AlphaViz User Guide

Developed by: *Eugenia Voytik, Sander Willems*.

This step-by-step guide helps you to get started with our software AlphaViz.

Table of Contents

Description

Software tools such as MaxQuant or DIA-NN identify and quantify high amounts of proteins and peptides. After downstream processing, for instance in Perseus, MSstats or the Clinical Knowledge Graph, differentially expressed proteins are possible candidates for biomarker discovery. Instead of rushing into conclusions, AlphaViz allows to link these identifications with the original unprocessed LC-TIMS-Q-TOF data and easily assess their individual quality or the overall quality whole samples.

Installation

To extend the tool's usability, AlphaViz is implemented in several modes: a user-friendly browser-based graphical user interface (GUI), thoroughly covered in this manual, a stand-alone command-line interface and a well-documented Python package, which are described on the GitHub repository (https://github.com/MannLabs/alphaviz).

Windows

Prerequisites: **Windows 10** (a system update might be necessary in case older versions do not work). If you install AlphaViz for all users, you might need admin privileges to run it (right-click on the AlphaViz logo on your desktop and select "Run as administrator").

Important: To prevent installation errors on Windows, we recommend uninstalling any previous AlphaViz version before installing a new one.

1. Download the latest release for Windows (alphaviz_gui_installer_windows.exe) from the GitHub repository and open the .exe file.
2. In the appeared "Select Setup Install Mode" dialog we suggest to select "Install for me only (recommended)" option.
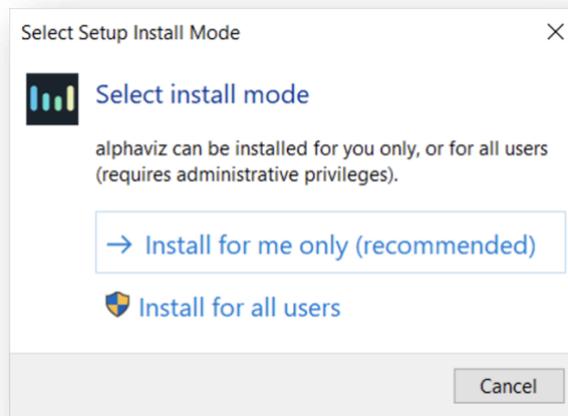
3. In the appearing "Setup – alphaviz version X.X.X" dialog window read and accept the License Agreement and press the "Next" button.
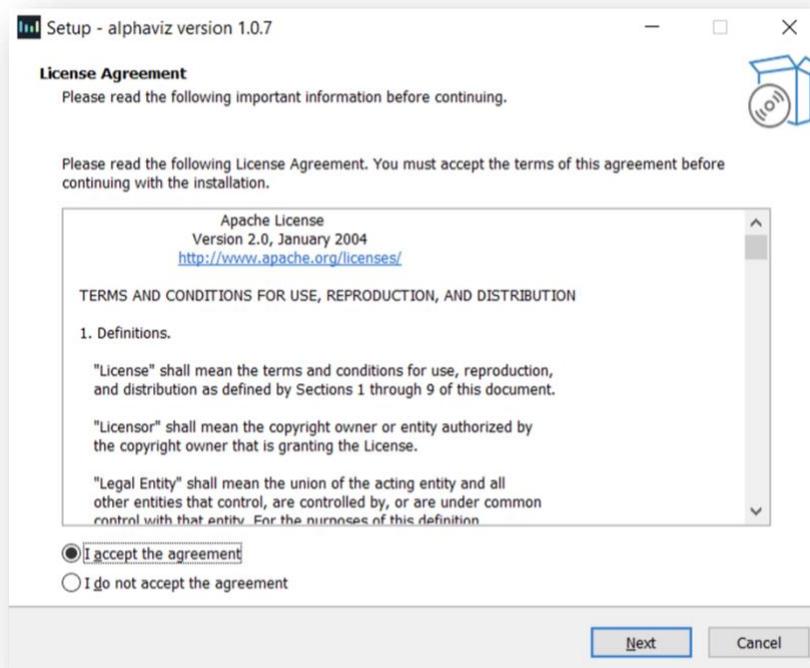
4. Select the destination location for the installation of AlphaViz software (the entire package is about 1.5 GB in size) (the size of the tool is increased due to the integration of deep learning predictions) and press the "Next" button.
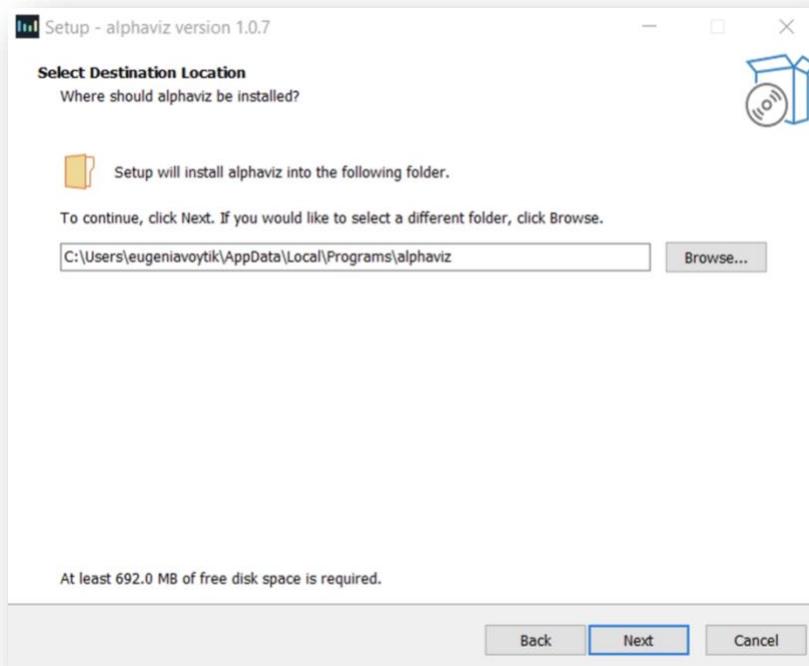
5. In the next dialog window mark the "Create a desktop shortcut" check box and press the "Next" button.
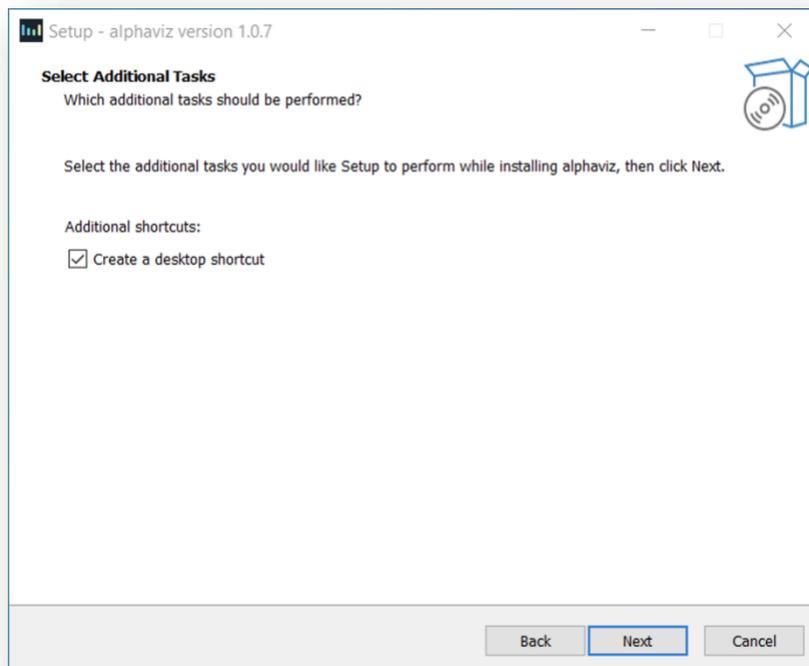
6. Check the setting and if everything is correct, press "Install" button. You may go back to change some settings using the "Back" button or "Cancel" the installation.
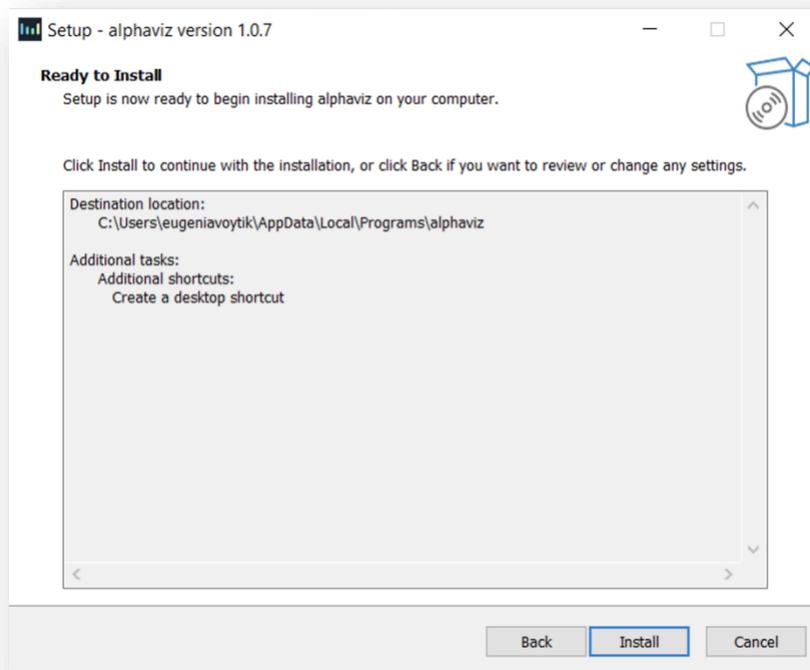
7. Wait till the installation process is finished and with the marked "Launch alphaviz" check box press the "Finish" button.
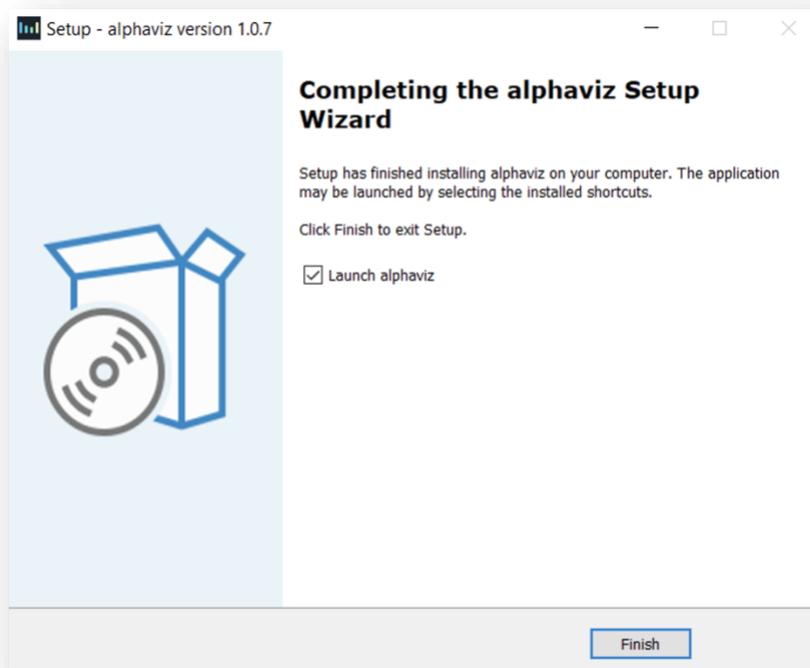
8. If a "Windows Security Alert" dialog window appears, press the "Allow access" button that will prevent the Windows Defender Firewall from blocking the AlphaViz tool on your PC.

MacOS

Prerequisites: at least **macOS Big Sur (11) or higher** (a system update might be necessary in case older versions do not work).

IMPORTANT WARNING: Since AlphaViz uses AlphaTims to read LC-TIMS-Q-TOF data from Bruker's timsTOF pro instrument (Bruker Daltonik), some calibration functions for it are provided by Bruker as libraries and are only available on Windows and Linux. Therefore, to avoid any problems with MS2 spectra quality assessment please use .hdf files into which .d folders can be converted using the AlphaTims's CLI on Windows or Linux machines as described in the AlphaTims GUI manual or in the AlphaTims CLI manual.

1. Download the latest release for macOS (alphaviz_gui_installer_macos.pkg) from the GitHub repository and open the .pkg file. If you receive a message that the file cannot be opened because it is from an unidentified developer, you can close the message by clicking "OK", go to the "Security & Privacy" section of the "System Preferences" menu and under the "General" tab press the "Open Anyway" button for the file "alphaviz_gui_installer_macos.pkg".

2. Click "Continue" on the appearing "Install AlphaViz X.X.X" dialog window.
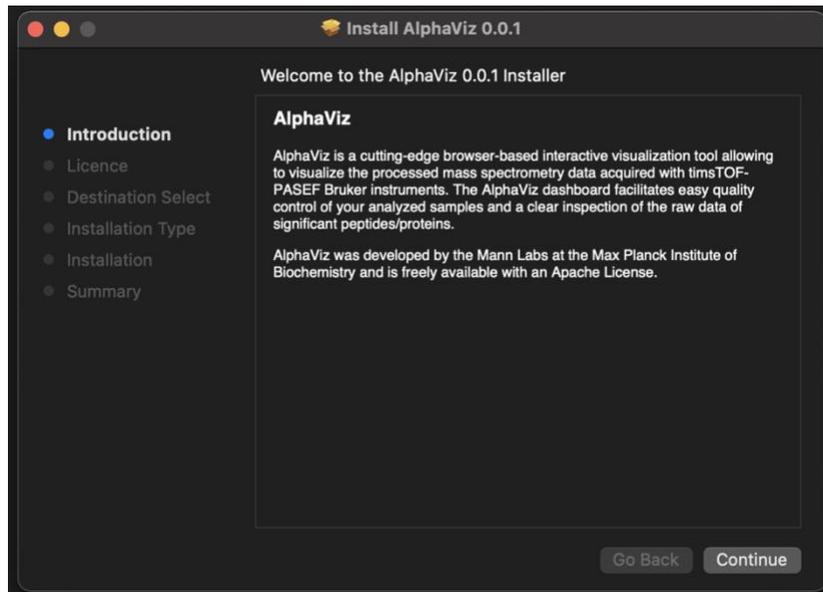
3. In the **License** section, the Software License Agreement (Apache License) will be shown for you. To continue the installation, press "Continue" button and in the appeared pop-up window you need to agree with the license.
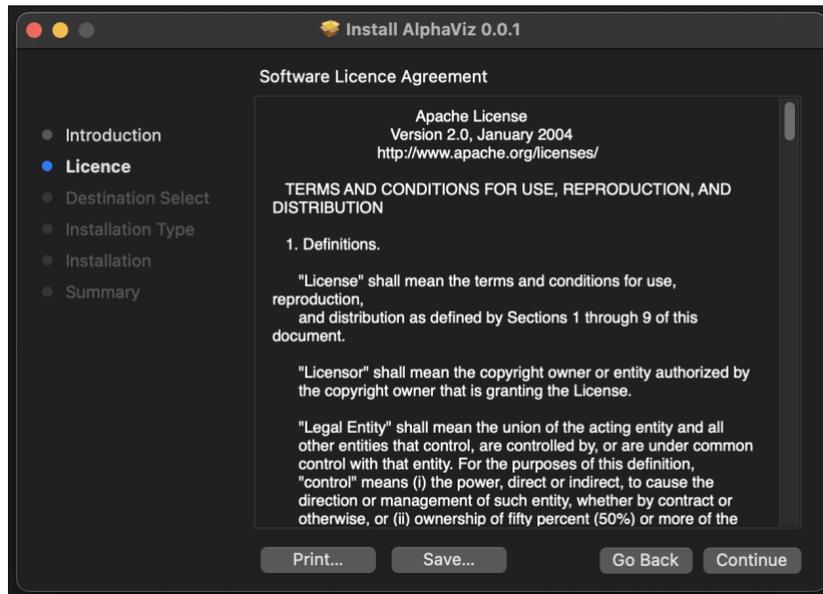
4. Press "Install" to start the installation (the entire package is about 1.8 GB in size) (the size of the tool is increased due to the integration of deep learning predictions). This might take a few minutes.
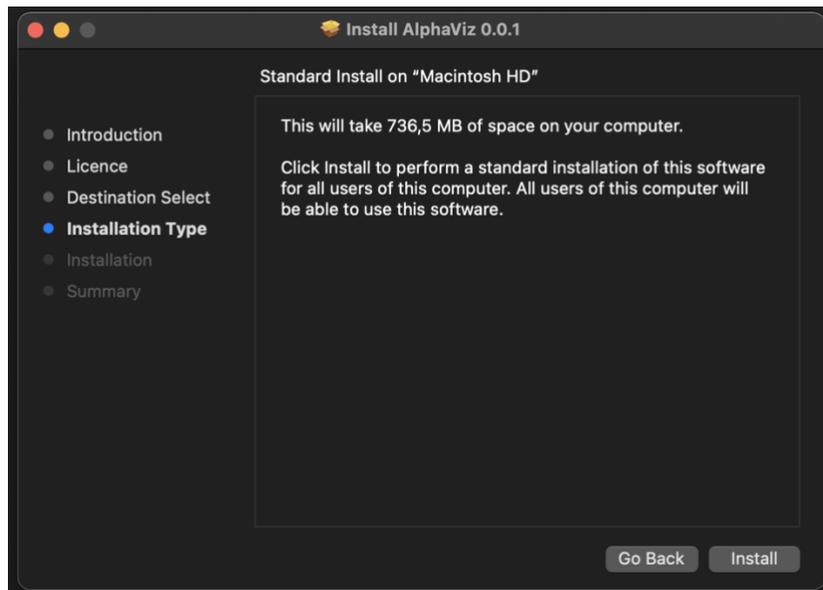

5. Click "Close" to quit the installation menu. AlphaViz is now available in the applications folder on your MacOS.
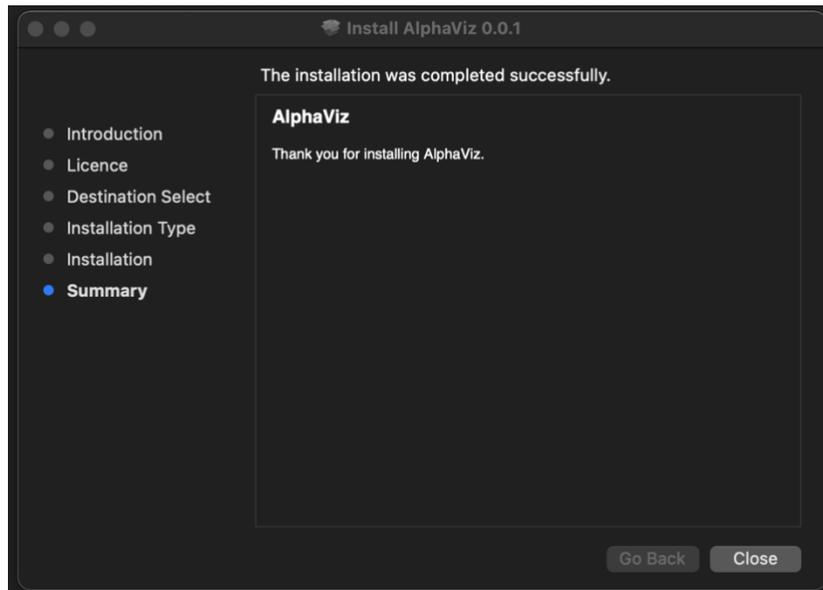
Linux

For Linux, AlphaViz can be installed as a Debian package. Please note that by using AlphaViz, you accept the terms of the Apache Licence and third-party licences!

1. Download the latest release for Linux (alphaviz_gui_installer_linux.deb) from the GitHub repository.
2. Run the installer either by double clicking on it, or by executing the command <sudo dpkg -i alphaviz_gui_installer_linux.deb> (copy everything between <>).

How to use AlphaViz
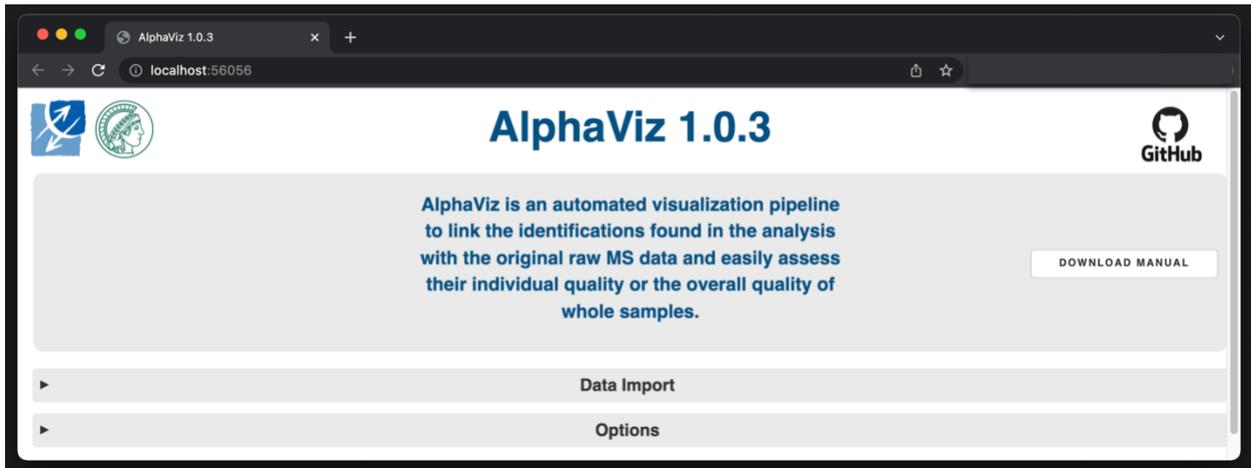
After launching AlphaViz, firstly a terminal window will be automatically opened containing the tool's information (in the background) as well as a new browser tab called "AlphaViz X.X.X" in your default browser with the "http://localhost:XXXXX" URL. To close the tool, just close the opened "AlphaViz X.X.X" tab or press "Ctrl + c" in the running terminal window.

For optimal performance and correct visualization, we recommend using Google Chrome or Mozilla Firefox browsers. Please note that an internet connection is not required to work with AlphaViz on your local machine, unless the tool checks for an updated version of AlphaViz on GitHub. If the newest version exists, you will get a notification about it in the terminal, as well as an additional button in the GUI (below the "Download manual" button) to download the latest version.

By pressing the "Download manual" button you can also download and check out this GUI manual.

Data import

To get started with AlphaViz, you need to use the "Data Import" panel to load the raw file itself (mandatory step). You may then optionally select the output of any supported software, such as currently MaxQuant or DIA-NN, along with the fasta file that was used to analyse the data, or you may skip this step by using AlphaViz in Targeted mode (see below).

a) Specify the path to the folder with the unprocessed .d or .hdf files in the "Specify the full path to the folder with unprocessed Bruker files:*" field, e.g. *"D:\bruker\21min_HELA_proteomics"* (Windows) or *"/Users/test/bruker/21min_HELA_proteomics"* (MacOS). File paths can easily be copied to the clipboard by "shift key + right mouse" (Windows), "option + command + c" (MacOS) or "ctrl + c" (Linux). **Note that the copied path should not contain quotes at the beginning or end and that this is a local path, i.e. not on a mounted drive.**

b) After completing the previous step, the "Select the raw file:*" field will automatically display all .d folders and .hdf files present in the specified folder and sorted in a natural order. By default, the first file of all present ones is selected (highlighted in grey), but you can click on any other file to activate it.

IMPORTANT! Steps a) and b) are mandatory for the tool to work.

IMPORTANT! Steps c) and d) are optional and can be skipped using AlphaViz in Targeted mode (see below).

c) In the "Specify the full path to the output folder of any supported software analysis tool:" field enter the path to the software output folder, e.g. *"D:\bruker\21min_HELA_proteomics"* (Windows) or *"/Users/test/bruker/21min_HELA_proteomics"* (MacOS).

MaxQuant output: If the output "txt" folder is inside the unprocessed .d folder specified at Step b), this path will be filled in automatically.

d) In the "Specify the full path to the fasta file:" field the path to the fasta file that was used for analysis should be entered. This field can be filled in automatically if the fasta file is within the .d folder specified at Step b).

e) Click the "Activate the prediction" check box if you want to use theoretical spectrum prediction for DDA data or ion mobility/retention time prediction for DIA data.

f) Press the "Load data" button. The loading process is indicated by a progress bar. Once the data has been uploaded, the "Data Import" panel will automatically collapse and new tabs will appear in AlphaViz.

If errors occur at any of the above steps, e.g. files have not been provided in the fields or data cannot be loaded, an error message will appear in the empty space (g) with detailed information.

Settings

To customize the visualization or change the settings of the graphs, refer to the "Settings" panel. Changing the parameters in the "Settings" panel automatically updates the already built plots in which they are applied and is subsequently applied for building new plots.

a) The "Tolerance settings" card allows setting tolerances for m/z (in ppm), ion mobility (in 1/K0) and retention time (in seconds). These values are used to build extracted ion chromatograms (XICs) or mobilograms (for DDA data) and 1D or 2D elution profile plots (for DIA data).

b) The "Heatmap options" card can be used to configure the MS1 and MS2 heatmaps by selecting:

- the x-axis or/and the y-axis labels ("X-axis label" & "Y-axis label" drop-down menus);
- the color scale of the plot containing a list of all color maps available in the Holoviews library which can be viewed here ("Color scale" drop-down menu);
- the background color of the plots ("Background color" drop-down menu);

- the size and color of the precursor sign on the heatmaps ("Precursor target size" integer input field & "Precursor target color" drop-down menu).

The selected color scale and background color are also applied to build 2D elution profile plots for DIA data.

c) In the "Customization options" card the names of qualitative and sequential color scales can be chosen and then applied to sequence coverage plots (coloring peptides of the protein of interest) or to 2D elution profile plots (coloring precursor and its fragments in DIA mode). By default, a qualitative color scale is used for the above-mentioned plots. In case the number of options, e.g. the number of peptides in the sequence coverage plot, exceeds the number of colors in the qualitative scale, the selected sequential color map will be used. Here the size of the saved plots (height, width) and their extension can be specified before the plot is generated.

Interactive tools

All interactive plots in AlphaViz are generated using two Python plotting libraries: Plotly (almost all plots) and Holoviews with Bokeh background (heatmaps only). Each of these libraries provides a number of interactive tools for their plots.

Almost all plots, except for heatmaps, are generated with Plotly and have the following tools:

- The '*Download plot' tool* allows you to download an image of the plot.
- The *'Zoom' tool* allows to click and drag the plot to zoom in and double-click to zoom out completely. To zoom in along one axis only, click and drag near the edges of either axis. Alternatively, to zoom along both axes simultaneously, click and drag near the corners of both axes.
- The *'Pan' tool* allows to click and drag on the plot to pan and double-click to reset the pan. It is also possible to pan along one axis by clicking and dragging from the center of the axis.
- The *'Zoom in/out' tool*s allow to zoom in and out by clicking the button or using the scroll wheel on the mouse and/or a two-finger scroll.
- The *'Reset axis' tool* restores the original plot view.
- The '*Show closest/Compare data on hover' tools* allow you to switch between different hover modes. In the "*Show closest data on hover*" hover mode a single hover label appears for the point directly underneath the cursor. In the "*Compare data on*

hover" hover mode a single hover label appears per trace, for points at the same x value as the cursor.

All heatmap generated using Holoviews:

- 🔍 The *box zoom tool* allows to select a rectangular area to be zoomed in.
- ⊖ The *wheel zoom tool* allows to zoom in and out by scrolling with the mouse scroll wheel. Note that the location of the mouse cursor determines where to scroll. Alse note that if you scroll outside the plot area along one of the axes, you can zoom in or out in one dimension, not two.
- ↻ The *reset tool* restores the original plot view, i.e. the x and y margins are set to the minimum and maximum values of the selected data.
- ▢ The *save tool* allows to save an image of the plot.

Data Visualization

Once the data has been loaded into AlphaViz, three new tabs will appear: "Main View", "Quality Control" and "Targeted Mode". By default, the "Main View" tab is opened (highlighted in white), but you can switch between tabs by clicking on them.

1. "Main View" tab
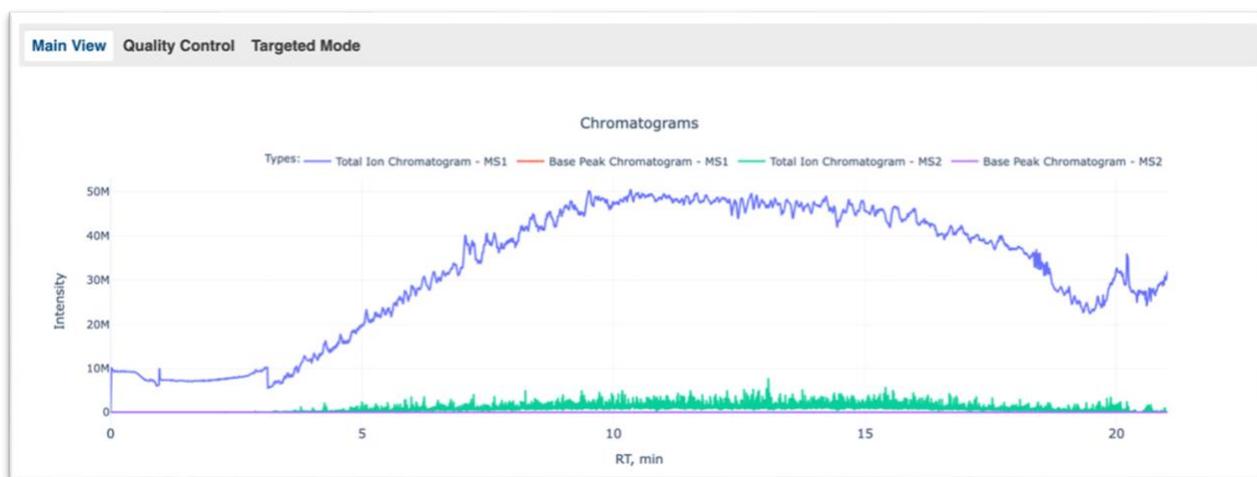
To assess the overall performance of the LC and MS instruments, various types of chromatograms, including total ion chromatograms (TICs), showing the summed intensity of all detected precursor (MS1) / fragment (MS2) ions against the retention time, and base peak chromatograms (BPIs), showing the intensities of the most abundant ions on MS1 and MS2 levels detected over time, are first shown in the "Main View" tab. By clicking on the different chromatogram types in the legend,

they can be (de)selected which will update the scale of the y-axis.

IMPORTANT! If only raw data without any identifications has been loaded (to work in "Targeted Mode" only), nothing else will be shown in the "Main View" tab. In order to subsequently explore the individual identified proteins together with their detected groups of peptides, either the output of DDA data analysis done by the MaxQuant software tool or the output of DIA data analysis done by DIA-NN must be loaded together with the fasta file.

All proteins detected in the sample are presented in the "Proteins table" (a1). Filtering the table and searching for the protein of interest can be done in several ways:

- **Manual search**: The table columns can be sorted by the "▲" or "▼" sign in the header. To navigate through all proteins, switch between the different pages of the table using the buttons (a2).
- **Search by gene name**: Start typing the gene name of the protein of interest in the field "Search the protein by its gene name:" (b). After entering the first three letters you will get a list of all available gene names of proteins to choose from. Once you select a gene name, the "Proteins table" will automatically filter based on that name.
- **Search using a list of predefined gene names**: To filter the "Proteins table" based on a list of proteins, load a list of gene names of pre-selected proteins by pressing the "Choose file" button. In doing so, you can load a .txt file containing the list of gene names one identifier per line as shown in the image.
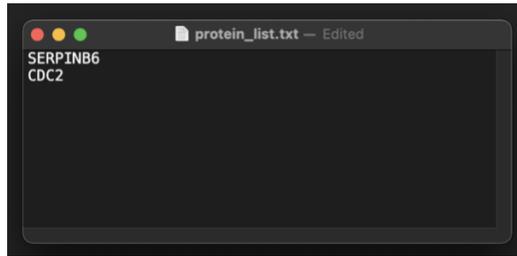
In all the ways described above, to undo any filtering press the "Reset proteins" button (d). In this case the table content will be returned to the original proteins table after data loading.

After finding an interesting protein candidate, click on it in the "Proteins table" (highlighted in blue) to select it. As a result, all identified peptides of this protein will appear in the previously empty "Peptides table" (e1). In addition, a protein coverage diagram (f) will be displayed below the "Peptides table" showing the position of all peptides on the protein sequence and the percentage of its amino acid coverage. If a protein in the "Proteins table" is not selected (by pressing the "Undo" button (b)) the "Peptides table" and the protein coverage diagram will be cleared. If you want to undo the peptide selection and display again the protein coverage diagram for all peptide of the protein of interest, prese "Deselect peptides" button (e2).

To assess the quality of the raw data for each peptide separately, select an individual peptide in the "Peptides table" (e1). In this case, the dashboard layout depends on the type of data that was loaded:

a) DDA data analyzed with MaxQuant software

The following MaxQuant output files must present in the output folder for AlphaViz to work correctly: 'allPeptides.txt', 'msms.txt', 'evidence.txt', 'proteinGroups.txt', 'summary.txt'.
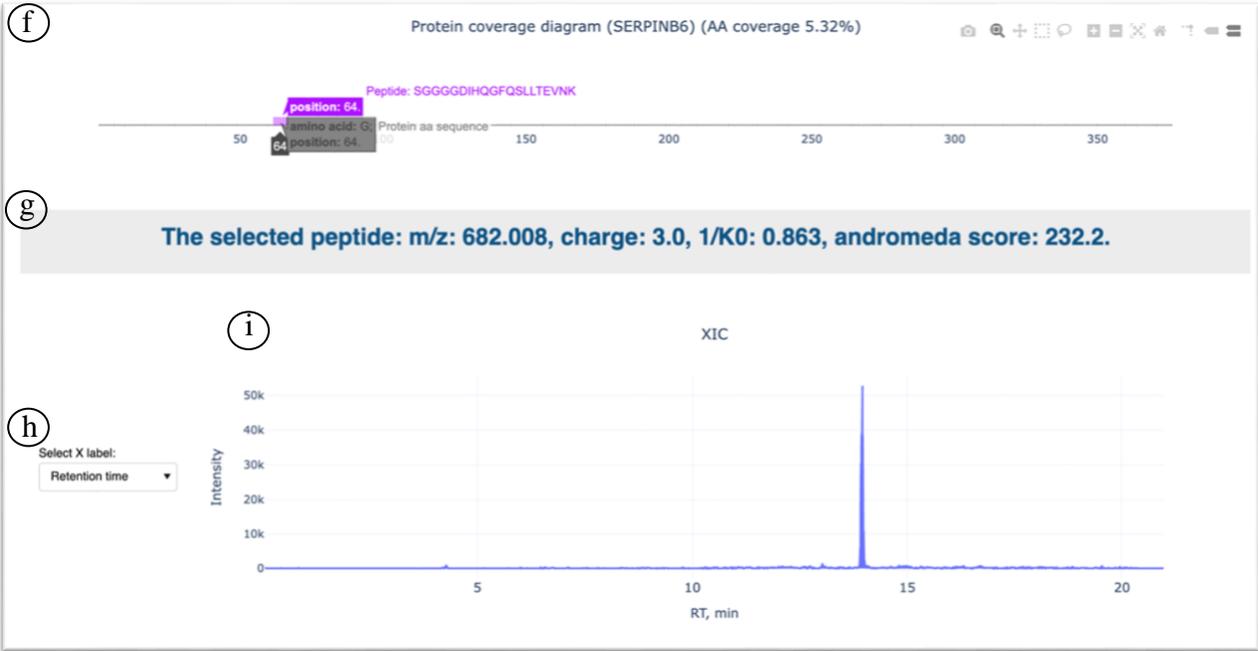
Below the "Peptides table", the following information will be shown for the selected peptide:

- an updated protein coverage diagram for the selected peptide only (f);
- a header (g) containing peptide information, such as its m/z, charge and ion mobility values, as well as the andromeda quality score calculated by MaxQuant;
- a "Select X label" drop-down menu (h) which allows to switch between the "Retention time", "Ion mobility" and "m/z" options, resulting in the type of line plot shown in (i), such as an extracted ion chromatogram (XIC), mobilogram or MS1 spectum respectively;

◀ PREVIOUS FRAME   NEXT FRAME ▶

j   MS1 frame(s) #7813   MS2 frame(s) #7819

Inversed IM, V·s·cm²   *Inversed IM, V·s·cm²*

200 400 600 800 1000 1200 1400 1600   200 400 600 800 1000 1200 1400 1600
m/z, Th   *m/z, Th*
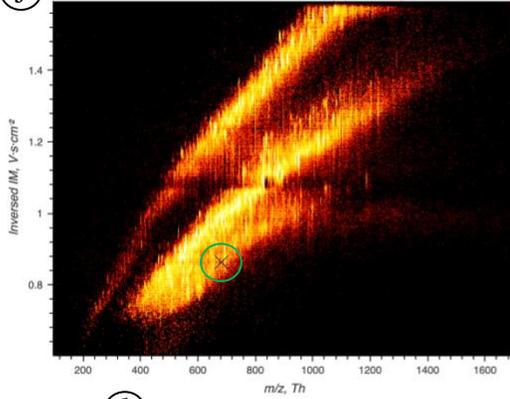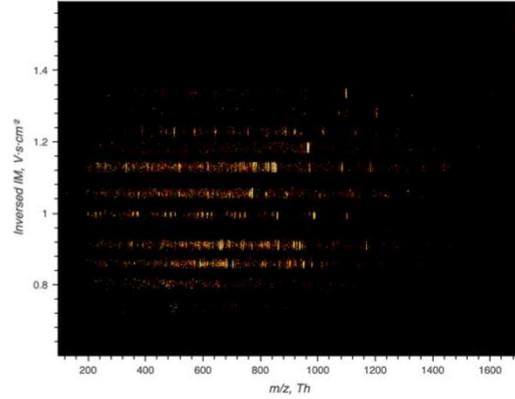
l   EXPORT AS .SVG   EXPORT AS .SVG

m   OVERLAP FRAMES

n   ☑ Show mirrored spectra

o   MS2 spectrum for Precursor: 40725
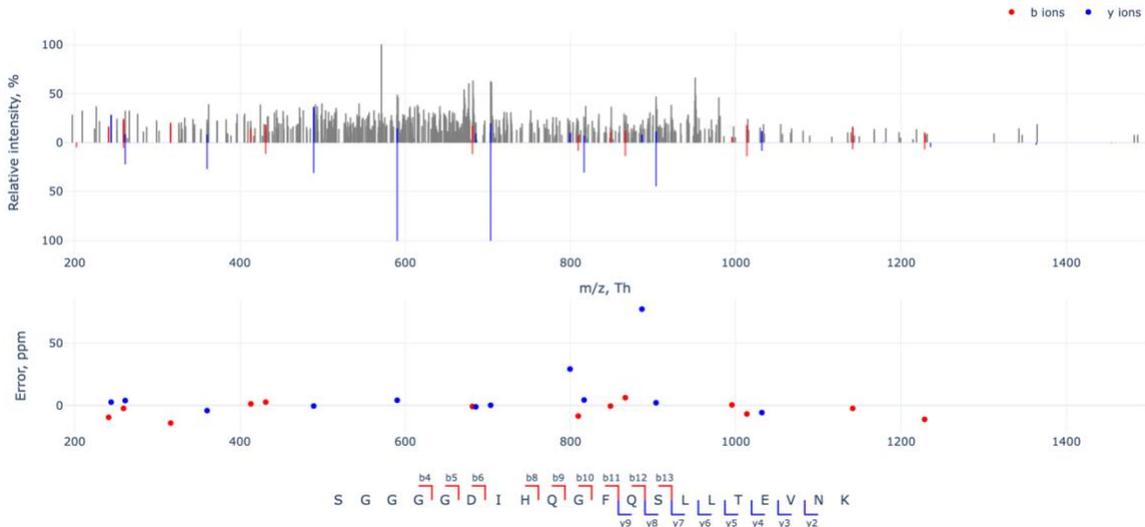
● b ions   ● y ions

Relative intensity, %

m/z, Th

Error, ppm

b4 b5 b6   b8 b9 b10 b11 b12 b13
S G G G G D I H Q G F Q S L L T E V N K
y9 y8 y7 y6 y5 y4 y3 y2

- two heatmaps (j) showing the intensity (colored) of the observed precursor (MS1) or fragment (MS2) masses (x-axis) against their ion mobility values (y-axis). The position of the peptide picked in the MS1 frame is marked as 'X'.
- "Previous frame" and "Next frame" buttons (k) allowing you to switch between the frames in which the peptide was analyzed, if there are more than one such frame;
- **(does not work in the current version)** the "Export as .svg" buttons (l) allowing to save heatmaps as .svg plots with the plot size specified in the input field "The size of the saved plots (height, width)" in the "Customization Options". Note that the plot will be saved in the same folder specified in the "Specify the full path to the folder with

unprocessed Bruker files:*" field on the "Data import" tab.

- an "Overlap frames" button (m) allowing overlapping of the frames in which the peptide was analyzed, if there are more than one such frame;
- a "Show mirrored spectra" check box showing theoretical mirrored MS2 spectra with predicted intensities coloring the b- and y-ions in the same color as in the experimental plot. This check box is enabled only if the "Activate the prediction" check box was checked during data import.
- a compound graph (o) showing the MS2 spectrum highlighting the b- and y-ions and the water and ammonia losses identified by MaxQuant, a mass error plot for the identified b- and y-ions and a peptide sequence showing which ions were identified for the selected peptide.

For the shown peptide, visualization of the mass spectrum confirms the high andromeda quality score, as many b- and y-ions were detected in the data with relatively low errors, which is essential for a confident identification. However, we may also observe many unidentified peaks as well, including high-intense ones.

b) DIA data analyzed with DIA-NN software

The following DIA-NN output files must present in the output folder for AlphaViz to work correctly: '{project_name}.tsv' and '{project_name}.stats.tsv'.

(f)

Protein coverage diagram (protein ID Q9NRG9) (AA coverage 1.28%)

100   200   300   400   500

(g)

The selected peptide: m/z: 428.21, charge: 2, 1/K0: 0.8, Quantity.Quality score: 0.99.

(i)

Precursor and fragment elution profiles of FS[Phospho (STY)]PVLGR in RT dimension (8.99 min)

(h)

Select dimension:

RT dimension ▼

intensity

150k
100k
50k
0

8.6   8.7   8.8   8.9   9   9.1   9.2   9.3   9.4

rt

— precursor — b2 (315.074) — b3 (412.127) — b4 (511.195) — b5 (624.279) — b6 (681.301) — y2 (232.14)
— y3 (345.224) — y4 (444.293) — y5 (541.346) — y6 (708.344)

(j)

MS1 frame(s) #5084

Inversed IM, V·s·cm⁻²

1.4
1.2
1
0.8

200 400 600 800 1000 1200 1400 1600

m/z, Th

MS2 frame(s) #5085

Inversed IM, V·s·cm⁻²

1.4
1.2
1
0.8

200 400 600 800 1000 1200 1400 1600

m/z, Th

(k)

EXPORT AS .SVG   EXPORT AS .SVG

Below the "Peptides table", the following information will be shown for the selected peptide:

- an updated protein coverage diagram for the selected peptide only (f);
- a header (g) containing peptide information, such as its m/z, charge and ion mobility values, as well as the "Quantity.Quality" score calculated by DIA-NN;
- a "Select dimension:" drop-down menu (h) which allows to switch between the "RT dimension" and "RT/IM dimension" options, resulting in the type of the plot shown in (i), such as an 1D or 2D elution profile plots respectively. For the 1D elution profile

plot, fragments can be (de)selected by clicking on the ions in the legend.

- two heatmaps (j) showing the intensity (colored) of the observed precursor (MS1) or fragment (MS2) masses (x-axis) against their ion mobility values (y-axis). The position of the peptide picked in the MS1 frame is marked as 'X'.
- **(does not work in the current version)** the "Export as .svg" buttons (l) allowing to save heatmaps as .svg plots with the plot size specified in the input field "The size of the saved plots (height, width)" in the "Customization Options". Note that the plot will be saved in the same folder specified in the "Specify the full path to the folder with unprocessed Bruker files:*" field on the "Data import" tab.

For the shown peptide, visualization of the 1D elution profile plot confirms the high "Quantity.Quality" score, forming together with the precursor one sharp peak of similar shape, which is necessary for a confident identification.

2. "Quality Control" tab

Hereby you can view summary statistics of the whole experiment, including how many proteins and peptides were identified in each sample, how many MS1 and MS2 scans were processed and so on. Furthermore, you can interactively examine quality control plots of the selected sample.

IMPORTANT! If only raw data without any software output has been loaded in the "Data Import" panel (to work in "Targeted Mode" only), the "Quality Control" tab is empty. Otherwise, the appearance of this tab will be slightly different for different types of data.
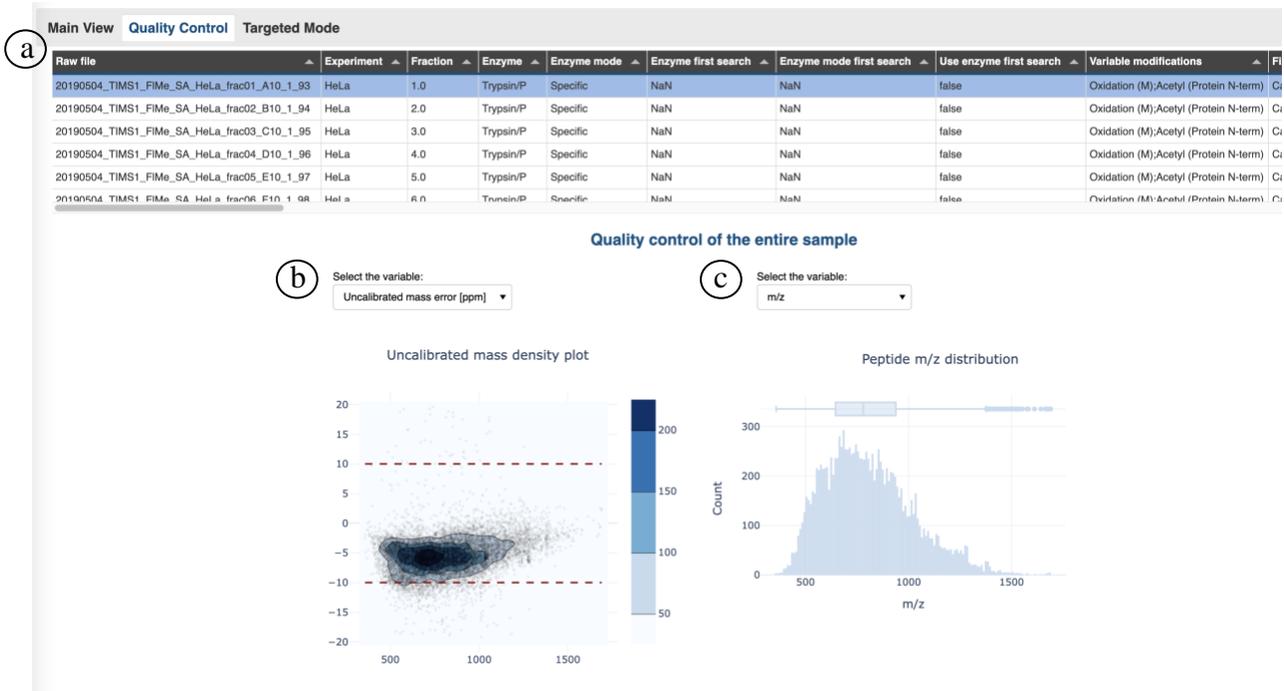
Figure 1A. The appearance of the "Quality Control" tab for the DDA data analyzed with MaxQuant.
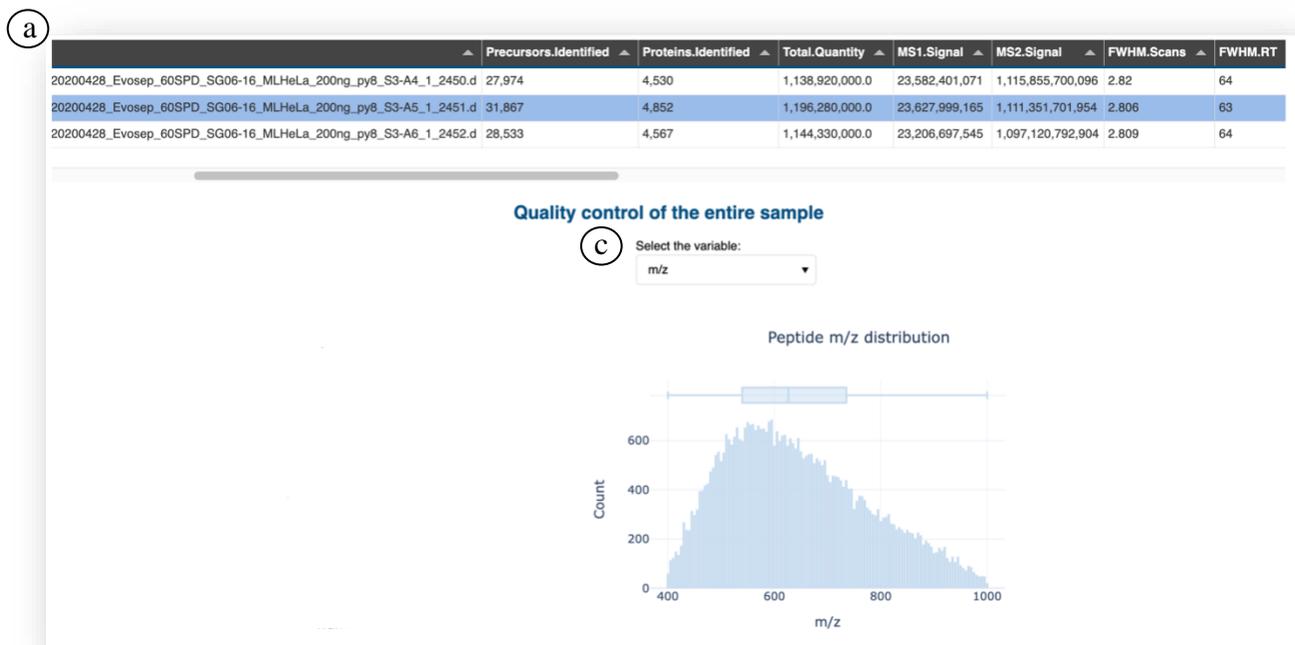

Figure 1B. The appearance of the "Quality Control" tab for the DIA data analyzed with DIA-NN.

For both DDA and DIA data the summary statistics table of the experiment (a) will be shown at the top of the "Quality Control" tab with a pre-selected row (highlighted in blue) corresponding to the loaded sample.

For DDA data analyzed with the MaxQuant tool, a section (b) displays the mass density plot before and after calibration (by selecting the variable in the drop-down menu). In the uncalibrated plot, the mass error currently selected in AlphaViz in ppm is displayed a dashed red line, guiding the user to set the correct mass error in the "Tolerance settings" section.

Using a drop-down menu in the part (c) the variable can be selected for the building a distribution plot, such as m/z, length, charge, scores (Andromeda score, Quality.Quantity score), etc.

3. "Targeted Mode" tab

The functionality implemented in this tab is designed to allow researchers to look into raw DIA data for a particular peptide of interest. If non-DIA data is loaded into AlphaViz, the tab will be empty and the message "To use this functionality please load DIA data." will be displayed.

- **"Manual Input" card**

The first card in the "Targeted Mode" tab allows the manual input of peptide information for further signal extraction from the raw data. This can be done in two ways:

- **Manually**: Set how many rows (peptides) to add to the table and using the integer input field (a) add some empty rows to the table. The table cells can then be modified manually.
- **By loading a file**: Load the table by pressing the "Choose file" button (b), containing columns with the same names as for the table (c) in the screenshot: "name", "sequence", "charge", "im" and "rt". The table can be saved in a file with any of the following extensions, e.g. .tsv, .csv or .txt. It is recommended to use this method.

The obligatory information about the peptide of interest, which must be entered in table (c), is:

- its name that can contain a "unique identifier" or any other information;
- amino acid sequence. Any modifications must be directly included in the peptide sequence, as for the demonstrated case. All modifications should be specified before the modified amino acid using the abbreviations used in the [modifications.tsv](#) file in AlphaViz (the "Identifier column). For example, the sequence for the peptide CKSAASSSNKR with deamidated asparagine will be "CKSAASSSdeamNKR";
- charge;
- ion mobility (im);
- retention time (rt).

The index column is automatically assigned in the table.

The table can be cleared by pressing the "Clear table" button (d).

Once the table row is filled with all necessary information and selected (highlighted in blue), elution profile plots and heatmaps appear in part (e). For the peptide shown, we observe elution profiles in the retention time (1D) and retention time / ion mobility (2D) dimensions confirming the presence of many fragment ions from the b- and y-series.

(does not work in the current version) The "Export as .svg" buttons (f) allowing to save heatmaps as .svg plots with the plot size specified in the input field "The size of the saved plots (height, width)" in the "Customization Options". Note that the plot will be saved in the same folder specified in the "Specify the full path to the folder with unprocessed Bruker files:*" field on the "Data import" tab.

- **"Prediction" card**

The second card allows to use prediction by the deep learning model built into AlphaViz. This will predict peptide information such as retention time and ion mobility. There are two ways to add the peptide data:

- **Manually (a)**: See the "Manual Input" card for instructions.
- **By loading a file**: Load the table by pressing the "Choose file" button (b) containing columns with the same names as for the table (c) in the screenshot: "sequence", "mods", "mod_sites" and "charge". The table can be saved as a .tsv file. It is recommended to use this method.

The obligatory information about the peptide of interest for the prediction that must be entered in table (c) is:

- sequence. In this case, it should be a stripped amino acid sequence with all modifications removed;
- mods. All modifications should be included as a string separated by semicolon. For example, "Phospho@T;Phospho@T" for two phosphorylated threonines. Examples of available modifications are 'Oxidation@M', 'Phospho@S', 'Phospho@T', 'Phospho@Y', 'Acetyl@Protein N-term', 'Carbamidomethyl@C' and so on;
- mod_sites. The index of modified amino acids. If the N-terminal of the peptide is modified, the modification index will be 0. If the first amino acid is modified, the index will be 1, and so on;
- charge.

After loading (preferably) or manually filling in the table (c), the "Run prediction" button should be pressed. The prediction process is indicated by the spinner. As soon as the prediction is completed, the table (c) is updated. Clicking on any peptide in the table, the visualization of its raw data will be presented below the table (c) (see the "Manual Input" card).