NIST EI Mass Spec Search Class
Part I: Spectral Searches with NIST MS Search

10/26/2024

James Little
tvasailor@gmail.com
https://littlemsandsailing.com/
Kingsport, TN

- *Retired* Research Fellow, Eastman Chem. Co.*
- *42 years experience unknown identification*
- *Now Consultant, MS Interpretation Services*
- *Specialties[1] EI GC-MS, LC-MS/MS, Chemical Ionization,[7] Accurate Mass, Derivatization,[8,9] MS library management, SciFinder,[10] Chemspider,[10] Surfactant ID,[11] NMR, GC-IR, organic synthesis, matrix ionization effects,[21] etc.*

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

>50 Mass Specs Networked
Worldwide

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

NIST EI Mass Spec Search Class

Part 0: Changes in EI NIST23 Program (V3.0)
Part I: Spectral Searches with NIST MS Search
Part II: Structure Searches with NIST Search and MS Interpreter
Part III: AMDIS (NIST) for Processing EI Mass Spectral Data Files
Part IV: Advanced NIST Hybrid Search of EI Spectra
Part V: Creating and Sharing User EI and MS/MS Libraries
Part VI: Creating and Using Retention Indices in NIST Software
Part VII: Tracking Complex Coelution with AMDIS and NIST Search

Link to Full Course

Table of Contents

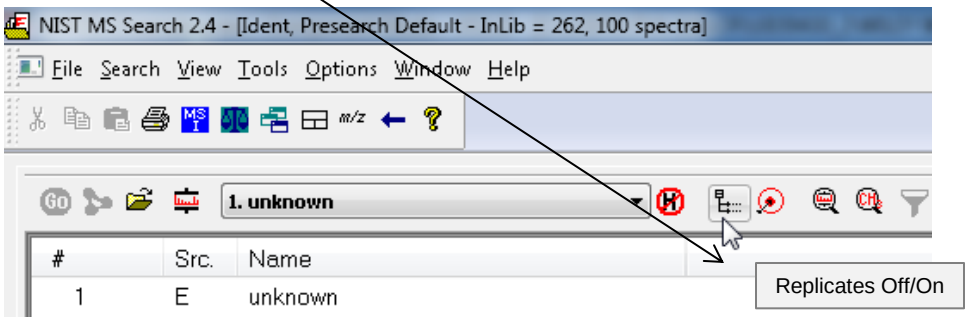
Topic	Slide #
NIST Mass Spectrometry Software and Wiley Libraries	4
Help Files for NIST Search	5
General Windows commands/functions	6
Customizing the windows	7
Tour of NIST Search Program	9
Critical parameters for searching library	18
Setting up presearch parameters	19
Primary Libraries of EI Spectra *and* Associated Structures	21
Other EI Libraries: Wiley/NIST Combined and specialty	22
Selecting the libraries used for spectral searches	23
Select the libraries used for similar structure searches	24
Selecting limits to avoid high match factors for partial spectra	25
Constraints applied to hit list after the list is determined	26
Saving setup of standard search parameters	27
Searching demonstration/***Tip*** No. Synonyms and databases	28
Reverse match for mixtures versus standard match	29
Display of related species	32
InChIKey field link to PubChem on web	33
Updating indices in older libraries	34
Library spectra *m/z* values by nominal not exact mass	35
Retention indices (Kovat) determined by users in AMDIS	37
Different derivative displays	38
Using NIST software to obtain proposed nominal MW and substructures	39
Webinar references	40
Acknowledgements	41

NIST Mass Spectrometry Software and Wiley Libraries

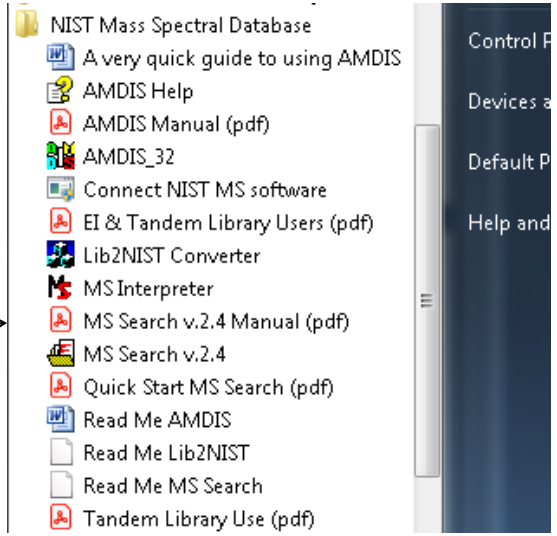
- Free software suite for processing mass spectral data
- Supplied with the purchased commercial libraries or by instrument manufacturers
- Utilized with Wiley Libraries for qualitative mass spectral analyses
- Both GC-EI and LC-MS unknown identification
- Import data[2,6] from variety of instrument manufacturers
- Searches by spectrum, structure, name, CAS No., peaks, MW, MF, etc.
- AMDIS[3,4,12] available for processing mass spectral files (*vendor and netCDF formats*)
- MS Interpreter[2,13-15] for correlating molecular substructures to ions
- Structure Export and Import using vendor Drawing Packages[2]

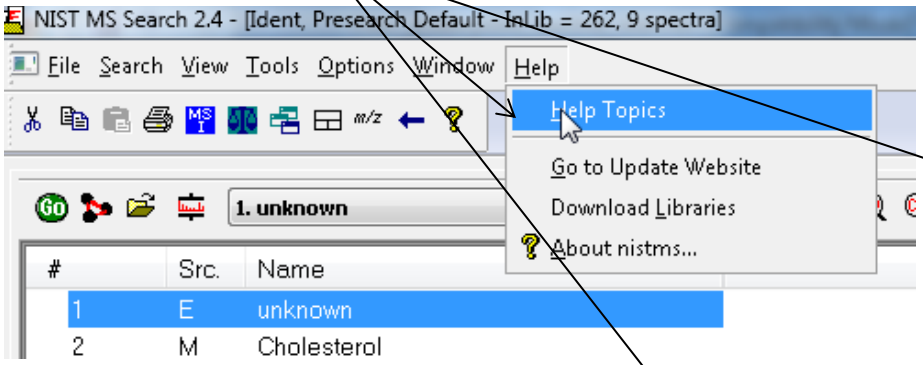
Help Files for NIST Search

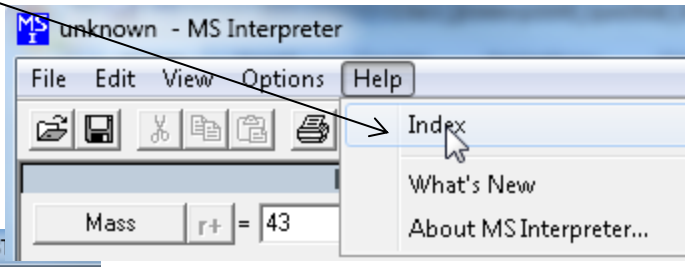
-"Hover" over Program Icon with mouse and function description displayed

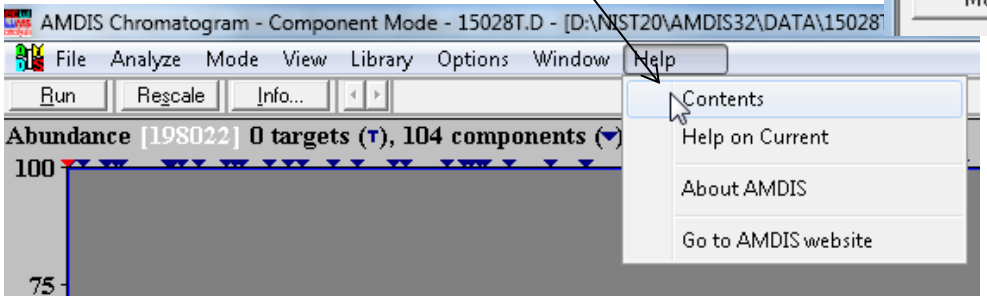


-Detailed documentation for NIST Search[2] and AMDIS[3,4,12]
-MS Interpreter included in NIST manual[2] and in posters[13-15]
-Windows Program Group
-"In program" assistance for all three programs





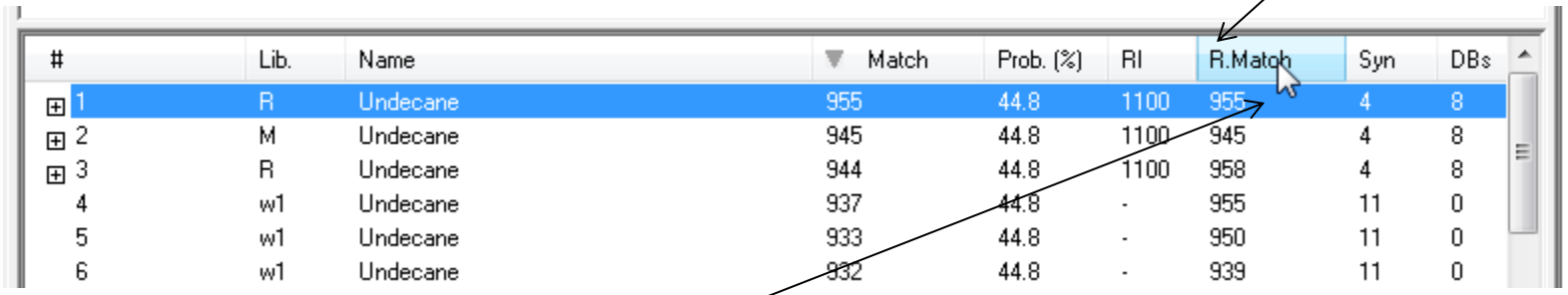




NIST Software in General is "Windows Compliant"

-***left click (LMB)*** to select an item, ***double LMB*** on that item to perform operation
-***right click (RMB)*** in area or item to see operations that can be performed or to change properties of window
-***LMB*** on first item and last item to select group ***while*** holding ***shift key***
-***LMB*** to select/deselect individual items ***while*** holding ***Ctrl button***
-use up and down arrows ***on keyboard*** to step between entries
-some NIST windows such as librarian have no delete button to delete ions, ***must*** use delete key on keyboard!
-***control a*** (select all), ***control x*** (delete selected), ***control c*** (copy); ***control v*** (paste)
-***control k*** copies entries into windows in tab-separated text format, e.g., ***paste*** into Excel
-F1 MS Search help
-F9 send spectrum to MS Interpreter
-***LMB*** and ***zoom*** mass spectral windows, ***RMB*** then ***LMB*** to ***zoom out***

Tip 3: ***LMB*** and **drag** to rearrange order of column headers



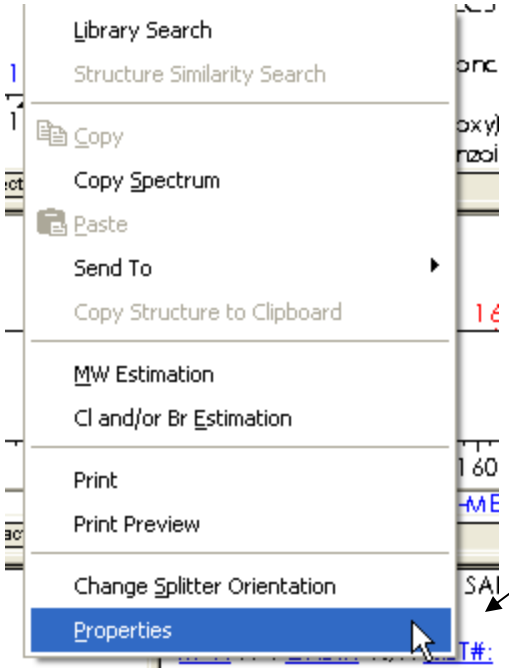
#	Lib.	Name	Match	Prob. (%)	RI	R.Match	Syn	DBs
1	R	Undecane	955	44.8	1100	955	4	8
2	M	Undecane	945	44.8	1100	945	4	8
3	R	Undecane	944	44.8	1100	958	4	8
4	w1	Undecane	937	44.8	-	955	11	0
5	w1	Undecane	933	44.8	-	950	11	0
6	w1	Undecane	932	44.8	-	939	11	0

-***LMB*** on column of interest
-Can sort in lower value first or higher
-Will show use in mixtures in example later in presentation

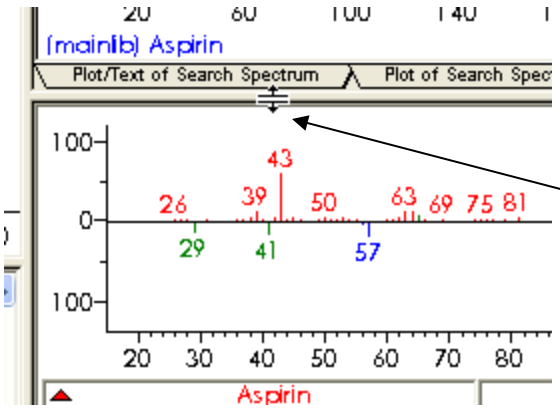
Tip 1: When reviewing search results, use up and down arrows on keyboard to quickly step through results!

Tip 2: When viewing structures in MS Interpreter, use left and right arrows on keyboard to quickly review results!

Customizing the Windows



-***RMB*** within any window of interest and select properties to change window information and/or layout

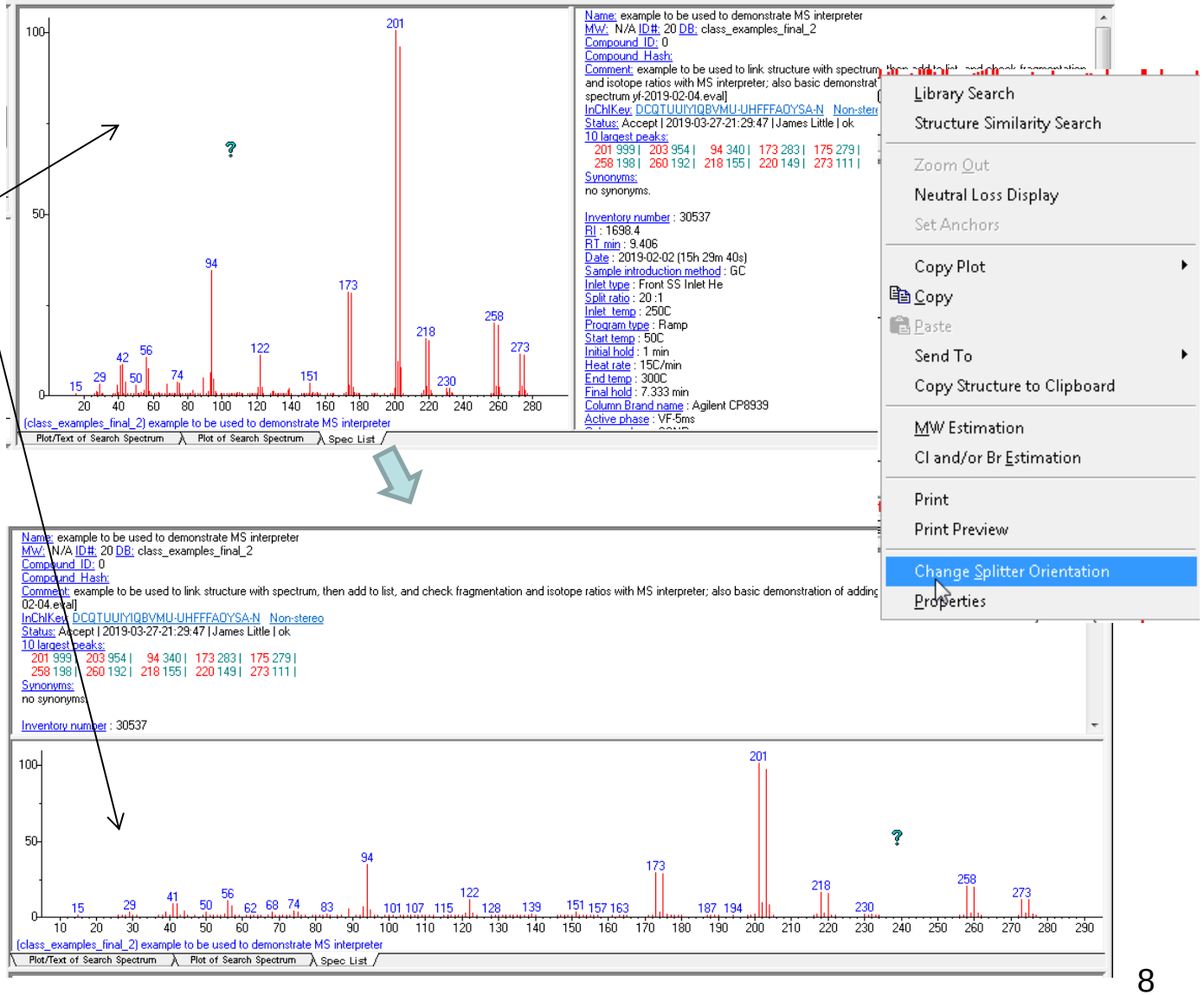


-place cursor over any bar between (***top or side***) windows and then ***LMB*** and drag to change the size of a window or make window so small it essentially disappears.

Customizing the Windows (continued)

-RMB in display windows then ***LMB*** to “Change Splitter Orientation”

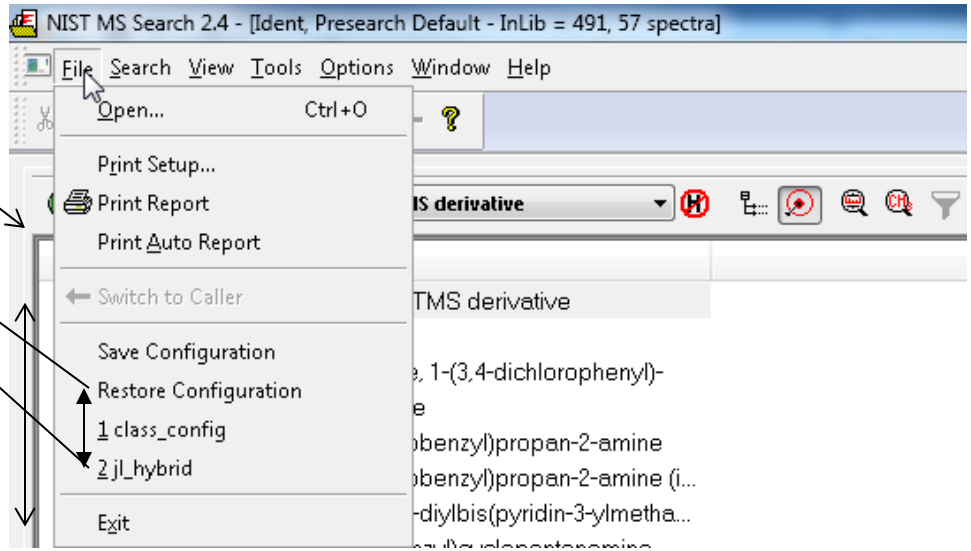
Tip: ***All*** mass spectra in windows can be ***“Boxed”*** to ***expand*** by ***LMB*** and dragging to expand/enlarge; ***restore*** to original by ***RMB*** in spectrum region and selecting ***“Zoom Out”***



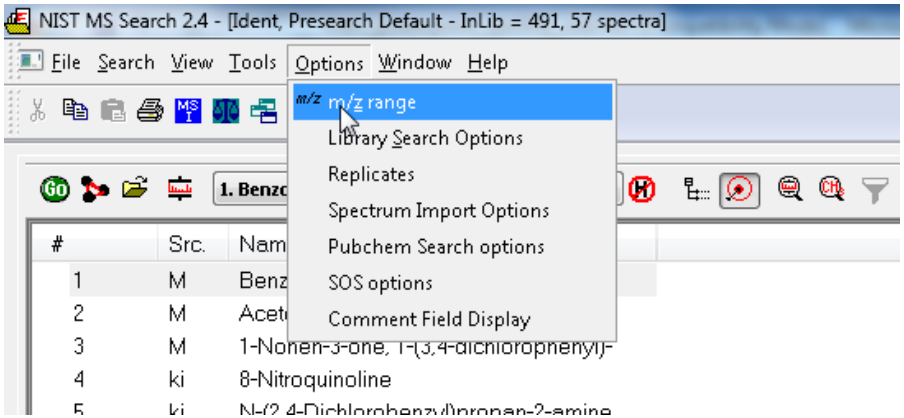
Tour of NIST Search Program
(Top Links)

-Restore and Create user Configurations

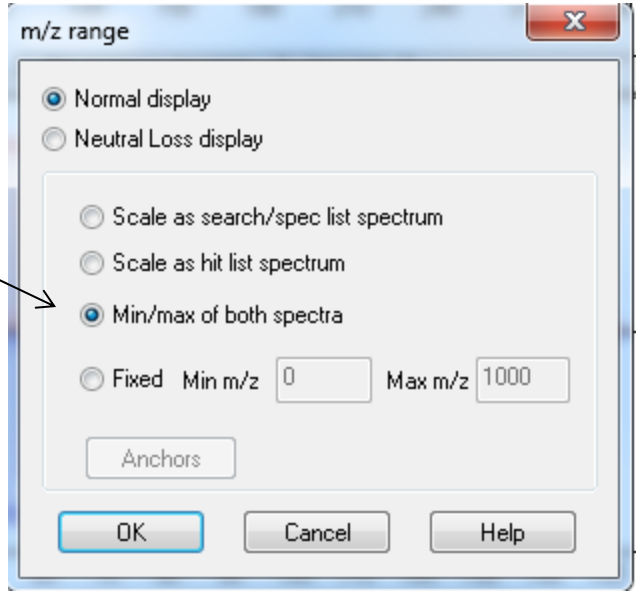
- Most used listed



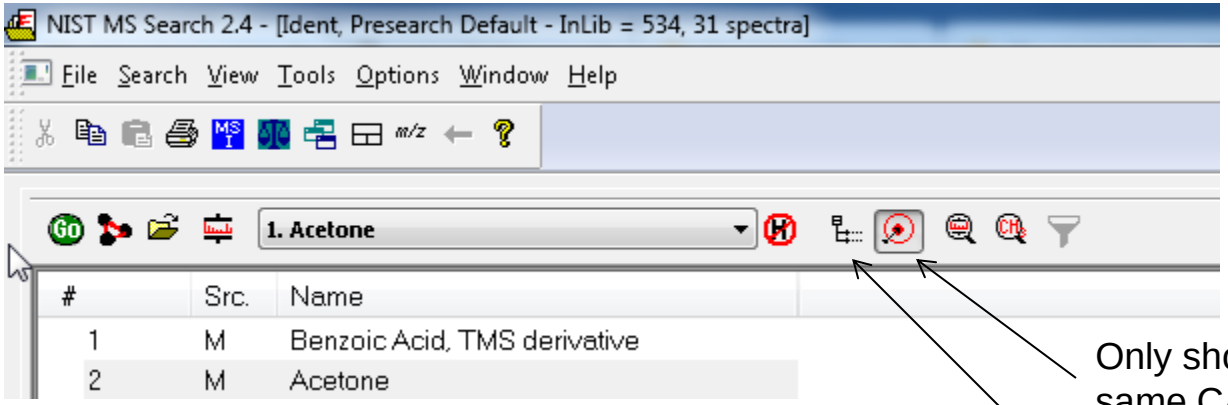
Tour of NIST Search Program (continued)
(Top Links)



-I prefer min/max of both spectra



Tour of NIST Search Program (continued)
(Top Links)



Tip: When reviewing results, use up and down arrows on keyboard to quickly review result

Only shows best hit with same CAS number

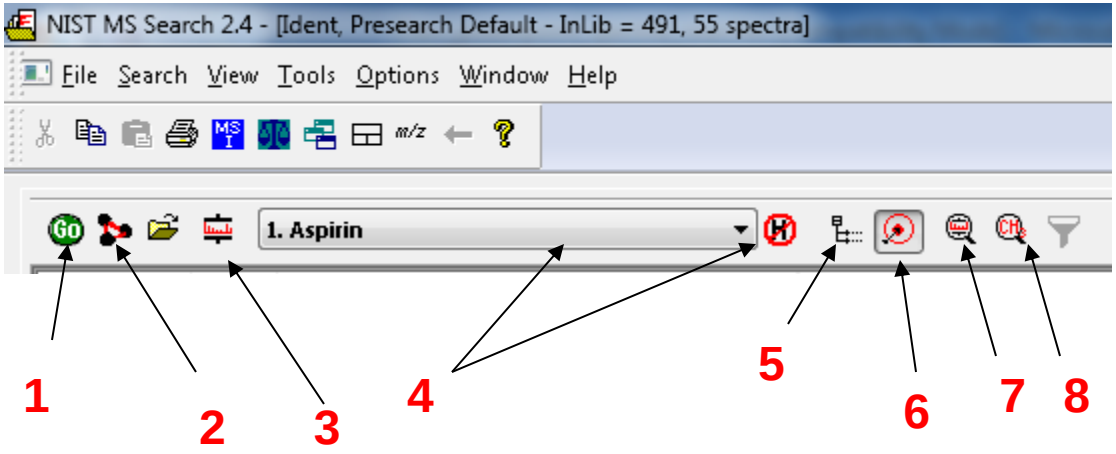
Turns on or off replicates in search results shown below

Demonstrated by Search Results for Acetone:

-Can toggle viewing replicates by "left-clicking" on + or – box here ***or*** icon on toolbar shown above

Tour of NIST Search Program (continued)

Icons in Lib Search Window

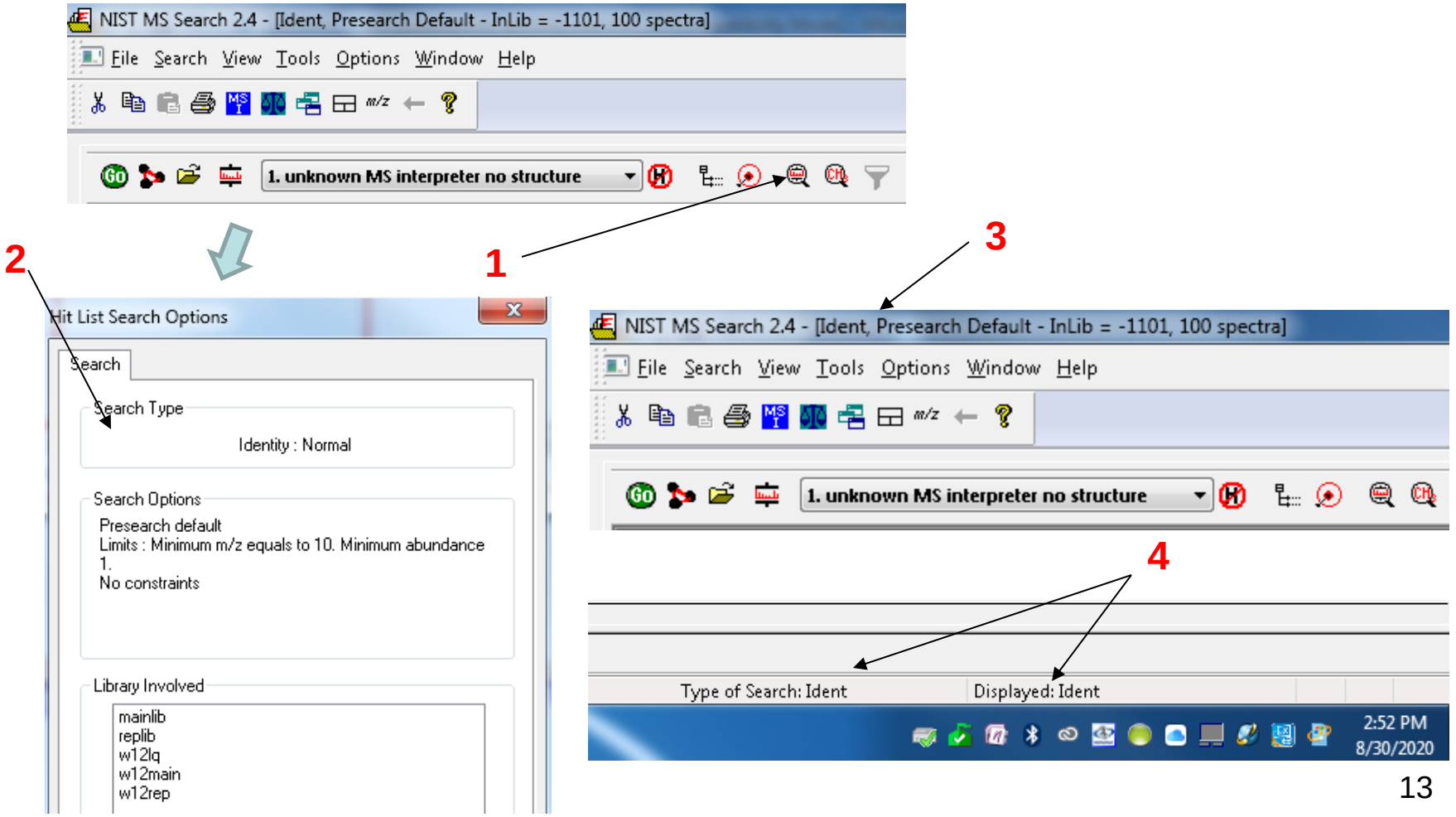


1. ***LMB*** to start search or ***double LMB*** on entry in spec list window
2. ***LMB*** to do structure search
3. ***Critical*** search criteria for structure and spectra search
4. Search results stored and ability to clear list
5. ***"Show/Don't Show"*** replicate entries in NIST library
6. Only show best hit of entry with same CAS number, minimizes looking at redundant entries in search window
7. View Hit List Search Options (*see details next slide*)
8. ***After*** search, shows sub-structural information based on search results

Tour of NIST Search Program

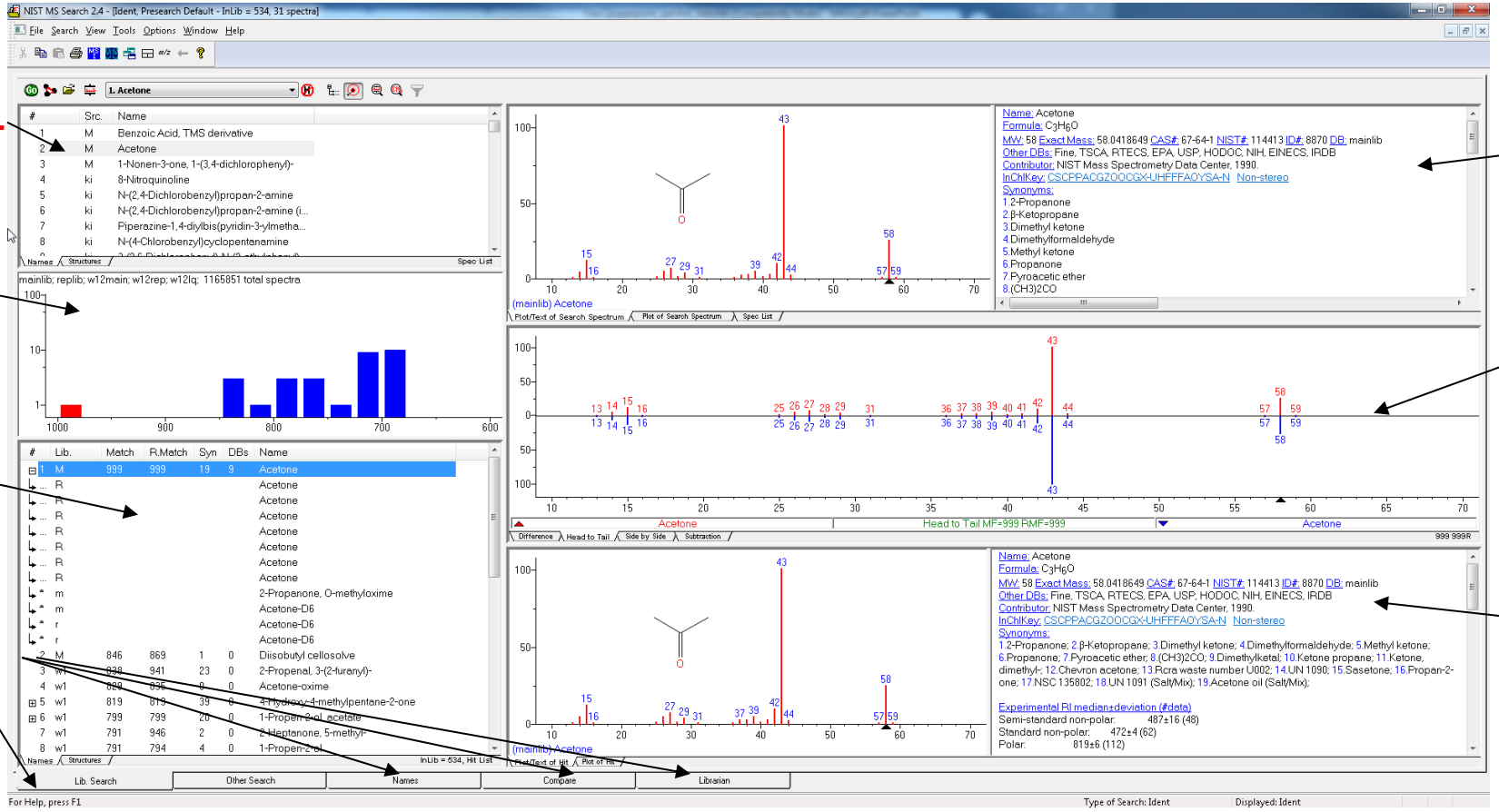
View of Hit List Normal Search Options *Utilized*

1. ***LMB*** the "View Hit List Search Options"
2. See details of the last search performed
3. Also, an abbreviated description of the last search noted at ***top*** of NIST search window
4. Furthermore, at the ***bottom*** of the screen, *respectively* displayed, are the types of searches that will be performed and the type of results currently displayed



Tour of NIST Search Program (continued)

(General Window in Lib Search View)

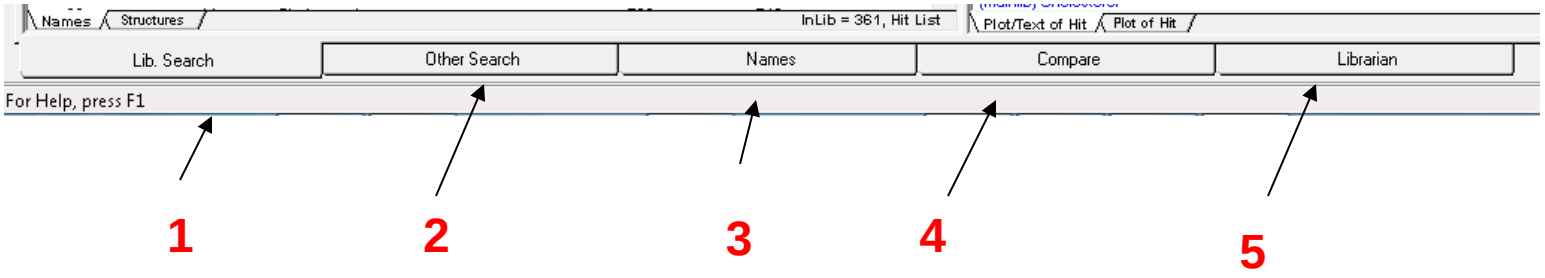


1. Spec list window for import of spectra and structures from other programs
2. Histogram, Statistics on search
3. Hits list, step through by ***LMB*** then ***up*** and ***down arrows on keyboard***
4. Unknown spectra and info
5. Comparison of unknown to selected hit result, ***many different display options*** with tabs at bottom left of window
6. Spectrum of hit and other associated information
7. Accessing other windows and associated functions

Tour of NIST Search Program (continued)

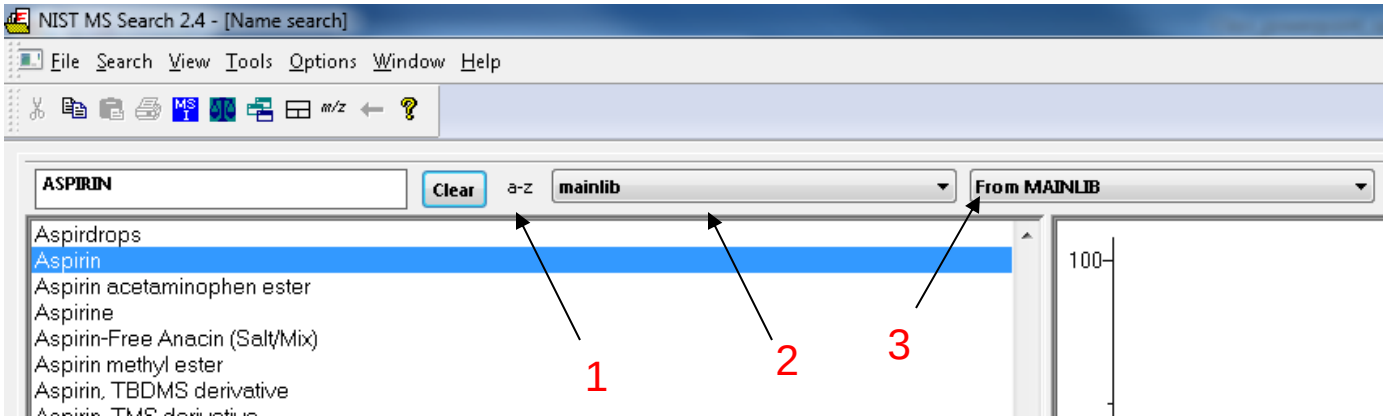
Tabs for Other Functions Accessed at Bottom of Main Library Page

(Many Detailed Discussions in Future Webinars)

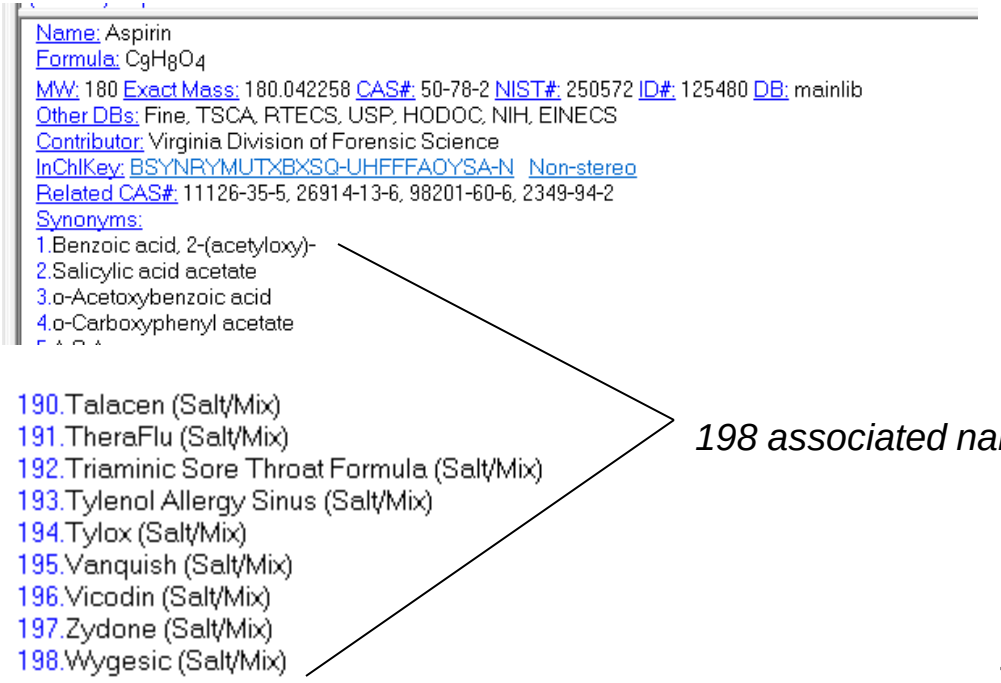


1. Lib Search-main window for searching spectra and structures
2. Other Search-search by CAS, MW, ID no., partial name, MF, etc.
3. Names-search by name, often 2-100 names for same compounds, e.g., check out *aspirin on next slide*
4. Compare-Window used to compare spectra, also can display best hits from search
5. Librarian-window used to edit spectra, correlate spectra with structure, create user libraries, type in spectra manually, etc.

Tour of NIST Search Program (continued)
Name Search Tab



1. *Toggle: letters only or letters/numbers*
2. *Search **one library** at a time*
3. *Replicates,derivatives,isotopes, stereoisomers in pull-down menu*



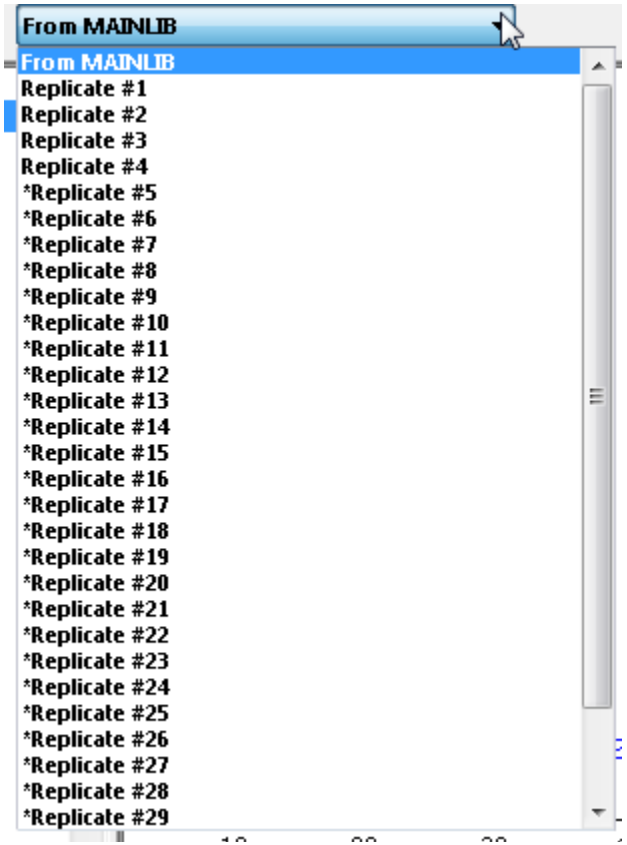
198 associated names!!

Tip: Difficult to see components in pull down menu, thus ***RMB*** on "Aspirin" in name list, then *either* **Send To*/* Spec List** and search *or* **Lib Search** directly; See improved display in Library tab on next page

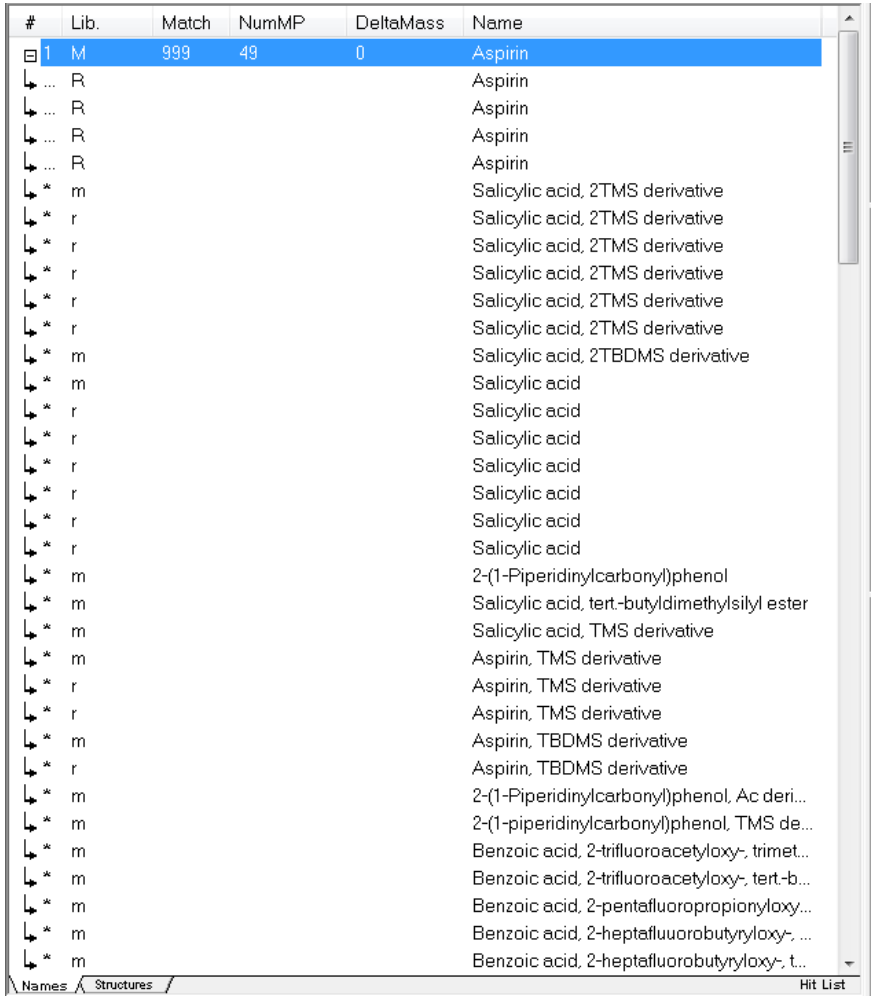
Tour of NIST Search Program (continued)
Improved View in Library Tab Compared to Name Tab

-***increase*** from just displaying replicates in both views to ***including*** isotopes, stereoisomers, and derivatives as explained in ***slide 32***

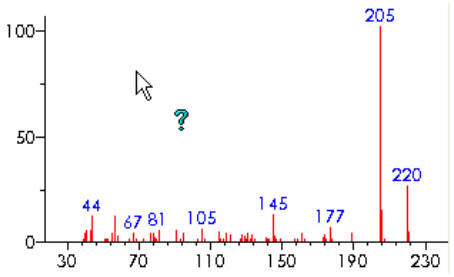
Limited view for aspirin in Name Tab

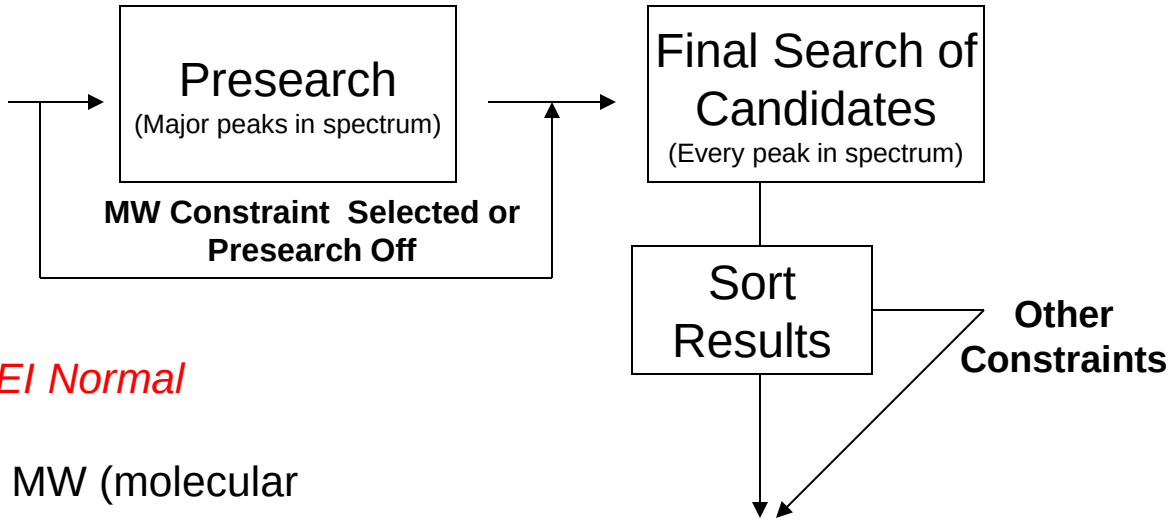


Improved view after sending to Spec List and searching in library window



Critical Parameters for Searching Libraries





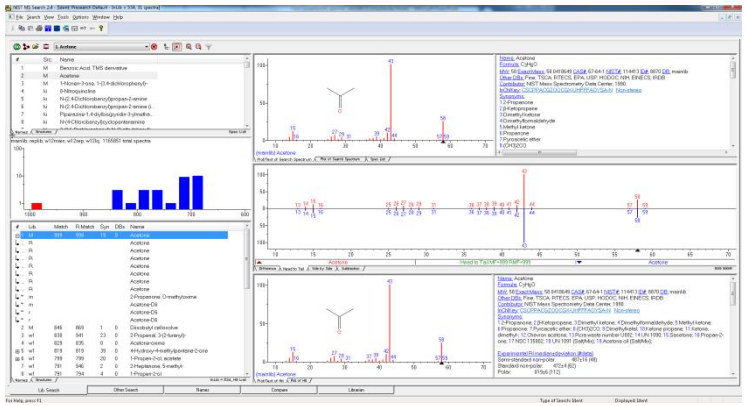
Spectrum Search Type-*Identity/EI Normal*

MW-skips presearch and limits to MW (molecular weight) specified

Presearch Default*-best setting for optimum results*

Presearch off-skips presearch, >1 minute per search

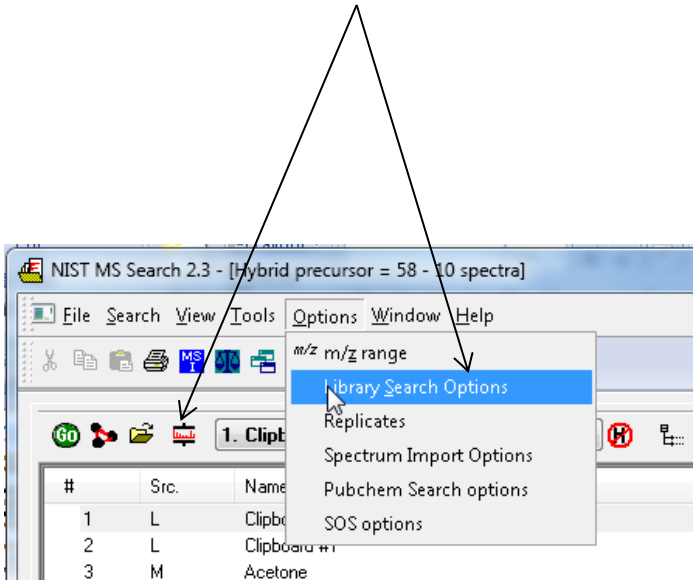
Constraints-results filtered ***after*** Final Search

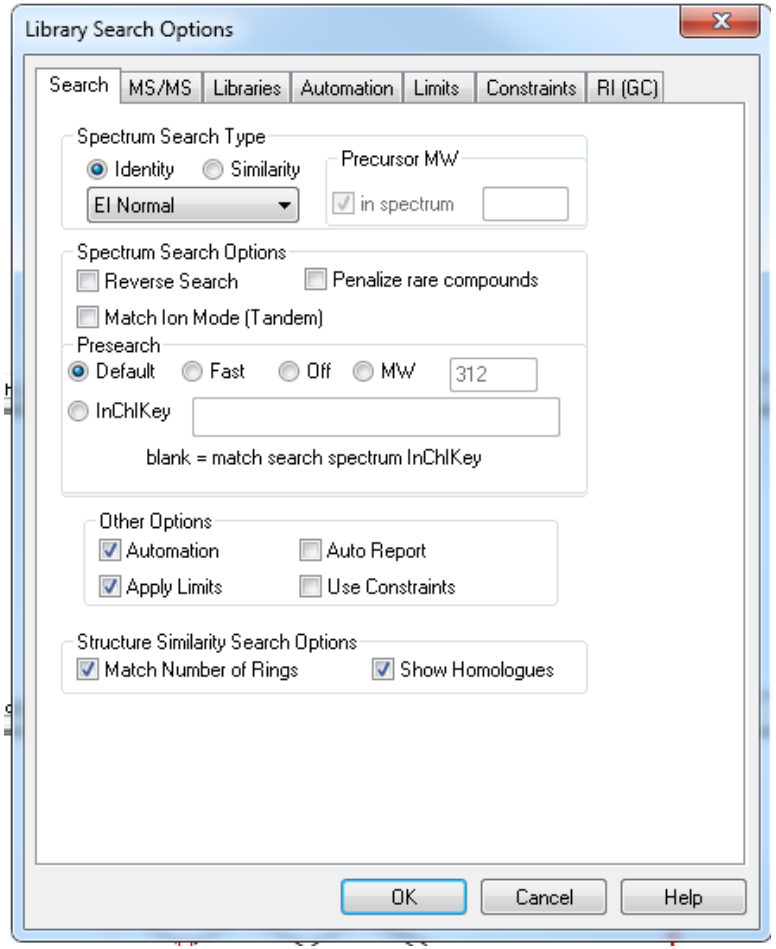


Setting Up Presearch Parameters

Critical Step

Two ways to access:

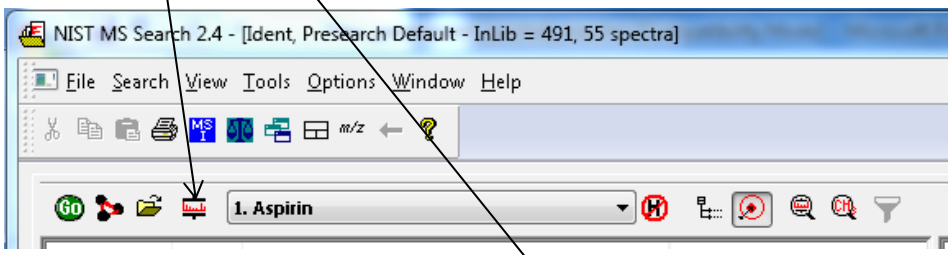




Setting Up Presearch Parameters

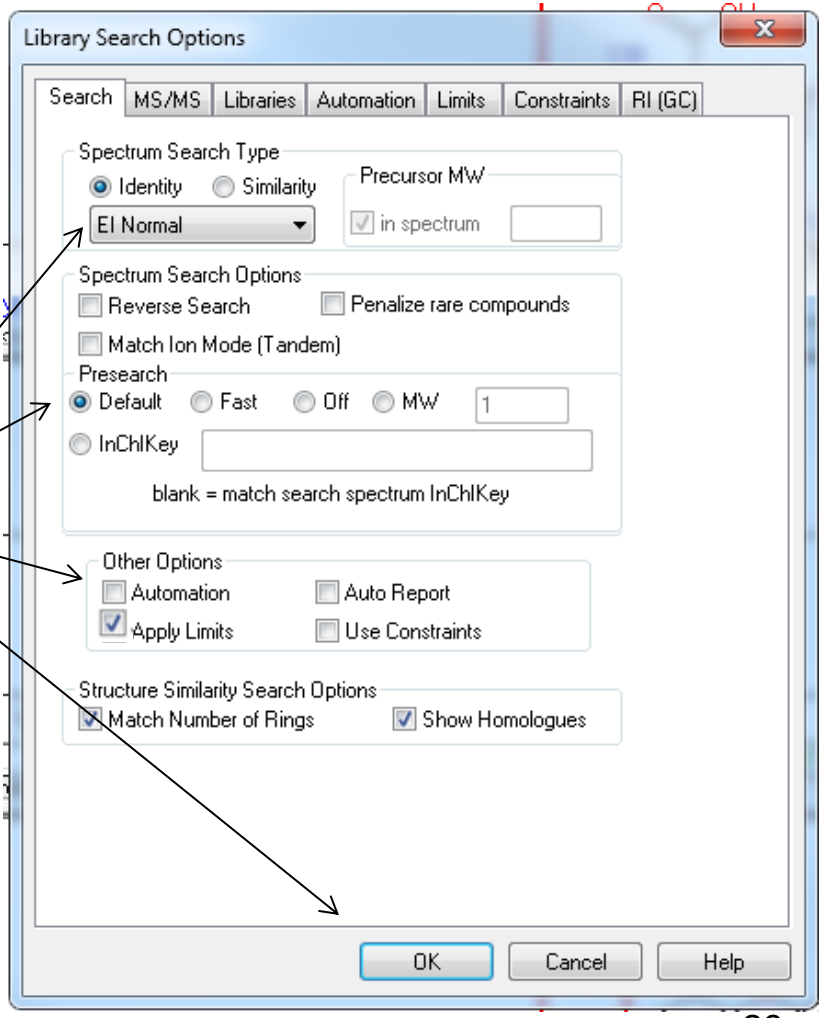
Critical Step

- ***LMB*** on icon to open ***Library Search Options*** window
- Then select appropriate options with ***LMB***



then…

1
2
3
4



Note: Previous to Version 2.4 (2020), EI Normal was just "Normal"**!**

Primary Libraries of EI Spectra ***and*** Associated Structures

Wiley: >815k entries

NIST: >350K entries

User Libraries: *e.g.* Eastman, >50K entries, automatically updated *nightly*[5]

Which Ones to Search*? (My opinion, all >1,100K* ***excluding*** *w12leg)*

-**w12main**: best spectra for component
-**w12rep:** replicates (spectra can be instrument dependent)
-**w12lq:** lower quality, <4 ions per spectrum
-**w12leg**: spectra once present in main and rep, but removed for various reasons
-**main (NIST):** one best entry selected (subjective) for component
-**rep (NIST):** replicates for main spectrum
-**user libraries**: individual's or company's personal libraries

Other EI Libraries: Wiley/NIST Combined and Specialty

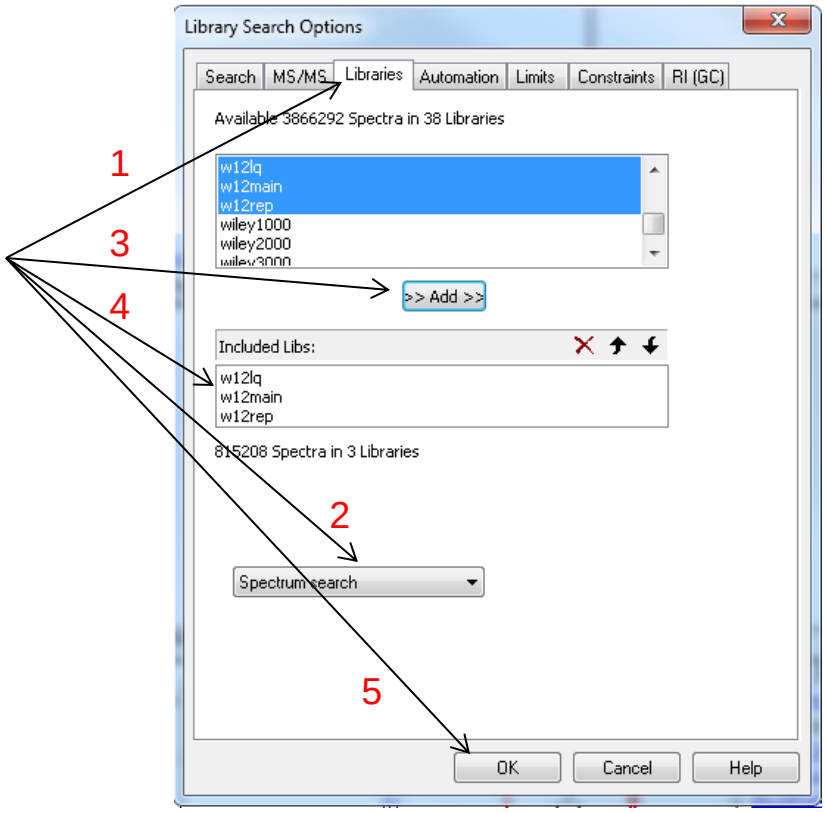
Combined Wiley/NIST (>1.1 M spectra)

Wiley Specialty (>91 K spectra)**:**

- Lipids (430)
- Pesticides from Prof. Mondello (1,300)
- Designer Drugs (28,032)
- Fragrances (3,462)
- Drugs, Poisons, Pesticides, Pollutants, and Their Metabolites (10,430)
- FAMES: Fatty Acid Methyl Esters (240)
- Physiologically Active Substances: Drugs, Steroid Hormones, and Endocrine Disruptors (4,182)
- Pesticides from Rolf Kuhnle (1,238)
- Androgens, Estrogens, Steroids and Derivatives (3,722)
- Organic Compounds for Drug Discovery/Combinatorial Synthesis (37,055)
- Petrochemicals and Biomarkers (1,100)
- Online Search Wiley and NIST Libraries (Check for Inclusion, free) https://www.sisweb.com/software/ms/nistsearch.htm

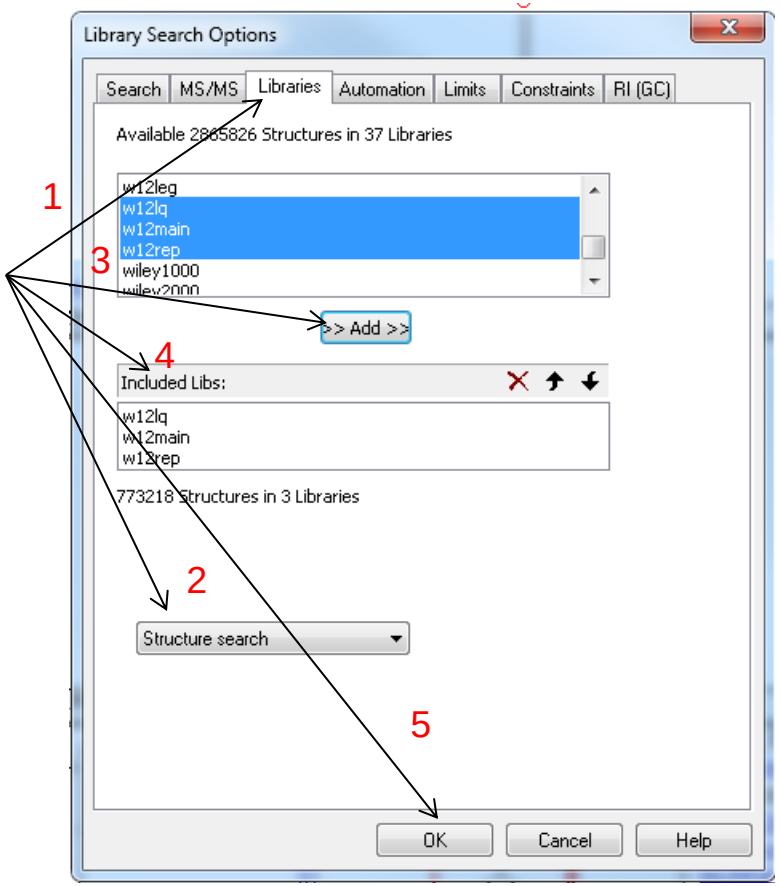
Select the Libraries Used for Spectral Searches

- ***LMB*** on the ***Libraries*** tab and make sure ***Spectrum Search*** is selected
- Select the group of libraries to be searched
- ***>>Add>>*** all libraries to be searched
- ***Order*** of libraries normally ***only important*** in Other Search Tab which return ***maximum no.*** of hits



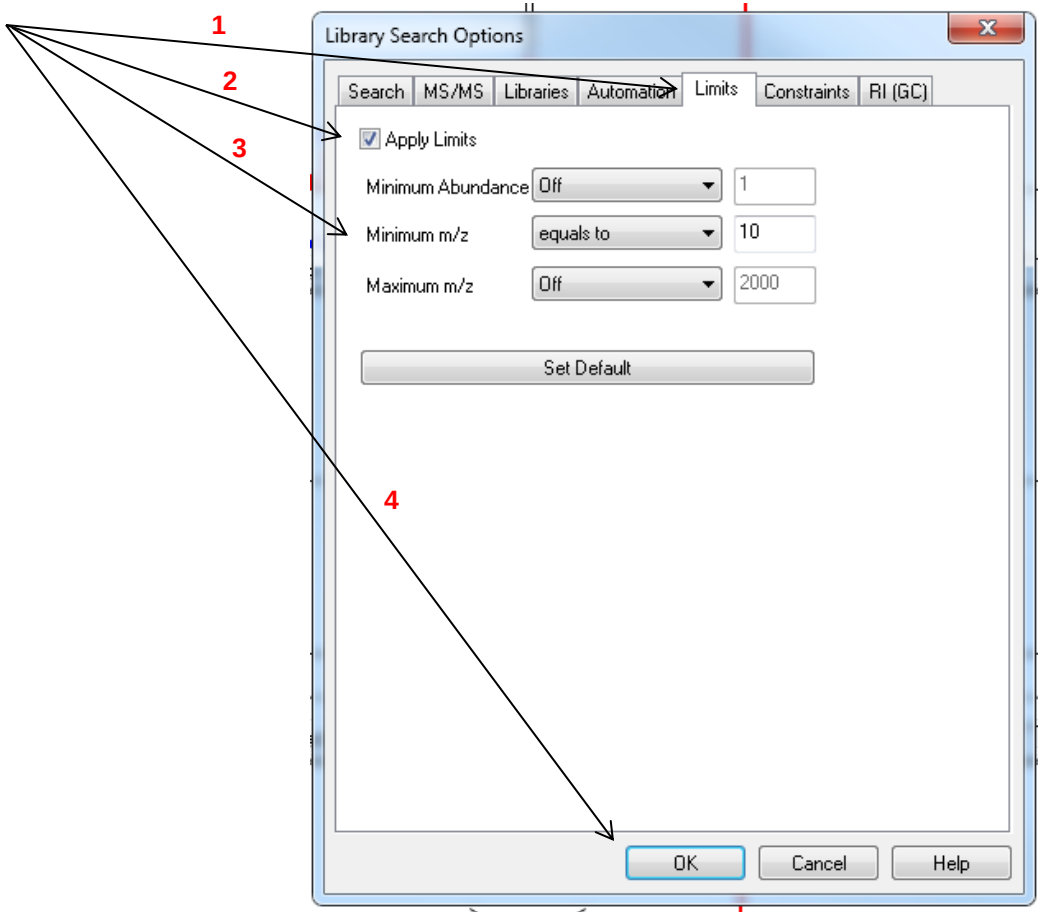
Select the Libraries Used for *Similar Structure* Searches

- ***Similar structure searches*** are driven by libraries in ***Structure Search***
- ***LMB*** on the ***Libraries*** tab and make sure ***Structure Search*** is selected
- Select the group of libraries to be searched by Similar Structure
- ***>>Add>>*** all libraries to be searched
- ***Order*** of libraries normally ***only important*** in Other Search Tab which return ***maximum no.*** of hits

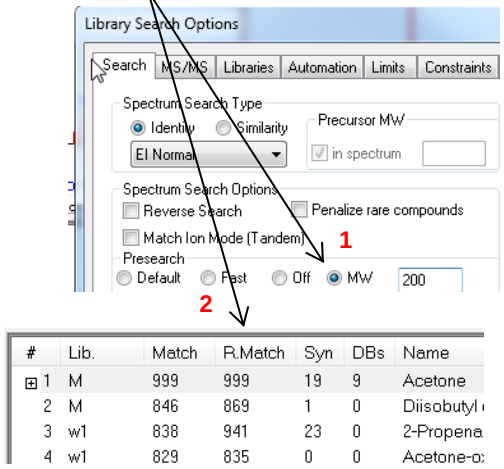


Selecting Limits to Avoid High Match Factors for Partial Spectra in Libraries

- ***LMB*** on the ***Limits*** tab
- Select ***Apply Limits***
- Enter the value for “Minimum m/z”
- This is necessary to force the search to compare the whole spectrum when calculating a match factor
- Library spectra with only a few ions ***are minimized*** in the search results by using limits



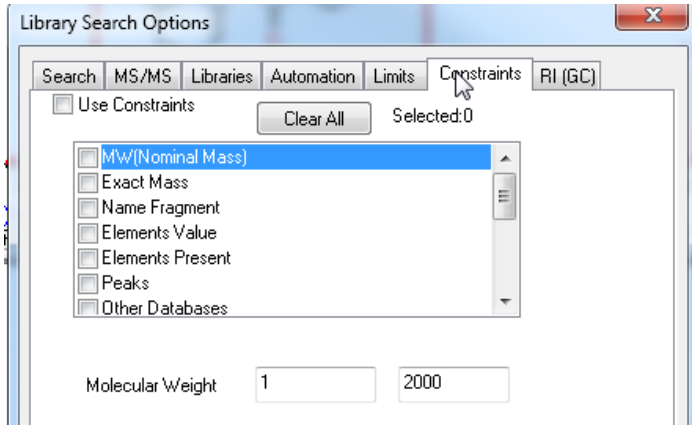
Tip: The partial spectra in libraries would be hard to find. One way to accomplish is to do a normal search limiting the MW. This will skip the presearch. Then ***sort*** the results by clicking on Rmatch to sort by “Reverse Search.”



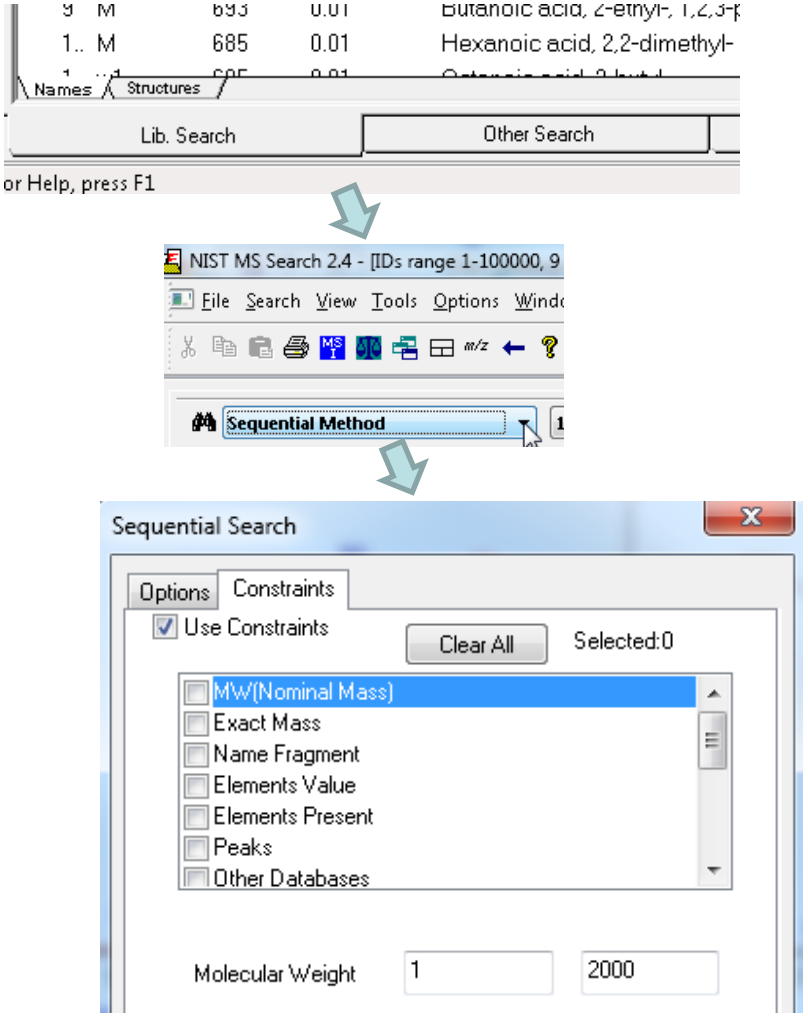
Constraints Applied to Hit List *After* the List Is Determined

1. I very seldom use these to refine the search results
2. However, very useful when finding spectra in the libraries using ***Sequential Method*** under ***Other Search*** Tab

1

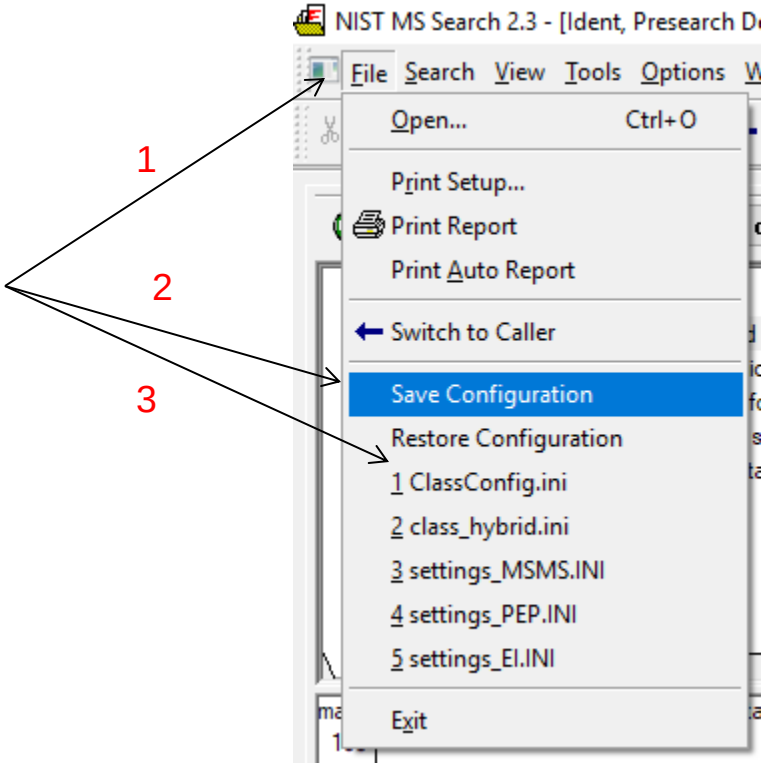


2



Saving Setup of Standard Search Parameters

- After all the parameters are setup for search, save them for future use
- Example setup below was named ***ClassConfig***
- In the future, this can be “re-called” by selecting it from this menu



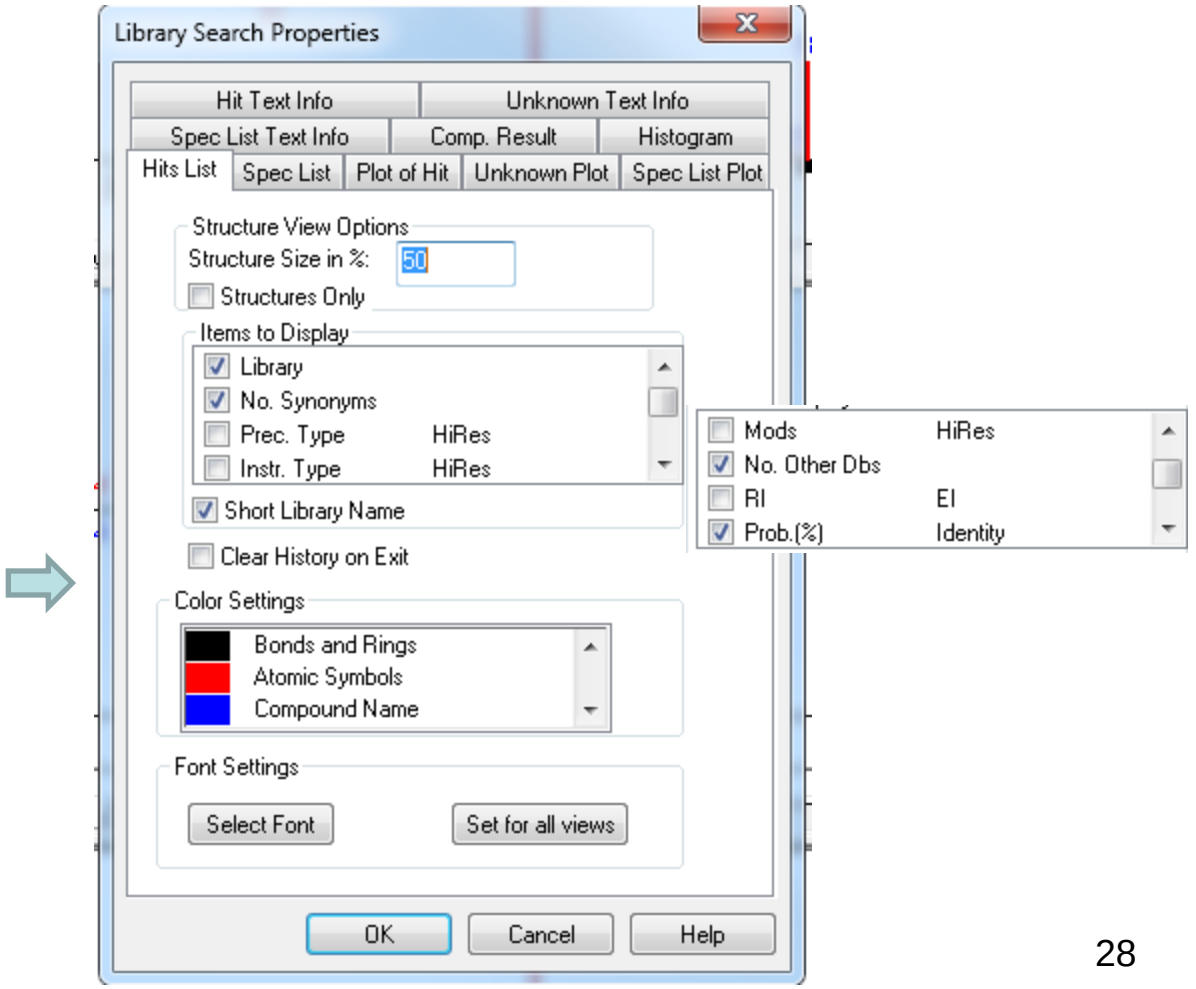
Searching Demonstration

-Example searches demonstrated in live presentation

Tip1: When reviewing results, use up and down arrows on keyboard to quickly review results!

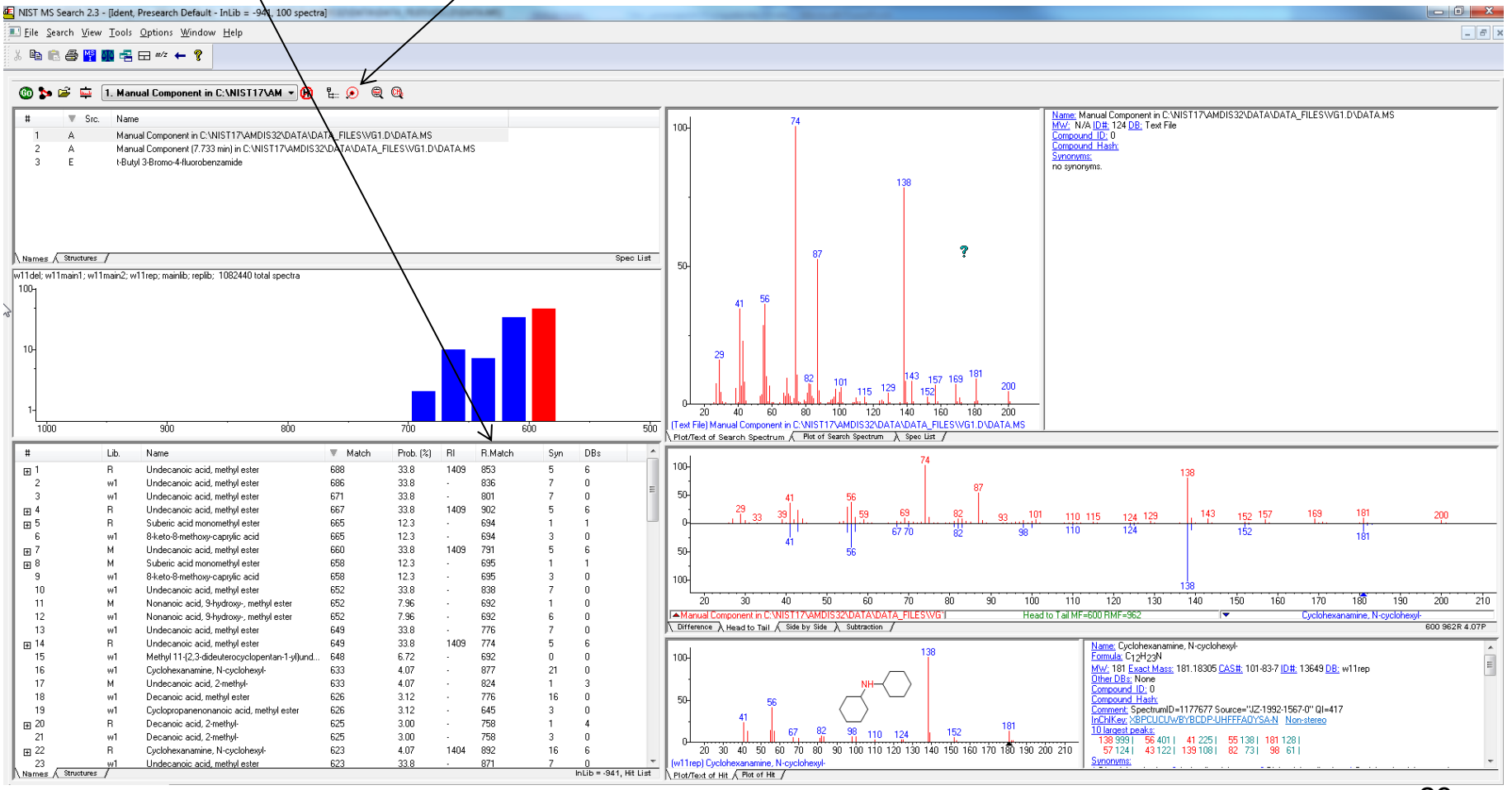
Tip2: When viewing results of search, consider adding **No. Synonyms** and **No. of Other Databases** to columns displayed.

The number of associated synonyms and databases makes a candidate more likely to be correct structure in many cases!



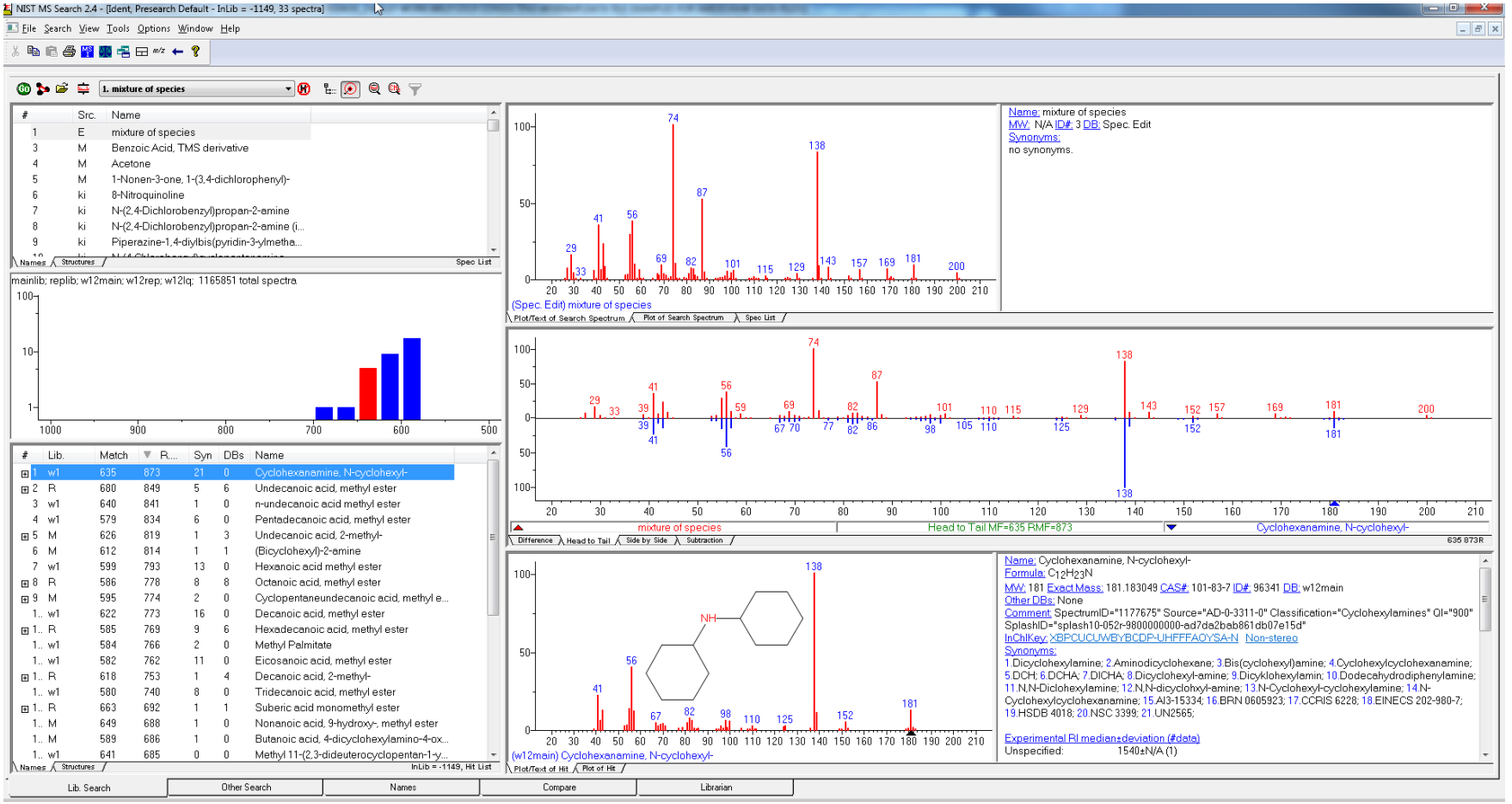
Reverse Match for Mixtures Versus Standard Match

-Also “Head to Head” display of unknown versus best hit shows many additional ions in unknown
-can ***LMB*** on R. Match (Reverse Header) to resort results
-Reverse match factor is calculated ***disregarding*** any peaks in the unknown that ***not*** in the library spectrum
-Make sure “Best Matching Only” is ***not selected*** on toolbar!



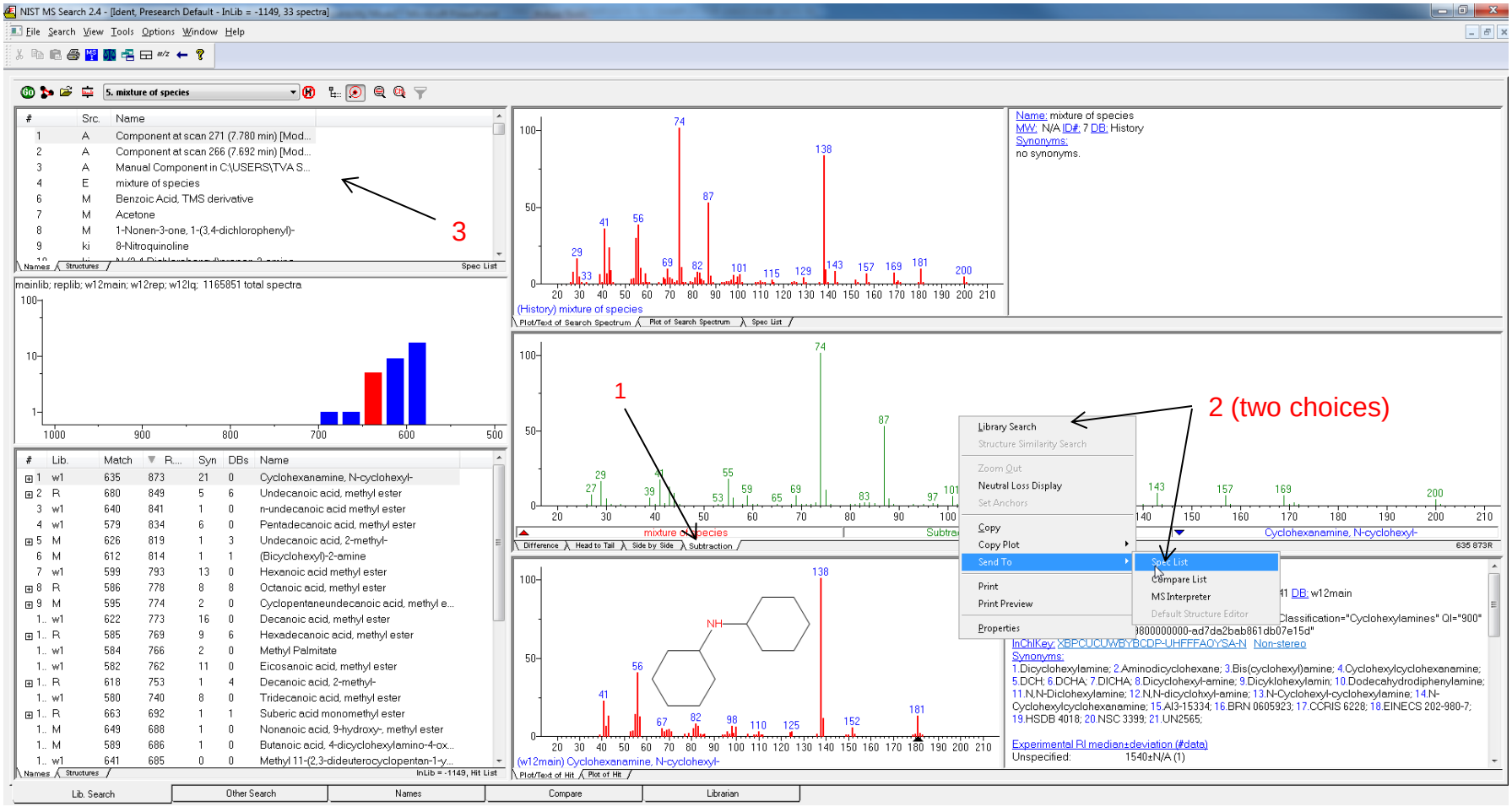
Reverse Match for Mixtures Versus Standard Match (continued)

-The top Reverse match is 873 for the dicyclohexylamine
-The next highest is 849 for methyl undecanoate
-These are the correct identities for the components in the mixture

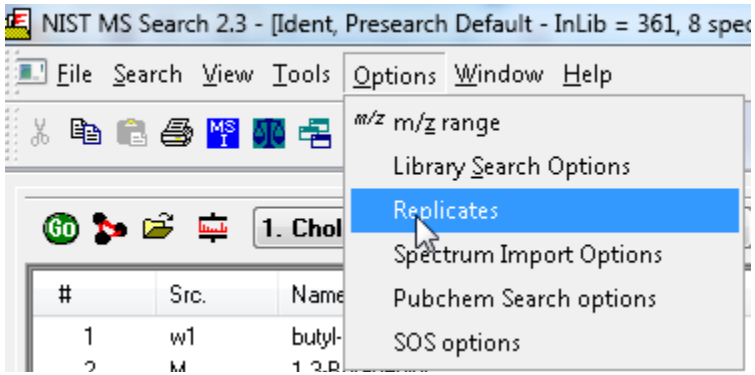


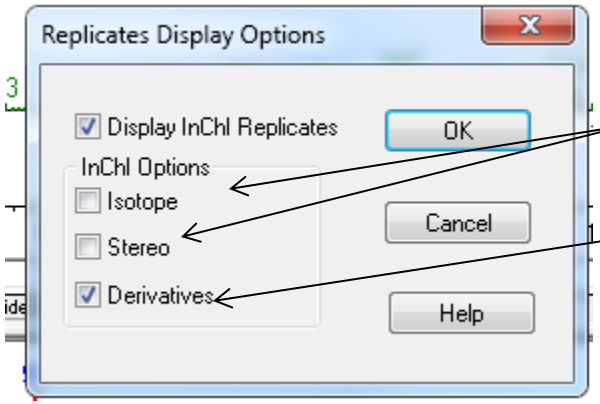
Reverse Match for Mixtures Versus Standard Match (continued)

-One can switch the display of unknown versus the selected hit to the “subtraction” mode (1)
-Thus the dicyclohexylamine result spectrum is *subtracted* from the unknown spectrum
-The difference or subtracted spectrum, using a “right-click,” can then be “Send to” “Spec List” window ***or*** Library Search (2)
-This subtracted spectrum can then be searched again (3) to yield a good “Match” for methyl undecanoate from SpecList



Display of Related Species





Inconsistent!

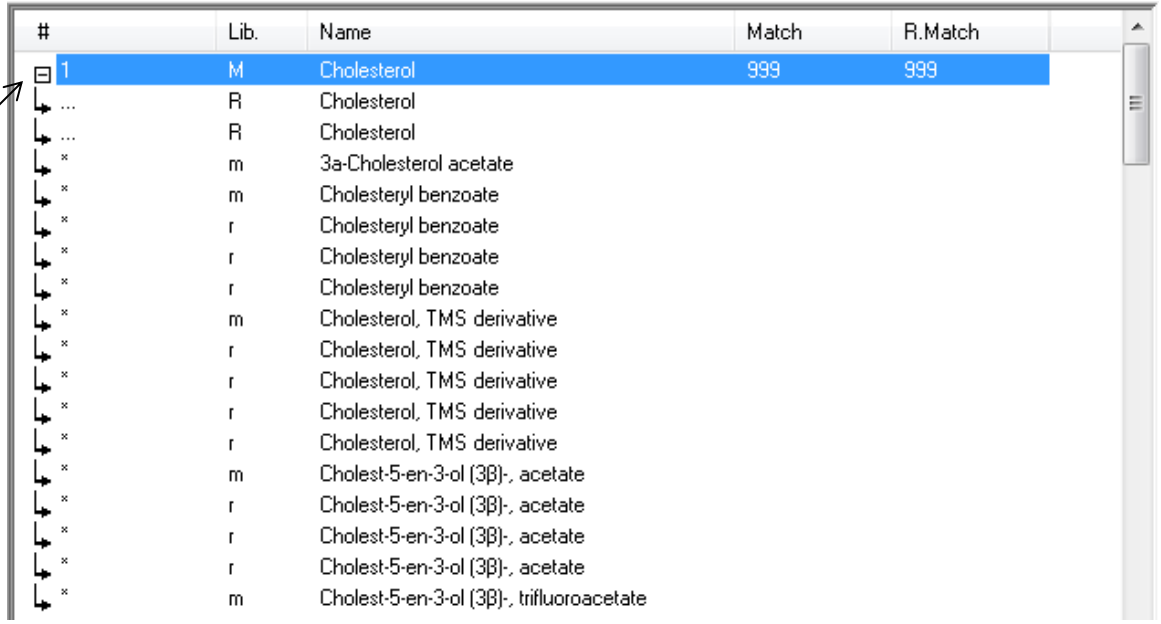
"Do see" ***unchecked***

"Do see" ***checked***

-Determines if replicates, isotopically labeled species, stereoisomers, and/or derivatives displayed with search results

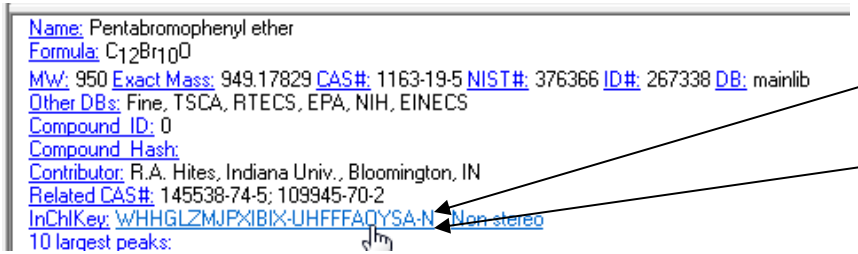
-Can toggle by "left-clicking" on + or – box

-***Also*** *affects Names Search Display*



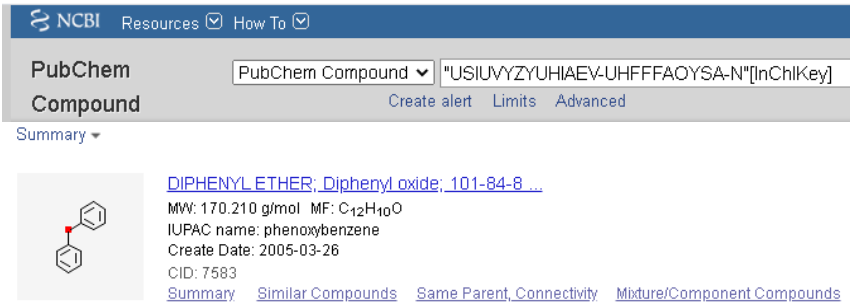
#	Lib.	Name	Match	R.Match
1	M	Cholesterol	999	999
...	R	Cholesterol		
...	R	Cholesterol		
*	m	3a-Cholesterol acetate		
*	m	Cholesteryl benzoate		
*	r	Cholesteryl benzoate		
*	r	Cholesteryl benzoate		
*	r	Cholesteryl benzoate		
*	m	Cholesterol, TMS derivative		
*	r	Cholesterol, TMS derivative		
*	r	Cholesterol, TMS derivative		
*	r	Cholesterol, TMS derivative		
*	r	Cholesterol, TMS derivative		
*	m	Cholest-5-en-3-ol (3β)-, acetate		
*	r	Cholest-5-en-3-ol (3β)-, acetate		
*	r	Cholest-5-en-3-ol (3β)-, acetate		
*	r	Cholest-5-en-3-ol (3β)-, acetate		
*	m	Cholest-5-en-3-ol (3β)-, trifluoroacetate		

InChIKey Field Link to PubChem on Web

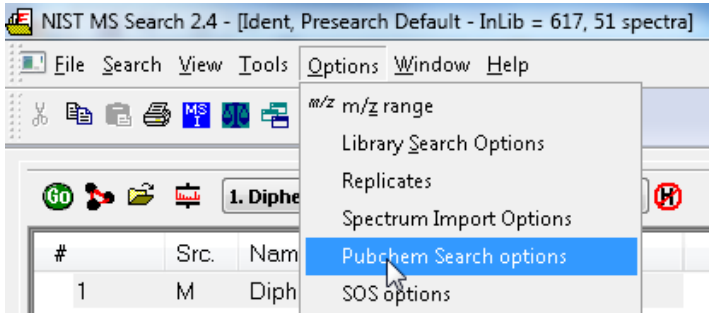


-InChIKey is ASCII “hashed” representation of structure* and is searchable

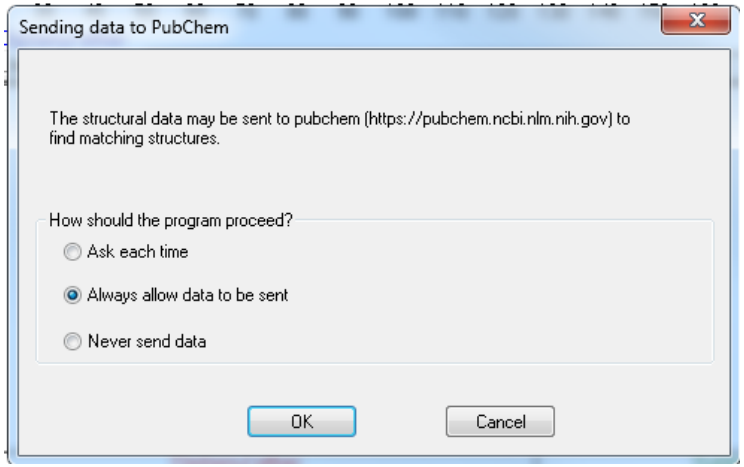
-Double ***LMB*** on ***any*** InChIKey in NIST software and you will be taken to PubChem Web Page on internet



Selecting PubChem Options:

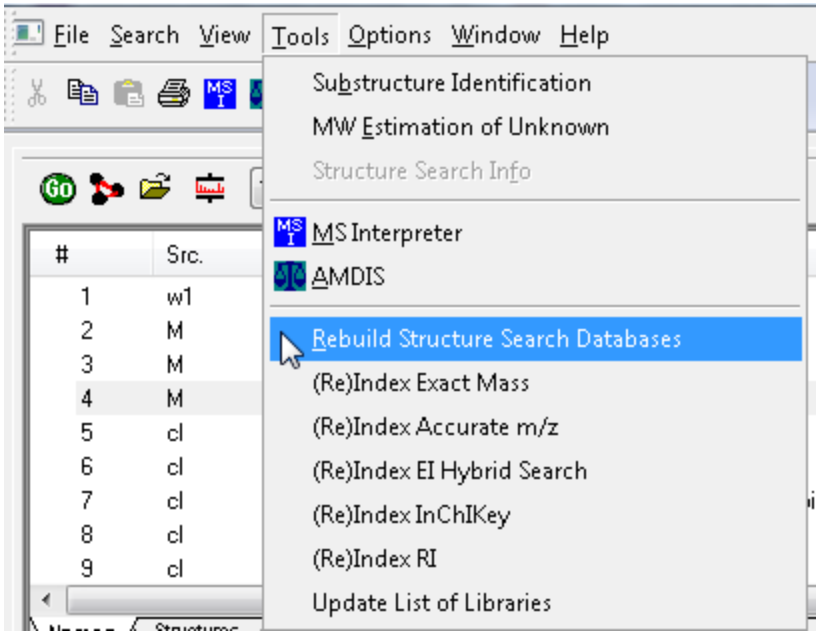


* https://en.wikipedia.org/wiki/International_Chemical_Identifier

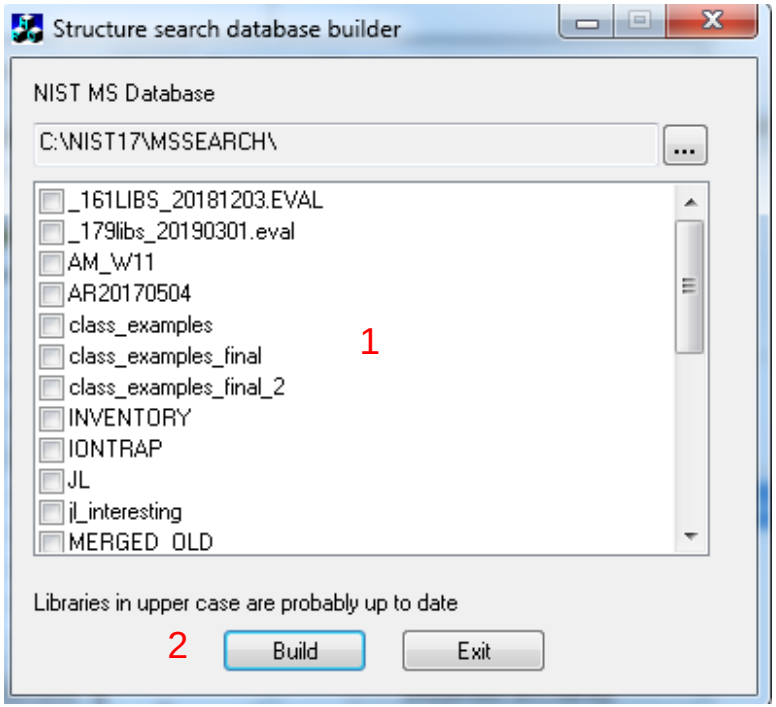


Updating Indices in Older Libraries

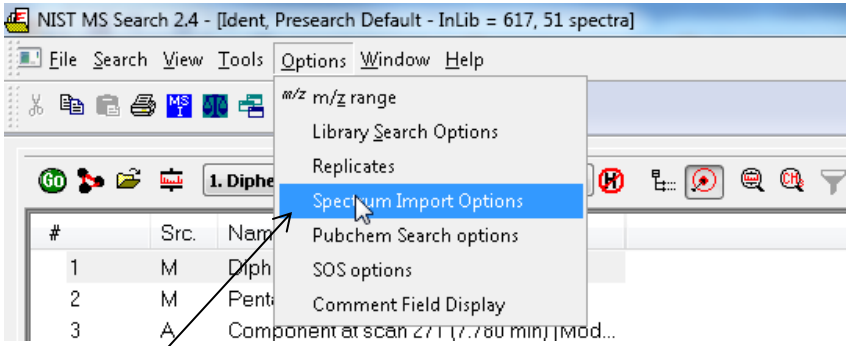
-Upgrades older versions of libraries to newer functions such as hybrid search, structures InChIKey, etc.



e.g. Updating indexes for a library
Ones in lower case probably ***not*** up to date, 1) select, 2) then Build



Library Spectra *m/z* Values by Nominal ***not*** Exact Mass



-Select “Spectrum Import Options”

-Spectra in NIST library are corrected to nominal mass*

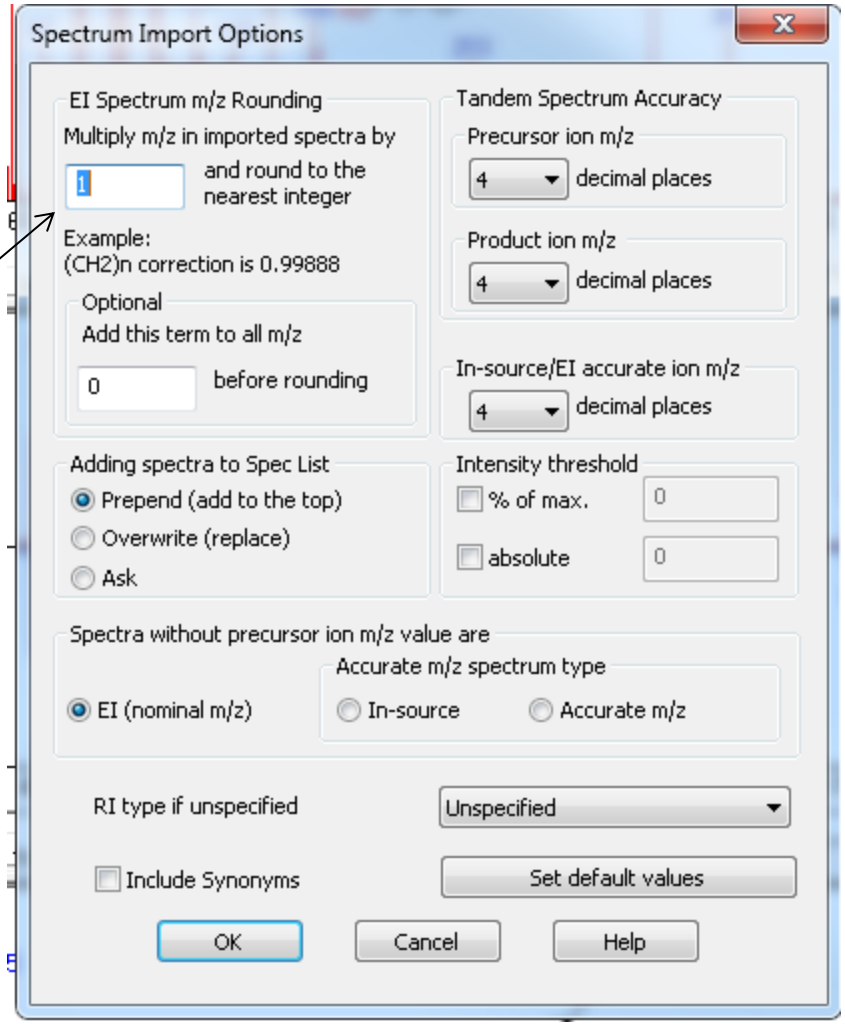
-Use ***this option*** to “Multiply *m/z* in imported spectra by..”

-More problematic for compounds MW>500, see examples ***next*** slide

-Negative mass defect a problem for multi brominated and chlorinated, species

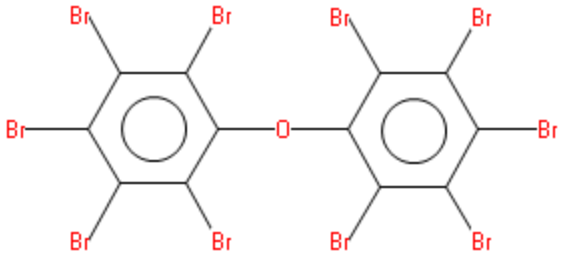
-Positive mass defect a problem for compounds with large number of hydrogens

* https://en.wikipedia.org/wiki/Mass_(mass_spectrometry)

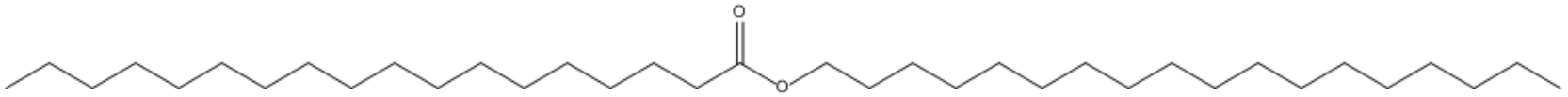


Nominal Molecular Weight Correction Examples

-NIST rounds up above 0.5, e.g. 536.5 converted to 537
-Some data systems round ***before*** exporting to NIST!



-Nominal MW needs to be ***increased*** to 950 when importing for proper searching!
-Library "thinks" bromine has MW of 79, mass spec "knows" MW is 78.9!
-Correct by multiplying *all m/z* values by 1.00087
-Problem ***only*** with higher MW species



Name: Eicosanoic acid, hexadecyl ester
Formula: $C_{36}H_{72}O_2$
MW: 536 Exact Mass: 536.55323 CAS#:

-Nominal MW needs to be ***decreased*** to 536 when importing for proper searching!
-Library "thinks" hydrogen has MW of 1, mass spec "knows" MW is 1.0078!
-Correct by multiplying observed *all m/z* values by 0.99888
-Problem ***only*** with higher MW species

Retention Indices (Kovat) Determined *by Users* in AMDIS

Name: Cholesterol
Formula: $C_{27}H_{46}O$
MW: 386 Exact Mass: 386.354866 CAS#: 57-88-5 NIST#: 332884 ID#: 7754 DB: mainlib
Other DBs: Fine, TSCA, RTECS, EPA, USP, HODOC, NIH, EINECS
Compound_ID: 0
Compound_Hash:
Contributor: NIST Mass Spectrometry Data Center
Related CAS#: 218965-24-3; 262418-13-3; 378185-03-6; 676322-57-9; 793670-51-6; 80356-14-5; 80356-33-8; 849593-11-9; 856708-55-9
InChIKey: HVYWMOMLDIMFJA-DPAQBDIFSA-N Non-stereo
10 largest peaks:
43 999 | 55 886 | 57 744 | 105 686 | 386 681 |
107 661 | 95 610 | 81 582 | 91 567 | 41 559 |
Synonyms:
1.Cholest-5-en-3-ol (3β)-; 2.(-)-Cholesterol; 3.Cholest-5-en-3β-ol; 4.Cholesterin; 5.Cholesterol base H; 6.Cholesteryl alcohol; 7.Cordulan; 8.Dusoline; 9.Dusoran; 10.Dythol; 11.Hydrocerin; 12.Kathro; 13.Lanol; 14.Nimco cholesterol base H; 15.Nimco cholesterol base No. 712; 16.Provitamin D; 17.Tegolan; 18.Wool alcohols B. P.; 19.3β-Hydroxycholest-5-ene; 20.5-Cholesten-3β-ol; 21.Cholestrin; 22.Cholestrol; 23.Super hartolan; 24.5,6-Cholesten-3β-ol; 25..DELTA.5-Cholesten-3-β-ol; 26.Cholesterine; 27.Dastar; 28.Fancol CH; 29.Cholest-5-en-3-ol, (3β)- #; 30.Cholest-5-en-3beta-ol; 31.Lidinite; 32.NSC 8798;

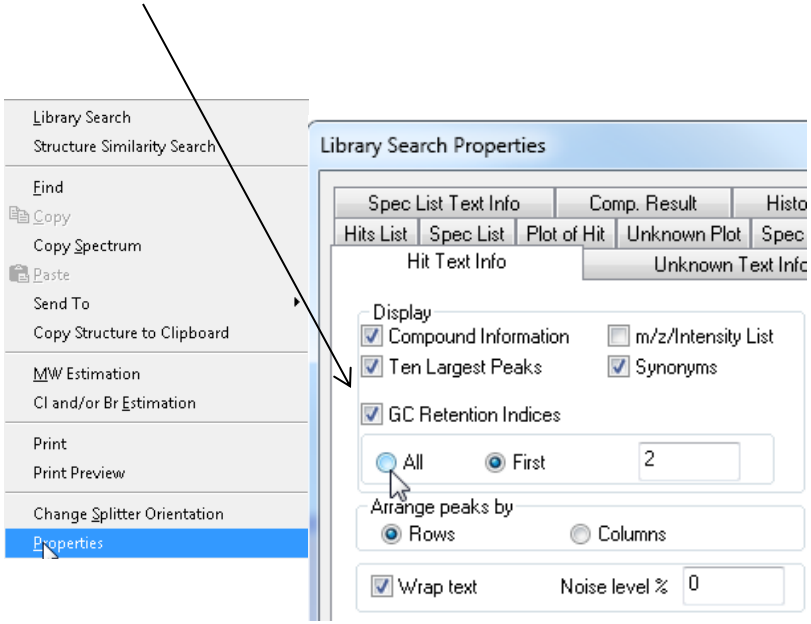
Experimental RI median±deviation [#data]
Semi-standard non-polar: 3087±12 (2)
Standard non-polar: 3052±29 (32)
Estimated non-polar retention index [n-alkane scale]:
Value: 2596 iu
Confidence interval (Low reliability): 174(50%) 752(95%) iu

Retention index
1. Value: 3098 iu
Column Type: Capillary
Column Class: Standard non-polar
Active Phase: DB-1
Column Length: 30 m
Carrier Gas: Helium
Column Diameter: 0.25 mm
Phase Thickness: 0.25 um
Data Type: Normal alkane RI
Program Type: Ramp
Start T: 50 C
End T: 250 C
Heat Rate: 10 K/min
Source: Steiger, S.; Haberer, W.; Muller, J.K., **Social environment determines degree of chemical signalling (Supplemented matherials)**, *Biol. Lett.*, *1*(6), **2011**, 822-824.

2. Value: 3098 iu
Column Type: Capillary
Column Class: Semi-standard non-polar
Data Type: Normal alkane RI
Program Type: Ramp
Source: Steiger, S.; Peschke, K.; Francke, W.; Muller, J.K., **The smell of parents: breeding status influences cuticular hydrocarbon pattern in the burying beetle Nicrophorus vespilloides**, *Proc. Roy. Soc. B*, *274*, **2007**, 2211-2220.

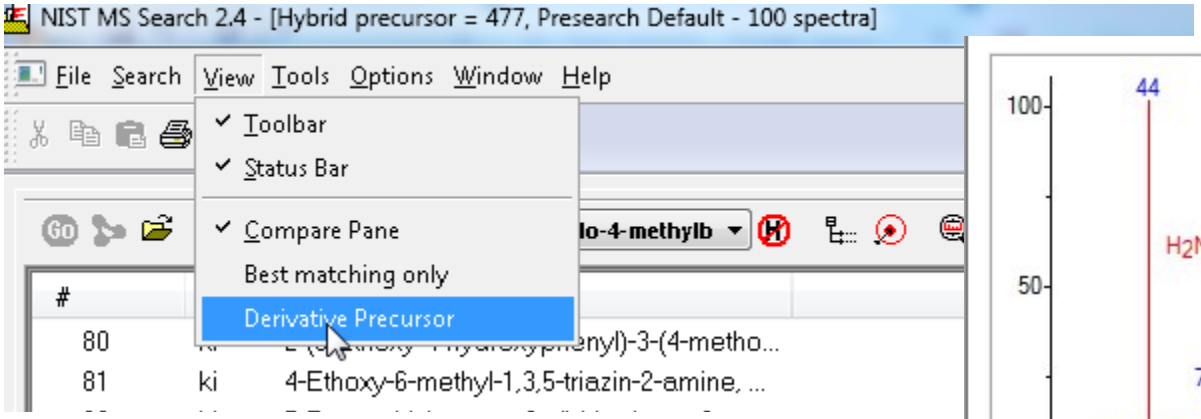
3. Value: 3075 iu
Column Type: Capillary

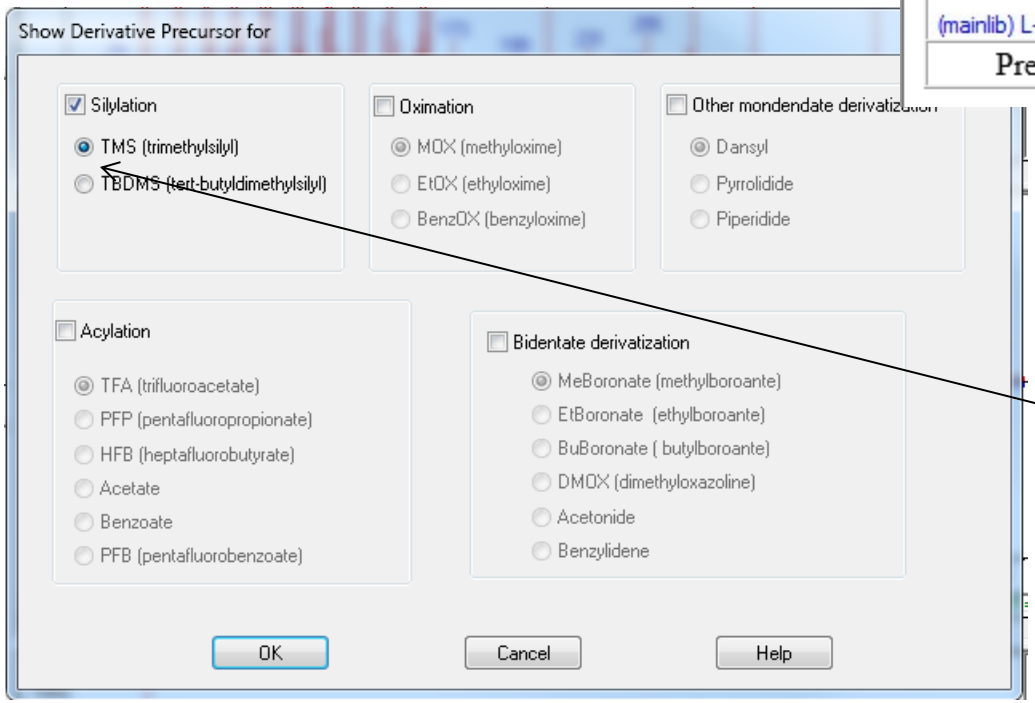
-NIST libraries have Retention/Kovat (RI) indices *
-Converts retention times into ***system-independent*** constants using a hydrocarbon calibration mixture
-RI's ***determined in*** NIST ***AMDIS*** software[3,4,12]
-Limit search, see Library Search Options/RI(GC) tab
-Additional orthogonal information for characterizing compounds
-MS Search results list methods and conditions for determination
-Standard display is top two to avoid "slowing" the display of search results
-Can expand to see **All** for a library entry, display First 0, 1…, or uncheck box to see none

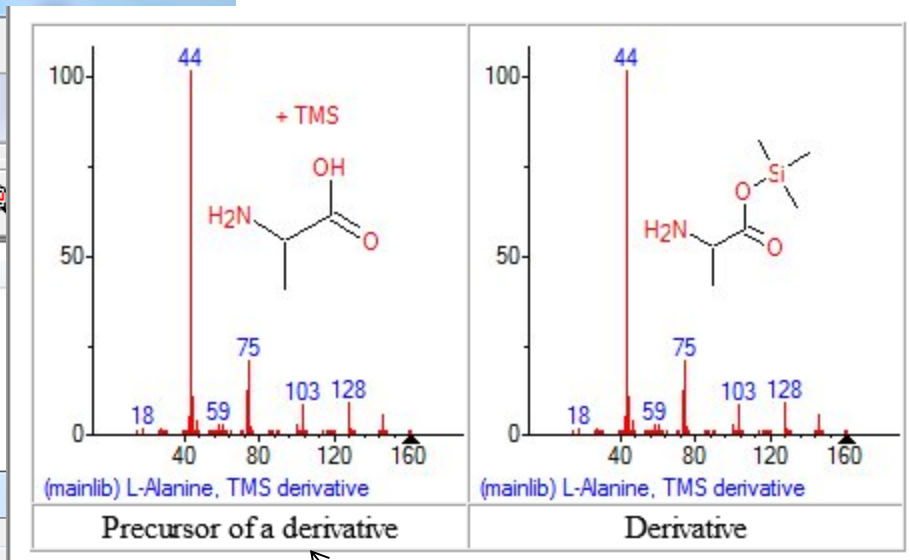


* https://en.wikipedia.org/wiki/Kovats_retention_index

Different Derivative Displays (*Affects Only Search Results Displays*)





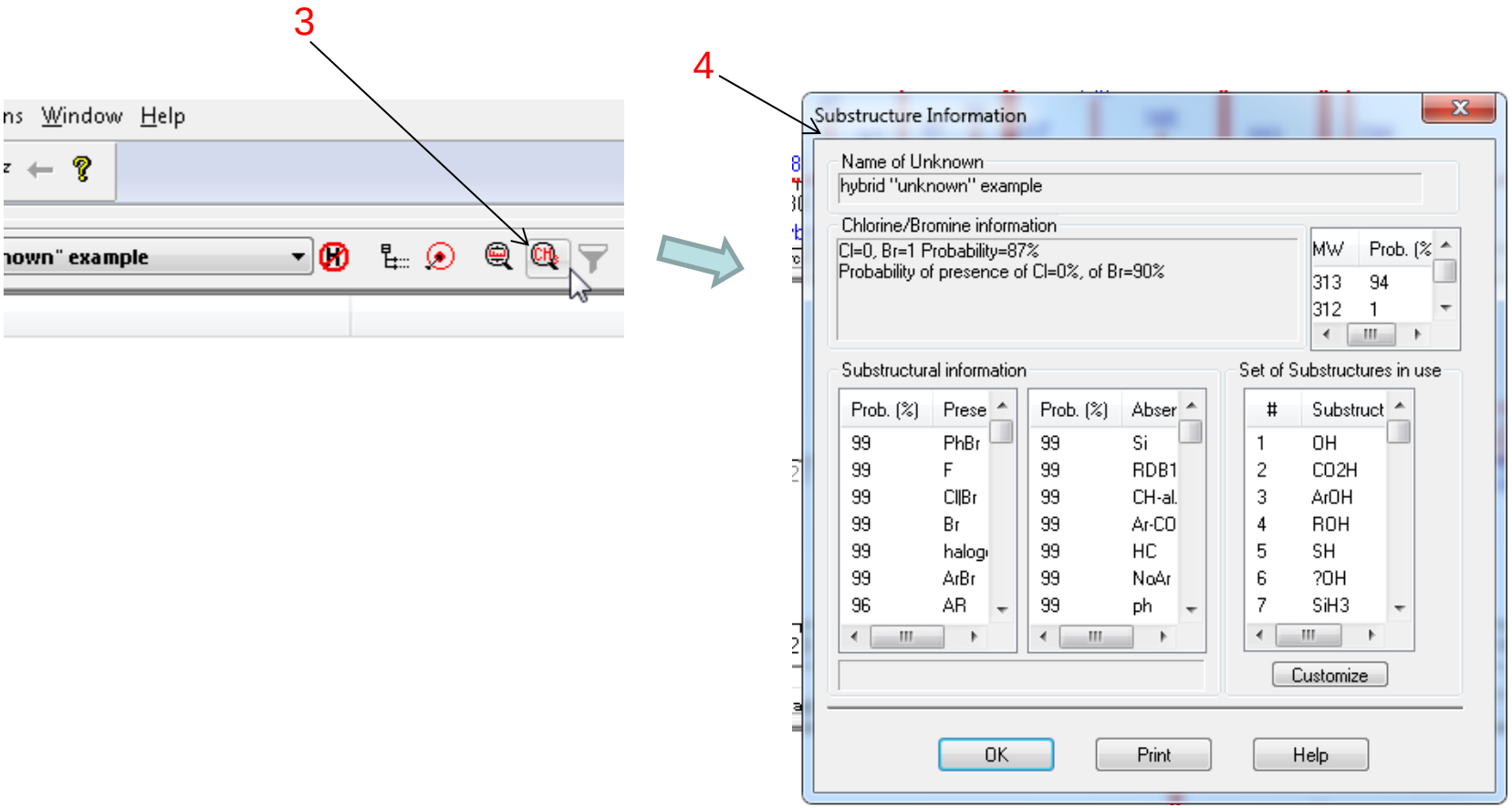


Original molecule or Derivative
Shown in Search Results

Limitation: *Can Only select one radio button under a class*

*Using NIST Software to Obtain **Proposed** Nominal MW and Substructural Information*

1. ***First*** Search spectrum in “Simple” Identity mode
2. Software uses components in hit list to determine probable nominal MW and substructural information
3. ***LMB*** on “View substructure Information”
4. ***Scroll*** through various windows in “Substructure Information” window



Webinar References (*Internet Links*)

1. *James Little Mass Spectral Resource Website*
2. *NIST Search Software Detailed Manual*
3. *AMDIS Program for Data Processing Detailed Manual*
4. *Basic Instructions for Using AMDIS with NIST Search*
5. *Nightly Automatic Update of Users' Libraries*
6. *Using NIST Search from Instrument Manufacturers' Software*
7. *Chemical Ionization for MW Determination*
8. *Trimethylsilyl Derivatives for GC-MS*
9. *Methyl Ester Derivatives for GC-MS*
10. *SciFinder/ChemSpider and Accurate Mass LC-MS Data for Unknown ID's*
11. *Surfactant Identification*
12. *QuickGuide.rtf Supplied with AMDIS Software Installation for Retention Indices*
13. *New Developments in the Modeling of Ion Fragmentation by MS Interpreter Software*
14. *Enhancements to NIST MS Interpreter for Modeling High Mass Accuracy Tandem Mass Spectra*
15. *An Automated Method for Verifying Structure-Spectral Consistency Based on Ion Thermochemistry*
16. *Combining Fragment-Ion and Neutral-Loss Matching during Mass Spectral Library Searching: A New General Purpose Algorithm Applicable to Illicit Drug Identification*
17. *The Hybrid Search: A Mass Spectral Library Search Method for Discovery of Modifications in Proteomics*
18. *Hybrid Search: A Method for Identifying Metabolites Absent from Tandem Mass Spectrometry Libraries*
19. *Structure Annotation of All Mass Spectra in Untargeted Metabolomics*
20. *Most Current Handouts for Webinar Series, Parts I-V*
21. *Lipid Matrix Ionization Effects in LC-MS*
22. Mass Spectral Similarity Mapping in Hybrid Searches Applied to Fentanyl Analogs
23. User Manual for Wiley Libraries

Acknowledgements

Wiley Webinar Production:

- Ryan McNaughton
- Kristen Makoski

Technical Advice:

- David Sparkman
- Stephen Stein