

# Getting Started with MASH Native

Version 1.1

## Overview

MASH Native, developed by the Ying Ge research group, is a comprehensive, unified, and user-friendly, freely available software environment for the analysis of native and denatured top-down proteomics data. MASH Native builds on previous work with MASH Suite, MASH Suite Pro, and MASH Explorer. The tutorials below provide a brief introduction to the basic functions of MASH Native in both **Discovery Mode** (for identification of an unknown protein) and **Targeted Mode** (for characterization of a known protein) for native and denatured top-down proteomics.

- **Tutorial 1:** Analysis of mass spectrometry data collected under native conditions for identification of an unknown protein.
- **Tutorial 2:** Analysis of mass spectrometry data collected under denaturing conditions for identification of an unknown protein.
- **Tutorial 3:** Analysis of mass spectrometry data collected under native conditions for the characterization of hemoglobin (*Bos Taurus*) using a Targeted Mode workflow
- **Tutorial 4:** Analysis of mass spectrometry data collected under denaturing conditions for the characterization of NADH dehydrogenase [ubiquinone] iron-sulfur protein 5-like (*Sus scrofa*) using a Targeted Mode workflow
- **Tutorial 5:** Post-Translational Modification Analysis using UniDec within MASH Native

Video tutorials are also available on [YouTube™](#). For a detailed description of the functions of the MASH Native software, please view the [User Manual](#).

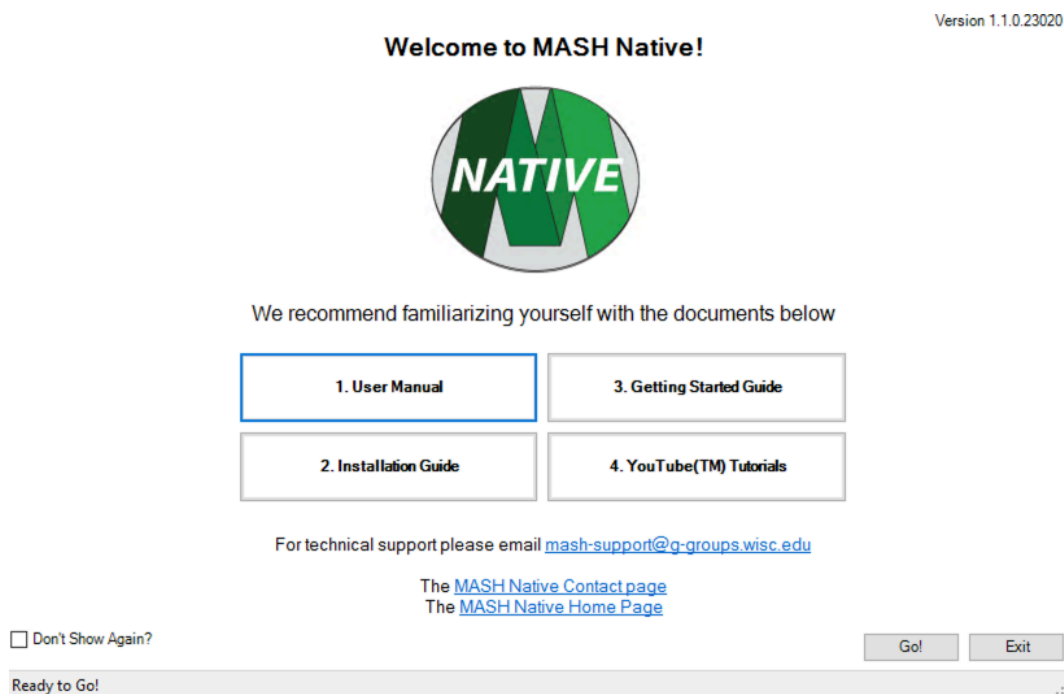
## Privacy/Usage Reporting

The MASH Native application may report data about your usage of the software to the MASH team. Each time you start a new version of the software, you will be asked to review your

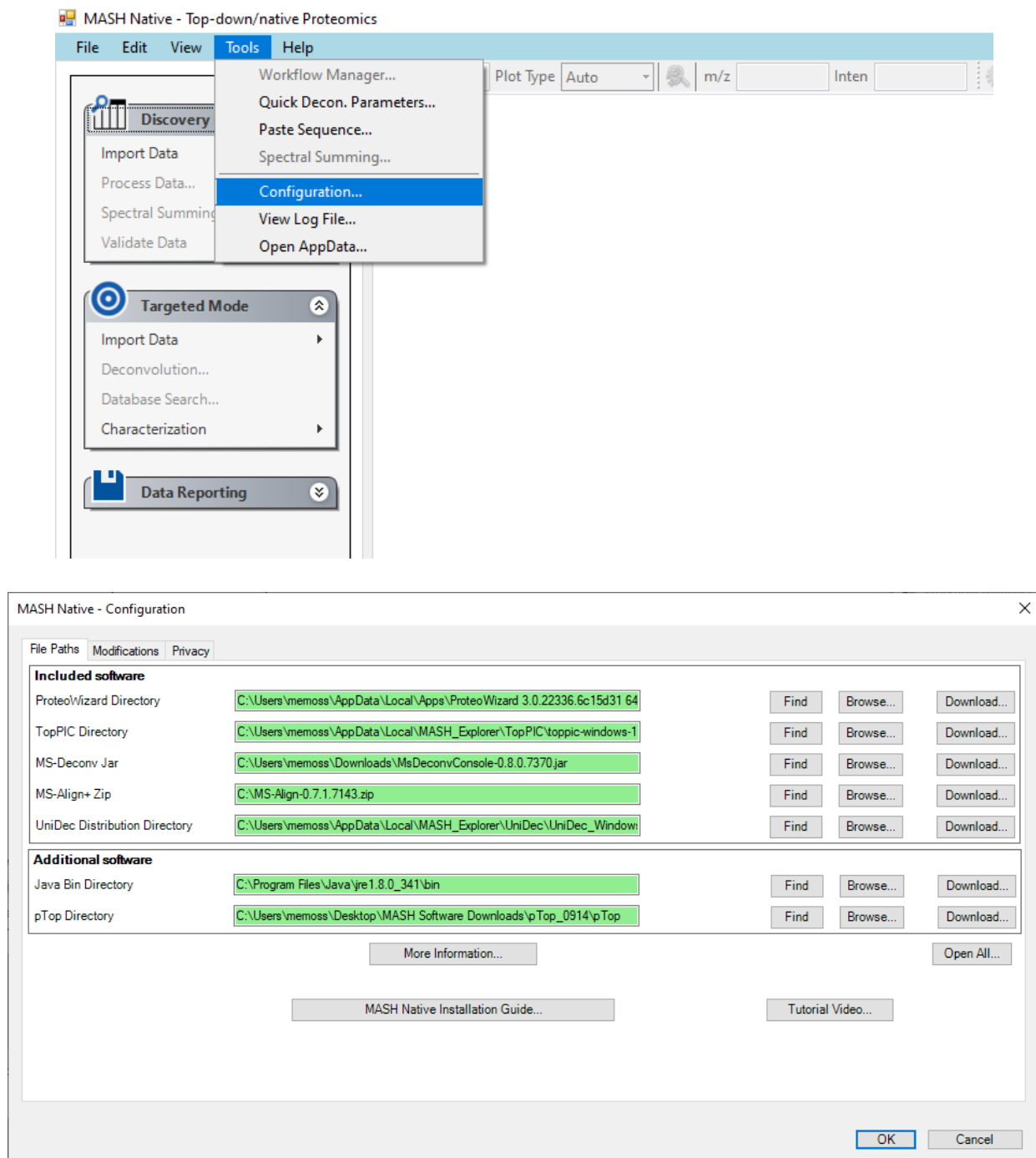
privacy/usage reporting settings. The GUI for this should be quite intuitive, but if you want more details, please see section 3 of the *User Manual*. Please see the [MASH Privacy Notice](#) for information about what data may be reported and what is done with that data.

## Installation and Configuration

To begin, please install MASH Native by referring to the [Installation Guide](#), also available in the software download folder. For this tutorial guide, all of the algorithms are included in the MASH Native application. MASH Native will open with a Welcome screen (**Figure 1**). At this time, UniDec and TopPIC Suite will be installed on the user's computer. MASH Native will prompt the user to install ProteoWizard, highly recommended for full functionality of MASH Native. MS-Deconv and MSAlign+ are included in the MASH Native App, but users will need to install and configure Java to use both software. Make sure that all associated software needed (for example, deconvolution and search tools) is installed and properly configured in the Configuration dialog (accessed via the Tools menu) (**Figure 2**).



**Figure 1. MASH Native's Welcome screen.**



**Figure 2. MASH Native Configuration dialog.** Users can access the Configuration dialog by the tools menu (top) where they are able to see the installation status of each software (bottom).

The download package of MASH Native includes TopPIC Suite, MS-Deconv., MS-Align+, and UniDec; the ‘Additional software’ deconvolution tools shown above are not required for the tutorial, but you may want to install them if you deem necessary for processing of your data. Please note that these tutorial uses TopPIC Suite version 1.5.4, MS-Deconv 0.8.0.7370, MS-ALign+ 0.7.1.7143, and UniDec 5.1.1.

## Interface Layout

Shown below is the main interface window for MASH Native (**Figure 3**). The components of the MASH Native main window are described below:

**Workflow and Parameters:** Users can follow the steps for LC-MS/MS and MS/MS workflow for data processing.

**Results View:** The mass spectra and chromatogram are shown in this panel

**Mass List:** The fragment ion information is presented in this panel.

**Experiment List:** This panel shows a list of the experiments that are currently loaded into MASH Native, and allows you to switch which experiment is active by double-clicking on the desired experiment. (The active experiment is highlighted in blue.)

**Log Book:** This panel displays processing and error messages. Note that the Experiment List and Log Book are tabs in the same panel of the MASH Native App window.

**Status Bar:** The progress of data processing is shown in this panel.

**Sequence Table:** This panel presents the protein sequence and top-down fragment ion information.

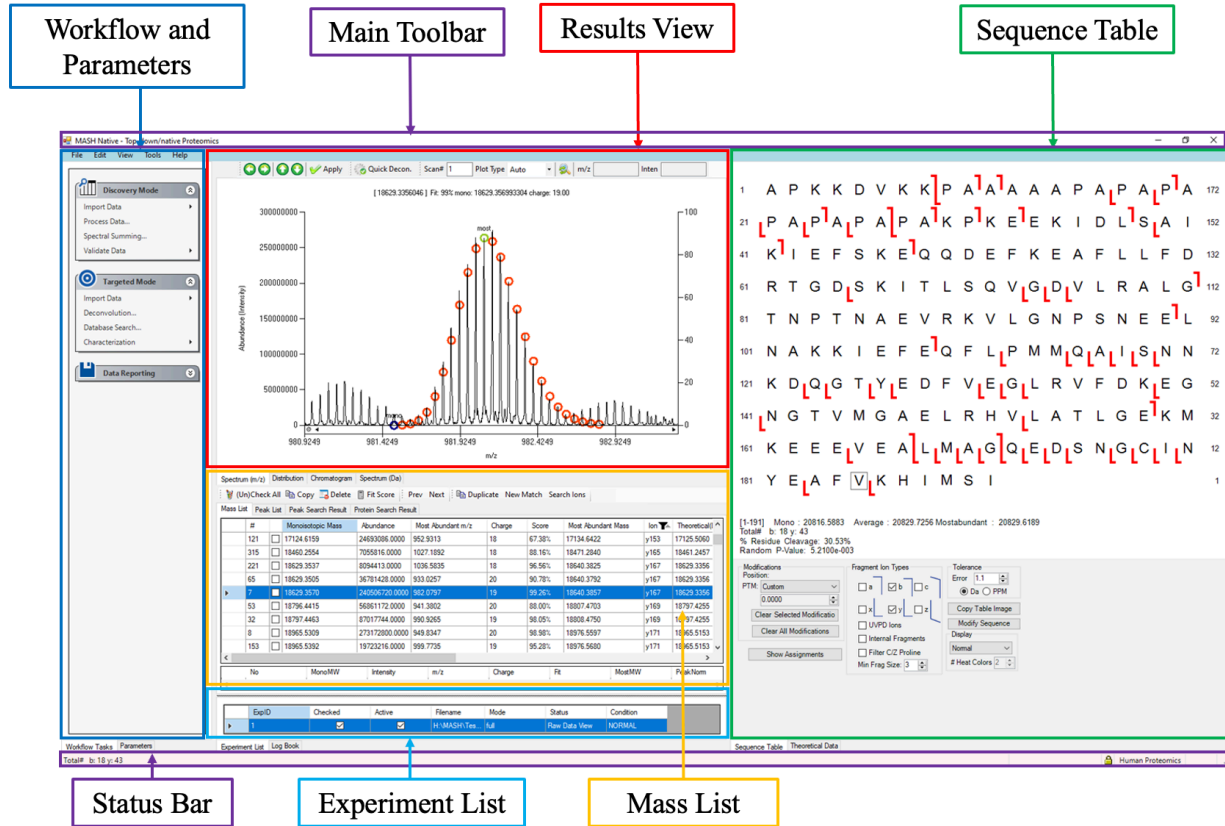


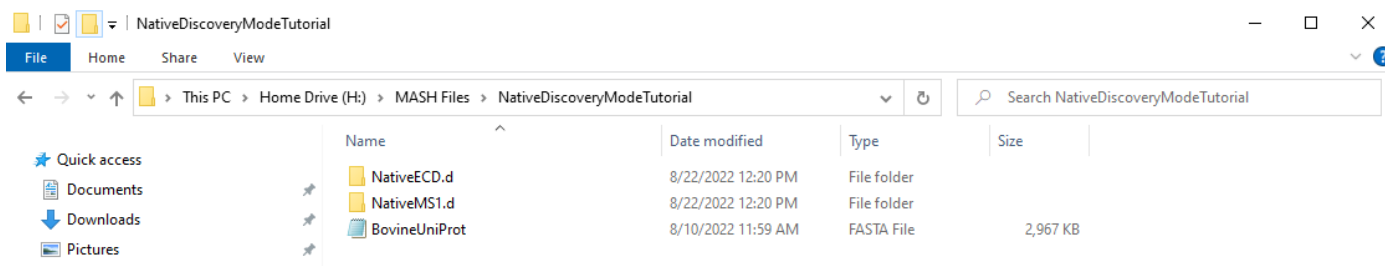
Figure 3. Main MASH Native User Interface.

## Discovery Mode Tutorials

The Discovery Mode workflow is intended to be used in the identification of an unknown protein(s). The following tutorials show MASH Native's capability for analysis of both native and denatured mass spectrometry data and includes spectral deconvolution and database search. Please follow the instructions in User Manual *Section 4 Discovery Mode: Protein Identification and Characterization from LC-MS/MS Native Mass Spectrometry Data*, if you require further assistance processing the data.

**Tutorial 1: Analysis of mass spectrometry data collected under native conditions for identification of an unknown protein.**

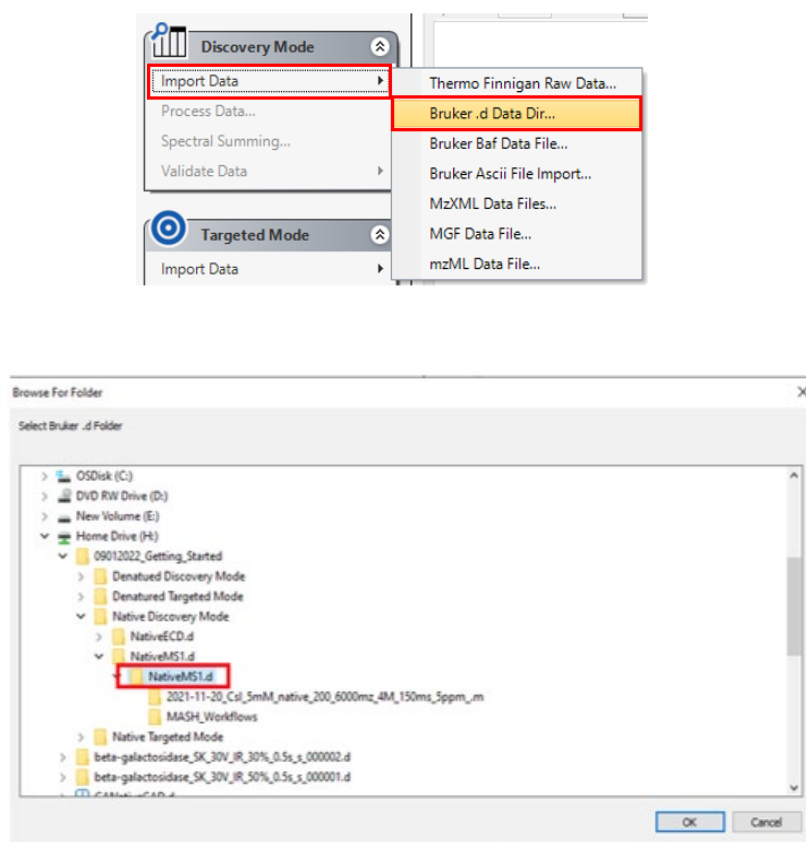
To begin the tutorial, download the [Native Discovery Mode Tutorial](#) folder which contains the Bruker NativeMS1 and NativeECD data files as well as the BovineUniProt.fasta file and save them to a new local folder on your computer as seen below (**Figure 4**). Proceed with the steps below for analysis of MS and MS/MS mass spectrometry data to identify an unknown protein from a bovine extract.



**Figure 4. Downloaded file folder containing needed files for Tutorial 1.**

### *Step 1: Import MS1 native data*

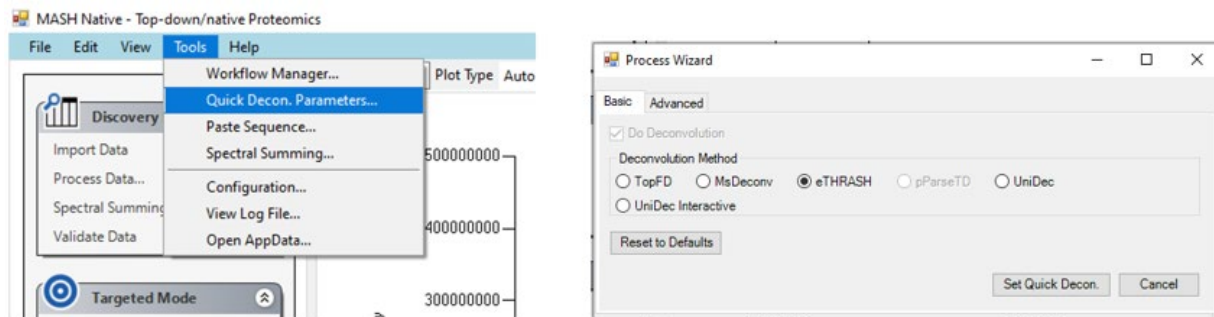
Use the Import Data function in the Discovery Mode panel to import the sample Bruker .d 'Native MS1' folder (**Figure 5**).



**Figure 5. Importing the MS1 data files for Tutorial 1.** Use the Import Data function in the Discovery Mode node to indicate the Bruker data .d Dir file type (top) and then select the .d data file from your local directory (bottom).

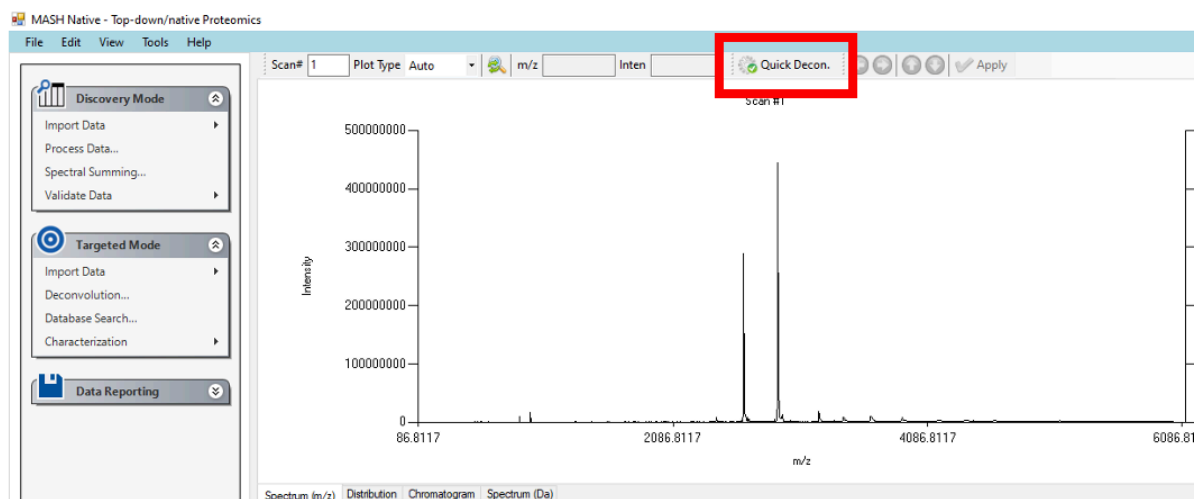
### *Step 2: Spectral deconvolution of the MS1 spectra and database search*

For this workflow, we will be using Quick Deconvolution with the eTHRASH algorithm. To ensure the Quick Decon. button will use the desired algorithm we will need to make sure to This is configured (**Figure 6**). Go to Tools and select “Quick Decon. Parameters.” This will open the Process Wizard dialog where Users can select their deconvolution method. For this tutorial, select eTHRASH and click “Set Quick Decon.”



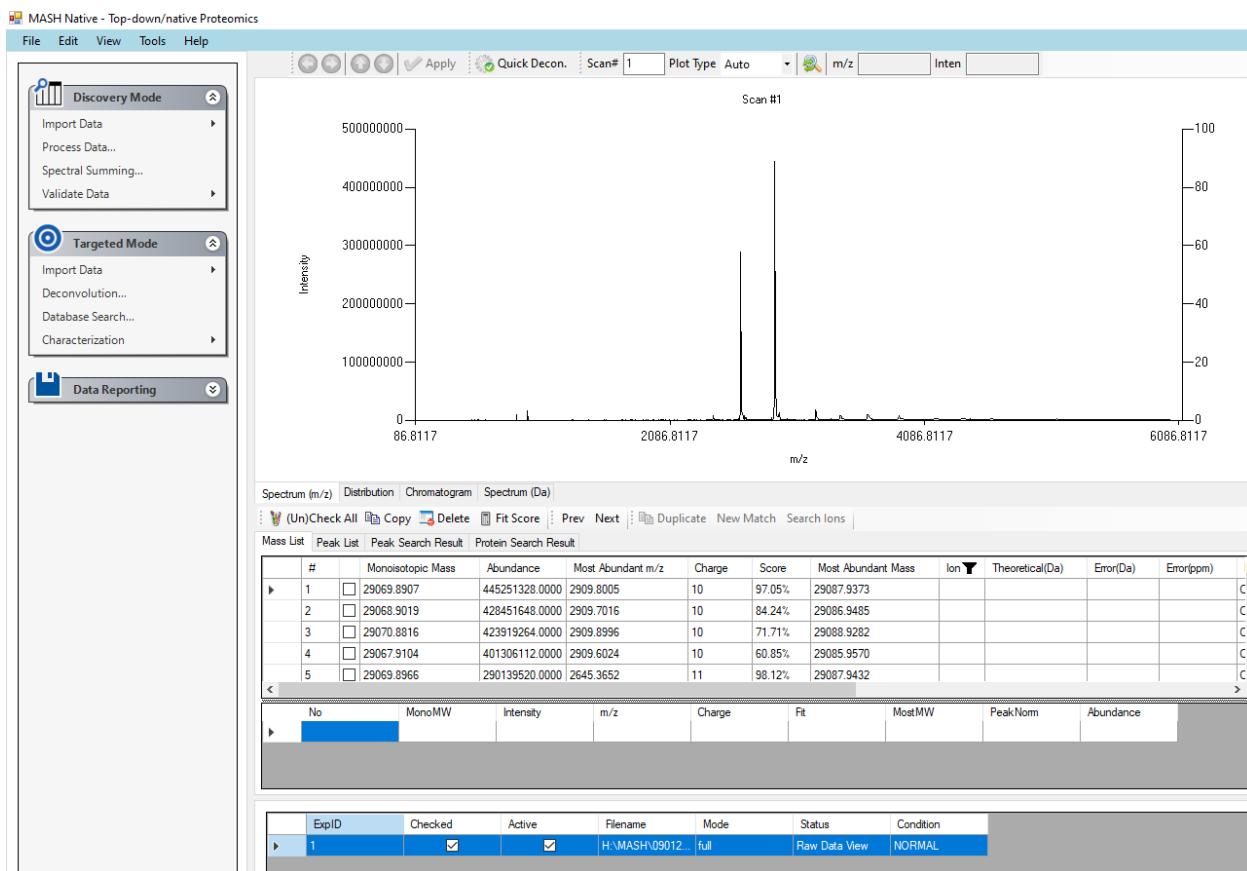
**Figure 6. Configuring the Quick Decon function in MASH Native.** Users can access this feature from the tools menu (left) and enter the desired parameters for deconvolution (right).

Once the configuration has been set, select the QuickDecon. (**Figure 7**) button to deconvolute the native MS1 spectrum. This will open the Workflow Manager window and the Terminal window. The Terminal window is where the deconvolution and search are running. You can monitor the Workflow Manager to check on the status of your run, but note that closing this window will stop your workflow! When your deconvolution is completed, you will see the workflow in the Completed Workflows tab of the Workflow Manager and the line “Workflow Finished” in the Log Book. Once the deconvolution is finished, the mass list will populate (**Figure 8**).



**Figure 7. Location of the Quick Decon. function in MASH Native.**





**Figure 8. Populated Mass List after deconvolution of native MS1 spectra in Tutorial 1.**

**CHECKPOINT:** Make sure your results match the photo above

Users can manually review the deconvolution and isotopic fit in the Results View panel, if needed. Use the arrows in the *Results View* to alter the theoretical fit and click the “Apply” button to save your adjustments (**Figure 9**). Using the information from the deconvolution results, users can then assign the charge states (**Figure 10**).

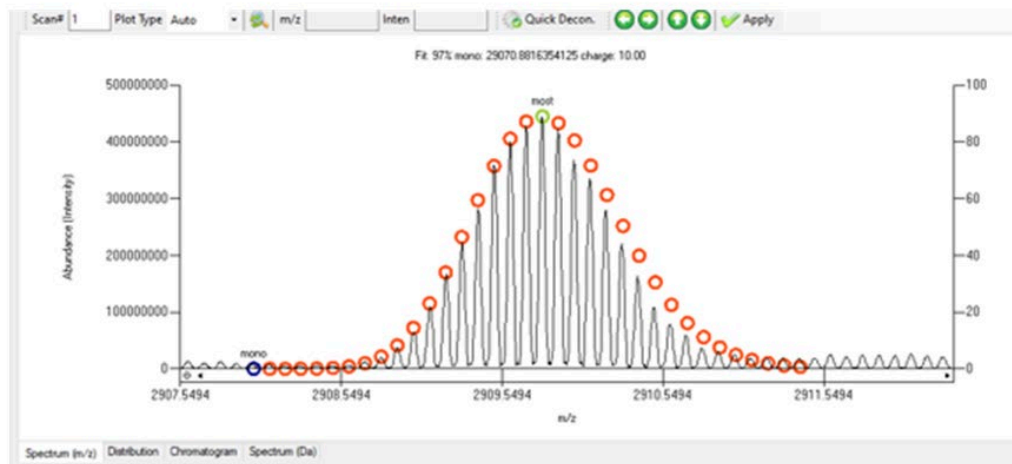
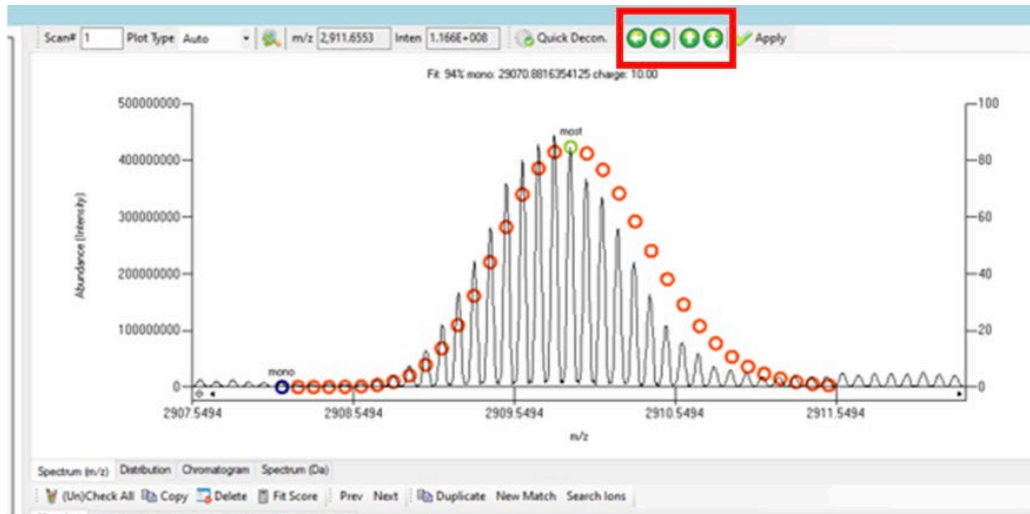
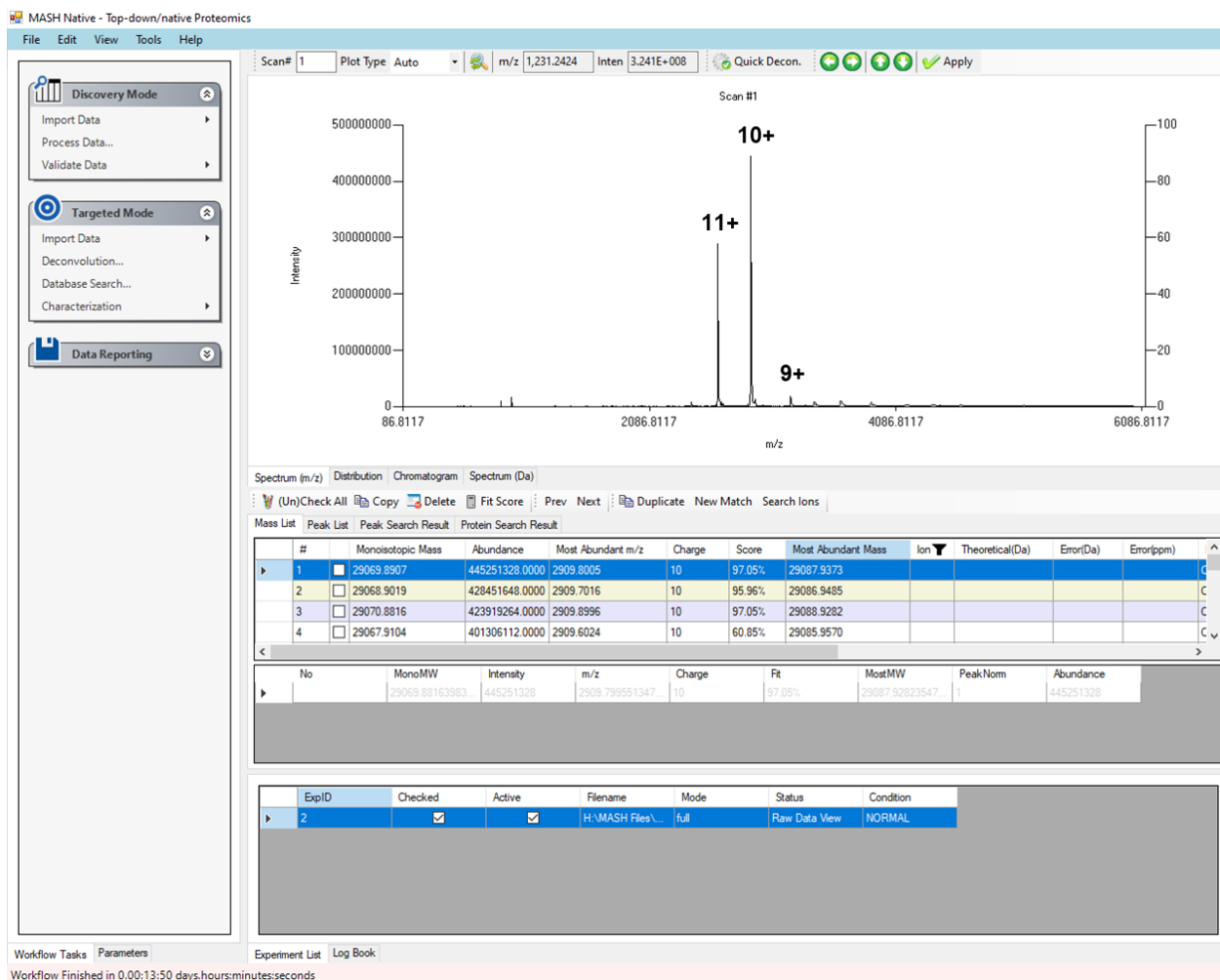


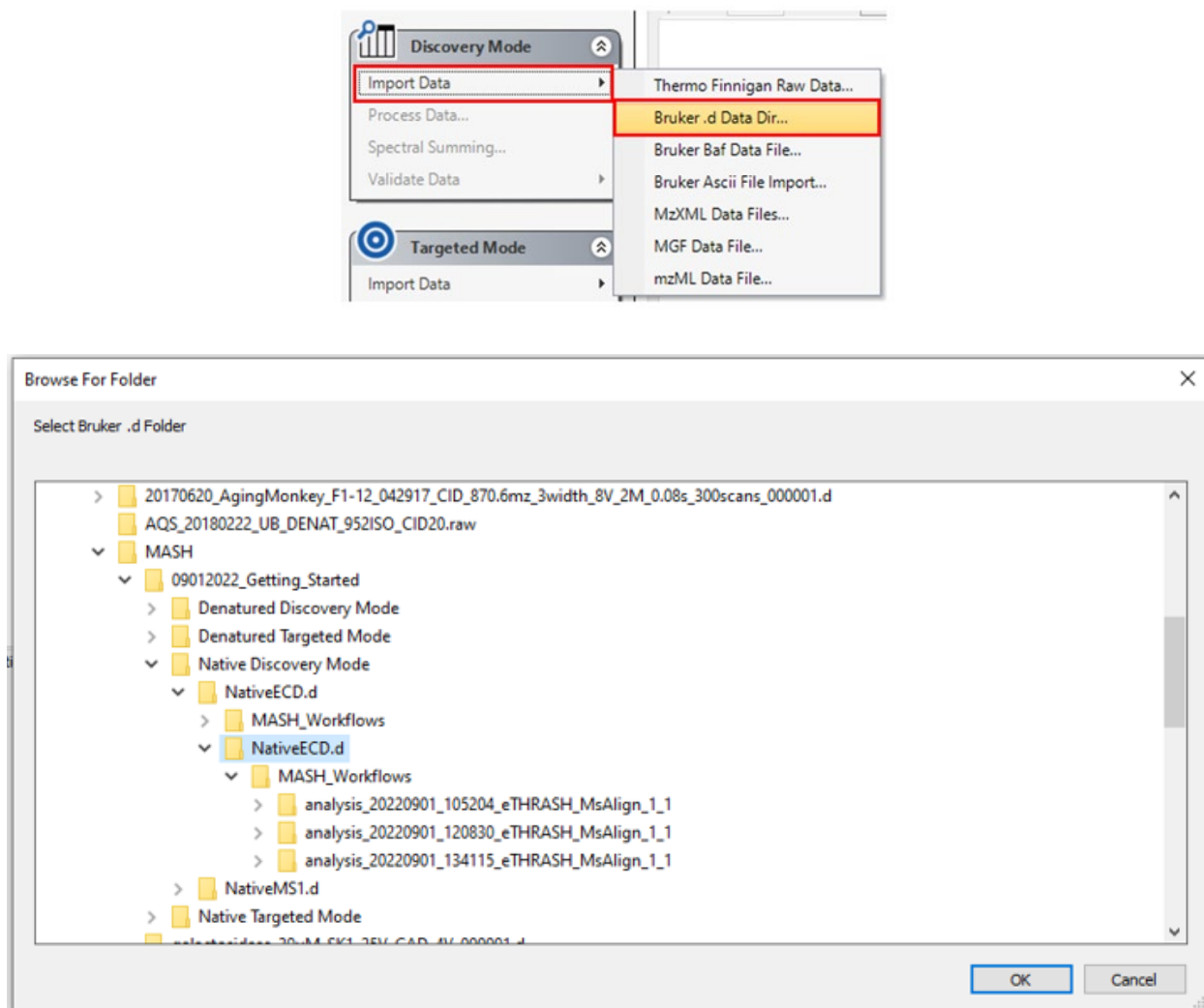
Figure 9. Adjusting of the isotopic fit in MASH Native.



**Figure 10. Assigned charge states for the native MS1 spectra in Tutorial 1.**

### Step 3: Import MS/MS data

Use the Import Data function in the Discovery Mode to import the sample Bruker .d 'Native ECD' folder (**Figure 11**) in the same manner as the NativeMS1 folder.

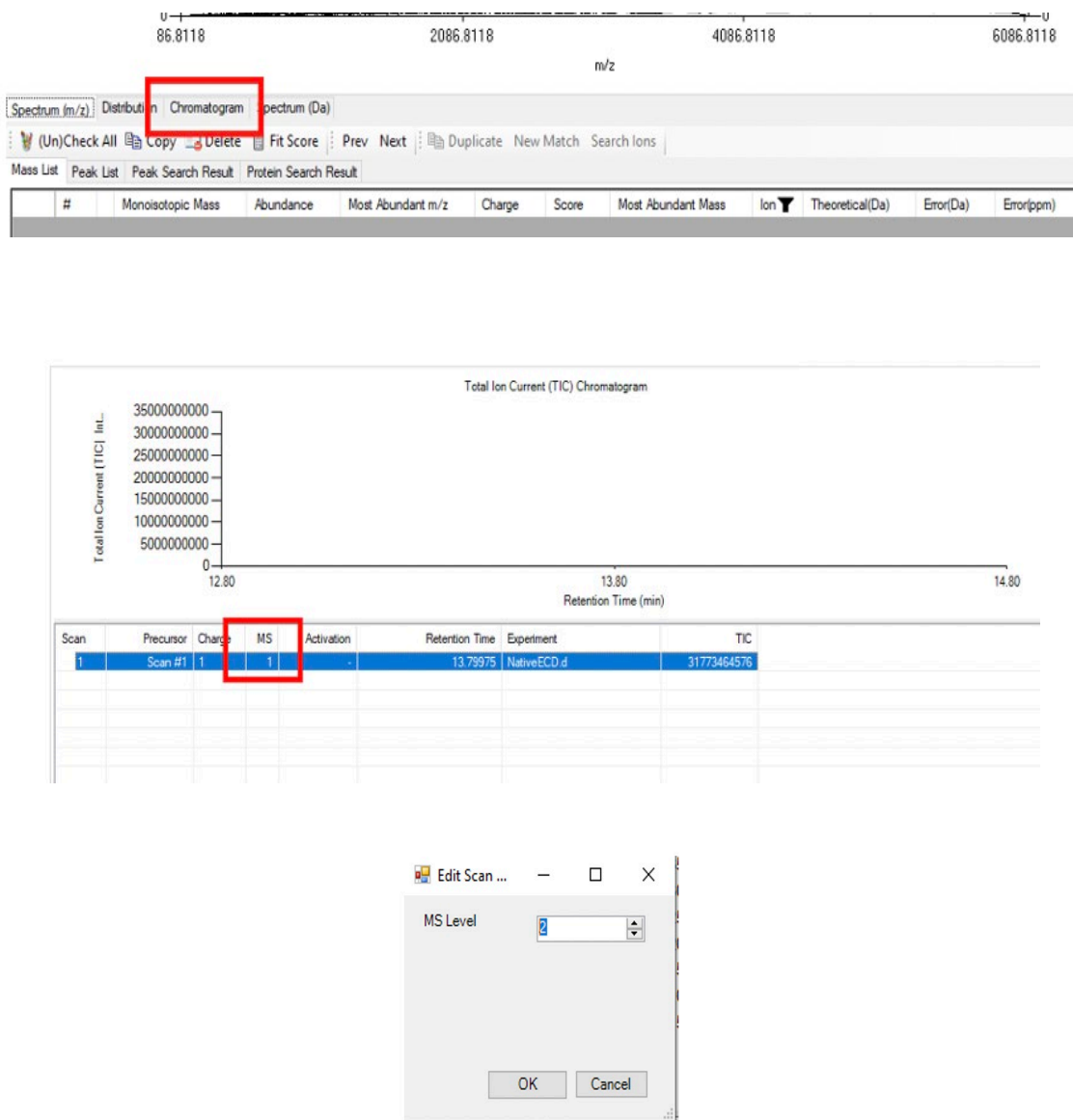


**Figure 11. Importing the MS/MS data files for Tutorial 1.** Users can use the Import Data feature in the Discovery Mode panel (top) and select the NativeECD file for analysis (bottom).

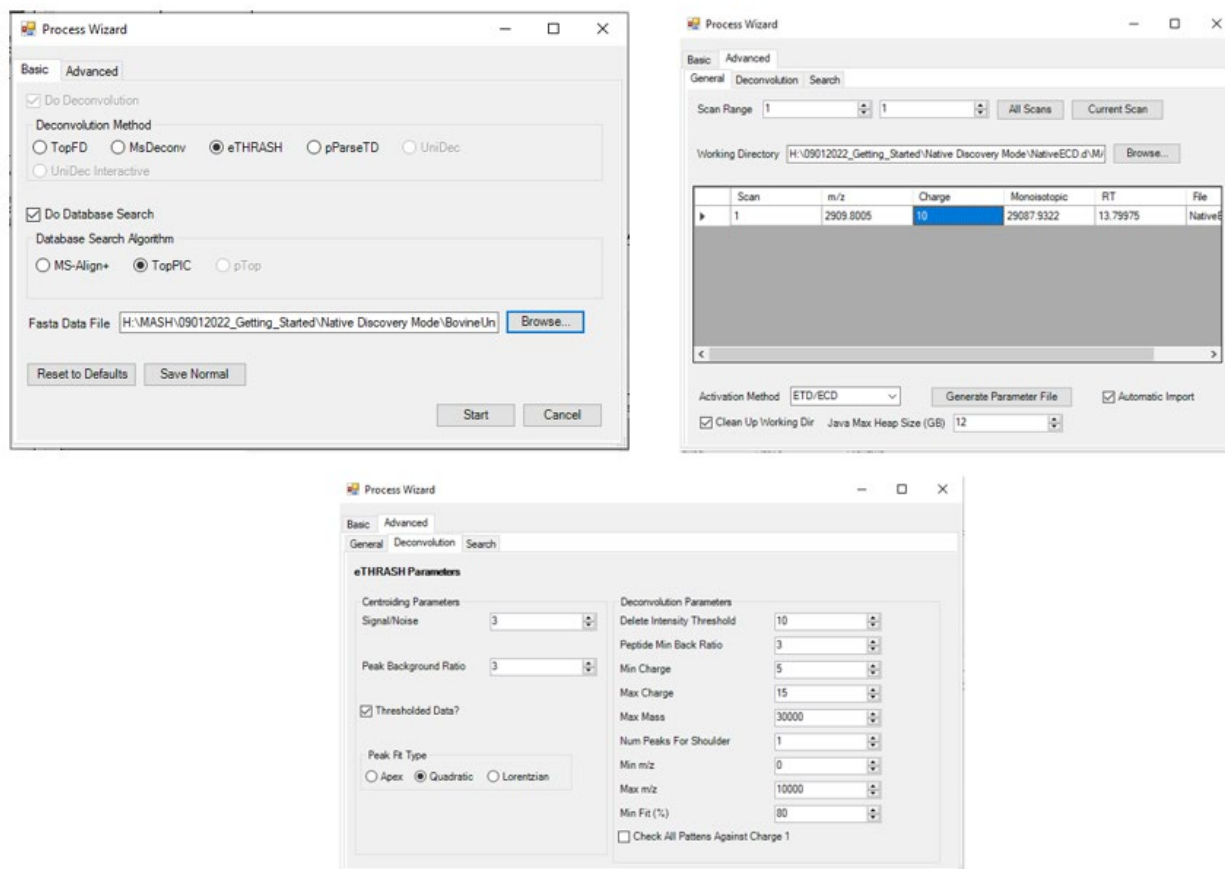
#### *Step 4: Spectral deconvolution of the MS/MS spectra and database search*

To process the fragmentation data, users must label the imported data as an MS/MS (MS2) file (**Figure 12**). To do so, click the “Chromatogram” button, then double click where it says “MS” and in the popup window, change the MS level to 2 and press “OK.” Next, click the Process Data menu item in the Discovery Mode panel to open the Process Wizard dialog. The Process Wizard will set up the series of tasks for selecting, configuring, and running the deconvolution and database search algorithms. Select the eTHRASH deconvolution and TopPIC database search

algorithms (**Figure 13**). Next, click on the Browse button and select the sample Bovine FASTA file to load into MASH Native. To set advanced parameters, open the Advanced tab and enter the precursor and charge state under the General tab as seen below. Next, enter deconvolution parameters under the Deconvolution tab as seen below. Click on the Start button to begin the analysis.



**Figure 12. Designating mass spectrometry files as MS/MS data in MASH Native.** Users can designate their data as a fragmentation file in MASH native under the chromatogram tab in the main interface (top), selecting the MS level (middle) and changing to MS2 (bottom).



**Figure 13. Deconvolution parameters for processing the ECD spectra for Tutorial 1.** In the Discovery Mode workflow, users can select the deconvolution and database search method and enter the database file (FASTA) in the Basic tab of the Processing Wizard (top left). In the general tab under the advanced section of the Processing Wizard, users can enter information for the  $m/z$  precursor and charge state (top right). In the deconvolution tab under the advanced section of the Processing Wizard, users can enter information for deconvolution (bottom).

### Step 5: Review results

Once the deconvolution and search are complete, the user will be prompted to either append or overwrite the results into the mass list. Select “Append” and your deconvolution and search results will automatically be imported into MASH Native if the run was successful. You can find your deconvolution results in the Mass List and your search results in the Protein Search Results

List (Figure 14). Your saved deconvolution results will be in your working directory under MASH\_Workflows (Figure 15). You may wish to select “Overwrite” to avoid viewing duplicated results.

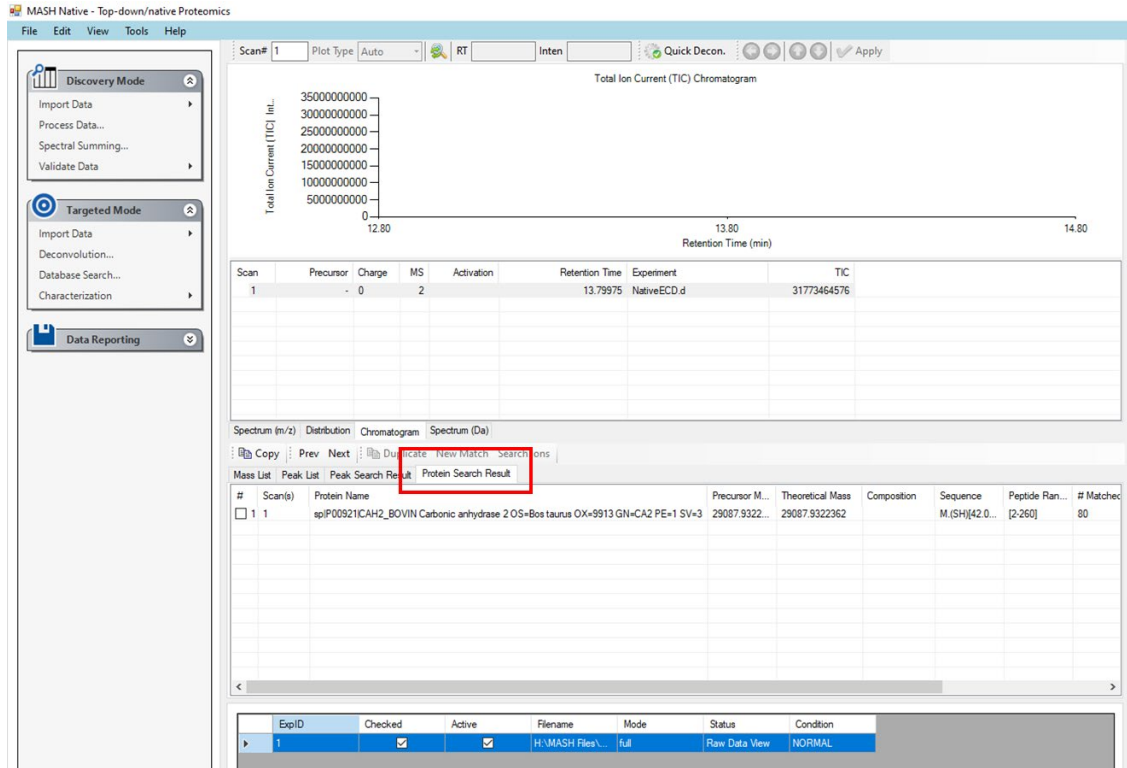
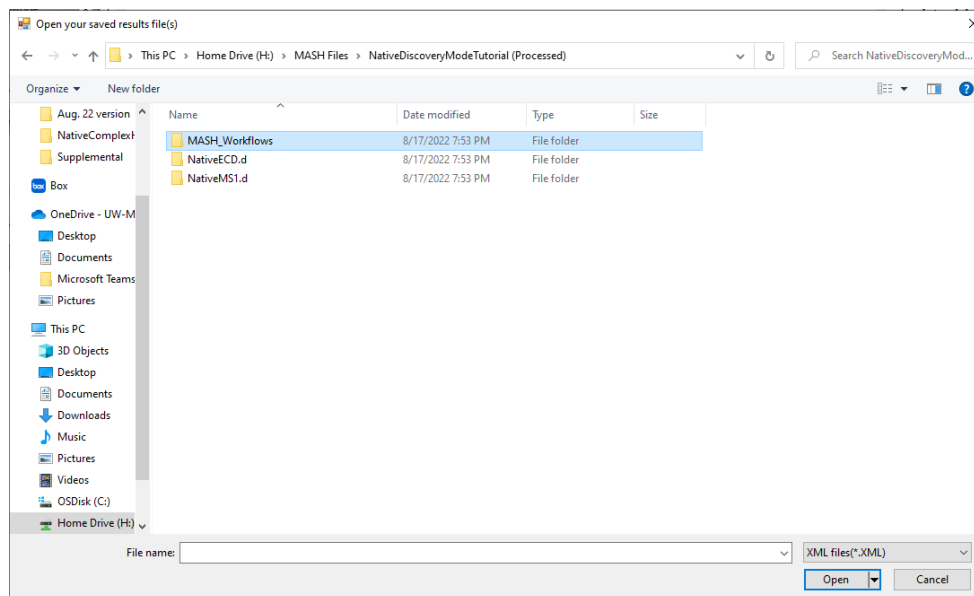


Figure 14. Protein search results view for Tutorial 1. Database search results of the ECD spectra for Tutorial 1 shows bovine carbonic anhydrase as the resulting protein match.

CHECKPOINT: Make sure your results match the photo above



**Figure 15. Workflow directory in MASH Native.**

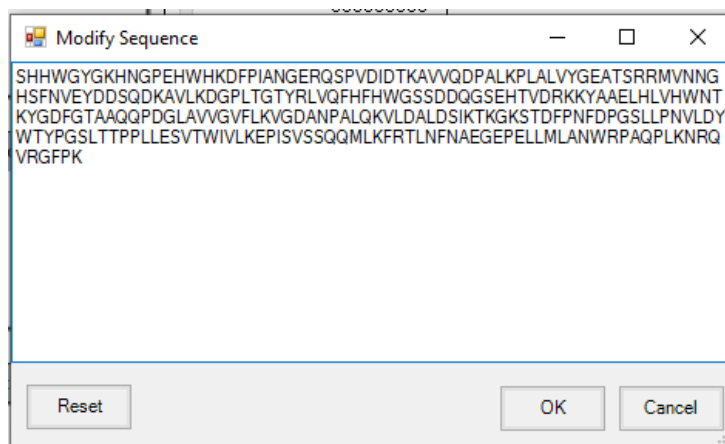
### *Step 6: Annotate results*

MASH Native allows you to view and annotate your imported results with the Mass List and Protein Search Results List, the Results View, and the Sequence Table. For more details on these functions, refer to the *User Manual*. In the Mass List, click on some of the loaded masses to view the fragment ions of a selected scan in the dataset. Use the arrows in the results view to alter the theoretical fit and click the “Apply” button to save your adjustments. You can click on the entry in the Protein Search Results List.

Known modifications for Bovine carbonic anhydrase can be found at [Uniprot.org](https://www.uniprot.org). With this, we can modify the protein sequence and add known modifications. In the Modify Sequence Dialog in the bottom right corner, paste the sequence below as show in **Figure 16**.

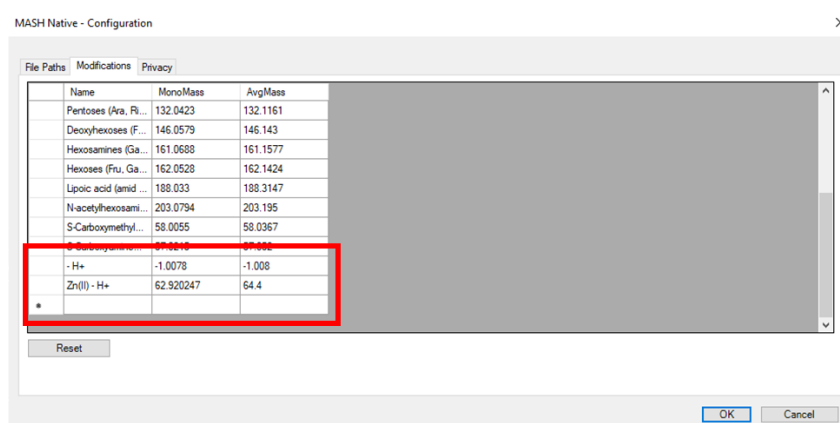
```
SHHWGYGKHNGPEHWHKDFPIANGERQSPVDIDTKAVVQDPALKPLALVYGEATSRR
MVNNGHSFNVEYDDSQDKAVLKDGPLTGTYRLVQFHFHWGSSDDQGSEHTVDRKKY
AAELHLVHWNTKYGDFGTAAQQPDGLAVVGVFLKVG DANPALQKVL DALDSIKTKGK
STDFPNFDPGSLLPNVL DYWTPGSLTTPPLESVTWIVLKEPISVSSQQLKFR TLNFNA
EGPELLMLANWRPAQPLKNRQVRGF PK.
```



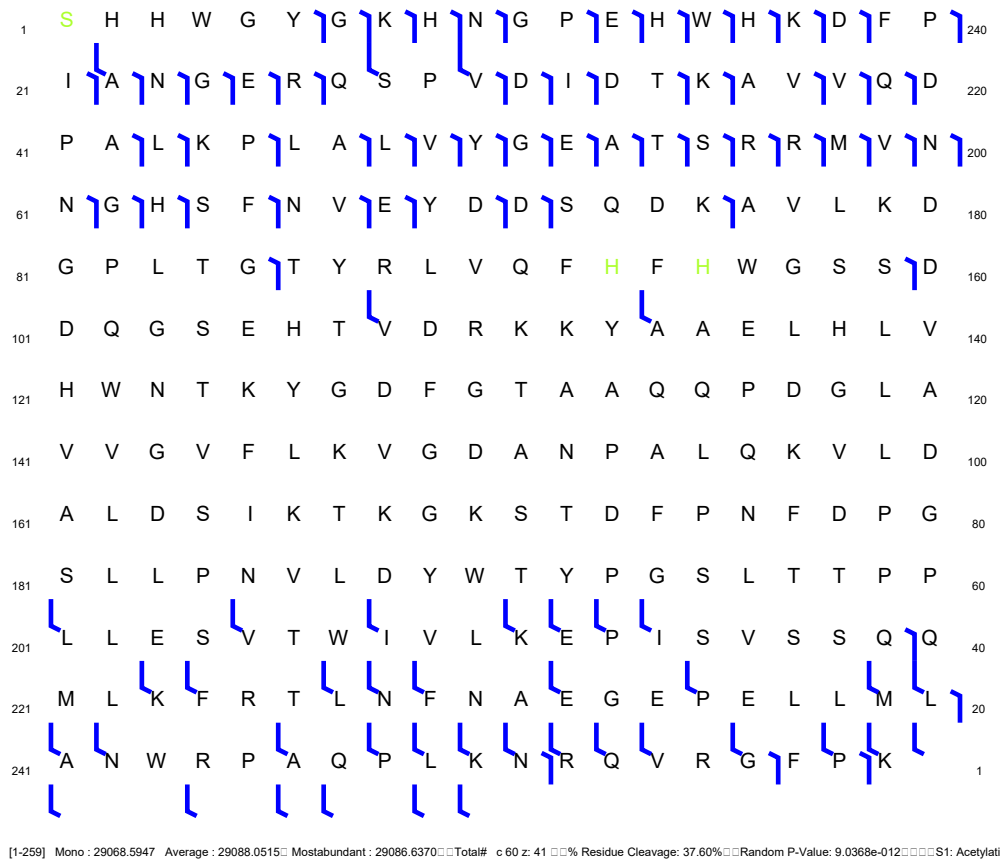


**Figure 16. Editing the sequence for bovine carbonic anhydrase in Tutorial 1.**

Users can then add the following modifications in the Configuration Panel under the Tools menu. Add -H<sup>+</sup> and Zn(II)-H<sup>+</sup> in the table with their respective masses shown in **Figure 17**. Then, using the PTM drop down box under the sequence table, add an acetylation to the N-terminal serine, -H<sup>+</sup> to H93, and Zn(II)-H<sup>+</sup> to H95. Check the *c* and *z*-ions and click the “Show Assignments” button. Try changing the mass tolerance from 1.1Da to 20ppm and see what happens. Users should see a sequence coverage map as shown in **Figure 18**. Additionally, users can use the internal fragment options to improve sequence coverage.



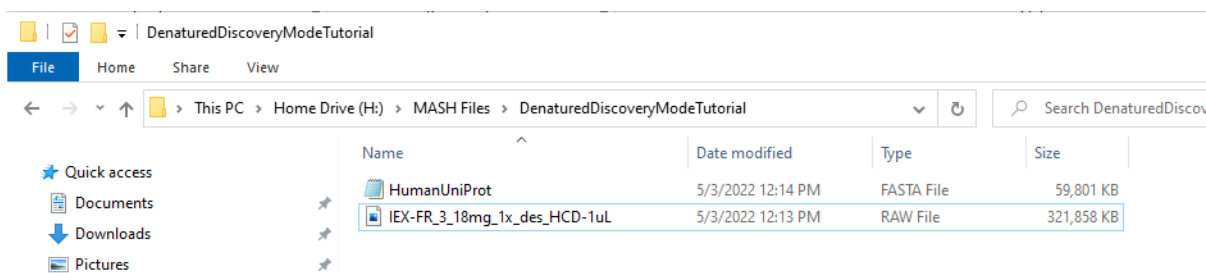
**Figure 17. Adding custom modifications into MASH Native.**



**Figure 18. Sequence coverage map for bovine carbonic anhydrase analyzed in Tutorial 1.**

## **Tutorial 2: Analysis of mass spectrometry data collected under denaturing conditions for identification of an unknown protein.**

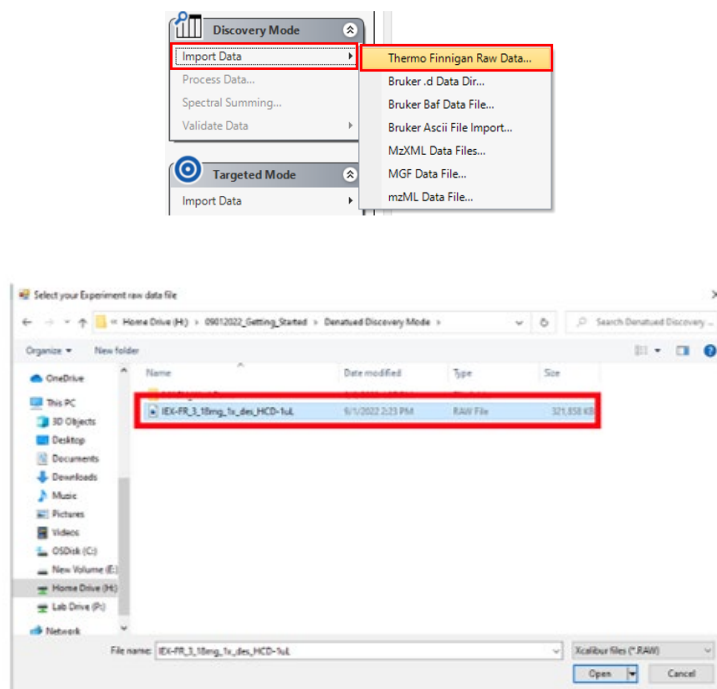
To begin the tutorial, download the [Denatured Discovery Mode Tutorial](#) folder which contains the Thermo IEX-FR\_3\_18mg\_1X-des-HCD-1uL raw file and Human UniProt fasta file and save them to a new local folder on your computer as seen below (**Figure 19**). This data is a data dependent acquisition file for proteins analyzed under denaturing conditions. Proceed with the steps below for analysis of the MS/MS mass spectrometry data to identify unknown proteins from a human cellular extract.



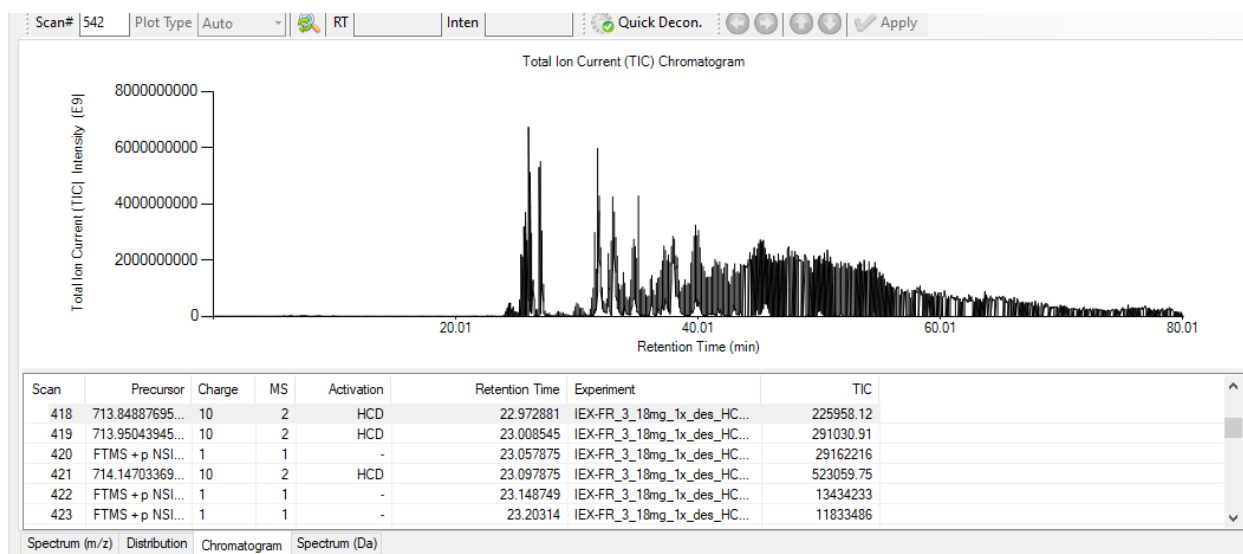
**Figure 19. Downloaded file folder containing needed files for Tutorial 2.**

### *Step 1: Import the MS/MS data*

Use the Import Data function in the Discovery Mode to import the sample Thermo RAW data file (**Figure 20**). Click on the chromatogram tab in the center panel. Note that this data was collected via liquid chromatography-MS/MS (LC-MS/MS) and includes multiple MS1 and MS2 scans (**Figure 21**).



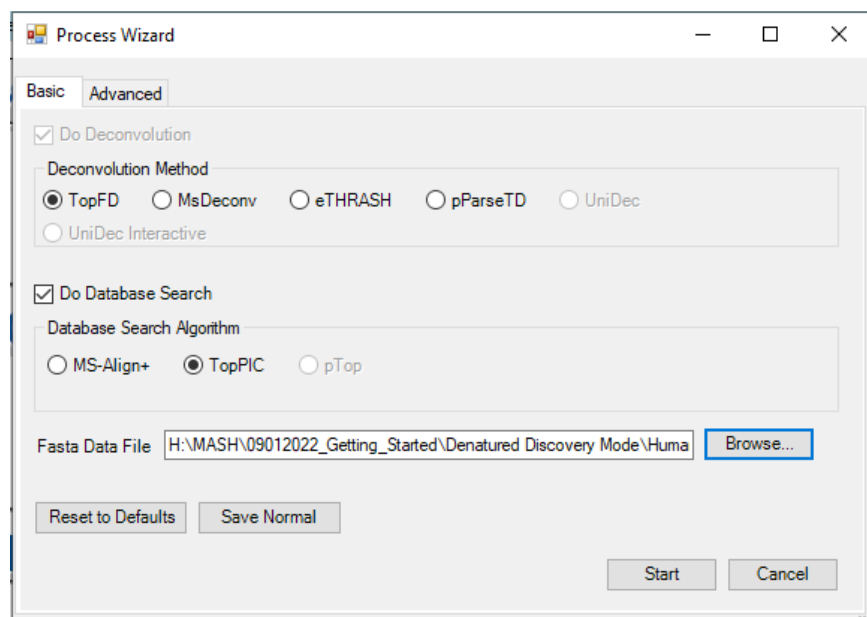
**Figure 20. Importing the MS1 data file for Tutorial 2.** Use the Import Data function in the Discovery Mode node to indicate the Thermo file type (top) and then select the RAW data file from your local directory (bottom).



**Figure 21. Chromatogram for the LC-MS/MS data file in Tutorial 2.**

### *Step 2: Spectral deconvolution of the MS/MS spectra and database search*

Click the Process Data menu item in the Discovery Mode section to open the Process Wizard dialog. The Process Wizard will set up the series of tasks for selecting, configuring, and running the deconvolution and database search algorithms. Select the TopFD deconvolution method and the TopPIC database search algorithm (**Figure 22**). Next, click on the Browse button and select the HumanUniProt.FASTA file to load into MASH Native. Including data filtering steps as in Tutorial 1 in the Advanced parameters will significantly reduce the deconvolution and database search times. To set advanced parameters, open the Advanced tab; however, it is not necessary to set these parameters to begin your workflow. Click the “Start” button. When your deconvolution is completed, you will see the workflow in the Completed Workflows tab of the Workflow Manager and the line “Workflow Finished” in the Log Book.



**Figure 22. Process Wizard showing the deconvolution and database search algorithm for Tutorial 2.**

### *Step 3: Review results*

Your deconvolution and search results will be automatically imported into MASH Native if the run was successful. You can find your deconvolution results in the Mass List and your search results in the Protein Search Results List. Users can manually review the deconvolution and isotopic fit in the Results View panel, if needed. Use the arrows in the *Results View* to alter the theoretical fit and click the “Apply” button to save your adjustments like Tutorial 1 (**Figure 9**).

Once the deconvolution and database search are complete, the user will be prompted to either append or overwrite the results into the mass list. Select “Append” and your deconvolution and search results will automatically be imported into MASH Native if the run was successful. You can find your deconvolution results in the Mass List and your search results in the Protein Search Results List (**Figure 23**). Users can view insightful information here regarding the protein name, sequence, number of matched fragments and the E Value score. From here, you can select each protein and review its sequence and modify as necessary as we did in Tutorial 1 for carbonic anhydrase. When done, be sure to save your analysis using the file save function in the top left

corner. Your saved can later be opened from your working directory under the MASH\_Workflows in your main data folder.

#	Scan(s)	Protein Name	Precursor M...	Theoretical Mass	Composition	Sequence	Peptide Ran...	# Matched Fr...	E Value	Search Algo
<input checked="" type="checkbox"/>	1 418	spIP05387RLA2_HUMAN 60S acidic ribosomal pr...	1417.6478	1417.632116		L.DSVGIEA...	[28-40]	8	3.16E-005	TOPPIC_P..
<input type="checkbox"/>	2 419	spIP05387RLA2_HUMAN 60S acidic ribosomal pr...	1417.6478	1417.632116		L.DSVGIEA...	[28-40]	6	3.38E-004	TOPPIC_P..
<input type="checkbox"/>	3 421	spIP63313ITYB10_HUMAN Thymosin beta-10 OS=...	2818.44529	2818.419611		K.TETQEK...	[21-44]	8	9.38E-009	TOPPIC_P..
<input type="checkbox"/>	4 439	spIP00441ISODC_HUMAN Superoxide dismutase [...]	16079.88441	16079.83941		M.(A)[Acetyl...	[2-154]	24	6.34E-021	TOPPIC_P..
<input type="checkbox"/>	5 440	trIB4DTY3IB4DTY3_HUMAN cDNA FLJ61403, hig...	10384.68951	10384.70951		M.(S)[Acetyl...	[2-88]	17	1.69E-003	TOPPIC_P..
<input type="checkbox"/>	6 442	spIP00441ISODC_HUMAN Superoxide dismutase [...]	16078.81324	16078.91824		M.(A)[Acetyl...	[2-154]	39	4.72E-011	TOPPIC_P..
<input type="checkbox"/>	7 443	spIP00441ISODC_HUMAN Superoxide dismutase [...]	16078.81324	16078.89324		M.(A)[Acetyl...	[2-154]	43	5.88E-013	TOPPIC_P..
<input type="checkbox"/>	8 445	spIP00441ISODC_HUMAN Superoxide dismutase [...]	15990.75894	15990.84394		M.(A)[Acetyl...	[2-154]	39	3.72E-011	TOPPIC_P..
<input type="checkbox"/>	9 446	spIP00441ISODC_HUMAN Superoxide dismutase [...]	15959.75884	15959.78384		M.(A)[Acetyl...	[2-154]	41	2.94E-012	TOPPIC_P..
<input type="checkbox"/>	1 448	spIP00441ISODC_HUMAN Superoxide dismutase [...]	16006.76272	16006.80772		M.(A)[Acetyl...	[2-154]	43	4.52E-013	TOPPIC_P..
<input checked="" type="checkbox"/>	1 449	spIP00441ISODC_HUMAN Superoxide dismutase [...]	16006.76272	16006.99272		M.(A)[Acetyl...	[2-154]	46	1.46E-014	TOPPIC_P..
<input type="checkbox"/>	1 451	spIP00441ISODC_HUMAN Superoxide dismutase [...]	15986.74124	15986.99124		M.(A)[Acetyl...	[2-154]	10	5.26E-009	TOPPIC_P..

**Figure 23. Protein search results view for Tutorial 2.** Database search results of the HCD spectra for Tutorial 2 showing matching protein results.

**CHECKPOINT: Make sure your results match the photo above**

#### Step 4: Annotate results

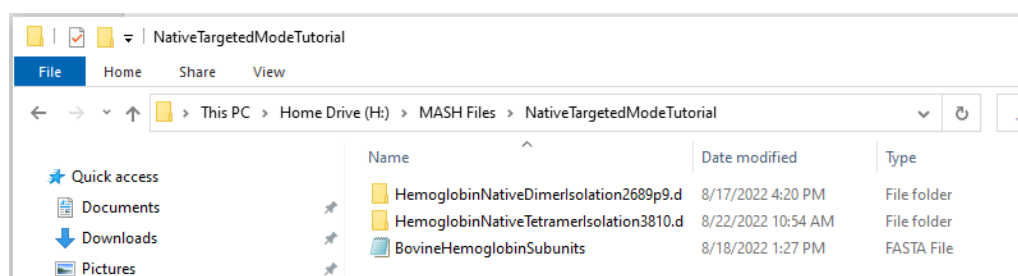
MASH Native allows you to view and annotate your imported results with the Mass List and Protein Search Results List, the Results View, and the Sequence Table. For more details on these functions, refer to the *User Manual*. In the Mass List, click on some of the loaded masses to view the fragment ions of a selected scan in the dataset. Change the scan number in the Chromatogram or at the top of the Results View to view the mass fragments of a different scan. Use the arrows in the results view to alter the theoretical fit and click the “Apply” button to save your adjustments. Like Tutorial 1, users can add modifications and use the internal fragment options to improve sequence coverage. Make sure to Save all your analysis when finished by using the File function in the top left corner of MASH Native’s main interface.

## Targeted Mode Tutorials

The Targeted Mode workflow is intended for the characterization of a known protein. These following tutorials show MASH Native's capability for analysis of both native and denatured mass spectrometry data and includes spectral deconvolution and database search. Please follow the instructions in User Manual *Section 4 Targeted Mode: Protein Identification and Characterization from MS/MS Data*, if you require further assistance processing the data.

### Tutorial 3: Analysis of mass spectrometry data collected under native conditions for the characterization of hemoglobin (*Bos Taurus*) using a Targeted Mode workflow

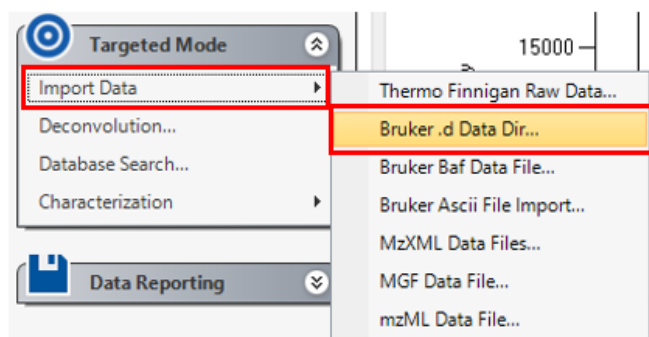
To begin the tutorial, download the [Native Targeted Mode Tutorial](#) folder which contains the Bruker data files for the analysis of Bovine hemoglobin as well as the Bovine Hemoglobin Subunits .fasta file and save them to a new local folder on your computer as seen below (**Figure 24**). Proceed with the steps below for characterization of this protein using MASH Native.



**Figure 24. Downloaded file folder containing needed files for Tutorial 3.**

#### *Step 1: Import data for analysis of hemoglobin tetramer*

Use the Import Data function in Targeted Mode node to import the HemoglobinNativeTetramerIsolation3810.d data file (**Figure 25**).



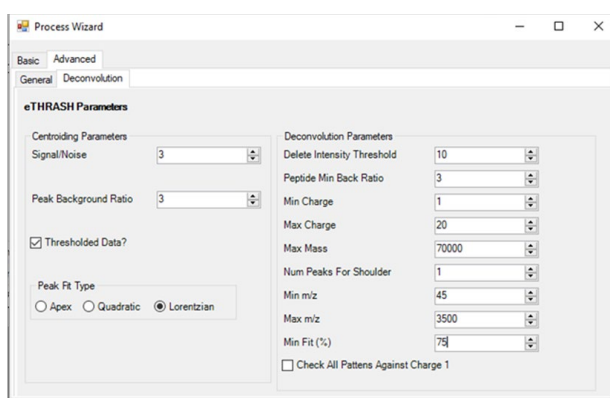
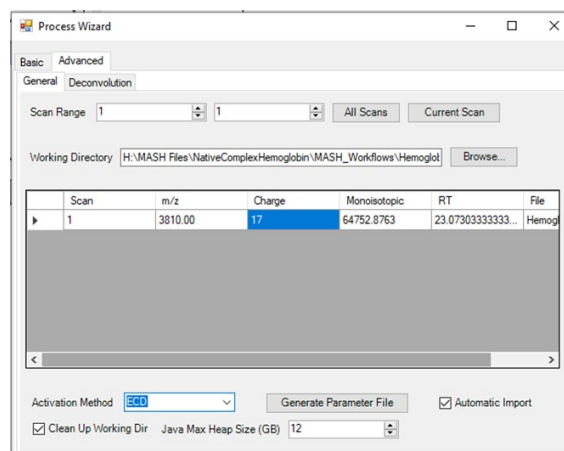
**Figure 25. Importing the data file for the analysis of the hemoglobin tetramer in Tutorial 3.** Use the Import Data function in the targeted Mode node to indicate the Bruker .d Dir (top) and then select the .d file for the analysis of the tetramer from your local directory (bottom).

### *Step 2: Spectral deconvolution of the tetramer data file*

This data shown here is generated from the isolation of the 17+ charge state for the hemoglobin tetramer and ECD fragmentation applied. Users can deconvolute the fragments and identify them using MASH Native's Targeted mode workflow. In the Targeted Mode node, selecting the Deconvolution function will bring up the Process Wizard where users are guided through the deconvolution step (**Figure 26**). Here, select the eTHRASH deconvolution method and input the parameters below for the General and Advanced parameters in the Deconvolution



settings. The Workflow Manager and terminal window will open. You can close the Workflow Manager if you wish, but do NOT close the terminal because this is where your computer is running the deconvolution. When the deconvolution is finished, the Mass List will automatically populate for fragment ion validation.

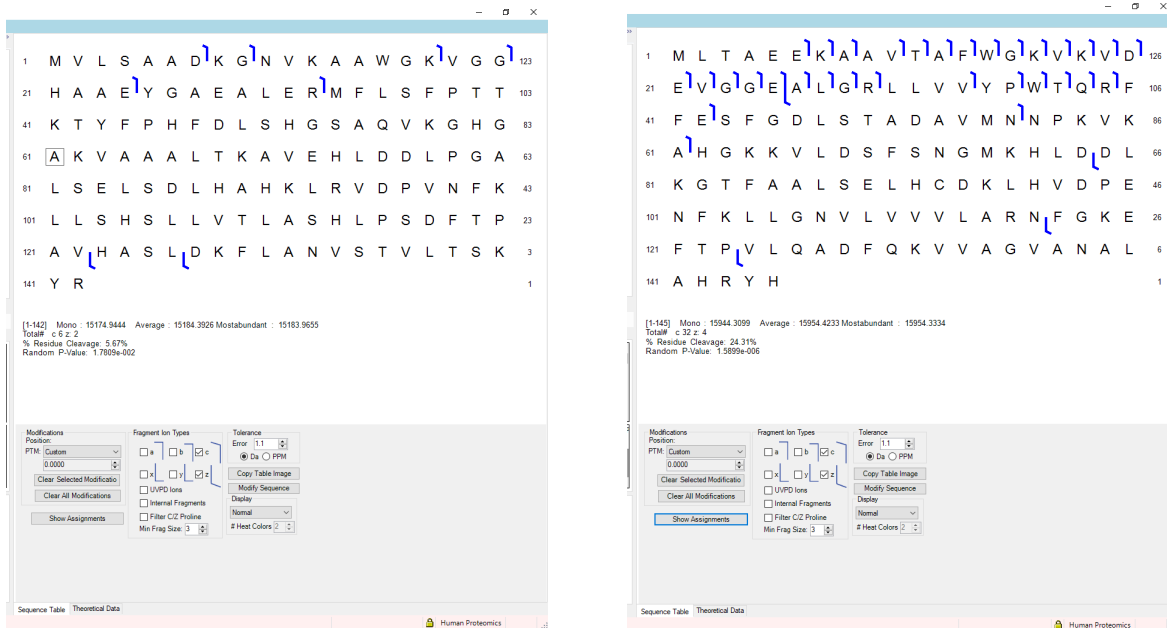


**Figure 26. Deconvolution parameters for processing the ECD spectra for Tutorial 3.** In the Targeted Mode workflow, users can select the deconvolution method and enter information for the  $m/z$  precursor and charge state in the general tab under the Advanced section of the Processing Wizard (top) and parameters for deconvolution in the general tab under the Deconvolution section (right).

### *Step 3: Review results for the analysis of the tetramer*

When the fragment ions are verified, users can use the protein sequences from the alpha and beta hemoglobin subunits to identify fragments (**Figure 27**). Open the

BovineHemoglobinSubunits FASTA file and copy the sequence for the alpha chain and paste it into the modify sequence box in the bottom right corner. Users can then select *c* and *z*-ions and “Show Assignments.” Similarly, users can identify fragments from the beta chain. Additionally, users can improve sequence coverage by adding non-covalently bound iron and other known modifications and include internal fragment analysis into their data analysis. The internal fragments will populate in the mass list. When you are finished, make sure to save your analysis using the file function in the top left corner of MASH Native’s main interface.

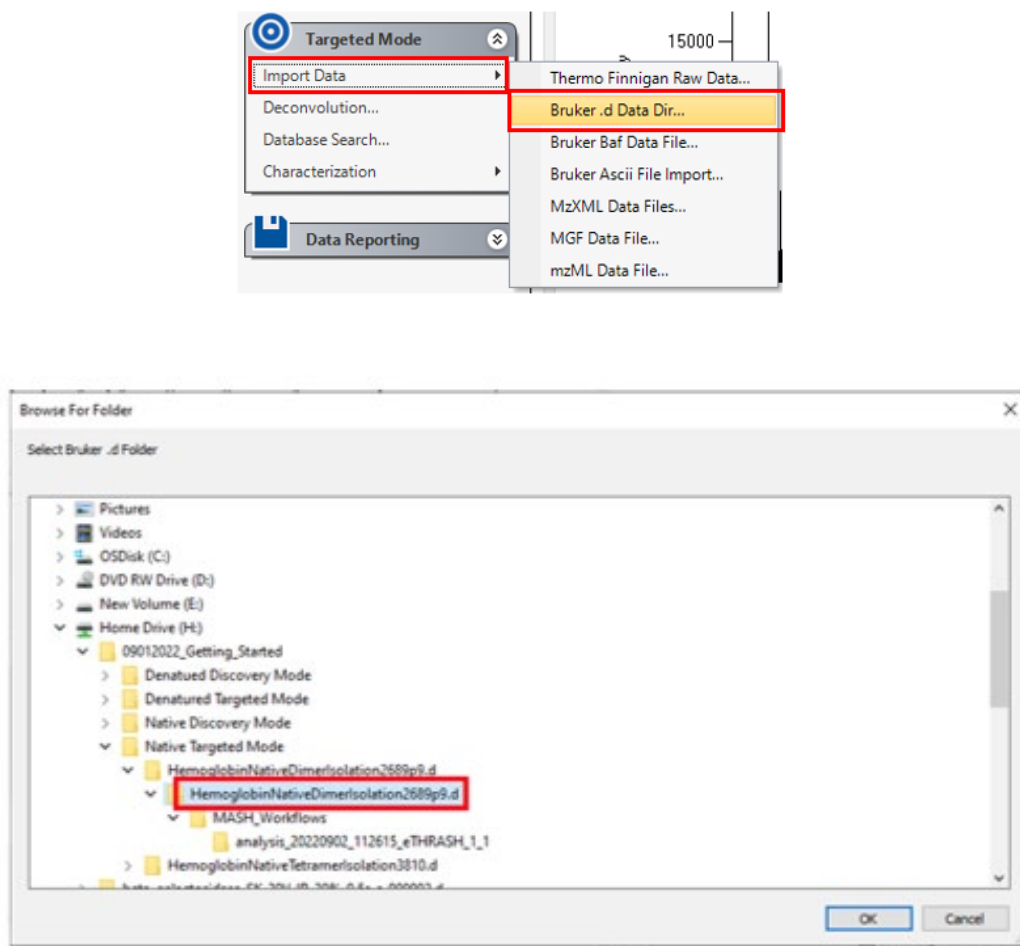


**Figure 27.** Sequence coverage map for hemoglobin analyzed in Tutorial 3. Sequence coverage map for the alpha (left) and beta (right) chains of bovine hemoglobin.

**CHECKPOINT:** Make sure your results match the photo

#### 4. Import data for dissociation of hemoglobin subunits

In a similar manner, import the HemoglobinNativeDimerIsolation2689p9.d data file for the isolation and the hemoglobin dimers and monomer dissociation experiment (**Figure 28**).



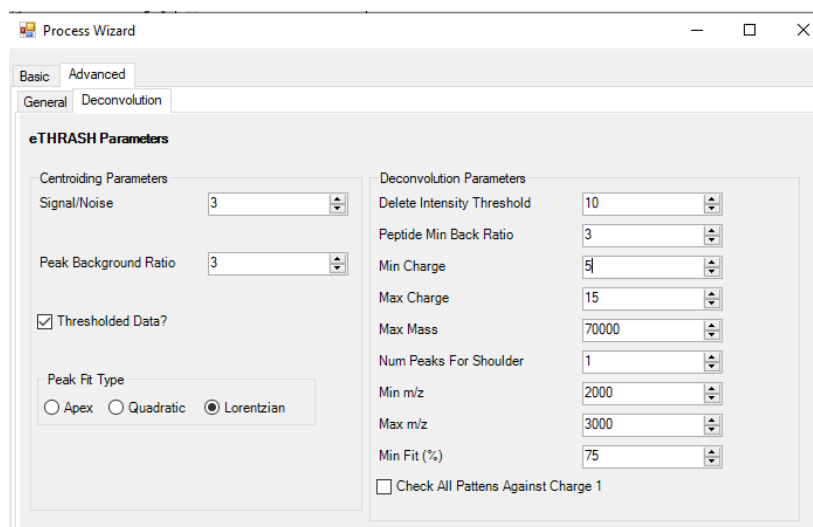
**Figure 28. Importing the data file for the analysis of the hemoglobin dimer in Tutorial 3.**

Use the Import Data function in the targeted Mode node to indicate the Bruker .d Dir (top) and then select the .d file for the analysis of the dimer from your local directory (bottom).

### 5. Spectral deconvolution of the dimer data file

This data shown here is generated from the isolation of the 12+ charge state for the hemoglobin dimer and CAD applied to dissociate the monomers. Users can deconvolute the spectra and identify the charge states of the dimers using MASH Native's Targeted mode workflow. In the Targeted Mode node, selecting the Deconvolution function will bring up the Process Wizard where users are guided through the deconvolution step. Here, select the

eTHRASH deconvolution method and input the parameters below (**Figure 29**). When complete, go back to the Basic tab and click start to begin the deconvolution. The Workflow Manager and terminal window will open. You can close the Workflow Manager if you wish, but do NOT close the terminal because this is where your computer is running the deconvolution. When the deconvolution is finished, the Mass List will automatically populate for fragment ion validation.



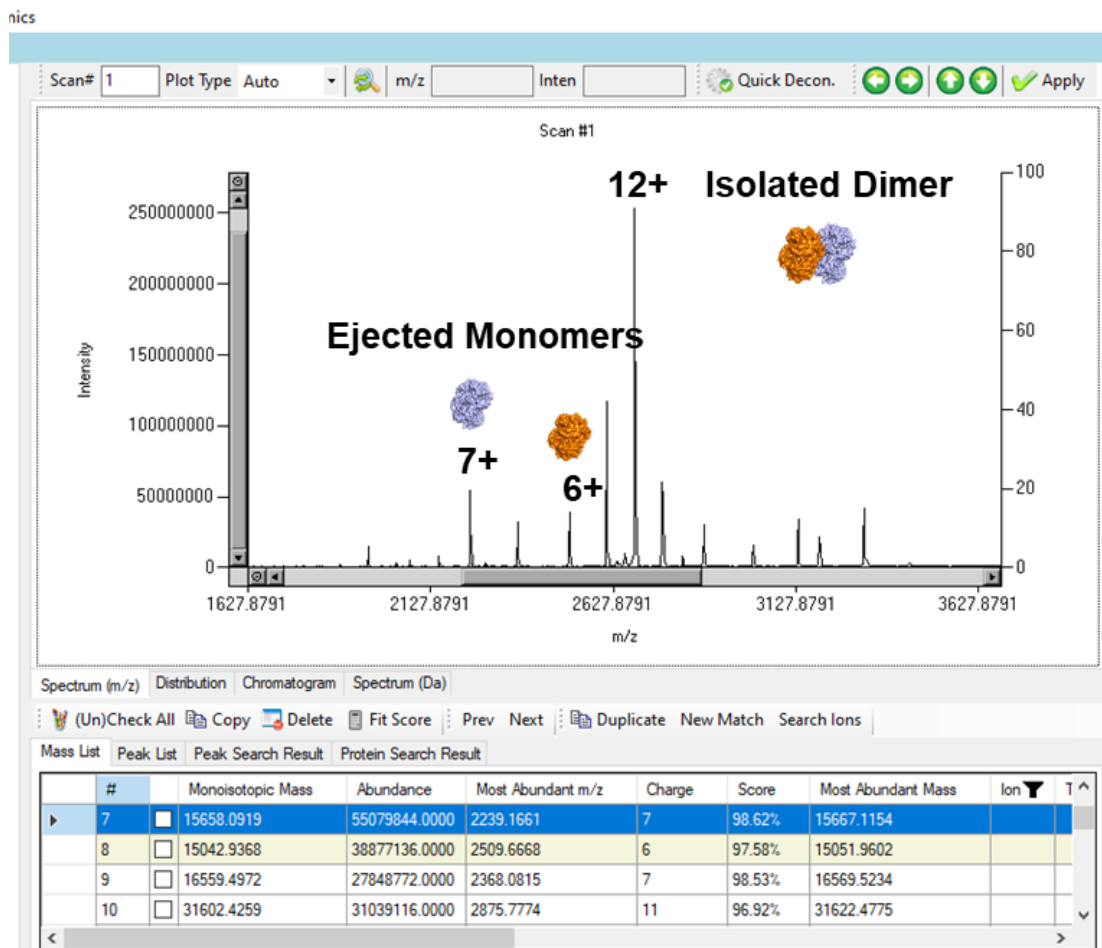
The image shows a screenshot of the 'Process Wizard' application window, specifically the 'Advanced' tab and 'Deconvolution' sub-tab. The 'eTHRASH Parameters' section is active, containing two columns of settings:

Centroiding Parameters	Deconvolution Parameters
Signal/Noise: 3	Delete Intensity Threshold: 10
Peak Background Ratio: 3	Peptide Min Back Ratio: 3
<input checked="" type="checkbox"/> Thresholded Data?	Min Charge: 5
Peak Fit Type: <input type="radio"/> Apex <input type="radio"/> Quadratic <input checked="" type="radio"/> Lorentzian	Max Charge: 15
	Max Mass: 70000
	Num Peaks For Shoulder: 1
	Min m/z: 2000
	Max m/z: 3000
	Min Fit (%): 75
	<input type="checkbox"/> Check All Patterns Against Charge 1

**Figure 29. Deconvolution parameters for processing the CAD spectra for Tutorial 3.** In the Targeted Mode workflow, users can select the eTHRASH algorithm and enter parameters in the Deconvolution tab under the Advanced section.

## 6. Review results

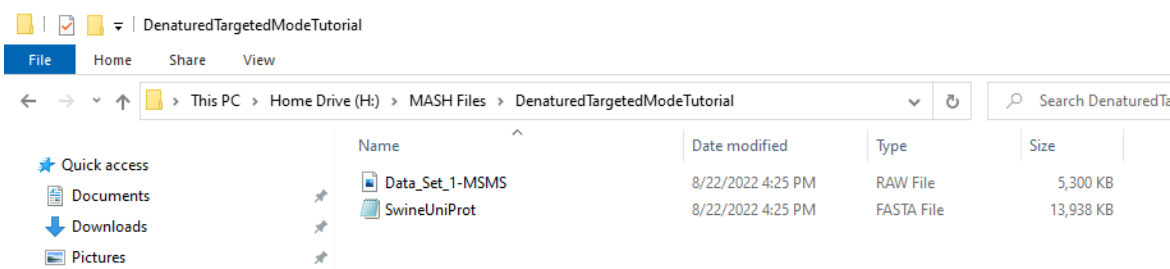
When the deconvolution finishes, results are automatically imported into the Mass List where they can be manually reviewed, and isotopic peak fit evaluated in a similar manner as reviewed in Tutorial 1. From this data, users can identify the dimer and ejected monomers as annotated below (**Figure 30**). Users can click and drag on the spectrum to zoom in on a desired area of the spectrum, as seen in the photo above. Be sure to save your results as MASH will not prompt you to do so. When you are finished, make sure to save your analysis using the file function in the top left corner of MASH Native's main interface.



**Figure 30. Annotated spectra for the analysis of the hemoglobin dimer in Tutorial 3.**

#### **Tutorial 4: Analysis of mass spectrometry data collected under denaturing conditions for the characterization of NADH dehydrogenase [ubiquinone] iron-sulfur protein 5-like (*Sus scrofa*) using a Targeted Mode workflow**

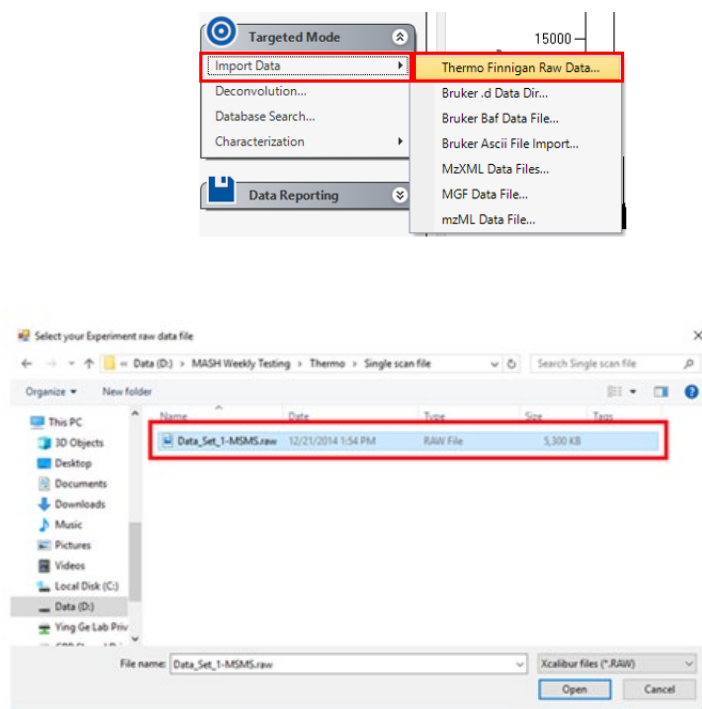
To begin the tutorial, download the [Denatured Targeted Mode Tutorial](#) folder which contains the Thermo Data\_Set\_1-MSMS.raw file for the analysis of NADH dehydrogenase [ubiquinone] iron-sulfur protein 5-like as well as the Swine UniProt .fasta file and save them to a new local folder on your computer as seen below (**Figure 31**). Proceed with the steps below for characterization of this protein using MASH Native.



**Figure 31. Downloaded file folder containing needed files for Tutorial 4.**

*Step 1: Import data for analysis of the MS/MS data*

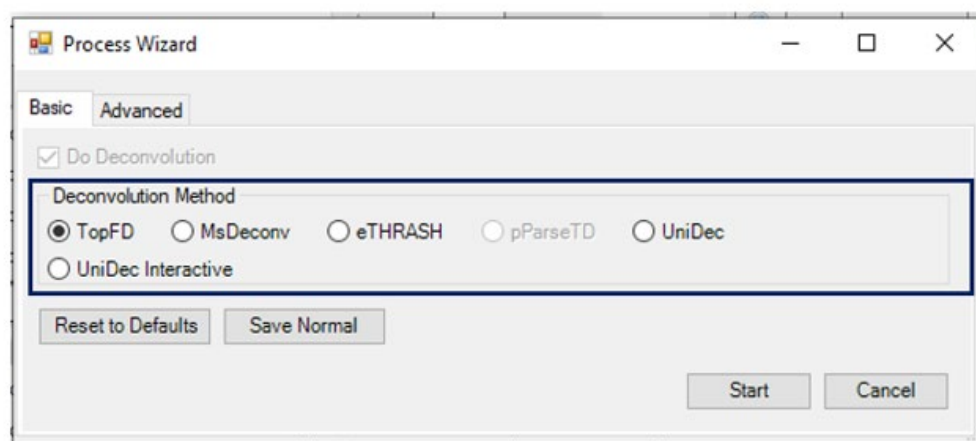
Use the Import Data function in Targeted Mode node to import the Thermo RAW data file using the Import Data function in the Targeted Mode node (**Figure 32**).



**Figure 32. Importing the data file for the MS/MS analysis of NADH dehydrogenase [ubiquinone] iron-sulfur protein 5-like in Tutorial 4.** Use the Import Data function in the targeted Mode node to indicate the Thermo Raw file type (top) and then select the RAW file for from your local directory (bottom).

### *Step 2: Deconvolution of the MS/MS data file*

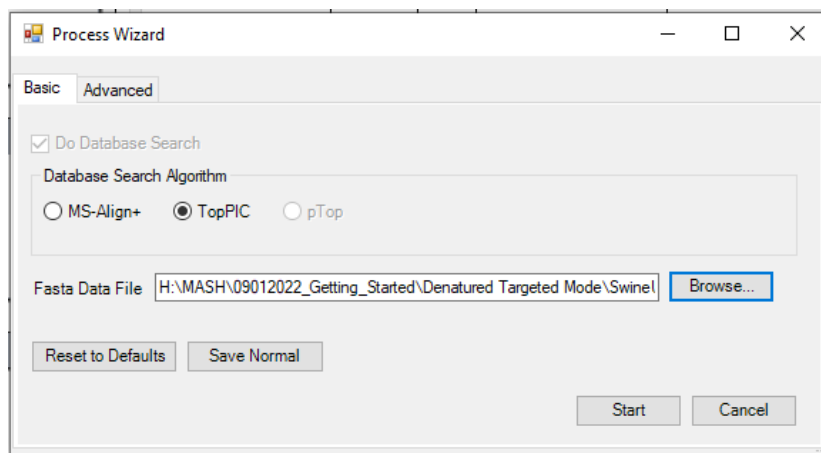
Under the Targeted Mode node, the Deconvolution function will bring up the Process Wizard which will guide users through the deconvolution step. Here, select the TopFD deconvolution method and click the Start button (**Figure 33**). You have the option to set more parameters in the Advanced tab, but this is not necessary to start your run. The Workflow Manager and terminal window will open. You can close the Workflow Manager if you wish, but do NOT close the terminal because this is where your computer is running TopFD. When the deconvolution is finished, the Mass List will automatically populate shown in the image below.



**Figure 33. Deconvolution parameters for processing the MS/MS spectra for Tutorial 4.**

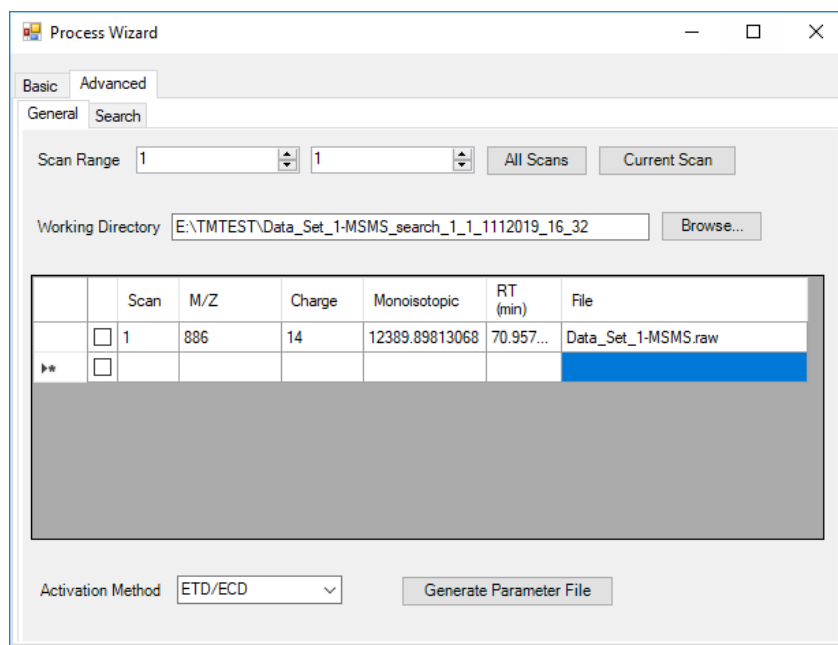
### *Step 3: Database search*

Following deconvolution, users can use the selected ion list for database search via the Process Wizard. In the Targeted Mode node, select the Database Search function, which will open the Process Wizard (**Figure 34**). Select the TopPIC database search to run on the selected ion list shown in the Mass List. Click on the Browse button to load the SwineUniProt.fasta file.



**Figure 34. Database search parameters for processing the MS/MS spectra for Tutorial 4.**

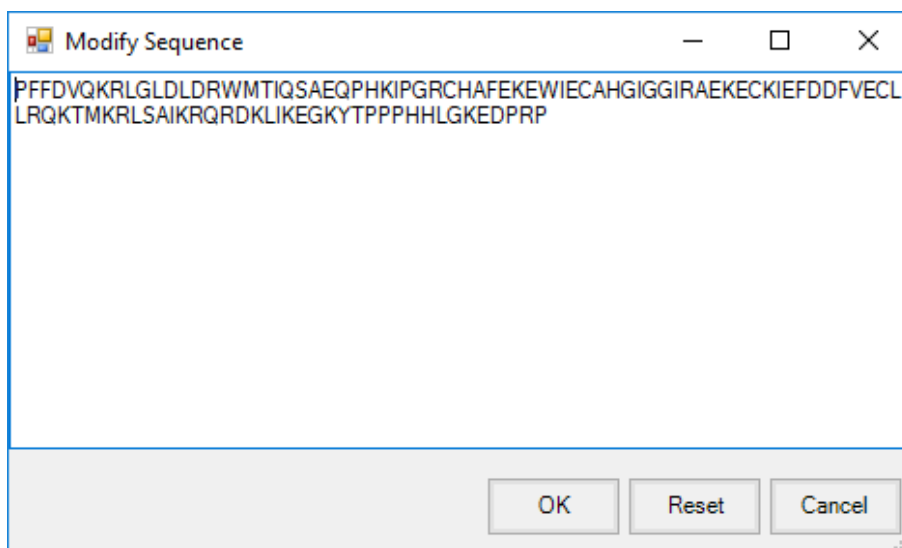
Processing of this data file requires Advanced parameters to be set (**Figure 35**). Please click on the Advanced Parameters tab. Change the charge from 0 to 14 and hit Enter on your keyboard, so that the parameters in the list are exactly as shown below. The monoisotopic mass will automatically be calculated. Ensure that the Activation is set at ETD/ECD. Navigate back to the basic tab and click on the start button to begin your search.



**Figure 35. Advanced parameters for processing the MS/MS spectra for Tutorial 4.**

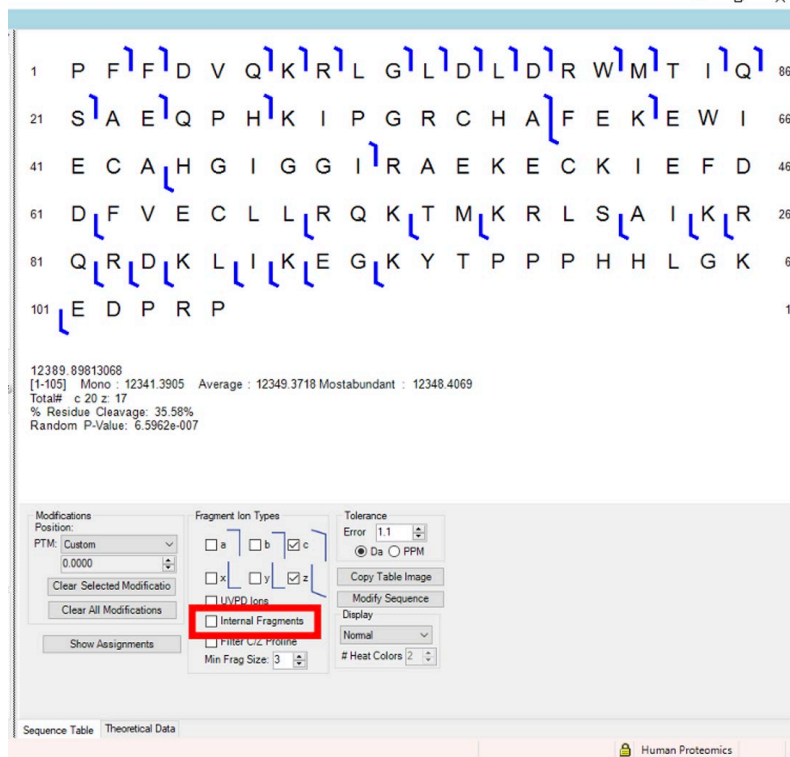






**Figure 37. Editing the sequence for bovine carbonic anhydrase in Tutorial 4.**

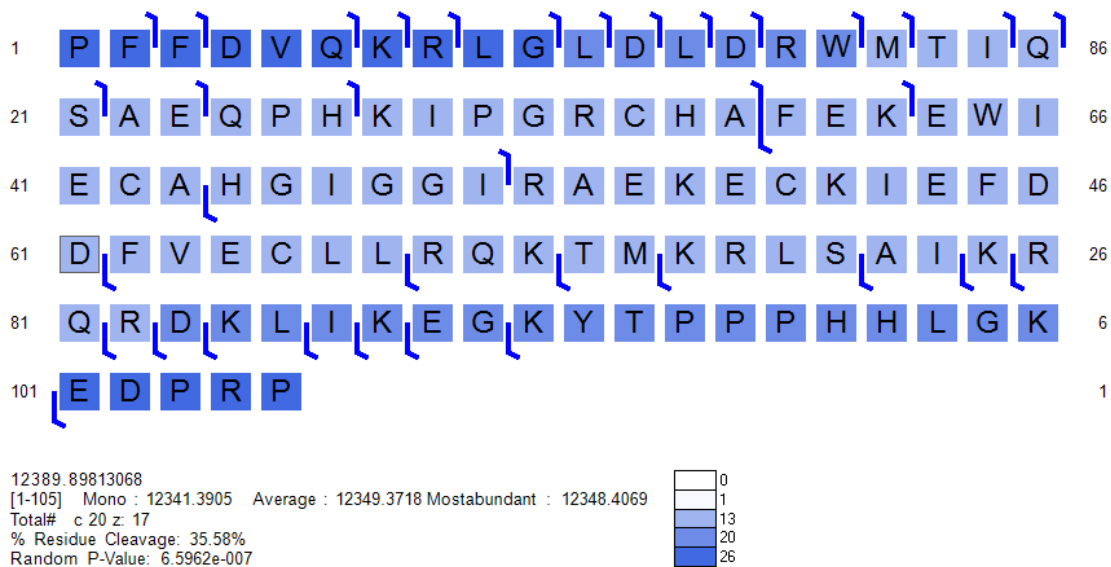
In the Sequence Table, check the c and z-ion types and click on the “Show Assignments” button to view the fragmentation matches for the protein. If users want to view internal fragmentation, select the Internal Fragments checkbox (**Figure 38**) before clicking the “Show Assignments” button, which matches monoisotopic masses with internal fragment ion types (e.g., c and z). With internal fragmentation enabled, expect a longer wait time before the GUI becomes responsive. Users can click on each entry in the Mass List to see the profile data.



**Figure 38. Internal fragments checkbox.**

### *Step 5: Using the Heat Map function*

To view the heat map, select Heat Map in the menu under Display in the bottom right corner. This gives a visual representation of the relative number of ions in the mass list in which each amino acid is present (**Figure 39**). You can change the number of heat colors for the map, we have shown the map with 5 heat colors below. Be sure to save all data when you are finished with your analysis using the File function in the top left corner of MASH Native's main interface.



**Figure 39.** Sequence table when using the heat Map display option in Tutorial 4.

### Tutorial 5: Post-Translational Modification Analysis using UniDec within MASH Native

To begin the tutorial, download the [Denatured Targeted Mode Tutorial](#) folder which contains the Thermo Data\_Set\_1-MSMS.raw file for the analysis of NADH dehydrogenase [ubiquinone] iron-sulfur protein 5-like as well as the Swine UniProt .fasta file and save them to a new local folder on your computer as seen in (**Figure 31**). Note, this is the same data file used for Tutorial 4. Proceed with the steps below for characterization of this protein using UniDec in MASH Native.

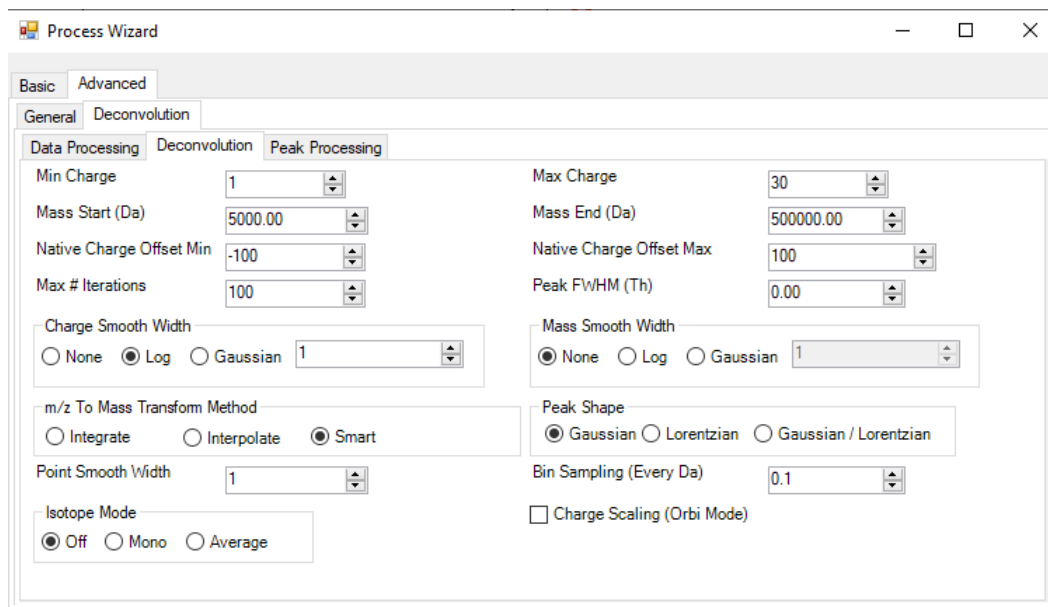
#### *Step 1: Import data for analysis of the MS/MS data*

Use the Import Data function in Targeted Mode to import the sample Thermo data file as seen in **Figure 32**.

#### *Step 2: Run UniDec deconvolution*

Under the Targeted Mode node, select the UniDec deconvolution Process Wizard will guide users through the deconvolution step. Select the UniDec deconvolution method and enter the same parameters as seen in **Figure 35**. Under the Deconvolution tab of the deconvolution tab

under the Advanced tab of the Process Wizard, change the max charge to 30 and bin sampling to 0.1 (**Figure 40**). Once these parameters have been entered, go back to the Basic tab and press “Start” to begin the deconvolution.



**Figure 40. Advanced parameters for the UniDec deconvolution of the MS/MS spectra for Tutorial 5.**

### *Step 3: Import Sequence*

In MASH, click on “Modify Sequence” in the bottom left corner and the “Paste Sequence Information” window will open. Paste the sequence below and click “OK” and the sequence will be displayed in the Sequence Table (**Figure 41**). In the table, it will calculate the most abundant mass of the sequence.

MPFFDVQKRLGLDLDRWMTIQSAEQPHKIPGRCHAFEKEWIECAHGIGGIRAEKECKIEF  
DDFVECLLRQKTMKRLSAIKRQRDKLIKEGKYTPPPHHLGKEDPRP

```

1  M P F F D V Q K R L G L D L D R W M T I 87
21 Q S A E Q P H K I P G R C H A F E K E W 67
41 I E C A H G I G G I R A E K E C K I E F 47
61 D D F V E C L L R Q K T M K R L S A I K 27
81 R Q R D K L I K E G K Y T P P P H H L G 7
101 K E D P R P 1

```

[1-106] Mono : 12472.4310 Average : 12480.570 Mostabundant : 12479.4474

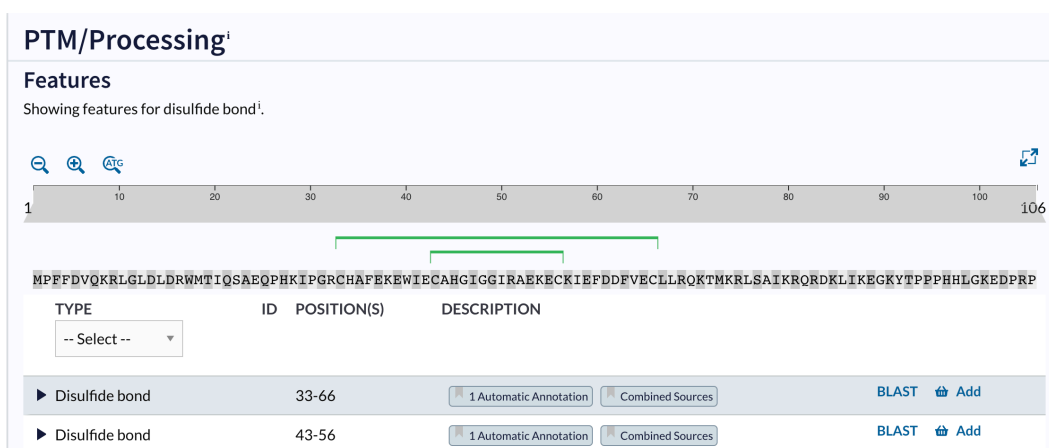
**Figure 41. Edited sequence table for Tutorial 5.**

*Step 4: Review mass list*

Once the deconvolution is complete, the Mass List will be populated. Manually review the deconvolution and isotopic fit to identify a high score charge assignment of most abundant masses that is closest to the Most Abundant Mass of the sequence above (**Figure 42**).

*Step 5: Post-Translational Modifications*

Known modifications for NDUFS5 can be found at [Uniprot.org](https://www.uniprot.org). With this, we can modify the protein sequence and add known modifications. Using the UniProt link, review possible post-translational modifications (PTMs) for the NDUFS5 protein. Note that there are two disulfide bonds in the sequence (**Figure 43**). This modification can be added to MSH Native's list of custom modifications (**Figure 44**). Users can then add the following modifications in the Configuration Panel under the Tools menu.



**Figure 43. Known disulfide bond positions in NDUFS5.**

Name	MonoMass	AvgMass
Deoxyhexoses (F...	146.0579	146.143
Hexosamines (Ga...	161.0688	161.1577
Hexoses (Fru, Ga...	162.0528	162.1424
Lipoic acid (amid ...	188.033	188.3147
N-acetylhexosami...	203.0794	203.195
S-Carboxymethyl...	58.0055	58.0367
S-Carboxyamine...	57.0215	57.052
Zn(II) -H+	62.920247	64.4
-H+	-1.0078	-1.008
Disulfide	-1.0078	-1.008


**Figure 44. Adding a disulfide bond modification into MASH Native.**

In the sequence table, add the modifications observed in UniProt. First, add a mass loss of 1.0078 Da to each of the four cysteine residues (C33, C43, C56, and C66) which are a part of disulfide bonds (to account for H loss upon disulfide bond formation). After adding the modifications, select “Show Assignments”. This will change the most abundant mass in the sequence table (**Figure 45**). Notice that the most abundant mass in the Mass List in the center panel does not match the Most abundance mass in the sequence table (**Figure 45**), indicating that this combination of PTMs is not correct for the protein that was measured. With this, the user can

adjust the modifications to match the most abundant mass observed in the Mass List. The difference between the Mass List and the calculated mass of the sequence is 133.047 Da. This mass shift is the removal of the N-terminal methionine and one disulfide bond. Using the “Modify Sequence” function, delete the methionine (M) in position one of the sequences and one of the disulfide bonds, either C43 and C56 or C33 and C66. When complete, select “Show Assignments” to observe the updated sequence table (**Figure 46**). Notice that the most abundant mass in the sequence table is not in agreement with the Mass List. Note, that MS2 experiments will be needed to confirm the exact location of the detected disulfide bond, but the MS1 results inform users of the need to search for disulfide bonds in MS2 or MS3 spectra.

1	M	P	F	F	D	V	Q	K	R	L	G	L	D	L	D	R	W	M	T	I	87
21	Q	S	A	E	Q	P	H	K	I	P	G	R	C	H	A	F	E	K	E	W	67
41	I	E	C	A	H	G	I	G	G	I	R	A	E	K	E	C	K	I	E	F	47
61	D	D	F	V	E	C	L	L	R	Q	K	T	M	K	R	L	S	A	I	K	27
81	R	Q	R	D	K	L	I	K	E	G	K	Y	T	P	P	P	H	H	L	G	7
101	K	E	D	P	R	P															1

[1-106] Mono : 12468.3998 Average : 12476.538 Mostabundant : 12475.4162  
 Total#  
 % Residue Cleavage: 0.00%  
 Random P-Value: 1.2806e-016  
 C33: -H+ C43: -H+ C56: -H+ C66: -H+



**Figure 45. Sequence table showing the incorrect combination of PTMs.**



1 P F F D V Q K R L G L D L D R W M T I Q 86  
21 S A E Q P H K I P G R C H A F E K E W I 66  
41 E C A H G I G G I R A E K E C K I E F D 46  
61 D F V E C L L R Q K T M K R L S A I K R 26  
81 Q R D K L I K E G K Y T P P P H H L G K 6  
101 E D P R P 1

[1-105] Mono : 12339.3749 Average : 12347.355 Mostabundant : 12346.3913  
Total#  
% Residue Cleavage: 0.00%  
Random P-Value: 5.3657e-003  
C42: Disulfide C55: Disulfide



**Figure 46. Sequence table showing the combination of PTMs.**