# **Getting Started with MASH Native**

## Version 1.1

## **Overview:**

MASH Native, developed by the Ying Ge research group, is a comprehensive, universal, and user-friendly, free software environment for the analysis of top-down native and denatured proteomics data. MASH Native builds on previous work with MASH Suite, MASH Suite Pro, and MASH Explorer.

The tutorial below provides a brief introduction to the basic functions of MASH Native in both **Discovery Mode** (MS/MS datasets) and **Targeted Mode** (single or targeted protein datasets) for top-down proteomics native and denatured analysis.

- Tutorial 1: Analysis of native MS and MS/MS data for an unknown protein using Discovery Mode.
- Tutorial 2: Analysis of denatured MS/MS data for an unknown protein using Discovery Mode.
- Tutorial 3: Analysis of native data for the characterization of hemoglobin (*Bos Taurus*) using Targeted Mode.
- Tutorial 4: Analysis of denatured data for the characterization of NADH dehydrogenase [ubiquinone] iron-sulfur protein 5-like (*Sus scrofa*) using Targeted Mode.
- Tutorial 5: Post-Translational Modification Analysis using UniDec within MASH Native

Download the sample data in each section and follow the instructions below to begin using MASH Native. Video tutorials are also available on <u>YouTube<sup>TM</sup></u>. For a detailed description of the functions of the MASH Native software, please view the <u>User Manual</u>.

## **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

(<u>https://labs.wisc.edu/gelab/MASH\_Explorer/doc/MASH%20Native%20Privacy\_Notice.pdf</u>) 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 <u>Installation Guide</u>, available on our <u>webpage</u> and in the software download folder.

Changes needed:

- Show configuration dialog with only the required software green. --> Insert screenshot when configuration dialog is finalized (I would insert this here and tell users this is the essential software needed to perform these basic familiarizations)
- Will add in splash screen to explain that the software gets downloaded then.

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).

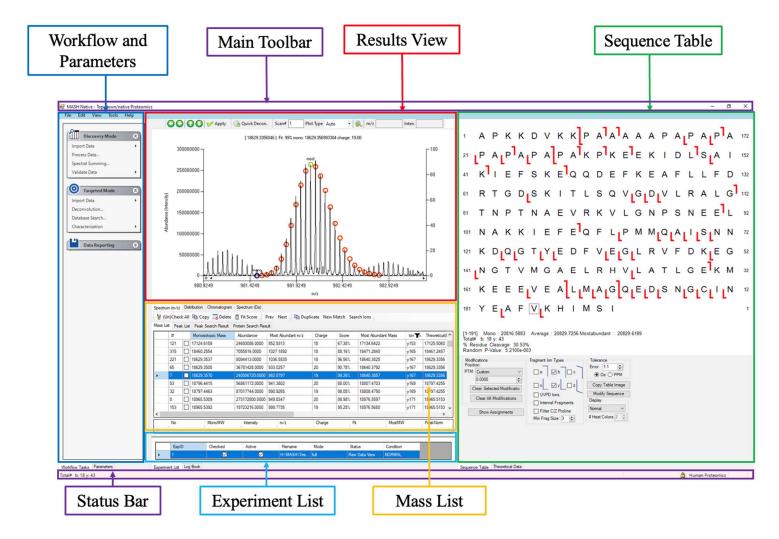
	Workflow Manager	Plot Type	Auto		m/z	Inten	
-	Quick Decon. Parameters	the type	Mato	300	110.2	mosn	
Discovery	Paste Sequence						
Import Data	Spectral Summing						
Process Data		_					
The second s	Configuration						
Spectral Summing	View Log File						
Validate Data	Open AppData						
Import Data Deconvolution	*						
Database Search							

File Paths Modifications Privacy				
Java Bin Directory	C:\Program Files\Java\jre1.8.0_341\bin	Find	Browse	Download
ProteoWizard Directory	C:\Users\memoss\AppData\Local\Apps\ProteoWizard 3.0.21229.9668f52 64-	Find	Browse	Download.
nformed-Proteomics Directory	C:\Program Files\Informed-Proteomics	Find	Browse	Download.
opPIC Directory	C:\toppic-windows-1.5.4\toppic-windows-1.5.4	Find	Browse	Download.
MS-Deconv Jar	C:\Users\memoss\Downloads\MsDeconvConsole-0.8.0.7370.jar	Find	Browse	Download.
MS-Align+ Zip	C:\MS-Align-0.7.1.7143.zip	Find	Browse	Download.
oTop Directory	C:\Users\memoss\Desktop\MASH Software Downloads\pTop_0914\pTop	Find	Browse	Download.
OpenMS Directory	C:\Program Files\OpenMS-3.0.0-pre-HEAD-2022-06-23	Find	Browse	Download.
IniDec Distribution Directory	C:\Users\memoss\Desktop\MASH Software Downloads\UniDec_Windows_2	Find	Browse	Download.
	More Information		D	ownload All
	MASH Native Installation Guide.	Tutori	ial Video	

This tutorial uses TopPIC Suite and UniDec; the other search and deconvolution tools shown above are not required for the tutorial, but you may want to install them. You also need to install ProteoWizard.

Please note that this tutorial uses TopPIC Suite version 1.5.4 and UniDec 5.1.1.

## **Interface Layout**



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.

## **Discovery Mode Tutorials:**

Discovery Mode is intended for MS/MS data analysis. These sample data show MASH Native's capability for analysis of both native and denatured MS/MS 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 MS/MS native top-down mass spectrometry data:

For the analysis of native mass spectrometry data, download the folder containing the native tutorial files and save them to a new local folder as seen below.

Download the folder below on the Google Drive for the native MS/MS tutorial

• <u>Native Discovery Mode Tutorial</u> folder which contains the Bruker NativeMS1 and NativeECD data files and BovineUniProt.fasta file.

ile Home Share	View						Y
🚽 👻 🛧 📙 > This P	C → Home Driv	e (H:) > MASH Files > NativeDiscov	veryModeTutorial	~	ō	,○ Search NativeDiscoveryMod	eTutorial
		Name	Date modified	Туре		Size	
<ul> <li>Quick access</li> <li>Documents</li> </ul>		NativeECD.d	8/22/2022 12:20 PM	File folde	2F		
	*	NativeMS1.d	8/22/2022 12:20 PM	File folde	21		
- Downloads	*	BovineUniProt	8/10/2022 11:59 AM	FASTA Fi	e	2,967 KB	
Pictures	*						

Proceed with the following steps for analysis of native MS/MS mass spectrometry data to identify an unknown protein from a bovine extract.

#### Step 1: Import MS1 native data

Use the Import Data function in the Discovery Mode to import the sample Bruker .d 'Native MS1' folder.

MASH Native - Top-down/native Prof File Edit View Tools Help Discovery Mode & Import Data · Process Data Spectral Summing Validate Data · Import Data · Deconvolution Database Search Characterization ·	Scan# Plot Type Auto	Browse For Folder  Select Bruker .d Folder  Select Bruker .def  Select Bruke	
Data Reporting 😵	2	betage indices indi	OK Cancel

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 Quick Decon. will use the desired algorithm, under Tools, select "Quick Decon. Parameters." This will open the Process Wizard dialog and the user can select their deconvolution method. Select eTHRASH for this tutorial and click "Set Quick Decon."

🛃 MASH Native - Top-	down/native Proteomics		
File Edit View	Tools Help		💀 Process Wizard — 🗆 🗙
	Workflow Manager	Plot Type Auto	
	Quick Decon. Parameters		Basic Advanced
Import Data Process Data Spectral Summing Validate Data	Paste Sequence Spectral Summing Configuration	400000000-	Do Deconvolution      Deconvolution Method      TopFD O MsDeconv      eTHRASH      pParseTD ProMEX O FLASHDeconv      UniDec O UniDec Interactive      Reset to Defaults
Targeted N	Node 🛞	 300000000 —	Set Quick Decon. Cancel

Select the QuickDecon. 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 use the Workflow Manager to check on the status of your run, but note that closing this window will stop your workflow. When your run 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.

## November 12, 2022

## MASH Native App

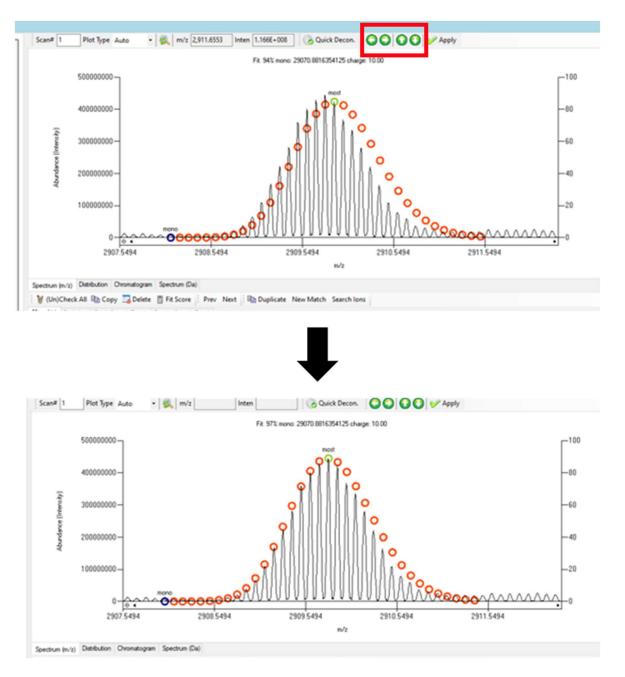
dit View Tools Help	omics				- 0
and the second se					
	Scan# 1 Plot Type Auto • 🕵 m/z	Inten 💦 🖓 Quick Decon.	Apply Apply		
Discovery Mode 🙁		Scan #1		1 C H A N G I	EME 1
ort Data +	50000000		L_1	100	
ral Summing		Т		[1-8] Mono : 889.3058 Average	889.9656 Mostabundant : 889.3058
Data +	40000000		-8	30	
rgeted Mode 🔶	30000000 -	1	-6	60	
ita •	Interaily				
earch	20000000 -		-4	40	
ation +					
	10000000		-2	20	
Reporting 😵					
	86.8117	2086.8117	4086.8117 6086.811		
		m/z			
	Spectrum (m/z) Distribution Chromatogram Spectrum (D	Ja)			
		Prev Next			
	Mass List Peak List Peak Search Result Protein Search	n Result			
	# Monoisotopic Mass Abundance	Most Abundant m/z Charge Score Most Abundant M	lass Ion Y Theoretical(Da) Елог(Da) Елог(ррл	n) Form	
	5				
	No MonoMW Intensity	y m/z Charge Fit	MostMW PeakNom Abundance	Modifications Fragm	ent Ion Types Tolerance
				Position: PTM: Custom	
				0.0000	
					x y z Copy Table Image UVPD Ions Modify Sequence
	ExpID Checked Active	Filename Mode Status	Condition		Internal Fragments Display
			NORMAL	Show Assignments	Filter C/Z Proline Normal V
	▶ 1	H:\MASH\09012 ful Raw Data View	NORMAL		
		H:\MASH\U9012 full Haw Data View	NURMAL	Min	Frag Size: 3 🗘 # Heat Colors 2 🗘
		H-MASHUGUIZ full Raw Data View		Min	
	> <b>1</b> S	S H:MASHUSUIZ. NJ Haw Data New	NORMAL	Mir	
		S H-MAGH-MODIZ_ Nal H-Ann Uada Wern		Mer	

Once the deconvolution is finished, the mass list will populate, as seen below.

🖷 MASH Native - Top-down/native Protec	omics								– 🗆 🗙
File Edit View Tools Help									
<i></i>		Quick Decon. Scan# 1	lot Type Auto	• 🕵 m/z	Inten				
Discovery Mode 🙁			Scan #1						1 CHANGEME 1
Import Data	50000000-								
Process Data			1						[1-8] Mono : 889.3058 Average : 889.9656 Mostabundant : 889.3058
Spectral Summing Validate Data	40000000-							-80	
Validate Deta									
Targeted Mode (*)	30000000							-60	
Import Data	Intensity		1						
Deconvolution	울 200000000-							-40	
Database Search	20000000							-40	
Characterization +	100000000000000000000000000000000000000		I I					0000	
	10000000							-20	
Data Reporting 📎			11						
	0	2086.6	117	<u></u>	86.8117	2 (S	Ē	086.8117	
	00.0117	2000.0	112	40 m/z	00.0117		0	000.0117	
		6							
	Spectrum (m/z) Distribution Chromatog		Franke Marin Makala	Consult Televil					
	Mass List Peak List Peak Search Res	ete 📳 Fit Score   Prev Next   🖻 Dup	licate New Match	search lons					
	# Monoisotopic Mass		Charge Score	Most Abundant Mas	is lon 🝸	Theoretical(Da)	Error(Da)	Error(ppm)	
	▶ 1	445251328.0000 2909.8005	10 97.05	<ul> <li>A state of the state of the state of the state</li> </ul>				C	
	2 🗌 29068.9019	428451648.0000 2909.7016	10 84.245					C	
	3 29070.8816	423919264.0000 2909.8996	10 71.715				_	c	
	4 29067.9104 5 29069.8966	401306112.0000 2909.6024 290139520.0000 2645.3652	10 60.85% 11 98.12%	C. Contraction of Contract				c c~	
	<	230133320.0000   2043.3032	30.12	20007.0402		1		>	Modifications Fragment Ion Types Tolerance
	No MonoMW	Intensity m/z	Charge	Fit Mos	stMW	PeakNorm	Abundance		Position: Error 2.1
									PTM: Custom
									Clear Selected Modificatio
									Clear All Modifications UVPD Ions Display
	ExpID Checked		Mode		indition				Show Assignments Filter C/2 Proline Normal V
	> 1	H:\MASH\0901	2 ful	Raw Data View NO	RMAL				Min Frag Size: 3 + #Heat Colors 2 0
Workflow Tasks Parameters	Experiment List Log Book								Sequence Table Theoretical Data
Vorkflow Finished in 0.00:16:39 days.hour	s:minutes:seconds								🔒 Human Proteomics

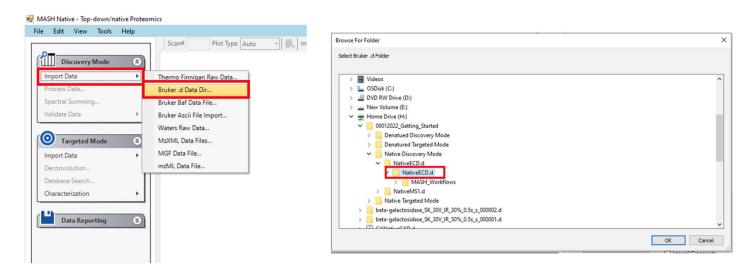
## CHECKPOINT: Make sure your results match the photo above.

in the *Results View* to alter the theoretical fit and click the "Apply" button to save your adjustments, like the example below.



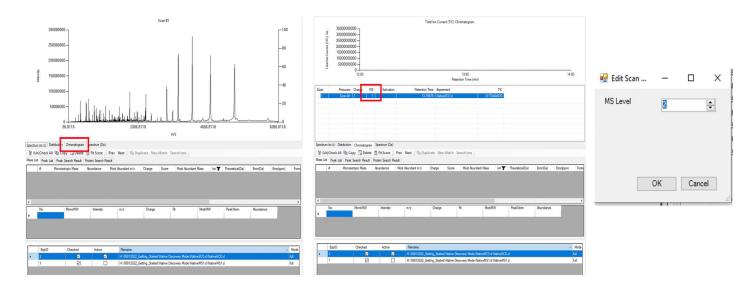
### Step 3: Import MS/MS data

Use the Import Data function in the Discovery Mode to import the sample Bruker .d 'Native ECD' folder.



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

To complete the database search, the user must label the ECD data as an MS2 file. 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."



Now, click the Process Data menu item in the Discovery Mode section to open the Process Wizard dialog. The Process Wizard will setup the series of tasks for selecting, configuring, and running the deconvolution and database search algorithms. Select the eTHRASH deconvolution and TopPIC database search algorithms. Next, click on the Browse button and select the sample 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.

🛃 Process Wizard	-		×	🖳 Proce	ss Wizard				-	o ×
(Preise)				Basic Ad	dvanced					
Basic Advanced				General	Deconvolution S	earch				
Do Deconvolution				Scan Ra	ange 1		•	All Scans	Current Scan	
Deconvolution Method										-
O TopFD O MsDeconv ● eTHRASH O pr	ParseTD O ProMEX O FLA	SHDecony	*	Working	Directory H:\09	012022_Getting_Sta	rted'\Native Discov	ery Mode \NativeECD.	d\M/ Browse	B
O UniDec O UniDec Interactive										
					Scan	m/z	Charge	Monoisotopic	RT	File
Do Database Search				•	1	2909.8005	10	29087.9322	13.79975	Native8
Database Search Algorithm										
O MS-Align+  TopPIC O pTop O MSPa	thFinderT									
Fasta Data File H:\09012022_Getting_Started\Native Dis	covery Mode\BovineUniProt.fas	rowse								
				<						>
Reset to Defaults Save Normal										
				Activati	on Method ETD.	ECD 🗸	Generate	Parameter File	Automati	c Import
	Start	Cancel		Clea	an Up Working Dir	Java Max Heap S	Size (GB) 12	•		
				1						
	🖳 Process Wizard						x u			
	Basic Advanced									
	General Deconvolution Search									
	eTHRASH Parameters									
	Centroiding Parameters		D	econvolution Parameters						
	Signal/Noise 3		÷ D	elete Intensity Threshold	10	\$				
			P	eptide Min Back Ratio	3	\$				
	Peak Background Ratio 3		÷ M	fin Charge	5	\$				
			м	lax Charge	15	-				
	Thresholded Data?		м	lax Mass	30000	\$				
			N	lum Peaks For Shoulder	1	\$				
	Peak Fit Type		м	fin m/z	0	÷				
	○ Apex	zian	м	fax m/z	10000	\$				
			м	fin Fit (%)	80	÷				
				Check All Pattens Agai	nst Charge 1	Long .				
							.8			

### Step 5: Review results

Your deconvolution and search results will be automatically imported into MASH Native if the run was successful.

#### November 12, 2022

#### MASH Native App

	🛇 🔘 🔘 🔗 Apply 🛛 🗞 Quick Decon. Scan# 1 Plot Type Auto 🔹 🔍 m/z Inten	
Discovery Mode 🔹	Scan #1	1 C H A N G E M E 1
nport Data	300000000	-100
ocess Data pectral Summing		[1-8] Mono: 889.3058 Average: 889.9656 Mostabundant: 889.3058
lidate Data	25000000 -	-80
	20000000-	
Targeted Mode 🛞	2	-60
convolution	15000000-	
tabase Search	10000000-	-40
aracterization +		-20
Data Reporting 😵		-2U
Data nepoting	2 11 14 14 14 14 14 14 14 14 14 14 14 14	······
	86.8118 2086.8118 4086.8118	6086.8118
	m/z	
	Spectrum (m/z) Distribution Chromatogram Spectrum (Da) W (Un)Check All Ba Copy ⊒ Delete II Fit Score Prev Next IIIa Duplicate New Match Search Ions	
	i i unicheck All da Copy Delete rit score : Prev Next : un Duplicate New Match Search lons Mass List Peak List Peak Search Result Protein Search Result	
	# Monoisotopic Mass Abundance Most Abundant m/z Charge Score Most Abundant Mass Ion Y Theoretical(Da)	Error(Da) Error(ppm) ^
	1         29072.7410         252024064.0000         3637.3552         8         97.18%         29090.7876           2         2         29071.7383         276307072.0000         3637.2298         8         86.23%         29089.7849	c c
	3 29071.5793 198970160.0000 3233.1874 9 98.74% 29089.6259	c
	4 29073.7989 216741392.0000 4156.9847 7 95.96% 29091.8455	C_
	No MonoMW Intensity m/z Charge Pit MostMW PeakNom	Abundance Modifications Fragment Ion Types Tolerance
		Position:
		PIM: Lustom Y a D C C C B C
		PTM: Custom
		0.0000 ♀
	Ego[2 Checked Active Filename Mode Status Condition	0.0000      Clear Selected Modification     Clear All Modifications
	▶ 1 I I H:\MASH\09012 full Raw Data View NORMAL	Cocco     Ces Selected Modificatio     Ces Selected Modificatio     UVPD lors     Modify Sequence
		0.000     Clear Selected Modification     Clear All Modification     Clear All Modification     Clear All Modifications     Drote     Drote
	▶ 1 I I H:\MASH\09012 full Raw Data View NORMAL	0.000     Clear Selected Modification     Clear All Modification     Clear All Modification     Clear All Modifications     Drote     Drote
	▶ 1 I I H:\MASH\09012 full Raw Data View NORMAL	0.000     Clear Selected Modification     Clear All Modification     Clear All Modification     Clear All Modifications     Drote     Drote

You can find your deconvolution results in the Mass List and your search results in the Protein Search Results List.

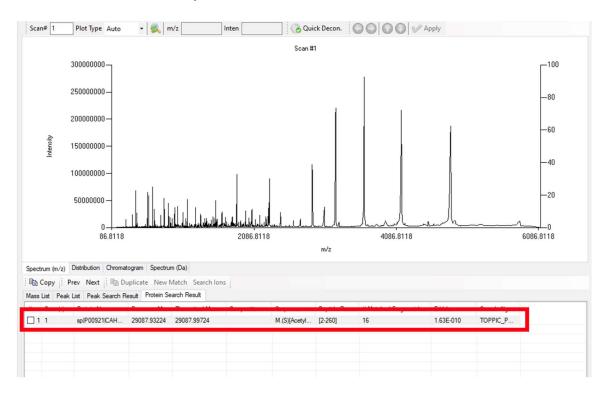
## CHECKPOINT: Make sure your results match the photo above.

Your saved deconvolution results will be in your working directory under MASH\_Workflows. You may wish to select "Overwrite" to avoid viewing duplicated results.

Open your saved results fil	le(s)						>
🗕 🔶 👻 🛧 🛄 > This	PC > Home Drive (H:) > MASH Fi	les > NativeDiscoveryModeTutorial	(Processed)		ٽ ~	🔎 Search NativeDiscove	eryMod
Organize 🔻 New folder						855 <b>-</b> (	. 0
Aug. 22 version 🔺	Name	Date modified	Туре	Size			
NativeComplex	MASH_Workflows	8/17/2022 7:53 PM	File folder				
Supplemental	NativeECD.d	8/17/2022 7:53 PM	File folder				
🔙 Box	NativeMS1.d	8/17/2022 7:53 PM	File folder				
OneDrive - UW-M							
Desktop							
🗄 Documents							
Microsoft Teams							
E Pictures							
This PC							
3D Objects							
E Desktop							
😫 Documents							
🕹 Downloads							
👌 Music							
E Pictures							
Videos							
🏪 OSDisk (C:)							
🛫 Home Drive (H:) 🗸							
File nar	me:				v	XML files(*.XML)	~
						Open 👻 Ca	ncel

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 the protein CAH2 can be found at <u>https://www.uniprot.org/uniprotkb/P00921/entry</u>. With this, we can modify the protein sequence and modifications. In the Modify Sequence Dialog in the bottom right corner, remove the following characters so that the sequence matches the image below: "M.", "(", ")", "[ACETYL]", "(", ")" "[81.3146]", and finally ".". Click OK to display these notation changes in the Sequence Table.

🖳 Modify Sequence			×
SHHWGYGKHNGPEHWHKDFPIANGERQSPVDIDTKAVV HSFNVEYDDSQDKAVLKDGPLTGTYRLVQFHFHWGSSD KYGDFGTAAQQPDGLAVVGVFLKVGDANPALQKVLDALD WTYPGSLTTPPLLESVTWIVLKEPISVSSQQMLKFRTLNF VRGFPK	DQGSEHTVDRKKY SIKTKGKSTDFPNF	AAELHL	VHWNT PNVLD
Reset	OK	1	ancel

You can also manually type in (or copy/paste) the following sequence:

SHHWGYGKHNGPEHWHKDFPIANGERQSPVDIDTKAVVQDPALKPLALVYGEATSRRMVNNGHSFN VEYDDSQDKAVLKDGPLTGTYRLVQFHFHWGSSDDQGSEHTVDRKKYAAELHLVHWNTKYGDFGT AAQQPDGLAVVGVFLKVGDANPALQKVLDALDSIKTKGKSTDFPNFDPGSLLPNVLDYWTYPGSLTT PPLLESVTWIVLKEPISVSSQQMLKFRTLNFNAEGEPELLMLANWRPAQPLKNRQVRGFPK

Next, add the following modifications in the Configuration Panel under the Tools menu. Add -H+ and Zn(II)-H+ in the table with their respective masses shown in the figure below. Then, using the PTM drop down box under the sequence table, add an acetylation to the N-terminal serine, -H+ to H93, and Zn(II)-H+ to H95. Check the c and z-ions and click the "Show Assignments" button. Additionally, users can restrict the mass tolerance and use the internal fragment options to improve sequence coverage.

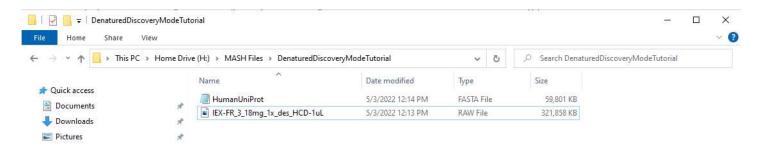
Name	MonoMass	AvgMass	
Pentoses (Ara, Ri	132.0423	132.1161	
Deoxyhexoses (F	146.0579	146.143	
Hexosamines (Ga	161.0688	161.1577	
Hexoses (Fru, Ga	162.0528	162.1424	
Lipoic acid (amid	188.033	188.3147	
N-acetylhexosami	203.0794	203.195	
S-Carboxymethyl	58.0055	58.0367	
0 Calbonyananon	07.0210	07.002	
- H+	-1.0078	-1.008	
Zn(II) - H+	62.920247	64.4	

#### Analysis of MS/MS denatured top-down mass spectrometry data:

For the analysis of denatured mass spectrometry data, download the folder containing the denatured tutorial folder and save them to a new local folder as seen below.

Download the folder below on the Google Drive for the denatured MS/MS tutorial

• <u>Denatured Discovery Mode Tutorial</u> folder which contains the IEX-FR\_3\_18mg\_1X-des-HCD-1uL raw file and HumanUniProt.fasta file.



Proceed with the following steps for analysis of denatured MS/MS mass spectrometry data to identify an unknown protein from a human extract.

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

Use the Import Data function in the Discovery Mode to import the sample Thermo raw data file.

🛃 MASH Native - Top-down/native Proteo	mics	🐱 Select your Experiment raw data file 🛛 🗙
File Edit View Tools Help	Scan# 1 Plot Type Auto	← → ✓ ↑ 📮 « Home Drive (H) > 00012022_Getting_Started > Denatued Discovery Mode > 🗸 δ 🖉 Search Denatued Discovery _
Discovery Mode 🙁		Organize   New folder  III   Date modified  Type  Size
Import Data  Process Data Spectral Summing Validate Data	Thermo Finnigan Raw Data Bruker .d Data Dir Bruker Baf Data File Bruker Ascii File Import	This PC     This PC     Destrop     Destrop     Destrop     Destrop     Destrop     Destrop     Destrop     Destrop     Destrop     Petures     Petures
Targeted Mode 😒 Import Data > Deconvolution	Waters Raw Data MzXML Data Files MGF Data File mzML Data File	Constraints Const
Database Search	20000000	Open 🔻 Cancel

Step 2: Spectral deconvolution and database search

Click the Process Data menu item in the Discovery Mode section to open the Process Wizard dialog. The Process Wizard will setup 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. Next, click on the Browse button and select the sample FASTA file to load into MASH. 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 and an "Activation File" popup window will appear, please click "Run Anyway" to begin the search.

Process Wizard		-		×
Basic Advanced				
Do Deconvolution				
Deconvolution Method				
TopFD OMsDeconv OeTHRASH OpParseTD (     UniDec Outpactive		() FLAS	SHDeconv	r.
☑ Do Database Search				
Database Search Algorithm				
O MS-Align+  TopPIC O pTop O MSPathFinderT				
	overy Mode∖H	luma Br	owse,	l
Fasta Data File H:\MASH\09012022_Getting_Started\Denatured Disco Reset to Defaults Save Normal		uma Br	rowse Cancel	
Reset to Defaults Save Normal	St	lart	Cancel	×

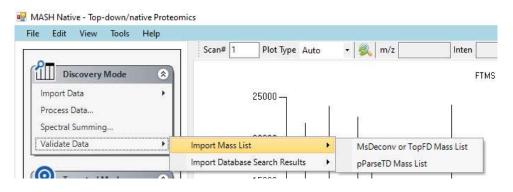
Starting your search will open the Workflow Manager window and the terminal window. The terminal window is where TopFD and TopPIC are running. You can minimize this window but closing this window will stop your workflow. You can close and open the Workflow Manager window as you wish, and you can use the Workflow Manager to check on the status of your run. When your run 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.

#### Step 3: Import 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.

## CHECKPOINT: Make sure your results match the photo above.

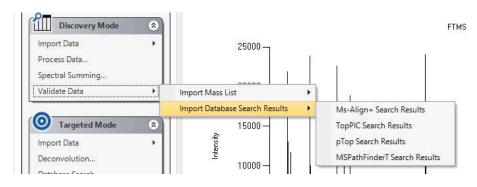
To import the deconvolution mass list for detailed analysis of the data set, select the Validate Data function in the Discover Mode panel, and click on Import Mass List.



Your saved deconvolution results will be in your working directory under MASH\_Workflows. Double click on the folder and select the file IEX-FR\_3\_18mg\_1x\_des\_HCD-1uL\_ms1.msalign as seen below. The loaded results will populate the Mass List. You may wish to select "Overwrite" to avoid viewing duplicated results.

ganize 🔻 New folder				EE ▼ 🔲 (
• OneDrive	Name	Date modified	Type	Size
	i_html	9/17/2022 8:36 PM	File folder	
This PC	input_topfd_file	9/16/2022 4:32 PM	File folder	
3D Objects	input_topfd_html	9/16/2022 3:59 PM	File folder	
🔜 Desktop	proteins.fasta_idx	9/16/2022 5:01 PM	File folder	
Documents	IEX-FR_3_18mg_1x_des_HCD-1uL_ms1.msalign	9/16/2022 4:20 PM	MSALIGN File	2,711 KB
	IEX-FR_3_18mg_1x_des_HCD-1uL_ms2.msalign	9/16/2022 4:22 PM	MSALIGN File	2,065 KB
Music	📄 input.msalign	9/16/2022 4:22 PM	MSALIGN File	2,065 KB
E Pictures	input_topfd_ms2.msalign	9/16/2022 4:22 PM	MSALIGN File	2,065 KB
Videos				
🏪 OSDisk (C:)				
New Volume (E:)				
🛫 Home Drive (H:)				
🛖 Lab Drive (P:)				
Network				
	ne: IEX-FR_3_18mg_1x_des_HCD-1uL_ms1.msalign		~ MSALIGN	files(*.MSALIGN)

To import the database search results for detailed analysis of the data set, select the Validate Data function in the Discovery Mode panel, and click on Import Database Search Results, and select TopPIC Search Results.



Navigate to your working directory and select the following file to import: IEX-FR\_3\_18mg\_1x\_des\_HCD-1uL\_ms2\_toppic\_prsm.tsv. You may wish to select "Overwrite" to avoid viewing duplicated results. Search results will be loaded into the Protein Search Results Table.

🚽 👻 🕇 📙 « MAS	H > IEX-FR_3_18mg_1x_des_HCD-1uL_20220916_155820_TopFD	_TopPIC_1_1734 > V	ල 🔎 Sea	rch IEX-FR_3_18mg_1x_d.
organize 🔻 🛛 New folder				
OneDrive	Name	Date modified	Туре	Size
💶 This PC	i_html	9/17/2022 8:36 PM	File folder	
Contractor and a	input_topfd_file	9/16/2022 4:32 PM	File folder	
3D Objects	input_topfd_html	9/16/2022 3:59 PM	File folder	
Desktop	📙 proteins.fasta_idx	9/16/2022 5:01 PM	File folder	
Documents	IEX-FR_3_18mg_1x_des_HCD-1uL_ms2_toppic_prsm	9/17/2022 8:28 PM	TSV File	250 KB
🕹 Downloads 🛛 🔤	input_toppic_proteoform	9/17/2022 8:33 PM	TSV File	83 KB
h Music	input_toppic_proteoform_single	9/17/2022 8:33 PM	TSV File	45 KB
Pictures	input_toppic_prsm	9/17/2022 8:28 PM	TSV File	250 KB
Videos	input_toppic_prsm_single	9/17/2022 8:27 PM	TSV File	112 KB
	Workflow_Timings	9/17/2022 8:38 PM	TSV File	1 KB
SDisk (C:)				
🛖 New Volume (E:)				
🛫 Home Drive (H:)				
🛖 Lab Drive (P:)				
Network				
File par	ne: IEX-FR_3_18mg_1x_des_HCD-1uL_ms2_toppic_prsm		CSV files	(*.CSV);TSV files(*.TSV)

#### 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.

## **Targeted Mode:**

Targeted Mode is intended for targeted mass spectrometry data analysis. Detailed instructions on using this mode can be found in the User Manual at *Section 4 Targeted Mode: Protein Identification and Characterization from MS/MS Data*.

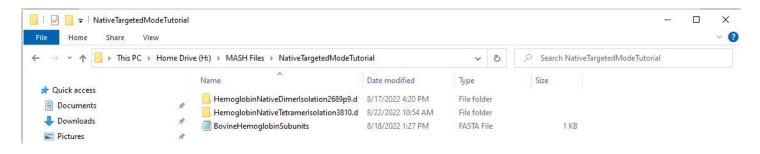
#### Analysis of native top-down mass spectrometry data:

For the analysis of native mass spectrometry data, download the folder containing the native tutorial files and save them to a new local folder as seen below.

Download the folder below on the Google Drive for the native targeted tutorial:

• Native Targeted Mode Tutorial folder which contains the Bruker

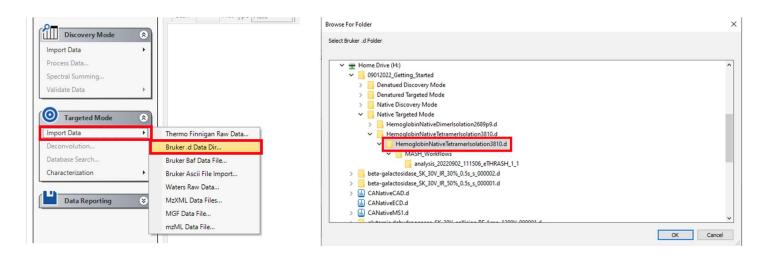
HemoglobinNativeDimerIsolation2689p9.d and HemoglobinNativeTetramerIsolation3810.d data files as well as the BovineHemoglobineSubunits.fasta file.



Proceed with the following steps for analysis of native MS/MS mass spectrometry data to characterize hemoglobin from a bovine extract.

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

Use the Import Data function in Targeted Mode to import the HemoglobinNativeTetramerIsolation3810.d data file.



#### Step 2: Deconvolution

This data shown here is generated from the isolation of the 17+ charge state for the hemoglobin tetramer and ECD fragmentation applied. You can deconvolute the fragments and identify them in targeted mode. Under Targeted Mode, Process Wizard will guide users through the deconvolution step. Select the eTHRASH deconvolution method and input the parameters below in the and click the Start button. Enter 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.

Process Wizard – 🗆 🗙	R Process Wizard	-	×
Basic Advanced General Deconvolution Scan Range 1 1 1 2 All Scans Current Scan	Basic Advanced General Deconvolution eTHRASH Parameters		
Working Directory       H:\MASH Files\NativeComplexHemoglobin\MASH_Workflows\Hemoglobin       Browse         Scan       m/z       Charge       Monoisotopic       RT       File         1       3810.00       17       64752.8763       23.07303333333       Hemoglobin	Centrolding Parameters     Deconvolution Parameters       Signal/Noise     3       Peak Background Ratio     3       Peak Background Ratio     3       Image: Convolution Parameters     10       Peptide Min Back Ratio     3       Min Charge     1       Max Charge     20       Max Mass     70000	4 4 4 4 4 4 4	
Activation Method Construction Method Const	Peak Rt Type     1       Apex     Quadratic       Image: Constraint of the second se		

Step 3: Annotate results

When the fragment ions are verified, users can use the protein sequences from the alpha and beta hemoglobin subunits to identify fragments. Open the BovineHemoglobinSubunits FASTA file and copy the sequence for the alpha chain and paste into the modify sequence box in the bottom right corner. Users can then select c and z-ions and "Show Assignments."

	Scan# 1 Plot Type Auto • 🔍 m/z Inten 🕜 Quick Decon. 📿 📿 📿 🏈	
Discovery Mode 🔹	Fit: 99% mono: 15659.2071596966 charge: 6.00	1 MVLSAADIKGINVKAAWGKIVGGI
Data 🔸	25000000	21 HAAE <sup>1</sup> YGAEALER <sup>1</sup> MFLSFPTT
Summing Data •	2000000	41 K T Y F P H F D L S H G S A Q V K G H G
geted Mode 🔹		_
•		61 A K V A A A L T K A V E H L D D L P G A 81 I S E I S D L H A H K I R V D P V N E K
,		
<b> </b>		101 L L S H S L L V T L A S H L P S D F T P
		121 A V H A S L D K F L A N V S T V L T S K
	2607.7370 2609.7370 2611.7370 2613.7370 2615.7370 m/x	141 Y R
	₩ (Un)Check All Ba Copy         Delete         Fit Score         Prev Next         Pab Duplicate         New Match         Search Ions           Mass Lat         Peak Search Rext         Preters Search Rext         Preters Search Rext         Nonestopic Mass         Nonestopic Ma	[1-142] Mono : 15174 9444 Average : 15184.3926 Mostabundant : 15183.9655 Total# c 5 z 2 % Residue Clavavage : 567% Random P-Value: 1.7809e-002
	3         1569 2072         2251818:0000         2612 3785         6         98.44.x         19668 2307           4         758.4295         22330810 0000         759.4363         1         98.00%         758.4295	
	No MonoMW Internaty m/z Charge Pt MontMW P 22518166 2 <	Modifications         Fragment fon Types         Tolerance           President         □
	EppID         Checked         Active         Flename         Mode         Status         Condition           1         Image: Condition of the state of t	Litter # Information     Internal Fragments     Display       Show Assignments     Filter C/2 Proline     Nomal       Min Frag Size     3 (a)     # Heat Colors (2 (b))

Similarly, users can identify fragments from the beta chain. Additionally, users can improve sequenece 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.

	sc Scanf 1 Plot Type: Auto • 🔍 m/z Inten 🦳 🌀 Quick Decon. 💿 💽 💽 🔿 "	»
Discovery Mode 🔹	Fit: 99% mono: 15659.2071596366 charge: 6.00	$ \  \  \  \  \  \  \  \  \  \  \  \  \ $
Indexery Mode     S     Import Data     Process Data Spectral Summing Validate Data     Process Data      Process Data     Deconvolution Database Search Characterization     Data Reporting     S	FE 593: morol. 19693.20/196808 charge 6.00         259000000-         19000000-         19000000-         19000000-         0         5000000-         0         25000000-         0         0         25000000-         0         0         25000000-         0         0         25000000-         0         25000000-         0         25000000-         0         25000000-         0         25000000-         0         25000000-         0         25000000-         0         25000000-         0         25000000-         0         25000000-         0         25000000-         10000000-         10000000-         10000000-         10000000-         10000000-         10000000-         10000000-         10000000-         10000000-         100000000-         100000000- <th>[1-145] Monc. 15944.309 Average: 15954.4233 Mostabundant: 15954.3334 <math display="block">[1-145] Monc. 15949-006</math> <math display="block">[2 + 1695 + 26 + 26 + 26 + 26 + 26 + 26 + 26 + 2</math></th>	[1-145] Monc. 15944.309 Average: 15954.4233 Mostabundant: 15954.3334 $[1-145] Monc. 15949-006$ $[2 + 1695 + 26 + 26 + 26 + 26 + 26 + 26 + 26 + 2$
	No MonoMW Internaty m/z Charge R MoutMW R	Modifications         Fragment for Types           President:
	< <tr>         Exp[D         Orecked         Adive         Rename         Mode         Satura         Condition           ▶         1         Image: Satura and Satura a</tr>	Clear Selected Modifications     □ V □ V □ Z     Ccepy Table Image       □ UVPD lons     □ Modify Sequence       □ UVPD lons     □ Drpby       Show Assignments     □ Filter 02 Position       Min Frag Size (2) ⊕     # Heat Colors 2 ⊕

Save your results and close this experiment.

## 4. Import data for dissociation of hemoglobin subunits

In a similar manner, import the HemoglobinNativeDimerIsolation2689p9.d data file for the isloation and the hemoglobin dimers and monomer dissociation experiment.

		Browse For Folder	×
Discovery Mode       Import Data         Import Data       Process Data         Spectral Summing       Validate Data         Validate Data       Validate Data         Import Data       Import Data         Deconvolution       Database Search         Characterization       Import Data         Data Reporting       Import Data	Thermo Finnigan Raw Data Bruker.d Data Dir Bruker Ascii File Import Waters Raw Data MzXML Data Files MGF Data File mzML Data File	Select Bruker .d Folder	
		ОК	Cancel

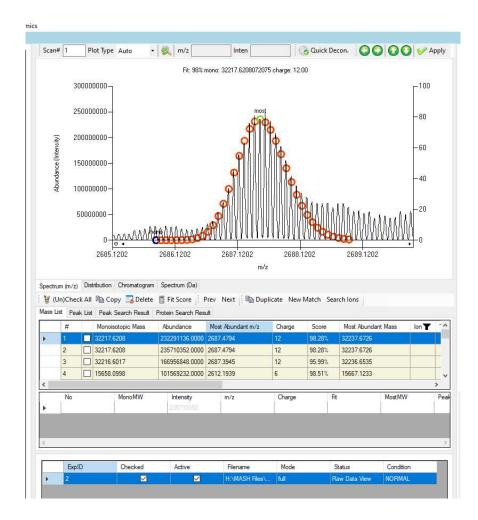
#### 5. Deconvolution

Deconvolute the spectra using the eTHRASH deconvolution algorithm and the advanced parameters shown below.

🖳 Process Wizard			8 <u>—</u>	×
Basic Advanced				
General Deconvolution				
eTHRASH Parameters				
Centroiding Parameters	Deconvolution Parameters			
Signal/Noise 3	Delete Intensity Threshold	10	÷	
	Peptide Min Back Ratio	3	-	
Peak Background Ratio 3	Min Charge	5	<b> </b>	
	Max Charge	15	÷.	
Thresholded Data?	Max Mass	70000	<b>.</b>	
	Num Peaks For Shoulder	1	<b> </b>	
Peak Fit Type	Min m/z	2000	÷	
O Apex O Quadratic      Lorentzian	Max m/z	3000	•	
	Min Fit (%)	75	¢	
	Check All Pattens Against C	Charge 1	Privant)	

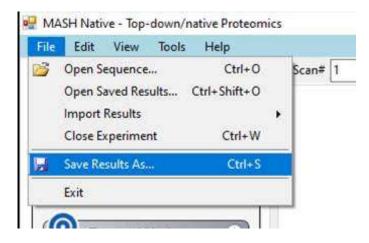
#### 6. Review results

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 above.



					Scan #1						
			0		1	2+	solat	ed Dim	ner	۲ <sup>100</sup>	
	2	50000000 -						Notes			
	2	00000000-					4	1		-80	
eitu	<b>;</b> 1	50000000-	<u> </u>	Ejected	Monomer	s				-60	
Inteneitu				1		1					
	1	00000000		8	1 A	1				- 40	
					7+ 🖤	1					
	3	50000000 -			6+	11				-20	
								1.1			
		0-	<u>-</u>		hand hand had	w.l.p				0	
		162	7.8791	2127.879		791 /z	3127	8791	3627.	8791	
¥ (L	Jn)Chee	Distribution	Chromatogram	Spectrum (Da)	rr Prev Next   🗈 Dug	/z			3627.	8791	
¥ (L	Jn)Chee ist Pea	Distribution ck All 🗈 Co sk List Peak	Chromatogram py 🔁 Delete Search Result	Spectrum (Da)	rr Prev Next ∣i @m Duj sult	/z olicate New	Match Se	arch lons			
₩ (U ass L	Jn)Chee	Distribution tk All 🗈 Co sk List Peak Monois	Chromatogram py 🛃 Delete Search Result notopic Mass	Spectrum (Da)	r Prev Next : Dug ault : Most Abundant m/z	/z		arch Ions Most Abundar		8791	
🥑 (L	Jn)Cheo ist Pea #	Distribution ck All 🗈 Co ak List Peak Monois	Chromatogram py Joelete Search Result totopic Mass 1919	Spectrum (Da)	Prev Next  : I Dup suit Most Abundant m/z 2239.1661	/z plicate New Charge	Match Se Score	arch lons			
₩ (U ass L	Jn)Cheo ist Pea # 7	Distribution ck All Co ak List Peak Monois	Chromatogram py Delete Search Result totopic Mass 1919 3368	Spectrum (Da) Fit Score Protein Search Re Abundance 55079844.0000	Prev Next  : I Dup suit Most Abundant m/z 2239.1661	v/z blicate New Charge 7	Match Se Score 98.62%	Most Abundar 15667.1154			
¥ (L	Jn)Cheo ist Pea # 7 8	Distribution ck All Concerning C	Chromatogram py Delete Search Result totopic Mass 1919 3368 1972	Spectrum (Da) Fit Score Protein Search Re Abundance 55079844.0000 38877136.0000	rrev Next Dup Dup suit Most Abundant m/z 2239.1661 2509.6668	v/z blicate New Charge 7 6	Match Se Score 98.62% 97.58%	Most Abundar 15657.1154 15051.9602			
₩ (L ass L	Jn)Cheo ist Pea # 7 8 9 10	Distribution ck All I Co ak List Peak Monois 15653. 15042. 16559.	Chromatogram py Delete Search Result totopic Mass 1919 3368 1972 1259	Spectrum (Da) Fit Score Frotein Search Re Abundance 55079644.0000 28877136.0000 27848772.0000 31039116.0000	m Prev Next 📴 Dup ault Most Abundant m/z 2509.6668 2569.668 2368.0815 2875.7774	/z olicate New Charge 7 6 7 11	Match Se Score 98.62% 97.58% 98.53% 96.92%	Most Abundar 15667.1154 15051.9602 16569.5234 31622.4775	nt Mass		>
🤘 (U ass L	Jn)Cheo ist Pea # 7 8 9	Distribution ck All I Co ak List Peak Monois 15653. 15042. 16559.	Chromatogram py Delete Search Result totopic Mass 1919 3368 1972	Spectrum (Da) Fit Score Frotein Search Re Abundance 55079844.0000 38877136.0000 27848772.0000 31039116.0000 Intensity	rr Prev Next : @ Dup suit Most Abundant m/z 2239.1661 2509.6668 2368.0815	Vz Dicate New Charge 7 6 7	Match Se Score 98.62% 97.58% 98.53% 96.92%	Most Abundar 15667.1154 15051.9602 16569.5234			
₩ (U ass L	Jn)Cheo ist Pea # 7 8 9 10	Distribution ck All I Co ak List Peak Monois 15653. 15042. 16559.	Chromatogram py Delete Search Result totopic Mass 1919 3368 1972 1259	Spectrum (Da) Fit Score Frotein Search Re Abundance 55079644.0000 28877136.0000 27848772.0000 31039116.0000	m Prev Next 📴 Dup ault Most Abundant m/z 2509.6668 2569.668 2368.0815 2875.7774	/z olicate New Charge 7 6 7 11	Match Se Score 98.62% 97.58% 98.53% 96.92%	Most Abundar 15667.1154 15051.9602 16569.5234 31622.4775	nt Mass		>
₩ (U ass L	Jn)Cheo ist Pea # 7 8 9 10	Distribution ck All I Co ak List Peak Monois 15653. 15042. 16559.	Chromatogram py Delete Search Result totopic Mass 1919 3368 1972 1259	Spectrum (Da) Fit Score Frotein Search Re Abundance 55079844.0000 38877136.0000 27848772.0000 31039116.0000 Intensity	m Prev Next 📴 Dup ault Most Abundant m/z 2509.6668 2569.668 2368.0815 2875.7774	/z olicate New Charge 7 6 7 11	Match Se Score 98.62% 97.58% 98.53% 96.92%	Most Abundar 15667.1154 15051.9602 16569.5234 31622.4775	nt Mass		>
₩ (U ass L	Jn)Cheo ist Pea # 7 8 9 10	Distribution ck All I Co ak List Peak Monois 15653. 15042. 16559.	Chromatogram py Delete Search Result totopic Mass 1919 3368 1972 1259	Spectrum (Da) Fit Score Frotein Search Re Abundance 55079844.0000 38877136.0000 27848772.0000 31039116.0000 Intensity	m Prev Next 📴 Dup ault Most Abundant m/z 2509.6668 2569.668 2368.0815 2875.7774	/z olicate New Charge 7 6 7 11	Match Se Score 98.62% 97.58% 98.53% 96.92%	Most Abundar 15667.1154 15051.9602 16569.5234 31622.4775	nt Mass		>
₩ (U ass L	Jn)Cheo ist Pea # 7 8 9 10	Distribution kk All @ Co kk Ust Peak 5559. 31602.	Chromatogram py Delete Search Result totopic Mass 1919 3368 1972 1259	Spectrum (Da) Fit Score Frotein Search Re Abundance 55079844.0000 38877136.0000 27848772.0000 31039116.0000 Intensity	m Prev Next 📴 Dup ault Most Abundant m/z 2509.6668 2569.668 2368.0815 2875.7774	/z olicate New Charge 7 6 7 11	Match Se Score 98.62% 97.58% 98.53% 96.92%	Most Abundar 15667.1154 15051.9602 16569.5234 31622.4775	nt Mass		>

From this data, users can identify and annotate the ejected monomers. 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.

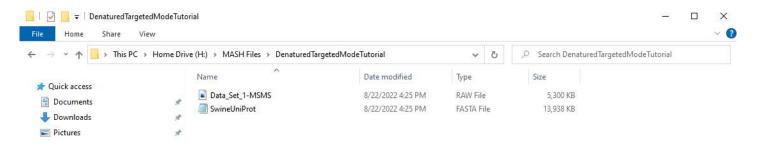


### Analysis of denatured top-down mass spectrometry data:

For the analysis of denatured mass spectrometry data, download the folder containing the native tutorial files and save them to a new local folder as seen below.

Download the folder below on the Google Drive for the native targeted tutorial:

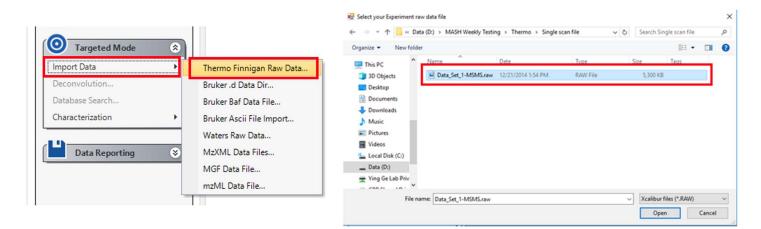
• <u>Denatured Targeted Mode Tutorial</u> folder which contains the Thermo Data\_Set\_1-MSMS.raw file and the SwineUniprot.fasta file.



Proceed with the following steps for analysis of denatured MS/MS mass spectrometry data for protein characterization.

#### Step 1: Import data

Use the Import Data function in Targeted Mode to import the sample Thermo data file.



Step 2: Deconvolution

Under the Targeted Mode, Process Wizard will guide users through the deconvolution step. Select the TopFD deconvolution method and click the Start button. You have the option to set more advanced 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.

			🖳 Pr	rocess Wizard					_		×		
			Basic	Advanced									
				o Deconvoluti	on								
			•			HRASH	⊖ pPar	seTD O ProMEX	⊖ FL	ASHDeconv			
			Re	set to Defaults	Save Normal			S	tart	Cancel			
lass Li	st Pea	ık List	Peak Search Result	Protein Search R	esult								
	#		Monoisotopic Mass	Abundance	Most Abundant m/z	Charge	Score	Most Abundant Mass	lon <b>T</b>	Theoretical(Da	) Error(Da)	Error(ppm)	F ^
•	1		12339.3900	117.5993	1029.8745	12		12346.4064					С
	2		12338.3777	105.9986	950.6530	13		12345.3942					С
	3		12339.4035	51.8926	1123.4091	11		12346.4199					С
	4		12322.3681	42.4147	1028.4560	12		12329.3846					С
	5		22024.5185	119.7045	882.5493	25		22038.5514					С
	6		12294.3720	29.4114	1026.1230	12		12301.3885					С
	7		12294.3782	29.0924	1119.3159	11		12301.3947					С
	8		12322.3807	32.8464	1121.8616	11		12329.3972					c ~

#### Step 3: Database search

Following deconvolution, users can use the selected ion list for database search via the Process Wizard. In the Targeted Mode panel, select the Database Search function, which will open the Process Wizard. 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.

🖳 Process Wizard	_		×
Basic Advanced			
☑ Do Database Search			
Database Search Algorithm			
O MS-Align+      TopPIC O pTop O MSPathFinderT			
Fasta Data File       H:\09012022_Getting_Started\Denatured Targeted Mode\Swinet         Reset to Defaults       Save Normal	JniProt.1 B	rowse Cancel	

This file requires Advanced Parameters to be set. Please click on the Advanced Parameters tab. Change the charge from 0 to 14 and hit Enter, 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.

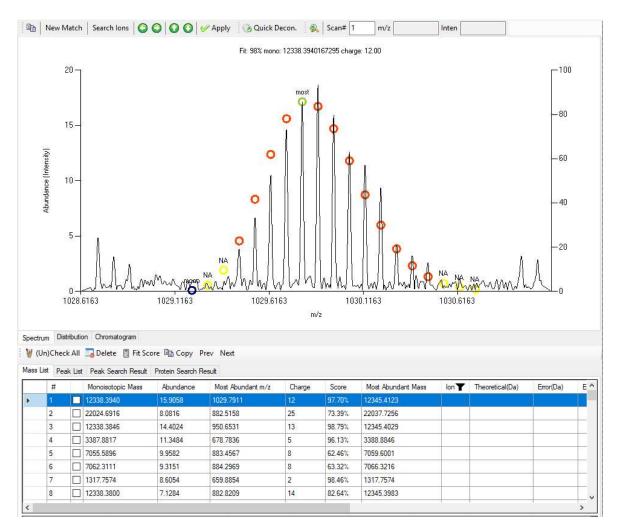
e ta_Set_1-MSMS.raw
e
ta_Set_1-MSMS.raw

As before, the Workflow Manager and the 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 TopPIC. When the search is finished, the terminal window will automatically close, the Log Book will show the line "Workflow Finished" and the Workflow Manager will move the workflow into the "Completed Workflows" Tab.

Your search results will be automatically imported into MASH Native. Look under the Protein Search Results list to view your results. If you would like to manually import your search results, please read Step 3: Import Results from the Discovery Mode and import your search results file with the Targeted Mode Characterization button.

#### 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*.



Your TopPIC search results contain the protein sequence of the identified protein, which will be automatically imported. Go to the Protein Search Result list and click on the first result.

Mass	List Peak	List Peak Search Result Prote	n Search Result						_
#	Scan(s) 1	Protein Name gi 311258932 ref XP_0031278	Precursor M 12389.8981	Theoretical Mass 12390.06813068	Composition	Sequence M.PFFDVQ	Peptide Ran [2-106]	# Matched Fragment Io 28	E \ 4.3
<									>

The Sequence Table will display this sequence. The protein sequence may need to be modified to remove some of TopPIC's notation. To do this, please click on the Modify Sequence button under the Sequence Table.

Modifications Position:	Fragment Ion Types	Tolerance Error 2.1
PTM: Custom	∕abc	Da      PPM
0.0000		Copy Table Image
Clear All Modifications	UVPD lons	Modify Sequence
Show Assignments	☐ Internal Fragments ☐ Filter C/Z Proline Min Frag Size: 3 ♀	Normal V # Heat Colors 2

In the Modify Sequence Dialog, remove the following characters so that the sequence matches the image below: "M.", "(", ")" "[48.68]", and finally ".". Click OK to display these notation changes in the Sequence Table.

You can also manually type in (or copy/paste) the following sequence:

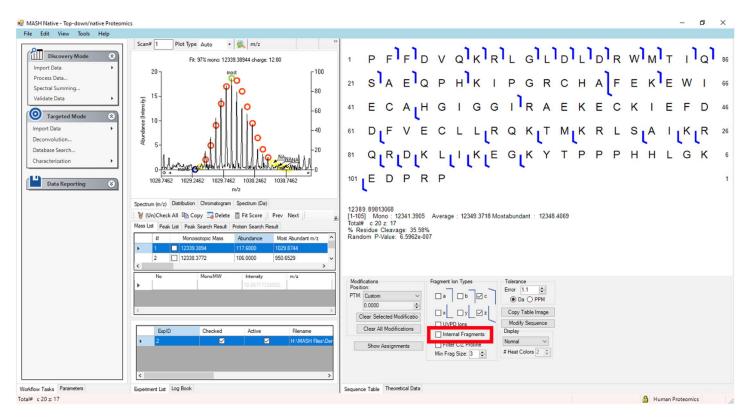
PFFDVQKRLGLDLDRWMTIQSAEQPHKIPGRCHAFEKEWIECAHGIGGIRAEKECKIEFDDFVECLLRQ KTMKRLSAIKRQRDKLIKEGKYTPPPHHLGKEDPRP

	Мо	dify	Sec	uen	ice											12	-				×
	DVQ													ECAH	IGIO	GIF	AEK	ECł	(IEFI	DDF	VECI
											[		ок			Re	set		C	Cance	el
quenc	e Table	F	F	Р	v	0	ĸ	P	ĩ	G		D	1	D	P	\٨/	M	т	1	Q	86
21			Ē	Q			ĸ		P		R			A		E	ĸ	Ē	w	4	66
41	-		A			ï	G		ï		A		ĸ		c	ĸ		E	F	D	46
	_	-			-		-													_	
61	D	F	V	E	C	L	L		Q				ĸ		L	s	A		ĸ	R	26
	0		D	K	L	I	K	E	G	ĸ	Y	1	Р	Р	Р	Н	Н	L	G	ĸ	6
81	Q		P	-																	1

For manually pasting protein sequence, click on Paste Your Sequence and the Paste Sequence Information window will be opened. Paste the target sequence and click OK and the pasted sequence will be displayed in Sequence Table.

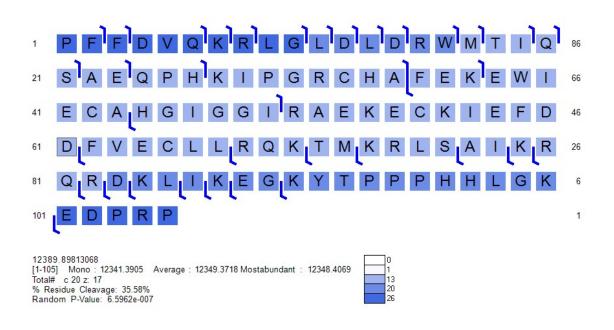
O Targeted Mode	۲	40	
Import Data Deconvolution Database Search		Intensity	
Characterization	•	Import Mass List Import Database Search Results Paste Your Sequence	• •

In the Sequence Table, check the c and z ion types and click on the Show Assignments button to view the fragmentation for that peptide sequence. If users want to view internal fragmentation, select the Internal Fragments checkbox before clicking the Show Assignments button, which matches monoisotopic masses with internal fragment ion types (e.g., cz). To search for internal ions such as c and z, the c and z ions should be checked on the interface. 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.



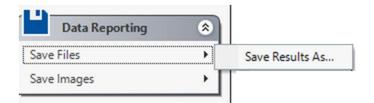
Step 5: Heat map

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. You can change the number of heat colors for the map, we have shown the map with 5 heat colors below.



#### Step 6: Save results

Saving annotated results is only available for Targeted Mode. In the Data Reporting panel click Save Files and then Save Results As. Your Mass List and Protein Search Results list will be saved as an XML file. Save this XML file in the same directory as your raw data file.



To load a saved XML file into MASH Native, in the top left-hand corner go to File and click Open Saved Results.

File	Edit View To	ols Help
	New Experiment	Ctrl+N
	Open Sequence	Ctrl+O
	Open Saved Results	Ctrl+Shift+O
	Import Results	•
	Close Experiment	Ctrl+W
	Save Results As	Ctrl+S
	Exit	

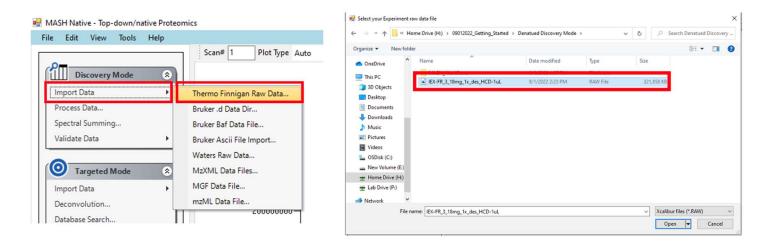
## **Spectral Summing:**

Sample Files: Use either of the files for discovery mode above.

- Sample Thermo raw data: <u>IEX-FR 3 18mg 1x des HCD-1uL.raw</u>.
- Sample 32-bit mzXML file: <u>IEX-FR 3 18mg 1x des HCD-1uL.mzXML</u>

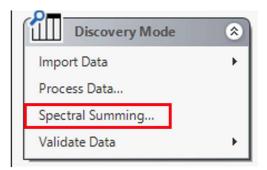
#### Step 1: Import data

Use the Import Data function in the Discovery Mode to import the sample Thermo raw data file from the Denatured Discovery Mode Tutorial.



Step 2: Open spectral summing tool

Under the Discovery Mode panel, select Spectral Summing"



The user can also open the spectral summing tool from the "Tools" Tab.

	-	
File Edit View	Tools	Help
	W	orkflow Manager
P	Q	uick Decon. Parameters
Discovery	Pa	aste Sequence
Import Data	S	pectral Summing
Process Data	C	onfiguration
Validate Data		ew Log File
		pen AppData

------ .

#### *Step 3: Set parameters*

In the Spectra Summing Tool dialog, select "Proteowizard" for both the Selection Strategy and Summing Algorithm choices. Observe the parameters for running the Proteowizard selection in the parameters tab.

Click OK and the summing process should be started as a workflow and can be accessed using the workflow manager. Should take approximately 2-3 minutes to complete.

🖶 Spectra Summing Tool	-		×	💀 Spectra Summing Tool 🛛 — 🗖	×
Main Parameters				Main Parameters	
Selection Strategy All Scans Select Ranges Proteowizard Summing Algorithm UniDec-Interpolate UniDec-Integrate Vendor Proteowizard	MS1 Handling Skip Copy Sum			Proteowizard Selection Parameters         Precursor Tolerance       0.0500         Scan Time Tolerance       10.00         Ion Mobility Tolerance       0.0100         Range Selection Parameters       Scan Start         Scan Stop       1734	
	OK	Ca	ancel	ОК	Cancel

#### Step 4: Observe/process new experiment

Once the summing algorithm completes, a new .mzML has been generated and automatically loaded into MASH Native. You can observe the summed and original version of the experiment under the "Experiment List" in the bottom center of the MASH Native application. You can then process this experiment as you would in the Discovery Mode step mentioned above.

	ExpID	Checked	Active	Filename	Mode	Status	Condition
•	1			E:\TestingFiles\Thermo\IEX-FR_3_18mg_1x_des_HCD-1uL	full	Raw Data View	NORMAL
	2			E:\TestingFiles\Thermo\MASH_Workflows\IEX-FR_3_18m	full	Raw Data View	NORMAL
		and the second sec					

The mzML file of the summed scans is placed in a workflow folder, which you can find the location through the workflow manager. See User Manual for more details on using the workflow system and for more details on the other summing routes that MASH Native provides.

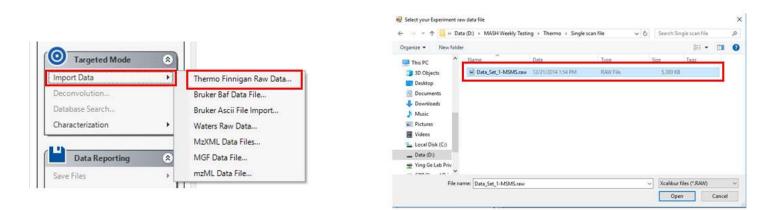
## **Spectral Deconvolution using UniDec:**

Sample Files: Use either of the targeted mode files above:

- Sample Thermo raw data: <u>Data\_Set\_1-MSMS.raw</u>.
- Sample 32-bit mzXML file: <u>Data\_Set\_1-MSMS.mzXML</u>.

#### Step 1: Import data

Use the Import Data function in Targeted Mode to import the sample Thermo data file from the Denatured Targeted Mode Tutorial.



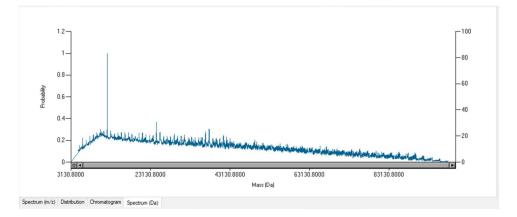
#### Step 2: Run UniDec deconvolution

Under the Targeted Mode, Process Wizard will guide users through the deconvolution step. Select the UniDec deconvolution method and click the Start button.

🛃 Process Wizard			-		×
Basic Advanced					
Do Deconvolution     Deconvolution Method     O TopFD O MsDeconv O eTHRASH     O UniDec O UniDec Interactive	○ pParseTD		() FLA	SHDecor	IV
Reset to Defaults Save Normal		St	art	Cance	el
184 3997	h84 399	/			1184

#### Step 3: Observe Mass List and deconvoluted spectrum results

Upon completion the Mass List will be populated using UniDec's results as previously described in the targeted mode spectrum above. The extracted deconvoluted spectrum can be observed by selecting the "Spectrum (Da)" tab in the spectrum view panel.



The zoom tools should work on the spectrum. Please see the "<u>Best Practices for UniDec Deconvolution in</u> <u>MASH</u>" file in the documentation for running UniDec on your particular type of targeted mode data.

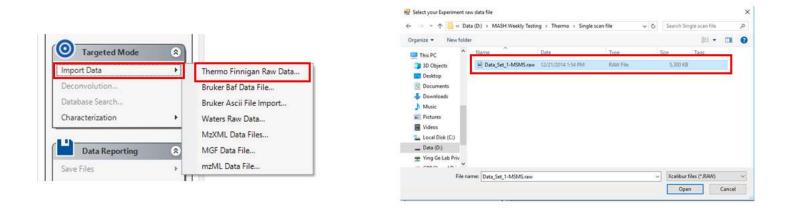
## **PTM Analysis using UniDec:**

Sample Files: Use either of the targeted mode files below:

- Sample Thermo raw data: <u>Data\_Set\_1-MSMS.raw</u>.
- Sample 32-bit mzXML file: <u>Data\_Set\_1-MSMS.mzXML</u>.

#### Step 1: Import data

Use the Import Data function in Targeted Mode to import the sample Thermo data file.



Step 2: Run UniDec deconvolution

Under the Targeted Mode, Process Wizard will guide users through the deconvolution step. Select the UniDec deconvolution method, under advanced, change the charge to 14, and change the activation method to ETD.

Pro	cess Wizard				- 0	
sic	Advanced					
enera	Deconvolution					
Scan	Range 1	<b>↓</b> 1	Ţ	All Scans	Current Scan	
Worki	ing Directory H:\	MASH\Testing Proto	col\Themo Single S	can\MASH_Workflows	Dat Browse	
	Scan	m/z	Charge	Monoisotopic	RT	File
•	1	886.00	14	12389.8981	70.957576666666	Data
•	1	000.00	17	12303.0301	70.337370000000	Dau
P	ļ	000.00		2200.0001	10.33737666666	Data
<	1	000.00		22303.0301	10.33737666666	
	ation Method			e Parameter File	Automatic Imp	

Then under Deconvolution  $\rightarrow$  Deconvolution, change the max charge to 30 and bin sampling to 0.1, seen in the figure below, then go back to the Basic Tab and press "Start".

Process Wizard	>
Basic Advanced	
General Deconvolution	
Data Processing Deconvolution Peak Processing	
Min Charge 1	Max Charge 30
Mass Start (Da) 5000.00	Mass End (Da) 500000.00
Native Charge Offset Min -100	Native Charge Offset Max 100
Max # Iterations 100	Peak FWHM (Th)
Charge Smooth Width	Mass Smooth Width
◯ None  ● Log  ◯ Gaussian 1 😫	● None ◯ Log ◯ Gaussian 1 📫
m/z To Mass Transform Method	Peak Shape
O Integrate O Interpolate O Smart	● Gaussian () Lorentzian () Gaussian / Lorentzian
Point Smooth Width	Bin Sampling (Every Da) 0.1
Isotope Mode	Charge Scaling (Orbi Mode)
Off O Mono O Average	

#### Step 3: Import Sequence

• Sequence: <u>NDUFS5 Sequence</u>.

Copy the NDUFS5 protein sequence from the sequence .txt file. In MASH, click on "Modify Sequence" in the bottom left corner and the "Paste Sequence Information" window will open. Paste the sequence and click "OK" and the sequence will be displayed in the Sequence Table. In the table, it will calculate the most abundant mass of the sequence.

1	М	Ρ	F	F	D	V	Q	K	R	L	G	L	D	L	D	R	W	Μ	т	I	87
21	Q	S	Α	Е	Q	Ρ	н	K	Т	Ρ	G	R	С	н	Α	F	Е	K	Е	W	67
41	Т	Е	С	Α	н	G	I	G	G	I	R	Α	Е	κ	Е	С	к	Ļ	Е	F	47
61	D	D	F	V	Е	С	L	L	R	Q	K	Т	Μ	ĸ	R	L	S	Α	I	K	27
81	R	Q	R	D	к	L	T	K	Е	G	K	Y	т	Ρ	Ρ	Ρ	н	Н	L	G	7
101	K	Е	D	Ρ	R	Ρ															1
[1-106	6] Mo	ono : 1	2472.4	310	Averag	je : 124	480.570	Mos	tabund	dant : 1	2479.4	1474	1								

Step 4: Review mass list

Upon completion the mass list will be populated using UniDec's results. Now, manually review the deconvolution and isotopic fit mass list to identify a high score charge assignment of most abundant masses that is closest to the Most Abundant Mass of the sequence.

#	Monoisotopic Mass	Abundance	Most Abundant m/z	Charge	Score	Most Abundant Mass
56	0.0000	0.0930	1123.4073	11	94.88%	12346.4004
57	0.0000	0.2607	1029.8740	12	97.48%	12346.4004
58	0.0000	0.2002	950.7304	13	87.20%	12346.4004

Step 5: Post-Translational Modifications

• NDUFS5 UniProt: <u>https://www.uniprot.org/uniprotkb/F1SV23/entry</u>

Using the UniProt link, identify the post-translational modifications (PTMs) for the NDUFS5 protein. From here, the user will see that there are two disulfide bonds in the sequence.

PTM/Proces	sing								
Features Showing features for di	sulfide bond <sup>i</sup> .								
1 10	20	30	40 50	60	70	80	90	100	106
			EWIECAHGIGGIRAEK	ECKIEFDDFVE	CLLRQKTMKR	LSAIKRQRDK	LIKEGKYTPPP	HHLGKEI	PRP
TYPE Select v	ID	POSITION(S)	DESCRIPTION						
Disulfide bond		33-66	1 Automatic An	notation Comb	ined Sources		BLAST 🕯	🗟 Add	
Disulfide bond		43-56	1 Automatic An	notation Comb	ined Sources		BLAST 🕯	🔂 Add	

Under Tools  $\rightarrow$  Configuration  $\rightarrow$  Modifications, add a disulfide bond into the table with the respective masses shown in the figure below.

Name	MonoMass	AvgMass		
Deoxyhexoses (F	146.0579	146.143 161.1577		
Hexosamines (Ga	161.0688			
Hexoses (Fru, Ga	162.0528	162.1424		
Lipoic acid (amid	188.033	188.3147 203.195		
N-acetylhexosami	203.0794			
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		

Then using the PTM drop down box under the sequence table, add a disulfide to C33, C43, C56, and C66, then select "Show Assignments." This will change the most abundant mass in the sequence table. If this mass matches the most abundant mass in the mass list, then these are the correct PTMs on your sample protein. If the masses do not match, then the combination of PTMs is not correct for your sample.

M P F F D V Q K R L G L D L D R W M T 1 1 87 Q S A E Q P H K I P G R C H A F E K E W 21 67 IECAHGIGGIRAEKECKI E F 41 47 DDFVECLLRQKTMKRLSAI K 61 27 RQRDKLIKEGKYTPPPHHLG 81 7 KEDPRP 101 1 [1-106] Mono : 12468.3998 Average : 12476.538 Total# Mostabundant : 12475.4162 % Residue Cleavage: 0.00% Random P-Value: 1.2806e-016 C33: -H+ C43: -H+ C56: -H+ C66: -H+

Because the masses do not match, the user must identify the PTMs based on the difference in masses. The difference between the mass in the mass list, and the raw sequence mass is 133.047 Da. From this, we can identify PTMs as the removal of the initiator methionine and one disulfide bond.

Under "Modify Sequence" delete M1, then in the PTM drop down box, add a disulfide bond to C43 and C56, and select "Show Assignments." Now the most abundant mass in the sequence table is extremely close to the mass in the mass list and can be assumed as the modified sequence.

