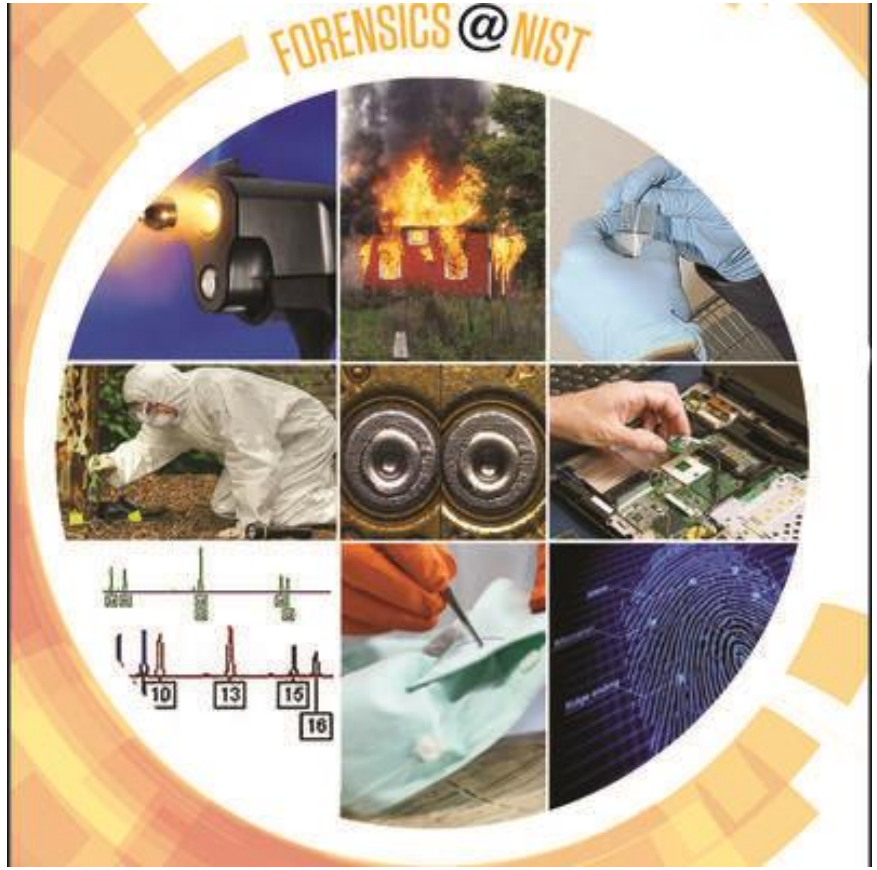
Unknown Identifications Using NIST EI Hybrid Search

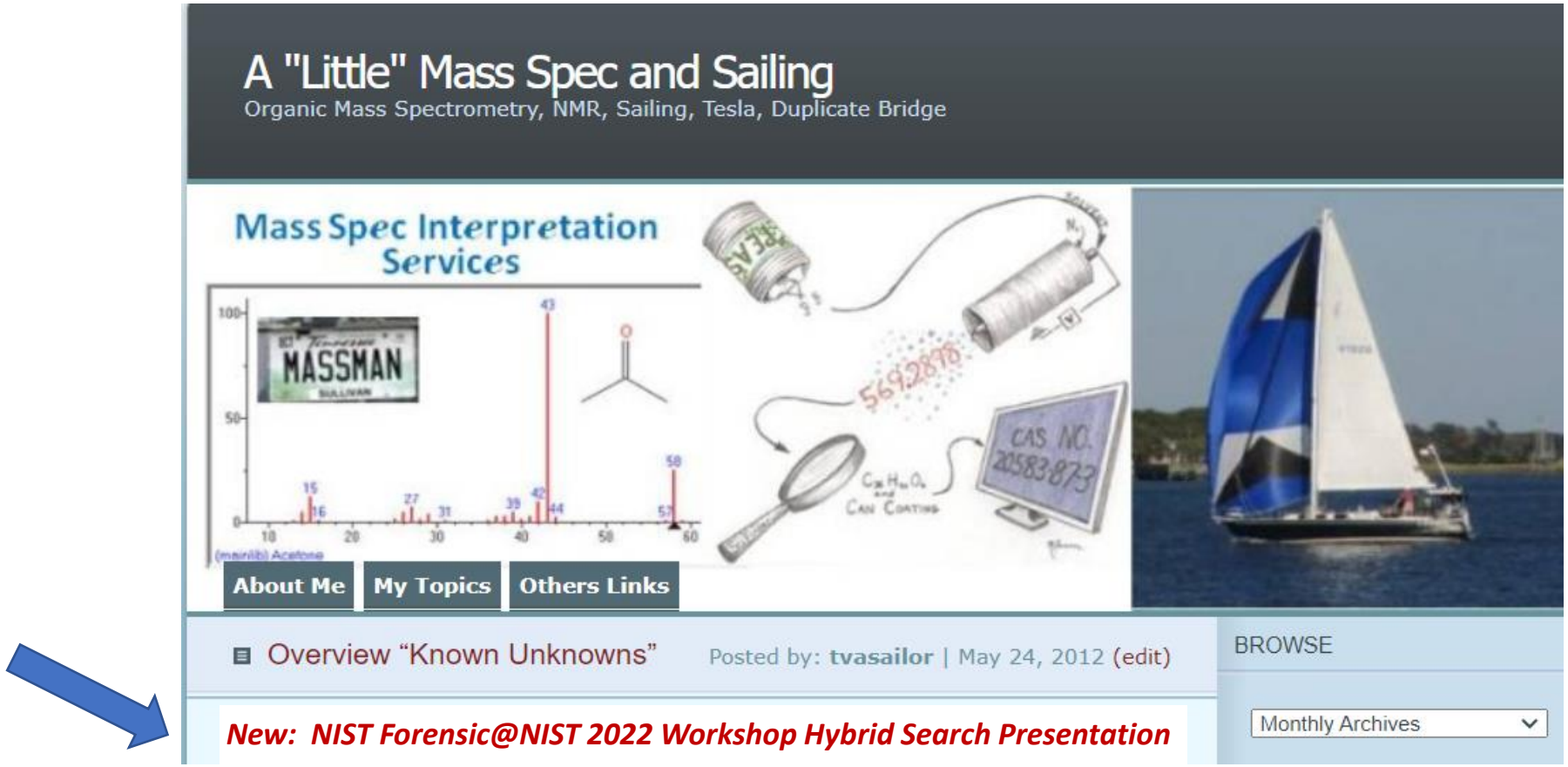
James Little
Mass Spec Interpretation Services



FORENSIC@NIST 2022 Workshop
Nov 15, 2022
Edward Sisco & Arun Moorthy
Mass Spectral Interpretation – Tips and Tools for GC-EI-MS and High-Resolution MS Data

Resources Found on Personal Website

- Todays workshop presentation
- Workshop handout with links to resources on my website
- ***Free*** training courses on NIST EI and MSMS hybrid searches
- Variety of other presentations on mass spec techniques



Experience in Unknown Identifications

- Retired Research Fellow, Eastman Chemical Company
- 42 Years Experience
- 6 Years Consulting
- Unknown Identification Using GC-MS, LC-MSMS, and ***Other*** Techniques

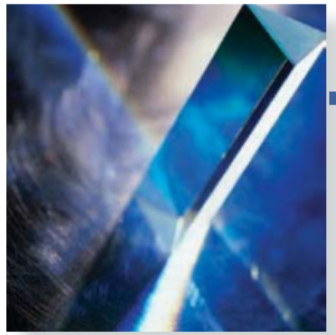
Eastman Chemical Company, Main Site, Kingsport, TN
50 Manufacturing Sites Worldwide, ~14,500 Employees

>50 Mass Specs Networked
Worldwide

LCGC Article Outlines Eastman's Approach for Unknown Identifications

In many cases, someone knows the identity of my unknown, I just need to find that information.. Thus the term "Known Unknowns.."

114 LCGC NORTH AMERICA VOLUME 31 NUMBER 2 FEBRUARY 2013 www.chromatographyonline.com



MS – THE PRACTICAL ART

Identifying "Known Unknowns" in Commercial Products by Mass Spectrometry

There are known knowns. These are things we know that we know. There are known unknowns. That is to say, there are things that we know we don't know. But there are also unknown unknowns. There are things we don't know we don't know.

Donald Rumsfeld

Former United States Secretary of Defense

"Tools in Toolkit" at Eastman to Identify "Known Unknowns"

- Sample history and known components in mixtures
- Commercial MSMS and EI libraries using NIST search
- Accurate mass analyses
- "Spectraless databases" such as CAS Registry, ChemSpider, etc.
- Chemical ionization (CI) GC-MS
- Training courses for NIST EI and MSMS search software
- Extensive EI and MSMS corporate user libraries ***shared and updated nightly***
- NIST MS Interpreter fragmentation program
- NMR (H, C, P, F), 1D and 2D
- IR
- Synthesis of enriched samples by ***known chemistry*** in 1-2 g quantities for confirmation by NMR and MS ***without purification***
- Deuterium exchange CI GC-MS for exchangeable protons
- Deuterium exchange infusion MSMS for exchangeable protons
- Derivatizations for GC-MS

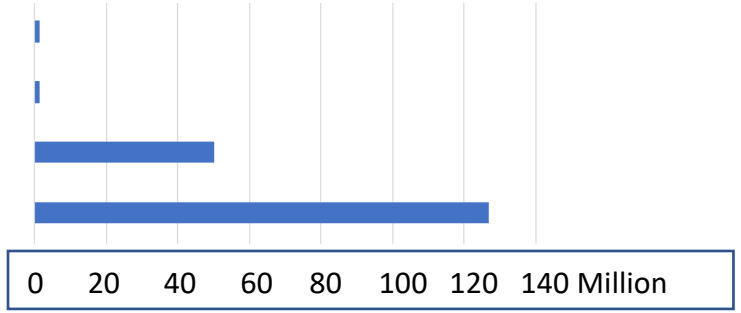
"Known Unknowns" Identified with "Spectraless" Databases

In general Ones with No Associated Spectra

SciFinder® A CAS SOLUTION

ChemSpider The free chemical database

~1.2M EI spectra in various libraries
~1.2M MSMS spectra in various libraries
~50M records in ChemSpider
~127 M records in CAS Registry, ~15K added each day



- ***Huge*** # of records in ChemSpider and the CAS Registry compared to spectra available in mass spectral libraries
- ChemSpider free, fee for CAS Registry (SciFinder)
- Search ChemSpider and CAS Registry by molecular weight or molecular formula
- CAS Registry by molecular formula only
- ChemSpider by both accurate mass and molecular formula
- Searches refined by key words, no. of associated references, substructure, *etc.*
- Final candidate(s) reviewed using fragmentation, sample history, *etc.*

Complimentary NMR Data Used with MS for Identifications

- NMR very useful for mixtures when ***utilized with*** structures proposed by mass spec
- Components confirmed can then be quantified *in mixtures* by NMR and used as standards for calibrating routine chromatographic techniques
- Primary standards are not needed for quantitation by NMR when an internal standard is added to the mixture

Low-Field Table Top NMR

Hi-Field NMR

- Hi field NMR's used at Eastman employ cryogens and expensive to purchase and maintain
- New low field table top systems use permanent magnets and much lower associated cost
- Latter type are limited in sensitivity and resolution, but should be very useful for many applications

Novel NIST EI Hybrid Search*

Program Description:

- Hybrid search generates a similarity score matching fragments ***and*** neutral losses
- Greatly extends the scope of the library
- Mass difference must be confined to a single region of molecule and no significant alteration of fragmentation behavior
- *DeltaMass* is the molecular weight difference between query and library compound and reflects the modification of the molecule

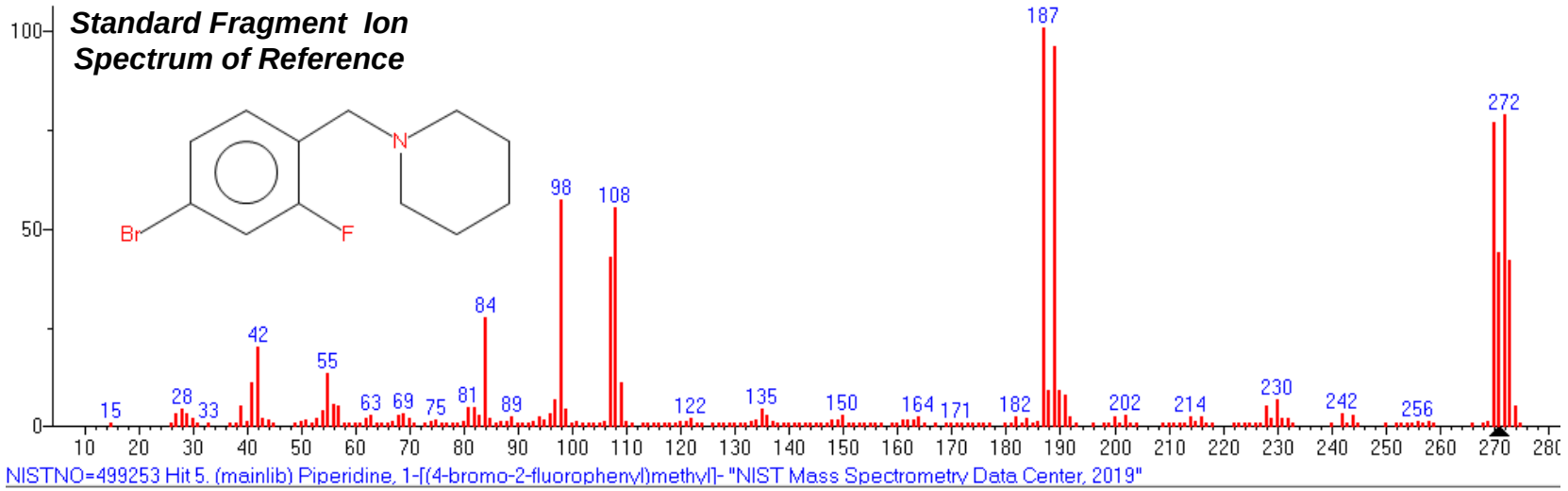
My Personal Experience:

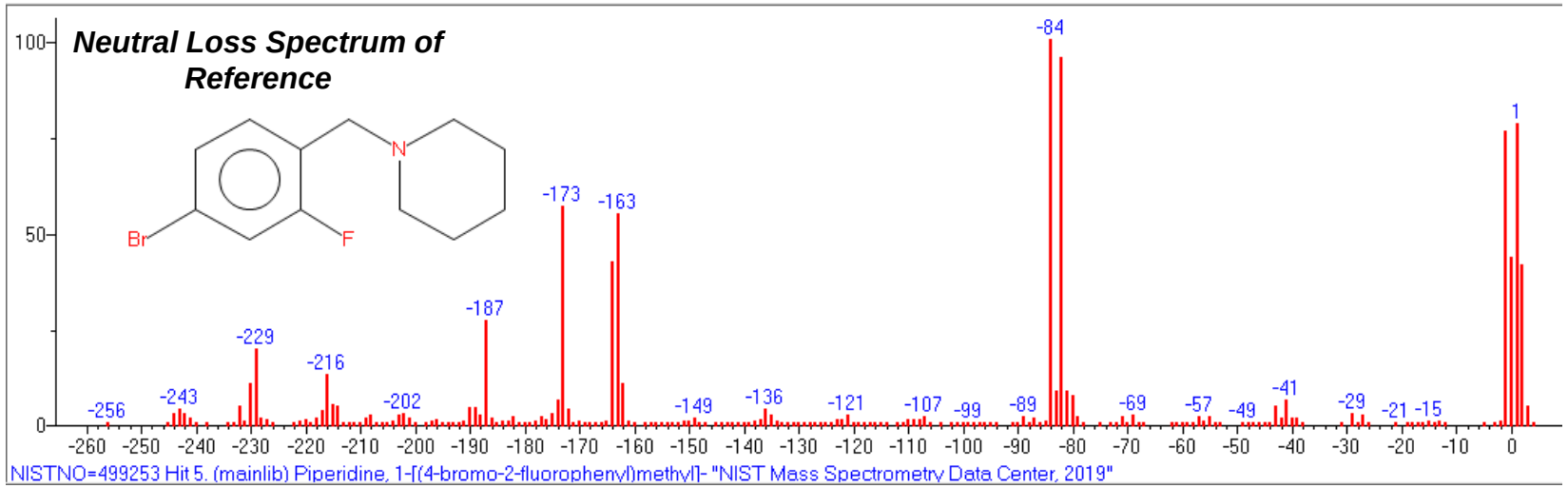
- Used for >40,000 searches in 3 years in evaluating new EI library spectra for NIST
- *Routinely amazed* by the types of similar compounds with high match factors
- Frequently yields useful results not noted in “simple” identity searches
- Very useful in identifying unknowns, finding similar model compounds, and supporting fragmentation mechanisms

*”Combining Fragment-Ion and Neutral-Loss Matching during Mass Spectral Library Searching: A New General Purpose Algorithm Applicable to Illicit Drug Identification,” A. Moorthy, W. Wallace, A. J. Kearsley, D. Tchekhovskoi, and S. Stein, *Analytical Chemistry* **2017** *89* (24), 13261-13268.

Hybrid Search *Combines Scores* of Standard Identity Search with Neutral Loss Search

- First searches unknown standard spectrum against library spectrum and generates a standard identity match factor
- Then searches unknown neutral loss against neutral loss spectra of library entries
- Generates a combined “hybrid” score



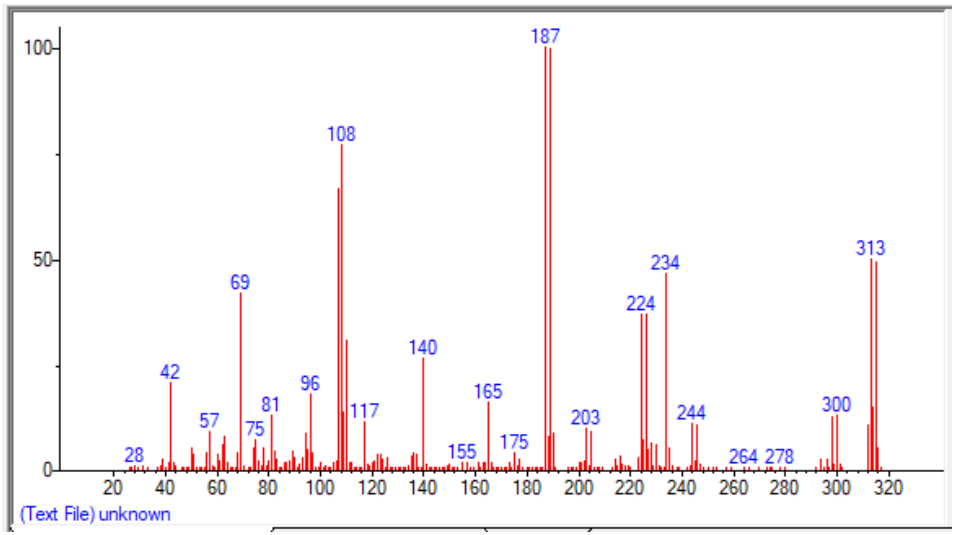


Hybrid Search Results for Unknown

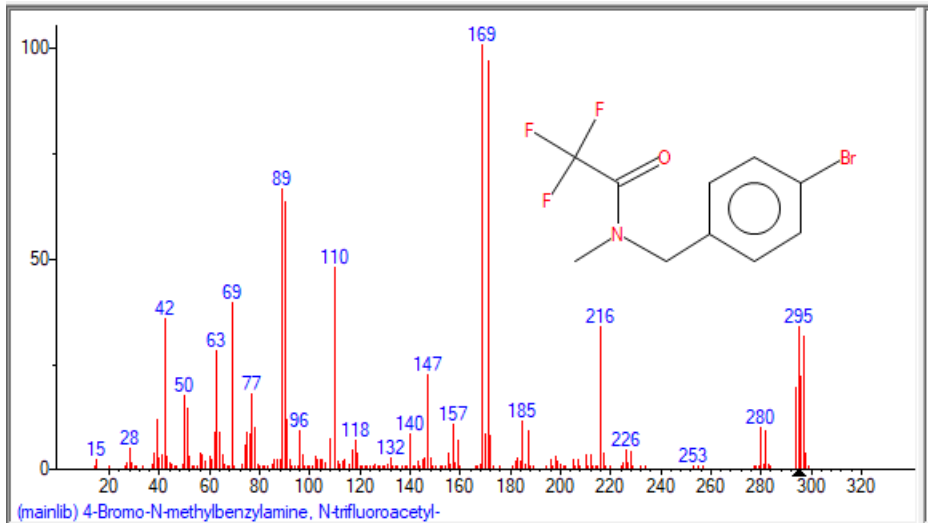
1) Best hybrid match factor is 908, next closest is 781
2) ***DeltaMass*** is 18, common for fluorine (mass 19) replacing hydrogen (mass 1) on ring, *i.e.* 19-1=18

#	Lib.	DeltaMass	Match	o.Match	Name
1	M	18	908	232	4-Bromo-N-methylbenzylamine, N-trifluoroacetyl-
2	M	70	781	559	N-[(5-Bromo-2-fluorophenyl)methyl]cyclopropanamine
3	M	91	766	437	1-Bromo-2-chloro-5-fluoro-4-methylbenzene
4	M	40	755	457	4-(4-Bromo-2-fluorobenzyl)morpholine

? Unknown

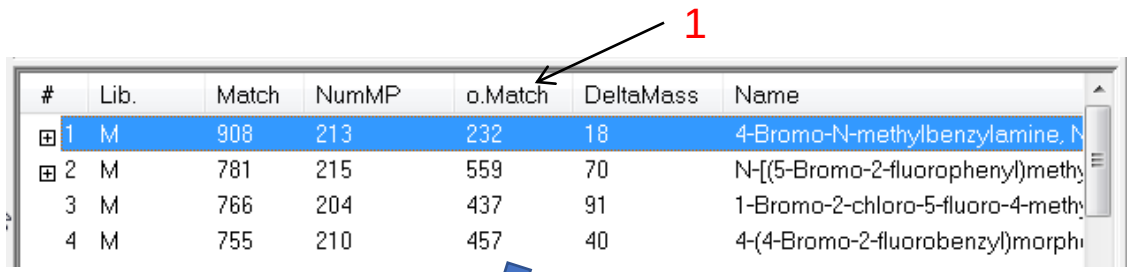


Hybrid Match 908

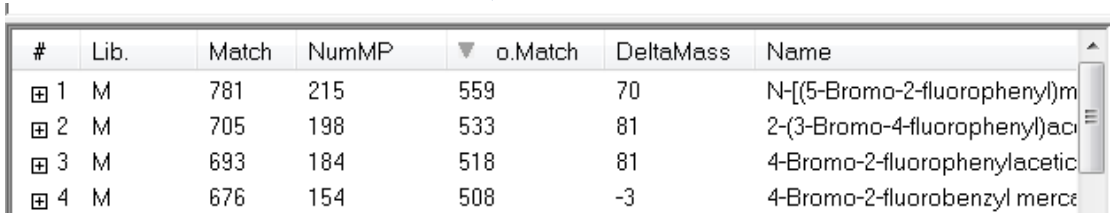


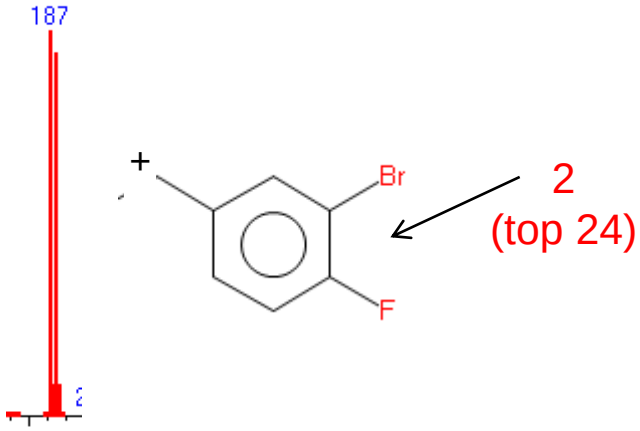
See Standard Identity Results *in Addition* to Hybrid

1. Resort “hybrid” search results by standard identity search match factor
2. Top 24 hits contain the subsubstructure with F and Br on a benzyl group, *m/z* 187



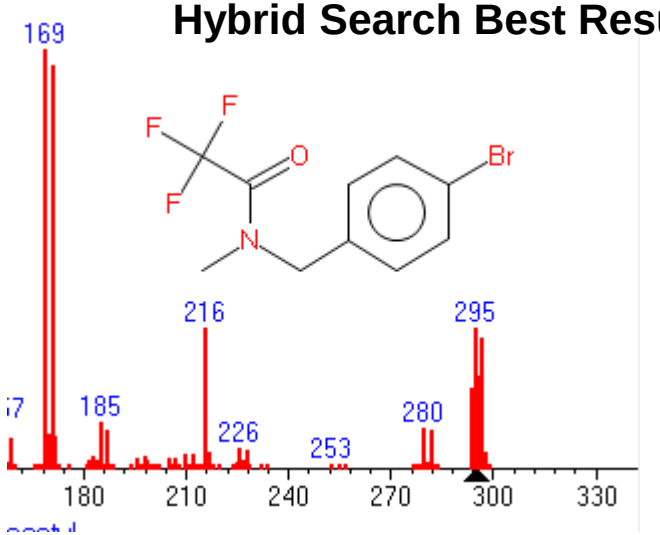
Resorted by o-match
EI Standard Identity Match Factor



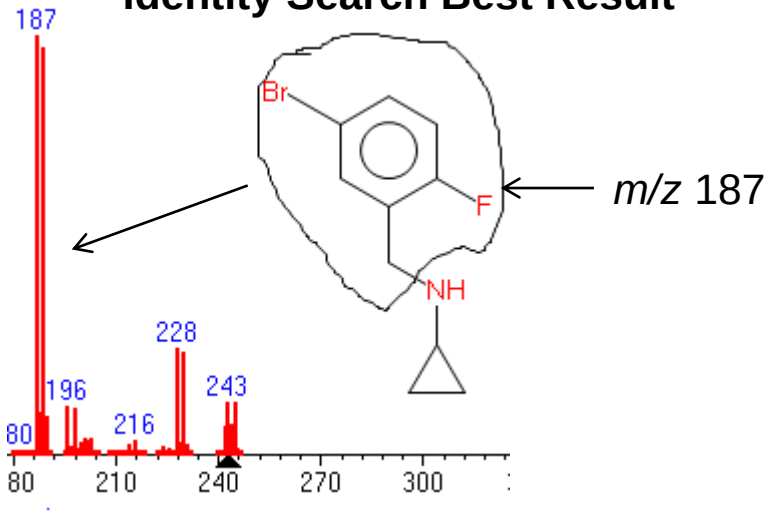


“Mentally” Merge Information of “Hybrid” and “Simple” Identity EI Search

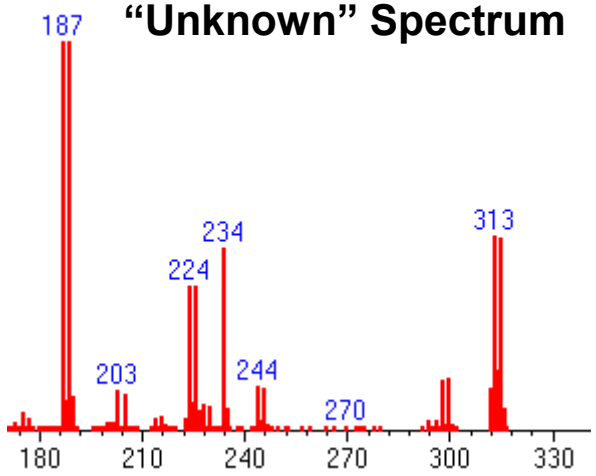
Hybrid Search Best Result



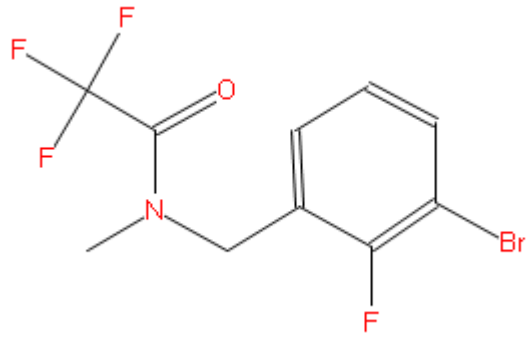
Identity Search Best Result



“Unknown” Spectrum



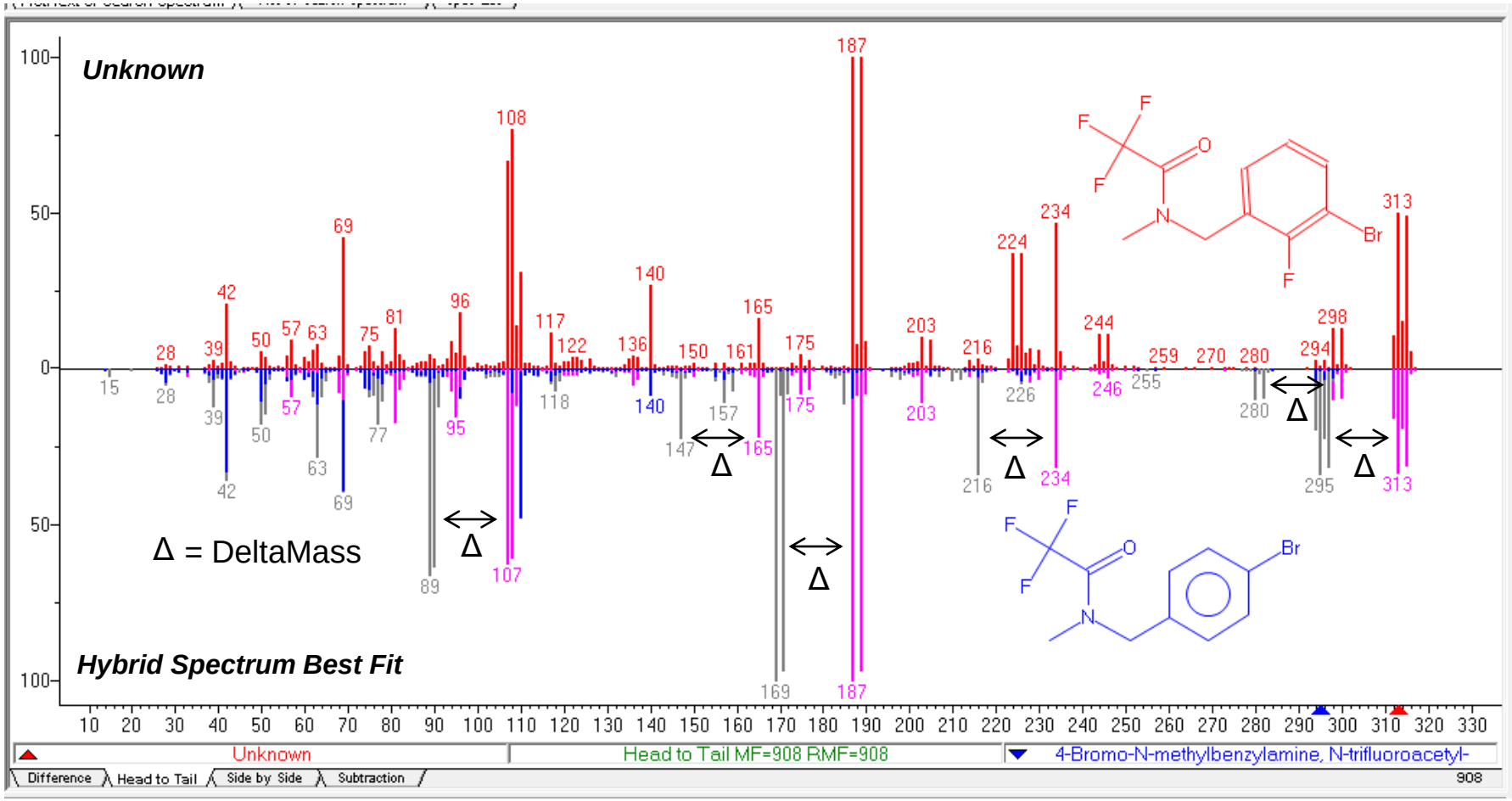
“Unknown Identity or Isomer”



Closer Look at Middle Display

Top is Unknown, Bottom is "Hybrid" Spectrum

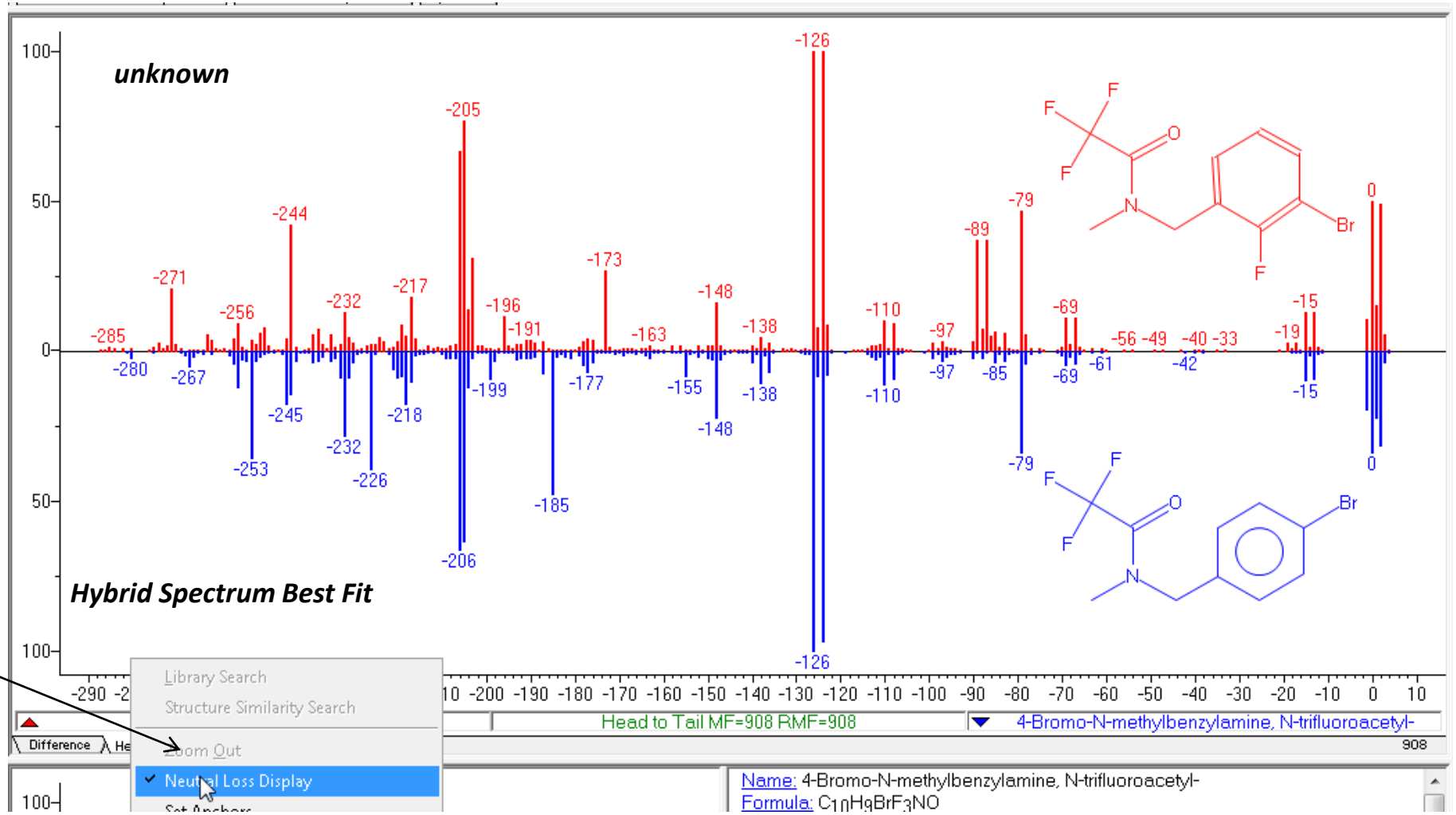
- In bottom spectrum, original ions in grey shifted by DeltaMass (Δ) 18 for user visual comparisons
- Can take a while to adjust to this view versus standard "Head to Tail" views



Alternate Comparison of Hybrid Spectrum: Neutral Loss Display

Top is Unknown, Bottom is "Hybrid" Spectrum

1. Can view alternatively as a neutral loss comparison
2. Shows whole spectrum of reference shifted by DeltaMass of 18
3. ***Easier*** and ***more efficient*** to ***look at "Hybrid" display*** with experienced eye!



1

*Hybrid Search **Needs** Nominal MW of Species to Work Properly!*

- Many EI spectra do not have molecular ions (~20% in NIST library)
- ***User*** must decide by:
  - Letting program determine automatically
  - User proposing from logical losses at higher m/z in spectrum
  - Chemical Ionization
  - Use value proposed by two different NIST algorithms within software

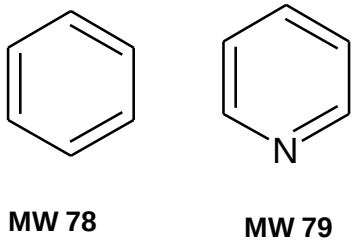
DeltaMass Table

- Useful Hybrid DeltaMass values noted in my spectral evaluations for NIST
- ~600 values in Excel spreadsheet
- Value note can be + /–, depending only species present/absent

DeltaMass Nominal	Group/Element (1)	Group/Element (2)
1	CH3SO2 group	NH2SO2 group
1	methyl on armoatic ring	amine on aromatic ring
1	CH2NH2 group	CH2CH3 group
1	nitrogen in heterocyclic aromatic ring	phenyl aromatic ring
1	amine on aromatic ring	phenol ring
1	N in six membered heterocylic aromatic	aromatic ring no nitrogen incorporated
1	insertion of nitrogen in place of carbon in 5-membered heterocyclic ring	no insertion
1	TBDMS derivative attached to two NH groups	TDDMS derivative inserted on O and one in group on aromatic ring
1	CH2NH2 onaromatic ring	CH2OH on aromatic ring
1	acid on aromatic ring	amide on aromatic ring
1	acid and amide on aromatic ring	two acids
1	methyl and methoxy on aromatic ring	dimethylamine on aromatic ring
1	methoxy on aromatic ring	N-methyl group on aromatic ring
1	methyl on pyrimidine ring	methyl on pyrimidine ring
2	MS2 of a M+2 isotope ion in compound such as Br, Cl	MS2 of a M+2 isotope ion in compound such as Br, Cl
2	two fluorines on aromatic ring	one chlorine on aromatic ring
2	cyclopentyl ring	furan ring
2	alkene	alkyne group
2	Sulfonyl fluoride	Sulfonic acid
2	OH on aromatic ring with nitrogen in ring	NH2 on aromatic ring with no nitrogen in ring
2	aromatic aldehyde	methoxy ether of phenol
2	3-methylbutylamide group	piperidine amide group

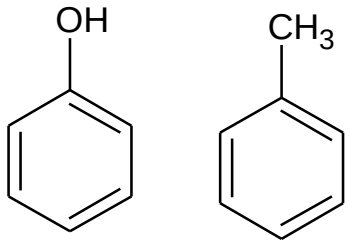
Associating Some *Simple* Structures with DeltaMass Values

- Some simple ***small*** MW compounds to illustrate types of substructural information
- ***Note:*** Odd values of DeltaMass contain one nitrogen change in structure, thus ***"Nitrogen Rule"***
- Isotope ratios and/or accurate mass helpful with redundancies
- Of course, these substructures can be a part of ***larger*** molecules



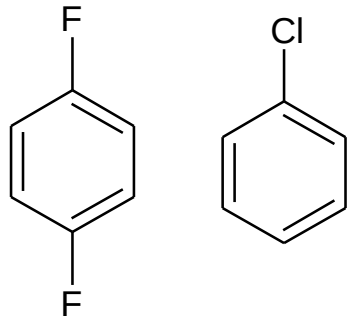
MW 78 MW 79

DeltaMass 1



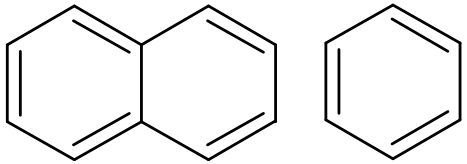
MW 94 MW 92

DeltaMass 2



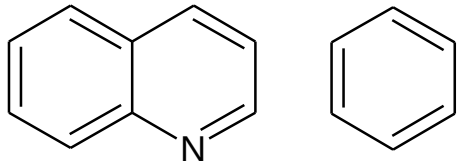
MW 112 MW114

DeltaMass 2



MW 128 MW 78

DeltaMass 50

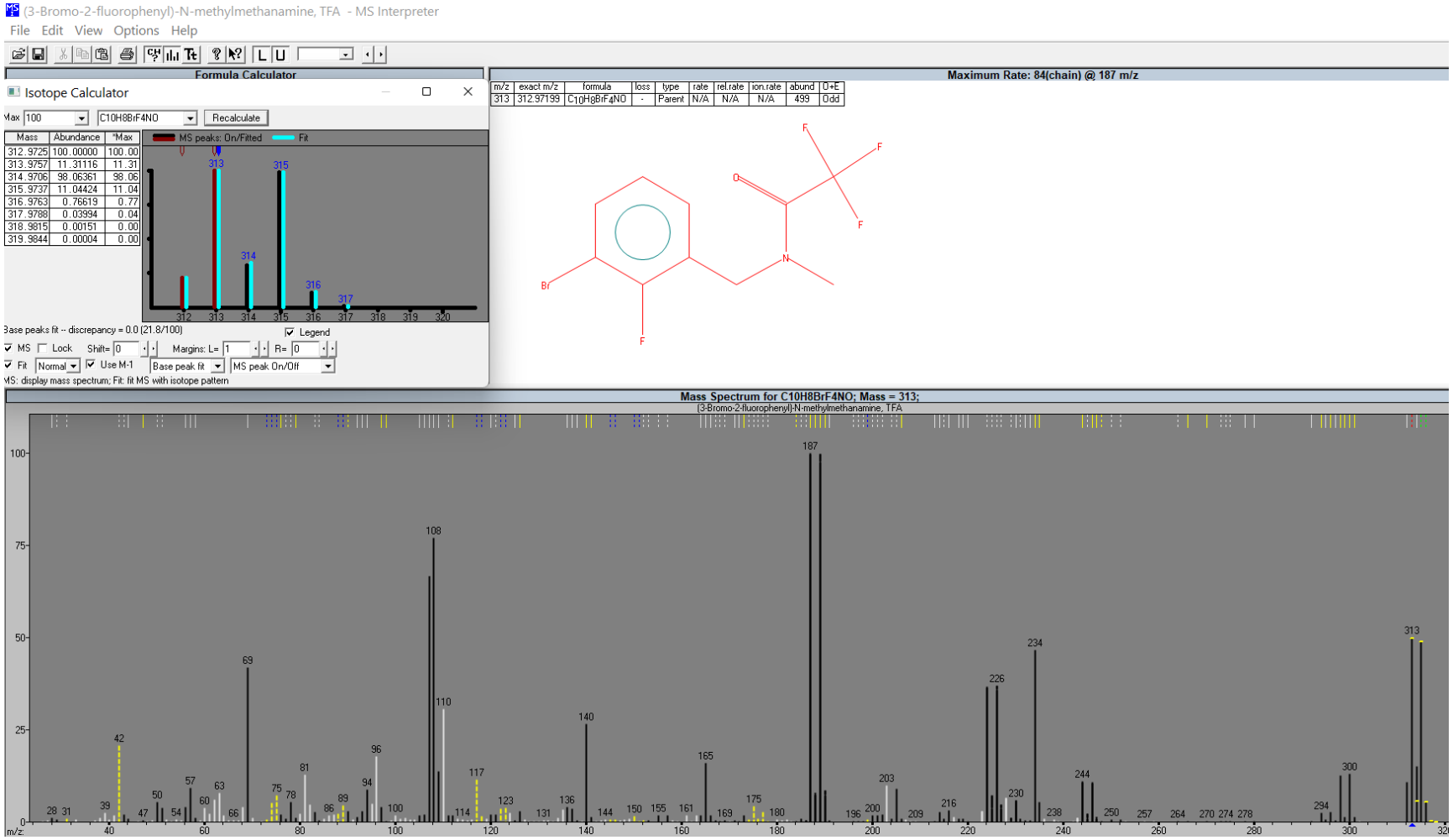


MW 129 MW 78

DeltaMass 51

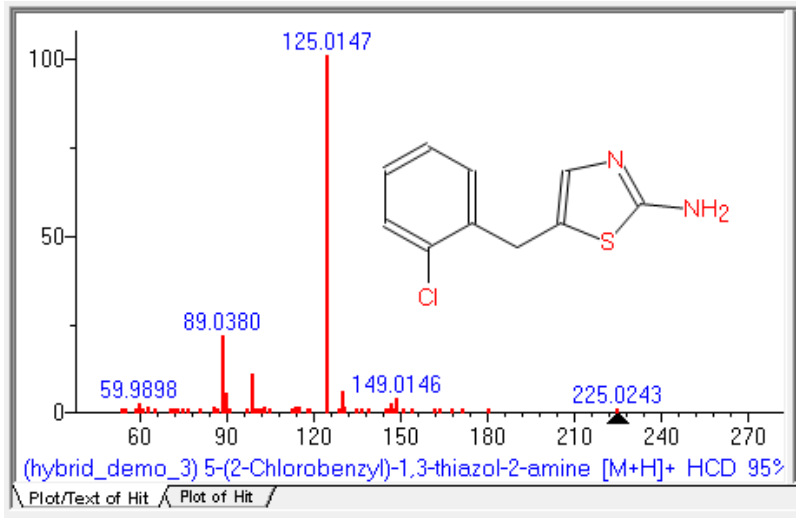
Powerful NIST MS Interpreter Program Correlates Ions to Structures

- Ions in “color and black” explained by program
- Isotope ratios
- Logical fragments
- Mechanisms and relative fragmentation rates
- Detailed descriptions in my ***free*** courses



Hybrid Search *Also* Used for LC-MSMS Unknowns

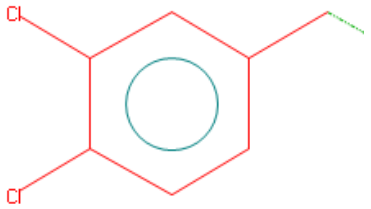
"Hybrid" Best Result
Accurate Mass DeltaMass, +Cl-H

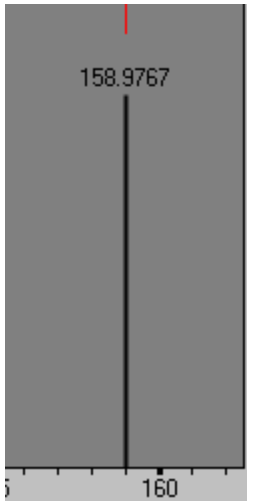


Top 20 **"Identity Searches" show 2 Cl's on benzyl ring**

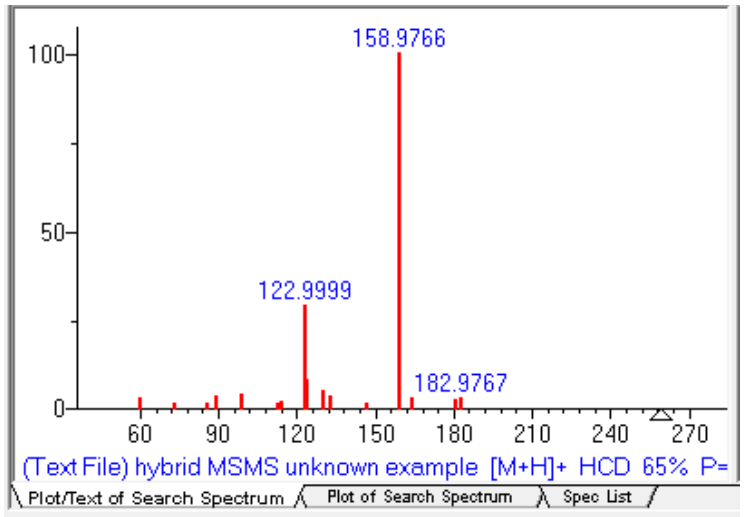
Maximum Rate: 65(chain) @ 158.9763 m/z

m/z	exact m/z	formula	loss	type	rate	rel.rate	ion.
159	158.976282	$C_7H_5Cl_2$	CH_4N_2S	H-Displacement	65	78	5

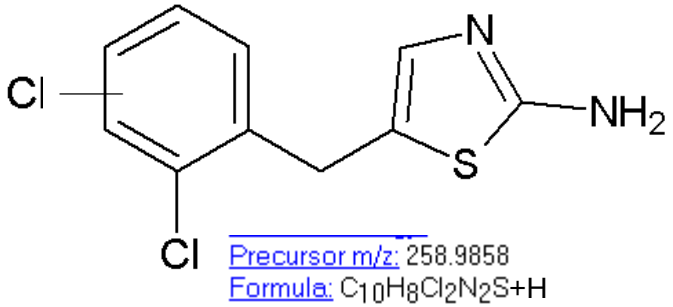




"Unknown" Spectrum



"Proposed Structure for Unknown"



DeltaMass Table

- Values in Excel DeltaMass Table only shown in nominal mass
- E.g. below see entries for DeltaMass = 34 nominal
- Manually added accurate DeltaMass column in example below

			Manually Calculated
34	pyridinyl group	dimethylamino group	33.9843
34	Sulfonamide group on aromatic	nitro on aromatic	34.9877
34	chlorine on aromatic ring	phenyl ring	33.9611
34	CF3 on aromatic	chlorine replacing	34.0263

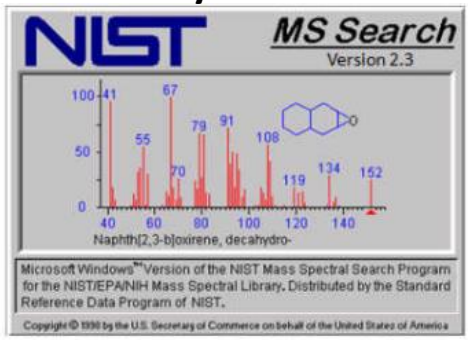
versus observed = 33.9610

“Real-World Example” of the Identification of PCP-Related Compound

O. David Sparkman, University of the Pacific
James Little, Mass Spec Interpretation Services

White Powder

NIST Hybrid Search

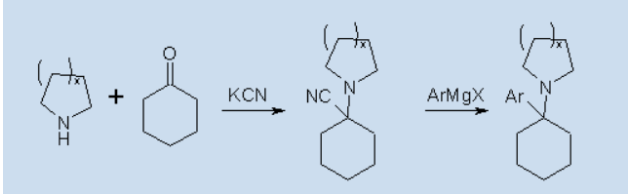


“Spectraless” Databases

SciFinder® A CAS SOLUTION

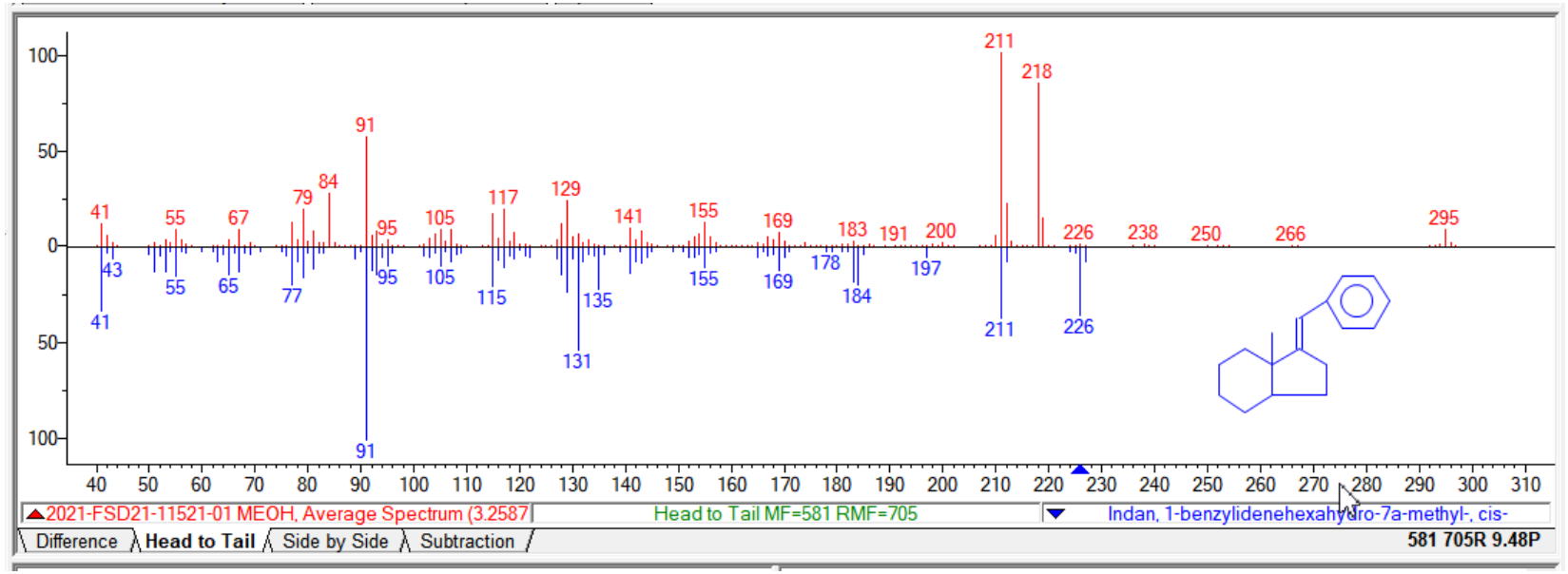
ChemSpider The free chemical database

Chemistry



Standard Identity Search Results for EI GC/MS

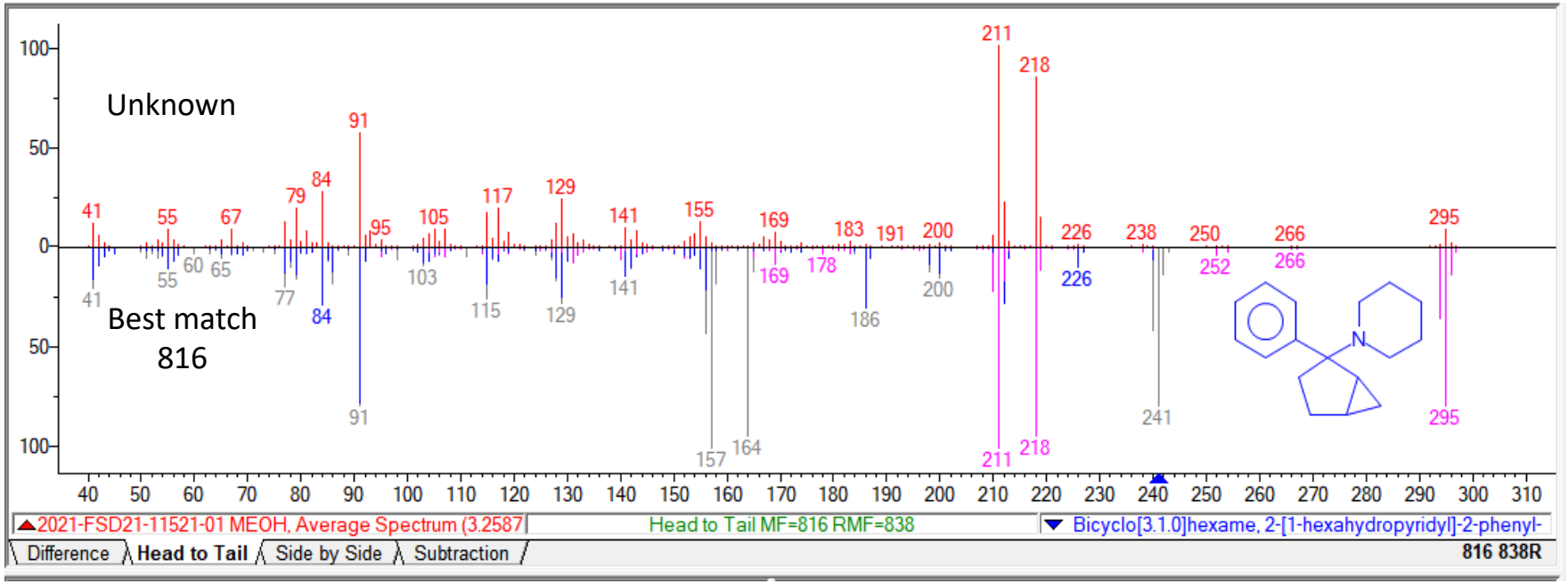
- Standard identity MS search *failed*
- Match is <600
- *No* useful results



#	Lib.	Match	Prob. (%)	R.Match	Syn	DBs	Name
⊞ 1	M	581	9.48	705	2	0	Indan, 1-benzylidenehexahydro-7a-methyl-, cis-
⊞ 2	M	580	9.11	706	1	0	1H-Indene, octahydro-7a-methyl-1-(phenylmethylene)-
3	M	570	6.43	661	1	0	9-Borabicyclo[3.3.1]nonane, 9-(1,2-diphenylethyl)-
4	M	563	4.92	665	1	0	1-Ethoxy-7-phenylvinylidene-bicyclo[4.1.0]heptane
5	M	563	4.92	662	0	0	(5R,6S)-6-Phenyl-5-(phenylsulfonyl)-1-(prop-2-yn-1-yl)...

Hybrid Search Results for EI GC/MS

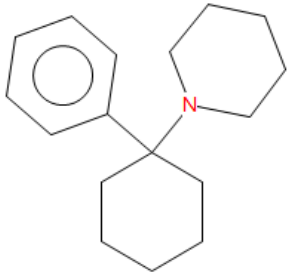
- Hybrid search yields one promising hit
- Match is >800
- DeltaMass = 54



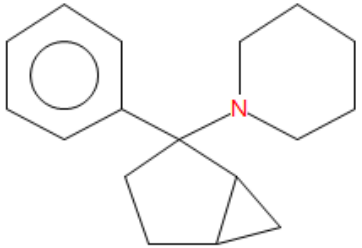
#	Lib.	DeltaM...	Match	R.Match	o.Match	NumMP	o.NumMP	Syn	DBs	Name
1	M	54	816	838	348	142	93	1	0	Bicyclo[3.1.0]hexame, 2-[1
2	M	77	756	761	337	161	100	3	0	1,3-Cyclohexanedione, 5-(
3	M	53	731	733	345	166	125	4	1	Methyl 4-hydroxy-3,5-dinitr
4	M	54	727	729	364	169	134	0	0	N-Allyl-N-methyl-1-phenylh
5	M	55	703	706	563	172	137	1	0	1-Ethoxy-7-phenylvinylider
⊞ 6	M	69	699	750	587	149	97	2	0	Indan, 1-benzylidenehexah

Summary of MS Results

- Hybrid search indicates PCP-related compound
- Δmass = 54 from the hybrid search *not* associated with values noted in my Deltamass table
- Accurate mass data for the unknown indicated a molecular formula of $C_{21}H_{29}N$ via DART MS analyses (E. Sisco, NIST)
- Best match has a molecular formula of $C_{17}H_{23}N$
- Indicates addition of C_4H_6
- Plus, additional double bond equivalent (DBE)



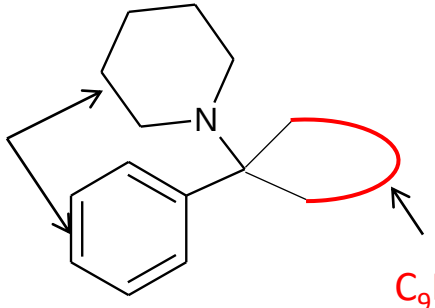
"Angel Dust"
Phenylcyclohexyl piperidine
PCP
Phencyclidine



Best Match
$C_{17}H_{23}N$
MW 241
7 DBE's

$+C_4H_6$

$C_{12}H_{15}N$
5 DBE's



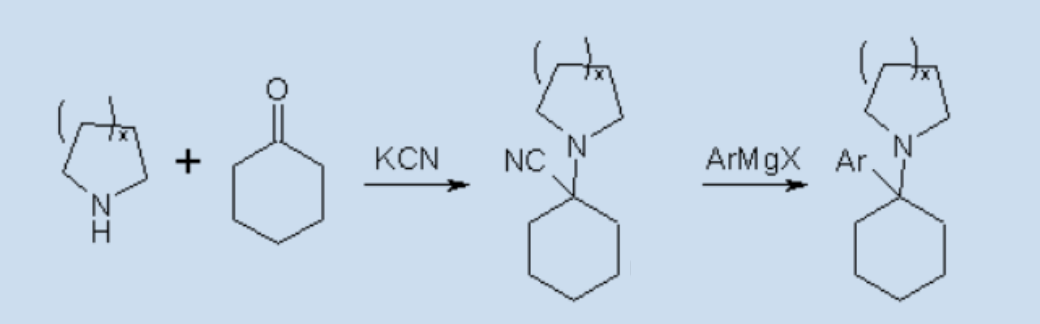
C_9H_{14}
3 DBE's

Total Structure
$C_{21}H_{29}N$
MW 295
8 DBE's

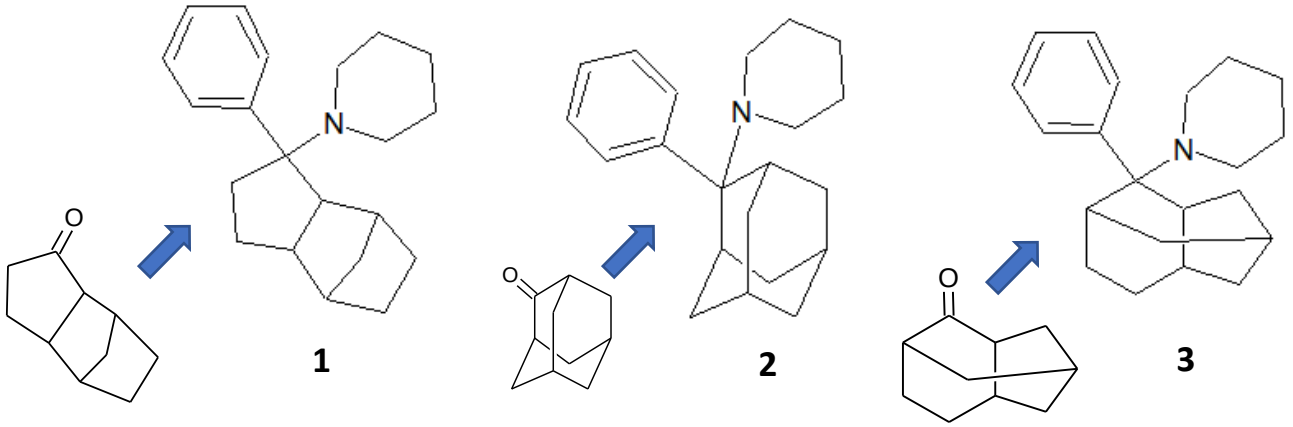
Structures Proposed from Chemistry

- Considered chemistry * to propose 3 structures
- PCP related species *could* be made with the reaction below from 3 commonly available ketones

Proposed Chemistry



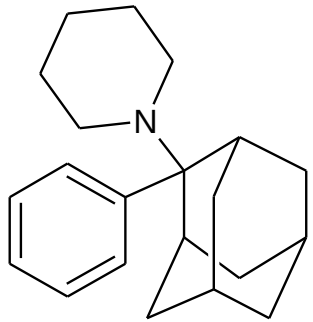
3 Proposed Structures from Chemistry and Easily Obtained Ketones



***Illicit Synthesis of Phencyclidine (PCP) and Several of Its Analogs**
by A.T. Shulgin and D.E. Mac Lean, *Clin. Toxicol. 9(4), 553-560 (1976)*
https://www.designer-drug.com/pte/12.162.180.114/dcd/chemistry/pcp.shulgin.html

Conclusions Leading to Identification of Unknown

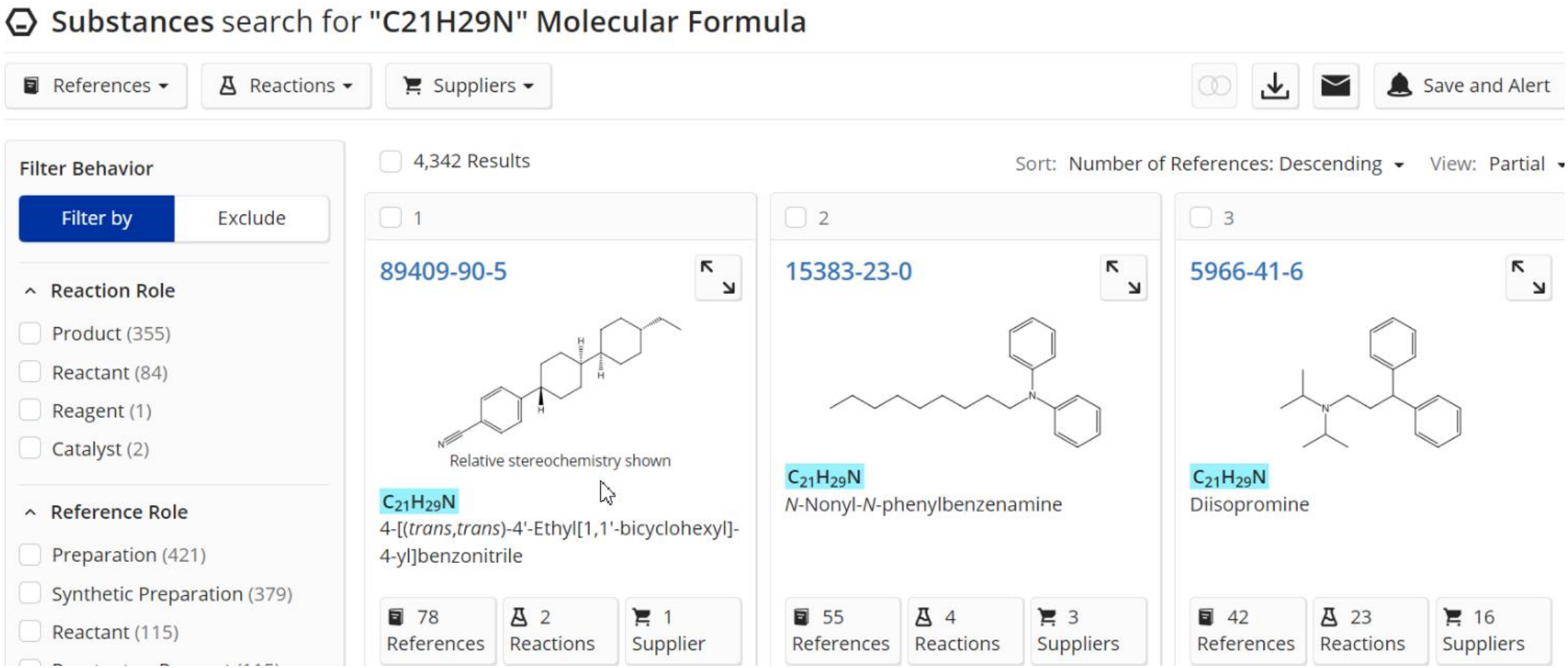
- Hybrid search was critical in suggesting PCP-related substructure
- DeltaMass was not easily associated to a definitive fragment in my Deltamass table
- Molecular formula from Accurate mass DART critical step (E. Sisco, NIST)
- Initially, molecular formula and DBE's plus chemistry to propose 3 structures
- Structure confirmed by proton NMR as structure below (A. Urbas, NIST)
- Later, "Spectraless" approach with ChemSpider and SciFinder demonstrated



Identity of PCP-Related Unknown
CAS No. 72241-00-7

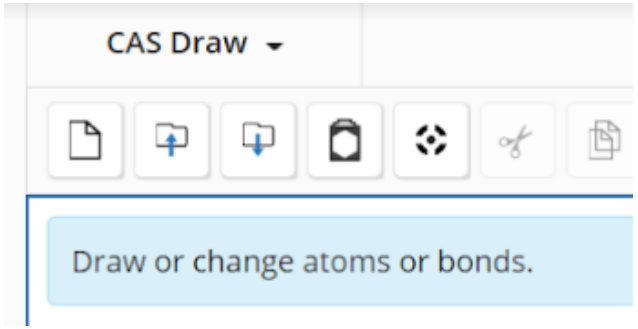
Alternate Approach Using SciFinder and the Hybrid Search Results

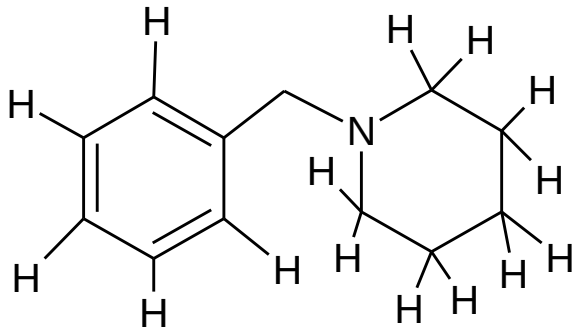
- SciFinder using ***molecular formula search***
- Results of SciFinder search for $C_{21}H_{29}N$ sorted by #'s of associated references
- *4,342 Too many* to be useful

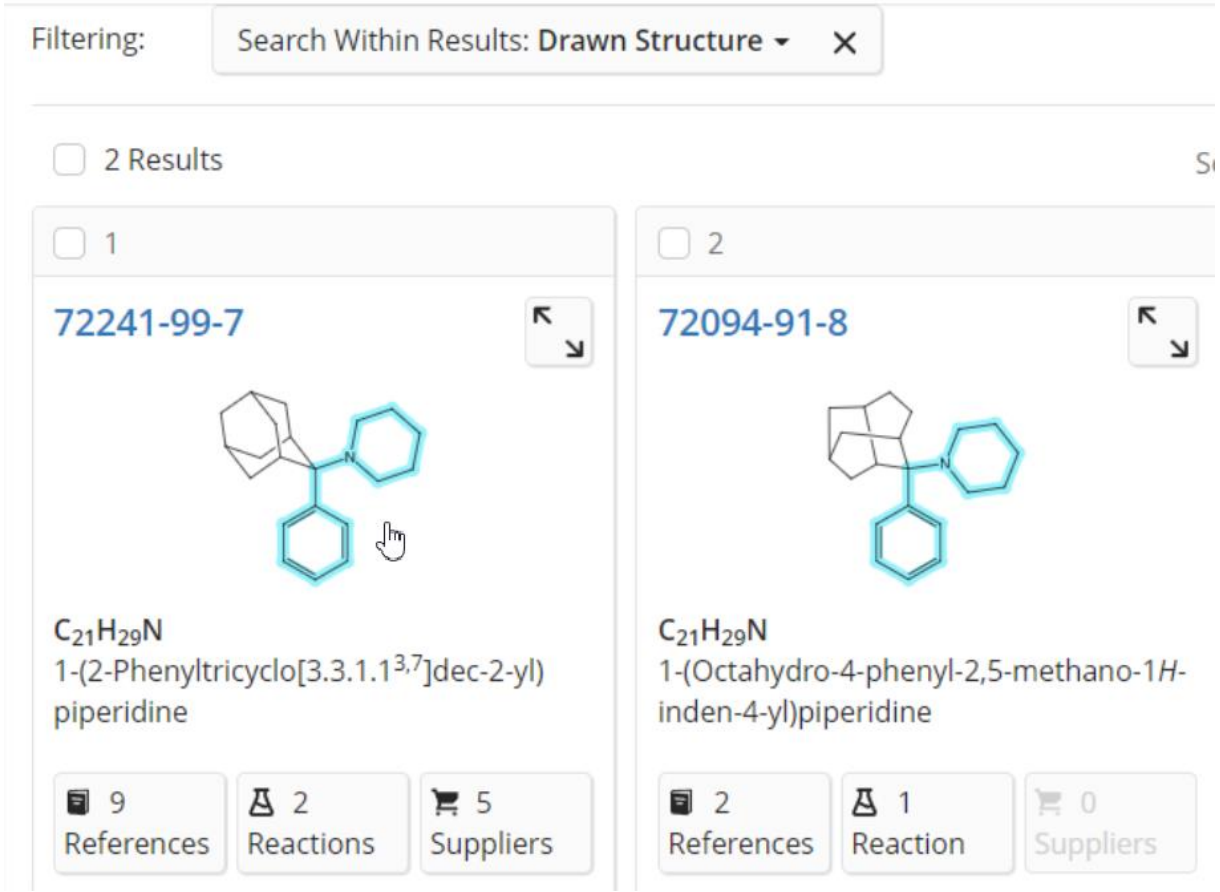


Refined Search by No Groups Attached to Piperidine Ring or Benzene in SciFinder (Adam Howard, Eastman)

- Approach *not* included in initial "known unknown" reference
- Allow no substitution on both rings by including hydrogens and a molecular formula of $C_{21}H_{29}N$
- Only 2 the of the 4,342 results had consistent structures
- Result #1 had 9 references and 4 suppliers, Result #2 had 2 references and 0 suppliers



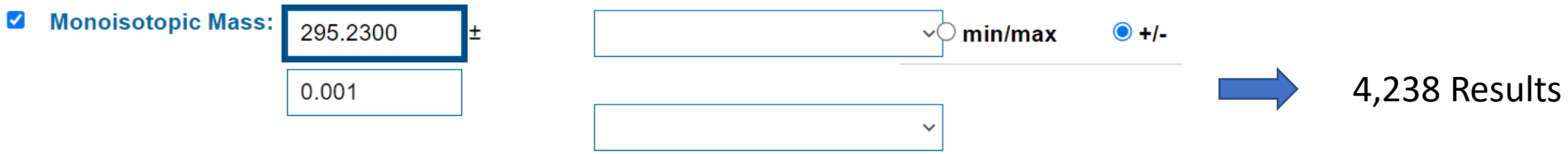




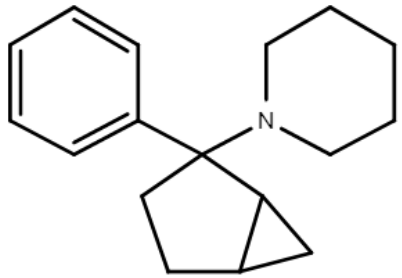
Alternate Approach Using ChemSpider Using Monoisotopic Mass and Similarity Structure Search

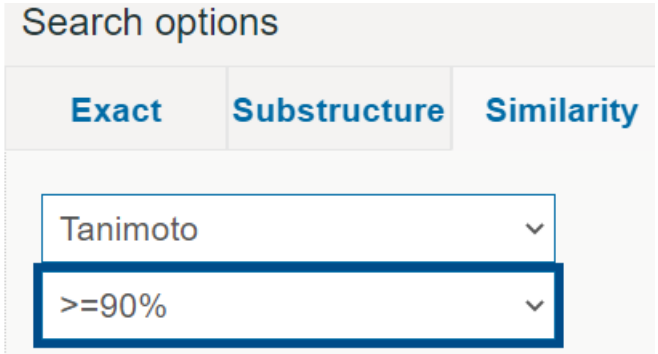
***Not* Included in Original ASMS Article**

Just Monoisotopic Mass Search with Window



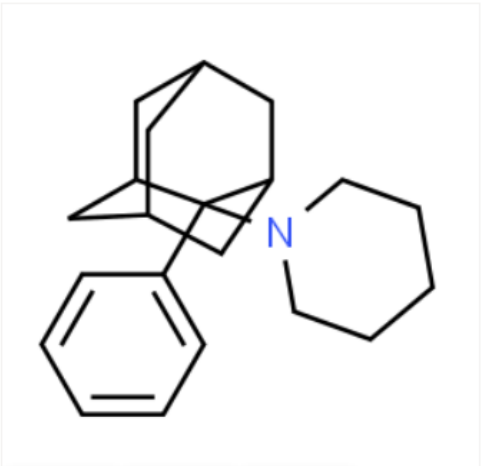
Both Monoisotopic Mass Search +/- Mass Error ***and*** Tanimoto (***Similarity***) Search





Found 1 result

Search term: **Structure Search - Similarity AND MM >= 295.229 AND MM <= 295.231 AND abs(Monoisotopic_Mass - 295.23) as mass_defect AND SingleComponent AND NonIsotopic**

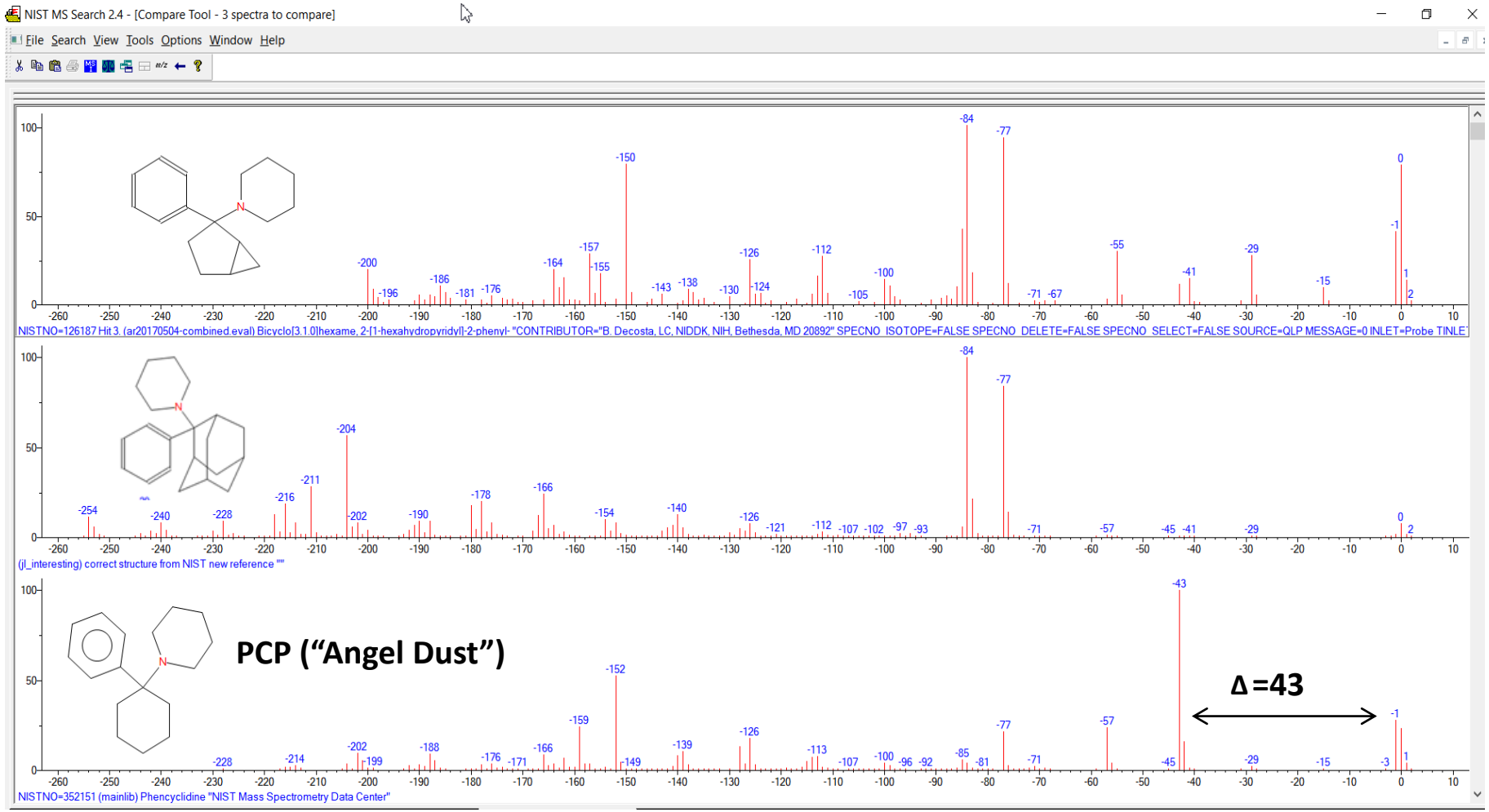


1-(2-Phenyladamantan-2-yl)piperidine

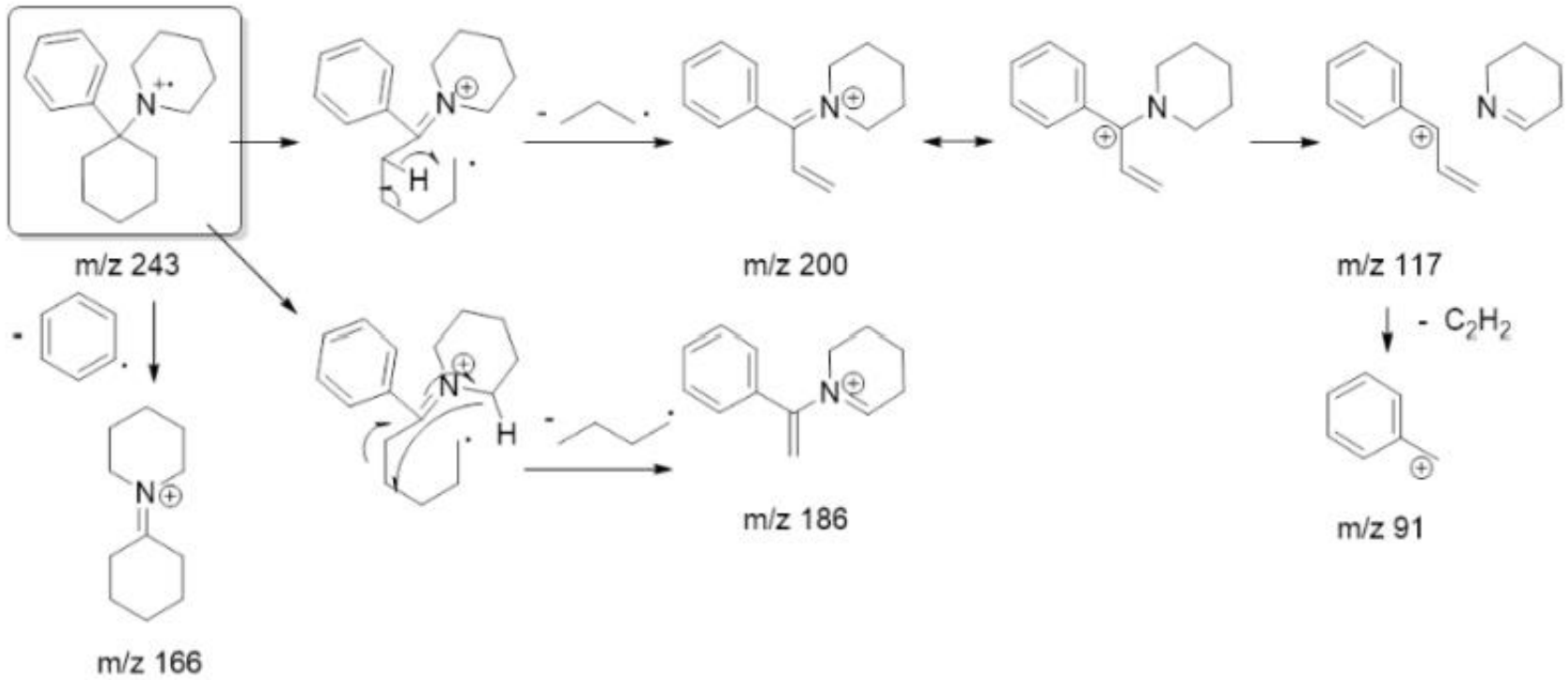
Molecular Formula	$C_{21}H_{29}N$
Average mass	295.462 Da
Monoisotopic mass	295.230011 Da
ChemSpider ID	58052

Addendum: Differences in EI Fragmentation of PCP-Related Compounds

- Major differences between neutral loss spectra of the three compounds below
- Thought the parent compound, PCP ("Angel Dust") would be more similar to other two
- The presence of fused cyclohexyl ring drives loss of 43, C_3H_7 shown in neutral loss spectra below
- Same mechanism for loss of 43 ***not*** accessible by other two compounds
- If best hit with MW of 241 not present, hybrid search ***would have failed*** to yield useful information



Explanation of 43 Loss in PCP EI Mass Spectrum (Martin Garraffo, NIST)



Conclusions

- Hybrid search is a very valuable addition to the identification process
- Extends the utility of all ***commercial*** and ***user*** EI and MSMS libraries
- Used in combinations with other approaches to identify unknowns
- Free detailed training for EI and MSMS hybrid searches

Internet Links to *Free* Resources in Workshop *Plus Others in* PDF Handout on Website

- NIST EI Search Software ***Free*** Training Course
- NIST MSMS Search Software ***Free*** Training Course
- Hybrid Search Delta Mass Table
- LCGC Review Article Eastman Approach to Identifications
- ChemSpider “Known Unknowns” (Spectra-less Database) Identifications
- CAS Registry “Known Unknowns” (Spectra-less Database) Identifications
- NMR Complimenting MS Searches for Unknown Identification and Quantitation
- Chemical Ionization (CI) GC-MS Resources
- NIST Literature Reference on Hybrid Search
- EI Fragmentation Mechanism for PCP-Related Species
- NMR Data for PCP-Related Unknown
- Trimethylsilyl Derivatives for GC-MS
- Methyl Ester Derivatives for GC-MS
- Differences in Standard EI and Orbitrap Spectra
- Helium Conversation Tips
- Simple Way to Monitor Lipid Matrix Effects in Biological Analyses
- Identification of Surfactants with LC-MS/MS
- Wiley KnowItAll Software Training Course for Identifications Using MS and IR

Contributors

- Stephen Stein (NIST)
- Dmitri Tchekhovskoi (NIST)
- H. Martin Garraffo (NIST)
- David Sparkman (NIST Contractor/Univ. of Pacific)
- Arun Moorthy (NIST)
- Ed Sisco (NIST)
- Aaron Urbas (NIST)
- Gary Mallard (NIST Contractor)
- Sandy Markey (NIST Contractor)
- Curt Cleven (Eastman)
- Adam Howard (Eastman)
- Laura Adduci (GSK, ***formerly*** Eastman)
- Antony Williams (***formerly*** ChemSpider, now EPA)
- Stacey Edwards (ETSU School of Pharmacy)

Any Questions?
(If not, will do real time demonstration of software)

A "Little" Mass Spec and Sailing
Organic Mass Spectrometry, NMR, Sailing, Tesla, Duplicate Bridge

Mass Spec Interpretation Services