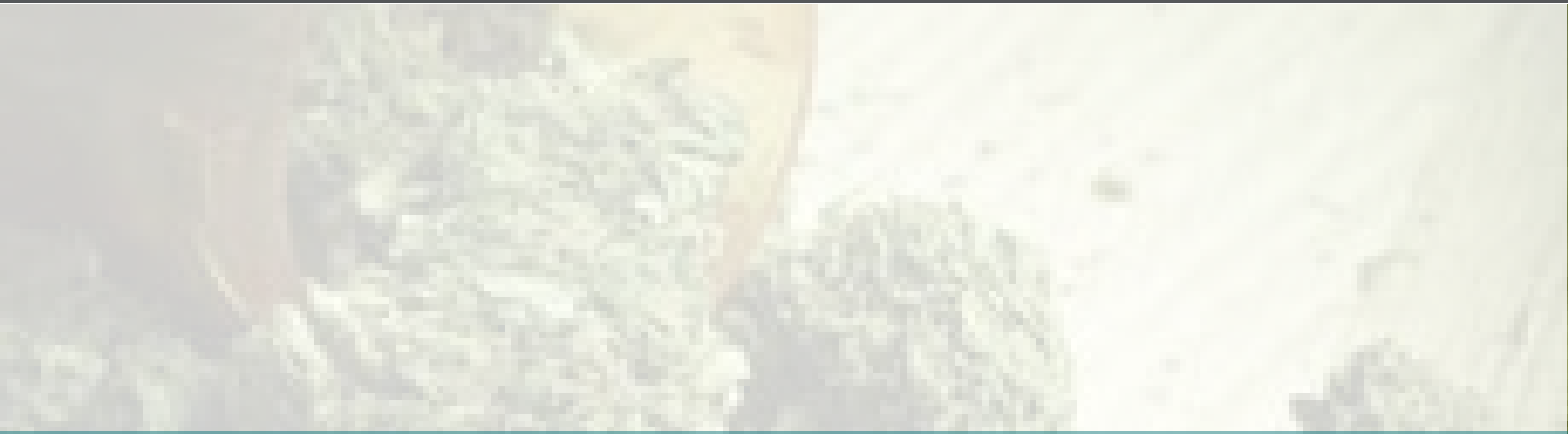
ThermoFisher SCIENTIFIC

Accelerate Unknown Detection in Emerging Drug Testing Using Thermo Scientific™ Compound Discoverer™ and mzCloud™

Ed Goucher
ClinTox Marketing Manager
Thermo Fisher Scientific

Emerging/Designer Drugs – “Fixing Bad”



Roughly 15-45 new ‘designer drugs’ hit the streets (globally) each year[1]

A trillion dollar business globally

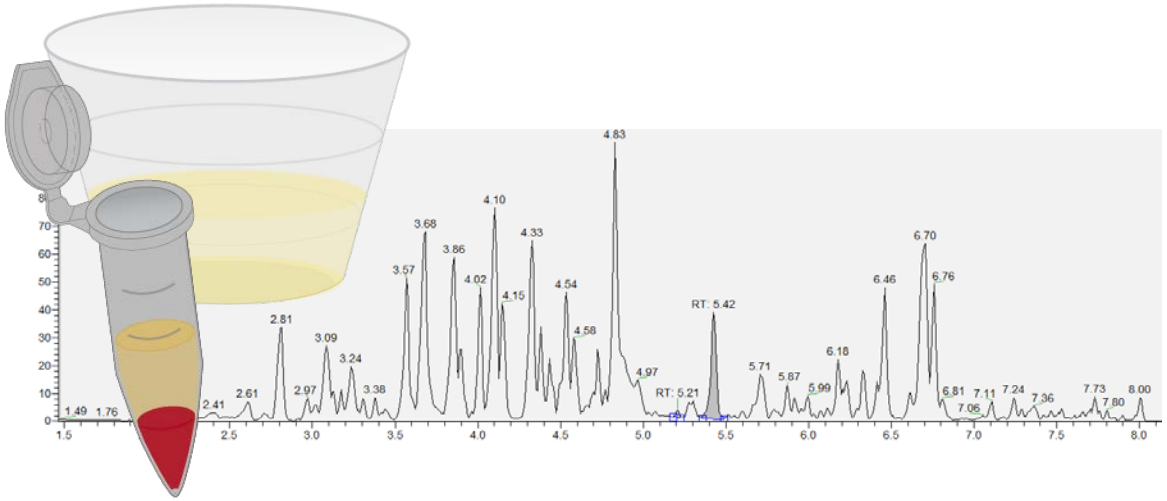
New designer drugs are NOT on current target analysis lists/ databases!

[1] US DEA data through 2015 and discovery/synthesis rates from Cayman and LGC (2016-2017)

How do you find something you don't know to look for?

- Sometimes we get lucky
  - "Bag of white powder"
  - Single (or small number) or analytes
  - Relatively pure

[1]



- Usually, we don't.
  - Complex sample matrices
  - Blood/Plasma – Urine – Hair
  - Classic 'needle in a haystack' problem

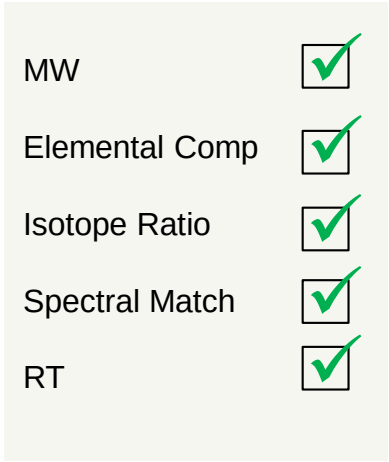
For forensic use only.

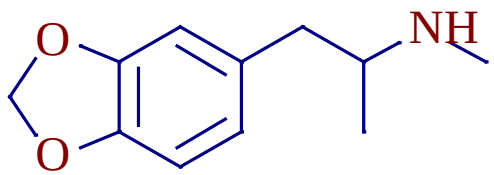
[1] The Control Center

What “Identifies” a Compound?

- Molecular weight (Accurate mass)
- Elemental Composition
- Isotope Ratio
- Fragmentation pattern (MS^2 or MS^n)
- Retention Time

Our Database of Drugs

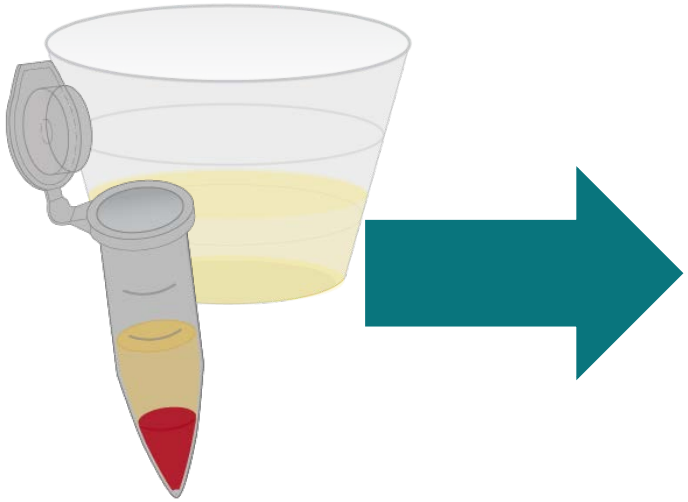




Typical Approaches – Playing ‘Catch-up’

- Rely heavily on targeted screening approaches using as complete a database as possible
  - Only can find what you already know.
- Rely on seizures of drugs during arrests which are transferred to reference libraries for identification
  - Only can find what others have identified
- Leaves most labs lagging far behind the real situation at street level.

How to Find What you Don't Know Is There



	Formula	Molecular Weight	RT [min]	Area (Max.)
10	C9 H18 F3 N2 O6 P S	370.05820	2.735	717604
11	C9 H9 N O2	163.06344	2.729	306788
12	C9 H12 N2 O2	180.09000	2.721	295418
13	C25 H47 F2 N O5 P2 S	573.26112	3.561	232104
14	C39 H50 F2 N6 O7 S	784.34272	4.117	213147
15	C17 H31 F2 N O2 P2	381.17982	4.113	203454
16	C40 H54 N2 O14	786.35790	4.093	2189681
17	C8 H9 N O	135.06855	1.406	10921051
18	C8 H14 N6 O2	226.11777	0.983	195996268
19	C8 H6	102.04721	2.401	1481012
20	C8 H9 N	119.07377	2.402	1449914
21	C28 H46 F N7 O14 P2	785.25624	6.668	1106722
22	C10 H22 N2 O3	218.16312	1.000	1434028
23	C18 H35 F2 N O6 P2	461.18987	5.044	875361
24	C2 H5 N O7	155.00691	0.870	112925

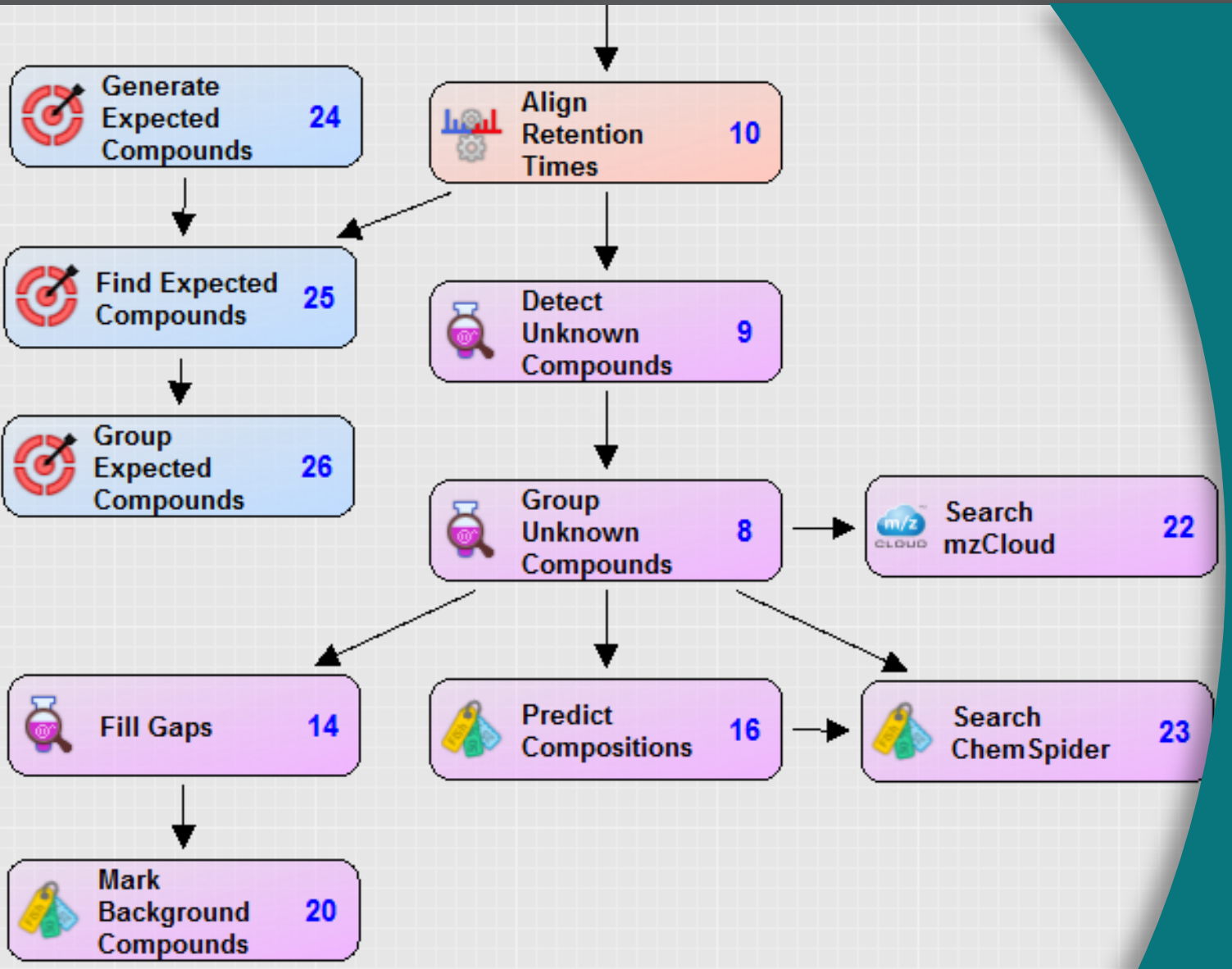
Need to take a complex biological sample and determine what (if any) compounds are suspect

Step 1 – Find out what is there

Create a list of unknown peaks (unique molecular weights)

Proceed to find and identify suspect compounds

Thermo Scientific Compound Discoverer: Destination Unknowns



Unknown Analysis

Identification

Metabolism

For forensic use only.

ThermoFisher SCIENTIFIC

Challenges: Terminology

- Too many terms:
  - A peak, a “feature”, a compound, a component….

Our (Thermo) Terminology

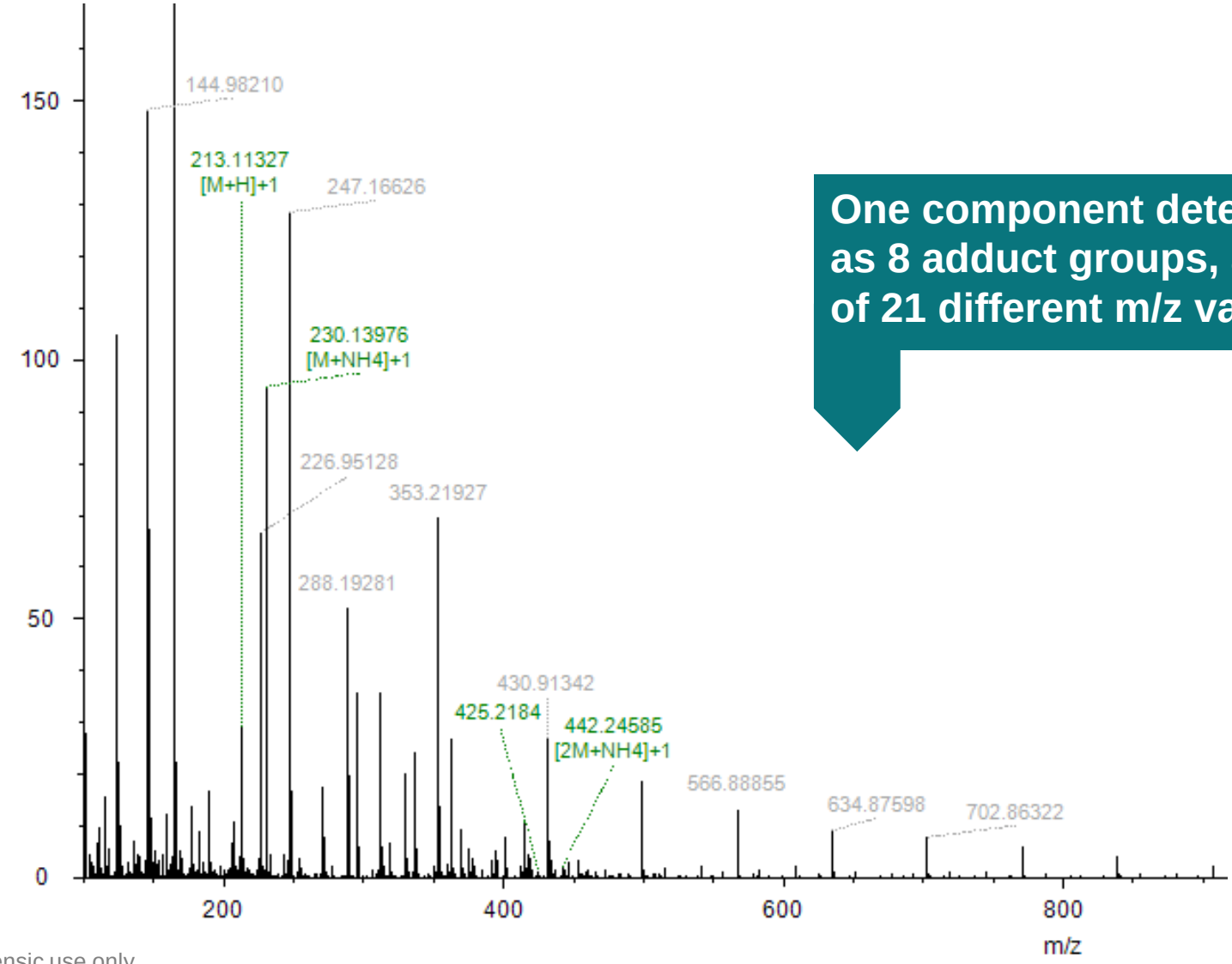
- A “peak”
  - Spectral peak – a single observed m/z in a spectra
  - Chromatographic peak – one or more m/z that create a time dependent elution
- A “feature”
  - A single m/z chromatographic peak
- A component
  - Collection of all observed features from isotopes and adducts
  - Assignable a single molecular weight
  - Analogous to a “compound”

One Compound isn’t just One Feature

Full MS at 70,000 resolution

Human plasma sample

Untargeted peak detection with Compound Discoverer



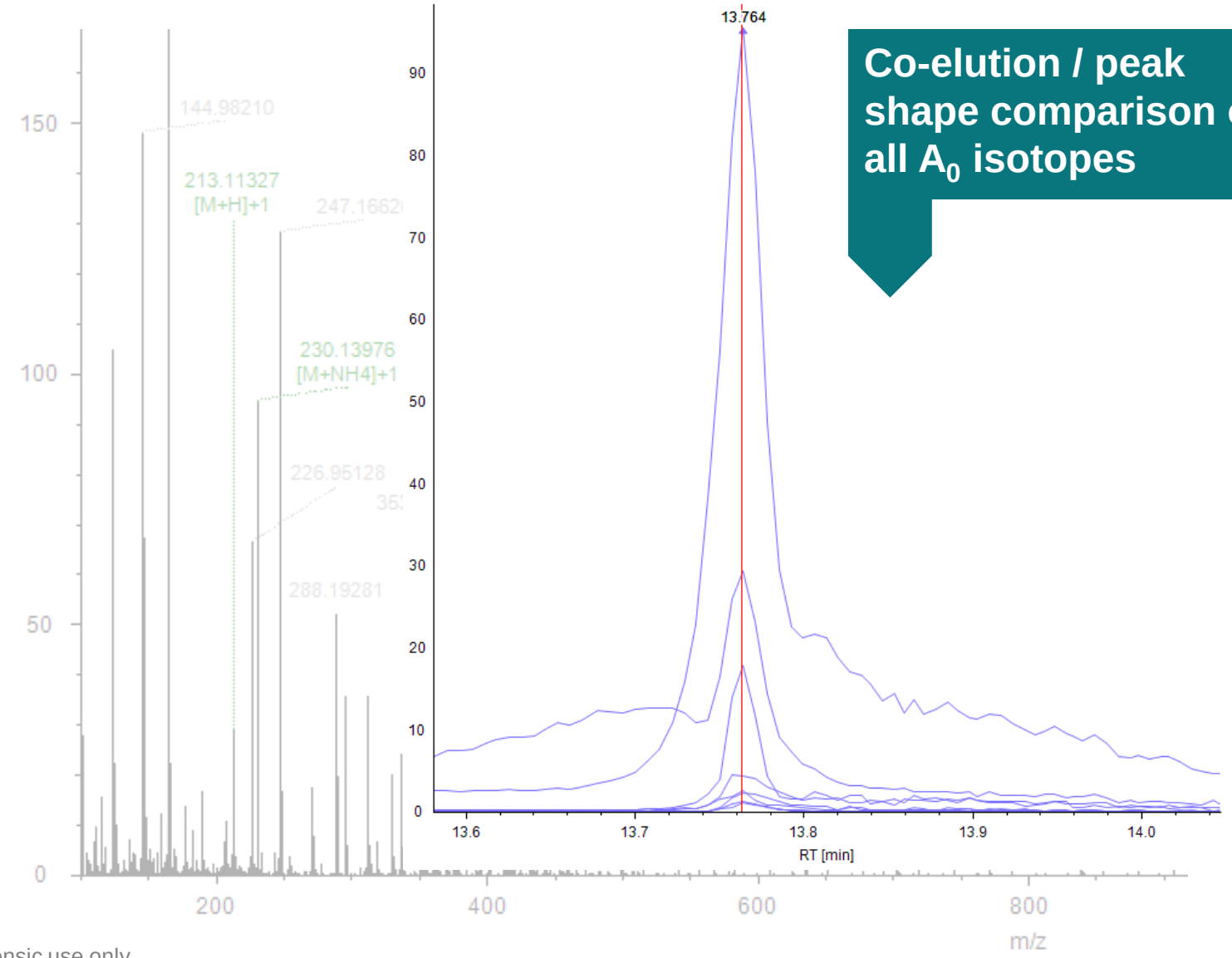
One component detected as 8 adduct groups, a total of 21 different m/z values.

One Compound isn’t just One Feature

Full MS at 70,000 resolution

Human plasma sample

Untargeted peak detection with Compound Discoverer



Co-elution / peak shape comparison of all A_0 isotopes

For forensic use only.

ThermoFisher SCIENTIFIC

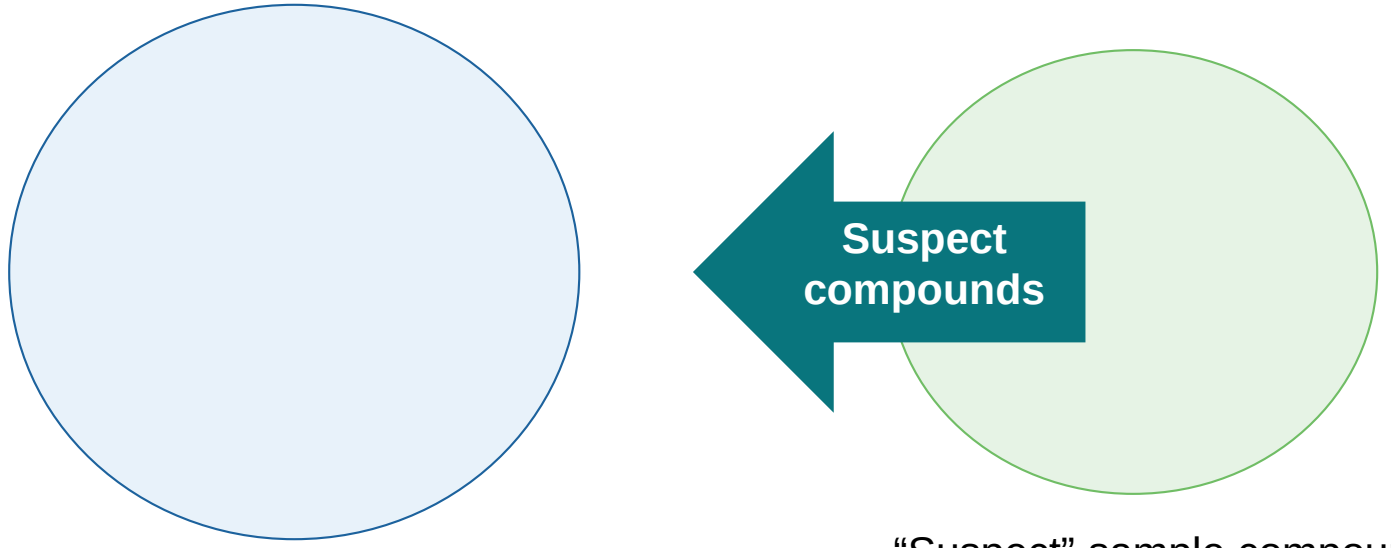
What's in a Peak…

- An 'average' sample will contain:
  - Between hundreds to tens of thousands of features
  - Dozens to several thousand unique compounds
- An 'average' compound is detected as:
  - Between 2 to 20+ individual isotopes
  - Multiple different adducts
- Source fragmentation adds to this complexity.

Differential Analysis – One Potential Tool

- With our list of components in our samples we can ask…

- Is there something in this sample (group) that isn't in my control?
- Is there something that I haven't seen before?
- Is there anything suspect?



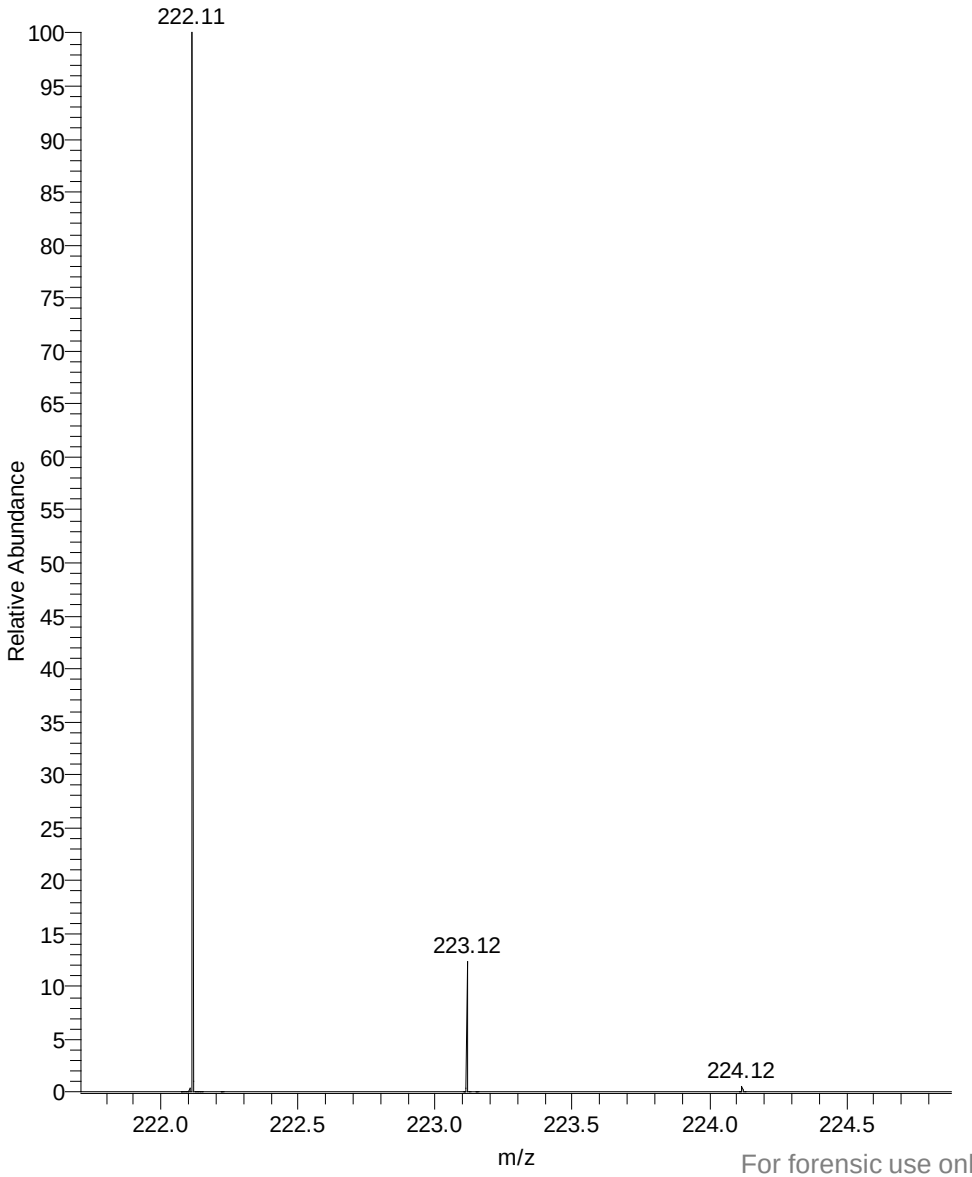
Historical or "clean" reference samples compounds

"Suspect" sample compounds

What “Identifies” a Compound?

- Molecular weight (Accurate mass)
- Elemental Composition
- Isotope Ratio
- Fragmentation pattern (MS^2 or MS^n)
- Retention Time

Molecular Weight – What Does it Tell us?



Single *m/z* value

Assume [M+H]

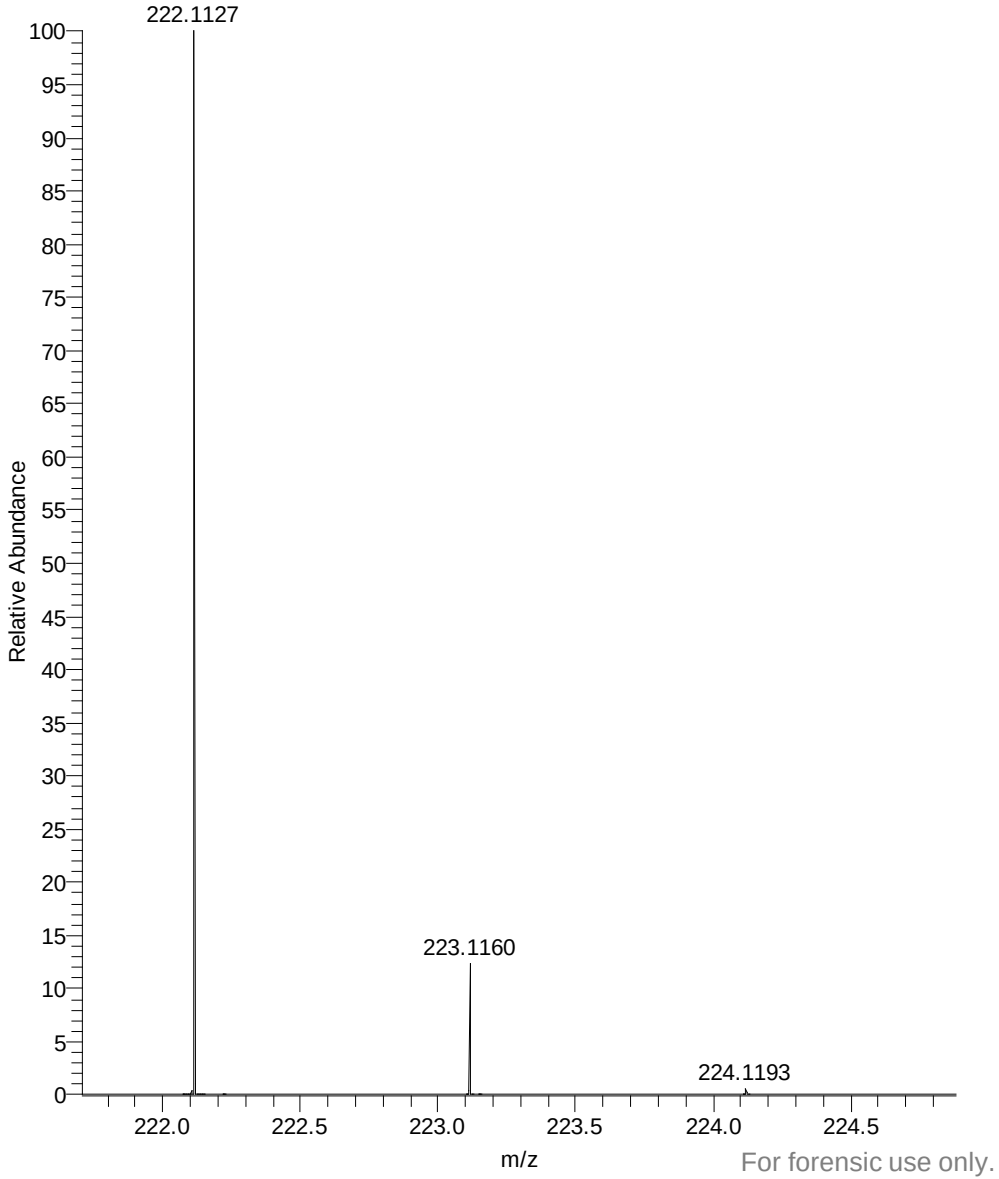
Calculated MW – 221.11

Search ChemSpider +/- 0.25

115,383 possibilities

…we have no clue what this is…

Molecular Weight – How about Accurate Mass!



Single *m/z* value

Assume [M+H]

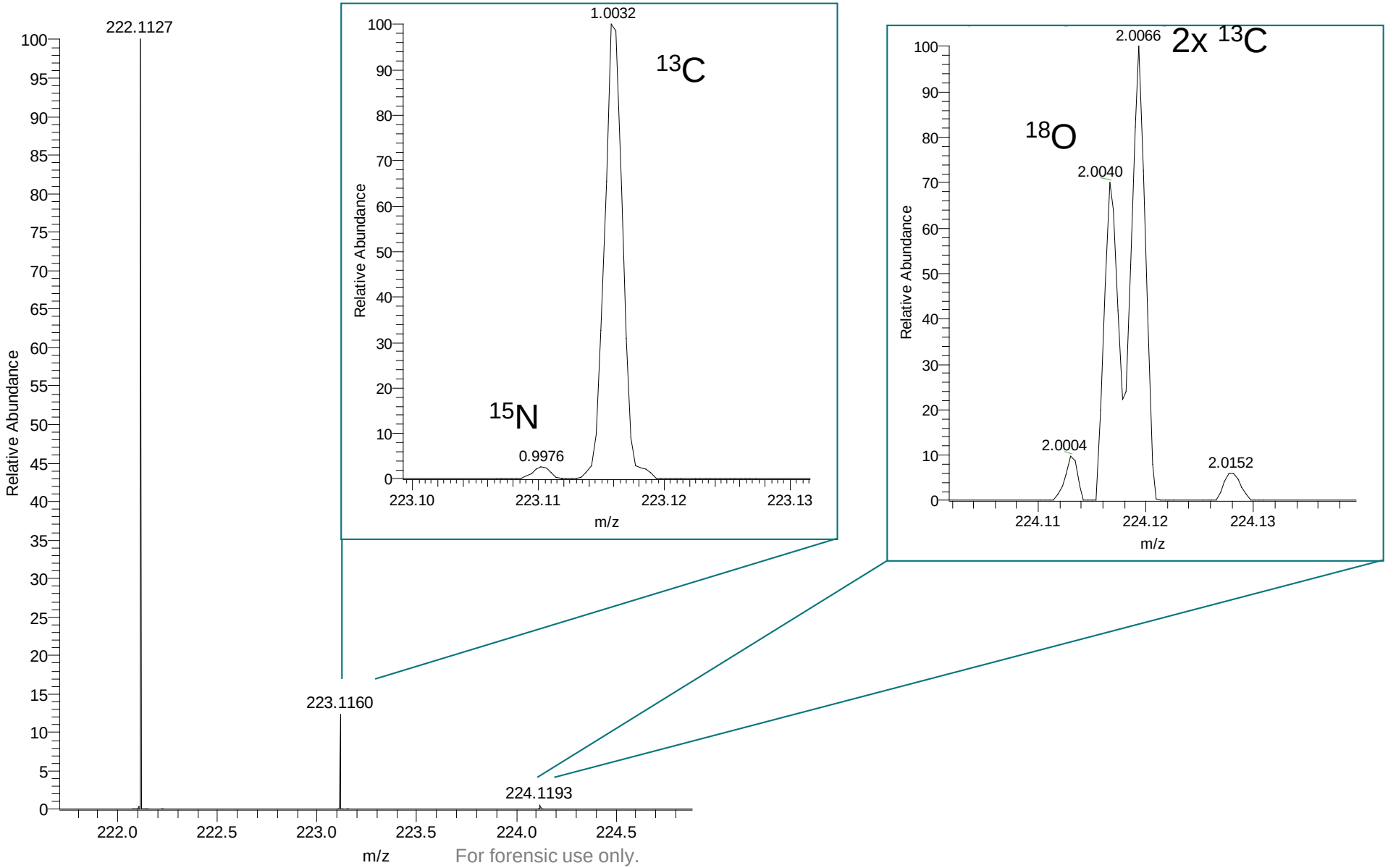
Calculated MW – 221.1054

Search ChemSpider +/- 5ppm

6,435 possibilities

…we still have no clue what this is…

Going Even Further – Elemental Composition with Very High Resolution



ThermoFisher SCIENTIFIC

Elemental Composition – Surely that will help…

Elements in Use

Isotope	Mass	Min	Max
14 N	14.0031	0	10
16 O	15.9949	0	15
12 C	12.0000	0	30
1 H	1.0078	0	60
32 S	31.9721	0	10
35 Cl	34.9689	0	4
31 P	30.9738	0	10
19 F	18.9984	0	6

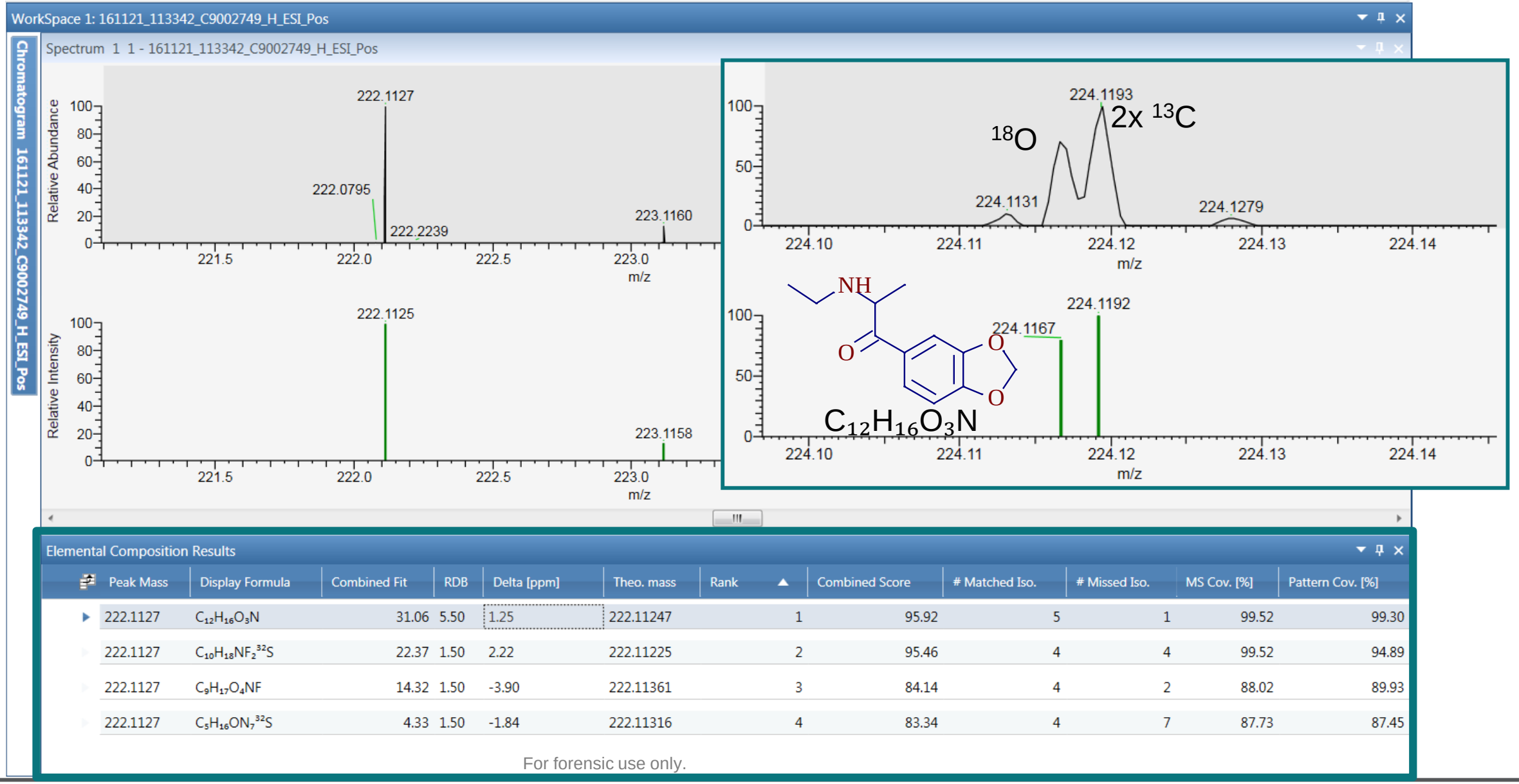
These are real unknown compounds

Must use an ‘open’ elemental composition set

Can’t ‘cheat’ and artificially limit elements and ranges

Must include fluorine (common synthetic trick to ‘hide’ designer drugs)

Elemental Composition – Surely that will help…

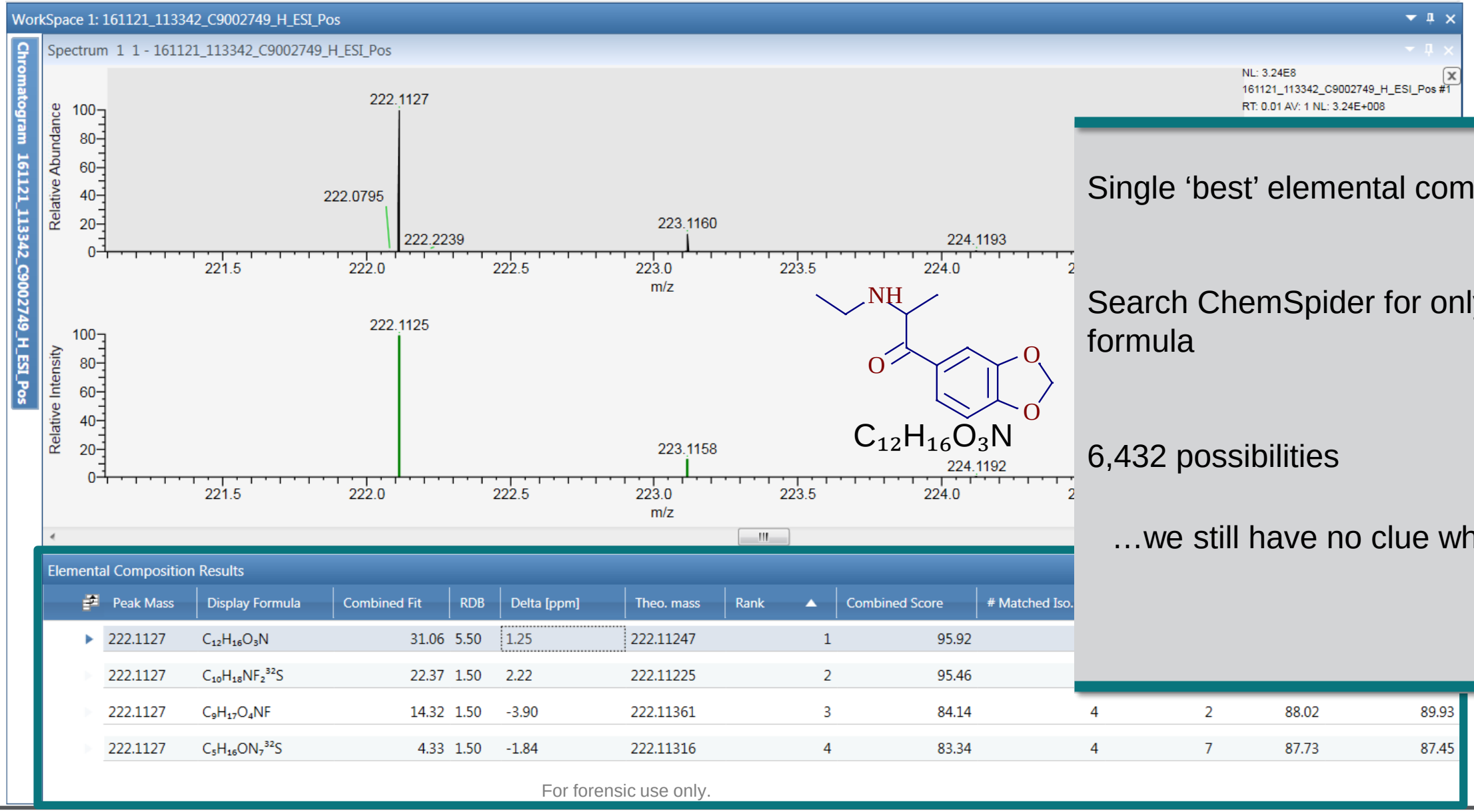


Elemental Composition Results

Peak Mass	Display Formula	Combined Fit	RDB	Delta [ppm]	Theo. mass	Rank	Combined Score	# Matched Iso.	# Missed Iso.	MS Cov. [%]	Pattern Cov. [%]
222.1127	$C_{12}H_{16}O_3N$	31.06	5.50	1.25	222.11247	1	95.92	5	1	99.52	99.30
222.1127	$C_{10}H_{18}NF_2{}^{32}S$	22.37	1.50	2.22	222.11225	2	95.46	4	4	99.52	94.89
222.1127	$C_9H_{17}O_4NF$	14.32	1.50	-3.90	222.11361	3	84.14	4	2	88.02	89.93
222.1127	$C_5H_{16}ON_7{}^{32}S$	4.33	1.50	-1.84	222.11316	4	83.34	4	7	87.73	87.45

For forensic use only.

Elemental Composition – Surely that will help…



Single ‘best’ elemental composition

Search ChemSpider for only that formula

6,432 possibilities

…we still have no clue what this is…

What “Identifies” a Compound?

- Molecular weight (Accurate mass)
- Elemental Composition
- Isotope Ratio
- Fragmentation pattern (MS^2 or MS^n)
- Retention Time

mzCloud - Library Solutions from Thermo Fisher Scientific

Fragmentation Libraries

Integrated solutions for small molecule research applications in LC- HRAM MS

mzCloud™

- High Quality – Extensive Curation
- HRAM MS/MS and MSn data
- Integrated search with Compound Discoverer

mzCloud (www.mzCloud.org)

Advanced Mass Spectral Database

Server location : US

search for compounds... Search

Home About Features App Database Partners Contact

mzCloud is a state of the art mass spectral database that assists analysts in identifying compounds in areas such as life sciences, metabolomics, pharmaceutical research, toxicology, forensic investigations, environmental analysis, food control and various industrial applications. mzCloud™ features a freely searchable collection of high resolution/accurate mass spectra using a new third generation spectra correlation algorithm.

Online access to the database is free of charge and no registration is required.

read more...

Enter Database

Google play New mzCloud App!

Search for Compounds by Name or ID

Search

6,845 (+52)	10,711 (+77)	2,398,307 (+29,645)	707,074 (+1,194)	view more statistics
compounds	trees	spectra	QM models	

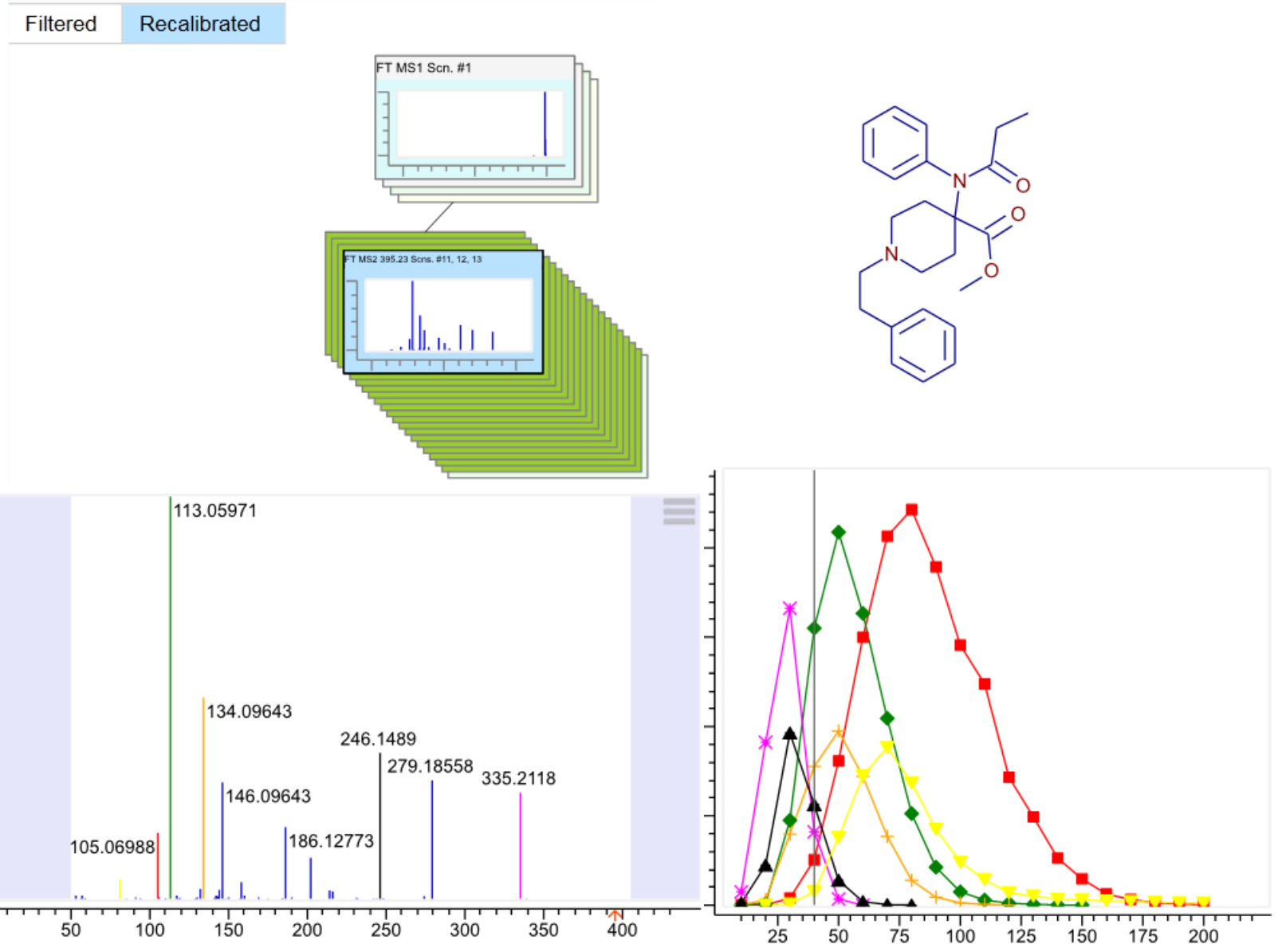
For forensic use only.

ThermoFisher SCIENTIFIC

More Information

Extensive MS/MS and MS^n

- **Dozens to thousands of spectra per compound**
- **10-20 Different HCD Energies**
- **Dynamically optimized trap CID energy**
- **No limits on how you run your instrument**



Diversity is Critical – Don’t Blind your Unknown Analysis

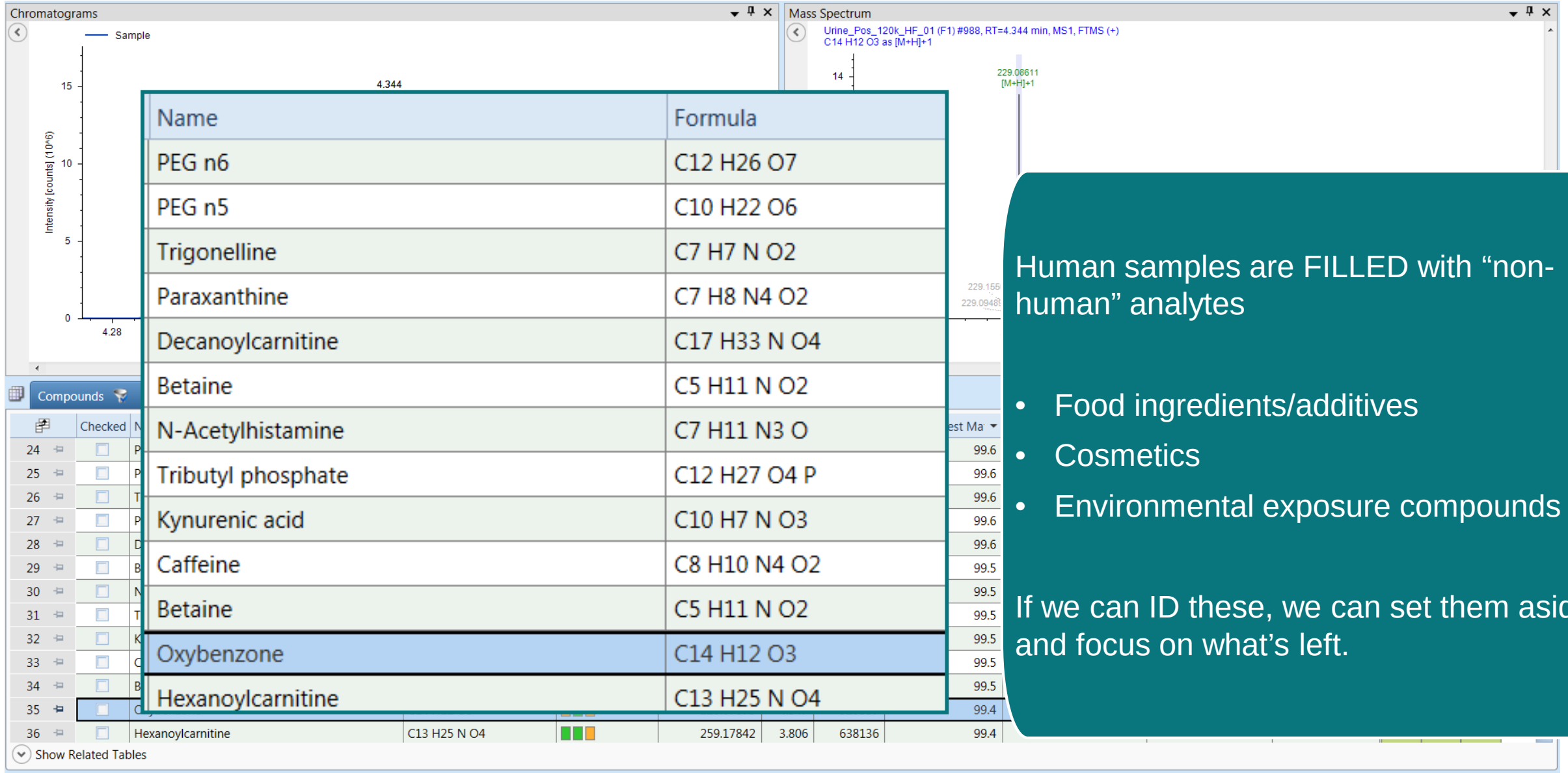
mzCloud contains a wide structure diversity

Compounds grouped into sixteen broad categories

“If it ionizes, we put it into mzCloud”

- Therapeutics/Prescription Drugs
- Drugs of Abuse/Illegal Drugs
- Sports Doping Drugs
- Steroids/Vitamins/Hormones
- Endogenous Metabolites
- Natural Products/Medicines
- Natural Toxins
- Counterfeit Drug (Therapeutic)
- Extractables/Leachables
- Pesticides/Herbicides
- Excipients/Additives/Colorants
- Illegal Additives
- Personal Care Products/Cosmetics
- Textile Chemicals/Auxiliary/Dyes
- Industrial Chemicals
- Perflourinated Hydrocarbons

“The Lineup” - Finding a Suspect by Process of Elimination



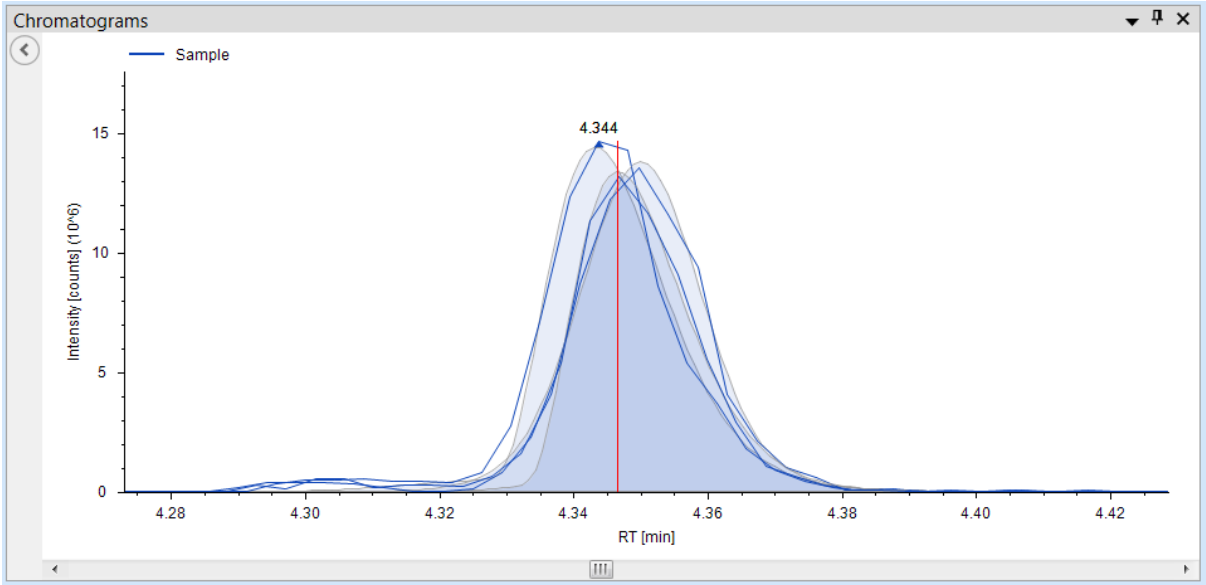
Name	Formula
PEG n6	C12 H26 O7
PEG n5	C10 H22 O6
Trigonelline	C7 H7 N O2
Paraxanthine	C7 H8 N4 O2
Decanoylcarnitine	C17 H33 N O4
Betaine	C5 H11 N O2
N-Acetylhistamine	C7 H11 N3 O
Tributyl phosphate	C12 H27 O4 P
Kynurenic acid	C10 H7 N O3
Caffeine	C8 H10 N4 O2
Betaine	C5 H11 N O2
Oxybenzone	C14 H12 O3
Hexanoylcarnitine	C13 H25 N O4

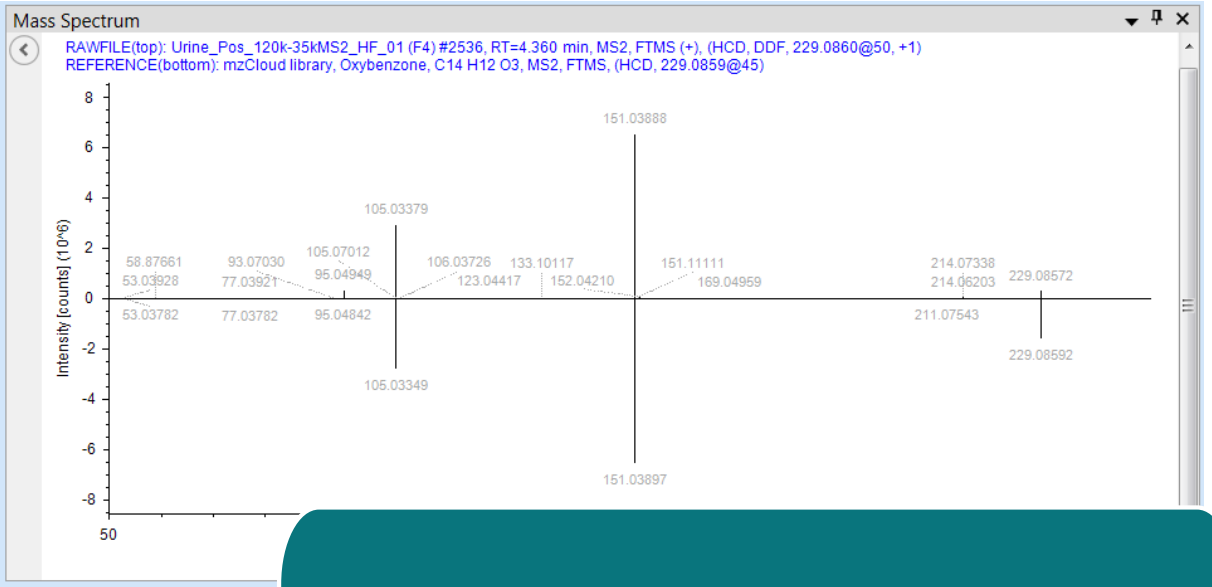
Human samples are FILLED with “non-human” analytes

- Food ingredients/additives
- Cosmetics
- Environmental exposure compounds

If we can ID these, we can set them aside and focus on what’s left.

“The Lineup” - Finding a Suspect by Process of Elimination



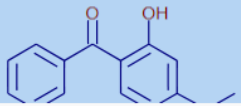


Compounds | Compounds per File | Features | mzCloud Results | ChemSpider Results | Input Files

	Checked	Name	Formula	Annotation Sc	Molecular Weight	RT [min]	Area (Max.)	mzCloud Bes
33		Decanoylcarnitine	C17 H33 N O4		315.24105	5.957	9157928	
34		Kynurenic acid	C10 H7 N O3		189.04268	2.926	8911403	
35		Caffeine	C8 H10 N4 O2		194.08052	3.117	20415729	
36		Betaine	C5 H11 N O2		117.07925	9.221	4173255	
37		Oxybenzone	C14 H12 O3		228.07875	4.347	320937	
38		Hexanoylcarnitine	C13 H25 N O4		259.17842	3.806	638136	

Hide Related Tables

Structure Proposals | Compounds per File | Predicted Compositions | mzCloud Results | ChemSpider Results

	Checked	Compoun	Structure	Name	Type	Formula	Best S	Molecular Weight	ΔMass [Da]	ΔMass [ppm]	Scan #	Match	Best Match	mzCloud ID	KEGG ID
1				Oxybenzone	Identity	C14 H12 O	97.1	228.07864	-0.00011	-0.47	2536	99.4	99.4	863	C14285

Show Related Tables

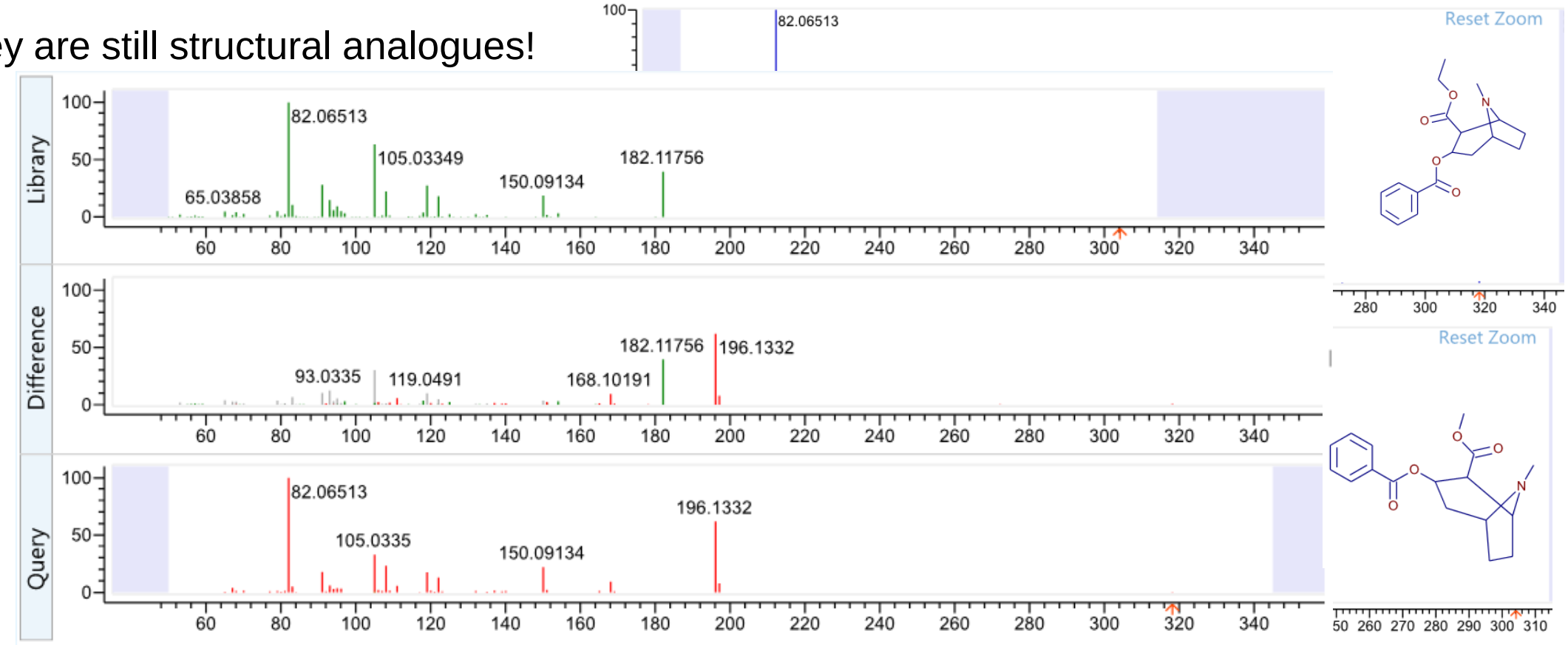
Oxybenzone, a sunscreen ingredient

Obviously we don’t need to worry about this suspect.

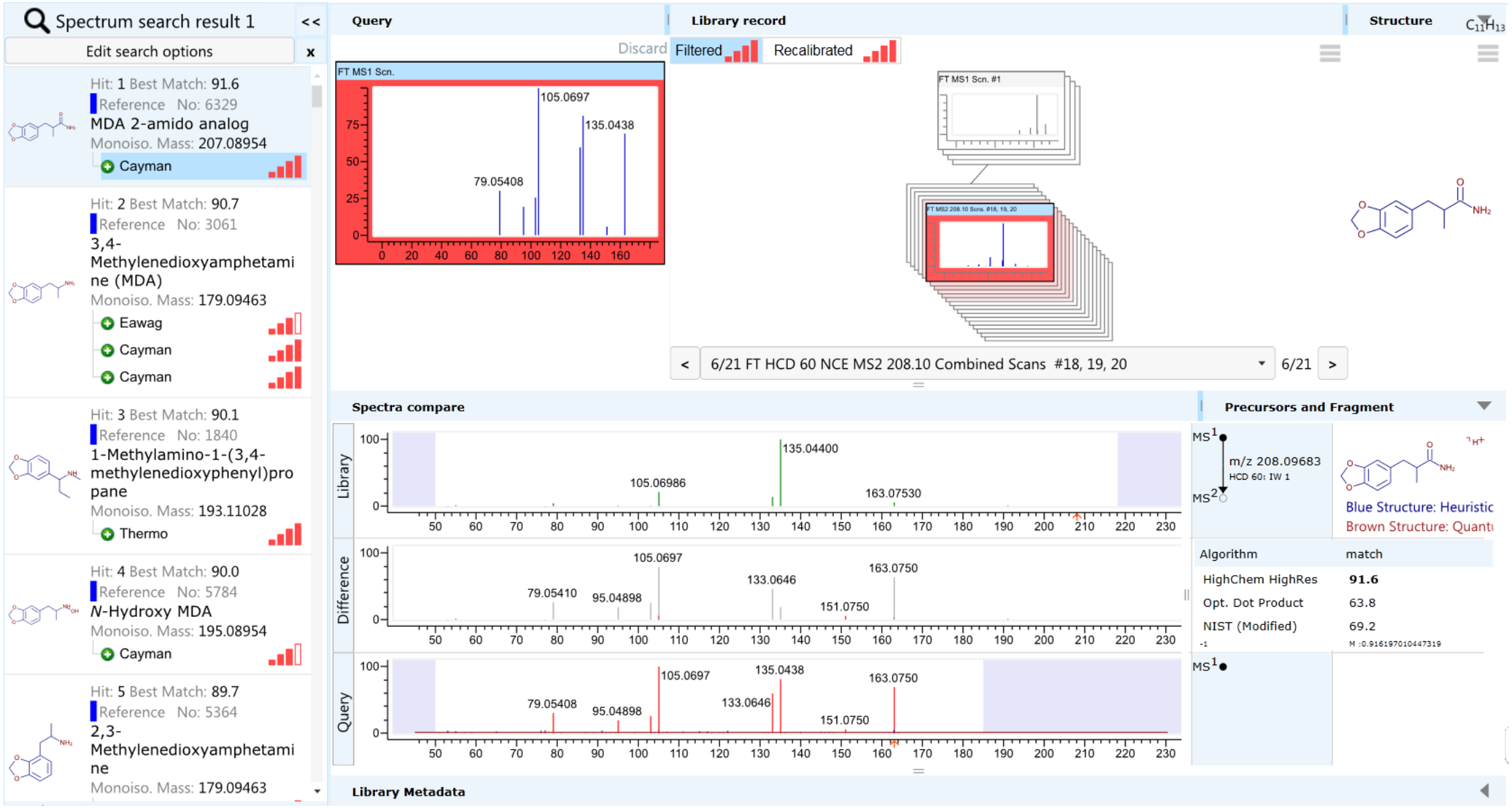
Using Libraries more Intelligently – Similarity Searching

- Designer drugs are often synthetic analogues of current compounds.
- Small change in structure = change in MW, formula, RT, and fragments = Totally missed!!

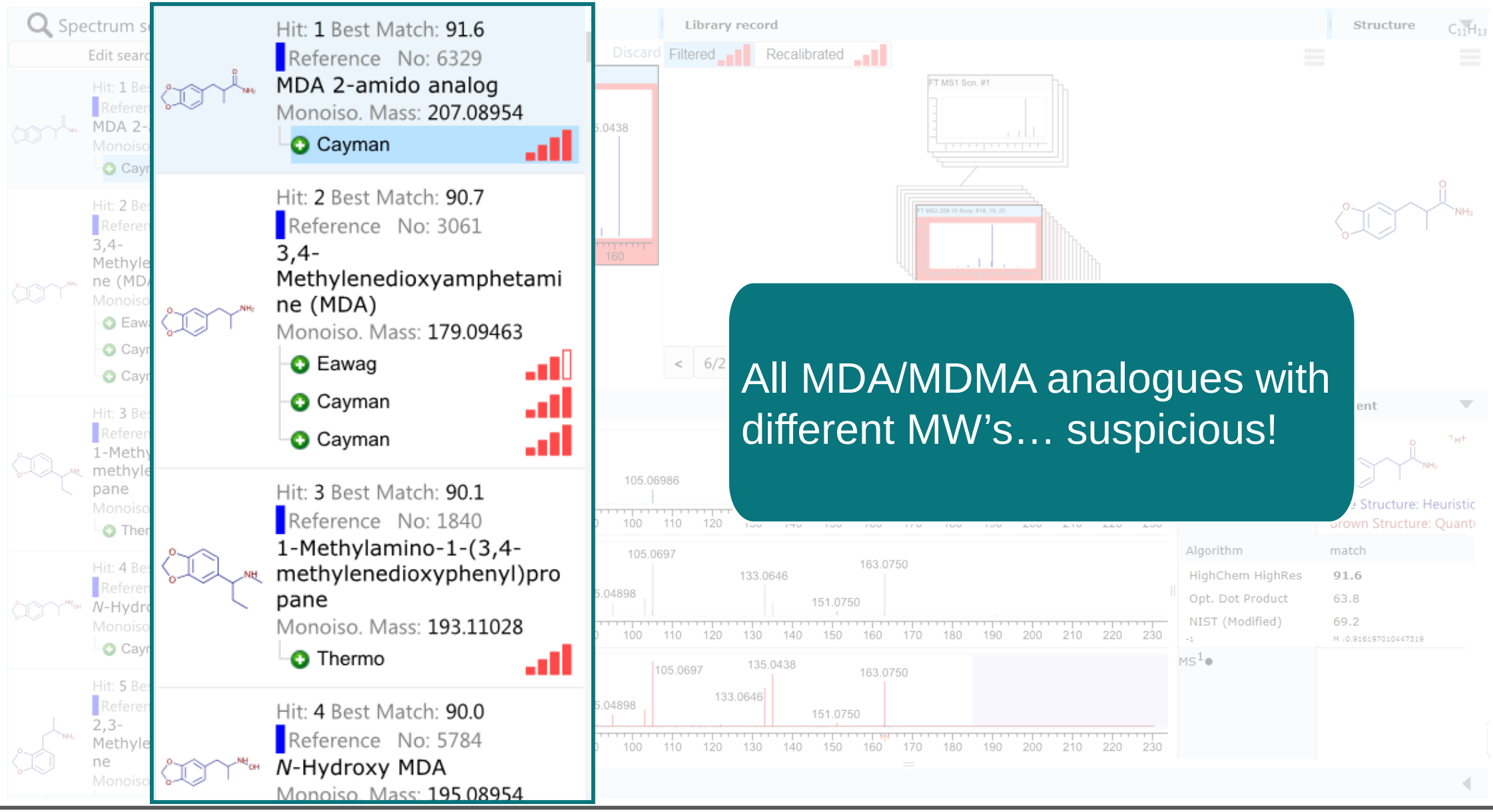
…but, they are still structural analogues!



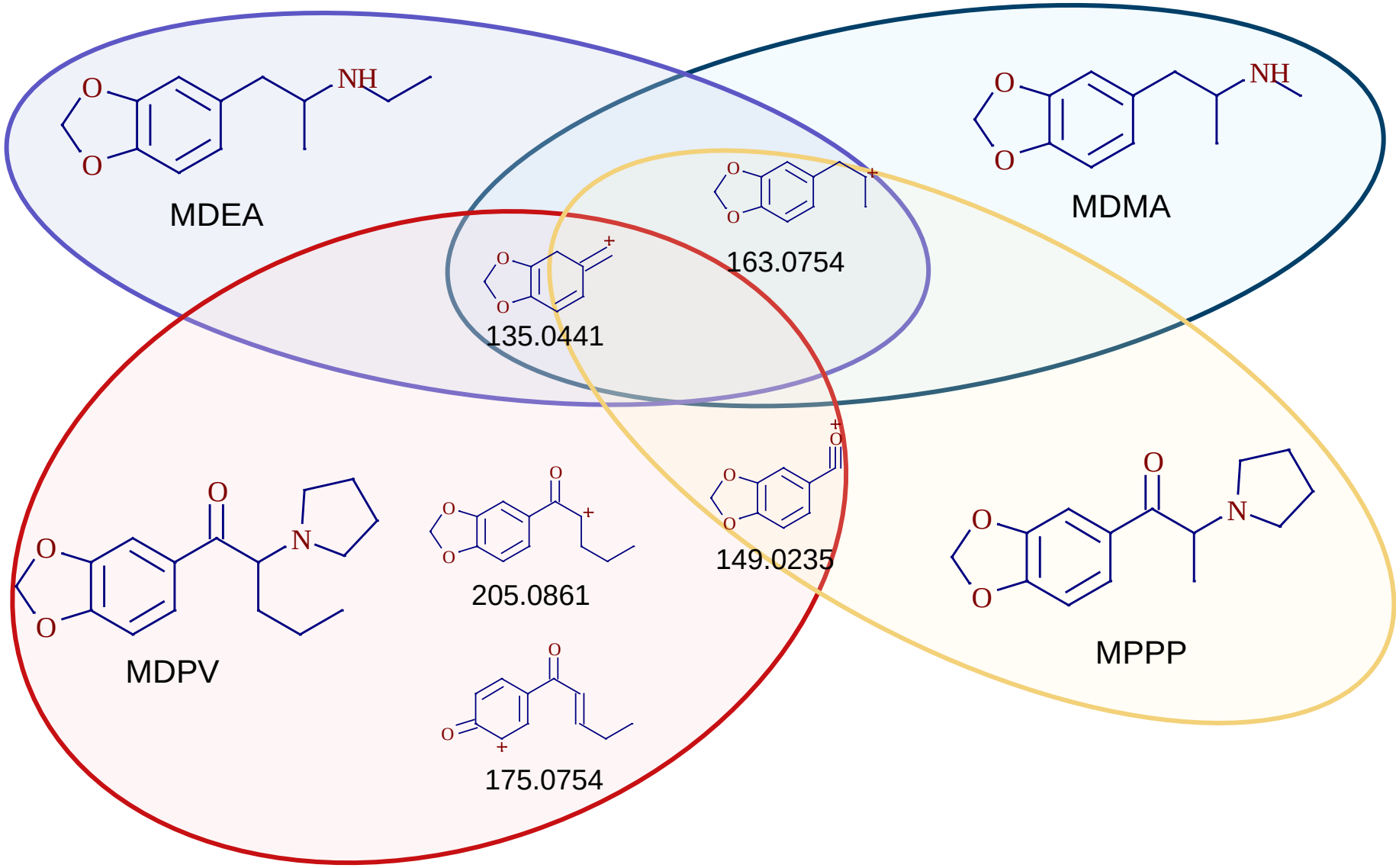
Libraries are for Learning – Not Just ID



Libraries are for Learning – Not Just ID



Getting Creative – Using What we Know in New Ways

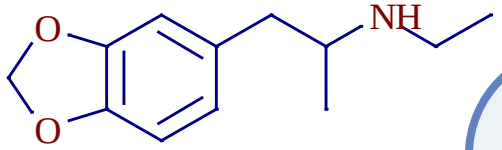


Many new designer drugs are synthetic analogues of existing drugs designed to escape targeted detection

Except for “entirely new” classes of drugs, we know at least a few analogues already.

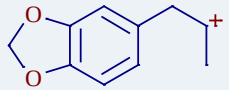
We can USE that knowledge!

Compound Class Fragment Ion Searching

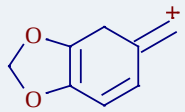


MDEA

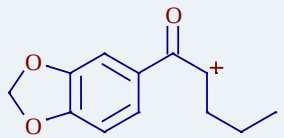
Class-Based Fragment Set



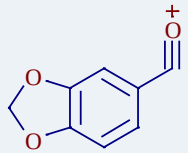
163.0754



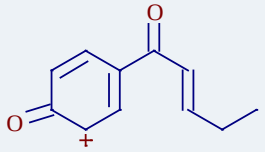
135.0441



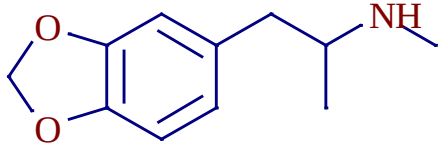
205.0861



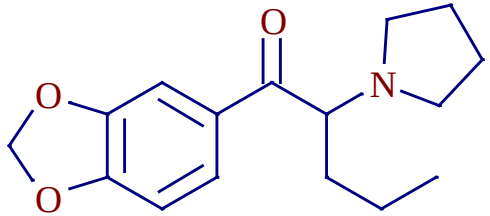
149.0235



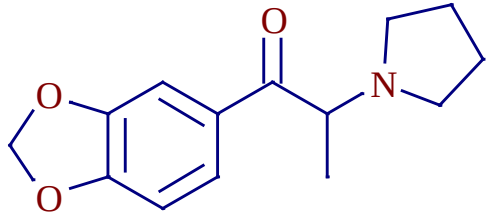
175.0754



MDMA



MDPV



MPPP

Build a representative ‘set’ of fragment ions from known compounds

Recruit high frequency fragment ions to the fragment set

Use this set of associated fragments to find structurally related analogues

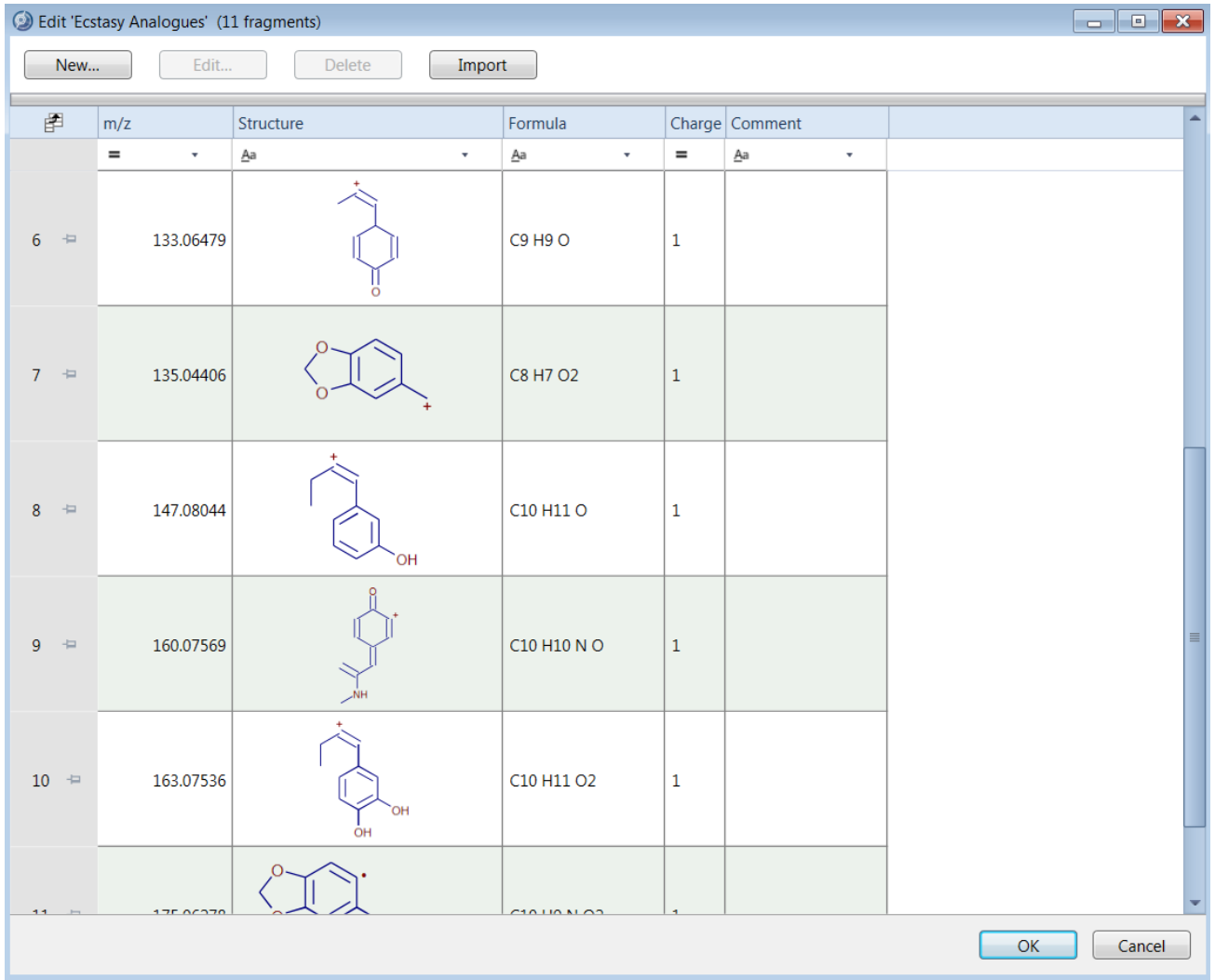
Using What We Know

Example scenario:

Eight total samples – 4 each for urine and plasma

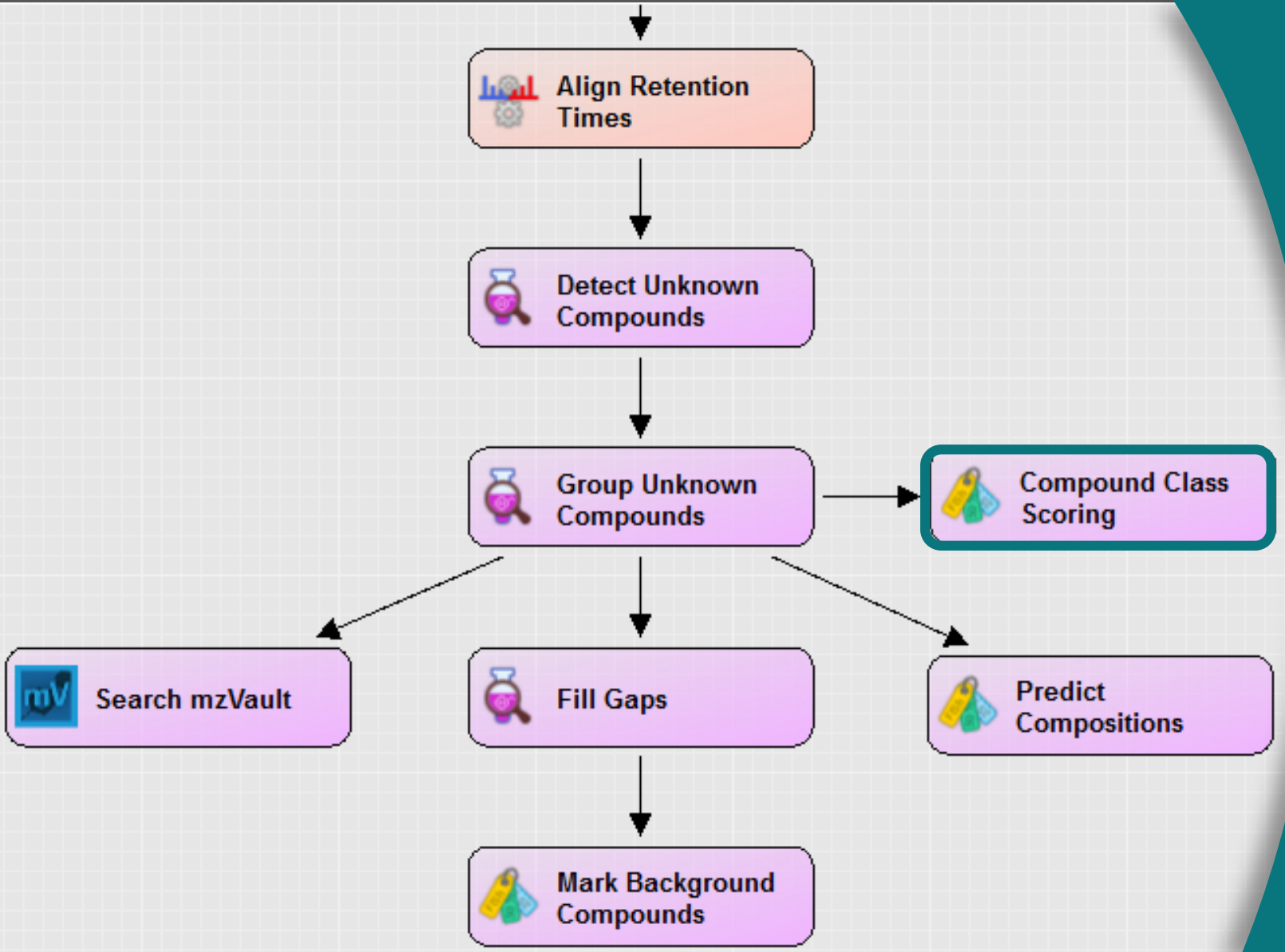
2 “suspect” samples and 2 “clean” samples for each matrix.

Search for ‘ecstasy-like’ compounds using fragments from known analogues (11 total fragments in the set).



ThermoFisher SCIENTIFIC

Our Workflow – Not “unknown” for long…

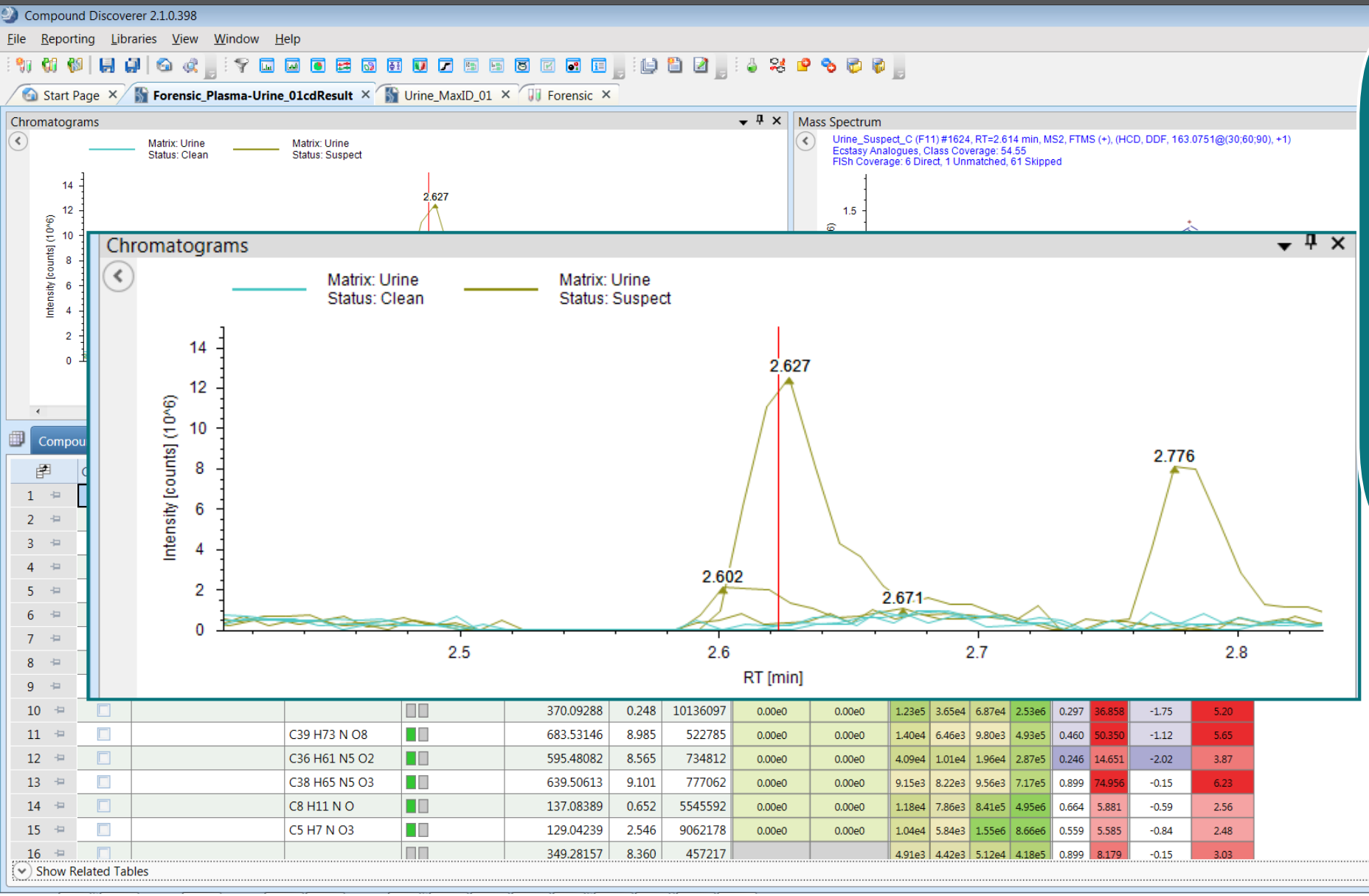


Combined approach

Unknown detection

Multiple tools

Use All the Tools Together – Seeing Everything at Once



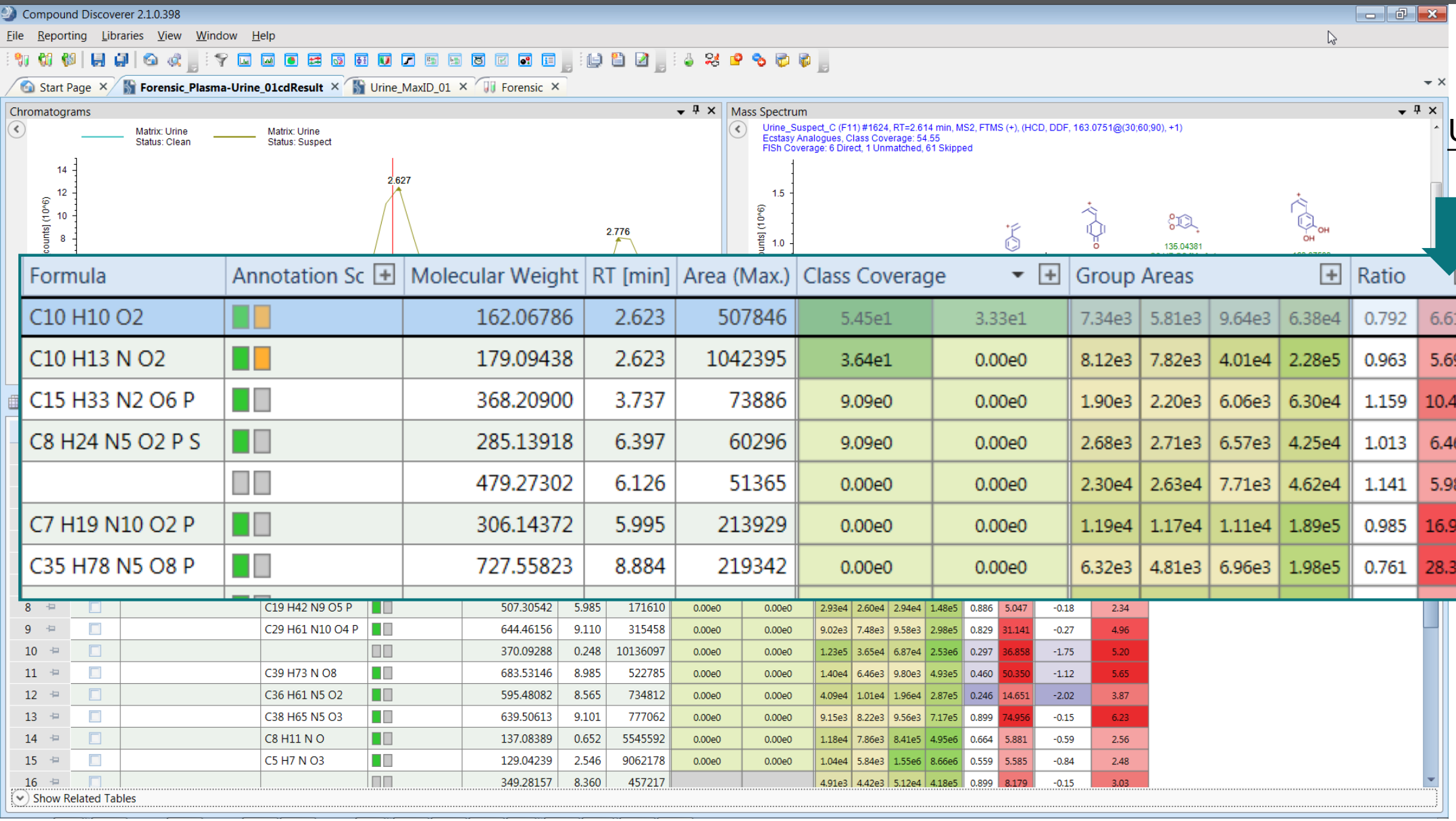
10			370.09288	0.248	10136097	0.00e0	0.00e0	1.23e5	3.65e4	6.87e4	2.53e6	0.297	36.858	-1.75	5.20
11	C39 H73 N O8		683.53146	8.985	522785	0.00e0	0.00e0	1.40e4	6.46e3	9.80e3	4.93e5	0.460	50.350	-1.12	5.65
12	C36 H61 N5 O2		595.48082	8.565	734812	0.00e0	0.00e0	4.09e4	1.01e4	1.96e4	2.87e5	0.246	14.651	-2.02	3.87
13	C38 H65 N5 O3		639.50613	9.101	777062	0.00e0	0.00e0	9.15e3	8.22e3	9.56e3	7.17e5	0.899	74.956	-0.15	6.23
14	C8 H11 N O		137.08389	0.652	5545592	0.00e0	0.00e0	1.18e4	7.86e3	8.41e5	4.95e6	0.664	5.881	-0.59	2.56
15	C5 H7 N O3		129.04239	2.546	9062178	0.00e0	0.00e0	1.04e4	5.84e3	1.55e6	8.66e6	0.559	5.585	-0.84	2.48
16			349.28157	8.360	457217			4.91e3	4.42e3	5.12e4	4.18e5	0.899	8.179	-0.15	3.03

Differential Analysis

Look first at ‘new’ peaks in our suspect sample

ThermoFisher SCIENTIFIC

Use All the Tools Together – Seeing Everything at Once



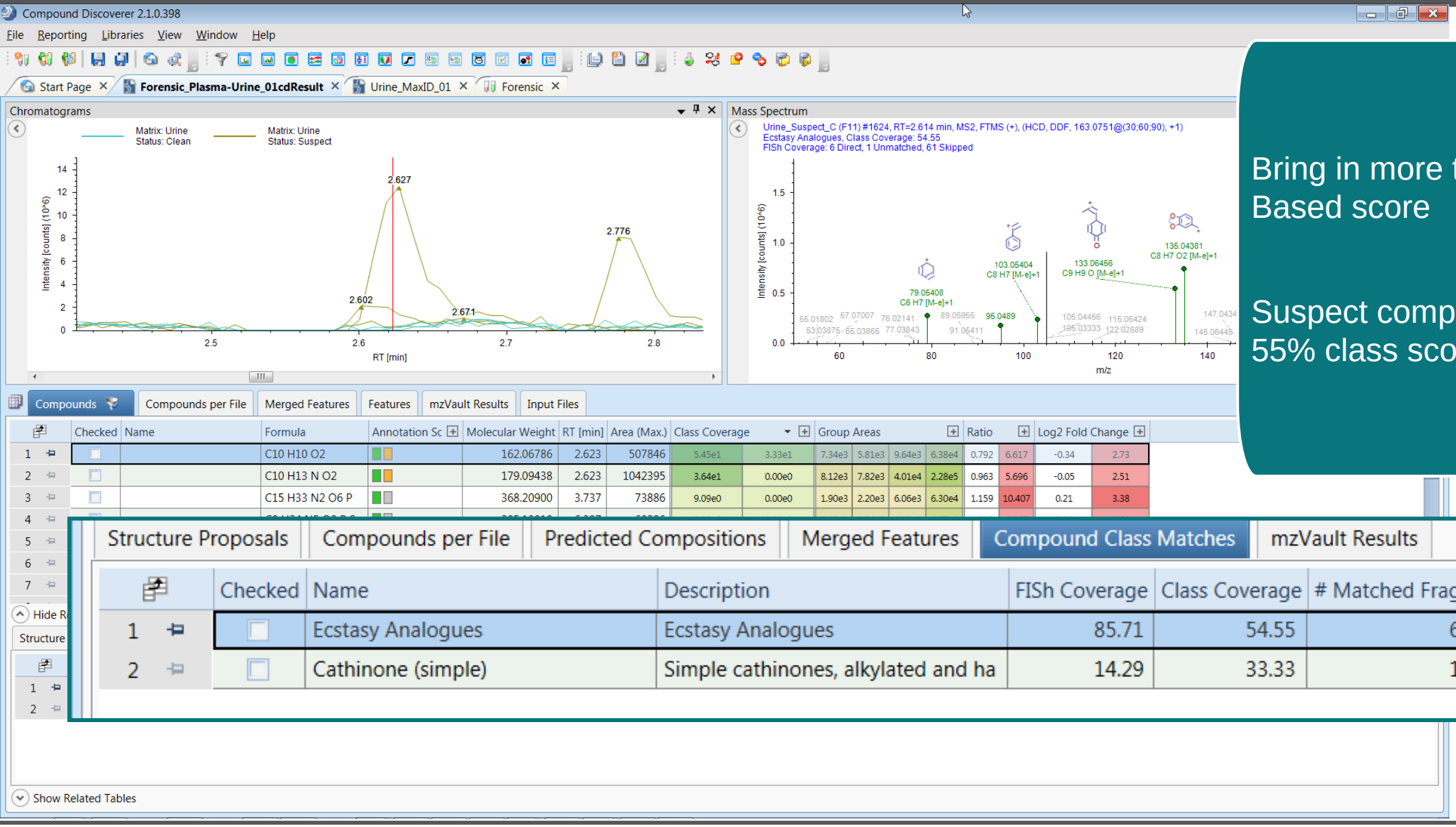
Urine (suspect)
Urine (clean)

Formula	Annotation Sc	Molecular Weight	RT [min]	Area (Max.)	Class Coverage		Group Areas				Ratio		Log2 Fold Change	
C10 H10 O2		162.06786	2.623	507846	5.45e1	3.33e1	7.34e3	5.81e3	9.64e3	6.38e4	0.792	6.617	-0.34	2.73
C10 H13 N O2		179.09438	2.623	1042395	3.64e1	0.00e0	8.12e3	7.82e3	4.01e4	2.28e5	0.963	5.696	-0.05	2.51
C15 H33 N2 O6 P		368.20900	3.737	73886	9.09e0	0.00e0	1.90e3	2.20e3	6.06e3	6.30e4	1.159	10.407	0.21	3.38
C8 H24 N5 O2 P S		285.13918	6.397	60296	9.09e0	0.00e0	2.68e3	2.71e3	6.57e3	4.25e4	1.013	6.467	0.02	2.69
		479.27302	6.126	51365	0.00e0	0.00e0	2.30e4	2.63e4	7.71e3	4.62e4	1.141	5.989	0.19	2.58
C7 H19 N10 O2 P		306.14372	5.995	213929	0.00e0	0.00e0	1.19e4	1.17e4	1.11e4	1.89e5	0.985	16.974	-0.02	4.09
C35 H78 N5 O8 P		727.55823	8.884	219342	0.00e0	0.00e0	6.32e3	4.81e3	6.96e3	1.98e5	0.761	28.395	-0.39	4.83

8		C19 H42 N9 O5 P		507.30542	5.985	171610	0.00e0	0.00e0	2.93e4	2.60e4	2.94e4	1.48e5	0.886	5.047	-0.18	2.34
9		C29 H61 N10 O4 P		644.46156	9.110	315458	0.00e0	0.00e0	9.02e3	7.48e3	9.58e3	2.98e5	0.829	31.141	-0.27	4.96
10				370.09288	0.248	10136097	0.00e0	0.00e0	1.23e5	3.65e4	6.87e4	2.53e6	0.297	36.858	-1.75	5.20
11		C39 H73 N O8		683.53146	8.985	522785	0.00e0	0.00e0	1.40e4	6.46e3	9.80e3	4.93e5	0.460	50.350	-1.12	5.65
12		C36 H61 N5 O2		595.48082	8.565	734812	0.00e0	0.00e0	4.09e4	1.01e4	1.96e4	2.87e5	0.246	14.651	-2.02	3.87
13		C38 H65 N5 O3		639.50613	9.101	777062	0.00e0	0.00e0	9.15e3	8.22e3	9.56e3	7.17e5	0.899	74.956	-0.15	6.23
14		C8 H11 N O		137.08389	0.652	5545592	0.00e0	0.00e0	1.18e4	7.86e3	8.41e5	4.95e6	0.664	5.881	-0.59	2.56
15		C5 H7 N O3		129.04239	2.546	9062178	0.00e0	0.00e0	1.04e4	5.84e3	1.55e6	8.66e6	0.559	5.585	-0.84	2.48
16				349.28157	8.360	457217			4.91e3	4.42e3	5.12e4	4.18e5	0.899	8.179	-0.15	3.03

ThermoFisher SCIENTIFIC

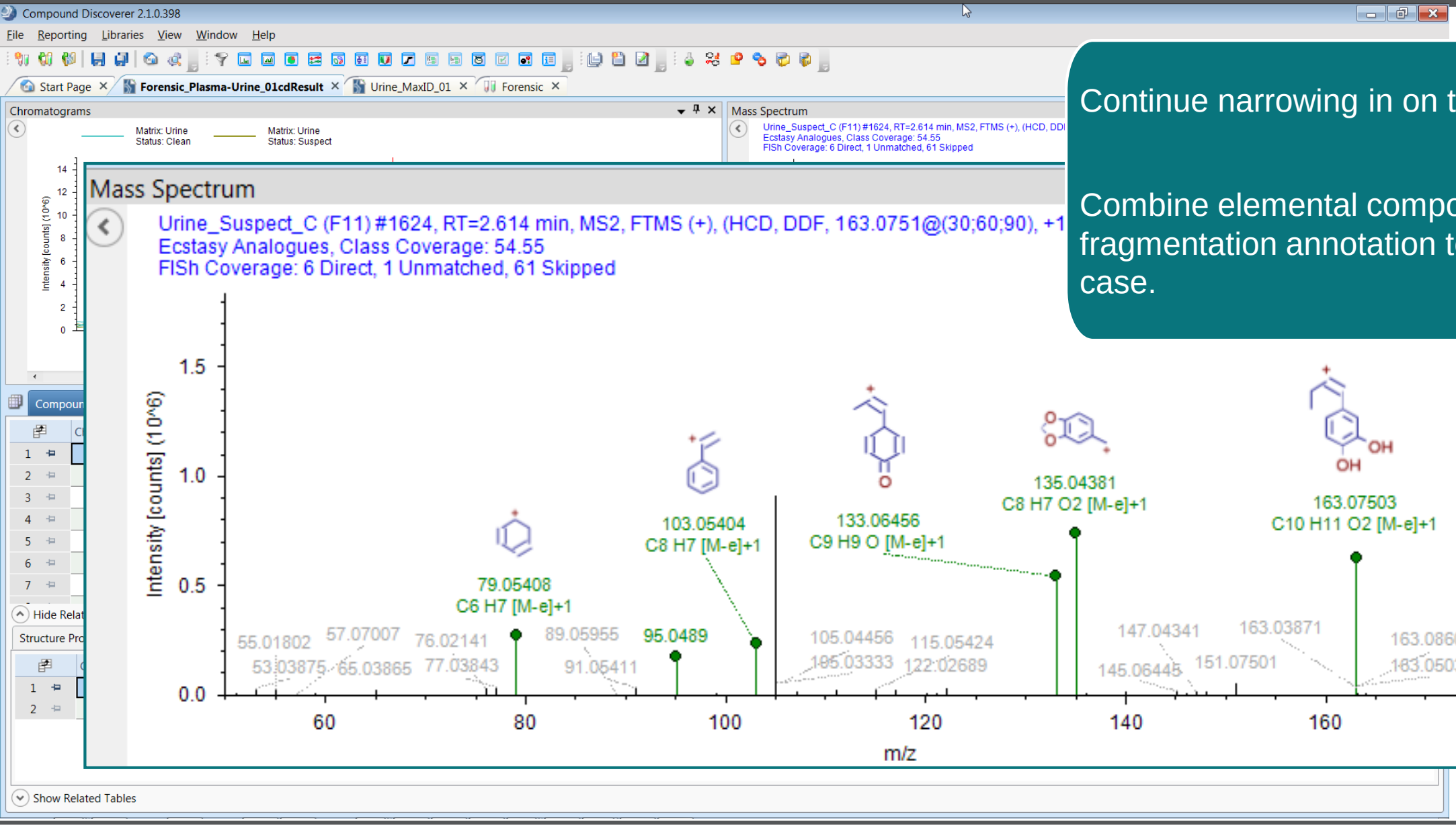
The New Tool in the Kit – Class Based Searching



Bring in more tools – Class Based score

Suspect compound has a 55% class score.

The New Tool in the Kit – Class Based Searching

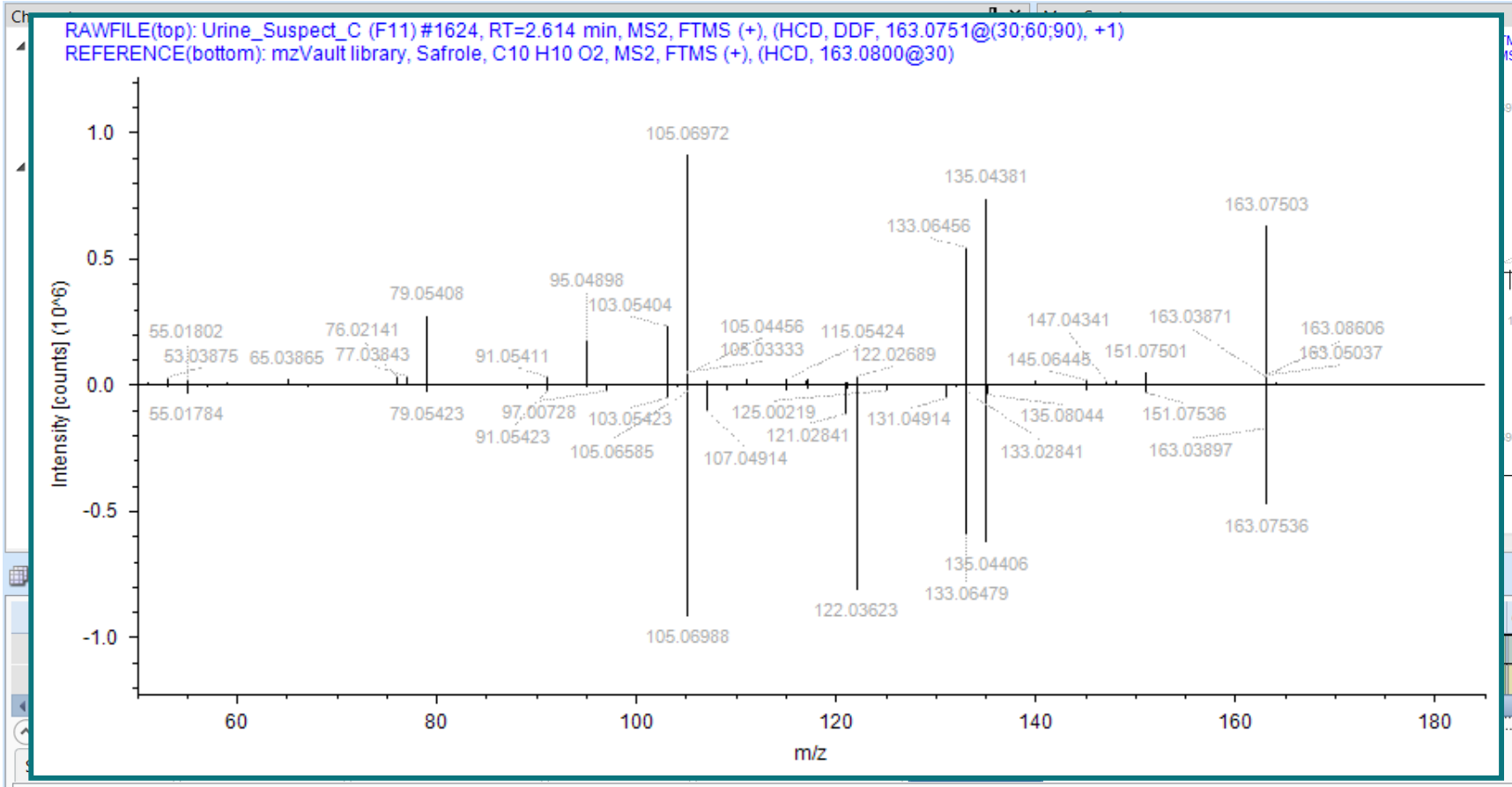


Continue narrowing in on the suspect.

Combine elemental composition and fragmentation annotation to build up the case.

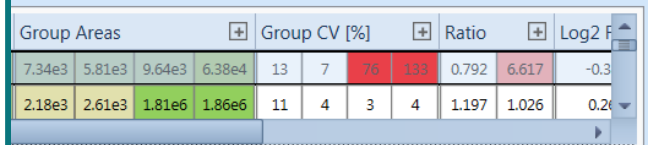
ThermoFisher SCIENTIFIC

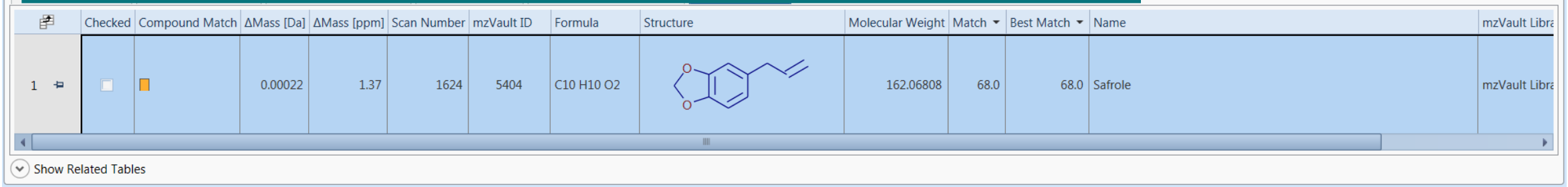
Use All the Tools Together – Seeing Everything at Once



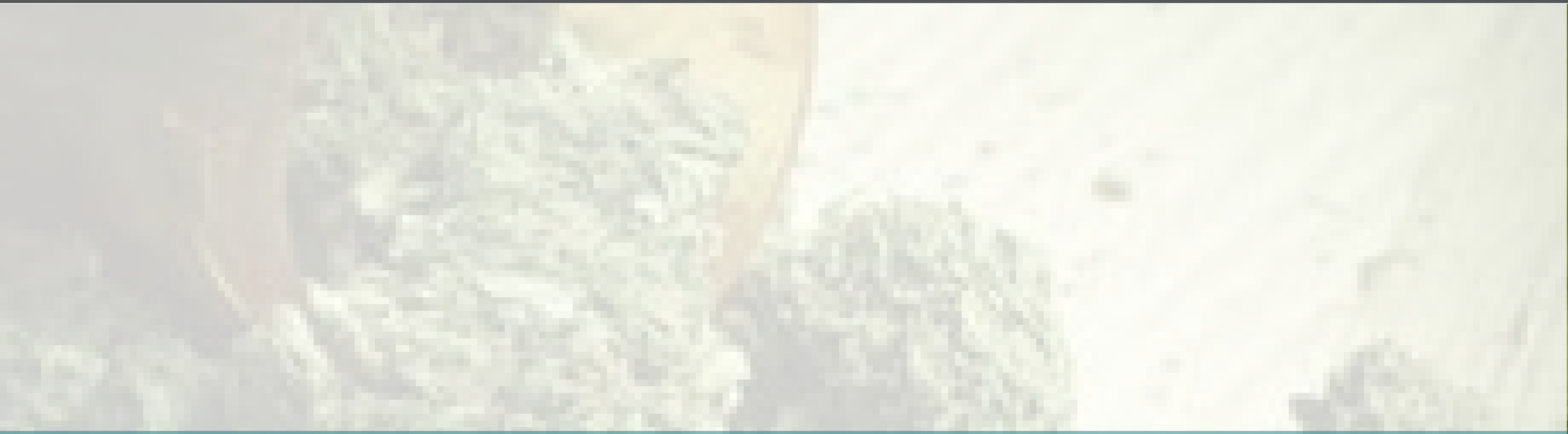
Use all the tools together

Similar spectra to safrole (and several other ecstasy analogues)





Emerging/Designer Drugs – “Fixing Bad”



Combine Multiple Tools

- Untargeted Detection and Differential Analysis
- Similarity Searching
- Class-Based Peak Annotation

[1] US DEA data through 2015 and discovery/synthesis rates from Cayman and LGC (2016-2017)

MyCompoundDiscoverer.com – Help and Information

Thank you.