LCMS Unknown Identifications Using MSMS Libraries
Part III: More Detailed Discussion of MSMS Hybrid Search
12/27/2020

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

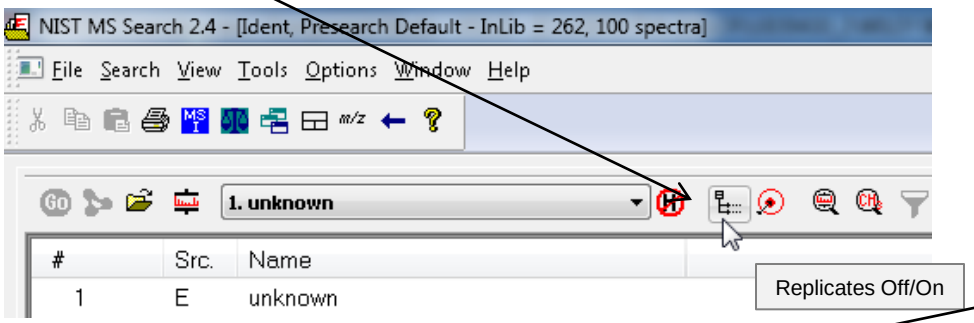
Table of Contents

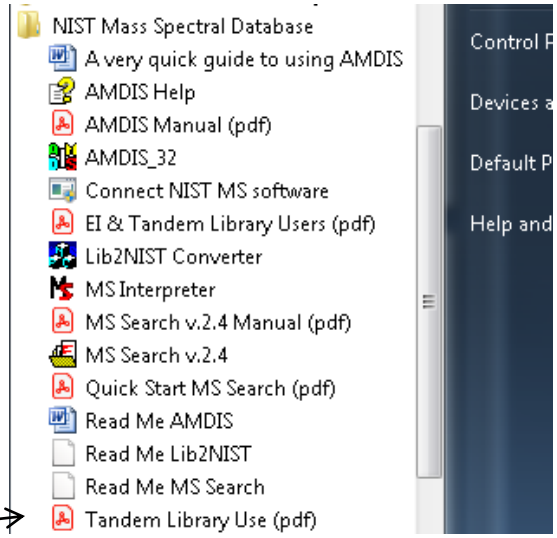
LCMS Unknown Identifications Using MSMS Libraries

Part 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

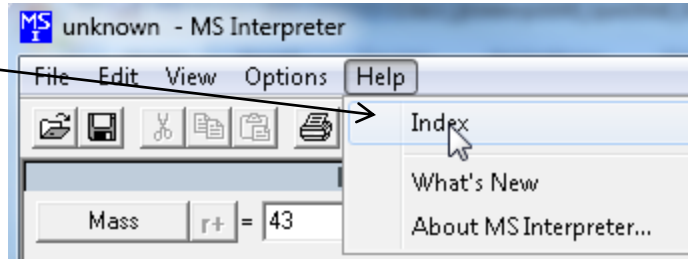
Help Files for NIST Search

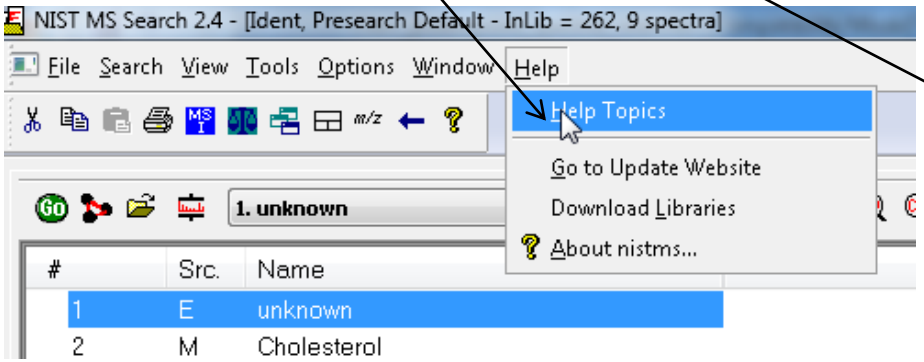
-"Hover" over Program Icon with mouse and function description displayed

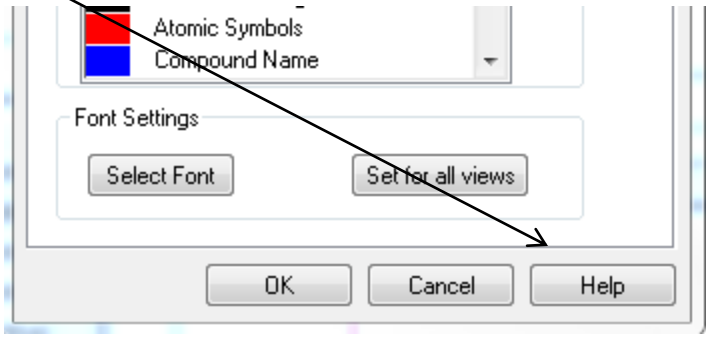




-***Detailed*** documentation for NIST Search[8]
-MS Interpreter included in NIST manual[8]
-Windows Program Group
-NIST Tandem Search ***Quick Start Guide***[9]
-"In program" assistance for both programs
-***Help button*** at corner of all windows



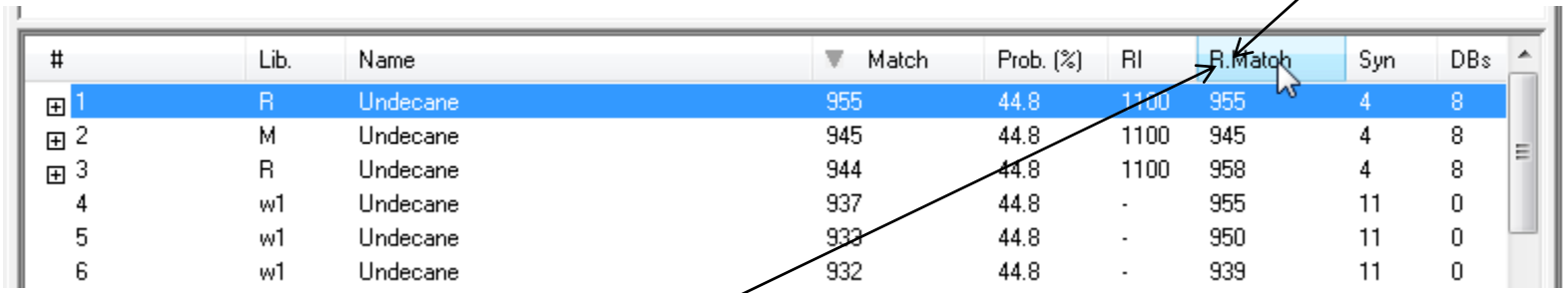




NIST Software in General is "Windows Compliant"

-***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!

LCMS Unknown Identifications Using MSMS Libraries

Part III: More Detailed Discussion of MSMS Hybrid Search

Presentation Topics:

- Description and personal experience
- Example of unknown identification with hybrid
- DotProd vs. Score in viewing search results
- DeltaMass table and examples
- Understanding head to tail plots
- Use of additional information in process
- Information on NIST MS/MS (Tandem) “hybrid” search
- Wiley MS/MS (Tandem) Library for Identity MS/MS searches

LCMS Unknown Identifications Using MSMS Libraries

Part III: More Detailed Discussion of MSMS Hybrid Search[9-12]

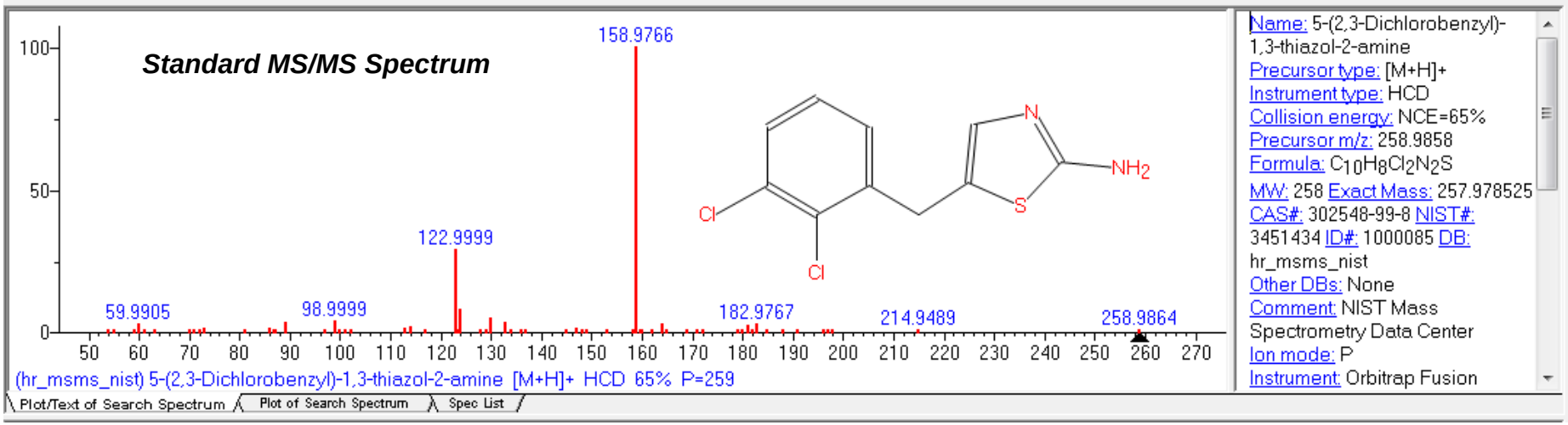
Program Description:

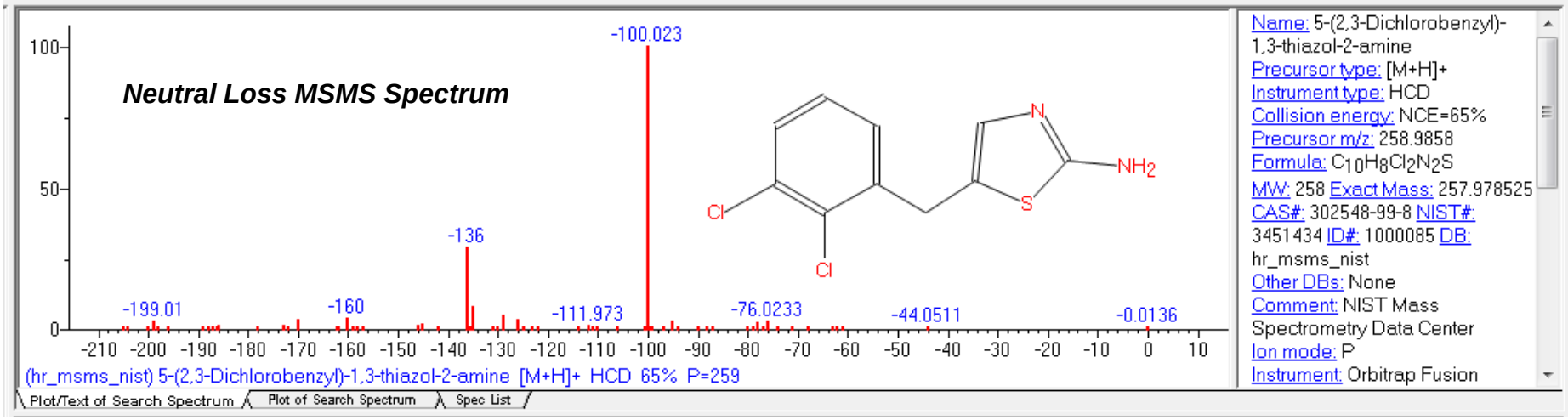
- Hybrid search generates a "hybrid" score matching fragments ***and*** neutral losses
- Extends the scope of the library by including "nearest neighbor"
- Success requires the presence of similar compounds in the library
- Mass difference must be confined to a single region of molecule and no significant alteration of fragmentation behavior
- *DeltaMass* is the mass difference between query and library compound and reflects the modification of the molecule

My Personal Experience:

- Personally used for over >20,000 EI searches[13-16]
- *Routinely amazed* by the types of similar compounds with high match factors
- Very frequently useful results not noted in "simple" (standard) similarity search
- Utility in finding similar model compounds, support for fragmentation mechanisms, and identification of unknowns
- Should be very useful in MS/MS searches for similar reasons
- ***In process, testing*** Hybrid MS/MS using ***many new entries*** present in NIST 20 version of library against NIST 17 library

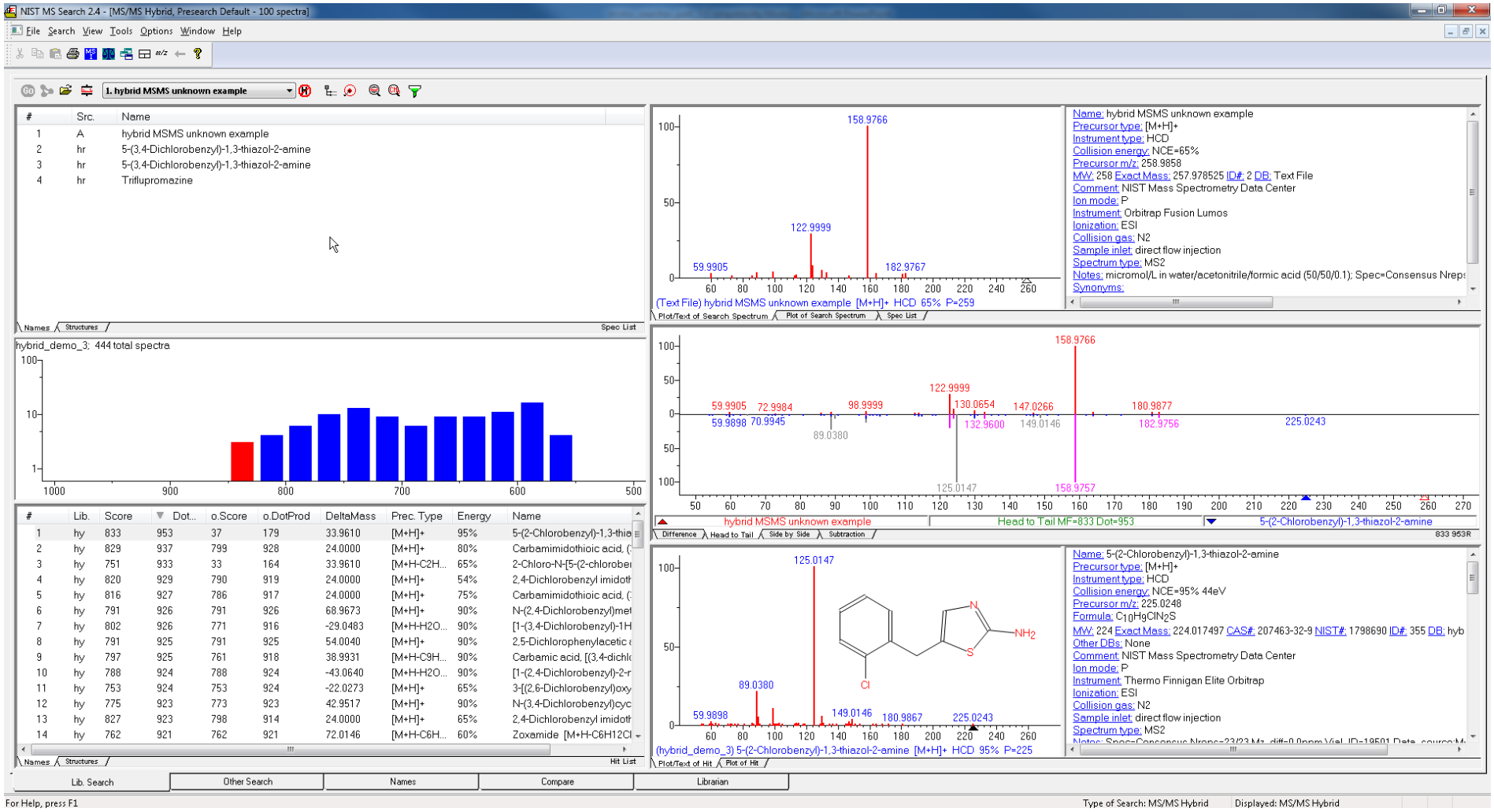
Hybrid Generates a *Similarity Score* on Matching *Both* Fragment Ions and Neutral Losses of Unknown to Library Reference Spectra[16-19,22]





Hybrid MSMS Search Results for Unknown*

1) Best match factor (DotProduct) is 953
2) ***DeltaMass*** is 33.9610
3) Chlorine (mass 34.9689) replacing hydrogen (mass 1.0078) on ring, i.e. 34.9689-1.0078=33.9611



*__Example__ using a spectrum from NIST 2020 hr_msms database (NIST# 1847874) as “unknown "and a small user created library of ~500 entries

Using Dot Product *Instead of* Score for Viewing Search Results

- Excerpt[17] from NIST Tandem Users Quick-Start
- Usually sort by **LMB** on DotProd instead of Score

Scores (reported in the **Lib Search** tab view's **Hit List** window)
When matching two product-ion mass spectra, **Scores** are reported as well as the **Dot Product**. The **Dot Product** is a simple mathematical measure widely used in reporting spectrum similarity. An exact match is reported as 999. The **Score** is adjusted for spectra with few peaks and is intended to adjust for the reduced selectivity when matching small numbers of peaks. For a single peak spectrum for example, a match will report a 999 for the **Dot Product** and significantly lower the **Score** depending on abundance. The difference between the **Dot Product** and **Score** will reflect an estimate of the uncertainty in identification due to the small number of peaks (generally less than four). This **Score** is simply a rough measure of identification confidence and has no well-defined statistical meaning.

LMB

#	Lib.	Score	DotProd	o.Score	o.DotProd	DeltaMass	Prec. Type	Energy	Name
1	hy	857	964	18	58	33.9610	[M+H]+	65%	5-(2-Chlorobenzyl)-1,3-thiazol-2-amine [...
2	hy	818	949	21	70	33.9610	[M+H]+	80%	5-(2-Chlorobenzyl)-1,3-thiazol-2-amine [...

DeltaMass Table[15]

- I created[15] a DeltaMass spreadsheet after looking at >20,000 spectral EI hybrid searches
- Also, useful for MSMS (tandem) searches
- E.g. below see entries for DeltaMass = 34 nominal
- Add accurate DeltaMass column ***in future***
- For now, user will need to manually use accurate mass

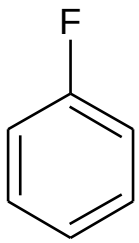
Need to add in future

34	pyridinyl group	dimethylamino group	33.9843
34	Sulfonamide group on aromatic	nitro on aromatic	34.9877
34	chlorine on aromatic ring	phenyl ring	33.9611
34	CF3 on aromatic	chlorine replacing	34.0263

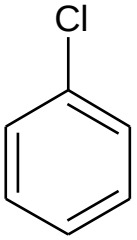
Only reasonable choice in our example

Associating Some *Very Simple* Structures with DeltaMass Values[15]

- Some simple ***small*** MW compounds to illustrate types of substructural information
- Of course, these substructures can be a part of ***much larger*** molecules
- ***Note:*** Odd values of DeltaMass contain one nitrogen change in structure, thus ***"Nitrogen Rule"***
- Isotope ratios and/or accurate mass helpful with redundancies

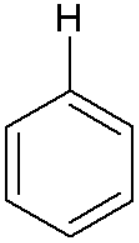


MW 96

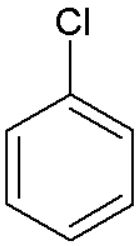


MW 112

DeltaMass 16

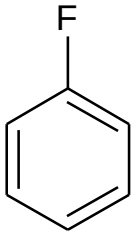


MW 78

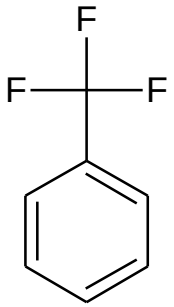


MW 112

DeltaMass 34

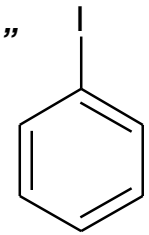


MW 96

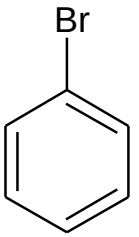


MW 146

DeltaMass 50

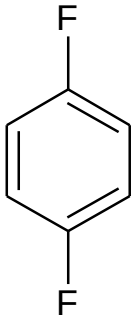


MW 204

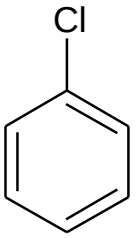


MW 156

DeltaMass 48

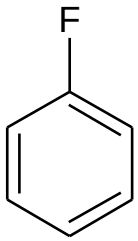


MW 114

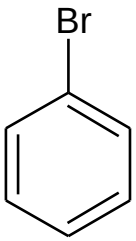


MW 112

DeltaMass 2

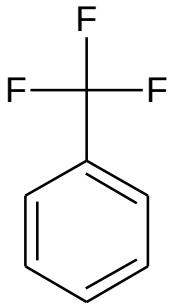


MW 96

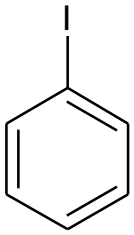


MW 156

DeltaMass 60

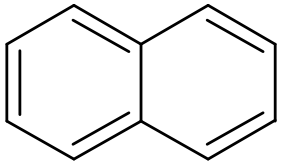


MW 146

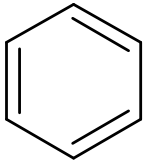


MW 204

DeltaMass 58

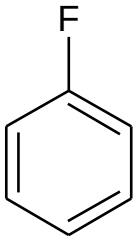


MW 128

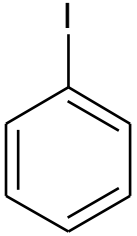


MW 78

DeltaMass 50

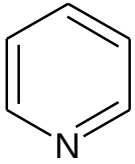


MW 96

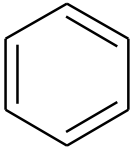


MW 204

DeltaMass 108

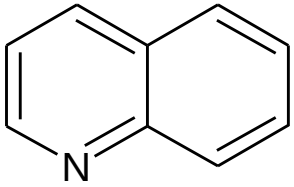


MW 79

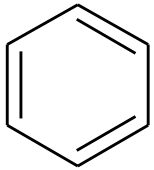


MW 78

DeltaMass 1

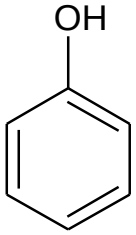


MW 129

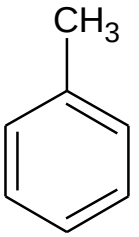


MW 78

DeltaMass 51



MW 94



MW 92

DeltaMass 2

Tip: See Normal "Peak Matching" Results in MSMS Hybrid Search Hit List

Add Column and Sort

1. **LMB** on *o-DotProd* column header to resort by "Peak Matching" part of Hybrid scoring
2. ***Top 30*** hits contain the substructure with ***2*** chlorines atoms on benzyl ring
3. Consistent with addition of chlorine on ring of best hybrid fit structure

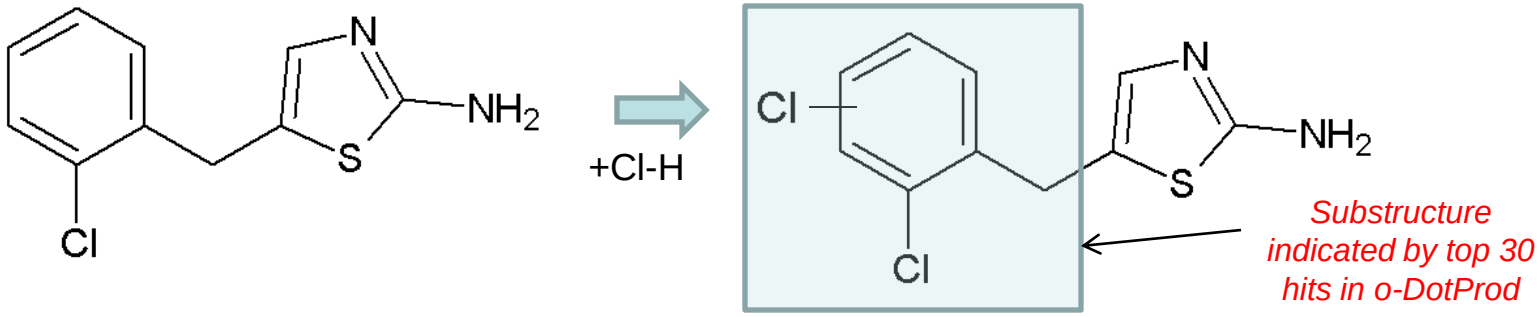
1

#	Lib.	Score	DotProd	o.Score	o.DotProd	DeltaMass	Prec. Type	Energy	Name
1	hy	833	953	37	179	33.9610	[M+H]+	95%	5-(2-Chlorobenzyl)-1,3-th
2	hy	829	937	799	928	24.0000	[M+H]+	80%	Carbamimidothioic acid
3	hy	751	933	33	164	33.9610	[M+H-C2H...	65%	2-Chloro-N-[5-(2-chlorob

Resorted by o-DotProd

#	Lib.	Score	DotProd	o.Score	o.DotProd	DeltaMass	Prec. Type	Energy	Name
1	hy	829	937	799	928	24.0000	[M+H]+	80%	Carbamimidothioic a
2	hy	791	926	791	926	68.9673	[M+H]+	90%	N-(2,4-Dichlorobenz
3	hy	791	925	791	925	54.0040	[M+H]+	90%	2,5-Dichlorophenyla
4	hy	788	924	788	924	-43.0640	[M+H-H2O	90%	[1-(2,4-Dichlorobenz
29	hy	662	904	662	904	-156.008	[M+H]+	80%	Miconazole [M+H]+
30	hy	608	904	608	904	-47.0225	[M+H]+	65%	1-(3,4-Dichlorobenzy

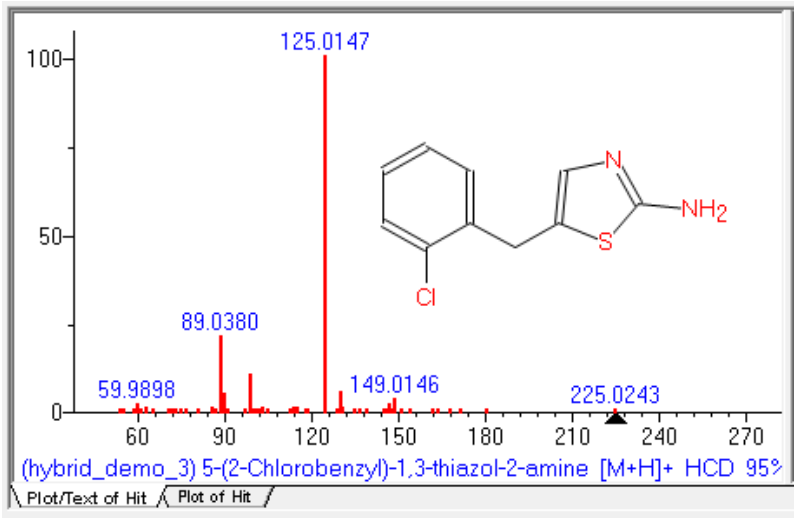
2



Substructure indicated by top 30 hits in o-DotProd

"Merge" Information of "Hybrid" and "Peak Matching Searches" *Using* Accurate Mass Data for Fragments, Precursor, and DeltaMass

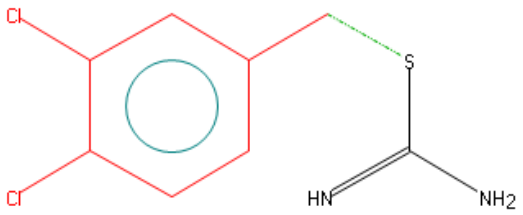
"Hybrid" Best Result
Accurate Mass DeltaMass, +Cl-H

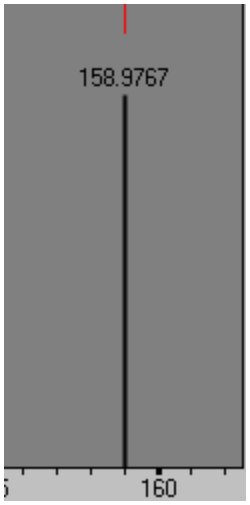


Top 30 "Peak Matching Hits' for 2 Cl's on benzyl ring

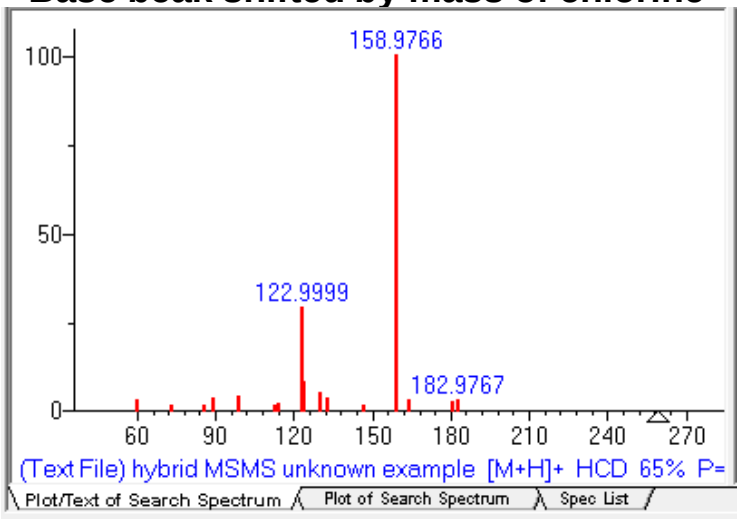
Maximum Rate: 65(chain) @ 158.9763 m/z

m/z	exact m/z	formula	loss	type	rate	rel.rate	ion.
159	158.976282	$C_7H_5Cl_2$	CH_4N_2S	H-Displacement	65	78	5

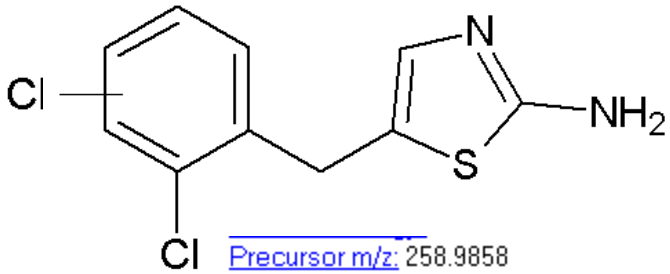




"Unknown" Spectrum
Base beak shifted by mass of chlorine



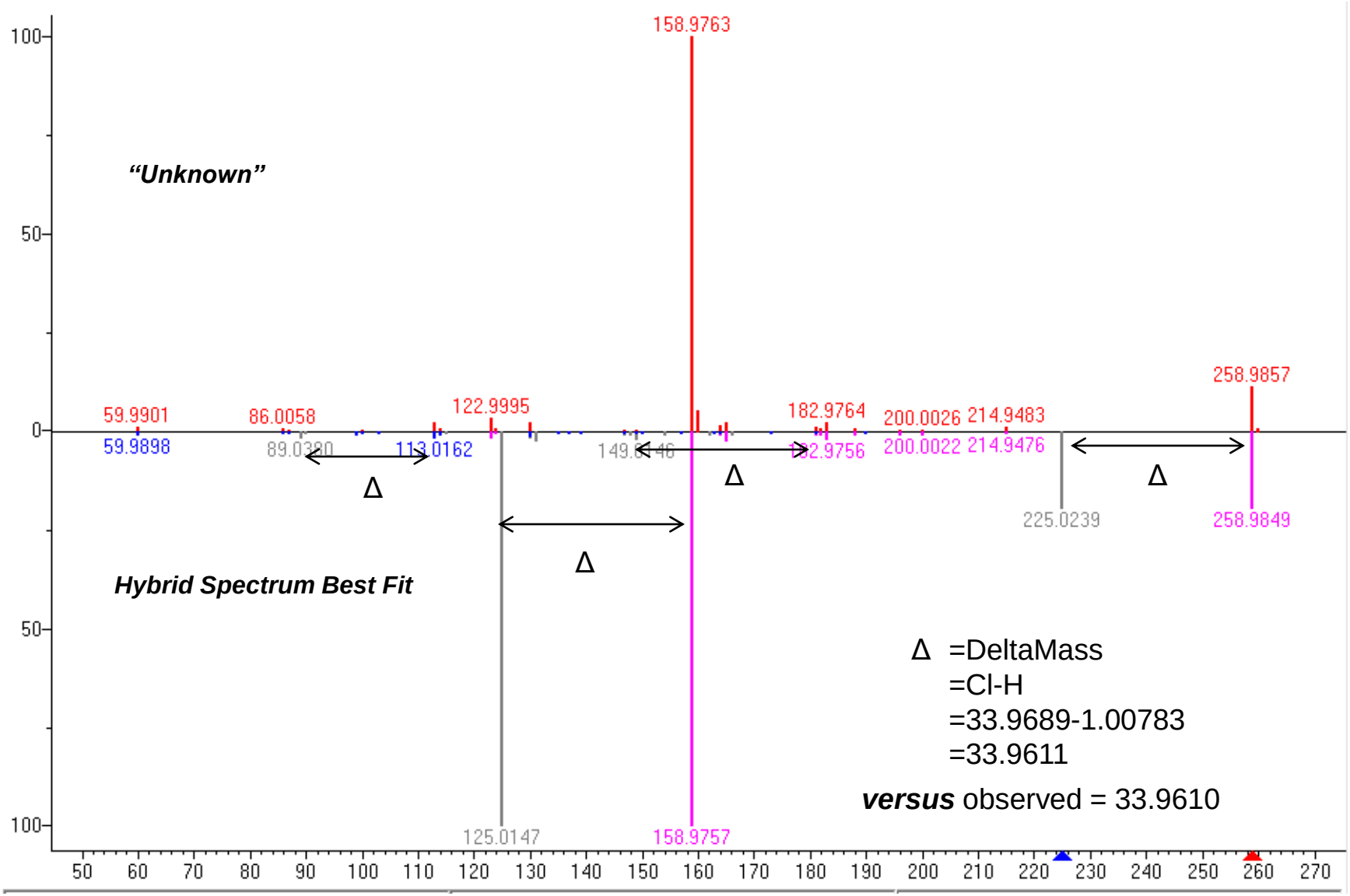
"Unknown Identity"



Precursor m/z: 258.9858
Formula: $C_{10}H_8Cl_2N_2S$+H

Closer Look at Middle Display
Top is Unknown, Bottom is "Hybrid" Spectrum

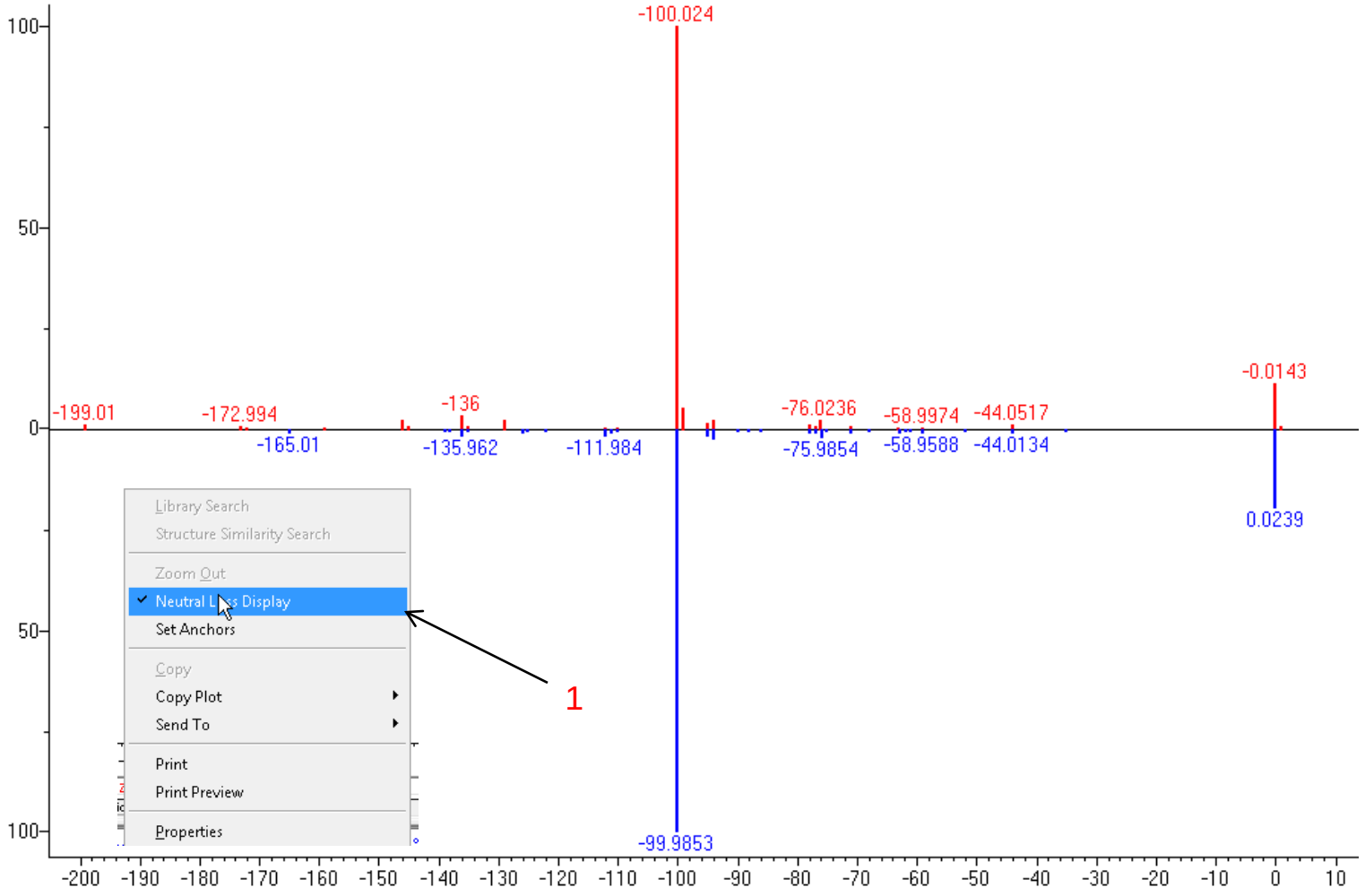
- In ***bottom*** spectrum, original ions in grey shifted by ***DeltaMass*** (Δ) to violet for user visual comparisons
- Ions ***not s***hifted in bottoms stay in "blue"
- Takes a while to adjust to this view versus standard "Head to Tail" views



Alternate Comparison of Hybrid Spectrum: Neutral Loss Display

Top is Unknown, Bottom is "Hybrid" Spectrum

1. **RMB** in spectrum window, **LMB** select Neutral Loss Display
2. Shows whole spectrum of reference shifted by DeltaMass of 33.9611
3. ***More efficient*** to ***look at "Hybrid" display*** with experienced eye!



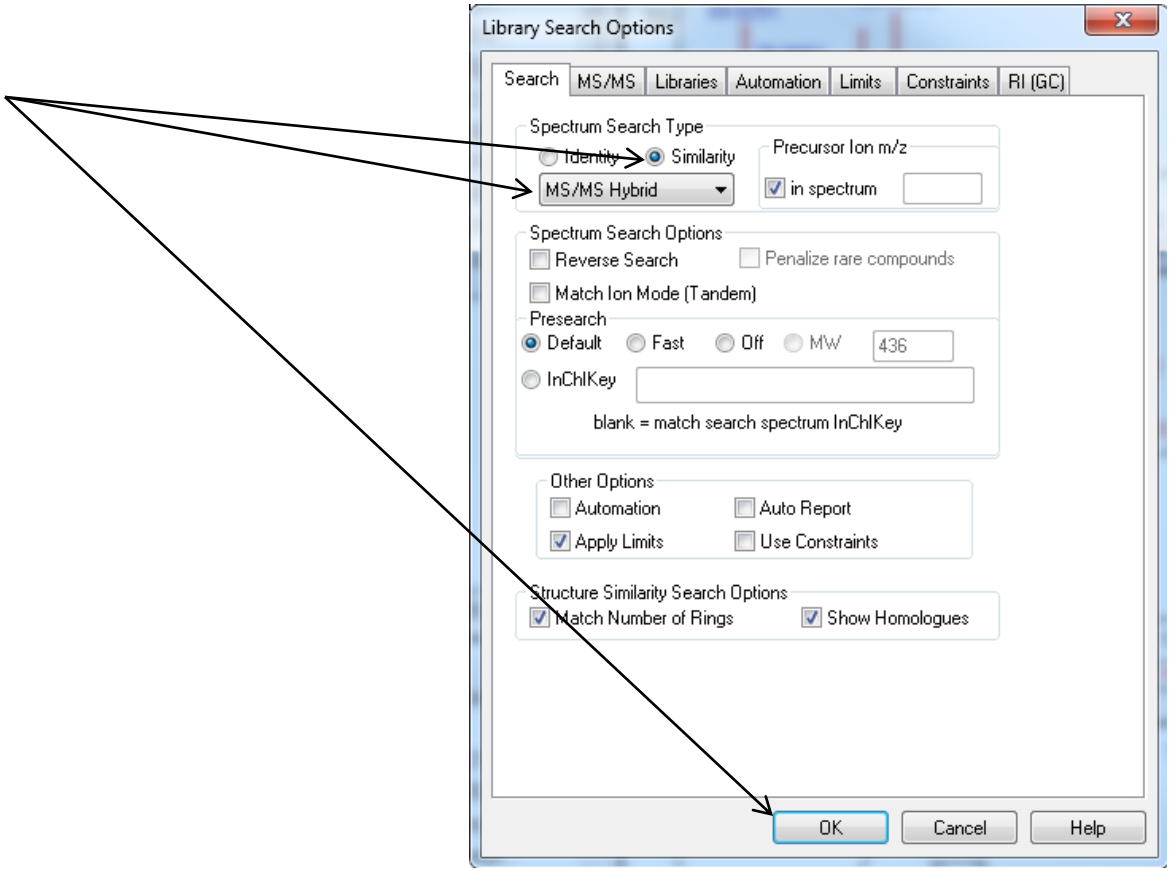
Other Things to Consider when Interpreting Hybrid Search for Unknown Identification

- Determining molecular formula using isotope ratios
- Sample history
- Logical fragments using MS interpreter
- Other identified/related components found in sample
- Information from organic chemist/sample submitter
- Other techniques such as NMR, IR, elemental analysis, deuterium exchange for active hydrogens (infusion), derivatization, etc..

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

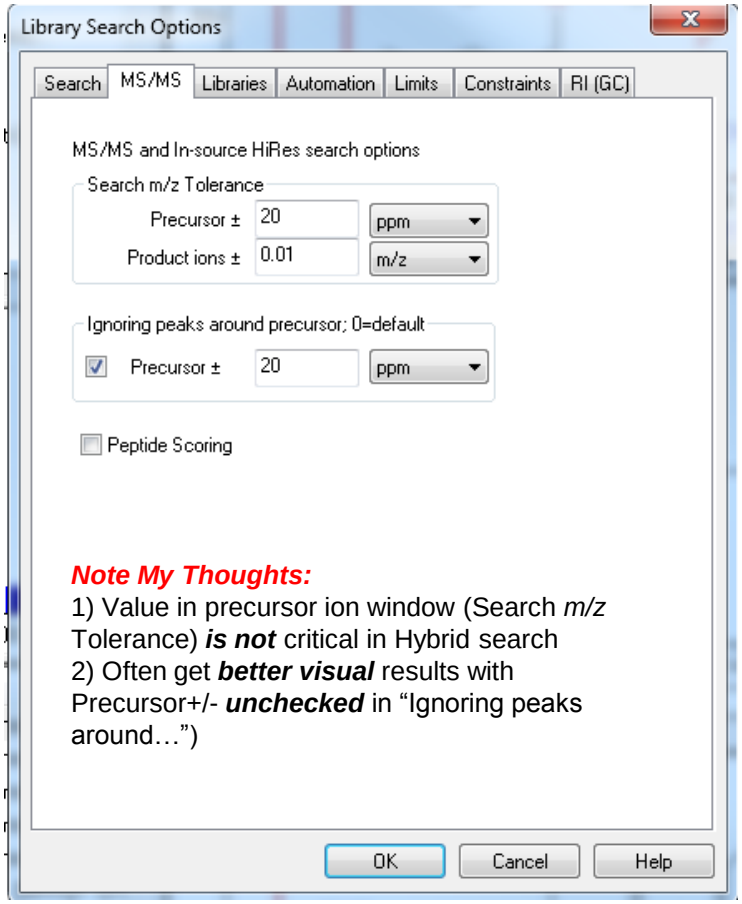


Setting Up Default Parameters Similarity MS/MS Hybrid Search (Step 2)

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.

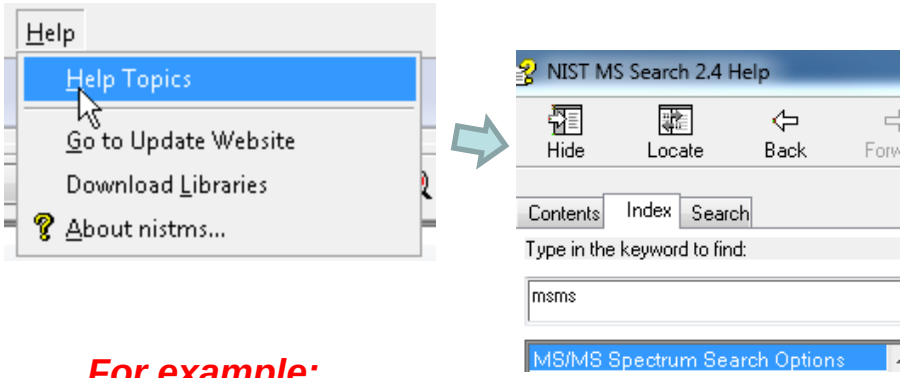


Note My Thoughts:

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



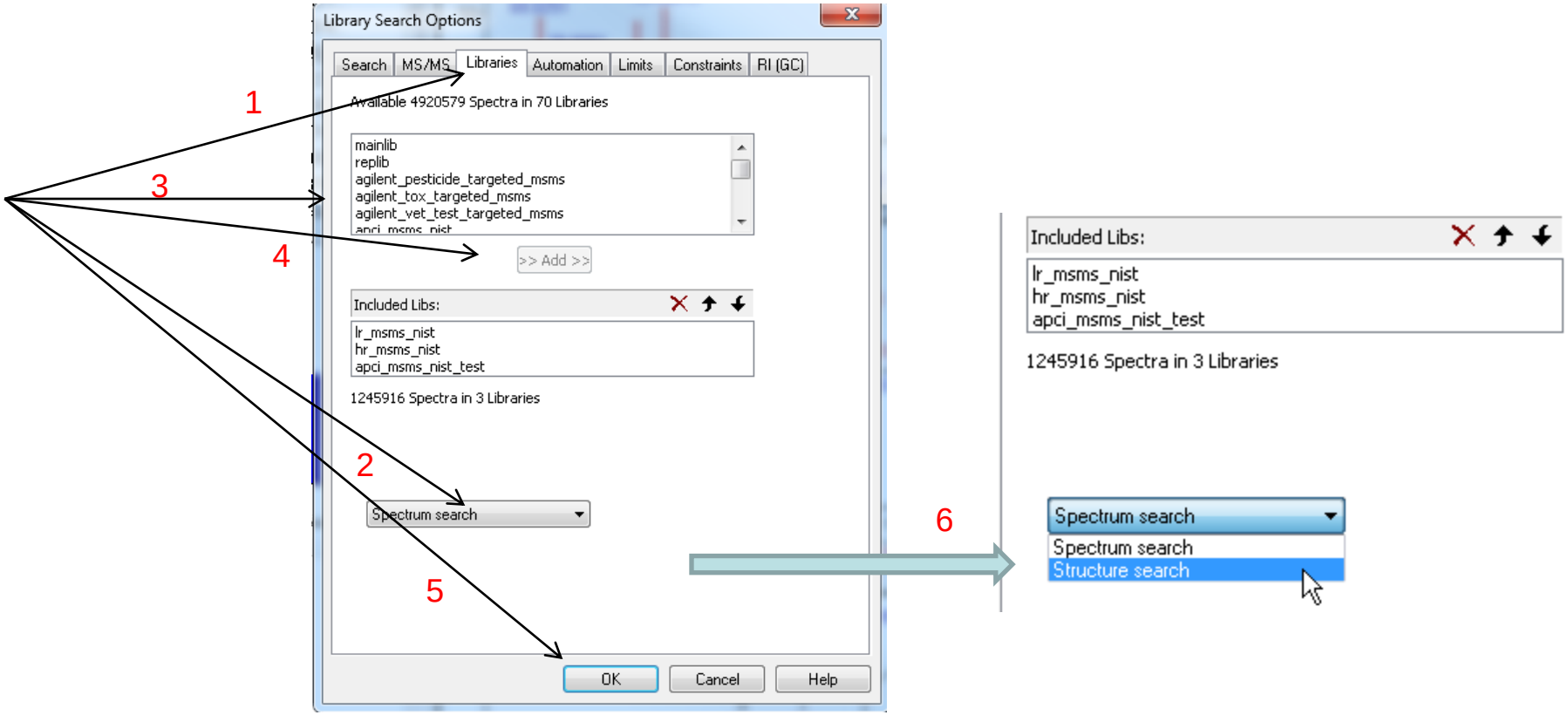
For example:

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)

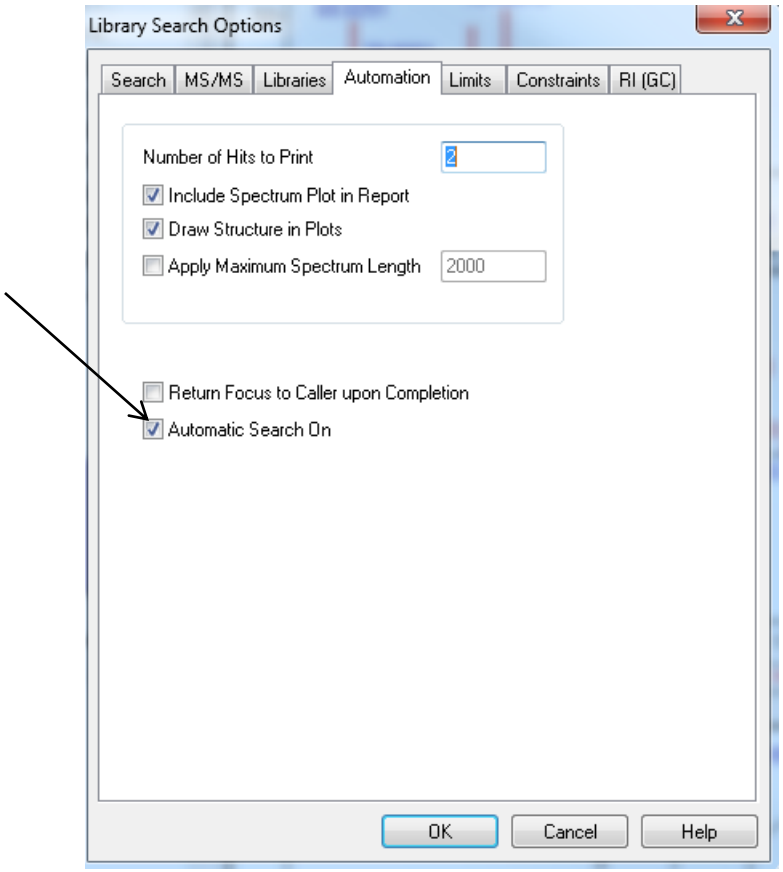
Select the Libraries Used for Spectral Searches in Library Search Options

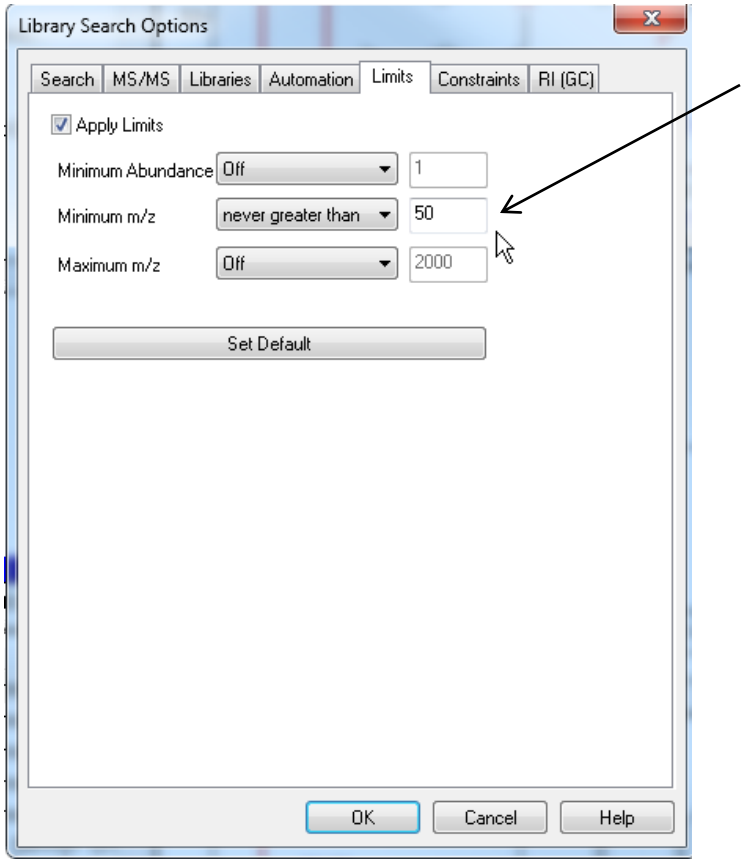
- ***LMB*** on the ***Libraries*** tab and make sure ***Spectrum Search*** is selected
- Select the group of libraries to be searched
- ***>>Add>>*** all libraries to be searched
- ***Order*** of libraries normally ***only important*** in Other Search Tab which return ***maximum no.*** of hits in Sequential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



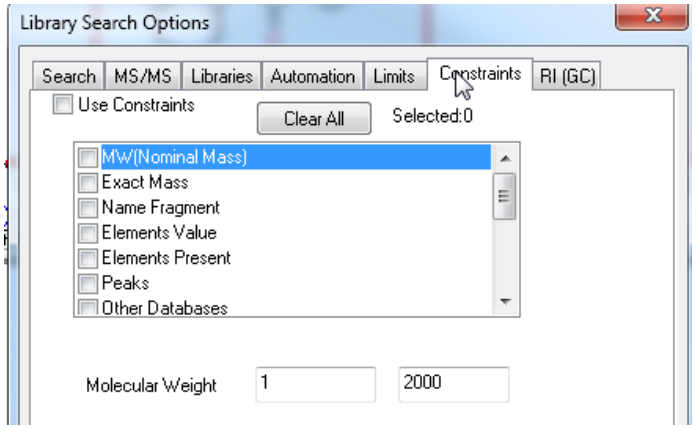


Setting Up Default Parameters for Similarity MS/MS Hybrid Search (Step 5)

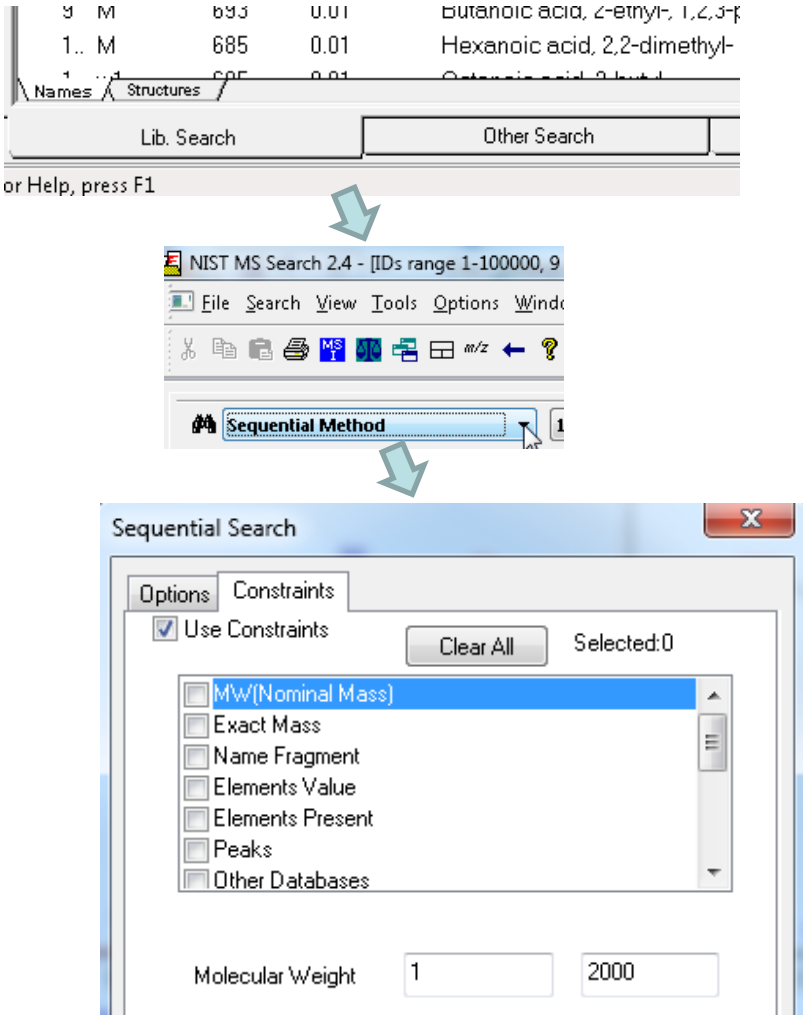
Constraints Applied to Hit List *After* the List Is Determined

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



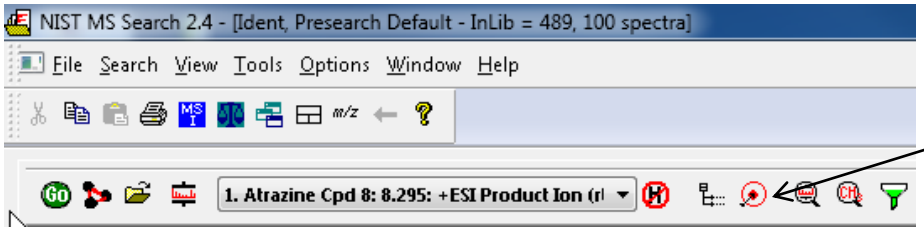
2



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

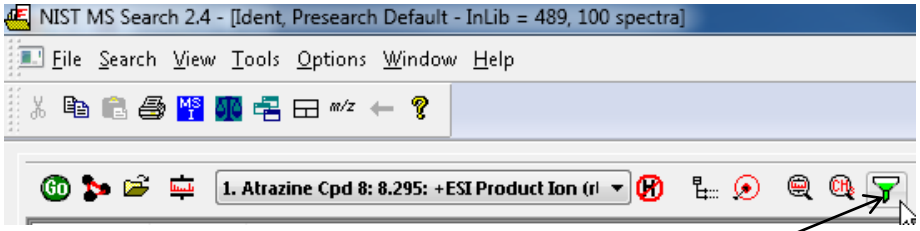
Ways of *Simplifying* Search Results Display

- 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



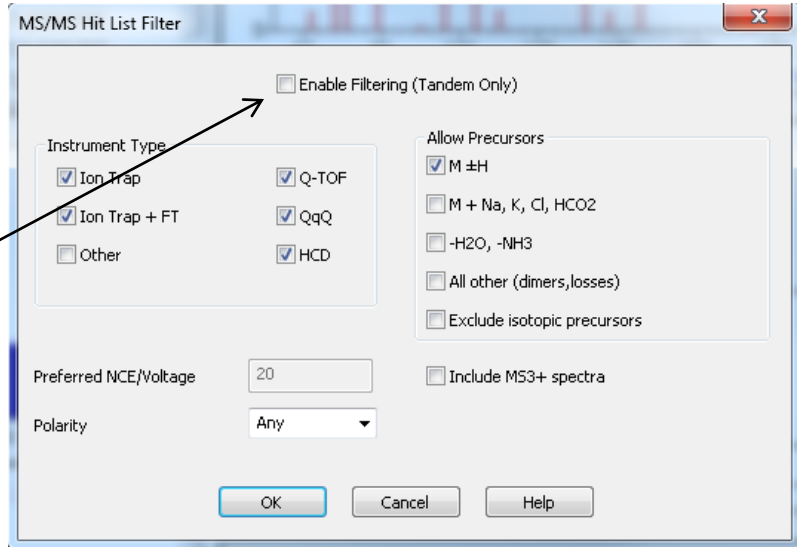
1st

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



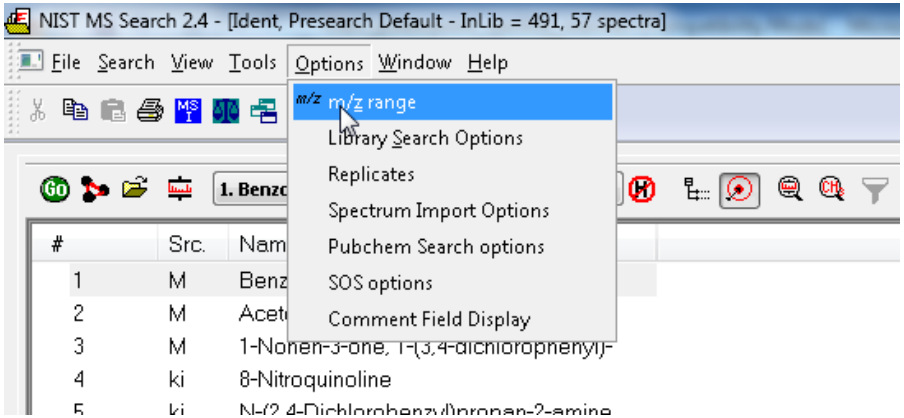
2nd

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

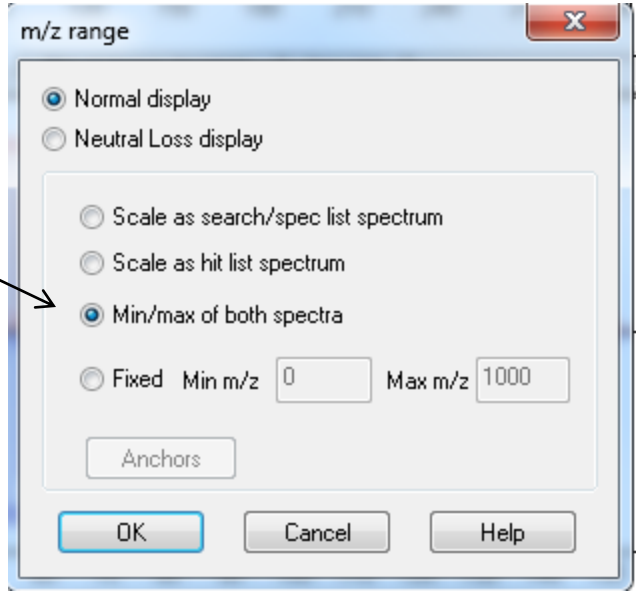


Setting Up Default Parameters for Similarity MS/MS Hybrid Search (Step 7)

Setting Default *m/z* Range Displayed in Spectra



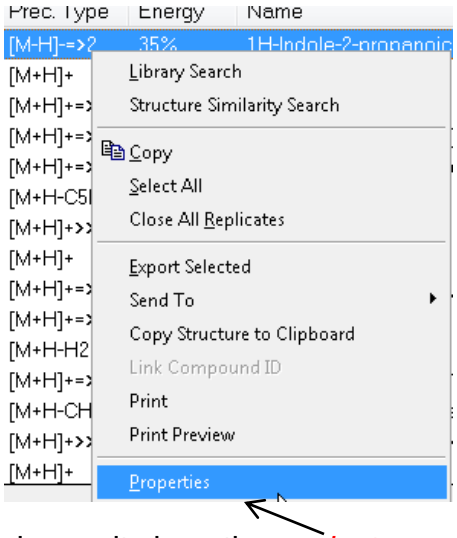
-I prefer min/max of both spectra



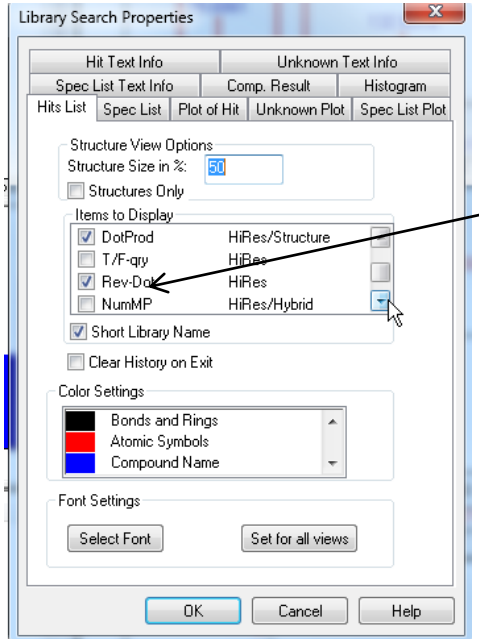
Setting Up Default Parameters for Similarity MS/MS Hybrid Search (Step 8)

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



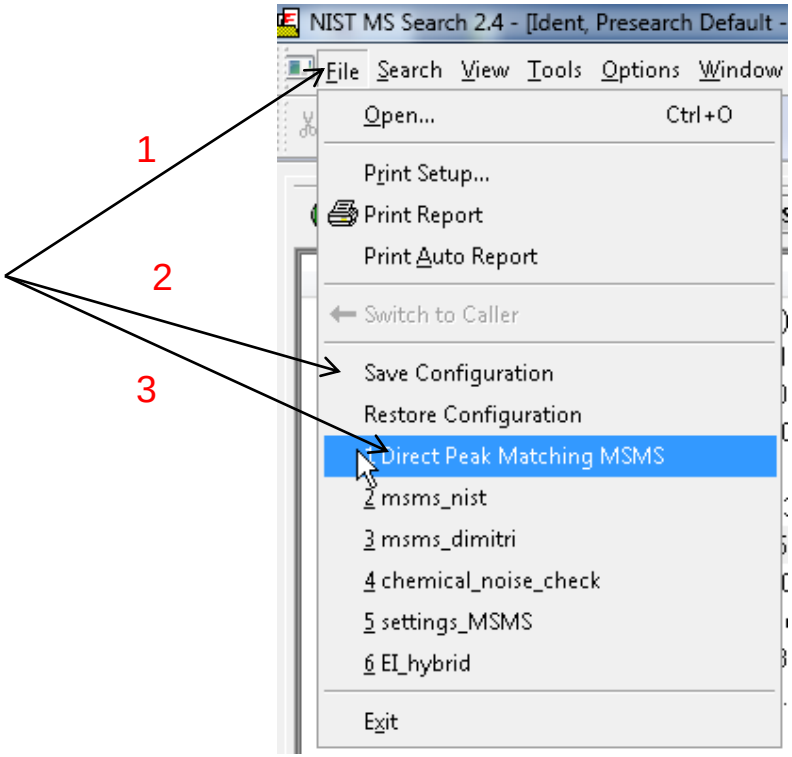
2) *Select* columns needed to reproduce display in bottom window, can *sort by any column* desired by clicking on heading after search is complete, the o-score is part of the hybrid score based on "direct peak matching" so similar to that obtained in first search, "EI Simple Similarity"

#	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



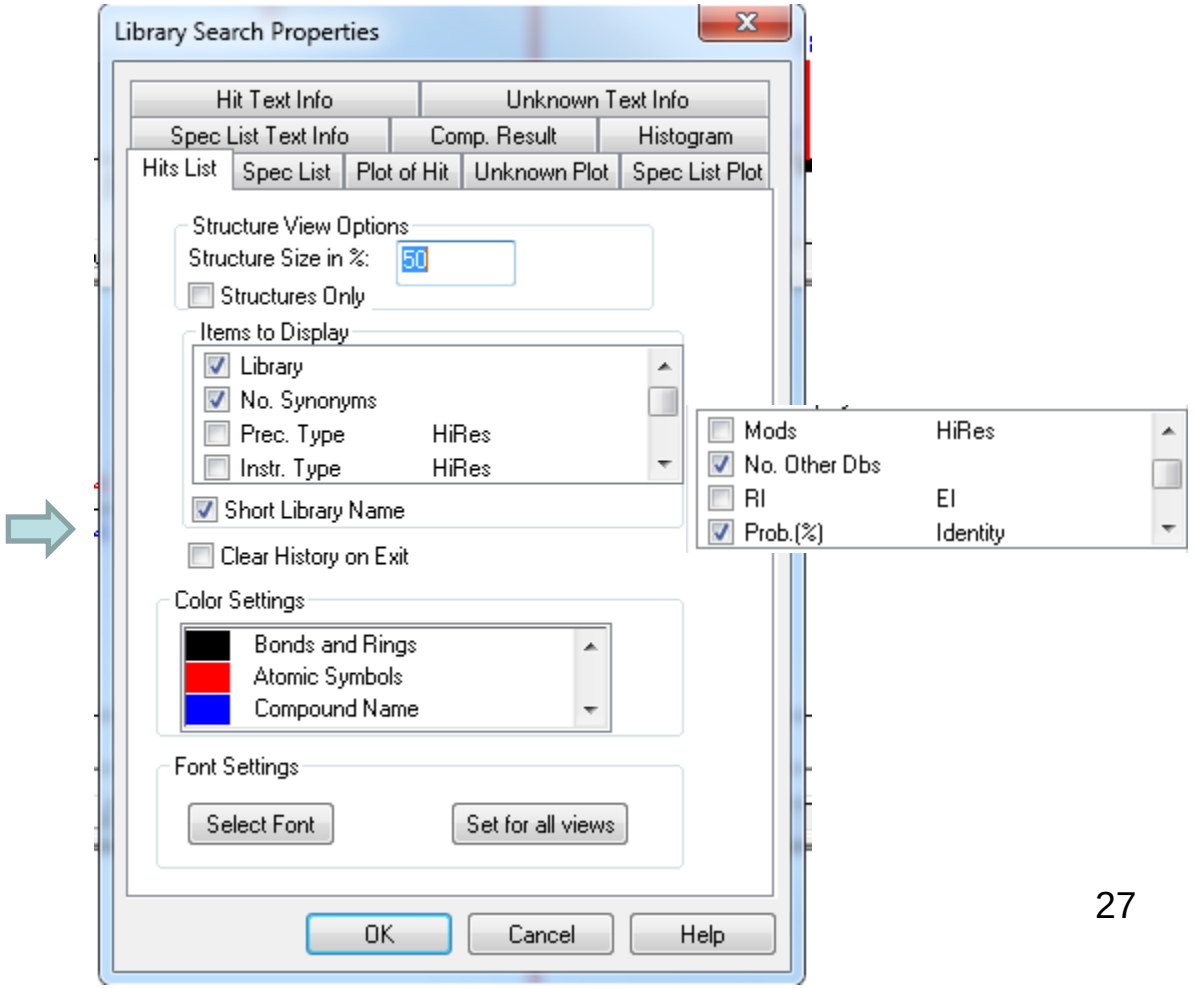
Searching Demonstration

-Example searches demonstrated in live presentation in 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. *Combining Fragment-Ion and Neutral-Loss Matching during Mass Spectral Library Searching: New General Purpose Algorithm Applicable to Illicit Drug Identification*
10. *The Hybrid Search: A Mass Spectral Library Search Method for Discovery of Modifications in Proteomics*
11. *Hybrid Search: A Method for Identifying Metabolites Absent from Tandem Mass Spectrometry Libraries*
12. *Structure Annotation of All Mass Spectra in Untargeted Metabolomics*
13. *YouTube Video for NIST EI Hybrid Search*
14. *Associated Handout for NIST EI Hybrid Search*
15. *Delta Mass Table Constructed from EI Hybrid Searches*
16. *Handouts and Other Resources for NIST EI Search*
17. *NIST Tandem Quick Start Guide*

Acknowledgements

- Stephen Stein (NIST)
- Dmitri Tchekhovskoi (NIST)
- David Sparkman (NIST Contractor)
- Arun Moorthy (NIST)
- Gary Mallard (NIST Contractor)
- Sandy Markey (NIST Contractor)