LCMS Unknown Identifications Using MSMS Libraries

Part II: NIST Search Software and Libraries

Updated 12/27/20

James Little
tvasailor@gmail.com
https://littlemsandsailing.wordpress.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,[3] Accurate Mass, Derivatization,[4,5] MS library management, SciFinder[7], Chemspider[7], Surfactant ID,[8] NMR, GC-IR, organic synthesis, matrix ionization effects,[6] 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

Table of Contents

LCMS Unknown Identifications Using MSMS Libraries

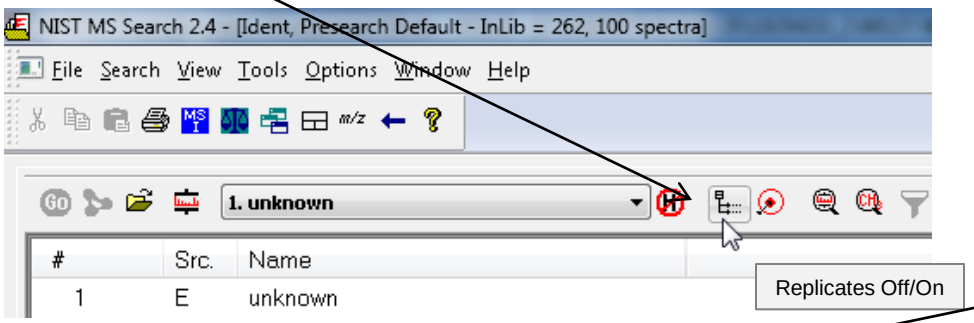
Part II: NIST MSMS Search Software and Libraries

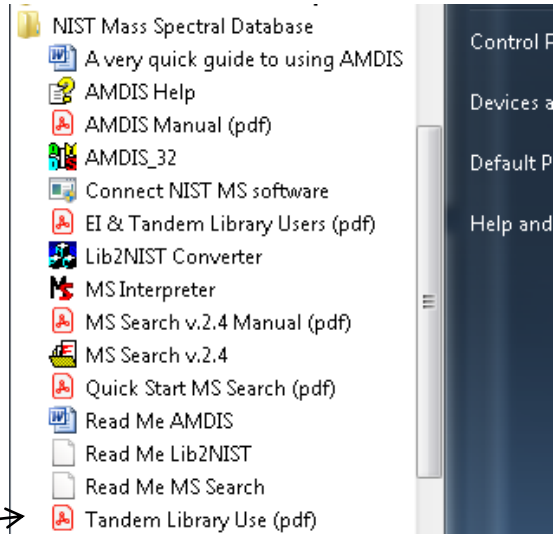
LCMS Unknown Identifications Using MSMS Libraries

Part I: Overview of Software and User Customized Configurations
Part II: NIST MSMS Search Software and Libraries
Part III: More Detailed Discussion of MSMS Hybrid Search
Part IV: Importing MSMS Spectra
Part V: NIST Structure Searches
Part VI: MS Interpreter Correlation of Substructure to MSMS Ions
Part VII: Using and Creating Other MSMS Libraries
Part VIII: Identification of Unknowns with "Spectraless" Libraries

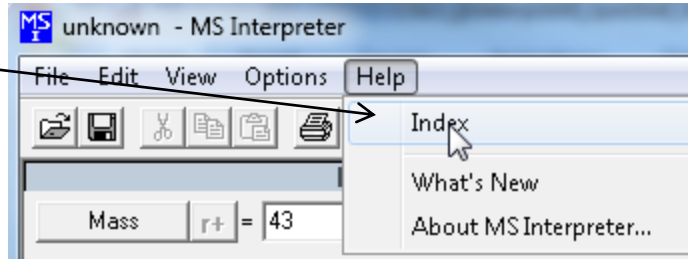
Help Files for NIST Search

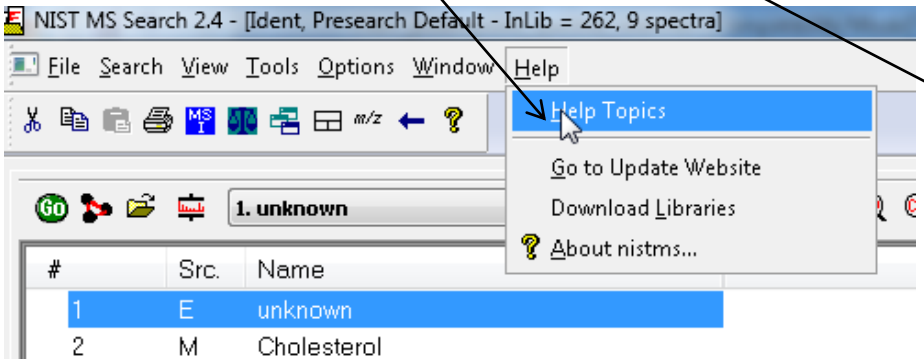
-"Hover" over Program Icon with mouse and function description displayed

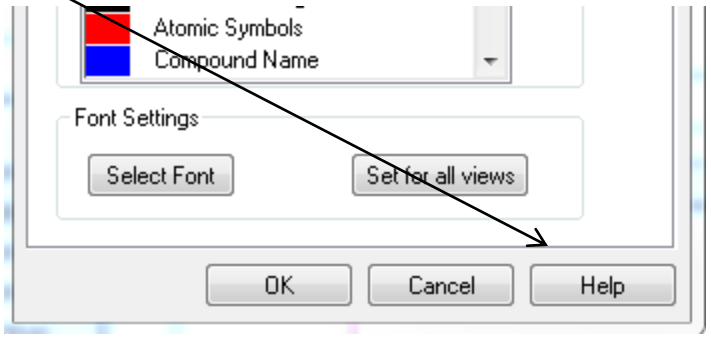




-***Detailed*** documentation for NIST Search[8]
-MS Interpreter included in NIST manual[8]
-Windows Program Group
-NIST Tandem Search ***Quick Start Guide***[9]
-"In program" assistance for both programs
-***Help button*** at corner of all windows



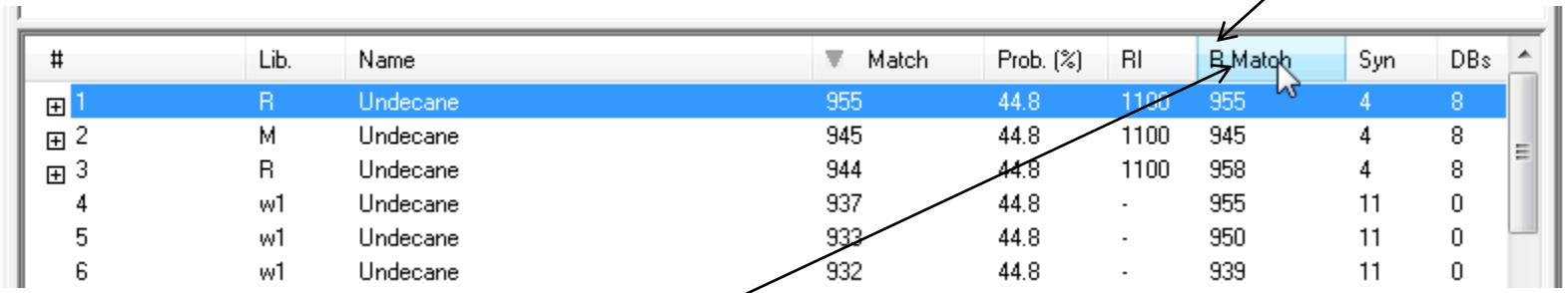




NIST Software in General is "Windows Compliant"

-***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

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

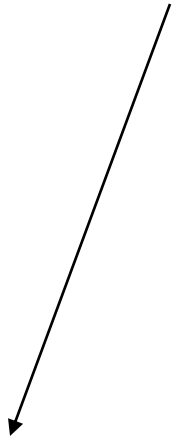
Tip 1: When reviewing search results, use up and down arrows on keyboard to quickly step through results!

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

Note: **Very Useful information in NIST Tandem Quick Start Guide[28] and NIST Users' Manual[2]**

Users' guide[28] includes:

- Search settings for normal applications
- Result scoring explanations
- Filtering information
- Suggested hit list columns to display
- Introductory paragraph below, *etc.*



Information for NIST 20 Tandem Library Users

This document describes special features of the **NIST Mass Spectral Search Program v.2.4** (NISTMS.exe) of interest to the **Tandem Library** users. General software features are presented in separate documentation.

NIST Mass Spectrometry Software and MSMS (Tandem) Libraries

- Free software program for searching and processing MSMS (tandem) libraries
- Aggressive NIST in-house program for obtaining new spectra of purchased compounds[11,13]
- Supplied with purchased NIST MSMS libraries
- Developed by NIST for curating and searching MSMS libraries
- Import data from variety of instrument manufacturers
- Searches by spectrum, structure, name, CAS No., peaks, MW, MF, etc.
- MS Interpreter[10,17-19] for correlating molecular substructures to fragment ions
- Structure export and import using vendor drawing packages
- Searches other libraries including user, Wiley, MoNA, etc.

Limitations of MSMS Searches for Unknown Identification:

- Overall process of LCMS more difficult than EI GC-MS due to large number of experimental variables of former method

- MSMS spectra are *much less* reproducible than EI spectra

- MSMS spectra *very* instrument dependent

- NIST MSMS search software[24,25] ***not as straightforward*** as EI

- Recording *many different* spectra for each species at a variety of collision energies *greatly* improves the usefulness of tandem searches

- High resolution LC-MS-MS with isotopic fidelity, mass accuracy, and accessibility very beneficial

Three Basic Types of NIST MSMS (Tandem) Searches

1) "EI Simple Similarity," no weighting of m/z intensities, ***direct peak matching approach***
   Use for finding similar compounds;

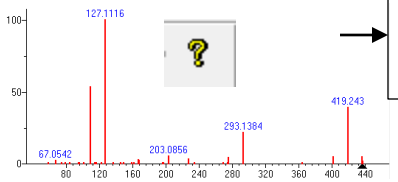
   however,[28] can use "Identity In-source HiRes," particularly good for accurate mass data, does not exclude low-resolution spectra in results

2) "MS/MS Identity," search of fragment ions with specified precursor ion
   Use for finding exact compound or isomer

3) "MS/MS Hybrid," ***novel search method*** [20-27]
   Use for identifying unknown components

- Finds compounds that differ by a single or multiple chemical moieties
- Key parameter is "*Delta-Mass"* which is mass difference between query and library components
- Significantly *extends* the scope of a library for identifications
- Combines "*direct peak matching*" score with the logical equivalent of "*neutral loss matching"* score to yield a "hybrid" score
- Precursor ion *must* be specified
- More discussion included in ***Part III*** of series

Overall Description for Searching Libraries and Displaying Results



MSMS tolerances

Presearch
(Major peaks in spectrum)

MW Constraint Selected or Presearch Off

Final Search of Candidates
(Every peak in spectrum)

Sort Results

Other Constraints

Search Results

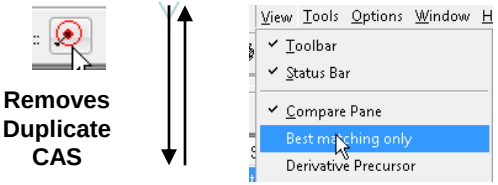
Spectrum Search Types-*3 Basic MSMS to Chose: Similarity/EI Simple, Identity/MSMS, Similarity/MSMS Hybrid*

MW-skips presearch and limits to MW (molecular weight) specified in “Identity/EI Simple”

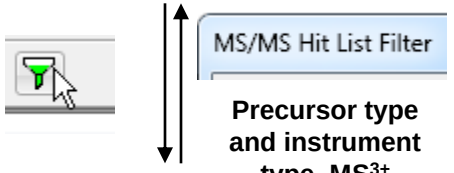
Presearch Default*-best setting for optimum results*

Presearch off-skips presearch, >1 minute per search

Constraints-results filtered ***after*** Final Search, ***not reversible***



Search Results



Search Results

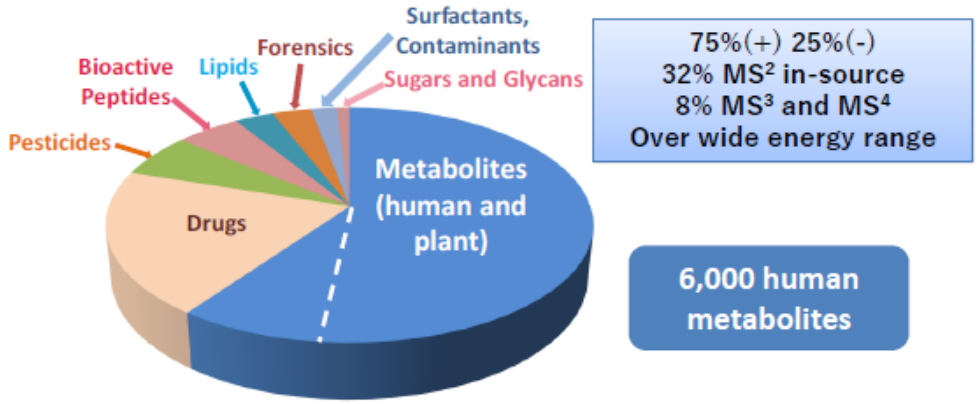
NIST MSMS (Tandem) Libraries 2020 Release[9-13]

hr_msms_nist: 1,026,712 high resolution accurate mass of small molecules
lr_msms_nist: 215,649 low resolution spectra of small molecules
apci_msms_nist: *3,555* high resolution APCI spectra of extractables and leachables

31K Compounds, 2X More than 2017
186K Precursor Ions - 1.3M Spectra

Fragmentation Methods	Precursor Ion Types
27,840 HRAM (High Res Accurate Mass) Compounds	26,575 Protonated
29,890 QTOF, HCD, IT-HRAM, QqQ Compounds	12,589 Deprotonated
29,444 Ion Trap Compounds (Low Res., up to MS^4)	10,032 Water/Ammonia Loss
246 APCI HRAM 'Extractables and Leachables'	24,167 Other In-Source Generated

Types of Compounds and Spectra



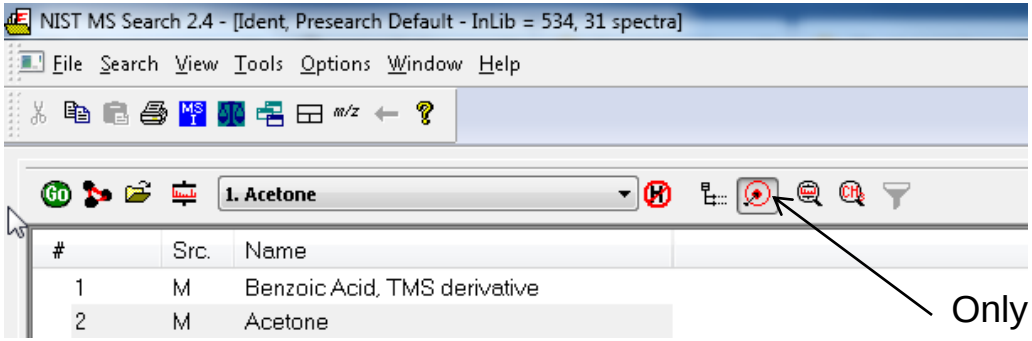
75%(+) 25%(-)
32% MS^2 in-source
8% MS^3 and MS^4
Over wide energy range

6,000 human metabolites

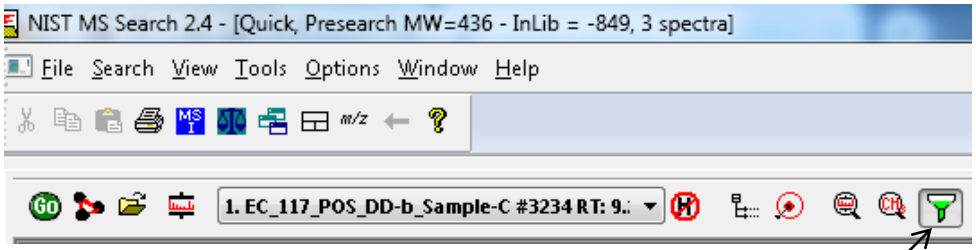
NOTE: Above numbers do not include **biopep_msms_nist**[14-16] *specialty library* of >4.3 million spectra of 1.26 millions species which utilizes special search program

Minimizing Number of Results *After* Search Performed

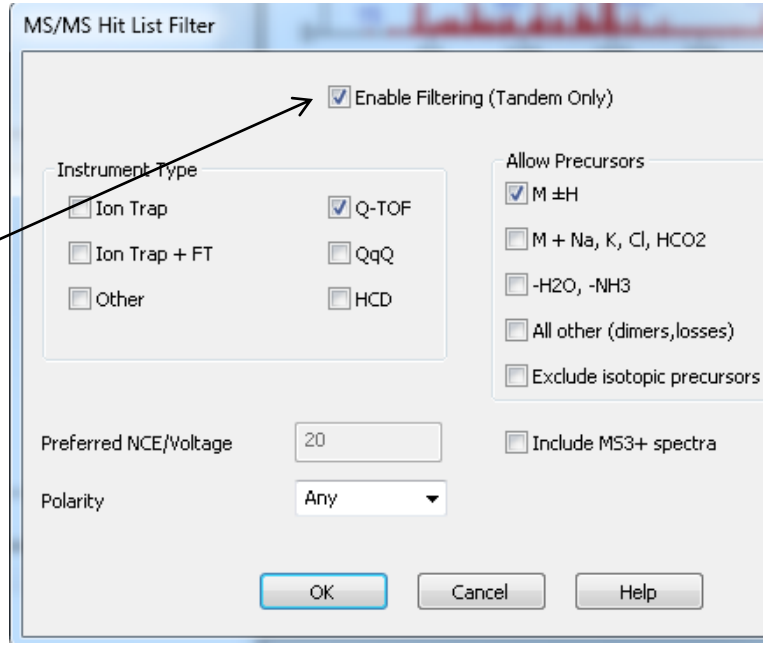
Process is Reversible without Repeating Search



Only shows best hit with same CAS number

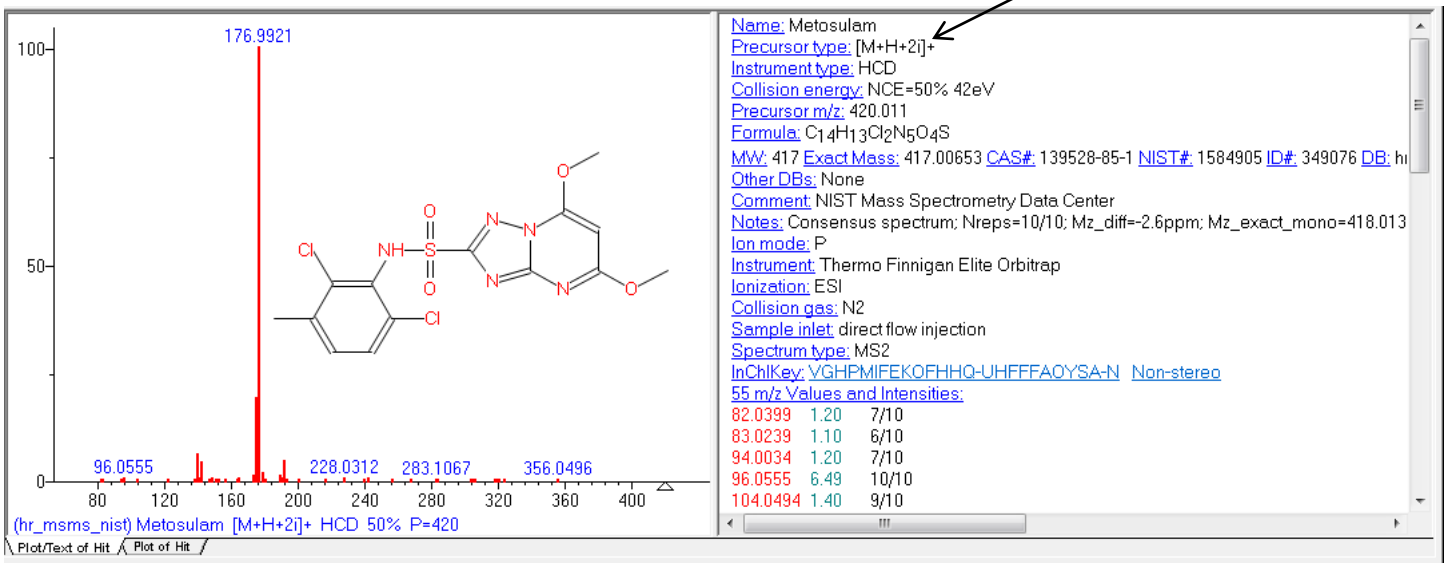


Filters types of spectra using "MS/MS Hit List Filter" option

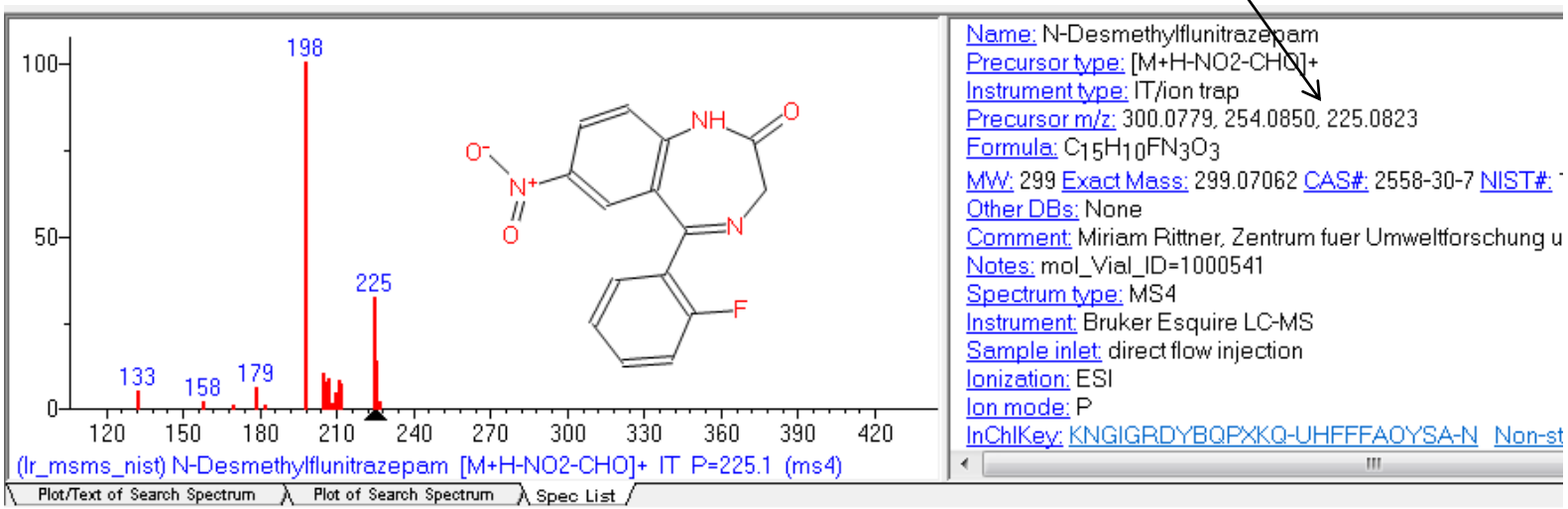


For Example: Two Types of Ions Filtered with MS/MS Hit List Filter Option

MS/MS of M+2 for ^{37}Cl isotope, select “Exclude Isotopic Precursors”



Multistage “MS4” Spectrum below, select “Include MS3+ Spectra”

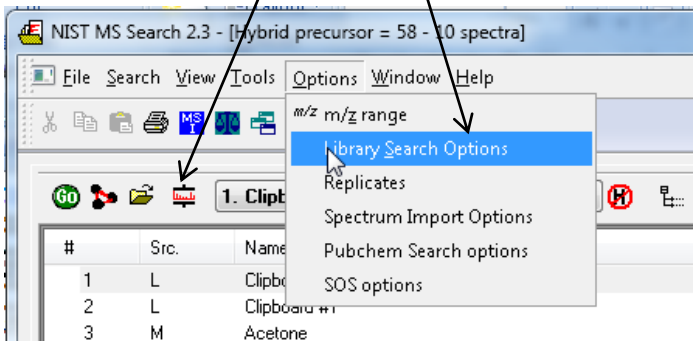


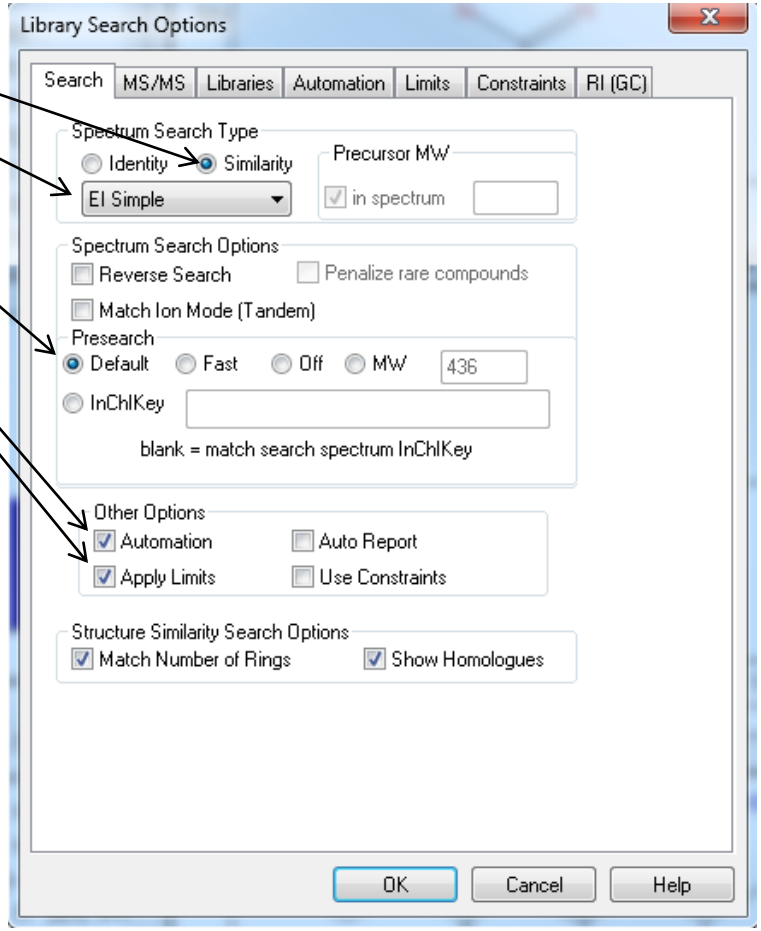
Setting Up Default Parameters for Similarity Search (Step 1)

Critical Step

-EI Simple Similarity Search, "*direct peak matching*"
-Basic search for ***similar*** spectra, no weighting of intensities
-Select the following settings

Two ways to access:



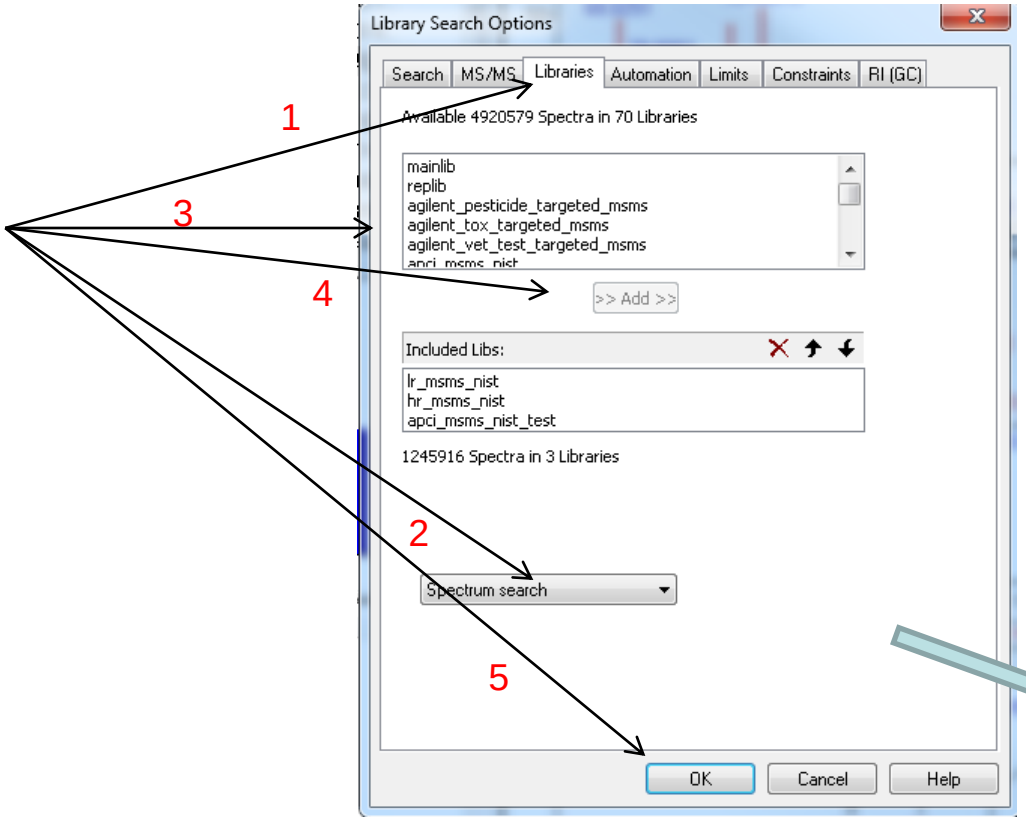


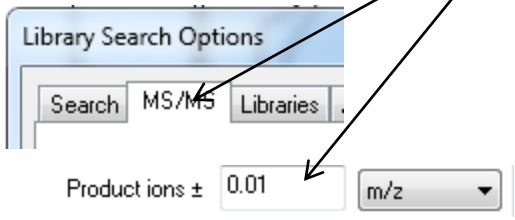
NOTE: If you want to use "Identity In-Source HiRes" instead of "Similarity EI simple*,"* ***must*** set-up values in MS/MS Tab,[28] use same value for Product ions +/- of 0.01 *m/z* found in slide 22, takes advantage of hi resolution accurate mass information

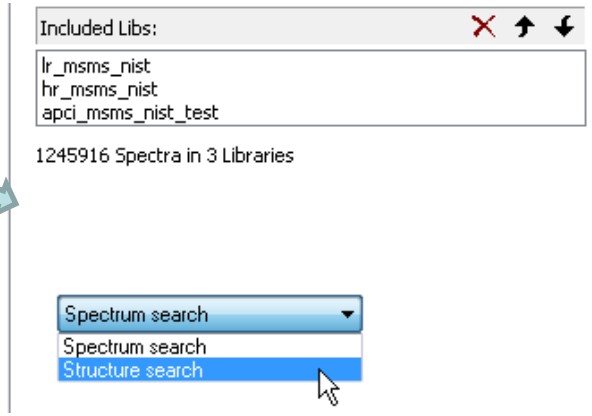
Setting Up Default Parameters for Similarity Search (Step 2)
Select the Libraries Used for Spectral Searches in Library Search Options

- ***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 in Sequential Search Function
- Add ***same*** libraries for ***structure*** search, discussed in detail in ***Part V*** of series

NOTE: If you want to use "Identity In-Source HiRes" instead of "Similarity EI simple," ***must*** set-up values in MS/MS Tab,[28] use same value for Product ions +/- of 0.01 *m/z* found in slide 22, takes advantage of hi-resolution accurate mass data

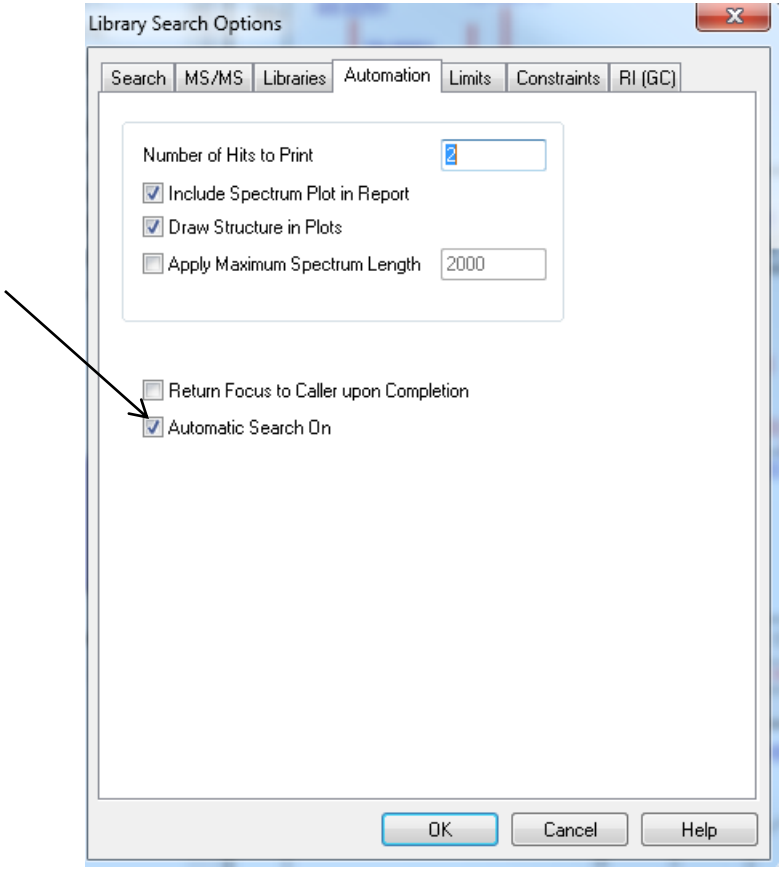


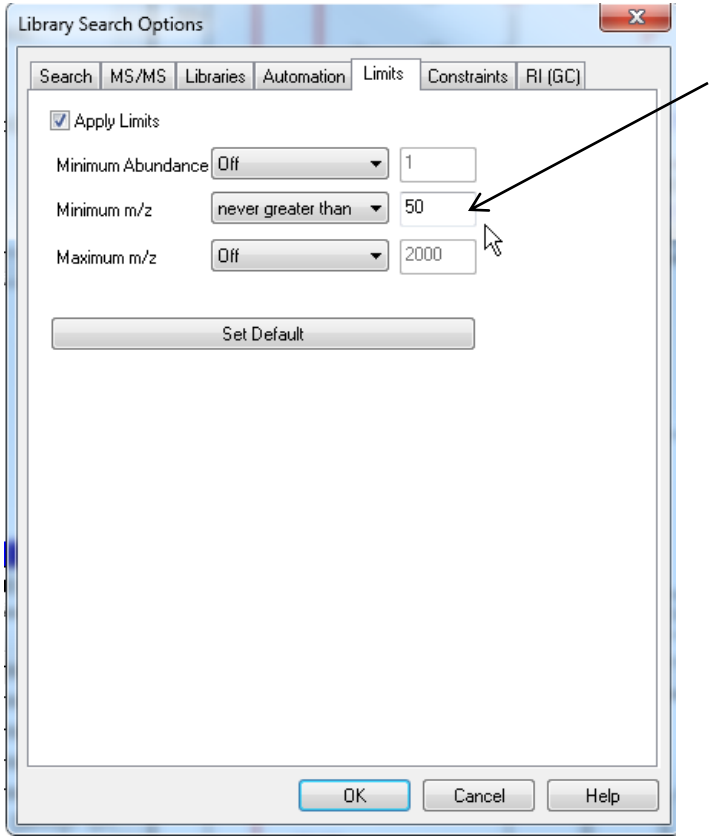




Setting Up Default Parameters for Similarity Search (Step 3)

Selecting Up Limits and Automation Settings in Library Search Options



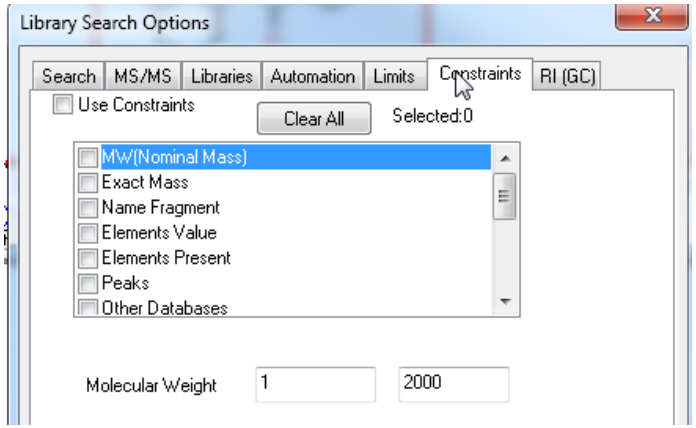


Setting Up Default Parameters for Similarity Search (Step 4)

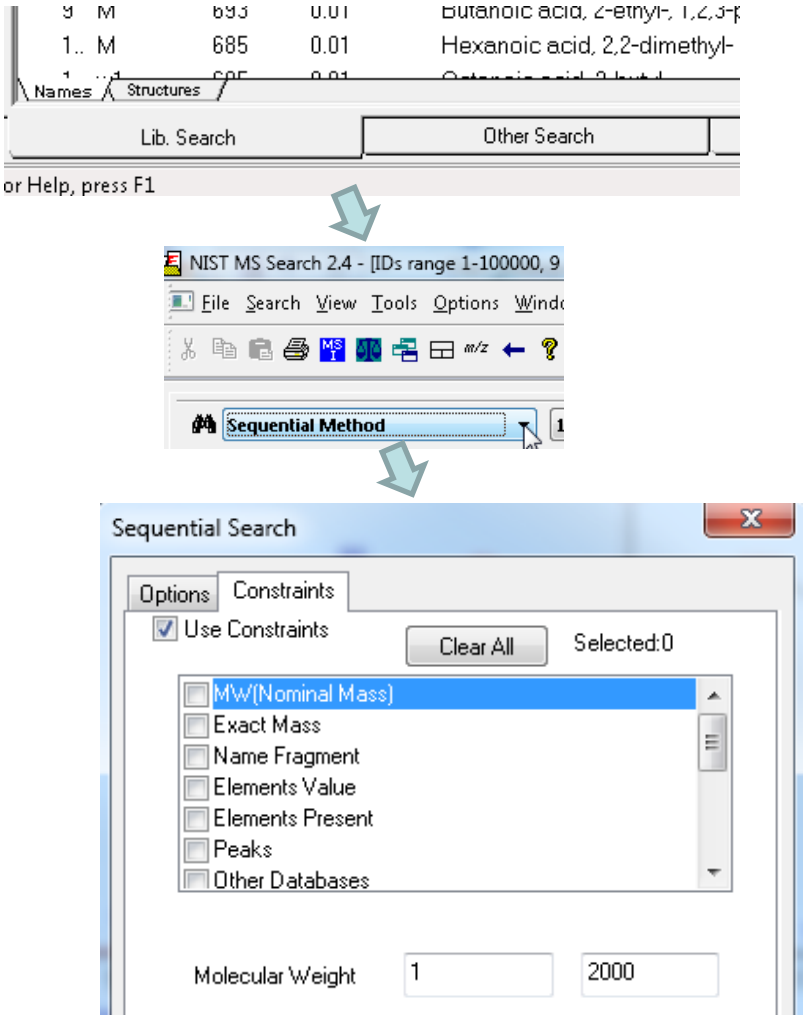
Constraints Applied to Hit List *After* the List Is Determined

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



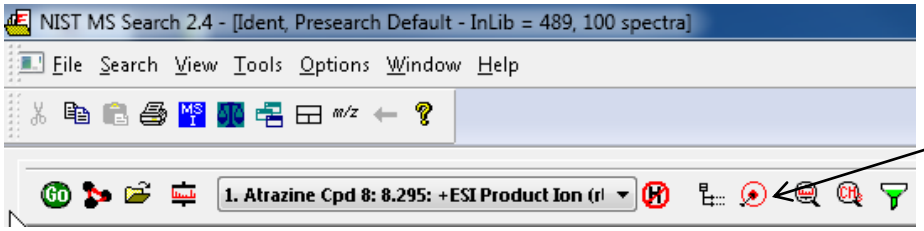
2



Setting Up Default Parameters for Similarity Search (Step 5)

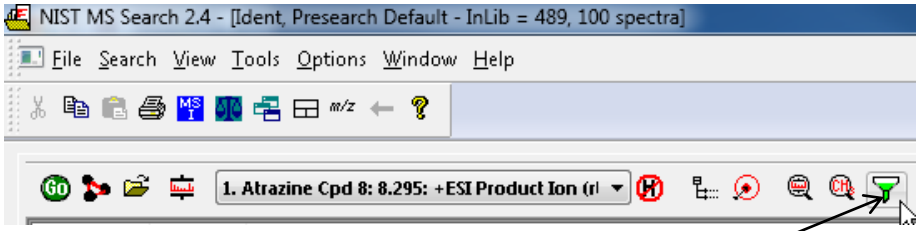
Ways of *Simplifying* Search Results Display

- These settings used to minimize spectra shown in library search results
- *Turn* them *off initially*, but use as needed to refine the results
- They can be used to limit the results *without repeating* the search
- The *1st* setting removes hits from list with same CAS number
- The *2nd* group of *filter settings* uses tags in library spectra to filter results



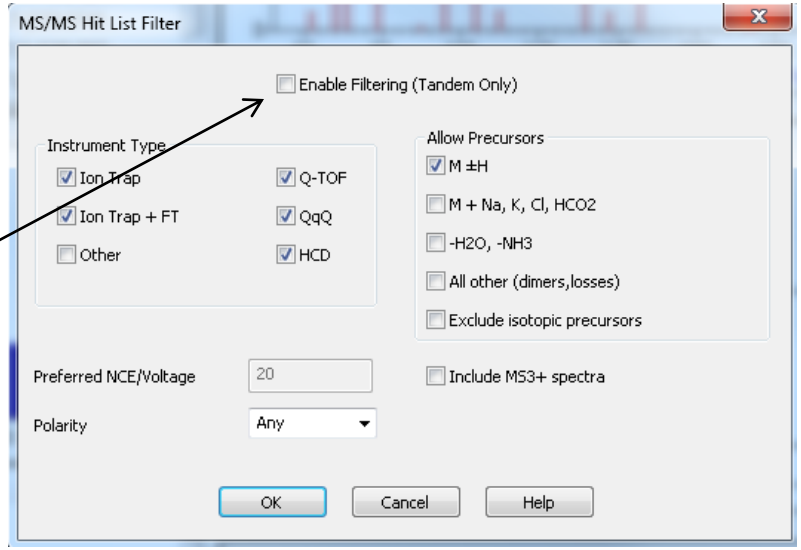
1st

-*Off* when not depressed
-*On* when depressed



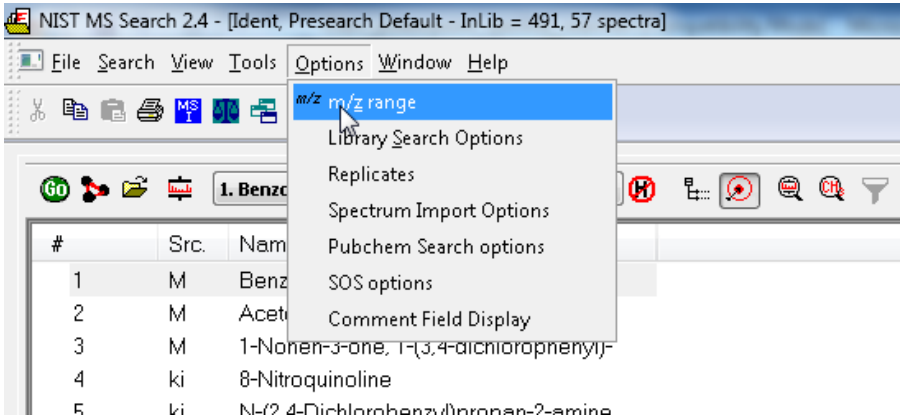
2nd

-*Access* Filter settings
-Initially *do not check* "Enable Filtering"

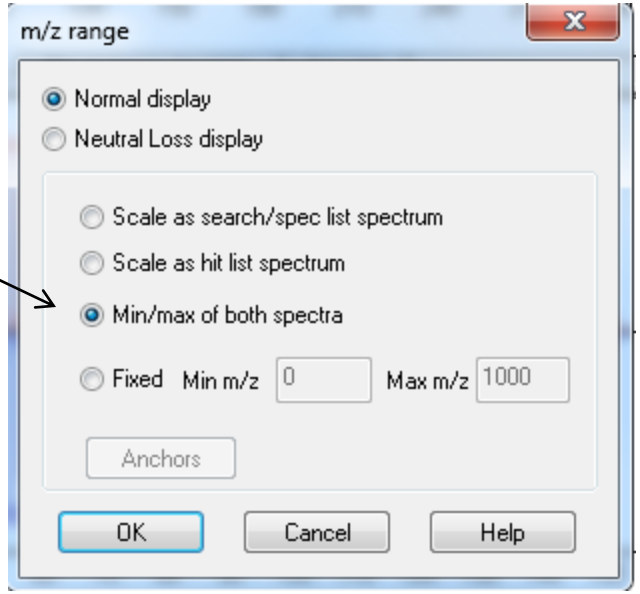


Setting Up Default Parameters for Similarity Search (Step 6)

Setting Default *m/z* Range Displayed in Spectra



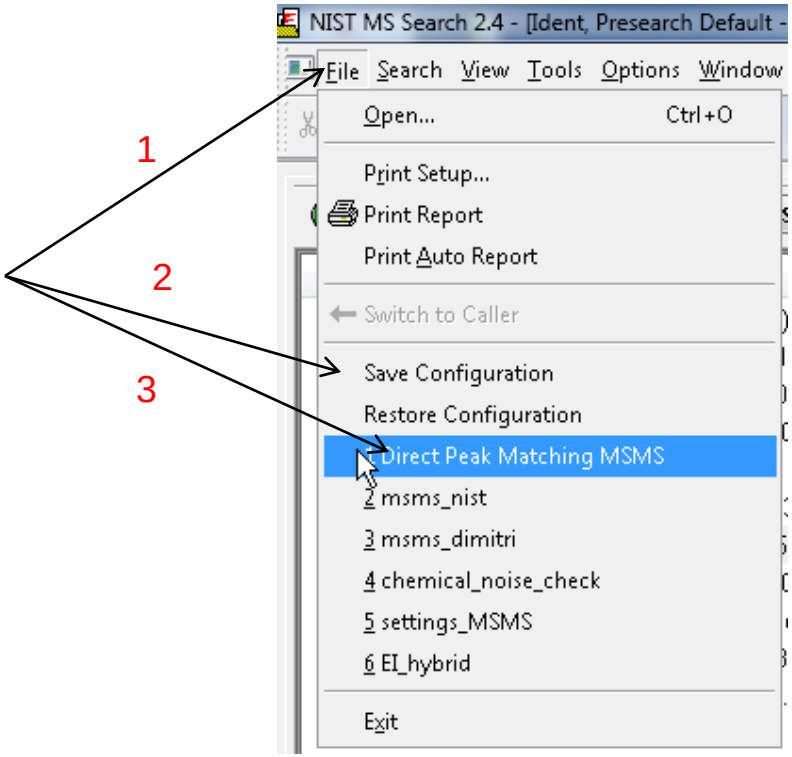
-I prefer min/max of both spectra



Setting Up Default Parameters for Similarity Search (Step 7)

Saving Setup of Standard Search Parameters

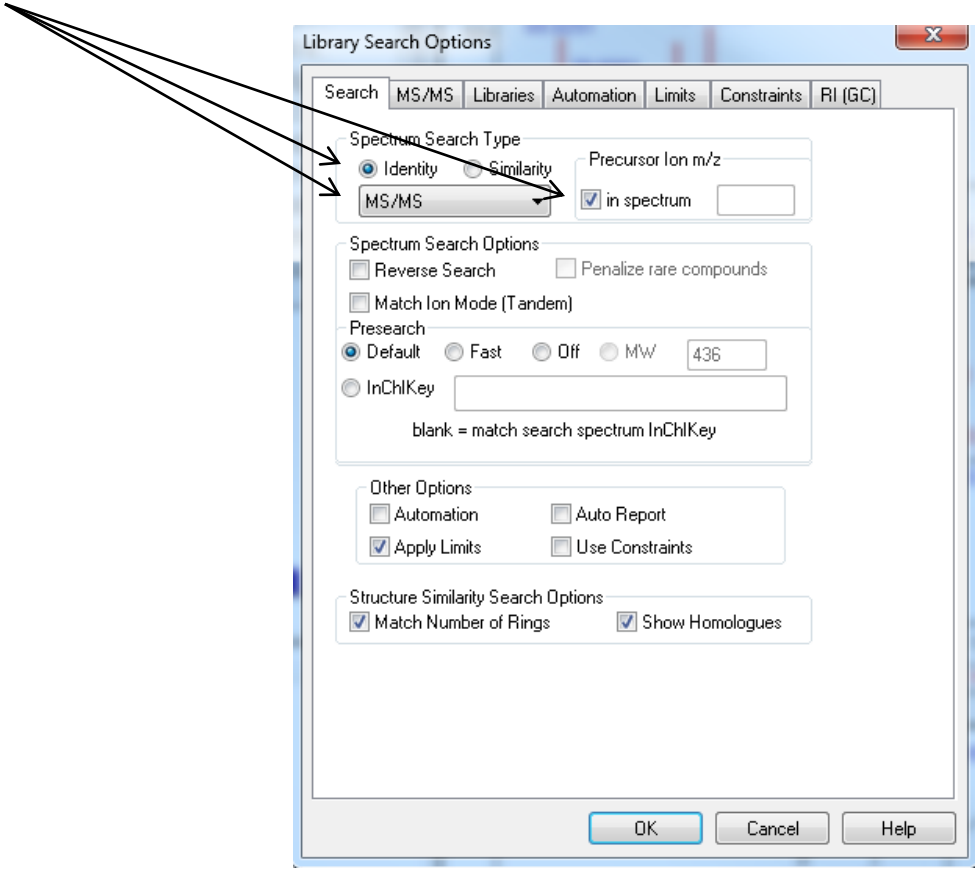
- After all the parameters are setup for search, save them for future use
- Example setup below was named ***Direct Peak Matching MSMS***
- In the future, this can be “restored” by selecting it from this menu



Setting Up Default Parameters for Identity MS/MS Search (Step 1)

Critical Step

- *Many* of settings *same* as in first type of search
- *Searches* are *restricted* by precursor ion m/z
- Check “in spectrum” if precursor field is *not listed* in spectrum header information
- User *must* then enter value in “Precursor Ion m/z” field appropriate for each spectrum searched

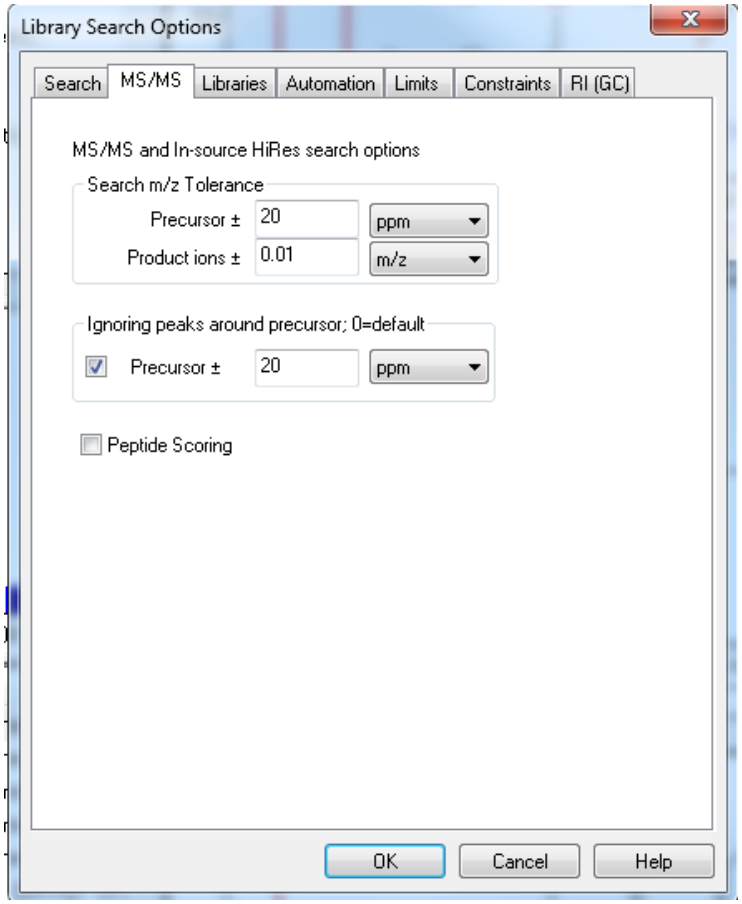


Setting Up Default Parameters for Identity MSMS Search (Step 2)

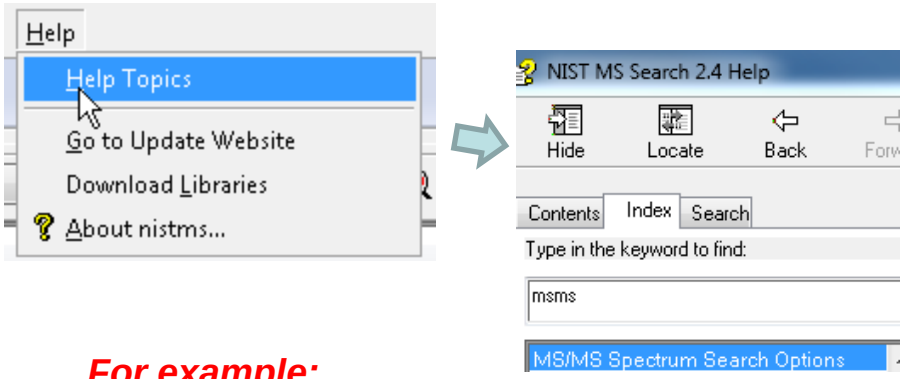
MS/MS Limits Tab

Excerpt from NIST Tandem Quick Start Guide[28]

The precursor mass tolerance should be set to reflect the accuracy of your instrument - NIST Tandem Library spectra always have the exact mass value for the precursor ion. It is generally recommended that the product-ion tolerance be set at 0.01 *m/z* units to ensure that lower mass peaks are matched. Due to the fact that spurious peaks commonly appear near the precursor ion, a setting of 20 ppm is recommended for the **Ignoring peaks around precursor** specification. **DO NOT** select **Peptide Scoring** unless using peptide libraries.



Additional information available within program "Help Topics" on MSMS parameters



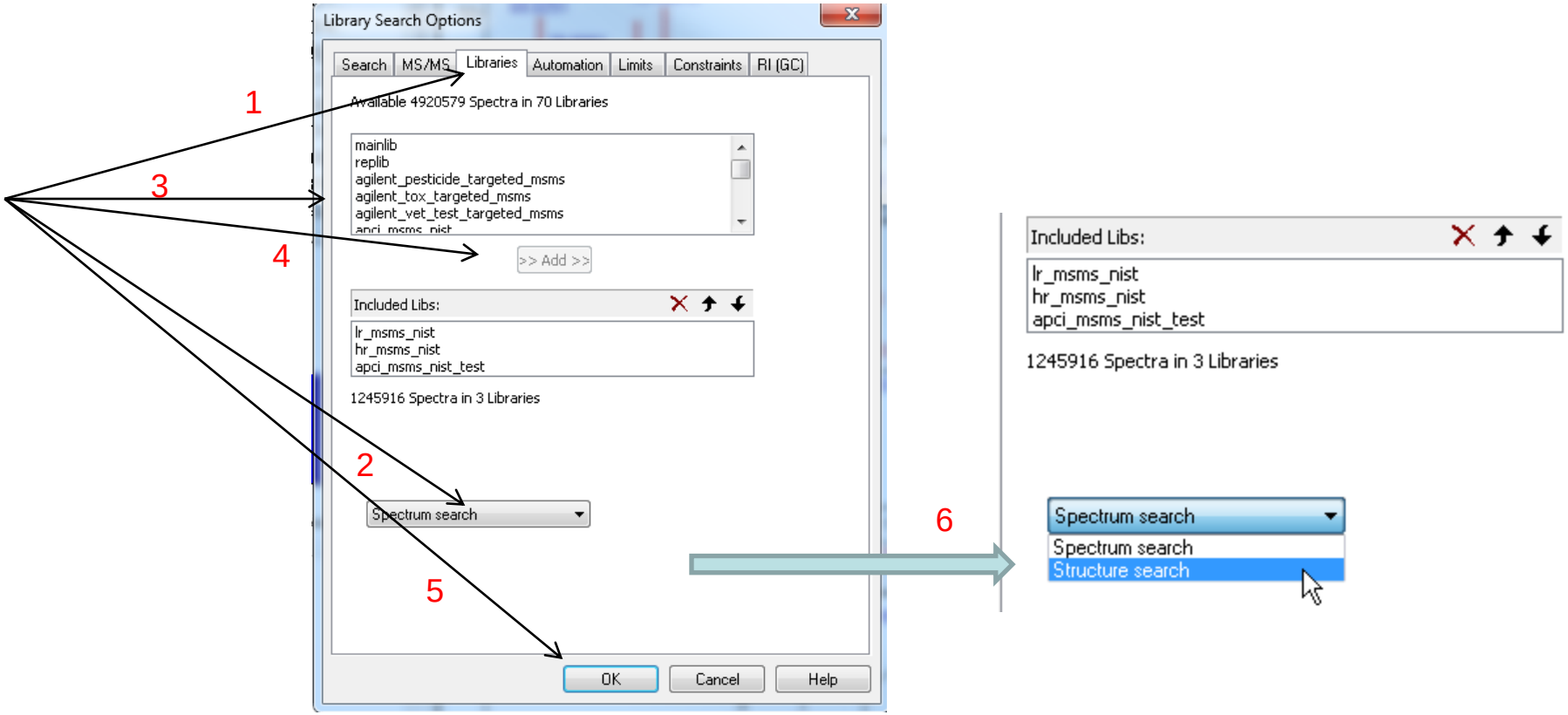
For example:

Note. For Low Resolution library (LR_*) product ions tolerance is fixed at unit mass resolution.

Setting Up Default Parameters for Identity MSMS Search (Step 3)

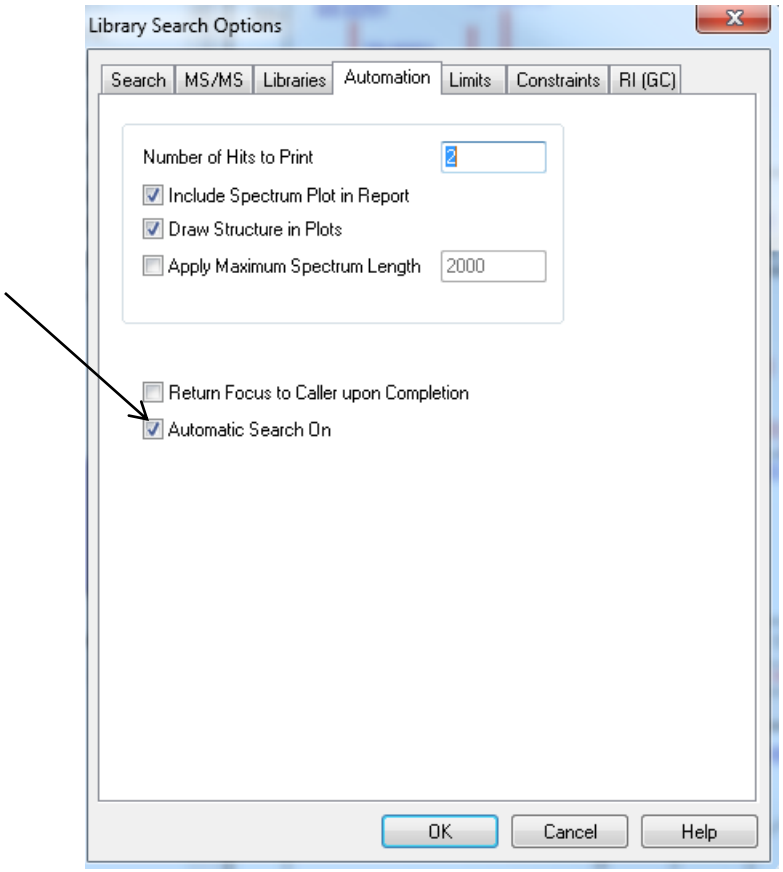
Select the Libraries Used for Spectral Searches in Library Search Options

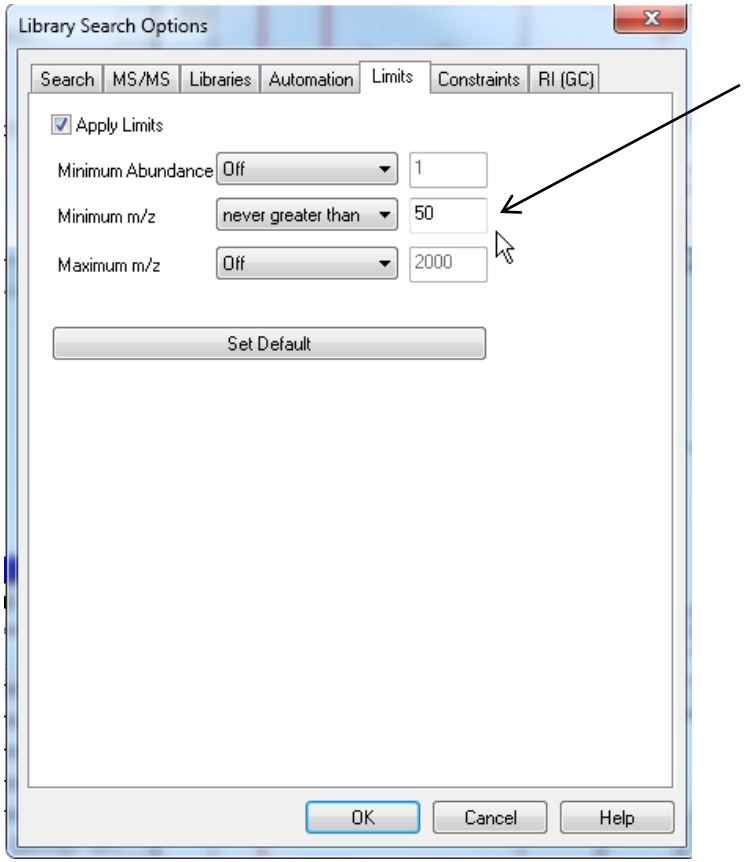
- ***LMB*** on the ***Libraries*** tab and make sure ***Spectrum Search*** is selected
- Select the group of libraries to be searched
- ***>>Add>>*** all libraries to be searched
- ***Order*** of libraries normally ***only important*** in Other Search Tab which return ***maximum no.*** of hits in Sequential Search Function
- Add ***same*** libraries for ***structure*** search, discussed in detail in ***Part V*** of series



Setting Up Default Parameters for Identity MSMS Search (Step 4)

Selecting Up Limits and Automation Settings in Library Search Options



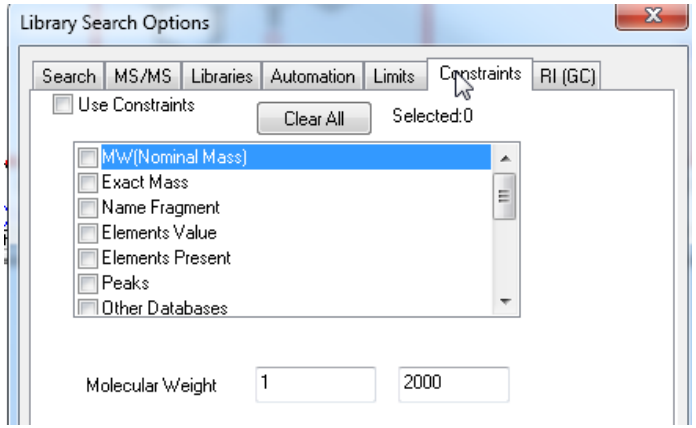


Setting Up Default Parameters for Identity MSMS Search (Step 5)

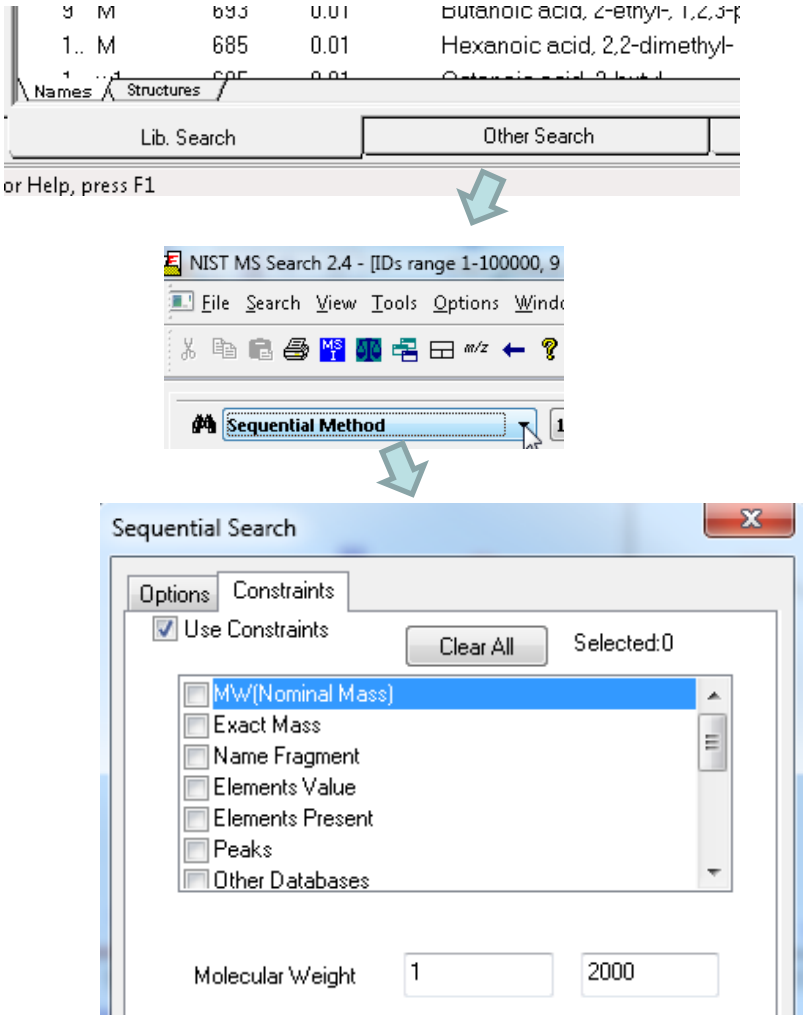
Constraints Applied to Hit List *After* the List Is Determined

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



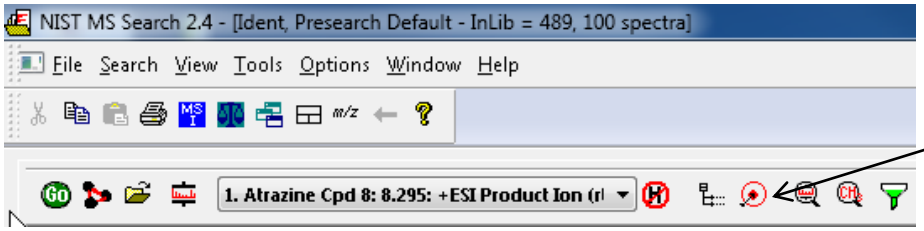
2



Setting Up Default Parameters for Identity MSMS Search (Step 6)

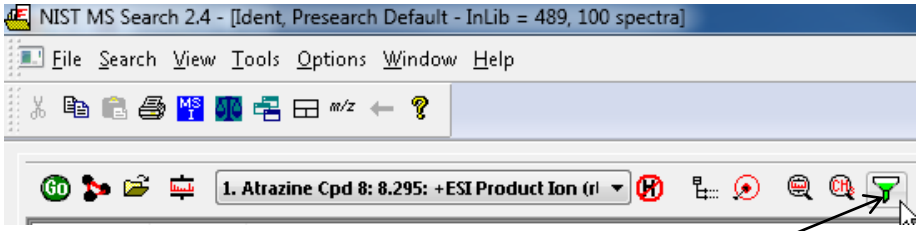
Ways of *Simplifying* Search Results Display

- These settings used to minimize spectra shown in library search results
- *Turn* them *off initially*, but use as needed to refine the results
- They can be used to limit the results *without repeating* the search
- The *1st* setting removes hits from list with same CAS number
- The *2nd* group of *filter settings* uses tags in library spectra to filter results



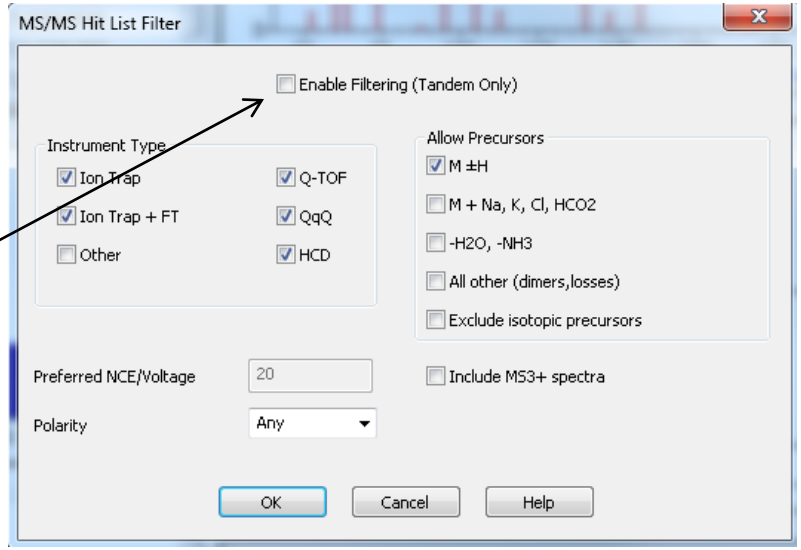
1st

-*Off* when not depressed
-*On* when depressed



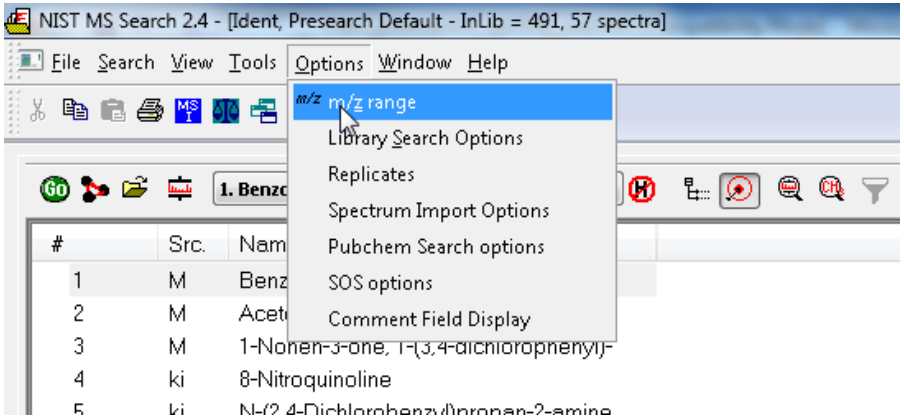
2nd

-*Access* Filter settings
-Initially *do not check* "Enable Filtering"

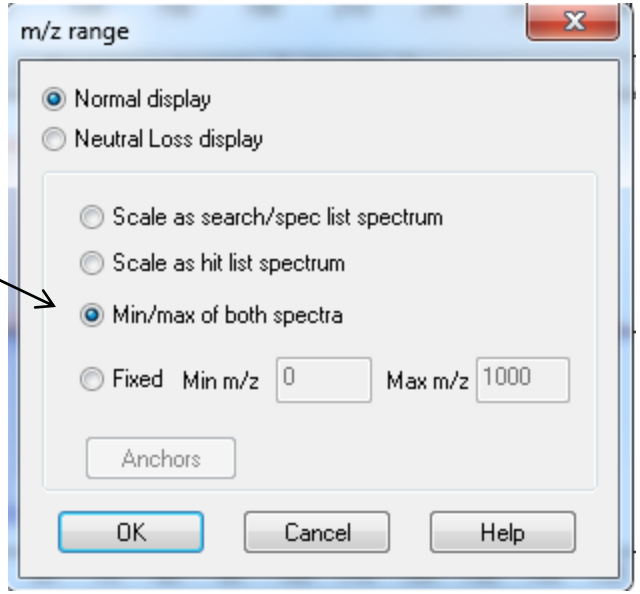


Setting Up Default Parameters for Identity MSMS Search (Step 7)

Setting Default *m/z* Range Displayed in Spectra



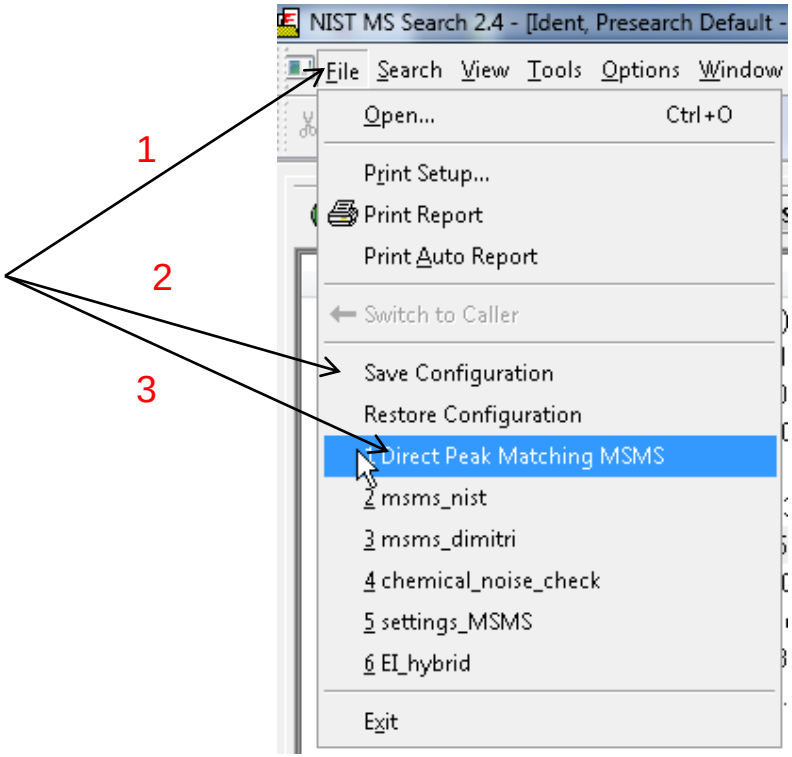
-I prefer min/max of both spectra



Setting Up Default Parameters for Identity MSMS Search (Step 8)

Saving Setup of Standard Search Parameters

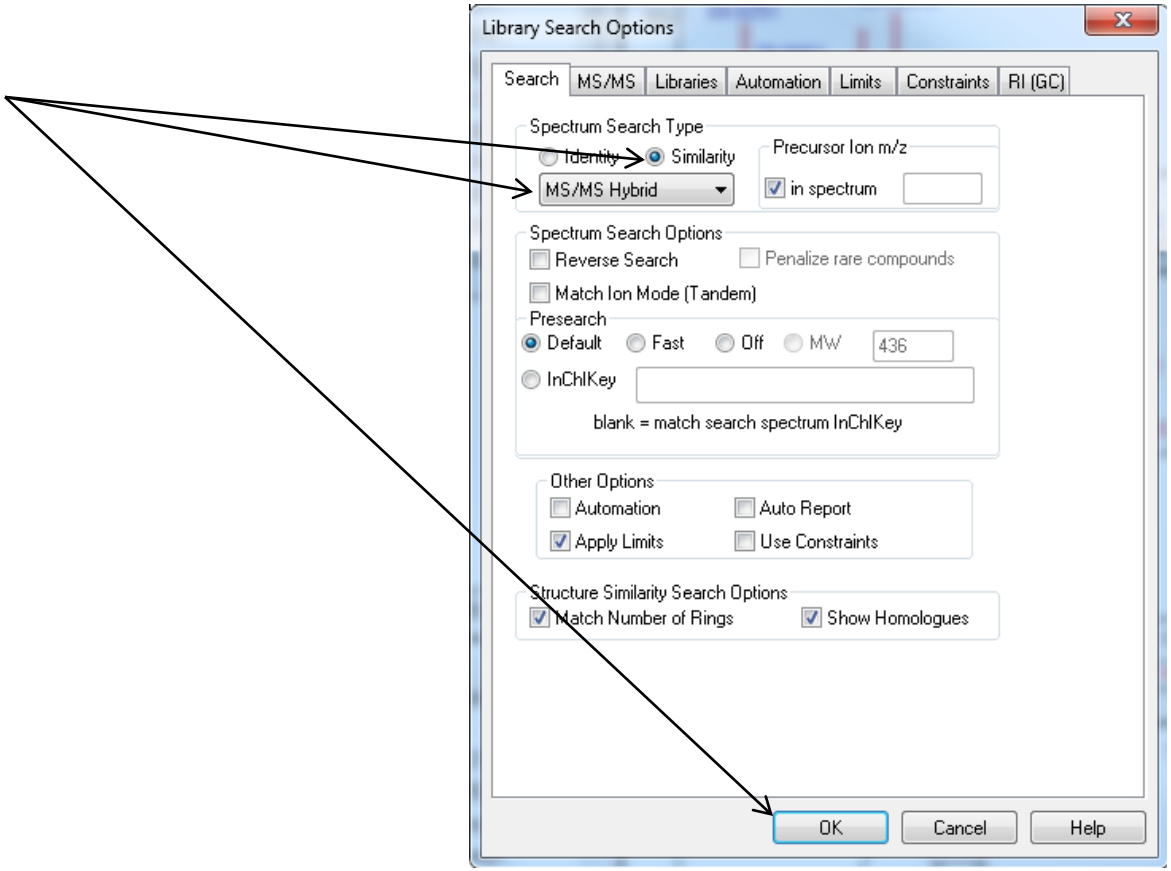
- After all the parameters are setup for search, save them for future use
- Example setup below was named ***Identity MSMS***
- In the future, this can be "restored" by selecting it from this menu



Setting Up Default Parameters for Similarity MS/MS Hybrid Search (Step 1)

Critical Step

- *Many* of settings *same* as in other types of MS/MS searches
- *Searches* are *restricted* by precursor ion m/z
- ***Un-***Check "in spectrum" if Precursor_m/z field is *not listed* in spectrum header information of unknown
- ***Some m***anufacturers import Precursor_m/z with their spectrum, ***others*** do ***not***
- If not, user *must* then enter value in "Precursor Ion m/z" field

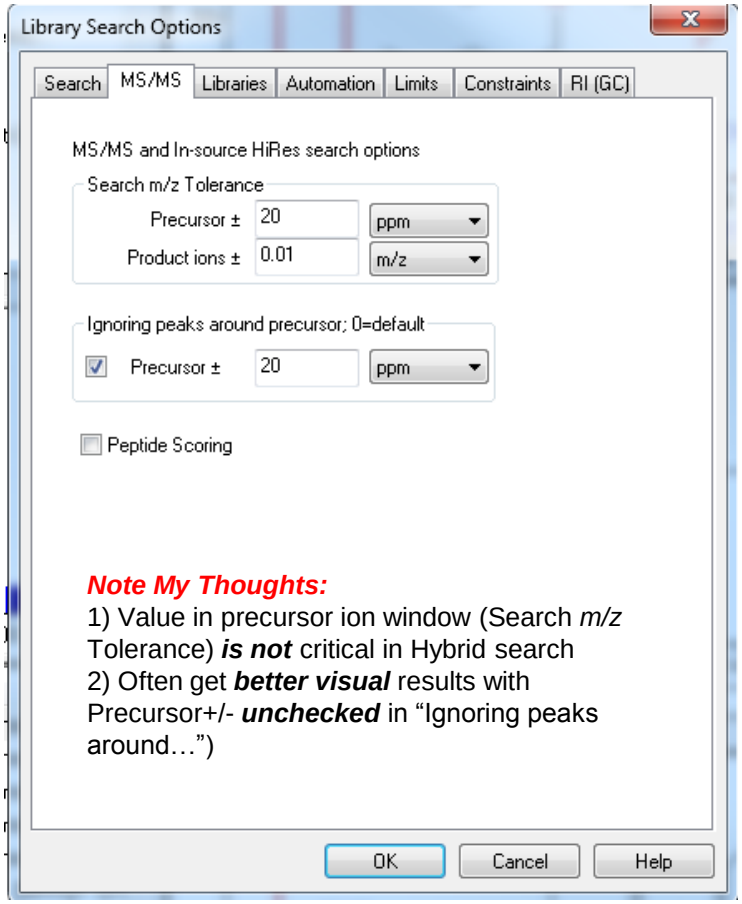


Setting Up Default Parameters Similarity MS/MS Hybrid Search (Step 2)

MS/MS Limits Tab

Excerpt from NIST Tandem Quick Start Guide[28]

The precursor mass tolerance should be set to reflect the accuracy of your instrument - NIST Tandem Library spectra always have the exact mass value for the precursor ion. It is generally recommended that the product-ion tolerance be set at 0.01 *m/z* units to ensure that lower mass peaks are matched. Due to the fact that spurious peaks commonly appear near the precursor ion, a setting of 20 ppm is recommended for the **Ignoring peaks around precursor** specification. **DO NOT** select **Peptide Scoring** unless using peptide libraries.

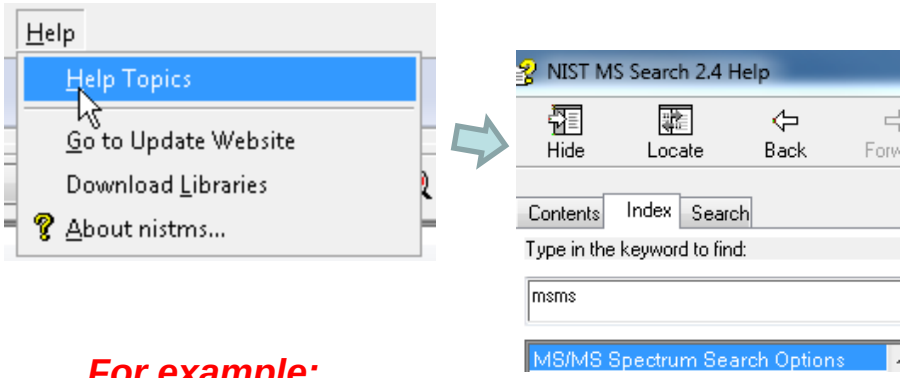


Note My Thoughts:

1) Value in precursor ion window (Search *m/z* Tolerance) ***is not*** critical in Hybrid search

2) Often get ***better visual*** results with Precursor+/- ***unchecked*** in "Ignoring peaks around…")

Additional information available within program "Help Topics" on MSMS parameters



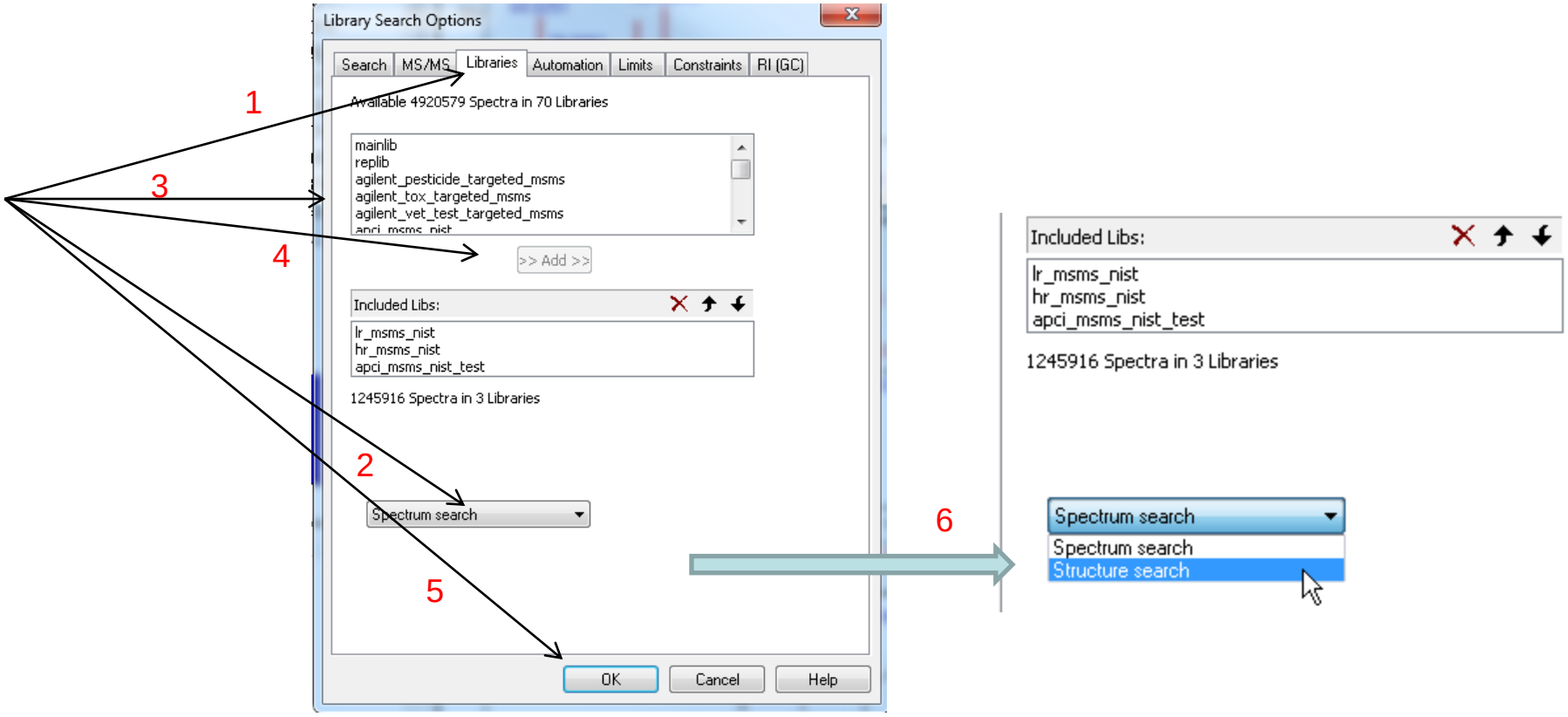
For example:

Note. For Low Resolution library (LR_*) product ions tolerance is fixed at unit mass resolution.

Setting Up Default Parameters for Similarity MS/MS Hybrid Search (Step 3)

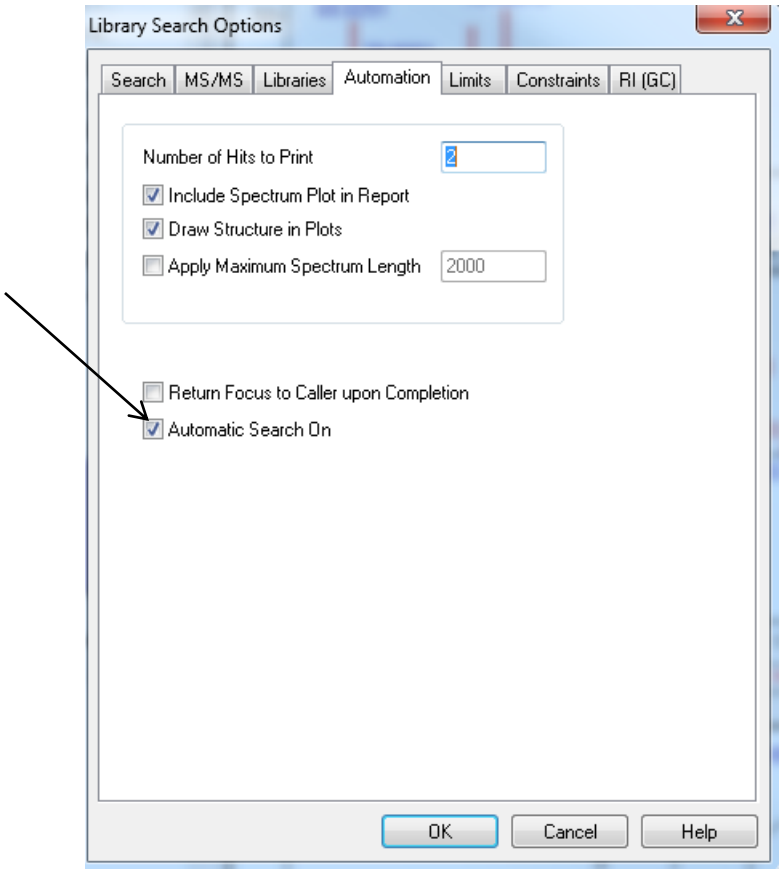
Select the Libraries Used for Spectral Searches in Library Search Options

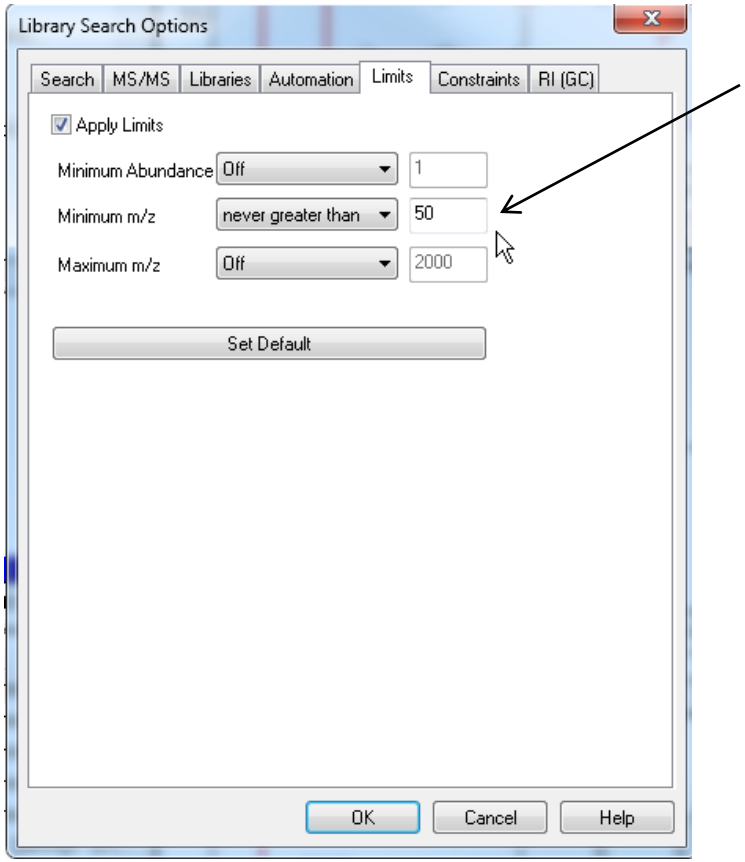
- ***LMB*** on the ***Libraries*** tab and make sure ***Spectrum Search*** is selected
- Select the group of libraries to be searched
- ***>>Add>>*** all libraries to be searched
- ***Order*** of libraries normally ***only important*** in Other Search Tab which return ***maximum no.*** of hits in Sequential Search Function
- Add ***same*** libraries for ***structure*** search, discussed in detail in ***Part V*** of series



Setting Up Default Parameters for Similarity MS/MS Hybrid Search (Step 4)

Selecting Up Limits and Automation Settings in Library Search Options



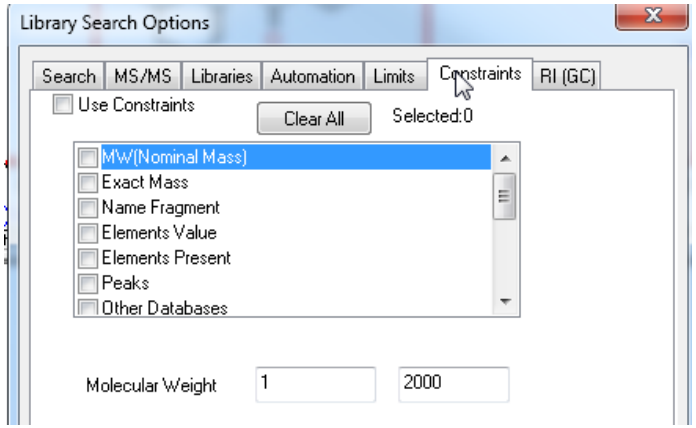


Setting Up Default Parameters for Similarity MS/MS Hybrid Search (Step 5)

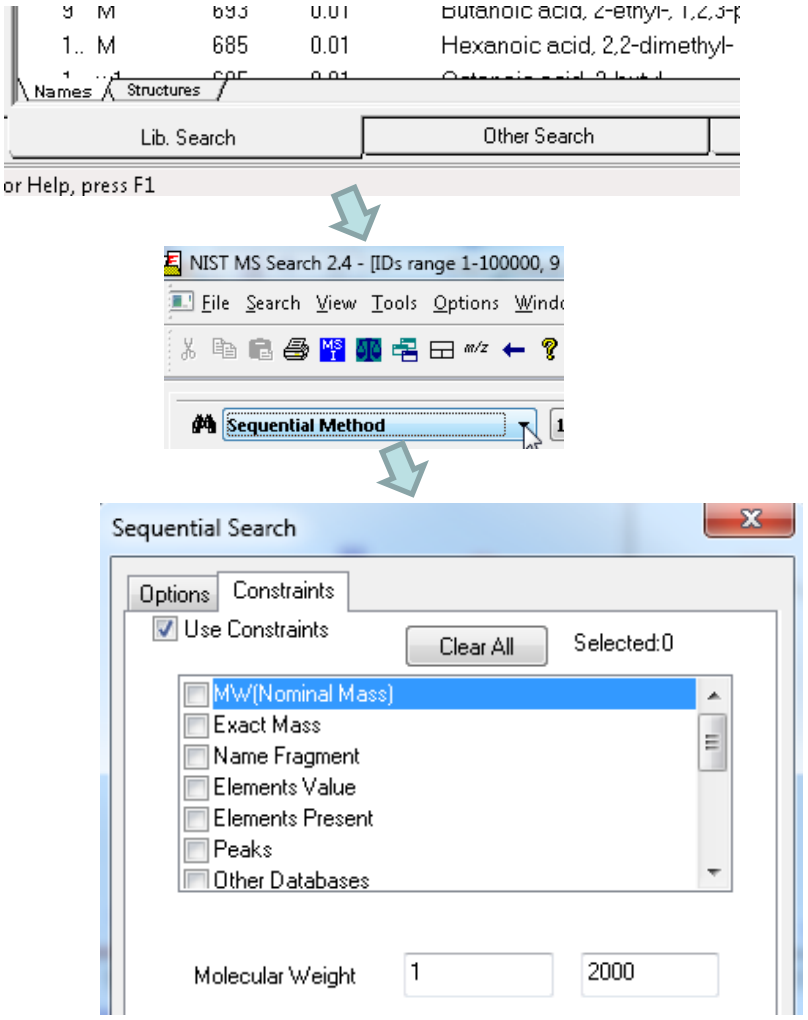
Constraints Applied to Hit List *After* the List Is Determined

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



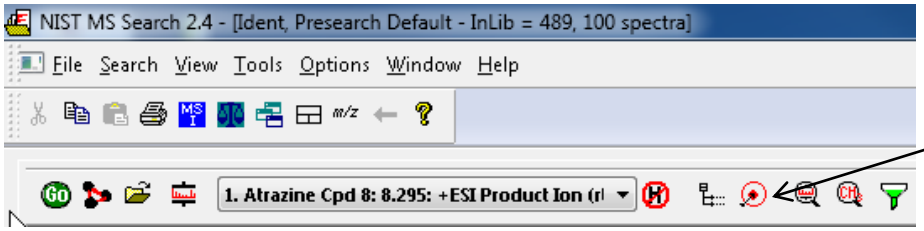
2



Setting Up Default Parameters for Similarity MS/MS Hybrid Search (Step 6)

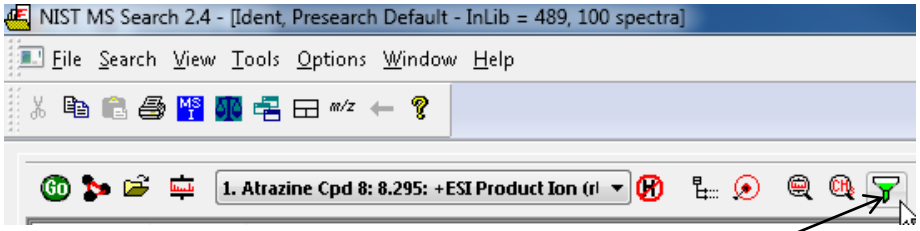
Ways of *Simplifying* Search Results Display

- These settings used to minimize spectra shown in library search results
- *Turn* them *off initially*, but use as needed to refine the results
- They can be used to limit the results *without repeating* the search
- The *1st* setting removes hits from list with same CAS number
- The *2nd* group of *filter settings* uses tags in library spectra to filter results



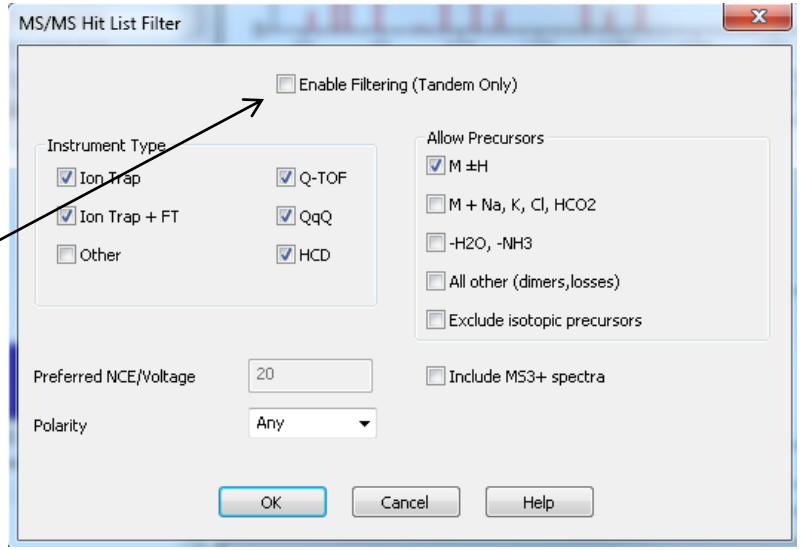
1st

-*Off* when not depressed
-*On* when depressed



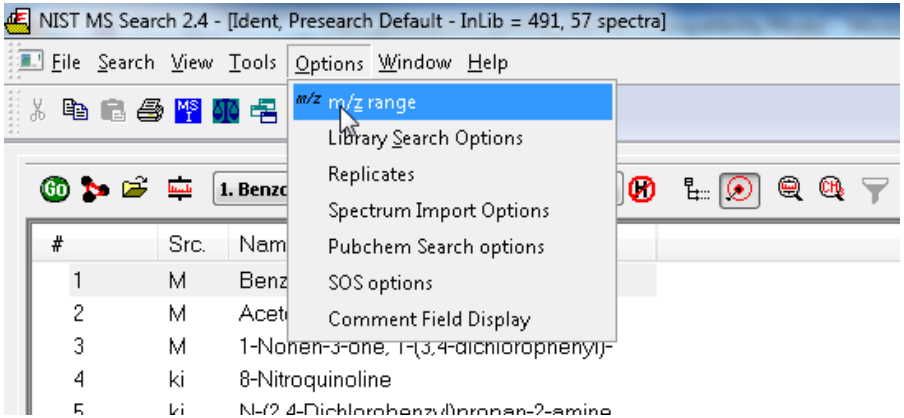
2nd

-*Access* Filter settings
-Initially *do not check* "Enable Filtering"

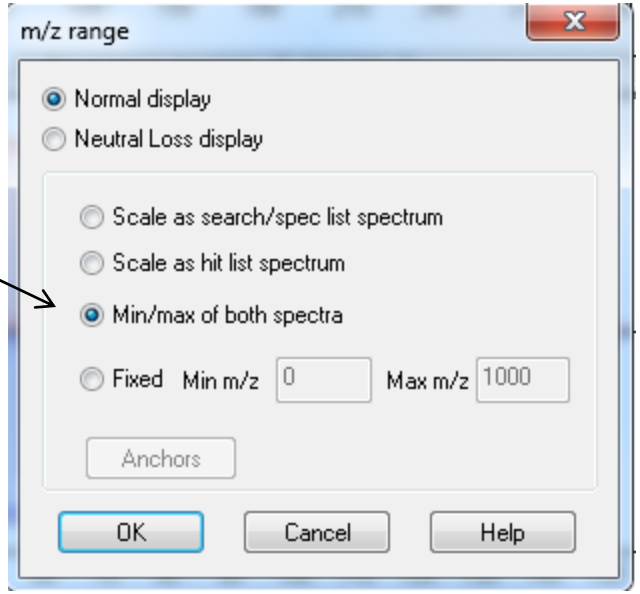


Setting Up Default Parameters for Similarity MS/MS Hybrid Search (Step 7)

Setting Default *m/z* Range Displayed in Spectra



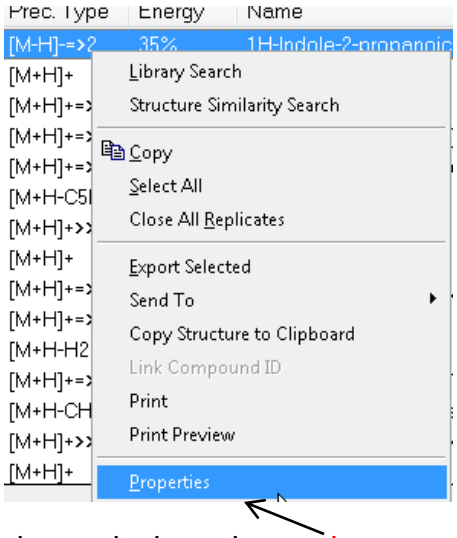
-I prefer min/max of both spectra



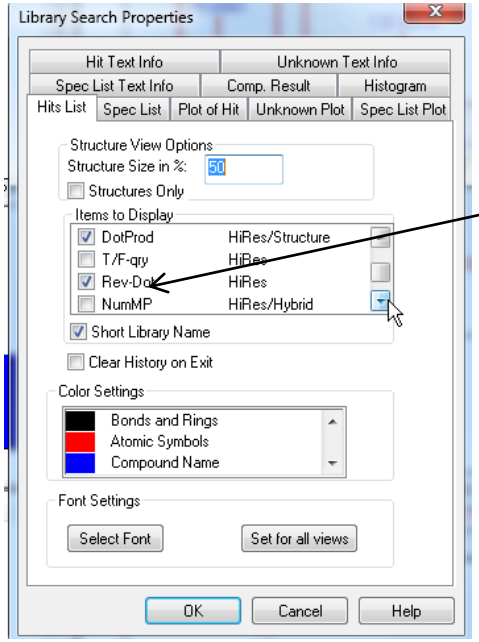
Setting Up Default Parameters for Similarity MS/MS Hybrid Search (Step 8)

Adding or *Changing* Columns in Results Display

- Useful to have "Rev-Dot" added as column in Identity MS/MS Search Results
- Can left click on column header and resort results by Rev-Dot
- Rev-Dot score is not penalized for ions found in unknown spectrum not found in reference spectrum



1) **RMB** anywhere in results box, then *select* properties



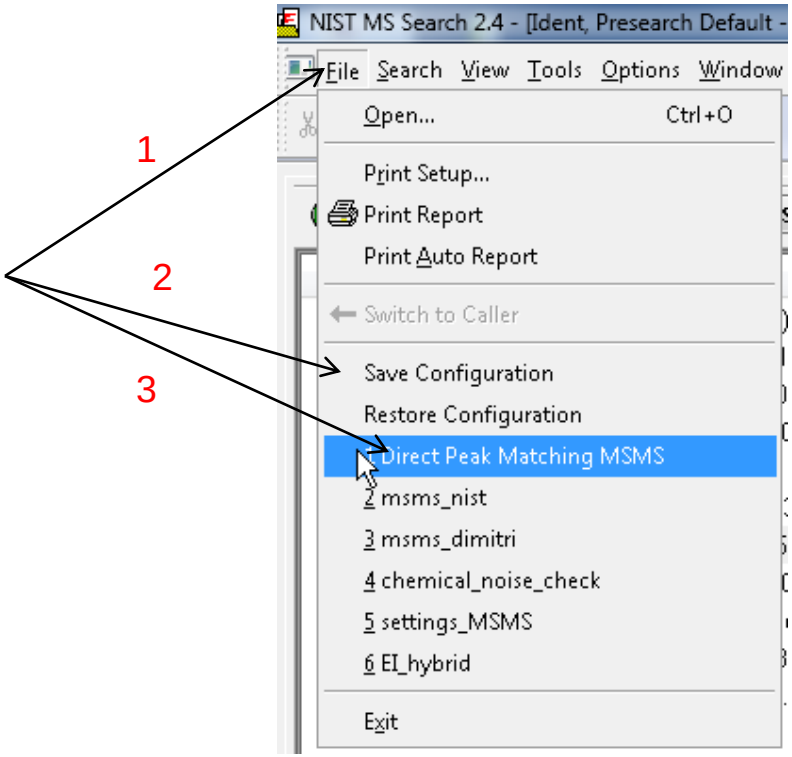
2) *Select* columns needed to reproduce display in bottom window, can *sort by any column* desired by clicking on heading after search is complete, the o-score is part of the hybrid score based on "direct peak matching" so similar to that obtained in first search, "EI Simple Similarity"

#	Lib.	Score	DotProd	o.Score	o.DotProd	DeltaMass	Prec. Type	Energy	Name
1	hr	999	999	999	999	-0.0000	[M+H]+	35%	Atrazine [M+H]+ IT-FT 35
2	hr	985	997	676	974	-14.0157	[M+H]+	35%	Sebuthylazin [M+H]+ IT-F
3	hr	965	979	965	979	-0.0000	[M+H]+	35%	Atrazine [M+H]+ HCD 35
4	hr	938	982	938	982	-0.0000	[M+H]+	40%	Atrazine [M+H]+ HCD 40
5	hr	936	990	3	49	-12.0267	[M+H]+	35%	Ametryne [M+H]+ HCD 3
6	lr	932	992	932	992	-0.0000	[M+H]+	35%	Atrazine [M+H]+ IT 35%

Setting Up Default Parameters for Similarity MS/MS Hybrid Search (Step 9)

Saving Setup of Standard Search Parameters

- After all the parameters are setup for search, save them for future use
- Example setup below was named ***Similarity Hybrid MSMS***
- In the future, this can be "restored" by selecting it from this menu



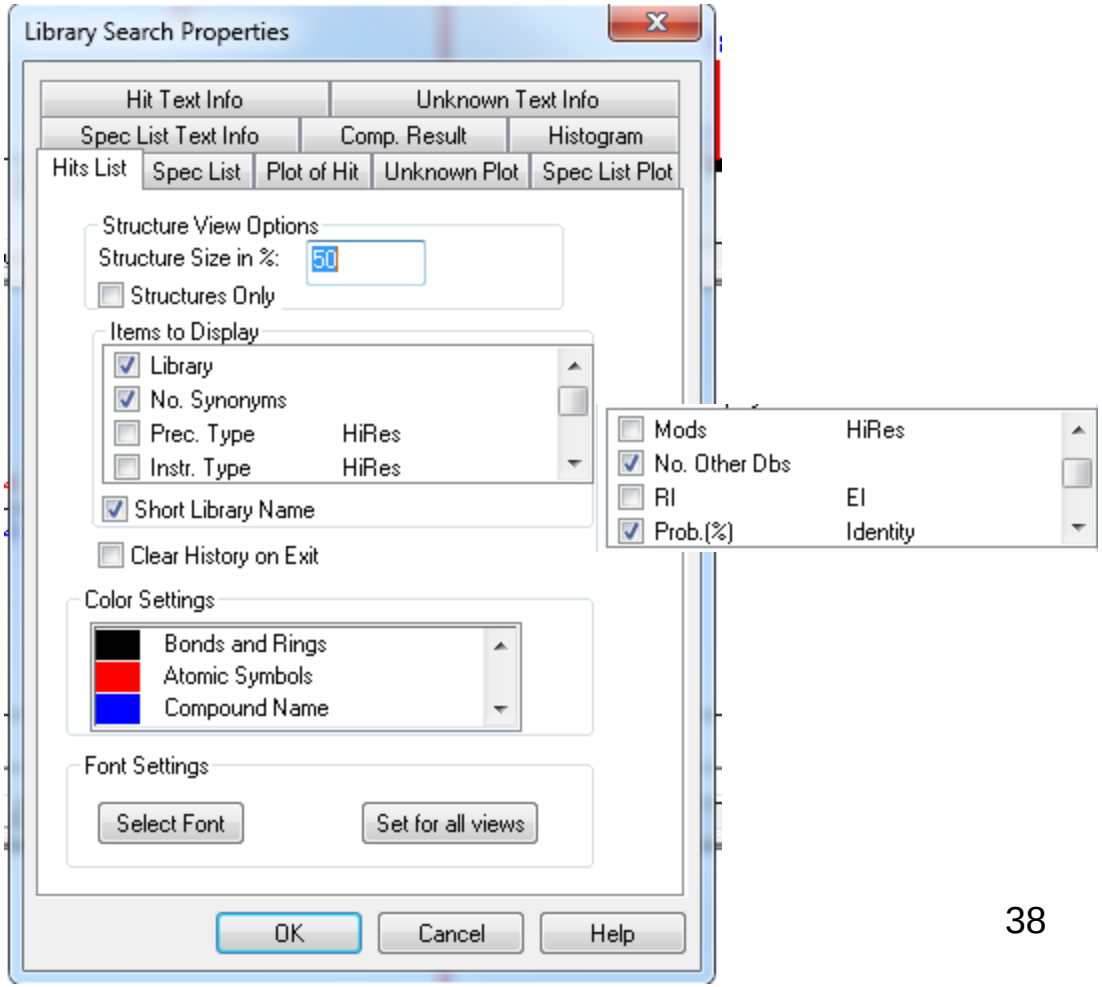
Searching Demonstration

-Example searches demonstrated in live presentation on YouTube video

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!



Presentation References (*Internet Links*)

1. *James Little Mass Spectral Resource Website*
2. *NIST Search Software Detailed Manual*
3. *Chemical Ionization for MW Determination*
4. *Trimethylsilyl Derivatives for GC-MS*
5. *Methyl Ester Derivatives for GC-MS*
6. *Lipid Matrix Ionization Effects in LC-MS*
7. *SciFinder/ChemSpider and Accurate Mass LC-MS Data for Unknown ID's*
8. *Surfactant Identification*
9. *NIST Flyer Detailing 2020 MSMS (Tandem) Library*
10. *NIST Flyer: New Hybrid Search and MS Interpreter for EI and Tandem Spectra*
11. *Description and Goals of NIST Tandem Libraries*
12. *Table Describing 2020 Tandem Libraries*
13. *Developing a Data Processing Pipeline for Extending a Comprehensive Tandem Mass Spectral Library*
14. *NIST Peptides Mass Spectral Program Description*
15. *Why are NIST Peptide Libraries Useful and How are They Used?*
16. *Detailed Table Listing Peptide Libraries and Downloads*
17. *New Developments in the Modeling of Ion Fragmentation by MS Interpreter Software*
18. *Enhancements to NIST MS Interpreter for Modeling High Mass Accuracy Tandem Mass Spectra*
19. *An Automated Method for Verifying Structure-Spectral Consistency Based on Ion Thermochemistry*
20. *Combining Fragment-Ion and Neutral-Loss Matching during Mass Spectral Library Searching: New General Purpose Algorithm Applicable to Illicit Drug Identification*
21. *The Hybrid Search: A Mass Spectral Library Search Method for Discovery of Modifications in Proteomics*
22. *Hybrid Search: A Method for Identifying Metabolites Absent from Tandem Mass Spectrometry Libraries*
23. *Structure Annotation of All Mass Spectra in Untargeted Metabolomics*
24. *YouTube Video for NIST EI Hybrid Search*
25. *Associated Handout for NIST EI Hybrid Search*
26. *Delta Mass Table Constructed from EI Hybrid Searches*
27. *Handouts and Other Resources for NIST EI Search*
28. *NIST Tandem Quick Start Guide*

Acknowledgements

- Stephen Stein (NIST)
- Dmitri Tchekhovskoi (NIST)
- David Sparkman (NIST Contractor)