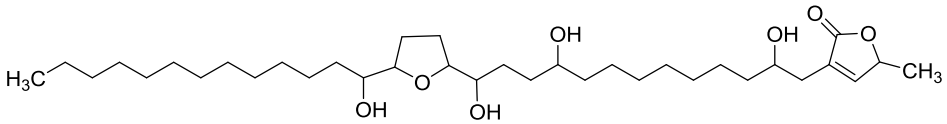




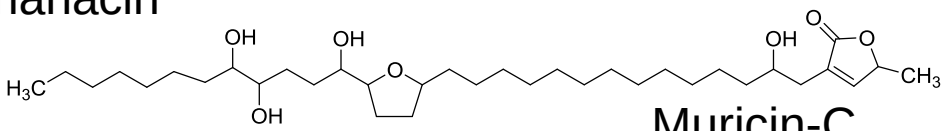
Gigantetrocin-B



Annonacin



Arianacin

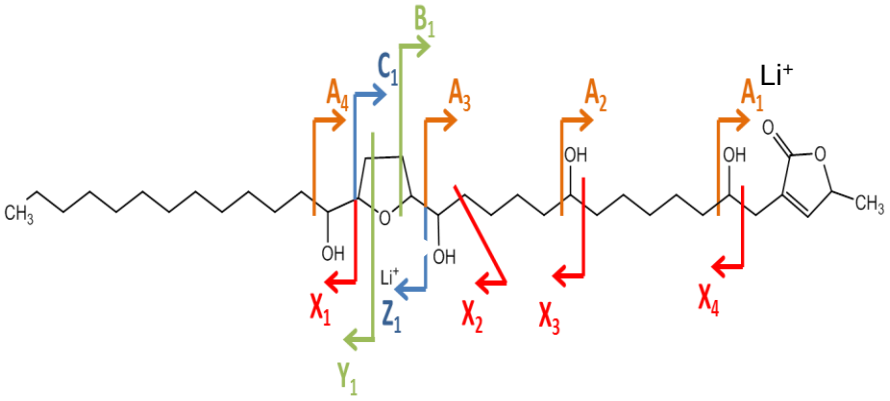


Muricin-C

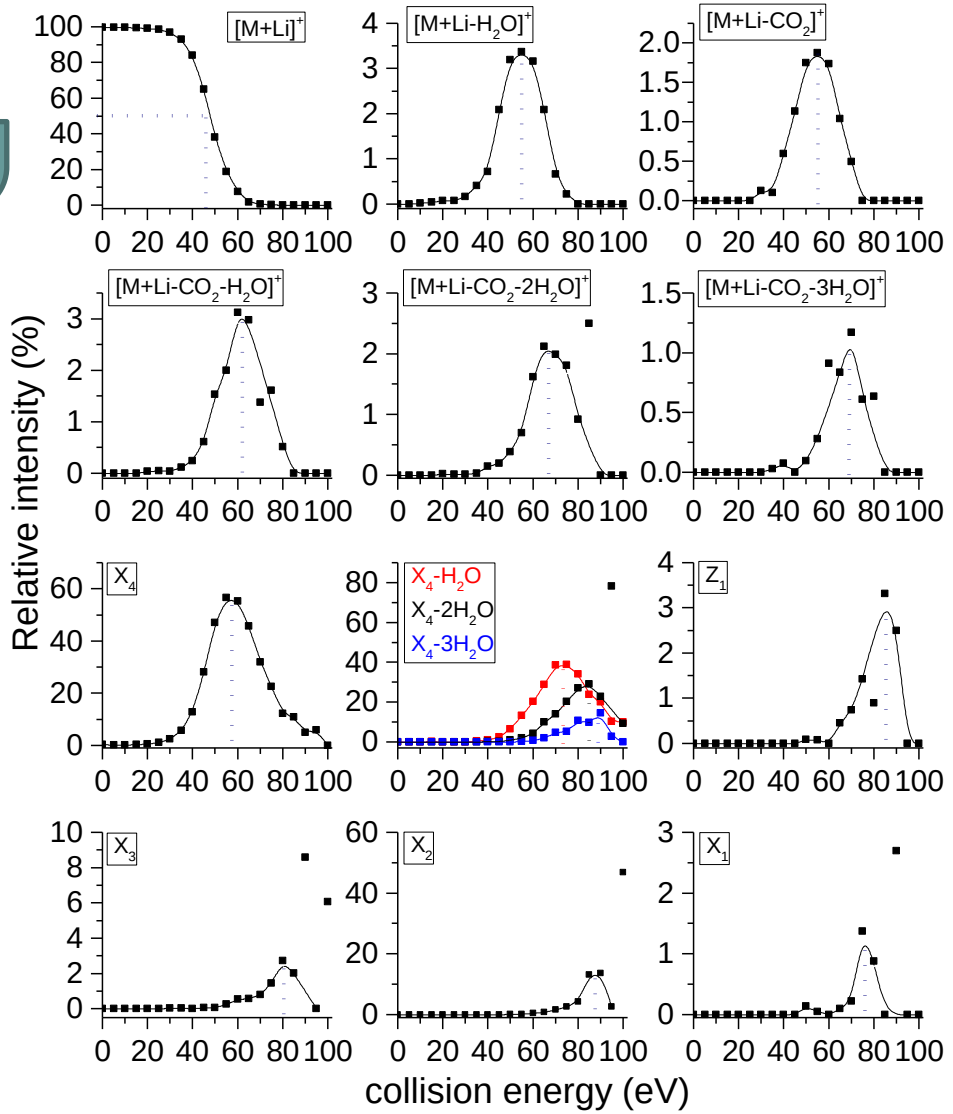
Structural diversity

MS/MS study

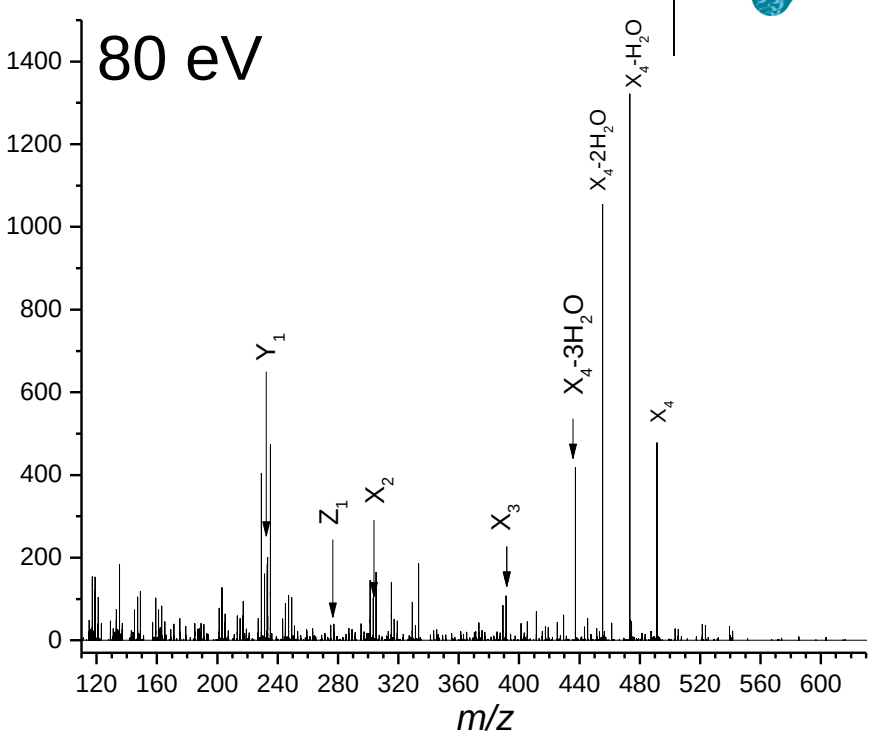
Optimization of the collision energy

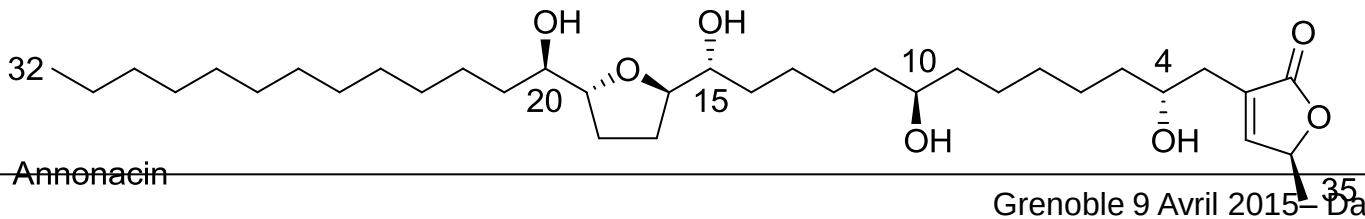


Two values are retained : 60 and 80 eV

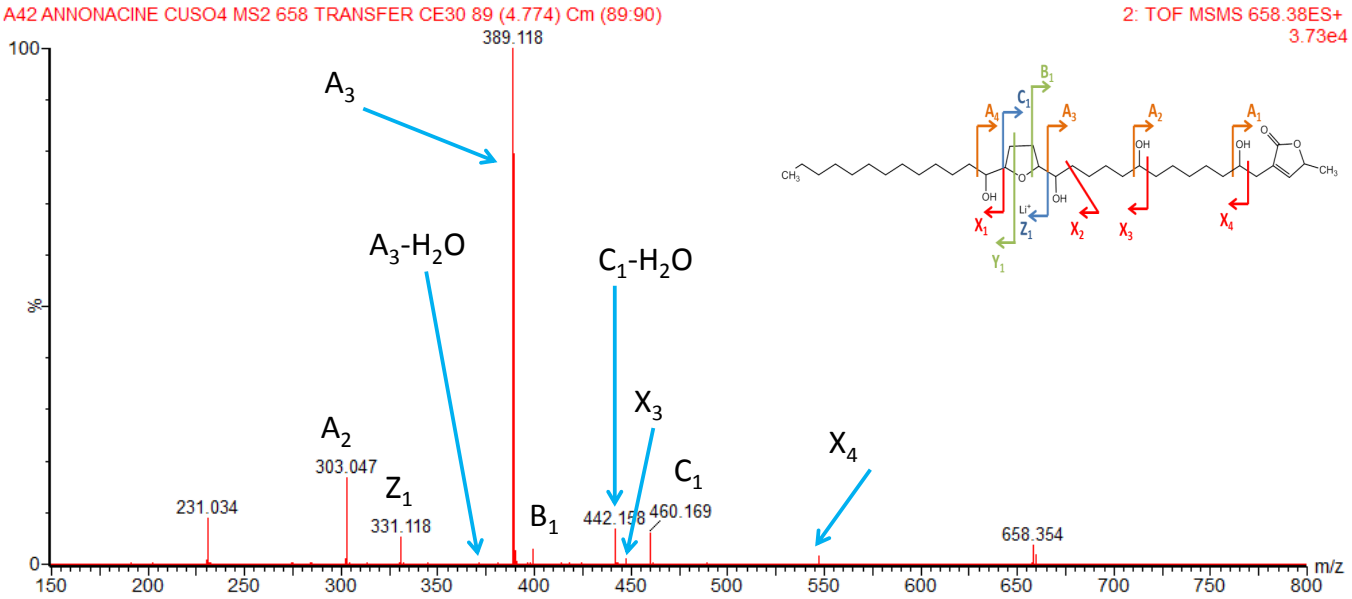


Structural diversity





Annonacin



Spectra acquired with an ion mobility mass spectrometer
MOBICS project (LCP, Orsay, G. Van der Rest, P. Maitre)