

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**!

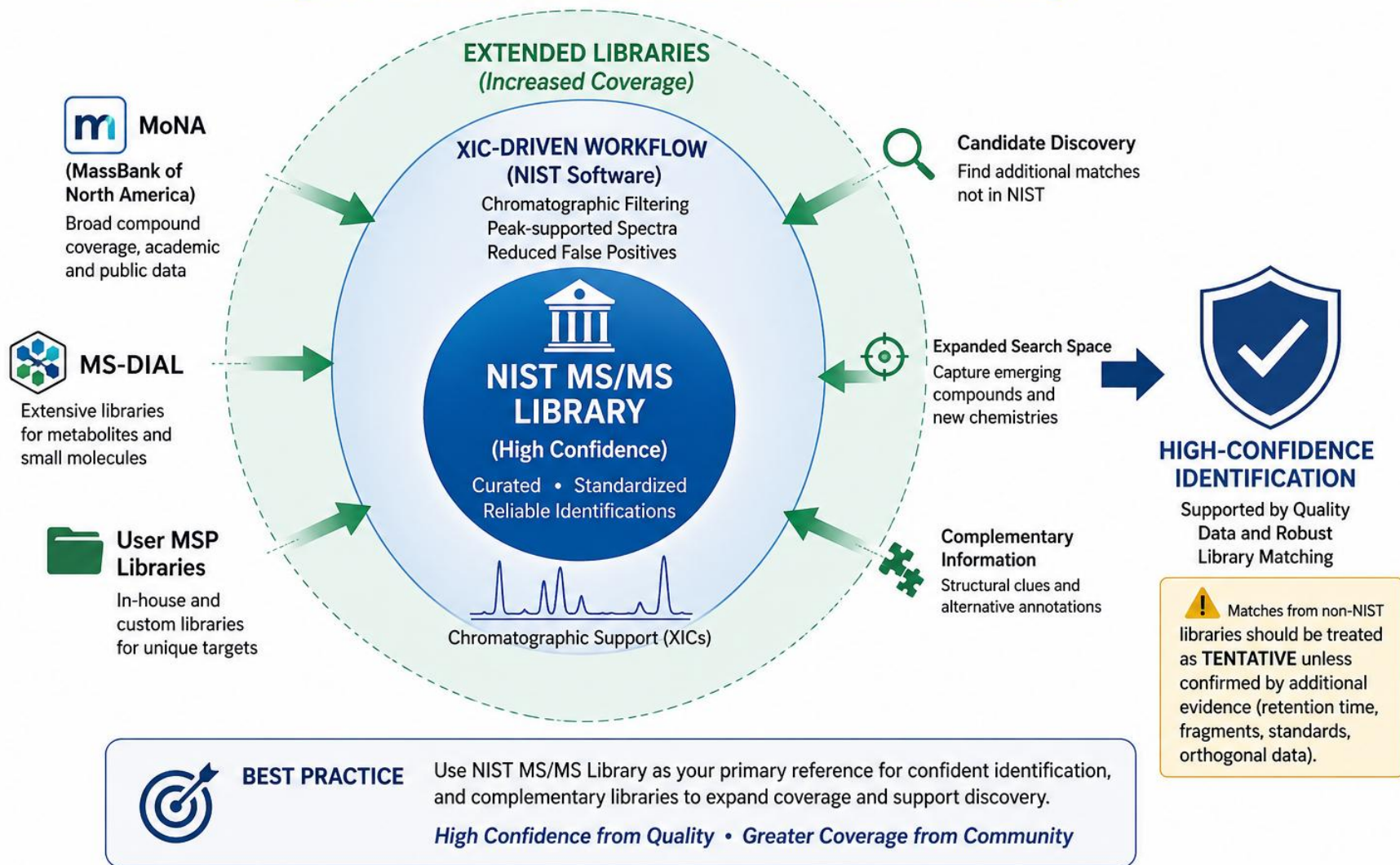
Mass Spec (m/z) Interpretation Services Organic Mass Spectrometry



NIST-CENTERED MS/MS LIBRARY STRATEGY

High Confidence Identifications with Maximum Coverage

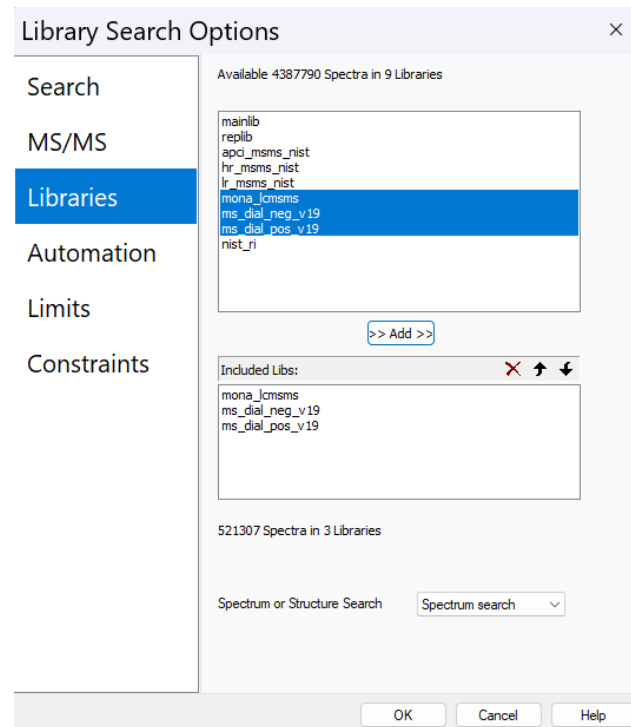
- NIST Library is the gold standard for MS/MS Identifications
- Price is very reasonable, shop around! Prices vary Dramatically, [see link](#)
- However, free public libraries and user libraries are useful



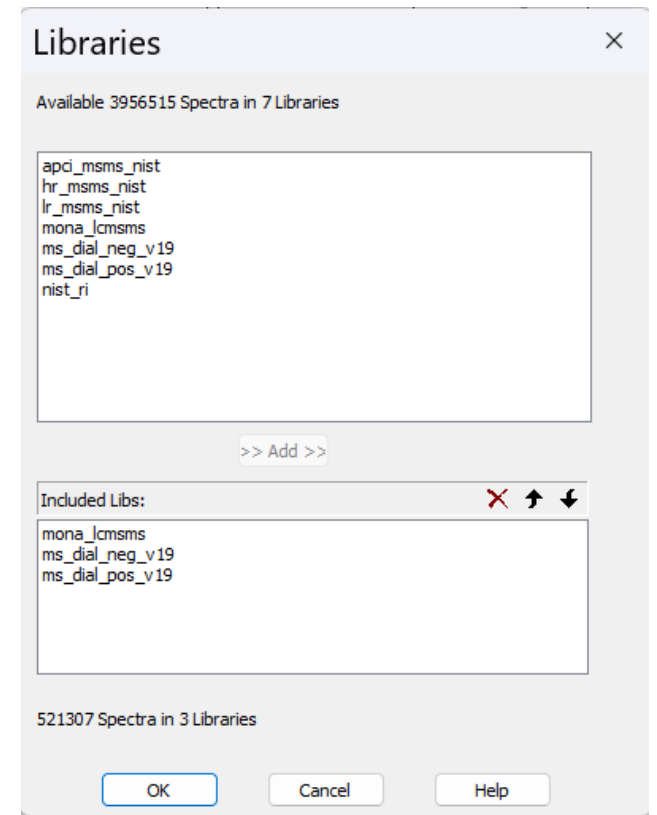
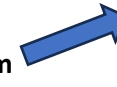
Free Public Libraries in NIST Format

- 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 public 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 `Ir_msms_nist` (low-resolution) library corrected automatically even with high-resolution tolerances
- Public 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

The 'Tolerance' dialog box shows the following settings:

- m/z Tolerance: Precursor ion ± 20 ppm, Product ions ± 40 ppm
- Min Abundance (%): 0

Buttons: OK, Cancel, Help

Alternate Low-Resolution
Less Preferred

The 'Tolerance' dialog box shows the following settings:

- m/z Tolerance: Precursor ion ± 1000 ppm, Product ions ± 1000 ppm
- Min Abundance (%): 0

Buttons: OK, Cancel, Help

The 'Library Search Options' dialog box shows the following settings:

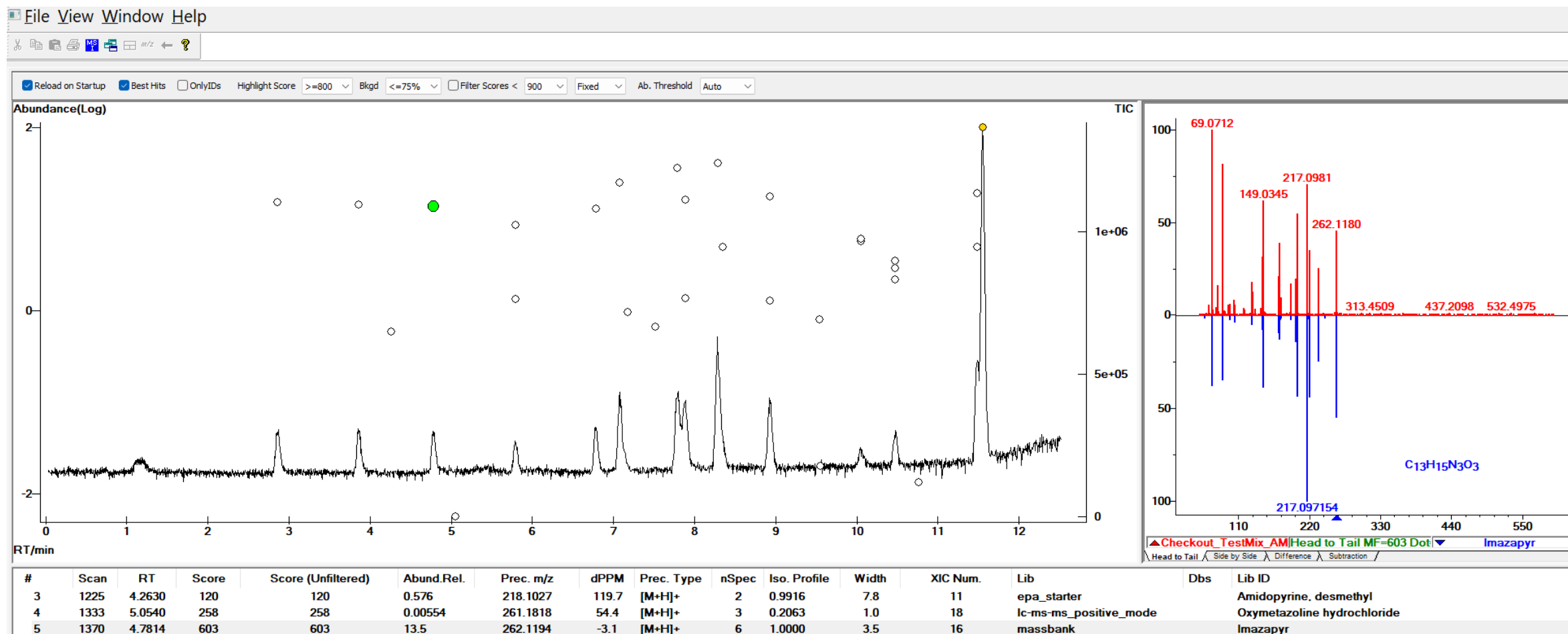
- Search button
- Spectrum Search Type: Identity (selected), Similarity
- MS/MS dropdown menu
- Precursor Ion m/z: In spectrum (checked)

The 'Library Search Options' dialog box shows the following settings:

- Search button
- Spectrum Search Type: Similarity (selected), Identity
- MS/MS Hybrid dropdown menu
- Precursor Ion m/z: In spectrum (checked)

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



Properties

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

PRE-RELEASE NIST MS Search 4.0 (Tandem mode) - precursor = 262.1194, Presearch Default - 64 spectra

File Search View Tools Options Window Help

#	Src.	MW	Formula	Name
1	A	0		Checkout_TestMix_AMSMS.1370.1...
2	A	0		Checkout_TestMix_AMSMS.2306.2...
3	L	0		Checkout_TestMix_AMSMS.3189.3...
4	L	0		Checkout_TestMix_AMSMS.3187.3...
5	L	0		Checkout_TestMix_AMSMS.3157.3...
6	L	0		Checkout_TestMix_AMSMS.2894.2...
7	L	0		Checkout_TestMix_AMSMS.2904.2...
8	L	0		Checkout_TestMix_AMSMS.2896.2...
9	L	0		Checkout_TestMix_AMSMS.2786.2...
10	L	0		Checkout_TestMix_AMSMS.2494.2...
11	L	0		Checkout_TestMix_AMSMS.2521.2...
12	L	0		Checkout_TestMix_AMSMS.2308.2...

Names Structures Spec List

epa_starter;lc-ms-ms_agilent_qtof;lc-ms-ms_negative_mode;lc-ms-ms_orbitrap;lc-ms-ms_positive_mode;lc-ms-ms_qtof;lc-ms-ms_spectra;lc-ms_spectra;massbank;massbank_mona;msms-public_experimentspectra-neg-vs19;msms-public_experimentspectra-pos-vs19; 1080707 total spectra

(Text File) Checkout_TestMix_AMSMS.1370.1370 File:Checkout_TestMix_AMSMS.mzML_purity:0.994969 precursor_mz

Head to Tail Side by Side Difference Subtraction

#	Lib.	Score	DotProd	Prob. (%)	Prec. Type	Instr. Type	Energy	DBs	PSS-Do
1	ma	543	685	100.0	[M+H] ⁺		20		699
2	lc	473	676	100.0	[M+H] ⁺	20.2-30...			721
3	lc	473	676	100.0	[M+H] ⁺	20.2-30...			721
4	lc	473	676	100.0	[M+H] ⁺	20.2-30...			721
5	lc	473	676	100.0	[M+H] ⁺	20.2-30...			721
6	ma	473	676	100.0	[M+H] ⁺	20.2-30...			720
7	ma	473	676	100.0	[M+H] ⁺	20.2-30...			721
8	ms	473	669	100.0	[M+H] ⁺	20.2-30...			710
9	lc	401	596	100.0	[M+H] ⁺				616
10	lc	401	596	100.0	[M+H] ⁺				616
11	lc	401	596	100.0	[M+H] ⁺				616
12	ms	399	555	100.0	[M+H] ⁺				567
13	ma	361	521	100.0	[M+H] ⁺		10		530
14	ms	357	569	100.0	[M+H] ⁺		30eV		602
15	lc	332	568	100.0	[M+H] ⁺		30eV		604
16	lc	332	568	100.0	[M+H] ⁺		30eV		604
17	lc	332	568	100.0	[M+H] ⁺		30eV		604
18	lc	332	568	100.0	[M+H] ⁺		30eV		604
19	ma	332	568	100.0	[M+H] ⁺		30eV		604
20	ma	332	568	100.0	[M+H] ⁺		30eV		604
21	ma	329	555	100.0	[M+H] ⁺		40		598
22	lc	243	474	100.0	[M+H] ⁺		20eV		507
23	lc	243	474	100.0	[M+H] ⁺		20eV		507
24	lc	243	474	100.0	[M+H] ⁺		20eV		507
25	lc	243	474	100.0	[M+H] ⁺		20eV		507
26	ma	243	474	100.0	[M+H] ⁺		20eV		507

Chemical structure of Imazapyr: CC1(C)C(=O)N=C1C2=CC=CC=C2C(=O)O

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

Library Search Options

Search

Spectrum Search Type

Identity Similarity

HiRes NoPrecursor

Standard Product ion Tolerance for High-Resolution Spectra

High Resolution Library Search Tolerances
(Low resolution libraries searched with 0.5 m/z tolerance)

Search m/z Tolerance

Precursor \pm ppm

Product ions \pm ppm

Standard Product ion Tolerance for Low-Resolution Spectra

High Resolution Library Search Tolerances
(Low resolution libraries searched with 0.5 m/z tolerance)

Search m/z Tolerance

Precursor \pm ppm

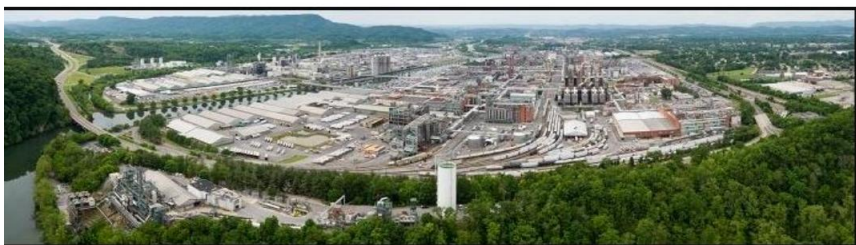
Product ions \pm ppm

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 &

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

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

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

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