

Version 3.5A user guide

Edited by Sohyeon Jeon (sjeon@wustl.edu)

TABLE OF CONTENTS

Chapter 1. Getting started	3
1.1 Files necessary for the AMS spectra database comparison panel	3
1.2 How to load the AMS spectral database and create the AMS spectral comparison panel in an Igor experiment	3
1.3 AMS/ACSM Database Export Tool	6
1.4 How to convert HR MS to UMR family grouped MS	7
Chapter 2. Unit mass resolution (UMR) mass spectra comparison	9
2.1 The selection of a mass spectrum of interest	9
2.1.1 User supplied 1D/2D UMR MS	9
2.1.2 MS Existing in the data base	11
2.2 A comparison result of the sample with reference mass spectrum and options	11
2.2.1 UMR Comparison Result Table	11
2.2.2 UMR Comparison Result Mass Spectrum Plot	12
2.2.3 Mass Spectra Rescaling Options	12
2.2.4 Comparison Constraints	13
2.3 Citation information	14
Chapter 3. High resolution to unit mass resolution (HR) spectra comparison	15
3.1 The selection of a mass spectrum of interest	15
3.1.1 User supplied 1D/2D HR MS	15
3.1.2 MS existing in the data base	16
3.2 A comparison result of the sample with reference mass spectrum and options	16
3.2.1 Result Table	16
3.2.2 Mass spectrum plot	17
3.2.3 HR family selection	18
3.2.4 Mass Spectra Rescaling Options	19
3.3.5 Comparison Constraints	19
3.4 Citation information	19
Chapter 4. Database library	20
Chapter 5. Export database	22

Chapter 1. Getting started

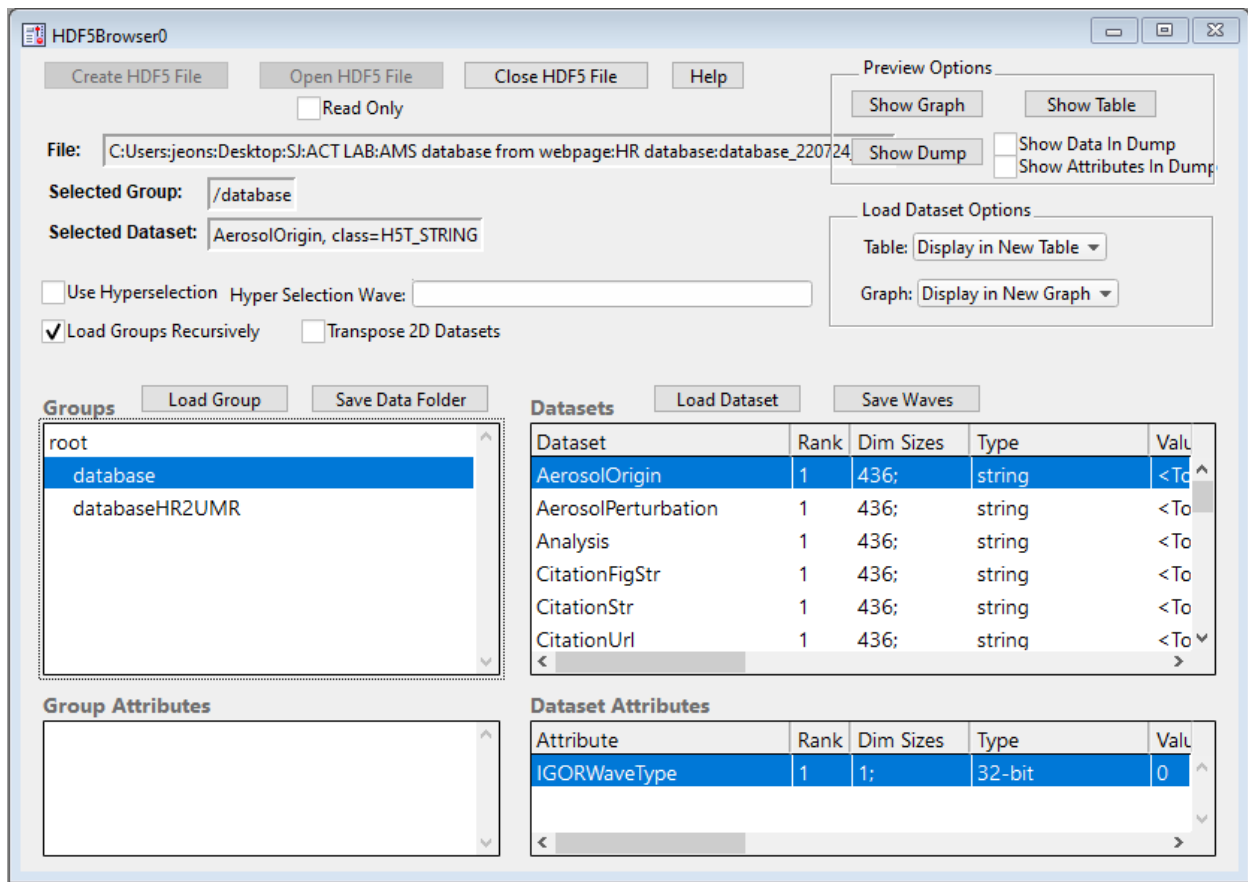
1.1 Files necessary for the AMS spectra database comparison panel

In order to use the AMS spectral database in Igor Pro, users need two sets of files: i) the HDF (Hierarchical Data Format, or h5) database file and ii) two database procedure files (ipf) named ‘AMS MS database_v**.ipf’ and ‘AMS MS database_HR2UMR_v**.ipf’. Users can download the ipf files from the existing webpage (<https://cires1.colorado.edu/jimenez-group/AMSsd/>). Alternatively, users can download and open the Igor template experiment file (AMS_database_v**_template.pxt) for the AMS database, which already has the database and procedure files already loaded. The template file can be downloaded on the same webpage mentioned above. The panel and all other code described herein are compatible with Igor versions 7.08, 8.04, and 9.02 for Macintosh and Microsoft operating systems.

1.2 How to load the AMS spectral database and create the AMS spectral comparison panel in an Igor experiment

The Igor Pro template experiment file (AMS_database_v**_template.pxt) already has the database and necessary procedure files loaded. However, when the user creates a new database h5 file or doesn't have the template experiment file, the user can follow the processes below

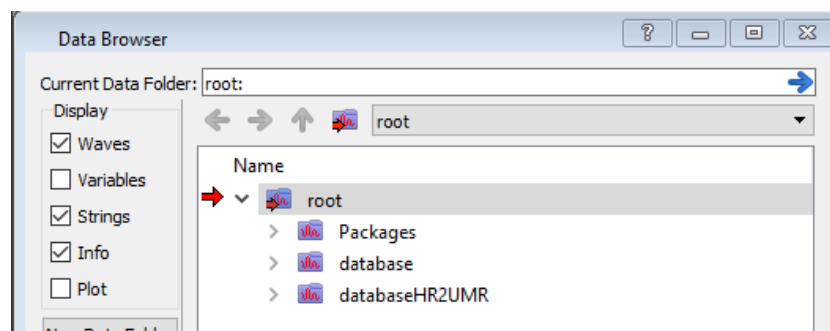
- 1) Open Igor pro program.
- 2) On the top Igor menu, click [File] > [Open File] > [Procedure...] and open both procedure files: ‘AMS MS database_v**.ipf’ and ‘AMS MS database_HR2UMR_v**.ipf’ (Or drag both procedure files to the Igor Pro experiment). Compile the procedure files.
- 3) On the menu, select [Data] > [Load waves] > [New HDF5 Browser]
- 4) On the panel, select the ‘Open HDF5 file’ button and select the database h5 file that you want to import <Figure 1.1> (Database h5 files must have ‘database’ and ‘databaseHR2UMR’ in the file name. Refer to ‘Chapter 5. Export the database’ to create a database h5 file).



<Figure 1.1. HDF5 Browser panel when AMS database.h5 file has been opened.>

- 5) On the 'Groups' section, select 'database' and click the 'Load Group' button and select 'databaseHR2UMR' and click the 'Load Group' button again. <Figure 1.1>

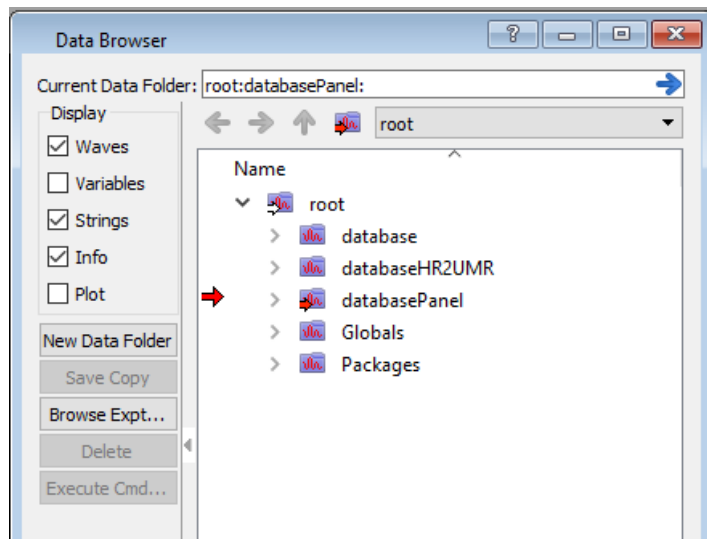
Caution: Both the 'database' and 'databaseHR2UMR' folders have to be uploaded in the 'root:' folder like below:



<Figure 1.2. Required locations of 'database' and 'databaseHR2UMR' folders>

- 6) Click the 'Close HDF5 File' button on the panel to close the HDF5 Browser panel.

- 7) On the top Igor menu, click [Database] > [Initialize] to create 'databasePanel' and 'Globals' folders in the 'root:' like <Figure 1.3>



<Figure 1.3. Correct final folder locations to use the database panel>

- 8) On the top Igor menu, select [Database] > [Database Panel] to open the database panel.
- 9) Database is now ready to use!

1.3 AMS/ACSM Database Export Tool

For a more rigid standardization of data format, it is highly recommended to use AMS/ACSM database export tool within Igor experiments to create an HDF file that combines many spectra into one file with metadata fields. The created HDF file will be used on the ‘UMR Data Comparison’ and ‘HR Data Comparison’ tabs of the panel as ‘New MS (refer to 2.1.1, 3.1.1)’. A new stand-alone ipf for this tool is being provided as a link on the existing website (<https://cires1.colorado.edu/jimenez-group/AMSSd/>).

Users should enter the metadata of their MS through the interface (Figure 1.4) and based on the entered metadata the HDF file will be named and created. Metadata categories on the export tool follow the ‘Comparison Constraints’ fields of the database panel (refer to 2.2.4). After the user has created an HDF file using the gold button in the panel, the HDF file generated from this tool will contain these entities (a wave in the Igor software is a matrix of numeric or text data):

- 1D numeric wave called ‘mz’
- 1D text wave called ‘mzlabel’
- 1D text wave called ‘SpectraComments’ as metadata
- 1D text wave called ‘SpectraName’ as metadata
- 2D numeric matrix called ‘Spectra’ as MS
- 2D text wave called ‘SpectraMetaData’ as metadata

Aerosol Data Base Export
AMS/ACSM Database Export Tool 2.0.3

UMR Database
 HR Database
 CV Database

Spectra DAQ metadata **Analysis metadata** **CV Database**

Aerosol origin: Select
 Type: Select
 Opt. comment: optional description Pop

Aerosol treatment: Select
 Type: Select
 Opt. comment: optional description Pop

Analysis: Select
 Instrument: Select
 Spectral resolution: Select
 Vaporizer type: Select
 One word descriptor for this file: test Verify metadata

Other DAQ descriptors
 Vaporizer temp. (deg. C): 600
 EI energy (eV): 70
 Opt. comment: i.e. special lens Pop

Citation
 Citation url: i.e. http://direct.sref.org/xxxx-yyyy/acp/2009-x-yyyy
 Citation fig #: Fig #(s)
 Comments: location, date, field project, misc info Verify metadata

Spectrum selection
 Add spectra from the top table or graph Verify Clear table Pop table

Full path of m/z axis wave
 Full path of m/z axis wave labels

Spectra name	Full path of wave	Optional comments

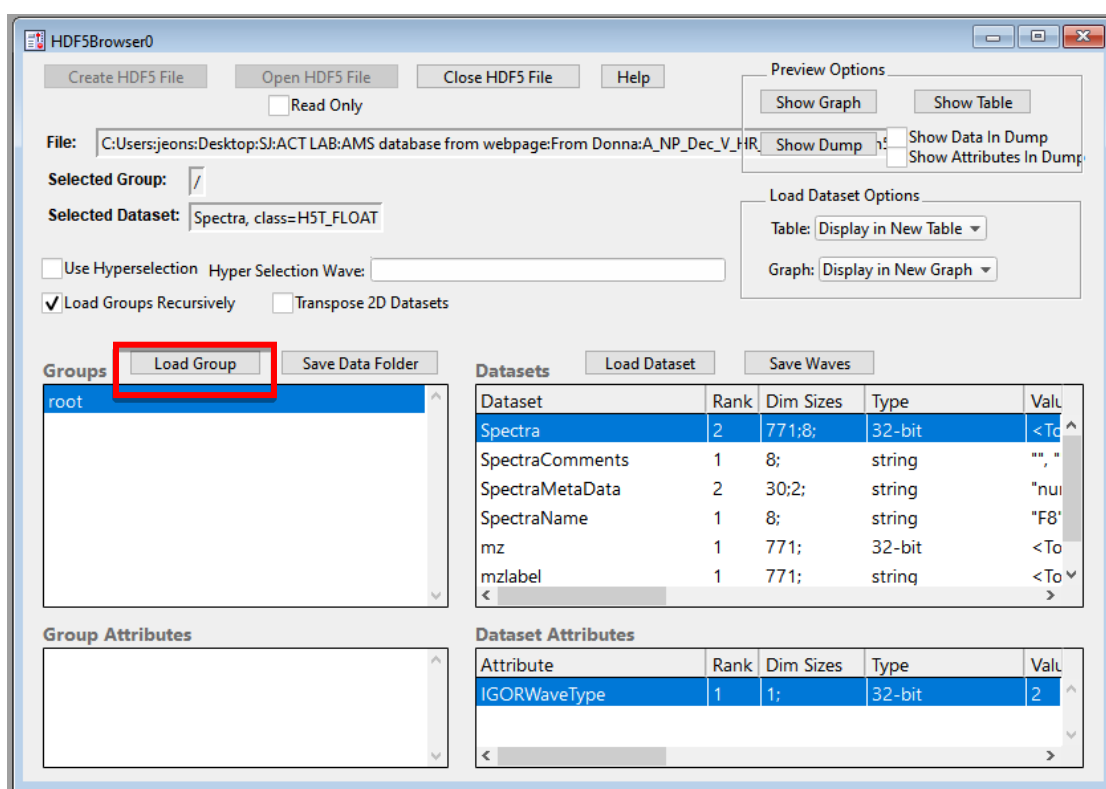
File creation
 File name: A_P_NDec_V_HR_S_test.h5
 Set file location Show
 Create data base h5 file Mail file to CUDatabaseManager@gmail.com

<Figure 1.4. Screenshot of AMS/ACSM Database Export Tool>

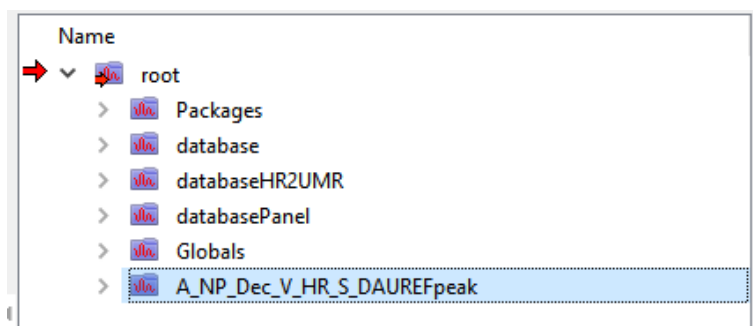
1.4 How to convert HR MS to UMR family grouped MS

For UMR data, users don't need additional pretreatment if one created the HDF file through the AMS/ACSM Database Export Tool. However, for HR data, the MS saved in the HDF file is in HR (with individual ion designations such as C₂H₃O) and has no family designation stored in the HDF file using AMS/ACSM Database Export Tool. This section describes the steps for converting the HR data to UMR family grouped MS, if needed.

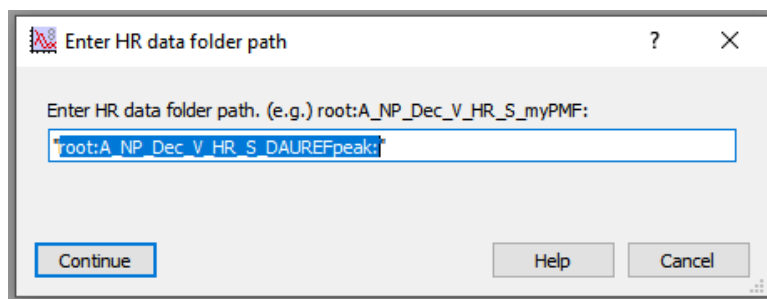
- 1) On the menu, select [Data] > [Load waves] > [New HDF5 Browser]
- 2) On the panel, select the 'Open HDF5 file' button and select the created HDF file that you created using the AMS/ACSM Data Export Tool.
- 3) On the 'Groups' section, click the 'Load Group' button.



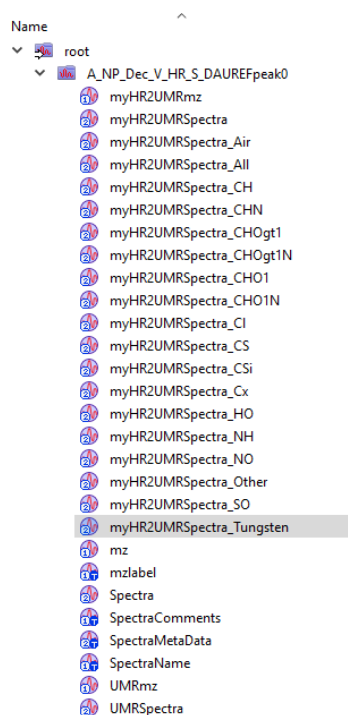
※ Caution: The uploaded folder has to be placed in the 'root:' folder like below:



- 4) Click the 'Close HDF5 File' button on the panel to close the HDF5 Browser panel (If the HDF file is UMR data users can stop here).
- 5) Within the Data Browser window, right-click on the created folder and select 'Copy Full Path'
- 6) In the main Igor menu, select [Database] > [Convert HR2UMR]
- 7) Enter the path of created folder like below:



- 8) Press the Continue button. The code will generate converted UMR spectra with HR family designation like below.



Chapter 2. Unit mass resolution (UMR) mass spectra comparison

2.1 The selection of a mass spectrum of interest

2.1.1 User supplied 1D/2D UMR MS

The user supplied 1D/2D UMR MS option is used to compare a user's mass spectrum with mass spectra in the database. To use the user supplied 1D/2D UMR MS option, 3 waves are needed: i) MS spectra wave, ii) wave of m/z value matching the MS spectra, and iii) Species wave. Detailed information about the waves is described below (Figure 2.1):

- MS wave: the 1D/2D wave of the mass spectrum from users. The wave values indicate the relative abundance of signal vs m/z in the mass spectrum. If the wave is in 2D, each column should correspond to a mass spectrum. If one created an MS through AMS/ACSM Database Export Tool (1.3), this wave corresponds to the wave named 'Spectra'.
- m/z value wave: the 1D wave of m/z values matching the MS spectra wave. The m/z values should be integers. The row number of this wave should match that of the MS spectra wave. If one created a UMR MS through AMS/ACSM Database Export Tool (1.3), this wave corresponds to the wave named 'mz'. If one converted HR to UMR MS through the internal function in the AMS database, this wave corresponds to the wave named 'myHR2UMRmz'.
- Species wave: the 1D text wave describing each mass spectrum. The row number of this wave should match the column number of the MS spectra wave. This wave corresponds to the wave named 'SpectraName' if one created the MS through AMS/ACSM Database Export Tool.

Row	mz	Spectra[[0]]	Spectra[[1]]	Spectra[[2]]	SpectraName	
		0	1	2		
0	12	0.0126627	0.00723526	0.00990823	BBOA	
1	13	0.00458538	0.0045581	0.00343991	LVOA	
2	15	0.0235422	0.0309041	0.0139332	COA	
3	16	0.00157571	0.00010379	0.00143828		
4	17	0.00984817	0.000648689	0.00898923		
5	18	0.0393927	0.00259476	0.0359569		
6	19	4.69687e-05	2.91742e-06	4.28359e-05		
7	20	7.88741e-05	4.8992e-06	7.19341e-05		
8	24	0.00145562	0.00150001	0.00146649		
9	25	0.00505095	0.00508161	0.00405383		
10	26	0.0242691	0.0249169	0.0195293		
11	27	0.0327079	0.041742	0.0405658		
12	28	0.175079	0.0115323	0.159809		
13	29	0.0622086	0.0564865	0.0305588		

<Figure 2.1. An example of m/z value wave ('mz'), MS wave ('Spectra'), and Species wave ('SpectraName') wave in Igor>

To use the User supplied UMR MS option, users can follow the below:

- 1) Check 'User supplied 1D UMR MS' to activate popup menus, if users have more than two mass spectra, check 'User supplied 2D UMR MS'
- 2) Select a 'Folder' including the waves mentioned above. This folder has to be placed in the root folder.
- 3) Select a MS Spectra wave (e.g. Spectra) from the 'MS' popup menu.
- 4) Select an m/z value wave (e.g. mz) from the 'mz' popup menu. The list automatically shows only waves having the same number of rows as the MS wave.
- 5) Select a species wave (e.g. SpectraName) from the 'Species wave' popup menu if you checked 'User supplied 2D UMR MS'.
 - 5-2) If you checked 'User supplied 2D UMR MS', select a mass spectrum that one wants to compare with the database in the 'MS wave' popup menu.
- 6) Click the 'Compare' button to calculate the cosine similarity of all MS in the database.

2.1.2 MS Existing in the data base

The existing MS option is used to compare a mass spectrum in the database with other reference mass spectra in the database. To use the Existing MS option, users can follow the steps below.

- 1) Check 'MS Existing in the data base' to activate the 'Sample Mass Spectrum' list
- 2) Click the mass spectrum to be compared in the list box.
- 3) As soon as the user clicks the mass spectrum in the list box, cosine similarity is automatically calculated.

2.2 A comparison result of the sample with reference mass spectrum and options

2.2.1 UMR Comparison Result Table

The result table consists of two columns: 'Reference MS' and 'Score'. The 'Reference MS' column indicates the reference mass spectrum name in the list. The 'Score' column shows the calculated cosine similarity between the mass spectrum of interest and the reference mass spectrum in the database. The result is sorted by the score from largest to smallest.

302_Beijing urban area_2010_COA			
Point	Reference MS	Cosine score	
0	302_Beijing urban area_2010_COA	0.9640	^
1	270_Xian and Beijing_HOA	0.9497	
2	297_Beijing urban area_2011_COA	0.9491	
3	247_MILAGRO Campaign_2006_HOA	0.9484	
4	268_Paris Summer_2009_HOA	0.9477	
5	303_Beijing urban area_2010_HOA	0.9403	
6	284_CalNex campaign_2010_HOA	0.9361	
7	313_Nonanal	0.9345	
8	261_SPC Research Station Po Valley_2008_HOA	0.9327	
9	305_Beijing urban area_2010_CCOA	0.9305	
10	1_Diesel Bus Exhaust	0.9290	
11	99_Paris_Winter_2010_HOA	0.9268	
12	116_ClusterMS_HOA	0.9242	
13	255_CARES campaign_2010_HOA	0.9233	
14	276_DIAPASON2014 campaign_HOA	0.9221	
15	111_EUCAARI_2008_HOA_median	0.9216	
16	77_HOA_avg	0.9195	v

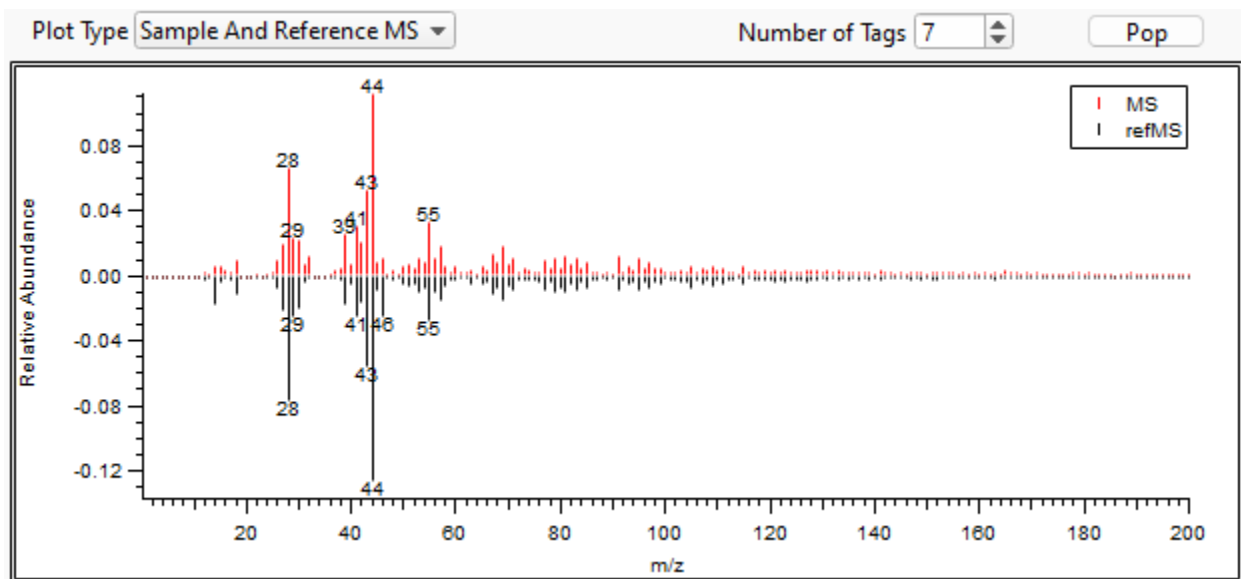
<Figure 2.2. Screenshot of result table on the 'UMR Data Comparison' tab>

2.2.2 UMR Comparison Result Mass Spectrum Plot

The UMR data comparison tab provides 5 different types of plots: i) Sample MS, ii) Reference MS, iii) Sample and Reference MS, iv) Sample MS – Reference MS, and v) Sample MS vs Reference MS. Users can select the plot type through the pop menu. A detailed description of plot types is shown below:

- Sample MS: mass spectrum of the selected mass spectrum to be compared in the selection region
- Reference MS: mass spectrum of the reference mass spectrum selected in the result table.
- Sample MS and Reference MS: mirror mass spectra of Sample MS (red) and Reference MS (black)
- Sample MS – Reference MS: subtracted mass spectrum between Sample MS and Reference MS
- Sample MS vs Reference MS: scatter plot between Sample MS (y-axis) and Reference MS (x-axis). It also provides regression information.

The ‘Number of Tags’ option above the plot window is used to label the mass value on the mass spectrum from largest to smallest for as many as entered. For example, if 5 is entered, the top 5 mass value is labeled in the plot. Users can open the plot in a new window by pressing the ‘Pop’ button.



<Figure 2.3. Screenshot of Plot window on the ‘UMR Data Comparison’ tab>

2.2.3 Mass Spectra Rescaling Options

Users can rescale the mass spectrum and its calculation by changing the mass range of the mass spectrum or increasing/decreasing m/z or intensity exponents. It rescales the mass spectrum based on equation 1 below.

$$\text{Weighted intensity} = [m/z]^m [\text{Peak intensity}]^n \text{ (Eqn. 1)}$$

As soon as the user changes the value in a variable box, the panel automatically calculates the cosine similarity with the reweighted mass spectrum and shows new results in the results table.

Mass Spectra Rescaling Options (all)

m/z Min	1	m/z Exponent	0
m/z Max	300	Int Exponent	1

<Figure 2.4. Screenshot of Scaling Options section on the 'UMR Data Comparison' tab>

2.2.4 Comparison Constraints

Users can use comparison constraints to filter the results in the database based on its metadata. It provides 6 categories to constrain the results: Database reference sample type, perturbation, deconvolution, instrument, MS resolution, and vaporizer type. The details of each category are described below.

- **Sample type:** Origin of the aerosol. It indicates if there were any instrumental or experiment set-up conditions that are not reflective of general atmospheric conditions. Sample type is classified into 'Ambient' (immediate ingestion of ambient air to the instrument) and 'Non-ambient' (including all chamber studies, any measurement whereby the spectrum is a specific source or a standard, or collected on filters as a preliminary step).
- **Perturbation:** It indicates if there was a perturbation of the sample due to some instrumental intermediary. If the sampling was carried out with an oxidation flow reactor, thermodenuder, a sample collected on filters, etc, it is classified as 'Perturbed', if not, 'Non-Perturbed'
- **Deconvoluted:** It denotes if there was an analysis step such as positive matrix factorization (PMF). If yes, 'Deconvoluted', if not, 'Non-deconvoluted'
- **Instrument:** It indicates the instrument used. It includes Air Quality ACSM, ToF-ACSM, Quad-AMS, C-ToF-AMS, HR-ToF V-mode, HR-TOF W-mode, L-ToF V-mode, L-ToF W-mode, and multiple instruments.
- **Resolution:** It shows the resolution of the sample spectrum. It includes unit mass resolution, high resolution, and combination (HR for low masses, UMR for high). Furthermore, if the spectrum is generated from HR to UMR, then the sample is assigned as HR that has been summed to UMR.
- **Vaporizer type:** It includes a capture vaporizer and a standard vaporizer.
- **Aerosol CO and CO₂ contributions:** This indicates whether the signals of CO⁺ at m/z 28 and CO₂⁺ at m/z 44 in the mass spectrum were adjusted to be equal during the data processing. In the case of UMR data, the default signal from aerosol CO⁺ at m/z 28 is set to be identical to the contribution from aerosol CO₂⁺ signal at m/z 44 (Allan et al., 2004). Especially in the case of high biomass burning sources, users may have adjusted the organic component at m/z 28 to be different from that at CO₂ at m/z 44. In the UMR case ('mz 28/mz 44 = 1 (for UMR)' on the UMR comparison tab of the panel), this option screens for cases when these two signals are equal, which is reflective of more ambient measurements and the default fragmentation table, and for cases when these signals are not equal ('Ratio \neq 1 (for both)' on the UMR comparison tab), which is very uncommon. In the case of HR data, the default setting is to not fit the CO⁺ ion, as it is generally unseparated from the dominant N₂⁺ ion at m/z 28. The default setting in the HR fragmentation table is to set the signal of aerosol CO⁺ to aerosol CO₂⁺. In terms of HR families, this default setting means that the family CHO1 signal at m/z 28 is equal to the family CHOgt1 signal at m/z 44. In the HR case ('CO (CHO1)

/ CO₂ (CHOgt1) = 1 (for HRtoUMR)’ on the UMR comparison tab or ‘CO/CO₂ =1’ on the HR comparison tab of the panel), this option screens for cases when these two signals are equal, which is reflective of more ambient measurements and the default HR fragmentation table, and for cases when these signals are not equal (‘CO/CO₂ ≠ 1’ on the HR comparison tab of the panel), which is very uncommon.

The user may select one or more desired constraint(s) and then press the ‘Apply’ button. At least one option must be selected in each category.

Comparison Constraints Apply

1. Database Reference Sample Type(s)
☒ Ambient ☒ Non-Ambient

2. Perturbation?
☒ Perturbed ☒ Non-Perturbed

3. Deconvoluted?
☒ Deconvoluted ☒ Non-Deconvoluted

4. Instrument
☒ Air Quality ACSM ☒ ToF-ACSM ☒ Quad-AMS
☒ C-ToF-AMS ☒ HR-ToF V-Mode ☒ HR-ToF W-Mode
☒ L-ToF V-Mode ☒ L-ToF W-Mode ☒ Multiple ☒ All

5. MS Resolution(s)
☒ Unit Mass Res. Data ☒ Combination
☒ HR to UMR ☒ All

6. Vaporizer Type(s)
☒ Standard Vaporizer
☒ Capture Vaporizer

7. Aerosol CO and CO₂ contributions
☒ mz 28/mz 44 = 1 (for UMR) ☒ CO (CHO1) / CO₂ (CHOgt1) = 1 (for HRtoUMR) ☒ Ratio ≠ 1 (for both)

<Figure 2.5. Screenshot of filtering option section on the ‘UMR Data Comparison’ tab>

2.3 Citation information

Citation information indicates the metadata of each mass spectrum. Users can obtain detailed information about the mass spectrum such as comments on each constraint category above, a short description of the selected mass spectrum, electron ionization (EI) energy, vaporizer temperature, experimenter’s name, provider’s group, citation URL, citation authors, and citation figure number.

In order to see the citation information, users can press the ‘Open metadata table’ button and then a table pops up that includes the descriptions above. When a user clicks the ‘Open’ button, it directly opens the paper on a web browser if the metadata has the citation URL. The top entry corresponds to the target mass spectrum and the bottom entry corresponds to the selected reference mass spectrum on the result table. If the target mass spectrum is ‘New MS’ from the user, the top ‘Get Citation’ button and ‘Open the paper’ buttons will be deactivated.

Mass Spectra Info

7_Beech_Flame
Open metadata table Open article URL

28_Ponderosa Pine Duff
Open metadata table Open article URL

Open UMR AMS database webpage

<Figure 2.6. Screenshot of citation section on the panel>

Chapter 3. High resolution to unit mass resolution (HR) spectra comparison

3.1 The selection of a mass spectrum of interest

3.1.1 User supplied 1D/2D HR MS

The User supplied 1D/2D HR MS option for HR data is used to compare a user's mass spectrum with reference mass spectra in the database. To use the User supplied 1D/2D HR MS option for HR data, users need i) MS spectra wave, ii) wave of exact m/z value matching the MS spectra, iii) wave of m/z value label matching the m/z value wave, and iv) Species wave. Detailed information about the waves is described below (Figure 3.1):

- MS wave: the 1D/2D wave of the HR mass spectrum from users. The wave values indicate the relative abundance of signal vs m/z in the mass spectrum. If the wave is in 2D, each column should correspond to a mass spectrum. If one created MS through AMS/ACSM Database Export Tool (1.3), this wave corresponds to the wave named 'Spectra'.
- m/z value wave: the 1D wave of m/z values matching the MS spectra wave. The m/z values should be an exact mass. The row number of this wave should match the MS wave. If one created an MS through AMS/ACSM Database Export Tool (1.3), this wave corresponds to the wave named 'mz'.
- m/z value label wave: the 1D string wave of m/z values label (chemical formula of HR fragments) matching the m/z value wave. If one created MS through AMS/ACSM Database Export Tool (1.3), this wave corresponds to the wave named 'mzlabel'.
- Species wave: the 1D string wave describing each mass spectrum. The row number of this wave should match with the column number of the MS wave. This wave corresponds to the wave named 'SpectraName' if one created the MS through AMS/ACSM Database Export Tool.

Row	mz	mzlabel	Spectra[[0]]	Spectra[[1]]	Spectra[[2]]	Spectra[[3]]	Spectra[[4]]	Spectra[[5]]	Spectra[[6]]	Spectra[[7]]	SpectraName
			0	1	2	3	4	5	6	7	
0	12	C	0.000434148	0.000235043	0.0152212	0.0132596	0.0142344	0.00041096	0.00839986	0.0142231	BBOA
1	13.0078	CH	0.00320853	0.000903705	0.00493116	0.00483059	0.00362393	0.00138969	0.00371362	0.00692391	HOA
2	14.0156	CH2	0.00594328	0.0318419	5.42059e-09	0.016034	0.0175614	0.0148814	0.047795	0.00925004	F6
3	15.0235	CH3	0.0256971	0.00833613	0.0211116	0.0225184	0.0110164	0.0106991	0.0338279	0.0252794	F5
4	15.9949	O	0.000337033	0.000240519	0.00123261	0.00116126	0.00226491	0.000574431	0.000860504	0.00165571	F4
5	16.0313	CH4	1.13919e-09	4.62853e-09	6.29949e-10	0.00360487	6.61393e-10	0.0814056	2.38171e-09	2.10033e-09	F3
6	17.0027	HO	0.002126	0.00152459	0.00770649	0.00708935	0.0141698	0.00369391	0.00549871	0.0103491	F2
7	18.0106	H2O	0.00849809	0.0061039	0.0308357	0.0283579	0.0566829	0.014772	0.0219912	0.0413916	F1
8	21.0053	C2H2Oplus2	7.64824e-05	3.39631e-05	7.5584e-05	0.000118971	9.58982e-05	5.43831e-05	5.38168e-05	0.000140855	
9	21.9949	CO2plus2	3.48583e-10	2.53057e-09	0.000675729	0.000496803	0.00099768	6.62498e-05	0.000825814	0.000741053	

<Figure 3.1 Examples of MS waves ('Spectra'), m/z value wave ('mz'), m/z value label wave ('mzlabel'), Species wave ('SpectraName') for HR MS comparisons>

To use the User supplied 1D/2D HR MS option, users can follow the below:

- 1) Check 'User supplied 1D HR MS' to activate popup menus, if you have more than two mass spectra, check 'User supplied 2D HR MS'.
- 2) Select a 'Folder' including the waves mentioned above. This folder has to be placed in the root folder.
- 3) Select a MS wave (e.g. Spectra) from the 'Spectra' popup menu.
- 4) Select an m/z value wave (e.g. mz) from the 'mz' popup menu. The list automatically shows only the wave having the same number of rows as the MS wave.
- 5) Select an m/z value label wave (e.g. mzlabel) from the 'mz label' popup menu. The list automatically shows only the wave having the same number of rows as the mz wave.
- 6) Select a spectra name wave (e.g. SpectraName) from the 'Spectra Name' popup menu if you have more than two mass spectra.
 - 6-2) Select a mass spectrum that one wants to compare with the database in the 'MS wave' popup menu.
- 7) Click the 'Compare' button to calculate the cosine similarity of all MS in the database.

3.1.2 MS existing in the data base

Please refer to 2.1.2.

3.2 A comparison result of the sample with reference mass spectrum and options

3.2.1 Result Table

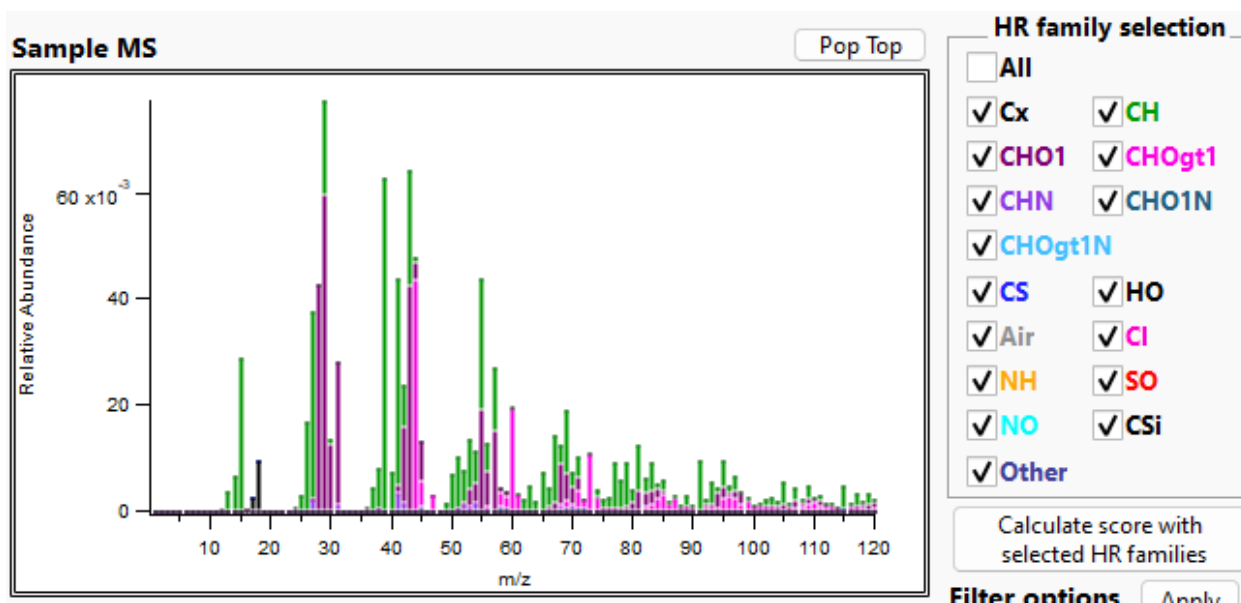
The result table for HR consists of three columns: 'Reference MS', 'Score', and 'Score with HR family'. For the 'Reference MS' and the 'Score' column, please see 2.2.1. The 'Score with HR family' column represents the cosine similarity calculated only by the selected HR ion families.

306_Beijing urban area_2010_COA			
Point	Reference MS	UMR Score	Score with HR fa
0	306_Beijing urban area_2010_COA	0.96352	0.96356
1	274_Xian and Beijing_HO	0.94985	0.94991
2	301_Beijing urban area_2010_COA	0.94823	0.94828
3	247_MILAGRO Campaign_	0.94805	0.94811
4	272_Paris Summer_2009_	0.9471	0.94716
5	307_Beijing urban area_2010_COA	0.93889	0.93889
6	288_CalNex campaign_20	0.93559	0.93561
7	317_Nonanal	0.93485	0.93491
8	265_SPC Research Station	0.93177	0.93183
9	309_Beijing urban area_2010_COA	0.92949	0.92924
10	259_CARES campaign_20	0.92219	0.92225
11	280_DIAPASON2014 camp	0.92215	0.9222
12	242_SOAR-1_Campaign_2	0.92008	0.92014
13	279_DIAPASON2014 camp	0.91928	0.91934
14	276_Xian and Beijing_CC	0.91589	0.91595
15	283_POPE2014 campaign	0.91395	0.91401
16	300_Beijing urban area_2010_COA	0.91342	0.91342
17	268_Paris Summer_2009_	0.9107	0.91076

<Figure 3.2. Example of a Result Table on 'HR Data Comparison' tab. 'UMR score' indicates cosine similarity calculated with UMR MS and 'Score with HR family' indicates cosine similarity calculated with summed and normalized MS of selected HR families >

3.2.2 Mass spectrum plot

The plot shows the stacked mass spectrum of the selected HR ion families such as Cx, CH, CHO1, CHOgt1, etc. Detailed information on HR ion families is described in 3.2.3. The top plot represents the target mass spectrum to be compared and the bottom does the selected reference mass spectrum in the database.



<Figure 3.3. Mass spectrum plot with the selected HR families on the ‘HR Data Comparison’ tab>

3.2.3 HR family selection

The HR ion families used here are the same names generated by PIKA which is the most commonly used AMS data analysis tool for HR. The panel also followed the same color legend from PIKA for consistency. The chemical formula for each HR ion family is described below. The ‘Other’ family indicates HR ions that are not a member of any other family.

Table 3.1. The chemical formula of HR families

HR family name	Chemical formula
Cx	C_x
CH	C_xH_y
CHO1	$C_xH_yO_z$ ($z=1$)
CHOgt1	$C_xH_yO_z$ ($z>1$)
CHN	$C_xH_yN_w$
CHO1N	$C_xH_yO_zN_w$ ($z=1$)
CHOgt1N	$C_xH_yO_zN_w$ ($z>1$)
CS	C_xS_y
HO	H_xO_y
Air*	N_x, O_x, Ar
Cl	H_xCl_y
NH	N_xH_y
SO	S_xO_y
NO	N_xO_y
Tungsten*	W_xO_y ($y \geq 0$)
Other	-
CSi	C_xSi_y

*On HR data comparison tab, the abundances of air and tungsten is reset to zero.

3.2.4 Mass Spectra Rescaling Options

Please see 2.2.3.

3.3.5 Comparison Constraints

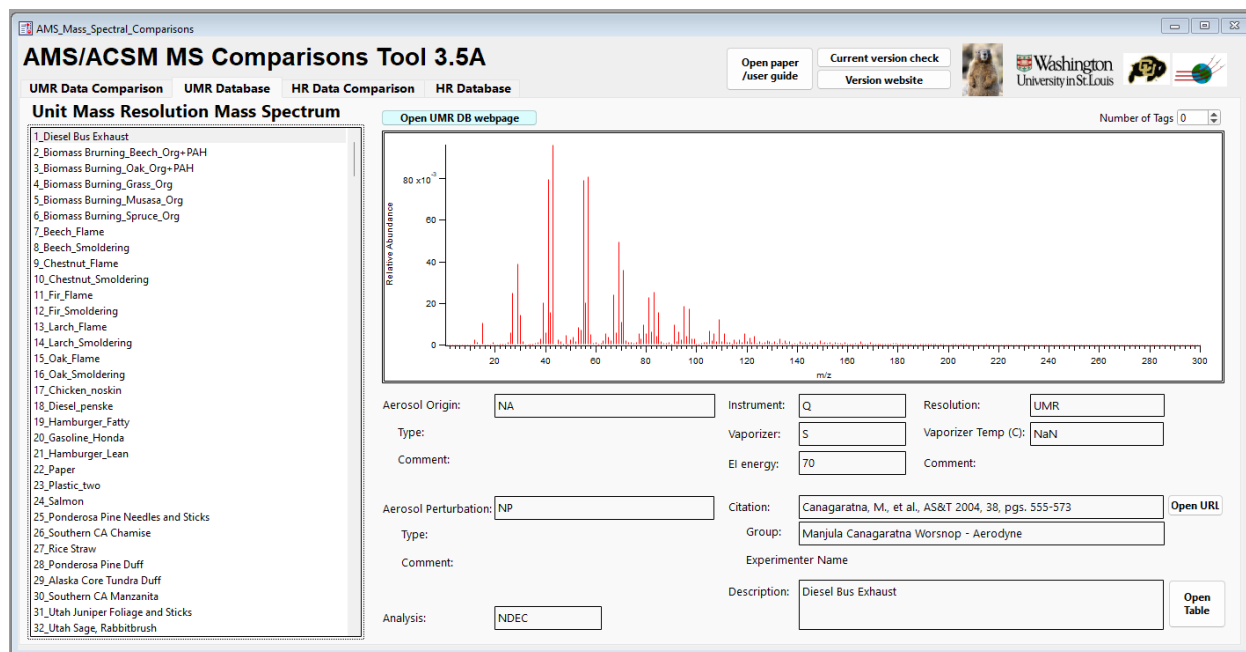
For an HR spectra comparison, the panel provides the same categories as the UMR data comparison tab except for the ‘Instrument’ and ‘MS resolution’ options. In the case of HR spectra, several non-applicable options are removed. Please refer to 2.2.4 for a detailed description.

3.4 Citation information

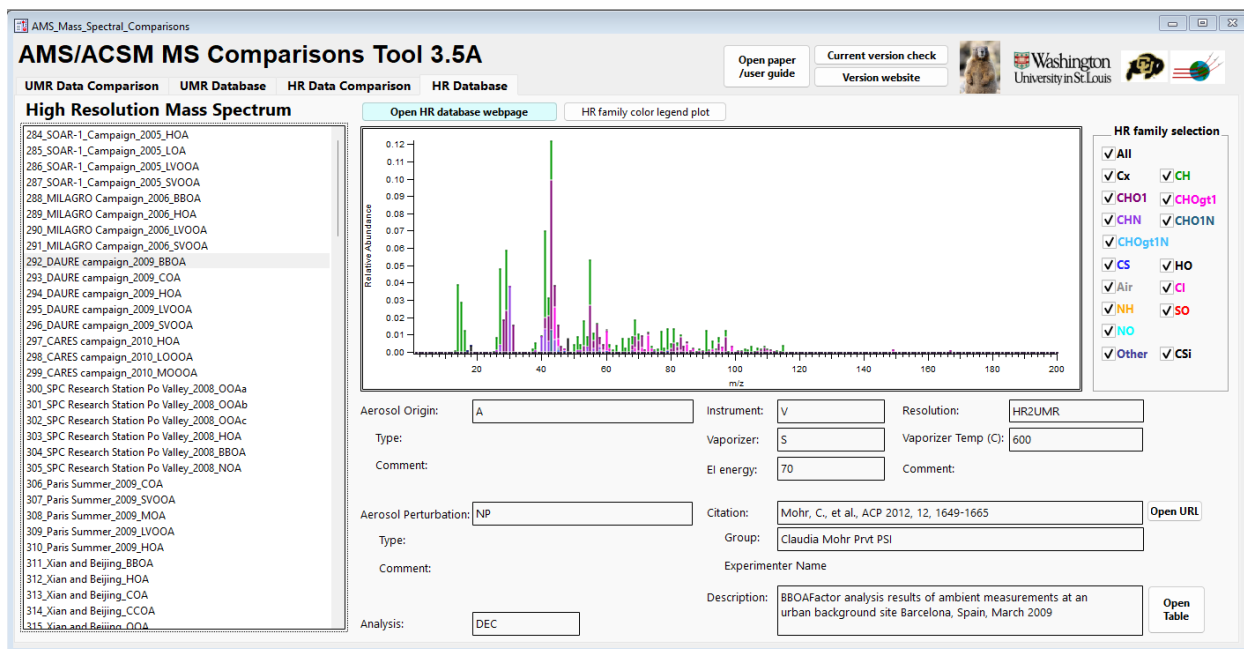
Please see 2.3.

Chapter 4. Database library

On the UMR Database and HR Database tabs, the panel plots the reference mass spectrum in the database and provides its metadata for users to look through easily. The list box presents the name of the reference mass spectrum in the database. When the user clicks one of them, the corresponding mass spectrum or stacked mass spectrum (HR) and metadata are displayed. The example displays for a UMR and an HR reference spectrum are below.



<Figure 4.1 The UMR Database tab>

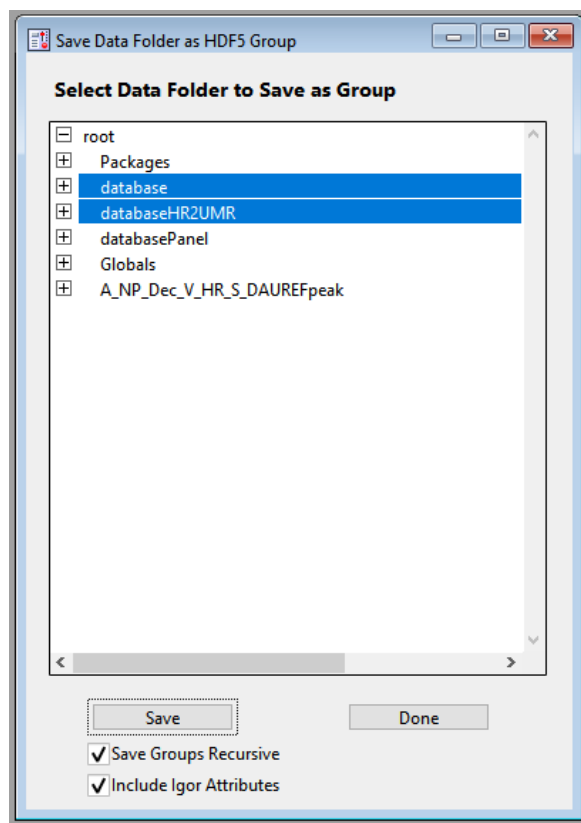


<Figure 4.2 The HR Database tab>

Chapter 5. Export database

To export a database from Igor Pro to an h5 file, users have to save all the waves and strings in the ‘database’ and ‘databaseHR2UMR’ folders. The following steps describe how to export a database to an h5 file:

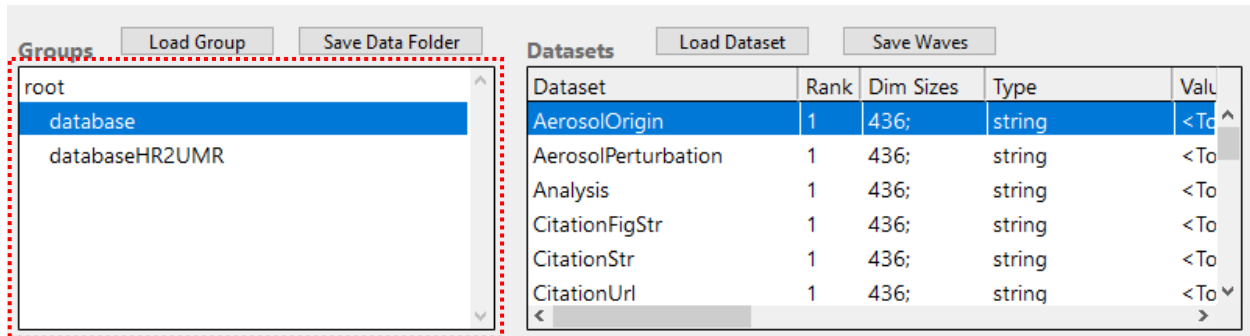
- 1) On the menu, select [Data] > [Load Waves] > [New HDF5 Browser].
- 2) On the HDF5 browser panel, click the ‘Create HDF5 File’ button.
- 3) Enter the name of the HDF5 file.h5 to be created. Users must include ‘.h5’ at the end of the file name.
- 4) On the panel, click the ‘Save Data Folder’ button.
- 5) On the ‘Save Data Folder as HDF5 Group’ panel, select both the ‘database’ and ‘databaseHR2UMR’ folders.



<Figure 5.1 The Igor window prompting the user to select data folders >

- 6) Check the ‘Save Groups Recursive’ and ‘Include Igor Attributes’ and click the ‘Save’ button.

- 7) Confirm that both folders are loaded in the ‘Groups’ window (on the left below) and check that everything is there on the right by clicking ‘database’ and then ‘databaseHR2UMR’.



<Figure 5.2 A section of the Igor window in the HDF5 Browser panel>

- 8) Click the ‘done’ button on the ‘Save Data Folder as HDF5 Group’.
- 9) Click the ‘Close HDF5 File’ button on the panel to complete the creation of the database library file.