NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

Welcome to the Official Release of the *NIST23* Mass Spectral Libraries

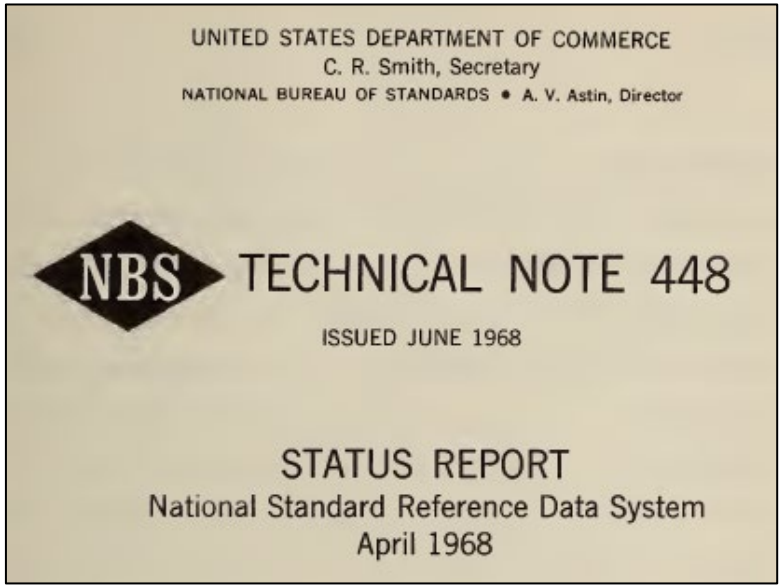
June 5, 2023

Mass Spectrometry Data Center
National Institute of Standards and Technology
Gaithersburg, Maryland

About the National Institute of Standards and Technology

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

- 1901: Founded as the National Bureau of Standards
- The national metrological institute (NMI) of the USA
- A non-regulatory agency of the US Department of Commerce
- Mission: To promote U.S. innovation and industrial competitiveness by advancing measurement science, standards, and technology in ways that enhance economic security and improve our quality of life
- Providing Standard Reference Data (SRD) one of the core functions of the Institute
- 1968: Congress passed the Standard Reference Data Act
- 2023: NIST currently provides more than 100 SRD products


UNITED STATES DEPARTMENT OF COMMERCE
C. R. Smith, Secretary
NATIONAL BUREAU OF STANDARDS • A. V. Astin, Director

NBS TECHNICAL NOTE 448

ISSUED JUNE 1968

STATUS REPORT
National Standard Reference Data System
April 1968

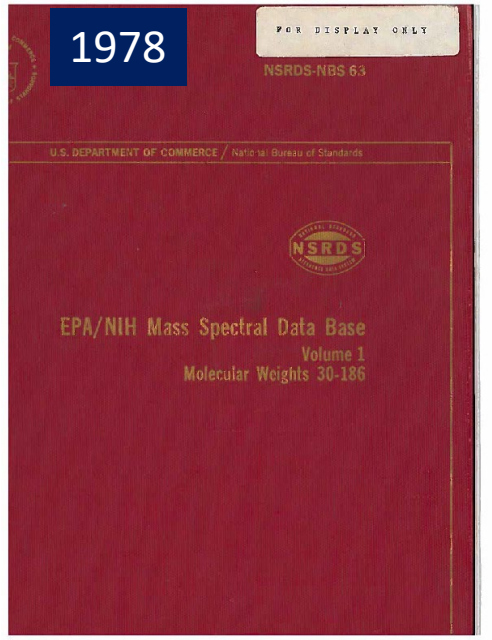
About the NIST Mass Spectrometry Data Center

National Institute of Standards and Technology
U.S. Department of Commerce

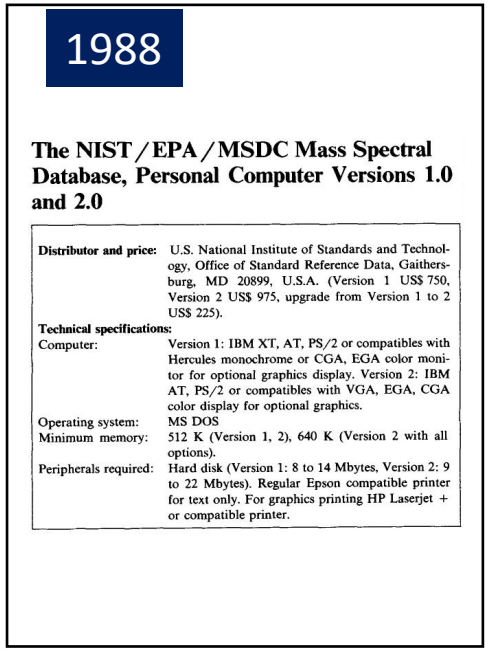
Mass Spectrometry Data Center

- 1945: NIST produced the first mass spectral reference spectra
- 1988: the NIH/EPA mass spectral library was handed over to NIST
- Since that time NIST has grown the library over seven-fold
- 2023: 12th release of the NIST/NIH/EPA library
- All measurements, evaluation, software development done in house
- Total staff of 37
- Divided evenly between chemists and programmers

1978

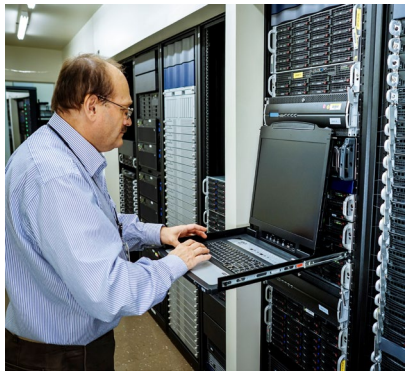


1988



The NIST / EPA / MSDC Mass Spectral Database, Personal Computer Versions 1.0 and 2.0

Distributor and price:	U.S. National Institute of Standards and Technology, Office of Standard Reference Data, Gaithersburg, MD 20899, U.S.A. (Version 1 US$ 750, Version 2 US$ 975, upgrade from Version 1 to 2 US$ 225).
Technical specifications:	
Computer:	Version 1: IBM XT, AT, PS/2 or compatibles with Hercules monochrome or CGA, EGA color monitor for optional graphics display. Version 2: IBM AT, PS/2 or compatibles with VGA, EGA, CGA color display for optional graphics.
Operating system:	MS DOS
Minimum memory:	512 K (Version 1, 2), 640 K (Version 2 with all options).
Peripherals required:	Hard disk (Version 1: 8 to 14 Mbytes, Version 2: 9 to 22 Mbytes). Regular Epson compatible printer for text only. For graphics printing HP Laserjet + or compatible printer.

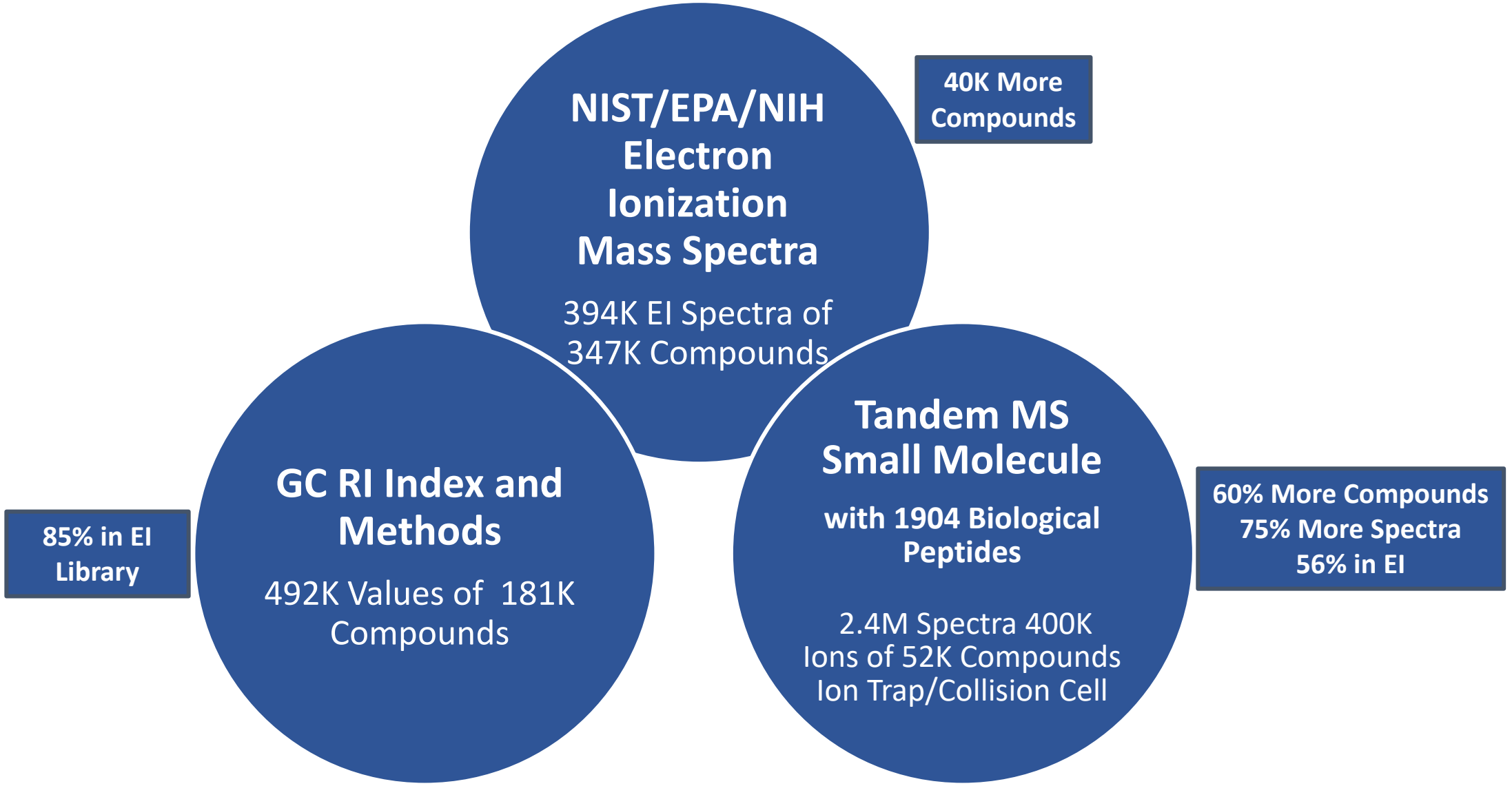


Triennial Lift Off

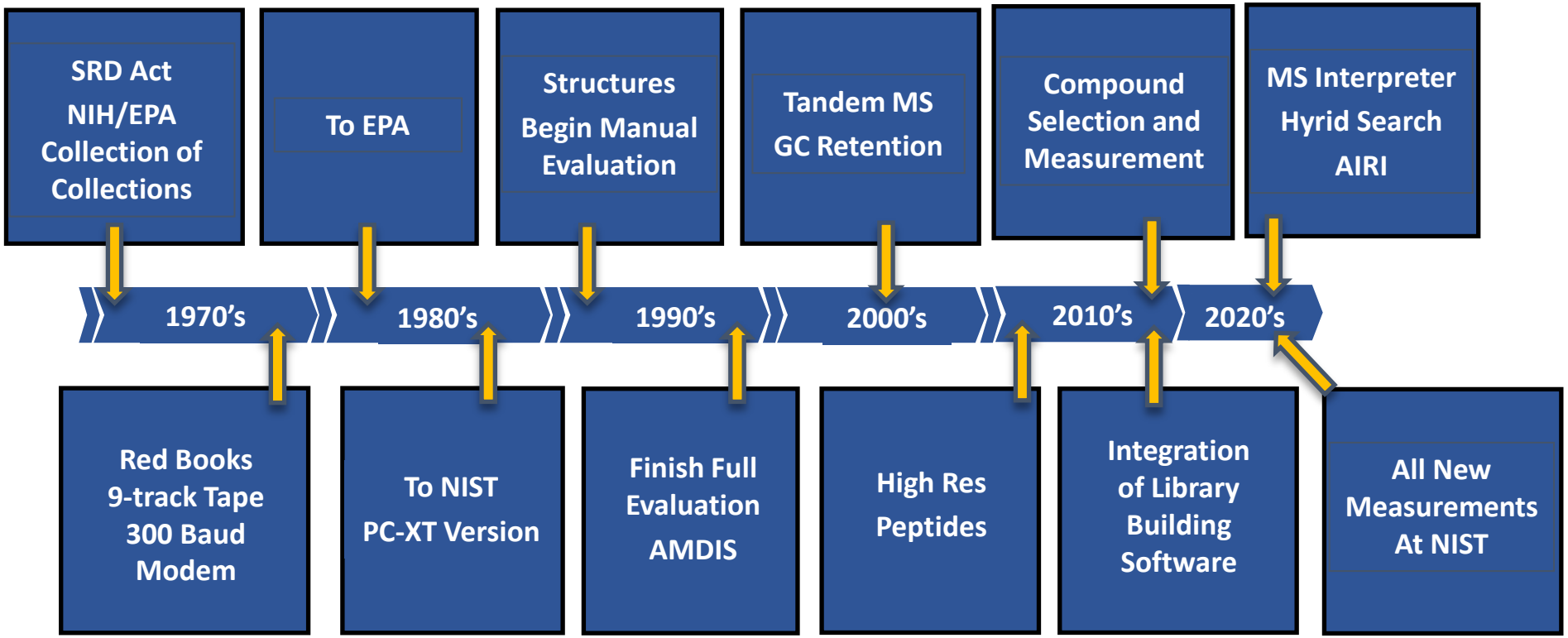
Presenters

Bill Wallace	Introduction
Steve Stein	Overview/Software
Tytus Mak	Selecting Compounds
Weihua Ji	EI Library
Yufang Zheng	Environmental/PFAS
Xiaoyu (Sara) Yang	Tandem Library
David Sparkman	Users

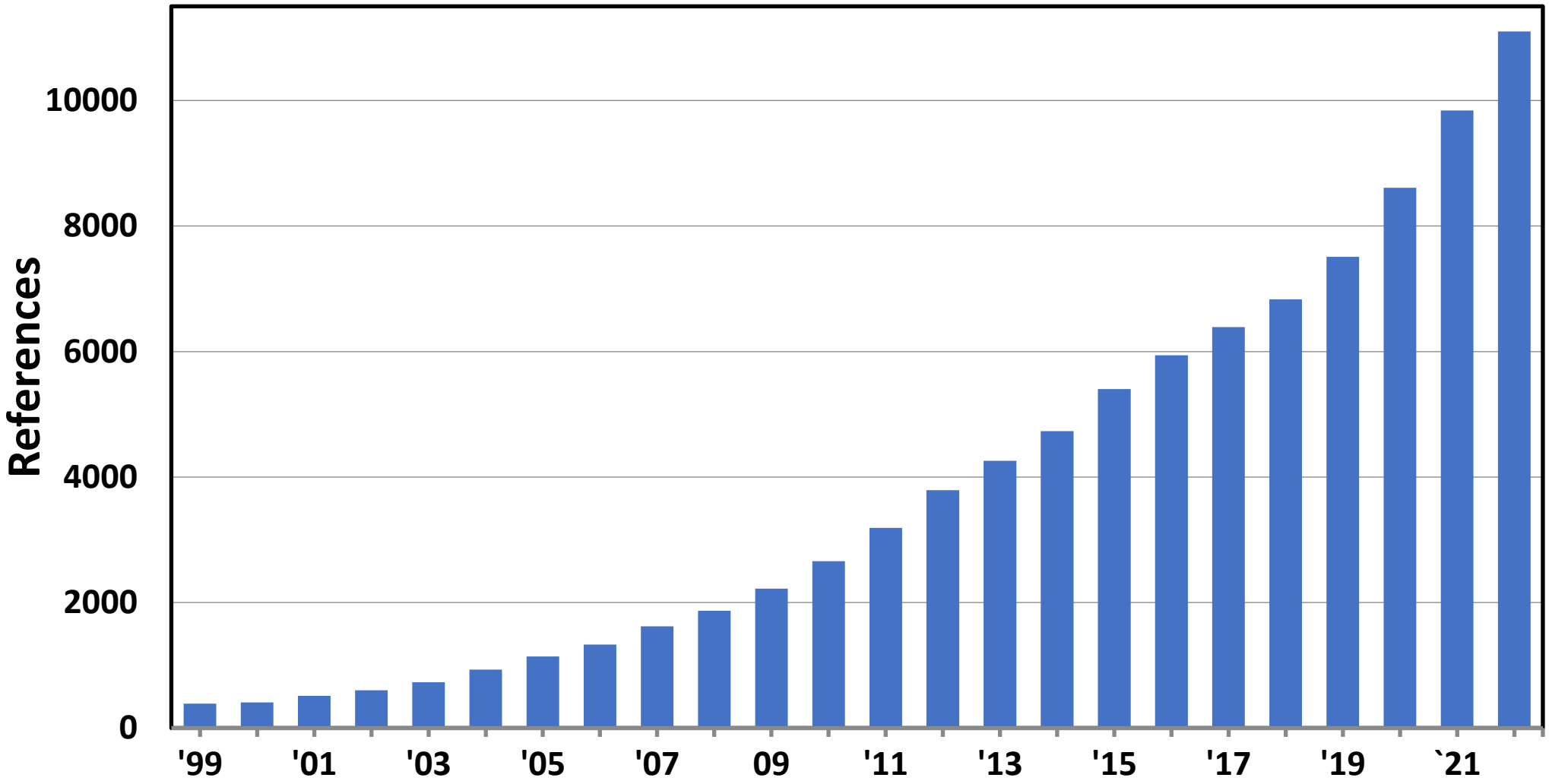
NIST 23 Mass Spectral Libraries



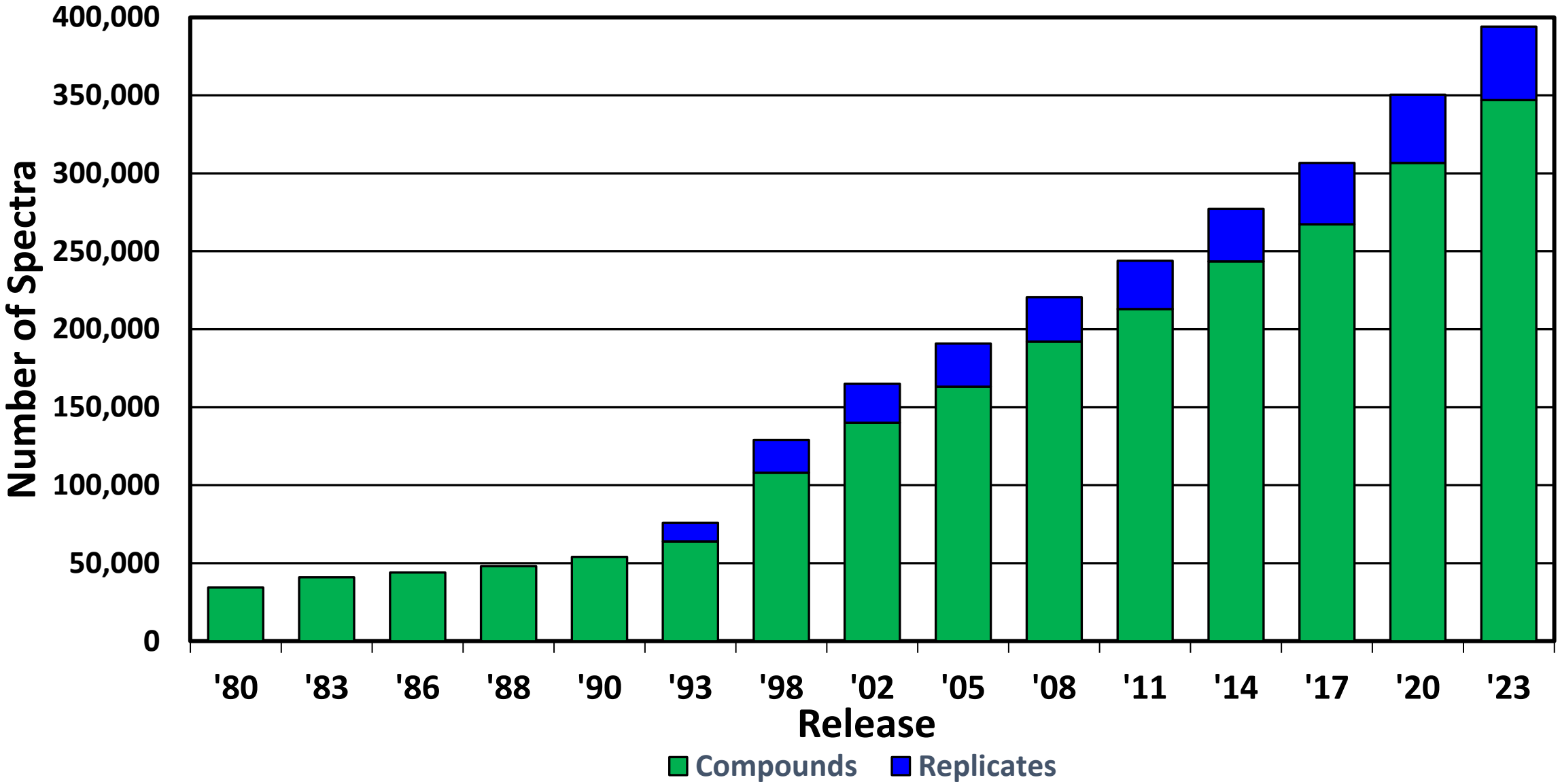
NIST/EPA/NIH EI MS Library Evolution



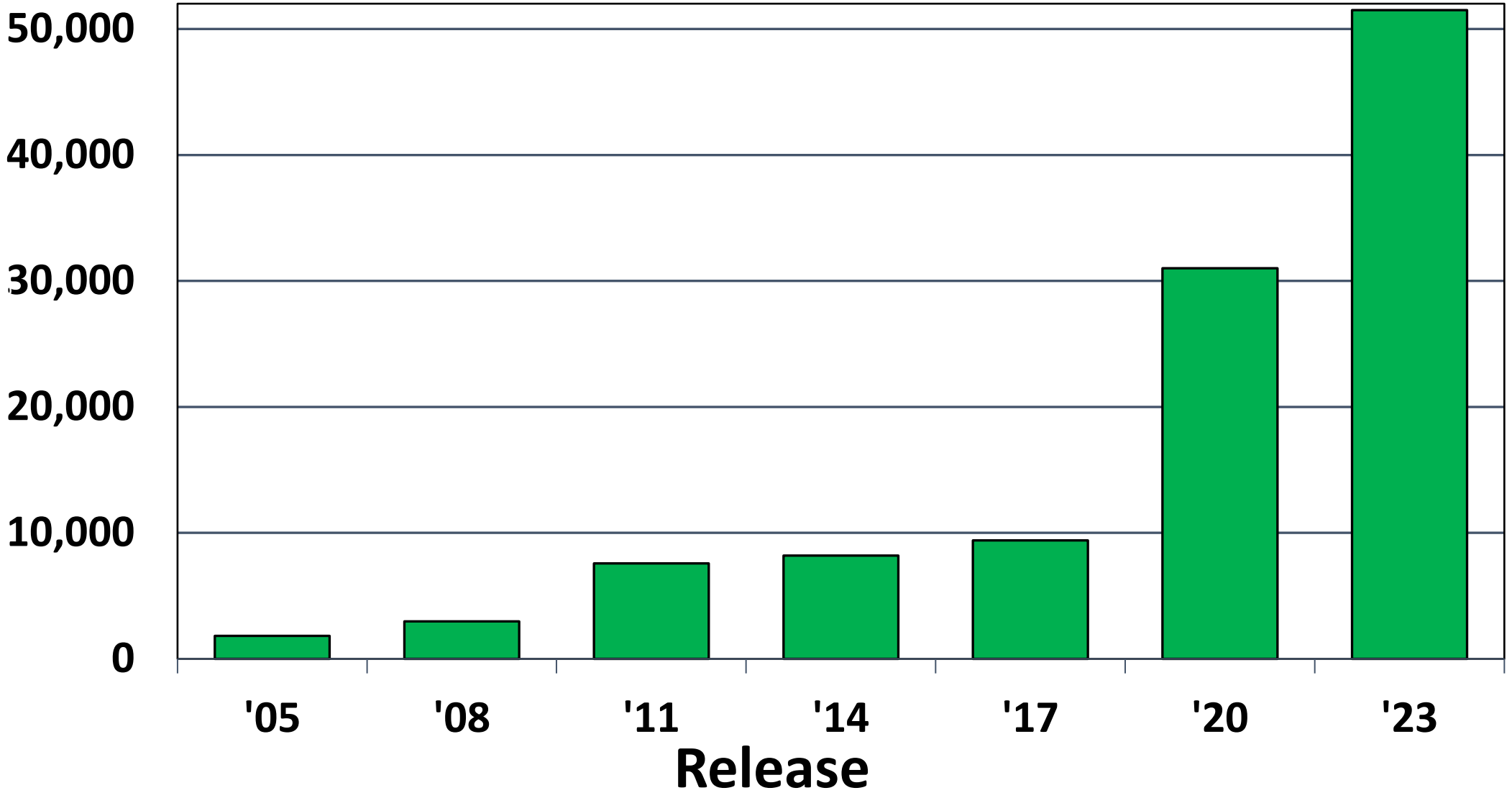
GC/MS References



EI Spectra over Time



Compounds in Tandem Library



How do we Select Compounds to Add to Our Libraries?

Tytus Mak

~~**What** new compounds are in NIST23?~~

How were new compounds selected for NIST23?

Why were these new compounds added to NIST23?

ZINC

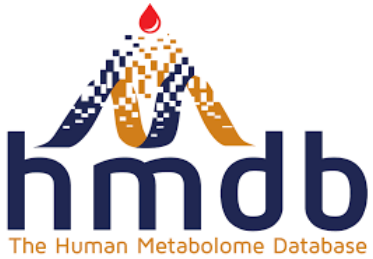
FOODB

DRUGBANK

FDA
Substances Added To Food

BioCyc
Genome Database Collection

EPA
Hazardous Substances Databank

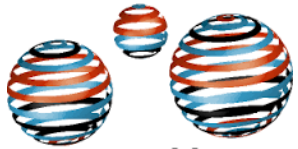


EPA
PARIS III

NutriChem

ChEBI

ChEMBL



npatlas

Blood Exposome Database

Automated Compound Selection Pipeline

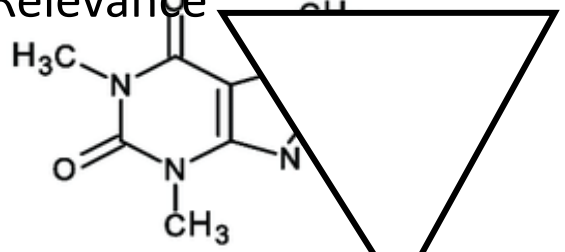
Merge via InChIKey

Meta-Compendium

500K Structures

Maximize Relevance

Most Compounds



Minimize Cost

Optimization Problem

InChI=1S/C8H10N4O2/c1-10-4-9-6-5(10)7(13)12(3)8(14)11(6)2/h4H,1-3H3 (caffeine)

character indicating the number of protons ('N' means neutral)

InChIKey=RYYVLZVUVIJVGH-UHFFFAOYSA-N

flag character for InChI version: 'A' for version 1

First block (14 letters)

Encodes molecular skeleton (connectivity)

Second block (8 letters)

Encodes stereochemistry and isotopes

flag character ('S') indicates standard InChIKey (produced out of standard InChI)

Purchasable?

Commercially Unavailable

Not Suitable

Coverage
Depends on Sample & Application

EI

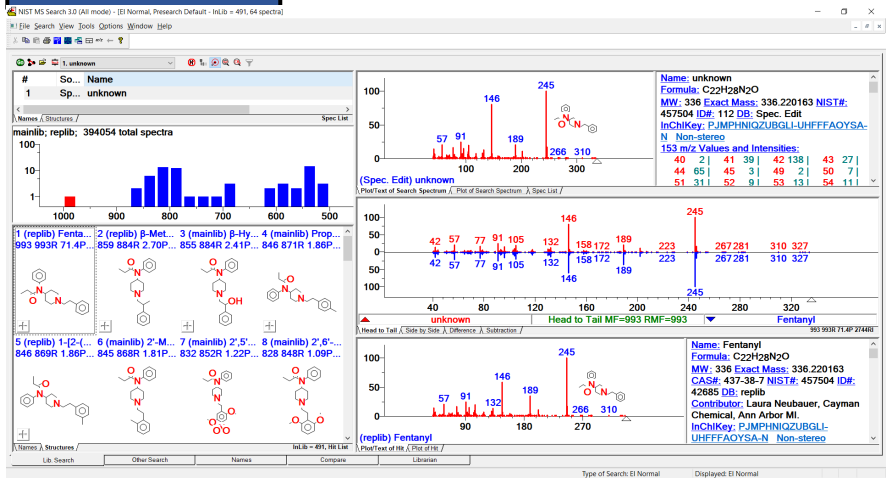
Citation	New	NIST23
Wikipedia	1570	6227
EPA Tox	2969	7117
Food DB	582	4273
EU Contaminants	6263	15149
Protein Data Bank	1194	4716
Human Metabolome DB	1992	9393
PFAS	161	749
Adams (Essential Oils)	2136	2136

Tandem

Citation	New	NIST23
Wikipedia	1618	6424
EPA Tox	3181	8146
Food DB	602	4491
EU Contaminants	6553	15818
Protein Data Bank	1246	4945
Human Metabolome DB	2071	9686
PFAS	90	116
Drug Bank	1130	3991

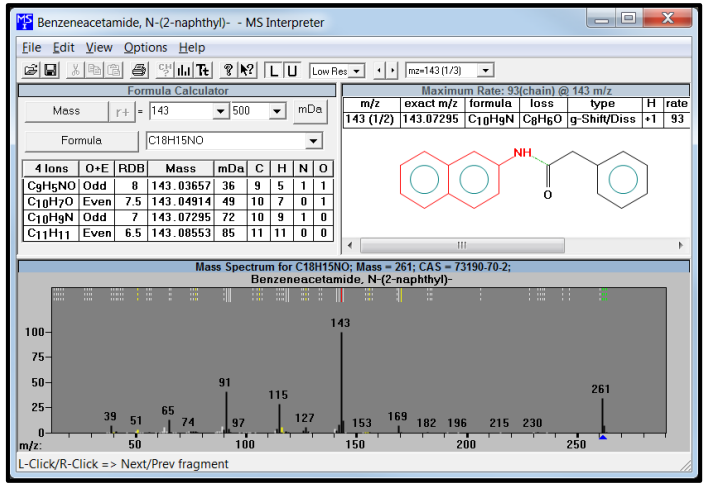
Software

MS Search



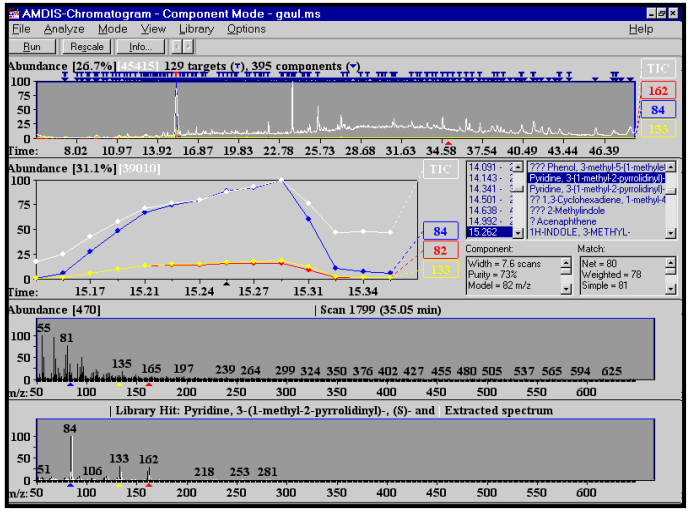
Callable by External Programs

MS Interpreter
EI or Tandem



API & MSPepsearch

AMDIS (gc/ms)

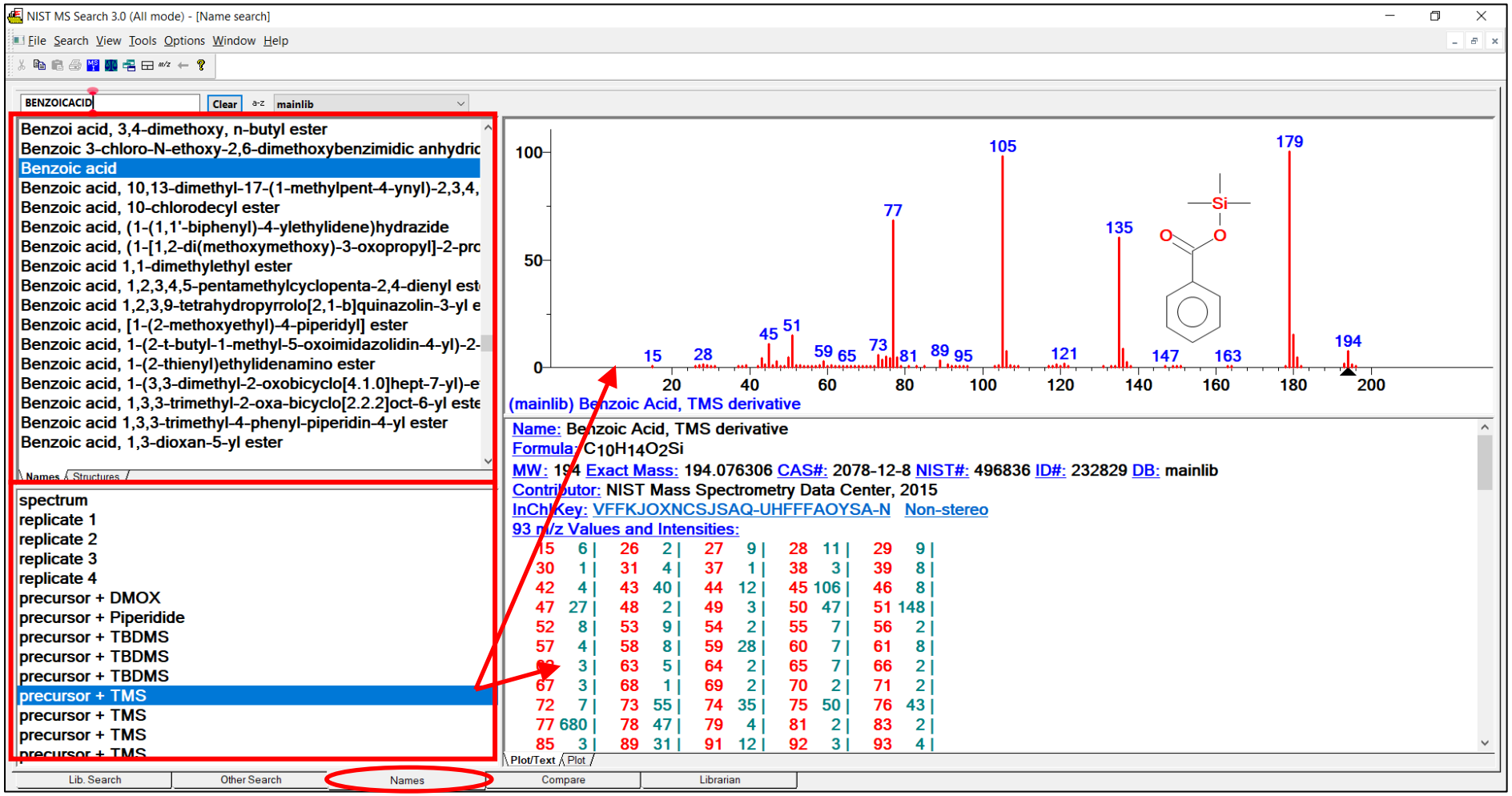


New Feature

Separate Names and Spectra - EI

Compound Name

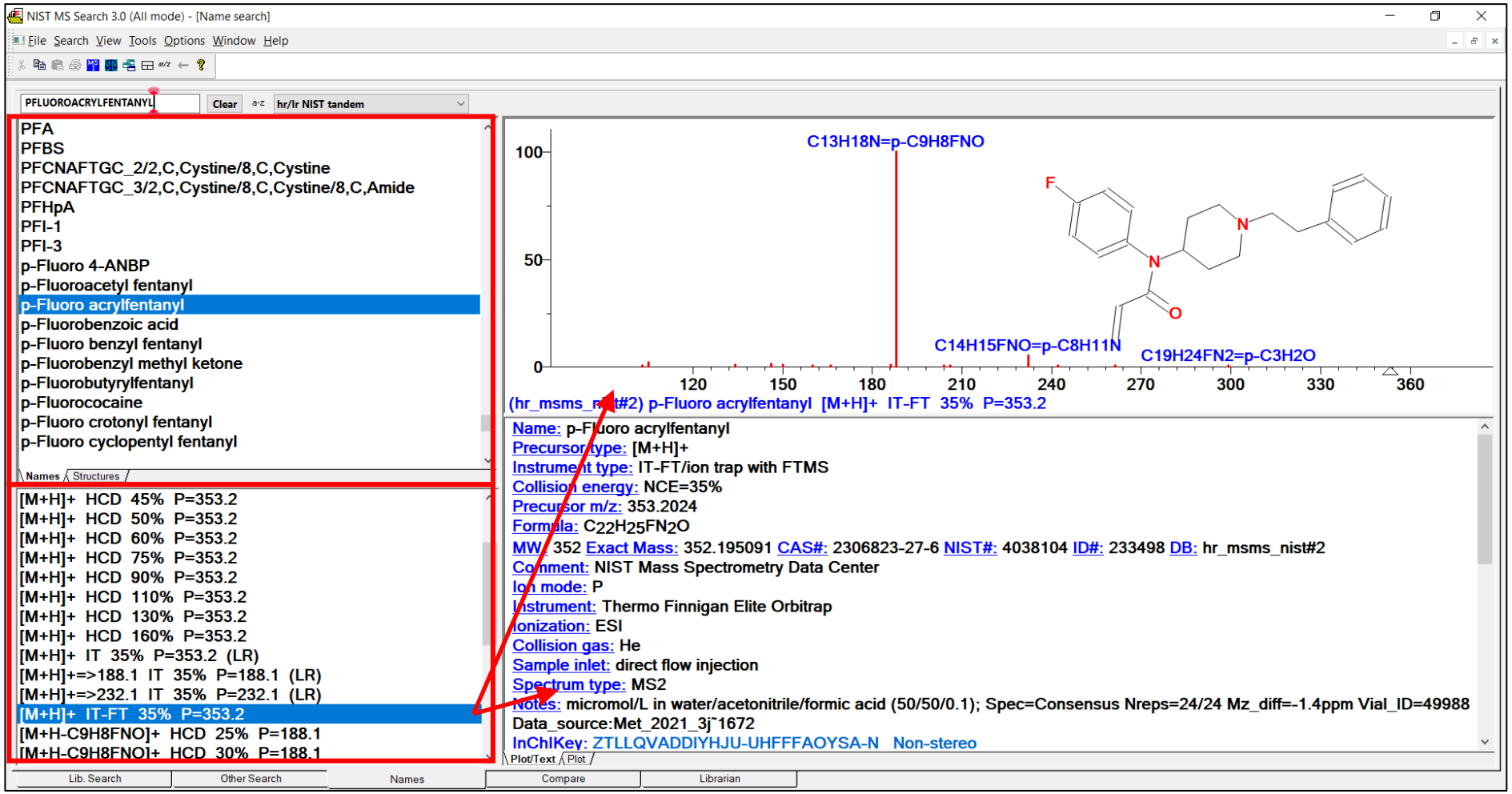
Spectra
Replicates
Derivatives
Stereiosomers



Separate Names and Spectra - Tandem

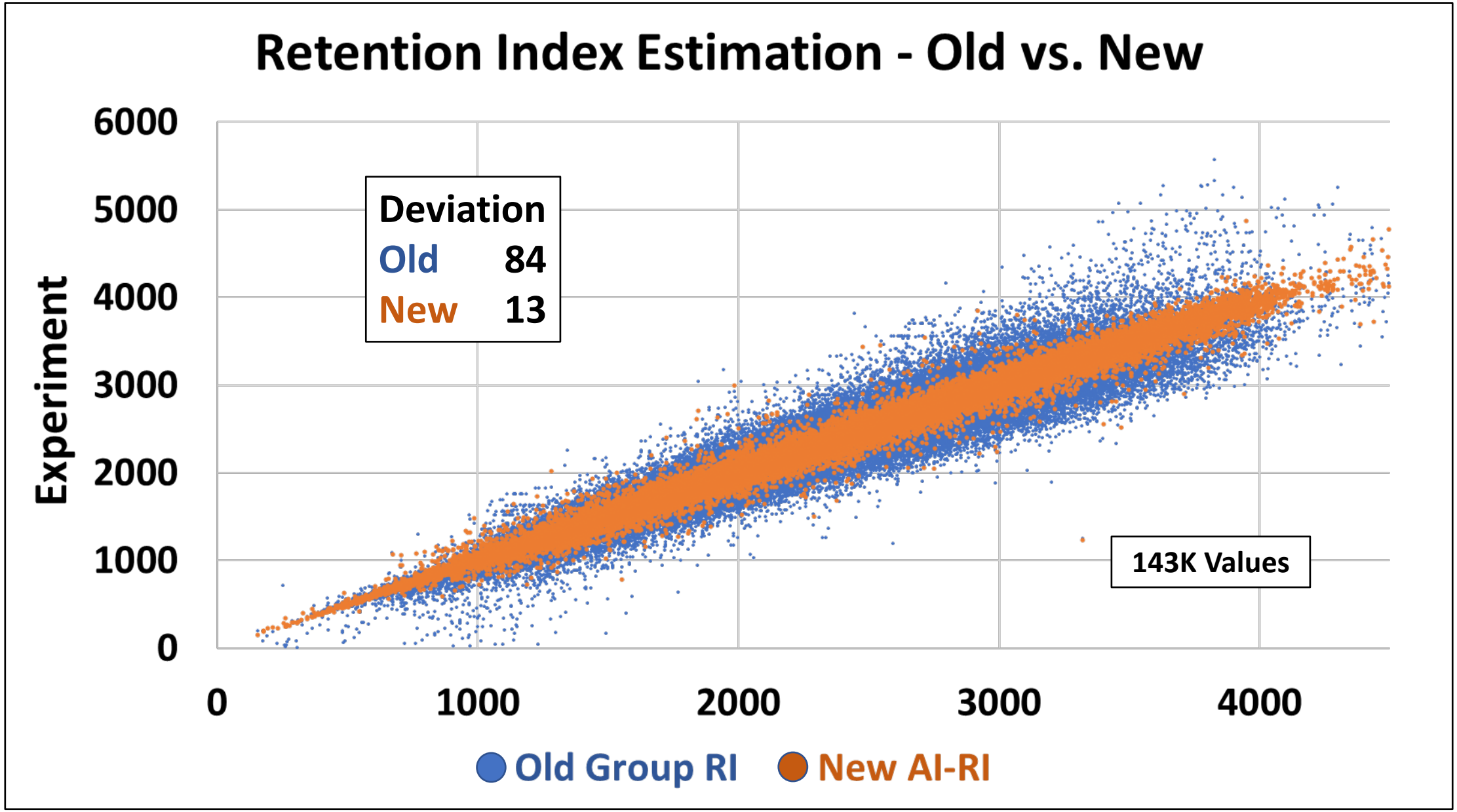
Compound Name

Spectra
Energies
Fragmentation
Resolution
Charge
In-Source



New non-Spectral Data
(distinguishes *isospectra*/compounds)

- **Reliable RI For All EI Library Compounds (40% => 100%)**
  - **AIRI Fills Gaps (Lewis Geer)**
  - **RI Can Be Used for Scoring**
- **'Other DB' Occurrences ('Prior Probability')**
  - **Non-Spectral Databases Citing Each Compound**
  - **61 DBs, Wide Applications (Environmental, Drug, Food, Metabolite, …)**
  - **EI 21%, Tandem 78%**


Retention Index Estimation - Old vs. New
Experiment
6000
5000
4000
3000
2000
1000
0
0
1000
2000
3000
4000
Deviation
Old 84
New 13
143K Values
Old Group RI
New AI-RI

New Search Type

Isomer ID by RI, Confirmation by 'Other DBs'

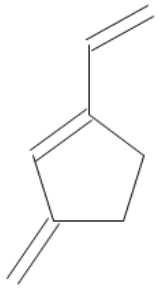
Search With o-Xylene Spectrum

No RI

#	Lib	MF	RMF	Prob	RI	DBs	Name
⊞ 1	M	941	942	29.7	865	32 ...	p-Xylene
⊞ 2	M	937	939	26.5	866	32 ...	Benzene, 1,3-dimethyl-
⊞ 3	M	924	925	18.0	888	36 ...	o-Xylene
⊞ 4	M	888	895	6.09	855	35 ...	Ethylbenzene
⊞ 5	M	878	881	4.55	121...		Benzeneethanol, α,β-dimethyl-

Use RI

#	Lib	MF	RMF	Prob	RI	origMF	DBs	Name
⊞ 1	M	924	925	90.6	888	924	36 ...	o-Xylene
2	M	826	877	6.70	878*	876		Cyclopentene, 1-ethenyl-3-methylene-
⊞ 3	M	767	939	1.22	866	937	32 ...	Benzene, 1,3-dimethyl-
⊞ 4	M	761	942	1.03	865	941	32 ...	p-Xylene
⊞ 5	M	699	750	0.17	898	749	1 G	1,3,5-Cyclooctatriene



Hybrid Search

For compounds not in the Library

NIST

Non-NIST

The Hybrid Search: A Mass Spectral Library Search Method for Discovery of Modifications in Proteomics

Meghan C. Burke,*[,†] Yuri A. Mirokhin,[†] Dmitrii V. Tchekhovskoi,[†] Sanford P. Markey,[†] Jenny Heidbrink Thompson,[‡] Christopher Larkin,[‡] and Stephen E. Stein[†]

Combining Fragment-Ion and Neutral-Loss Matching during Mass Spectral Library Searching: A New General Purpose Algorithm Applicable to Illicit Drug Identification

Arun S. Moorthy,*[,†] William E. Wallace,[†] Anthony J. Kearsley,[‡] Dmitrii V. Tchekhovskoi,[†] and Stephen E. Stein[†]

Hybrid Search: A Method for Identifying Metabolites Absent from Tandem Mass Spectrometry Libraries

Brian T. Cooper,*[,†,‡] Xinjian Yan,[‡] Yamil Simón-Manso,[‡] Dmitrii V. Tchekhovskoi,[‡] Yuri A. Mirokhin,[‡] and Stephen E. Stein[‡]

Increasing the Coverage of a Mass Spectral Library of Milk Oligosaccharides Using a Hybrid-Search-Based Bootstrapping Method and Milks from a Wide Variety of Mammals

Concepcion Africano Remoroza,* Yuxue Liang, Tytus D. Mak, Yuri Mirokhin, Sergey L. Sheetlin, Xiaoyu Yang, Joice V. San Andres, Michael L. Power, and Stephen E. Stein

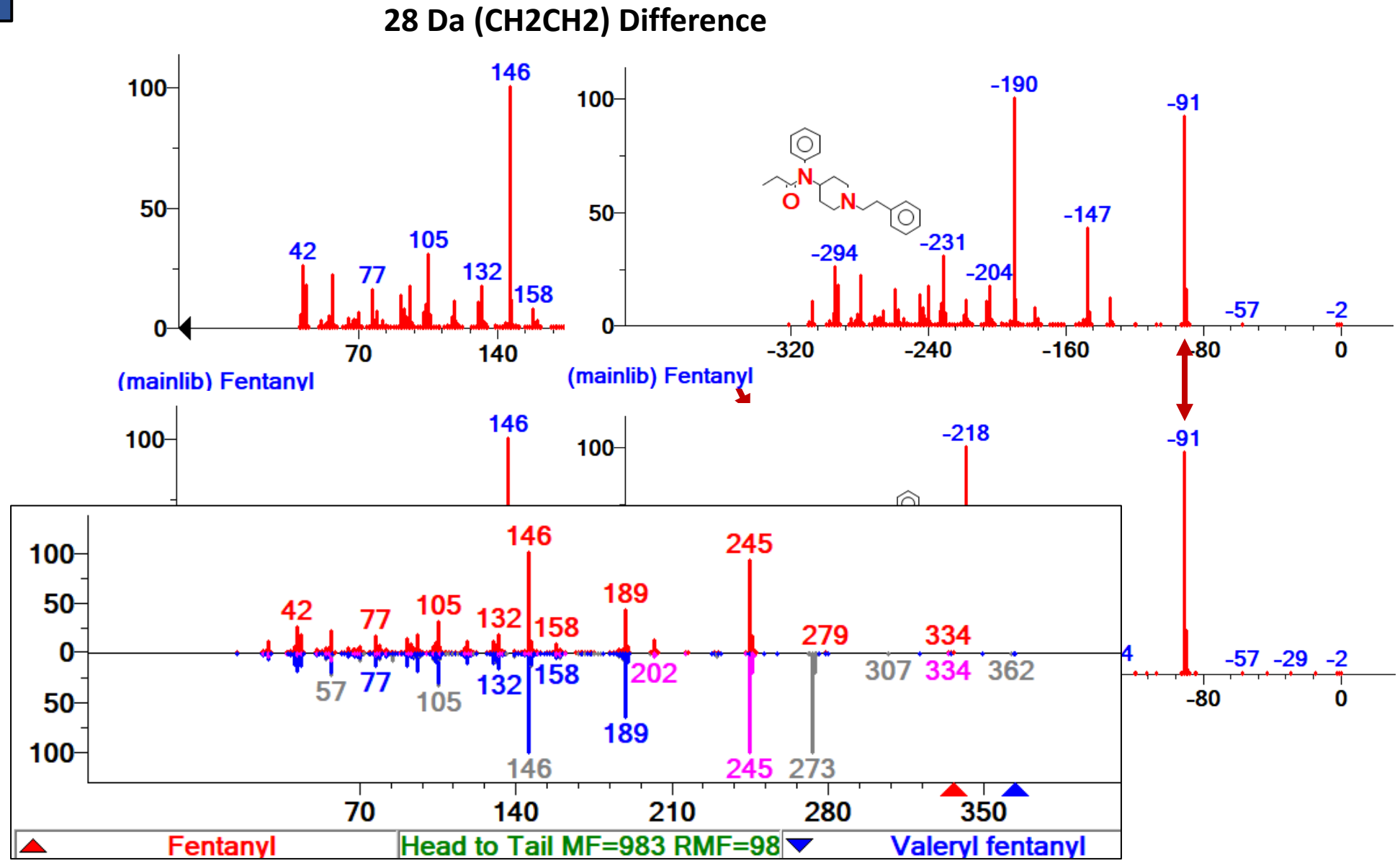
Structure Annotation of All Mass Spectra in Untargeted Metabolomics

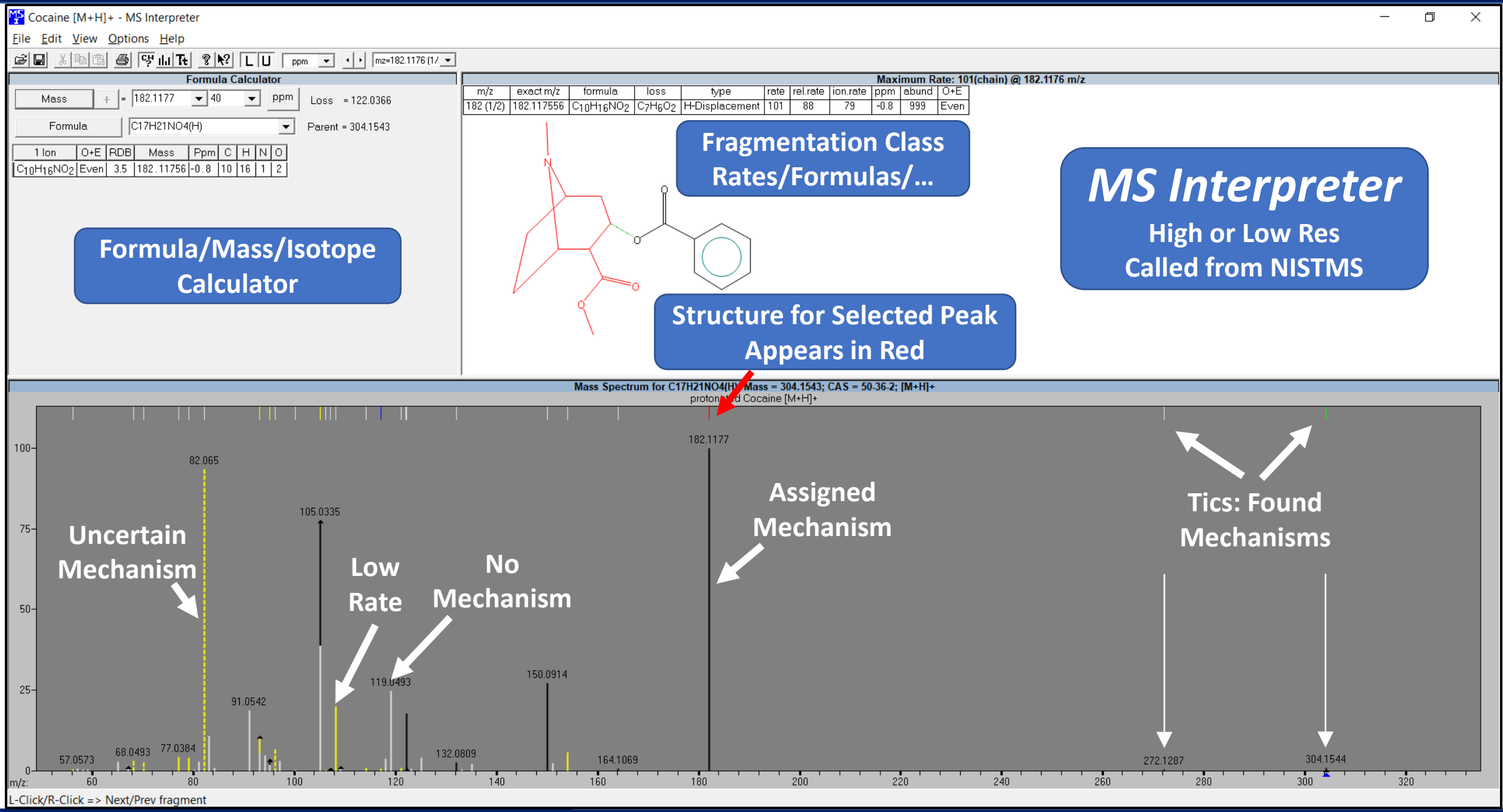
Ivana Blaženović,[†] Tobias Kind,[†] Michael R. Sa,[†] Jian Ji,[‡] Arpana Vaniya,[†] Benjamin Wancewicz,[†] Bryan S. Roberts,[†] Hrvoje Torbašinović,[§] Tack Lee,[∥] Sajjan S. Mehta,[†] Megan R. Showalter,[†] Hosook Song,[∥] Jessica Kwok,[†] Dieter Jahn,[⊥,#] Jayoung Kim,[∇,○,◆,¶] and Oliver Fiehn*[,†]

LC–MS/MS Software for Screening Unknown Erectile Dysfunction Drugs and Analogues: Artificial Neural Network Classification, Peak-Count Scoring, Simple Similarity Search, and Hybrid Similarity Search Algorithms

Inae Jang,[†] Jae-ung Lee,[†] Jung-min Lee,[†] Beom Hee Kim,[‡] Bongjin Moon,[†] Jongki Hong,*[,‡]

Hybrid Search

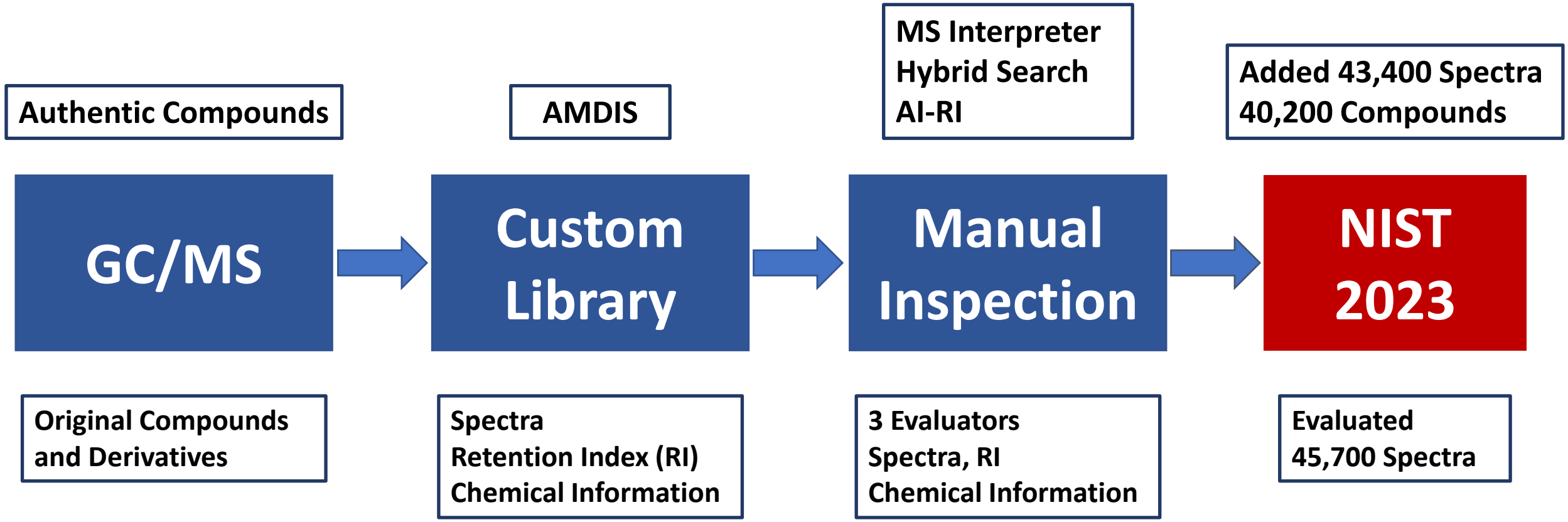



Cocaine [M+H]+ - MS Interpreter
File Edit View Options Help
Formula Calculator
Mass = 182.1177 40 ppm Loss = 122.0366
Formula C17H21NO4(H) Parent = 304.1543
1 Ion | O+E | RDB | Mass | Ppm | C | H | N | O
C10H16NO2 | Even | 3.5 | 182.11756 | -0.8 | 10 | 16 | 1 | 2
Formula/Mass/Isotope Calculator
Maximum Rate: 101(chain) @ 182.1176 m/z
m/z | exact m/z | formula | loss | type | rate | rel.rate | ion.rate | ppm | abund | O+E
182 (1/2) | 182.117556 | C10H16NO2 | C7H6O2 | H-Displacement | 101 | 88 | 79 | -0.8 | 999 | Even
Fragmentation Class Rates/Formulas/...
Structure for Selected Peak Appears in Red
MS Interpreter
High or Low Res
Called from NISTMS
Mass Spectrum for C17H21NO4(H) Mass = 304.1543; CAS = 50-36-2; [M+H]+
protonated Cocaine [M+H]+
182.1177
82.065
105.0335
91.0542
119.0493
150.0914
57.0573
68.0493
77.0384
132.0809
164.1069
272.1287
304.1544
Uncertain Mechanism
Low Rate
No Mechanism
Assigned Mechanism
Tics: Found Mechanisms
m/z:
L-Click/R-Click => Next/Prev fragment

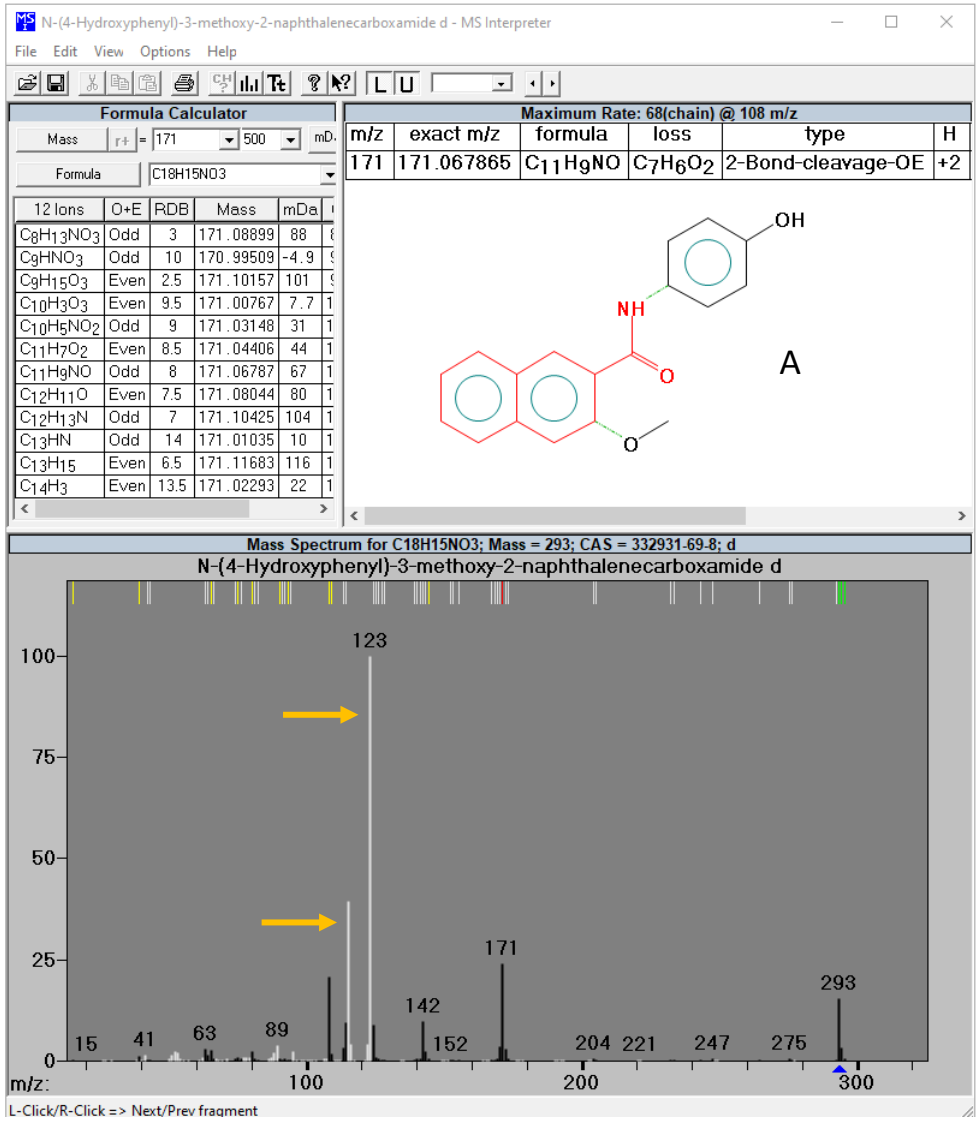
NIST/EPA/NIH EI Library: Evaluation

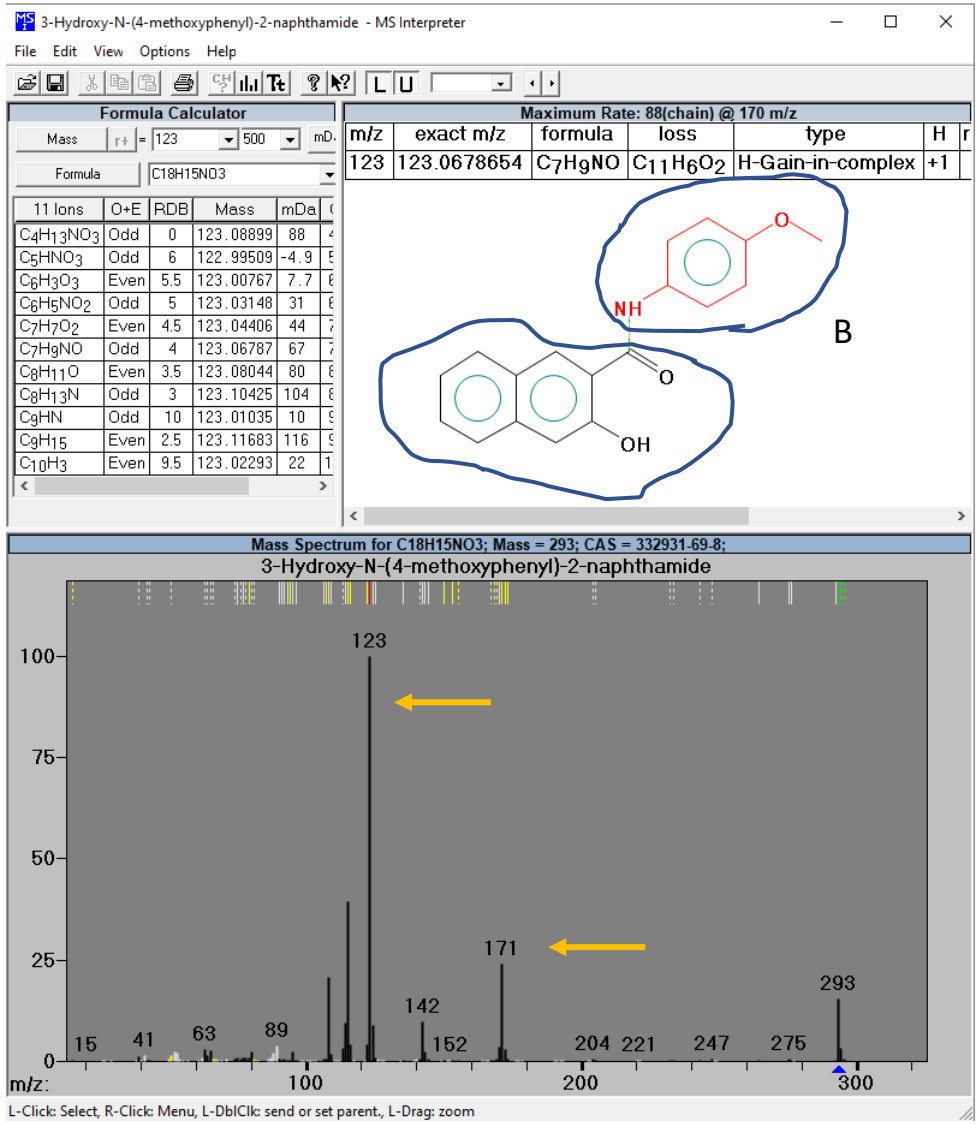
Weihua Ji

Expanding the EI MS Library

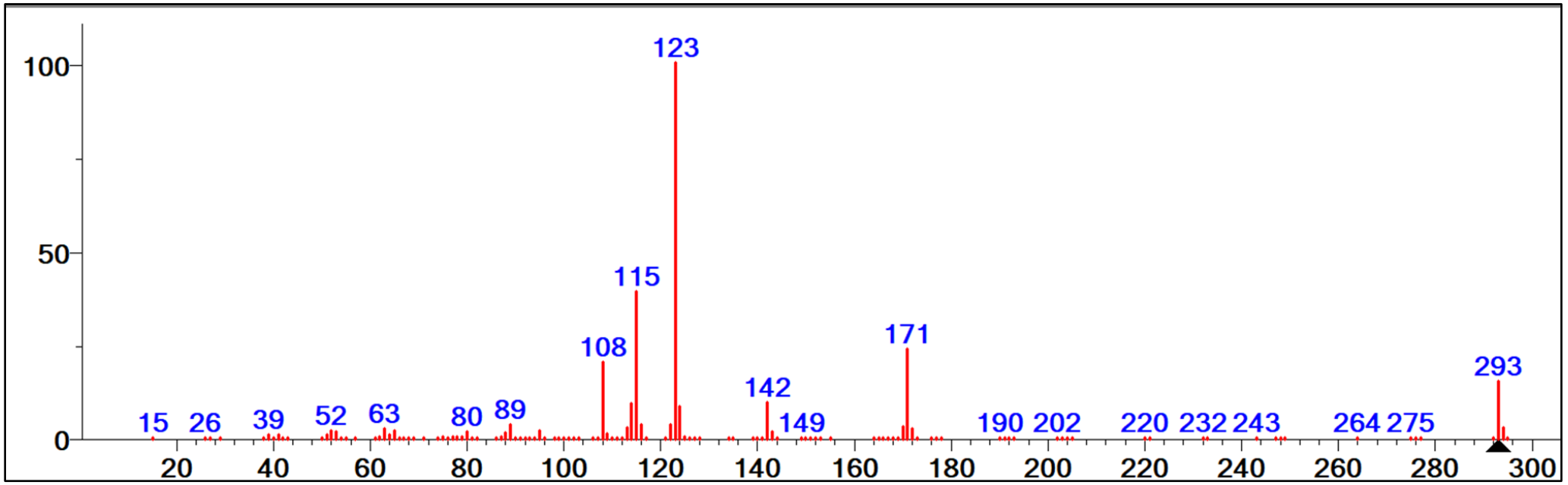


Evaluating Spectra with MS Interpreter



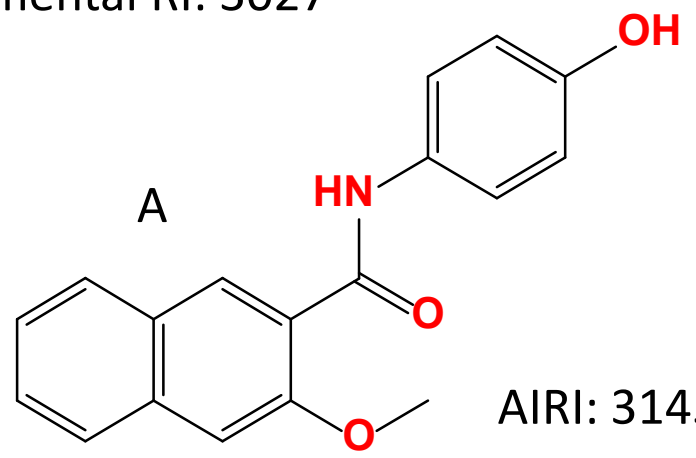


RI and AIRI Assist Spectrum Evaluation



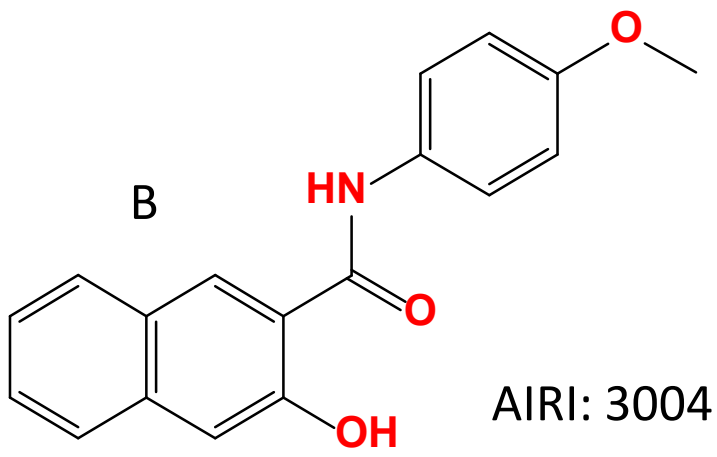
Experimental RI: 3027

A



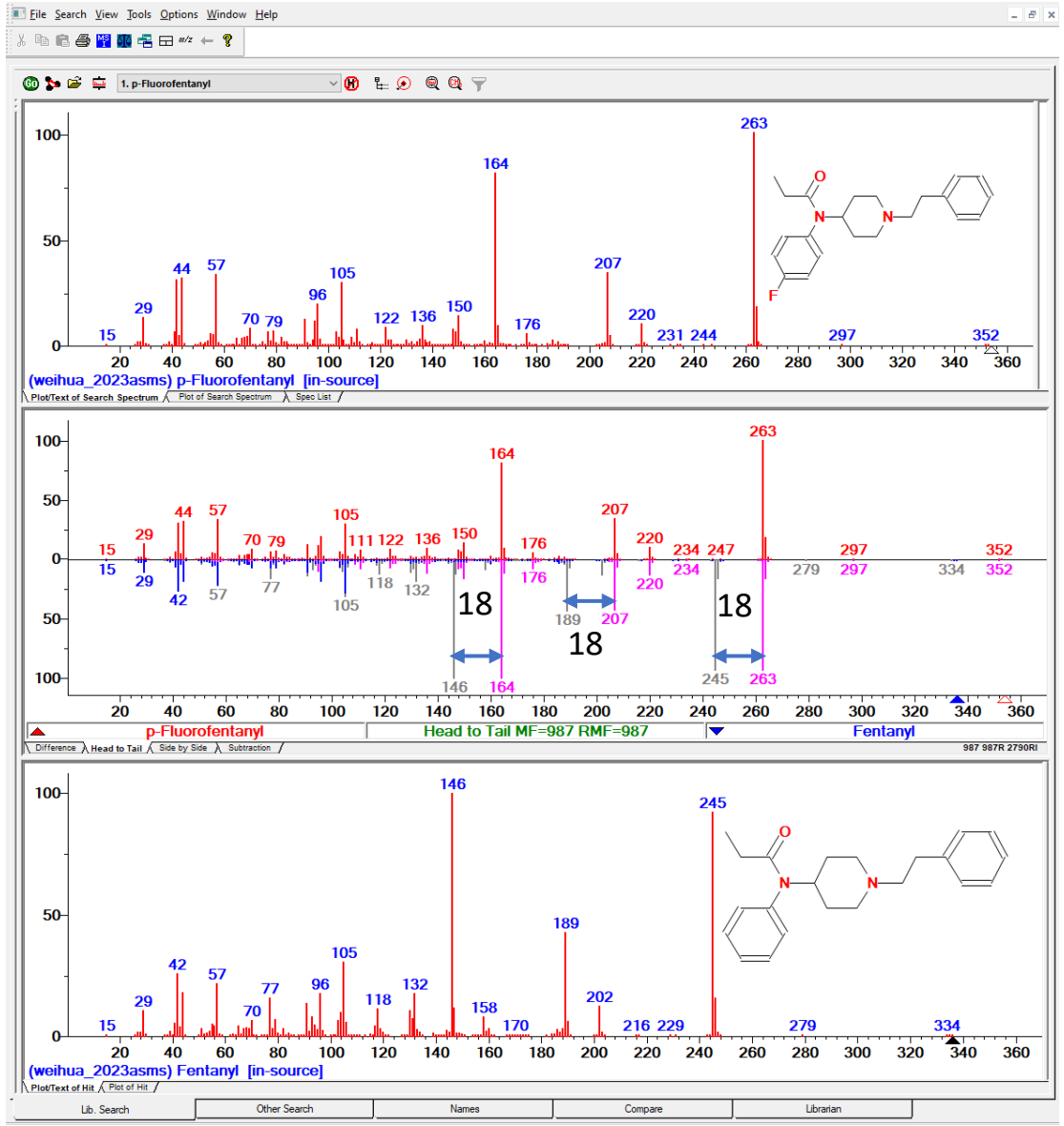
AIRI: 3145

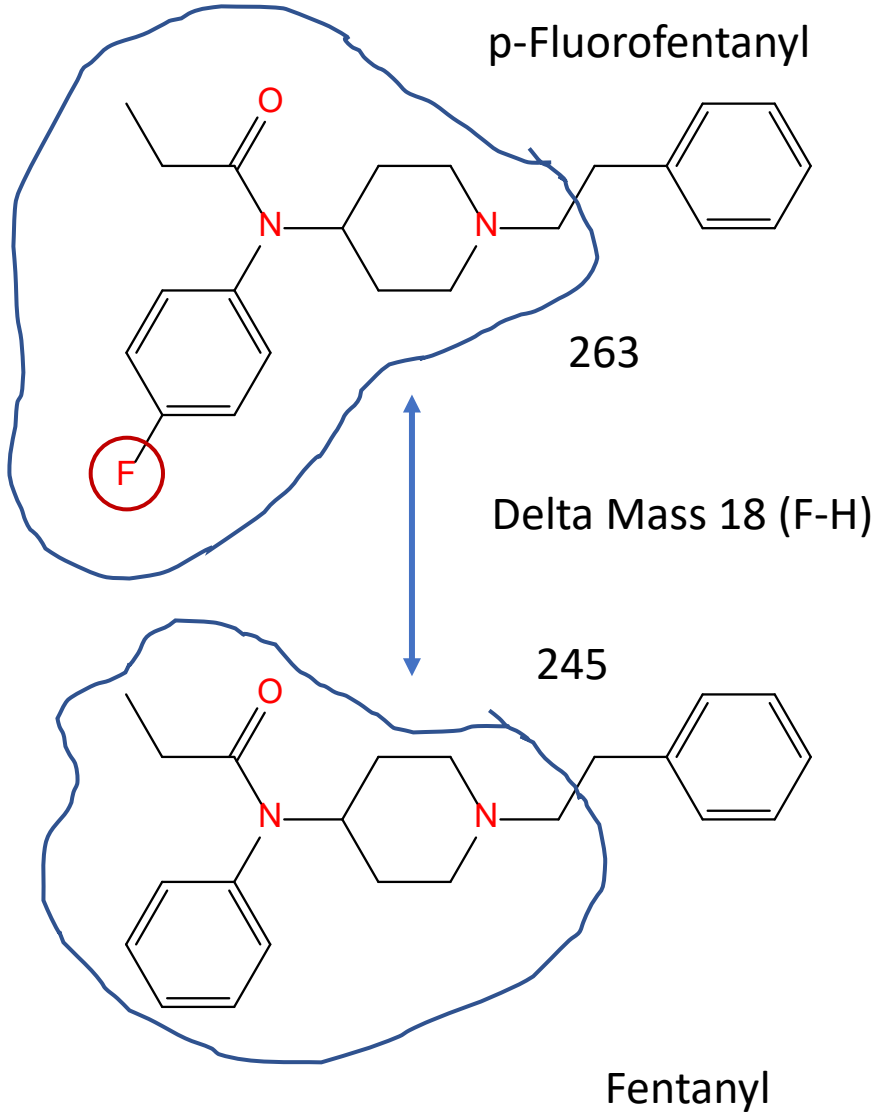
B



AIRI: 3004

Evaluating Spectra with the Hybrid Search





PFAS & Environmental

Yufang Zheng

PFAS Selection

Over 9,000 Compounds from PFAS Master List of EPA

Remove polymers and salts

6,695 Compounds of Interest

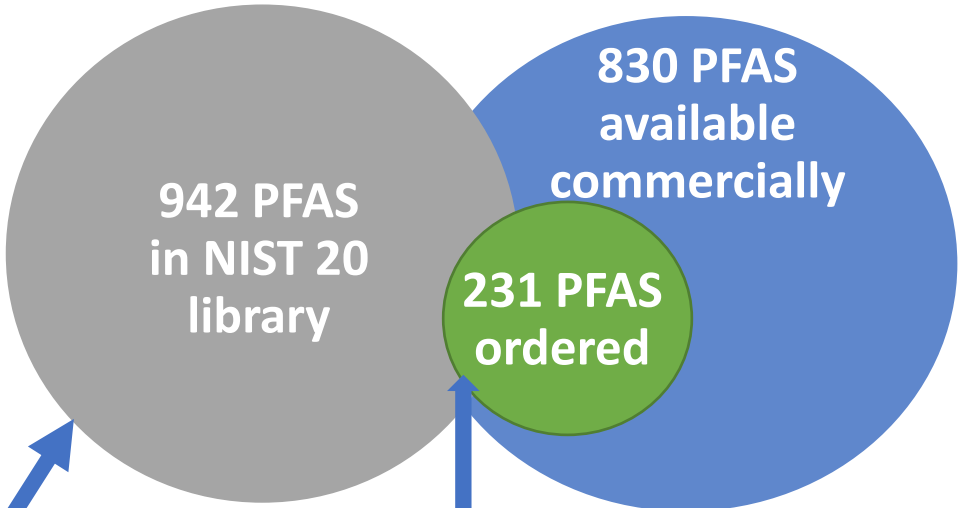
Find structures not in NIST 20

Find commercial sources

Generate a list of 830 PFAS

Order Compounds

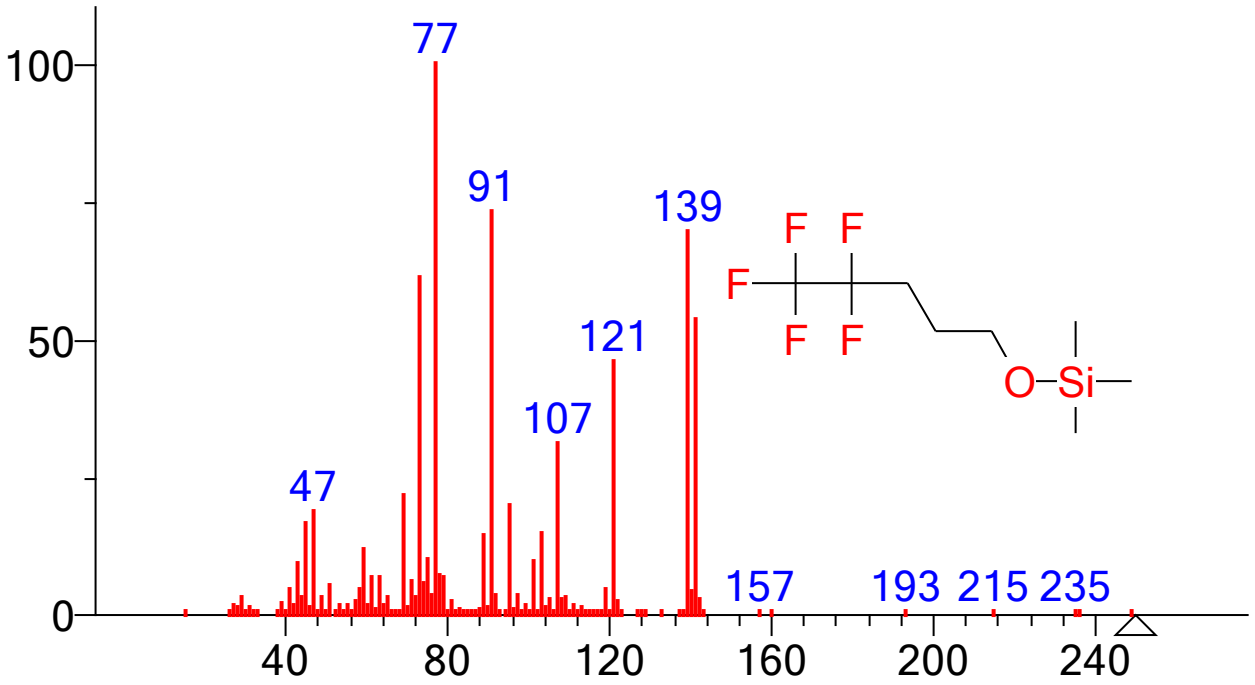
1st Batch of 231 PFAS Obtained



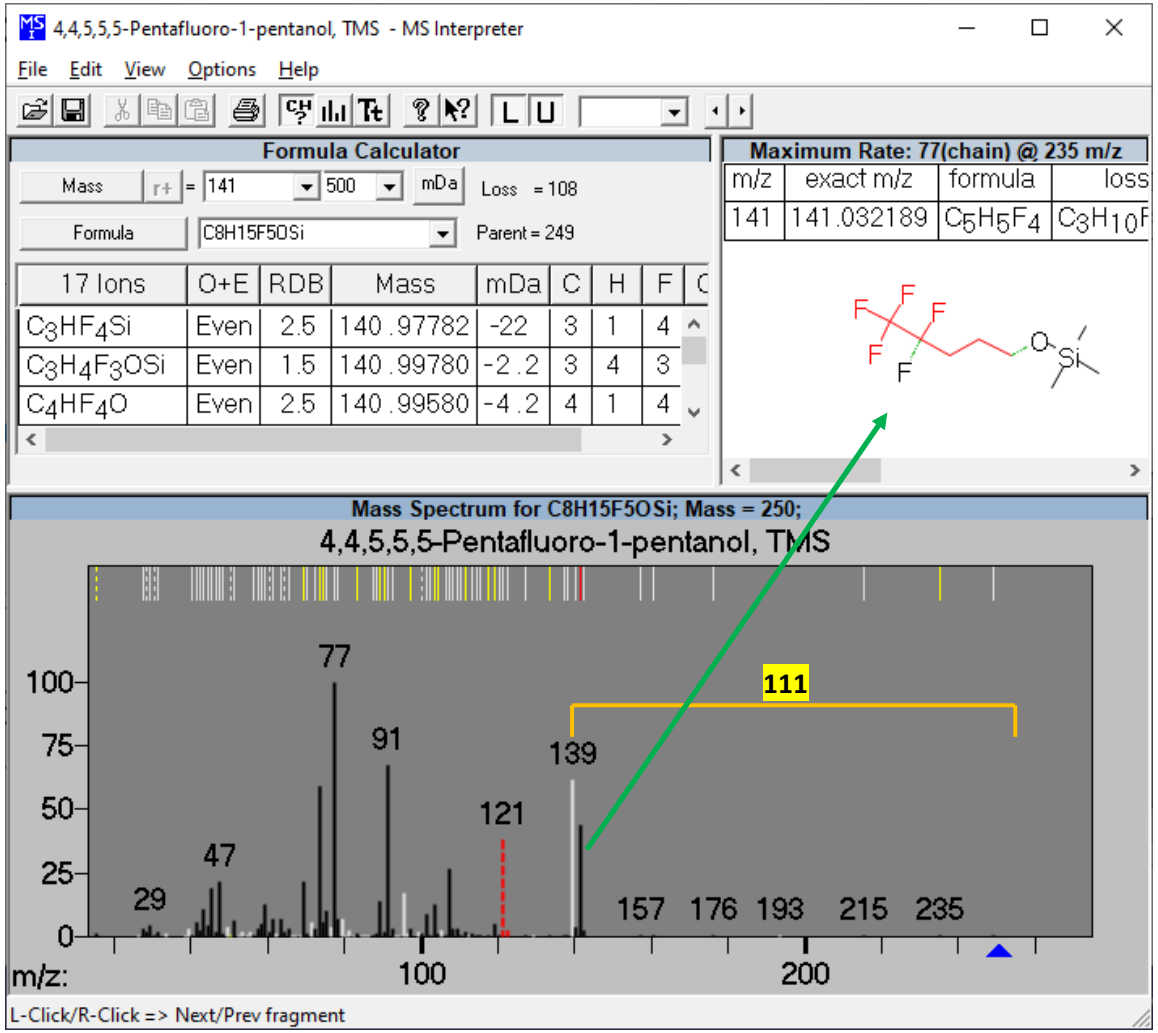
201 Compounds with low quality spectra or no retention index (RI) data

NIST MS Interpreter Aids Spectrum Evaluation

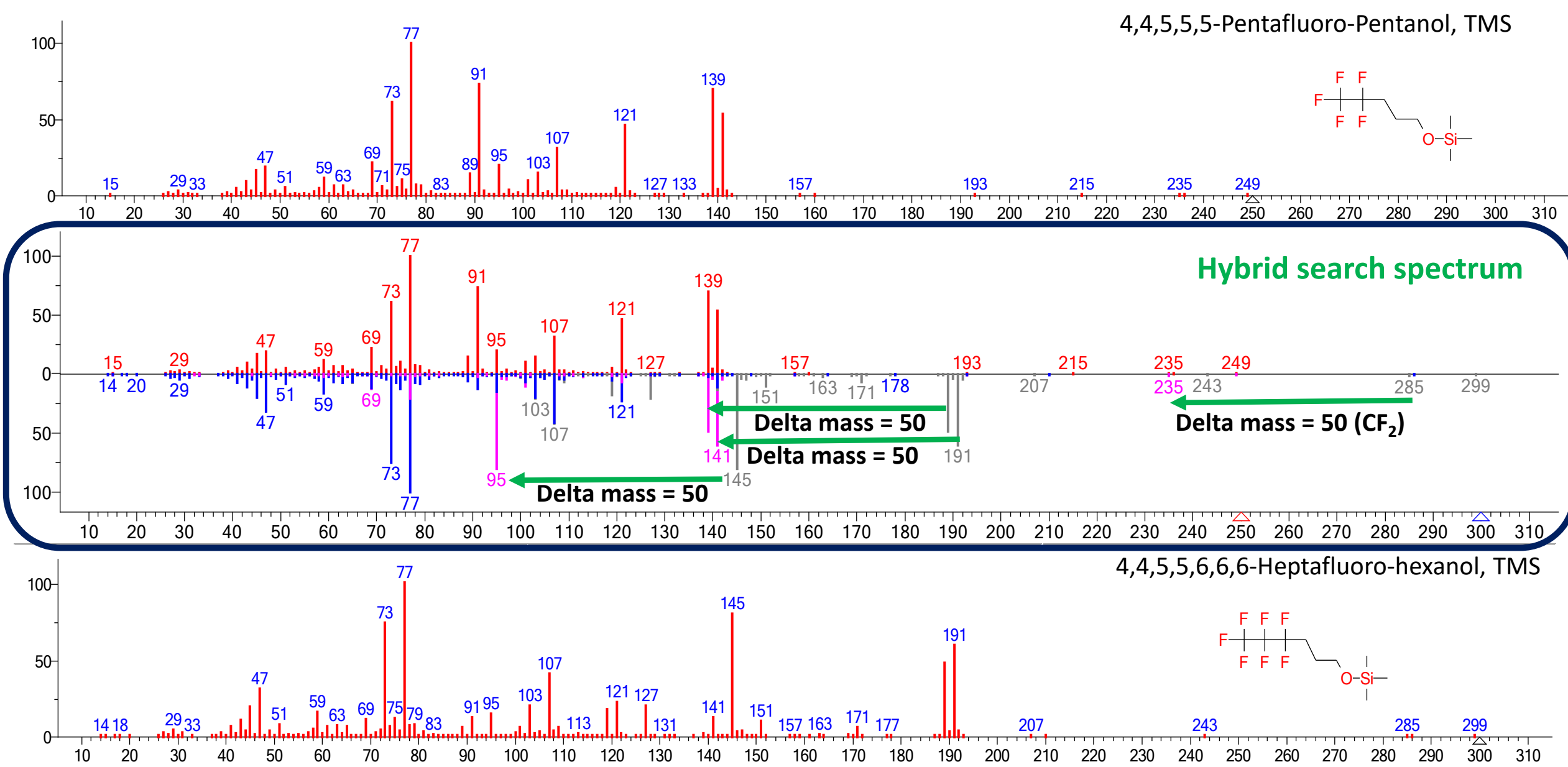
This program determines the probable origin of peaks using thermochemical rules

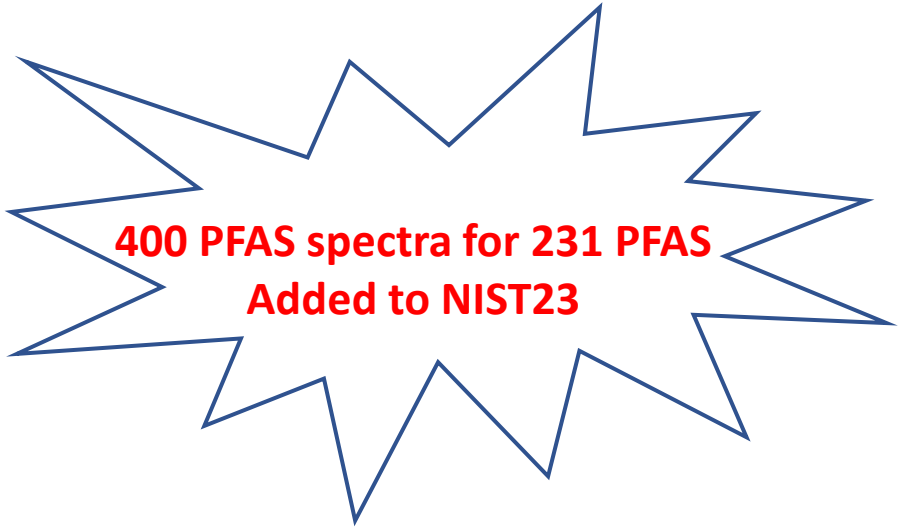


Name: 4,4,5,5,5-Pentafluoro-1-pentanol, TMS
Formula:C8H15F5OSi
MW:250



Hybrid Search to Aid Spectrum Evaluation





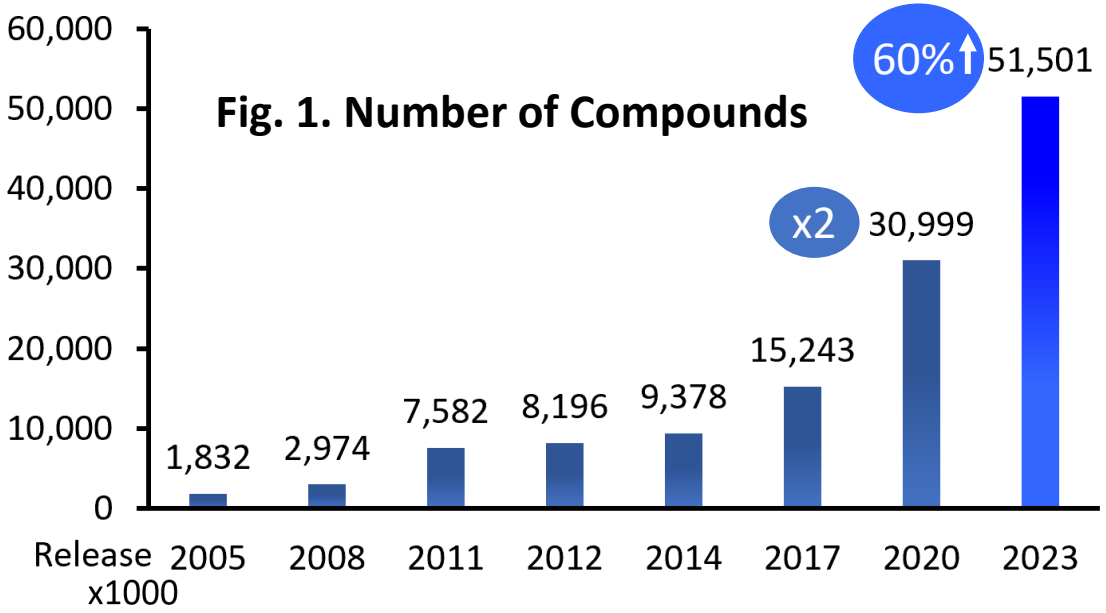
6/7/23 WOE 2:50PM GC/MS: Instrumentation and Applications; Room 332

Enhancing the Coverage and Quality of Spectra of per- and polyfluoroalkyl Substances in a Comprehensive Electron Ionization Mass Spectral Library

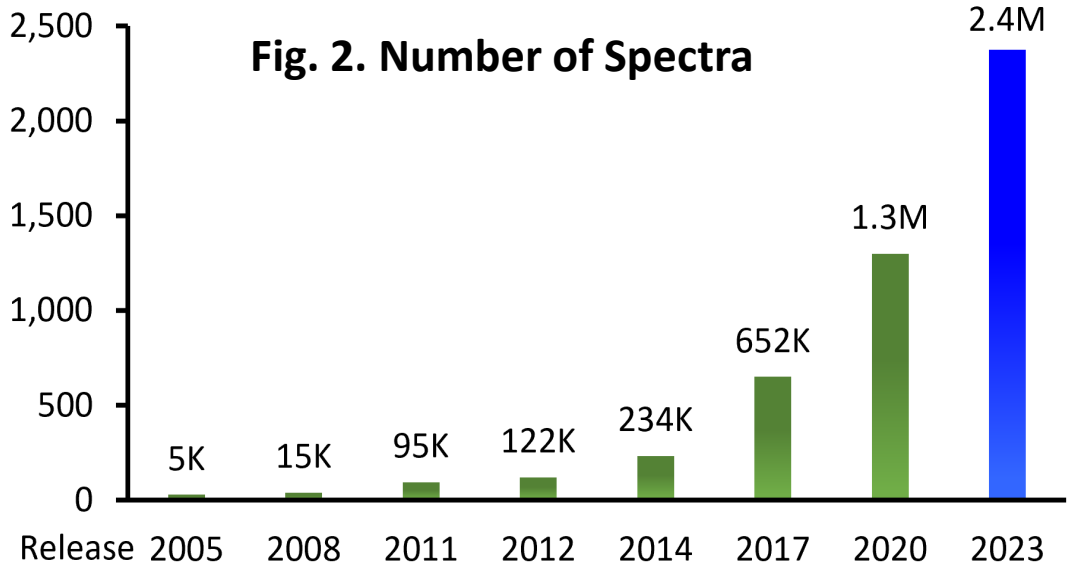
NIST23 Tandem Mass Spectral Library

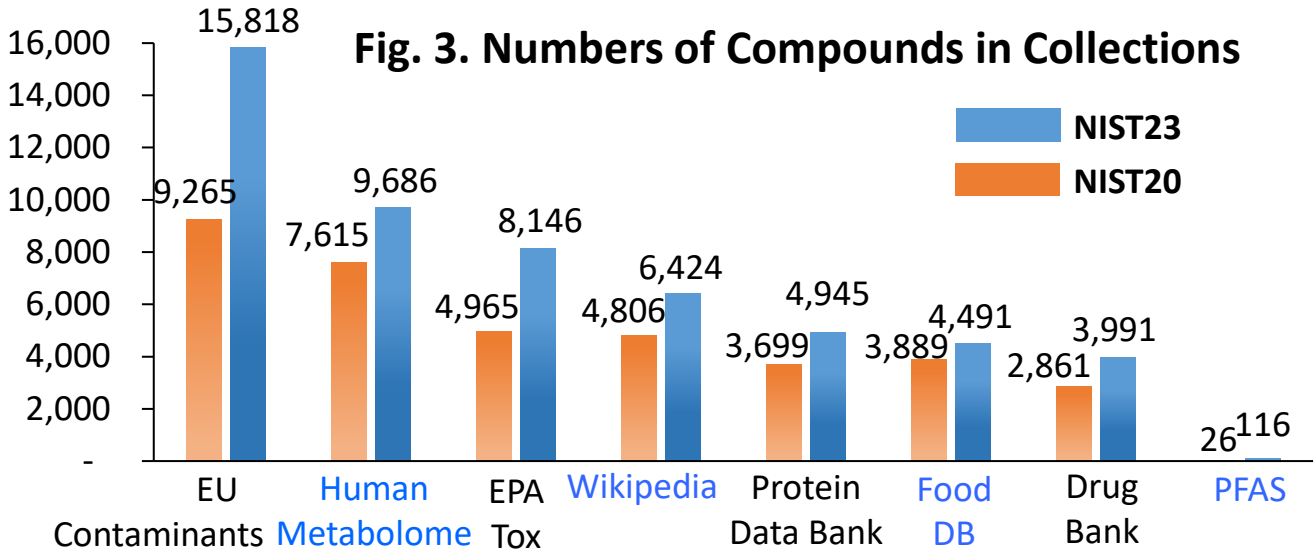
Xiaoyu (Sara) Yang

NIST23 Tandem Mass Spectral Library

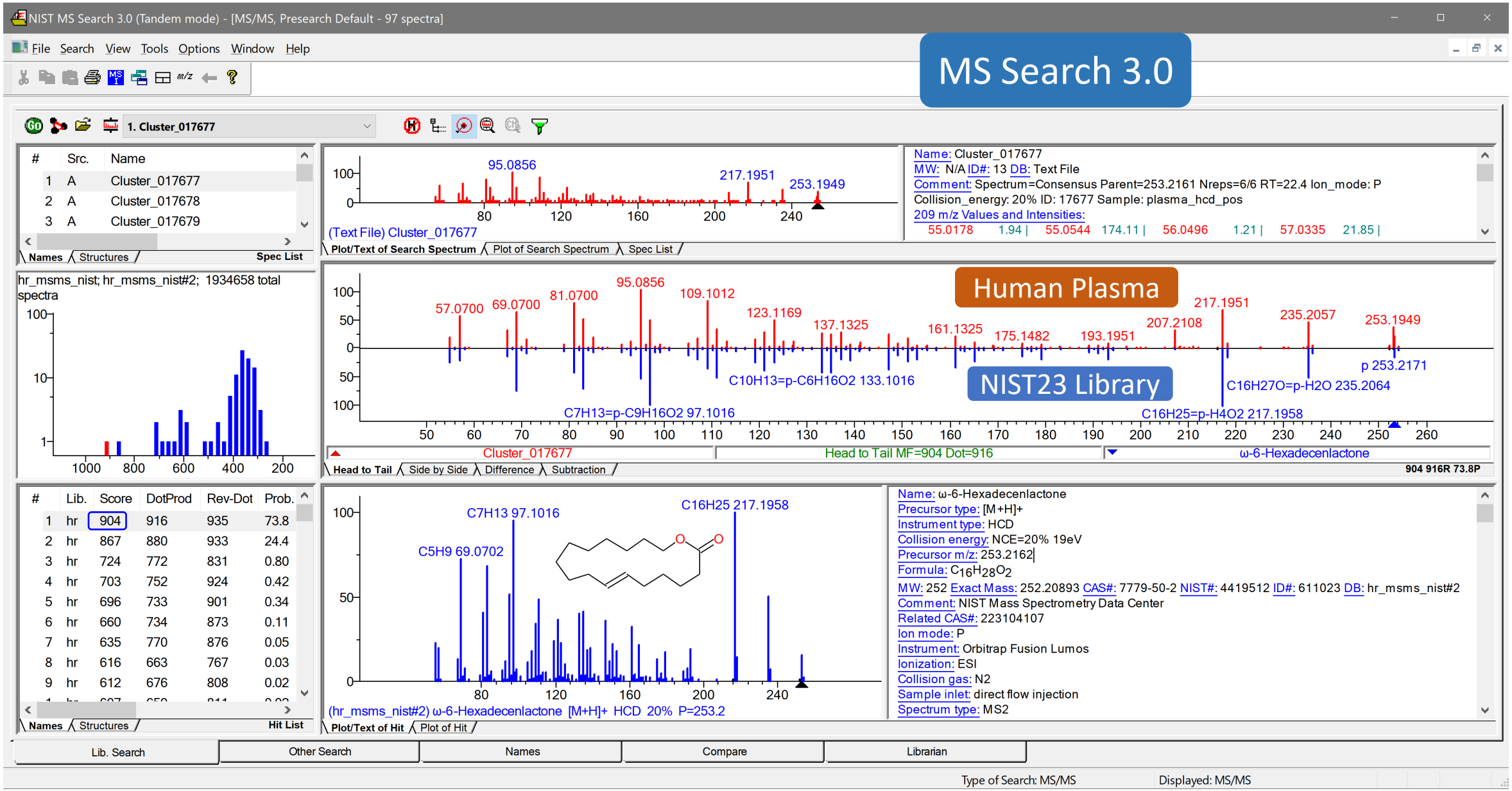


- Features: **high quality**
  comprehensive
  high and low resolution
  library searching software
- Application: metabolomics, health, environmental, pharmaceutical, forensics, food, and agriculture, etc.





Identifying Compounds by Searching the *NIST23* Tandem MS Library



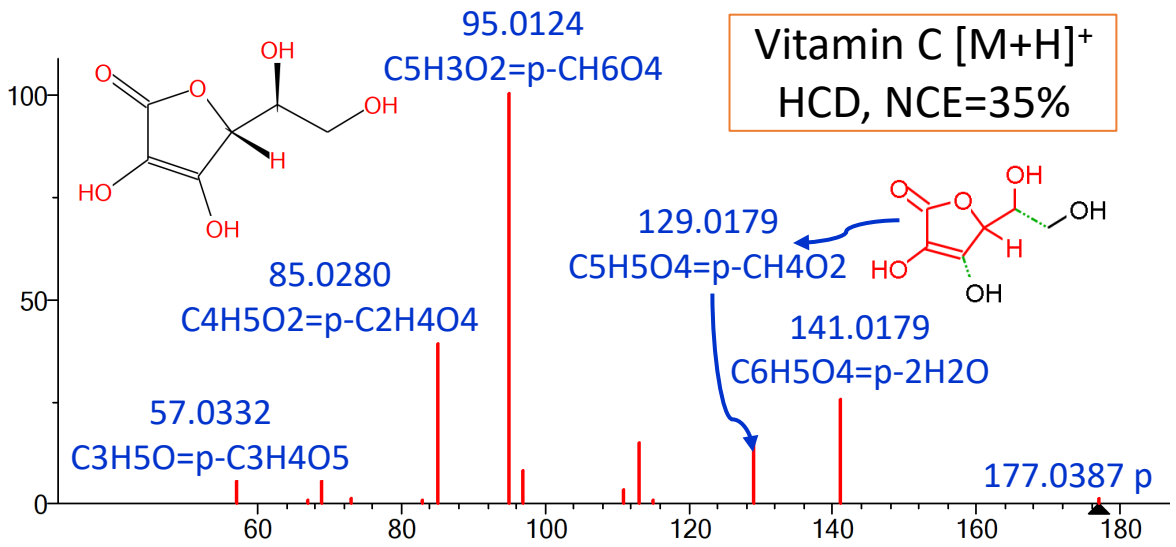
How Did We Create Our High Quality and Comprehensive Library?

1 Authentic compounds

2 Sample analysis
- ESI: HCD, FT, IT; QTOF
- APCI: QTOF
- Positive/Negative
- In-source
- 20 energy levels

Tandem mass spectra (MS^2, MS^3 and MS^4)

- Maximum coverage of compounds of analytical interest
- Widest range of fragmentation conditions



Annotating peaks with ***MS Interpreter***

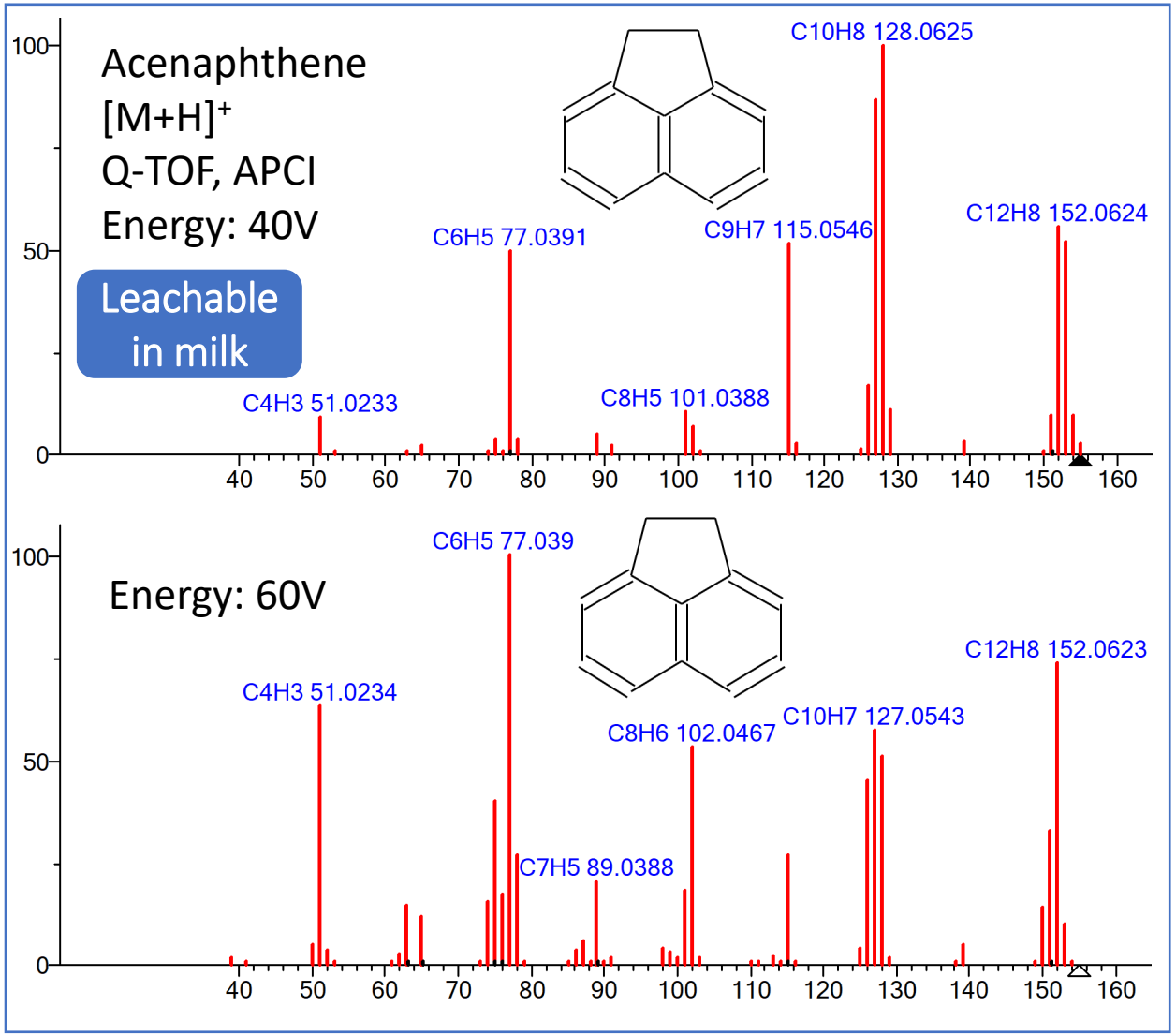
3 Quality control

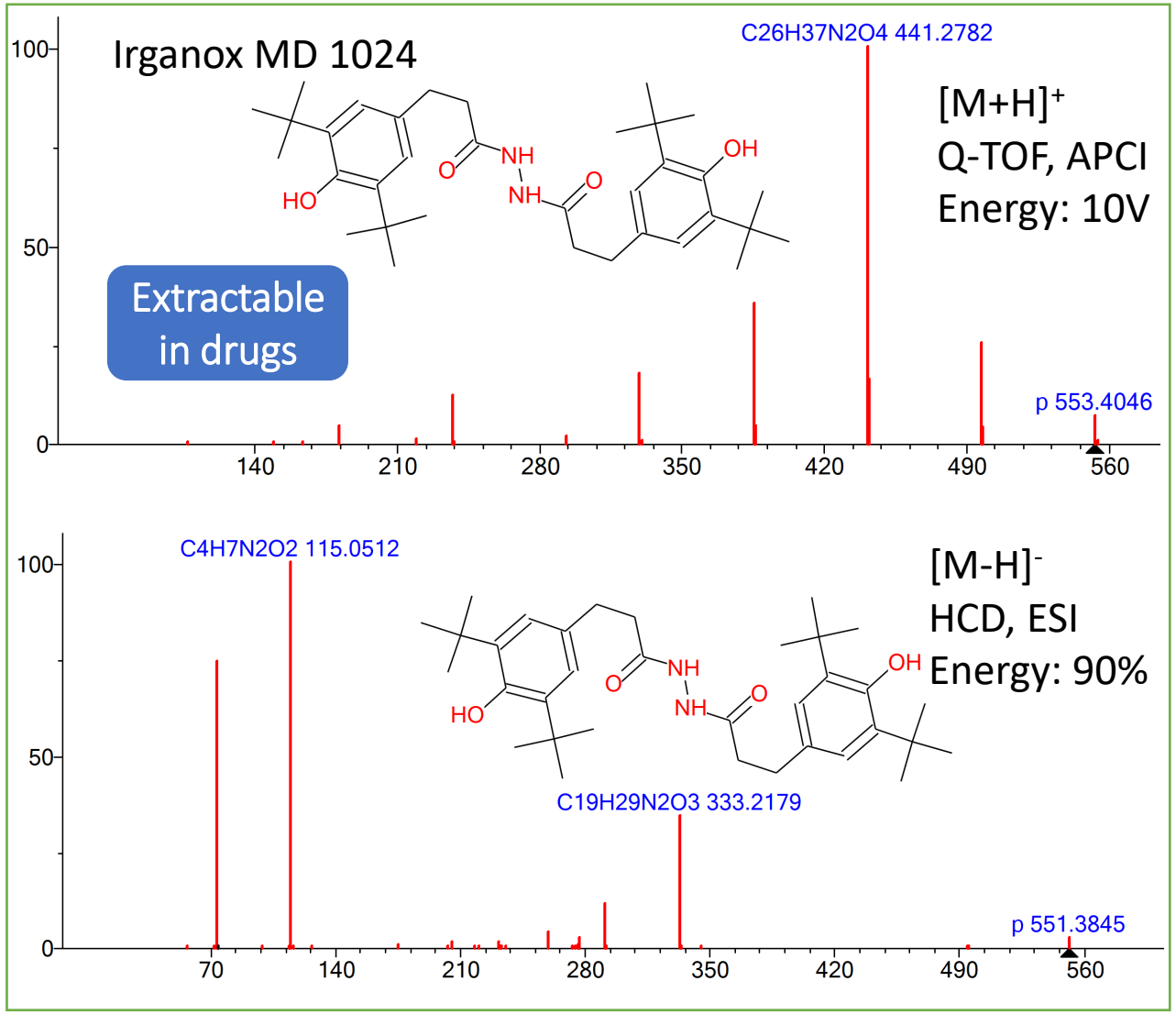
Full evaluation
Manual inspection
'Consensus spectra'
Extensive annotation

New
20,502 compounds
1 million spectra

NIST23 **Tandem MS Library**

561 Extractable & Leachable Compounds with APCI and 352 Compounds with APCI and ESI





NIST End-User and Distributor Support

Contractor

O. David Sparkman

massspec@nist.gov

Or

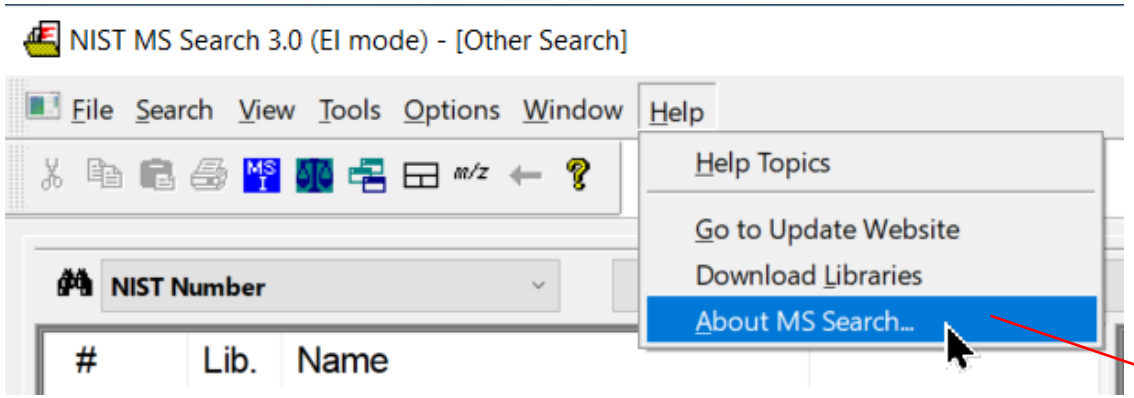
ods@csi.com

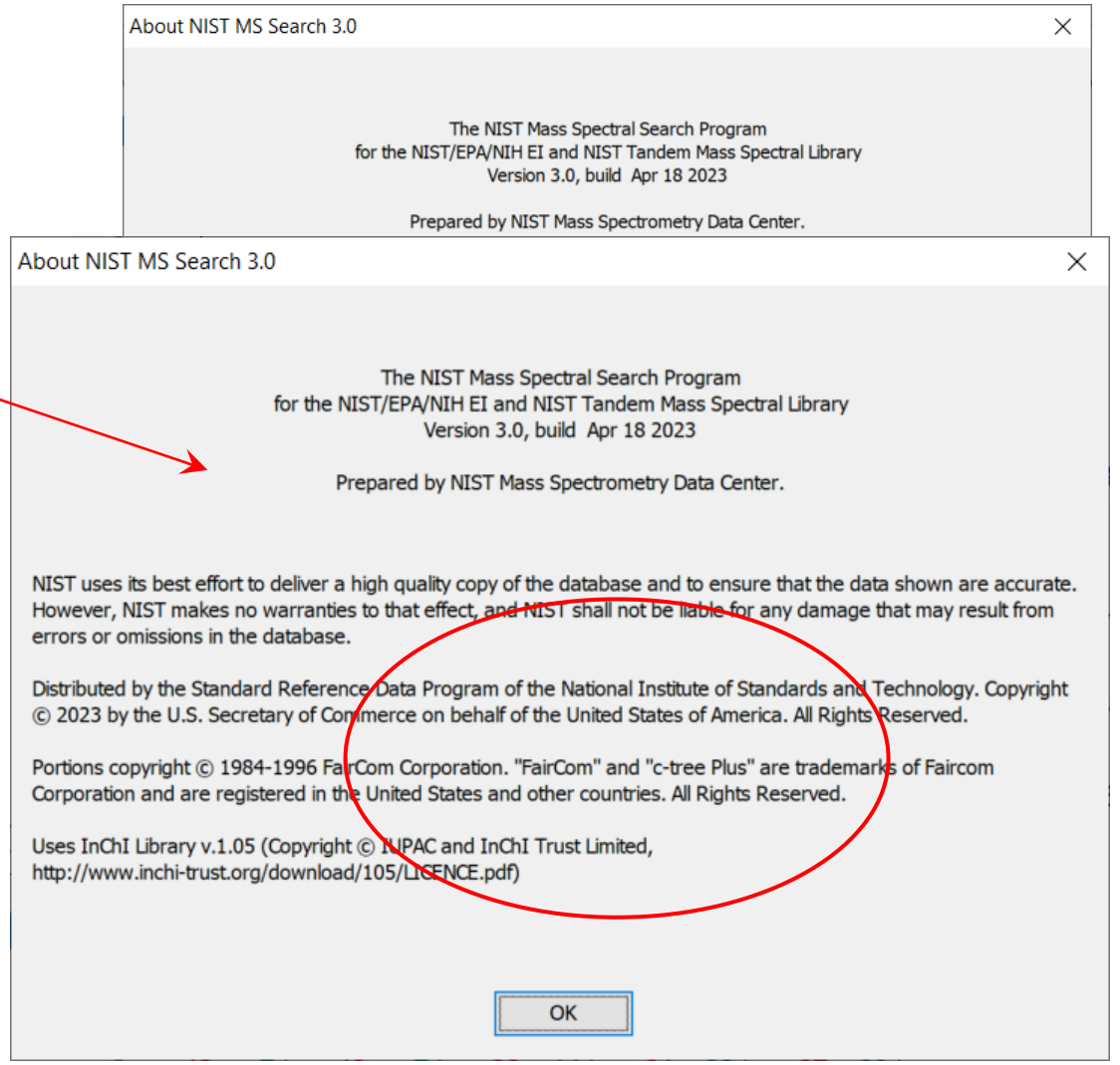
1-209-483-5740

Direct Benefit to Distributors

- **Compare NIST 23 with older versions and benefits of upgrading**
- **Interface MS Search to 3rd party software**
- **Use MS Search to perform tasks done by other software**
- **Clarify NIST MS Data licensing requirements**
- **Technical support for MS Search for your end users**
- **Submit inquires to massspec@nist.gov**

Custom Branding For Distributors



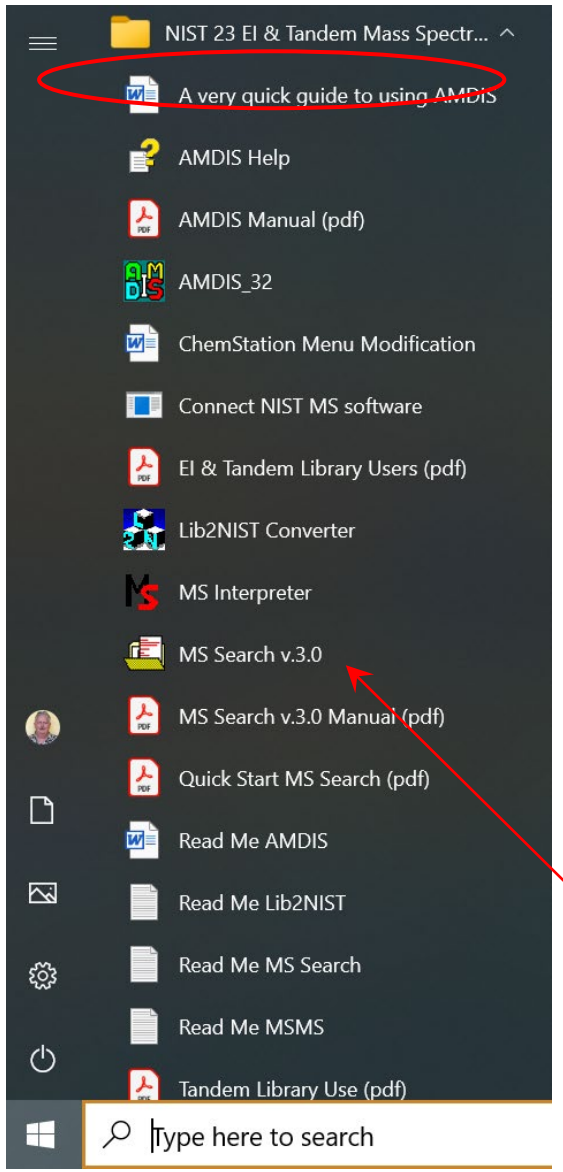


Branding Accomplished by:

Adding text file with name **Serial.dll** to installation media

Contact me for details

NIST 23 & MS Search Manual



Ver30Man.PDF
In
C:\NIST23\MSSearch\

Windows Start Menu

NIST Standard Reference Database 1A

Mass Spectral Libraries (NIST/EPA/NIH EI & NIST Tandem 2023) and NIST Mass Spectral Search Program (Version 3.0)

For Use with Microsoft® Windows

User's Guide

The NIST Mass Spectrometry Data Center

Stephen E. Stein, Technical Leader
William E. Wallace, Group Leader

Evaluation:

Weihua Ji (EI)
James L. Little (EI)
H. Martin Garraffo (Tandem)
Sanford P. Markey (EI)
W. Gary Mallard (EI)
Pedatsur Neta (Tandem)
Kirill V. Tretyakov (EI)
O. David Sparkman (EI)

Measurement/Evaluation:

N. Rabe Andriamaharavo (EI)
Tallat H. Bukhari (Tandem)
Edward P. Erisman (EI)
Yuxue Liang (Tandem)
Yi Liu (Tandem)
Yamil Simòn-Manso (Tandem)
Quan-Long Pu (EI)
Concepcion A. Remoroza (Tandem)
Nino G. Todua (EI)
Yufang Zhang (EI)

Hyperlinks in Manual and to External files

Hyperlinks

Contents

And
Delta_Mass_Guide.PDF

Delta Mass Guide for EI

The Group/Element agreements below have been observed in the eveluton of actual spectra using the Hybrid Search

There are **Group/Element (1)** and **Group Element (2)**. If the sign of δ is + (positive, assumed) then **Group Element (2)** is a possible replacement for what may be **Group/Element (1)** in the structure of the library compound used as the model. If the sign of δ is - (negative, specified) then Group Element (1) is a possible replacement for what may be Group/Element (2) in the structure of the library compound used as the model

DeltaMass	Nominal	Group/Element (1)	Group/Element (2)
	0	imidazole ring	furan ring
	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 N 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

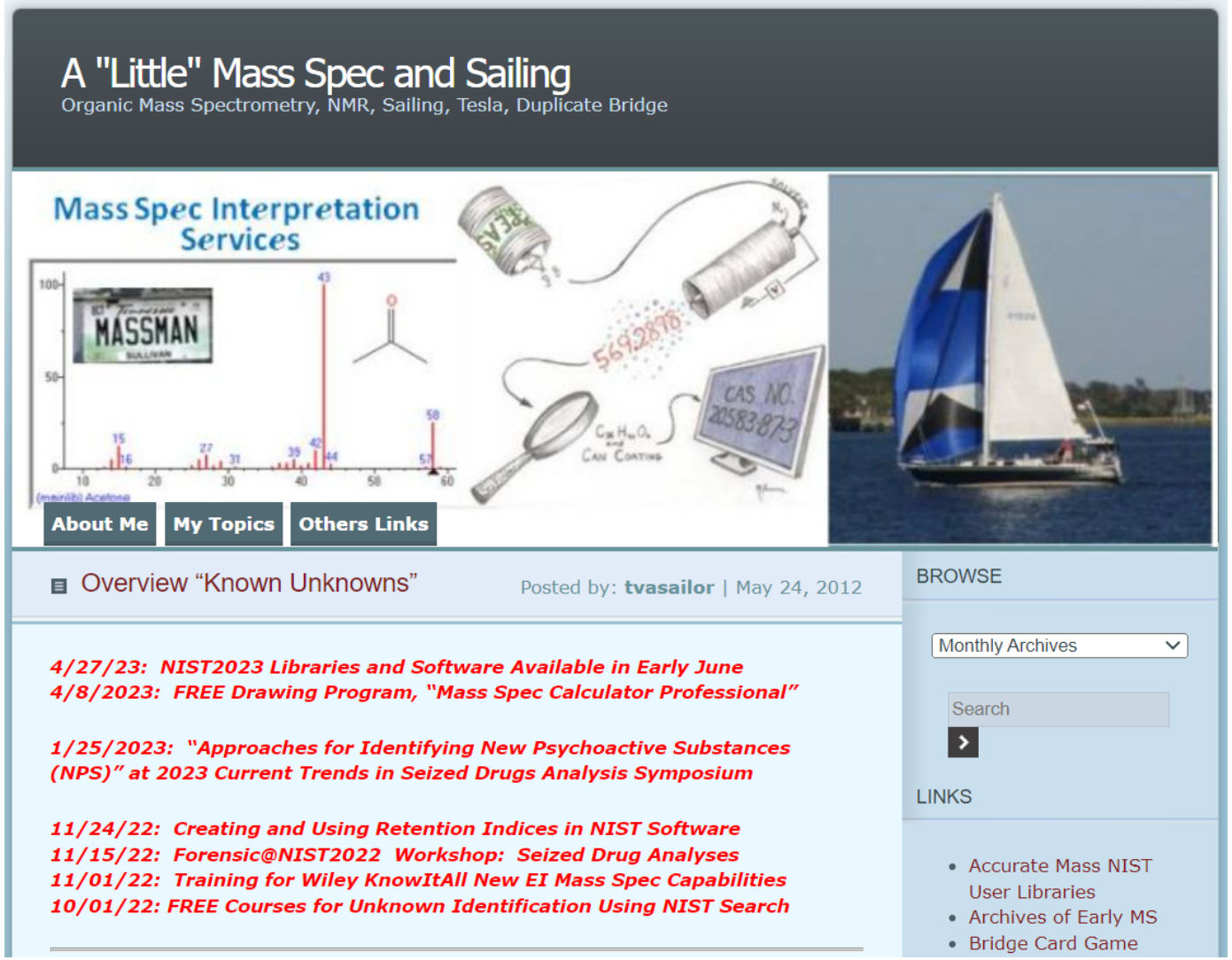
Hyperlink in Mar
OtherDBs.PDF

Catalog Name	URL
EPACOMPTOX_EPA official InertIngredients_Nov2019	c.epa.gov/dashboard/chemical-lists/PESTINERTS
EPACOMPTOX_Pesticides	
InertFinder_Nov2020	
EPACOMPTOX_PLASTICS NORMAN 2019	c.epa.gov/dashboard/chemical-lists/CPPDBLISTB
Thermo_AdditivesList2022	

Both in:
C:\NIST23\MSSearch

Training Videos for NIST 23 and beyond

https://littlemsandsailing.wordpress.com/



Training videos, handouts and instruction on NIST Libraries and NIST Mass Spectral Search Program

See You at Booth 713

- Speaker Order
  - Bill (2), Steve (7), Tytus (3), Steve (10-1,1,2,3,2,1)
  - Weihua (4), Yufang (4), Sara (4), Sparkman (5)

- Stein: Coverage, Software, Names/Spectra, Search Methods, Non-Spectral Data, EI-RI, Search Types, Hybrid Search, MS Interpreter