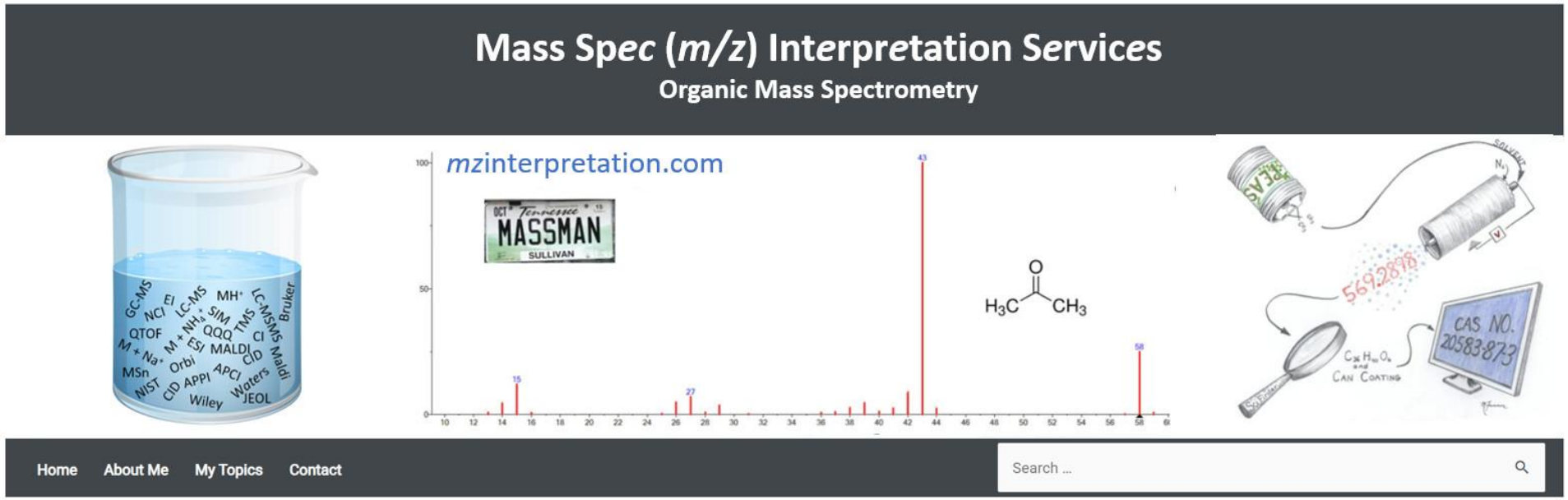
Creating and Using Free Public and User Libraries in NIST26 Chromatogram and Lib Search

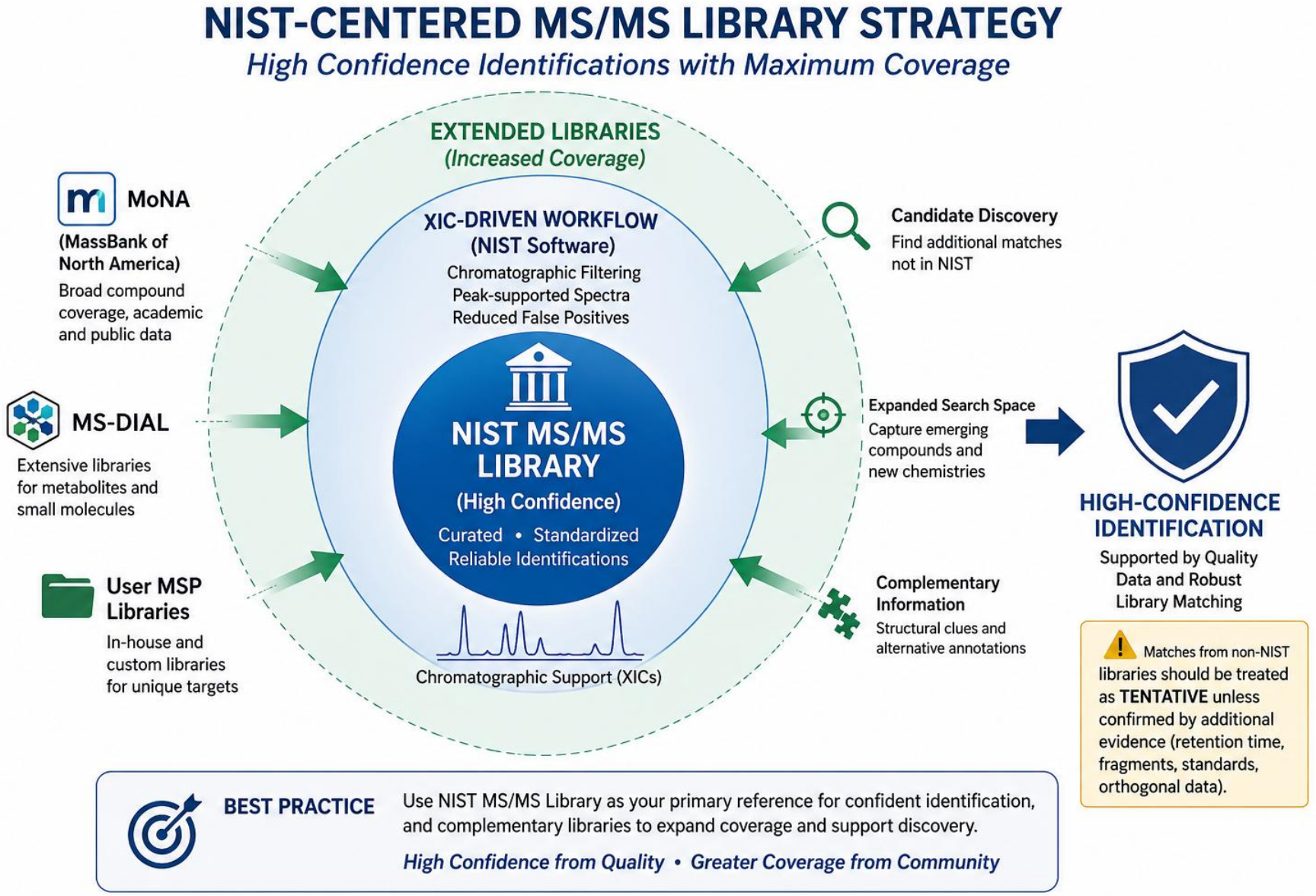
Video/Handout

James Little
Mass Spec Interpretation Services
April 25, 2026
mzinterpretation.com

See *Full Course* on NIST26 with new *Integrated* Deconvolution/Library Searching for *EI GC-MS* and *LC-MS/MS*!



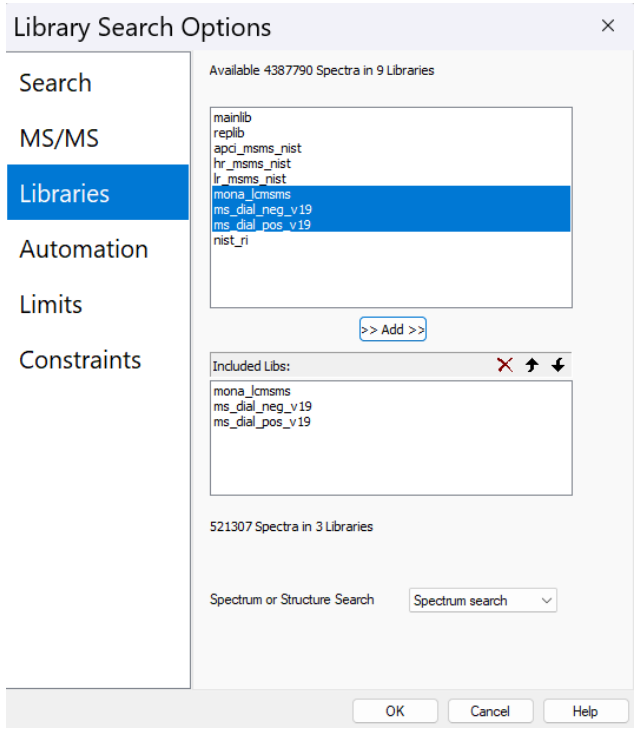
- NIST Library is the gold standard for MS/MS Identifications
- Price is very reasonable, shop around! Prices vary Dramatically, ***see link***
- However, free public libraries and user libraries are useful



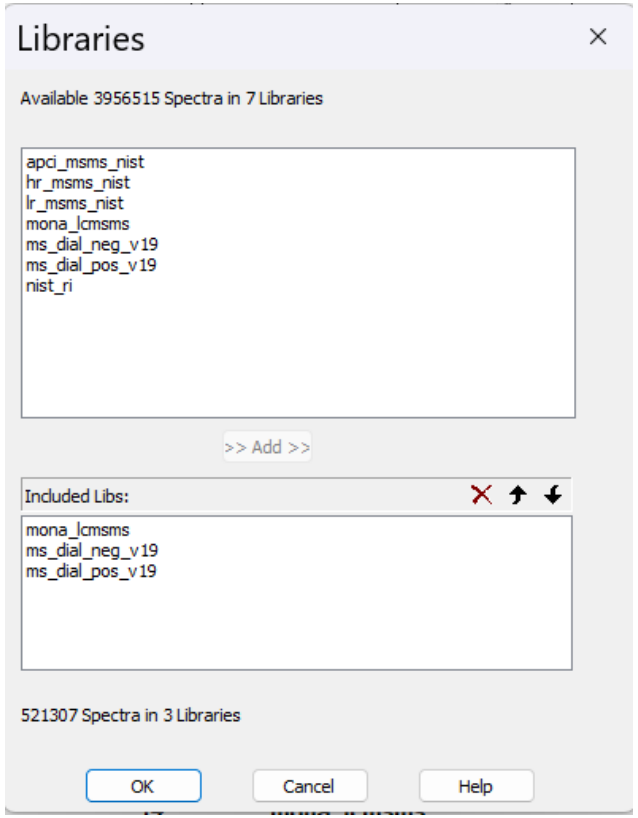
Free Public Libraries in NIST Format

- NIST commercial MS/MS (Tandem) Libraries are the “gold-standard”
- Majority of the most useful and diverse components in my identification evaluations
- However, there are some useful MS/MS spectra in free pubic libraries on the internet
- Initially, use free copies in NIST format found on my website
- If useful, learn to create your own copies from more current versions
- Definitely create your own unique user library
- >500,000 free public MS/MS spectra downloaded at the *following link*

Lib Search Library Selection



Chromatogram Library List



Source	File Name	MS/MS Spectra	Structures
MoNA (May 2025)	mona_lcmsms	153,635	150,404
MS-DIAL (2025-2026)	ms_dial_pos_v19	322,191	0
MS-DIAL (2025-026)	ms_dial_neg_v19	45,483	0
		Total 521,309	**Total** 150,404

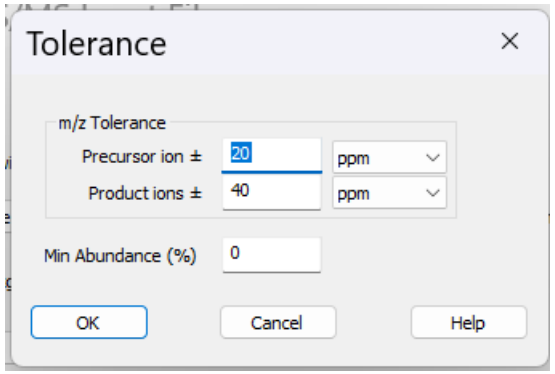
- **MoNA LC–MS/MS experimental download** provides the broadest, most generally useful set of *measured* tandem mass spectra in one file, while avoiding many specialized or lower-value subsets found on the MoNA website
- MoNA LC-MS/MS is a mixture of ~100,000 positive ion and ~50,000 negative ion spectra
- Combining the public MS-DIAL and MoNA LC–MS/MS libraries yields more than 500,000 tandem mass spectra, including roughly 450,000–550,000 spectra with exact fragment masses reported to three or more decimal places, representing approximately 30,000 unique compounds
- NIST 2026 tandem library substantially larger and more curated, containing ~3.2 million spectra from 68,635 compounds and offering deeper spectral coverage and more consistent identification performance
- Use NIST26 as primary MS/MS library, relatively very inexpensive
- Add MoNA and MS-DIAL libraries in NIST-compatible format to extend coverage further
- Benefit of NIST's curated core plus the breadth of major open-access libraries
- Integrated LC–MS/MS data processing in the 2026 NIST Search Software unlocks the full potential of these spectral libraries

Suggested Tolerances for Free Public MS/MS Library Searches

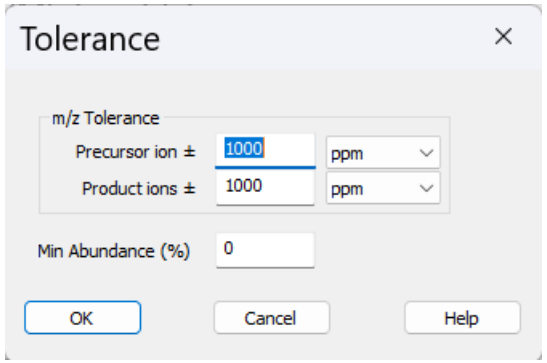
Identity MS/MS and Similarity MS/MS Hybrid

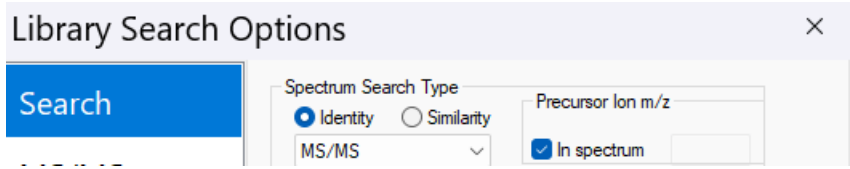
- ***Best analyzed*** with the high-resolution tolerances suggested by NIST for their libraries
- The NIST lr_msms_nist (low-resolution) library corrected automatically even with high-resolution tolerances
- Publice libraries contain many low-resolution spectra
- Can utilize the tolerances noted below to check that no useful low-resolution entries filtered out
- Normally, NIST suggested works best for both Identity MS/MS and Similarity MS/MS Hybrid (Precursor Ion *m/z* "In spectrum" selected)

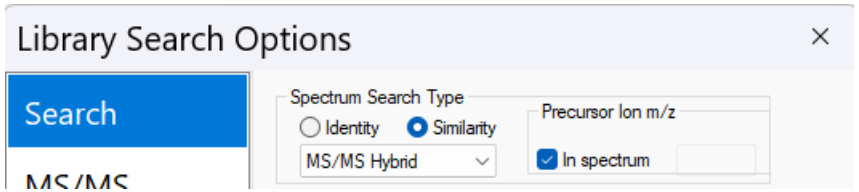
NIST Suggested
High Resolution
Best Approach



Alternate Low-Resolution
Less Preferred

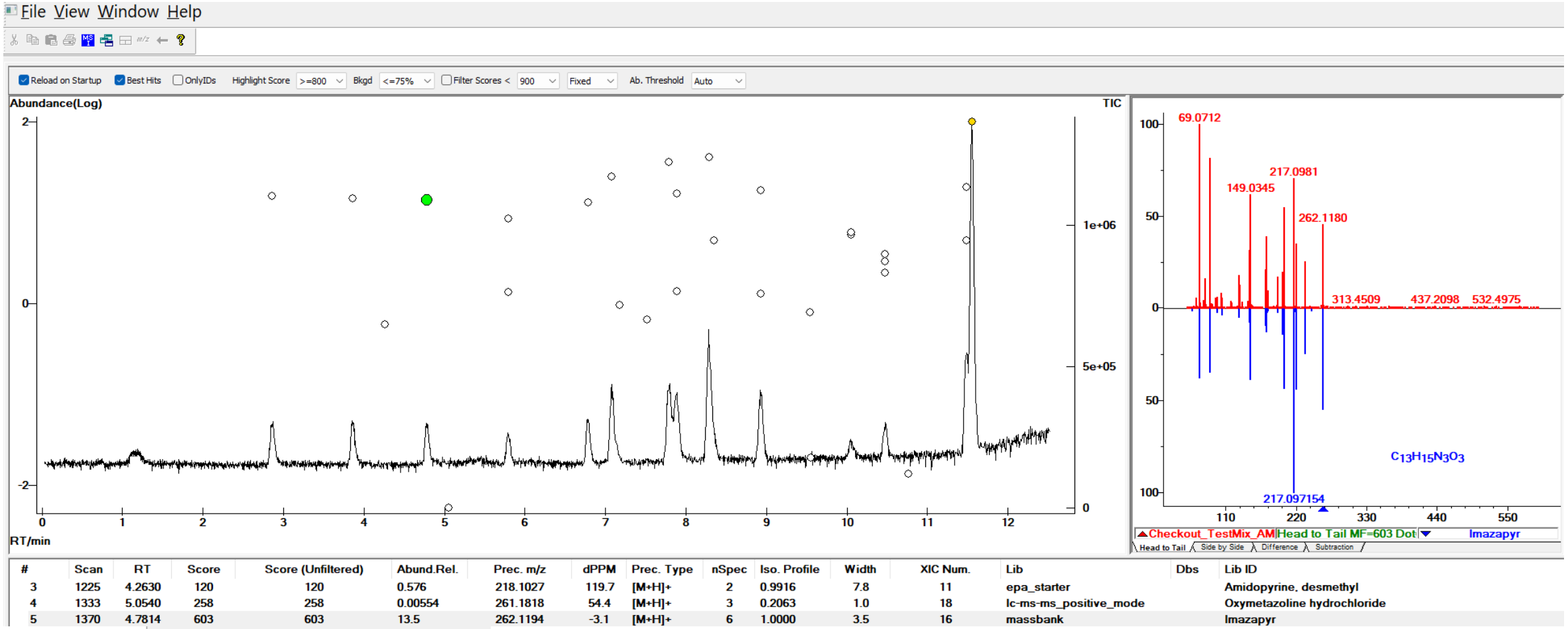






Compounds with No Structures for Many Free Public MS/MS Entries

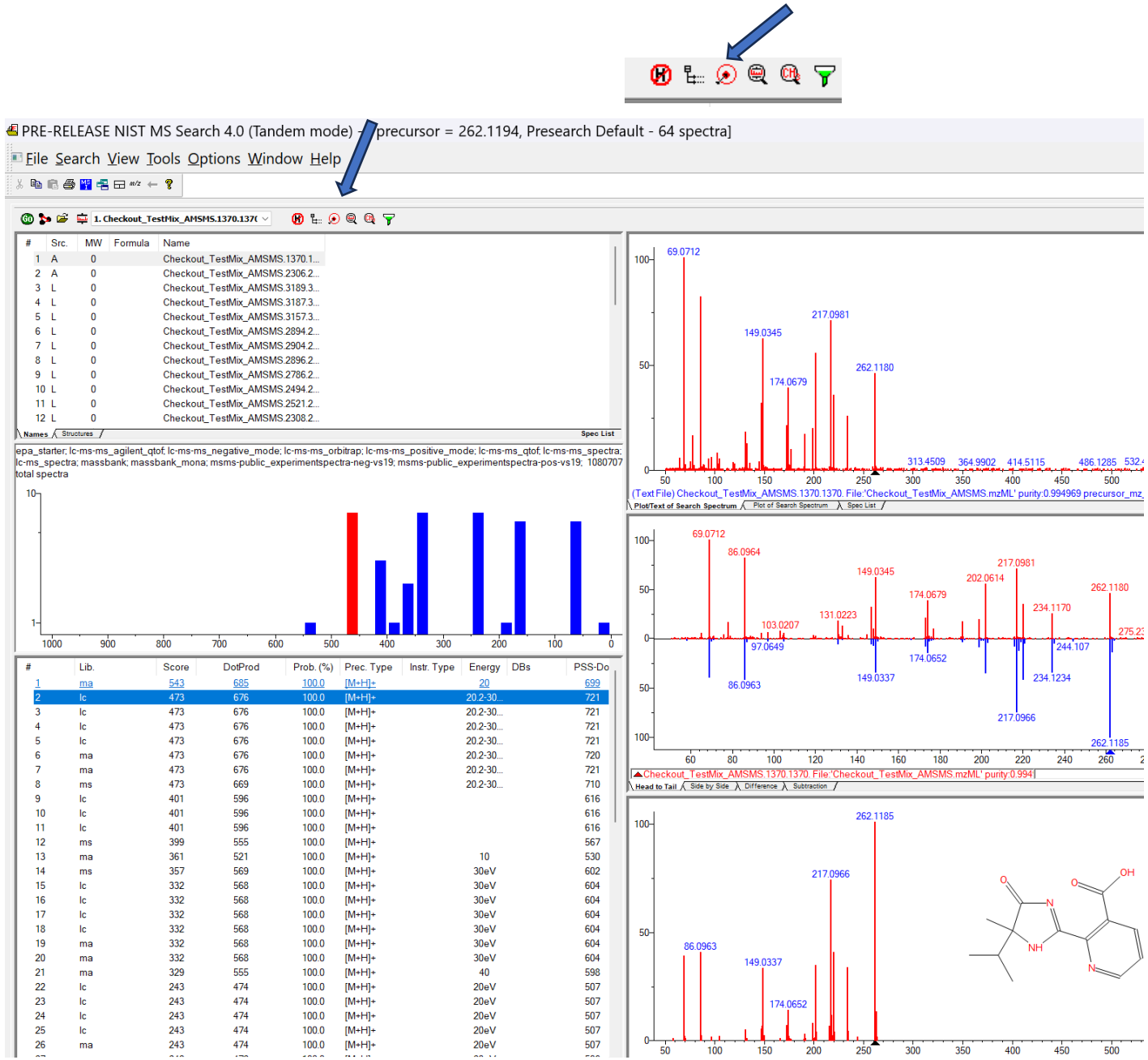
- Many of the compounds that will be found in Chromatogram Searches using the Free Public libraries might not have structures
- e.g. Imazapyr
- Imazapyr entry below, sent to Library Search with right click



#	Scan	RT	Score	Score (Unfiltered)	Abund.Rel.	Prec. m/z	dPPM	Prec. Type	nSpec	Iso. Profile	Width	XIC Num.	Lib	Dbs	Lib ID
3	1225	4.2630	120	120	0.576	218.1027	119.7	[M+H]+	2	0.9916	7.8	11	epa_starter		Amidopyrine, desmethyl
4	1333	5.0540	258	258	0.00554	261.1818	54.4	[M+H]+	3	0.2063	1.0	18	lc-ms-ms_positive_mode		Oxymetazoline hydrochloride
5	1370	4.7814	603	603	13.5	262.1194	-3.1	[M+H]+	6	1.0000	3.5	16	massbank		Imazapyr

Lib Search of Imazapyr showing structures

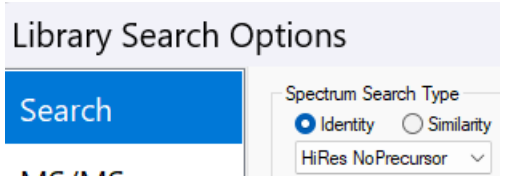
- Many of the hits had structures
- Only the best hit found in Chromatogram Search did not
- Make sure the “Best Matching Only” *is not* selected (depressed), or the duplicates with structure *might not be* found



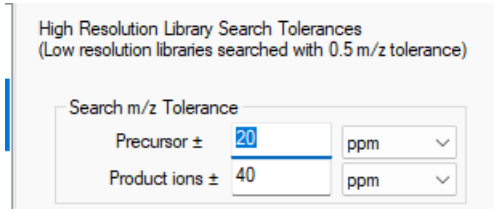
Suggested Tolerances for Free Public MS/MS Library Searches

Identity HiRes NoPrecursor

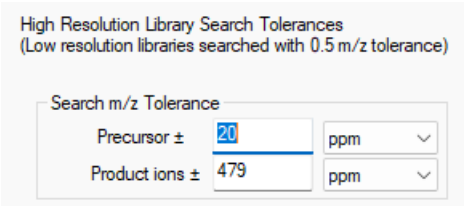
- This essentially matches the observed spectrum to the library MS/MS spectrum
- Can be useful for determining the class of a compound by looking at the top 2-5 hits
- Normally, use the Product Ion +/- 40 ppm, Precursor setting not important, not utilized
- For low-resolution spectra, thought it would be useful to increase to 1000 ppm, but values above 479 ppm yield 0 results! Thus, can try 479 ppm which might be useful?
- However, later tests showed increasing to 1000 ppm with
- Tolerances above 479 ppm are not a problem for the NIST hr_msms_nist library



Standard Product ion Tolerance for High-Resolution Spectra



Standard Product ion Tolerance for Low-Resolution Spectra



Create Your Own Updated Version of Free Public Libraries in NIST Format

- See NIST "List of Sources" of Free Public MS/MS Libraries in my website resources, ***see link***
- See video ***link*** and handout ***link*** discussing the approach and challenges of creating both Free Public Libraries and User libraries in NIST Format on my website
- At that point in time, I thought putting lr_ in front of the name of any user library would make the NIST software process with it with corrections utilized for their commercial low-resolution library, lr_msms_nist
- It does not, must use wide tolerances of 1000 ppm for precursor and fragment ions

James Little
tvasailor@gmail.com
mzinterpretation.com
Kingsport, TN

- *Retired* Research Fellow, Eastman Chem. Co.*
- *42 years experience unknown identification*
- *Now Consultant, MS Interpretation Services*
- *Specialties: EI GC-MS, LC-MS/MS, Chemical Ionization, Accurate Mass, Derivatization, MS library management, SciFinder, Chemspider, Surfactant ID, NMR, GC-IR, organic synthesis, matrix ionization effects, etc.*

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

*https://en.wikipedia.org/wiki/Eastman_Chemical_Company

>50 Mass Specs Networked
Worldwide

List of Sources Public Tandem Libraries/Data

NIST Compatible *msp* or *SDF* files (convertible to lib

MassBank-Europe - https://github.com/MassBank/

MoNA - https://massbank.us/downloads (includes S

MS-Dial - https://systemsomicslab.github.io/compr

mzInterpretation.com - https://mzinterpretation.co
spectra-in-nist-format/ (also tutorials and test files)

Other formats (*mgf*, *json*, *csv* or alternate *msp*)

GNPS - https://gnps.ucsd.edu/ProteoSAFe/libraries