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. After downstream processing in Perseus, MSstats or the Clinical Knowledge Graph, differentially expressed proteins become possible candidates for biomarker discovery. AlphaViz is an automated visualization pipeline to link these identifications with the original raw data and easily assess their individual quality or the overall quality whole samples.

Installation

Windows

Prerequisites: **Windows 10** (a system update might be necessary in case older versions do not work).

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 "User Account Control" dialog asking about permission for the app to make changes to your device press the "Yes" button.
3. In the appearing "Setup – alphaviz version X.X.X" dialog window accept the License Agreement and press the "Next" button.
4. Select the destination location for the installation of AlphaViz software (the size of the whole package is around 800 MB) 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. In the appearing "Windows Security Alert" dialog window press the "Allow access" button that will prevent the Windows Defender Firewall from blocking the AlphaViz tool on your PC.
9. Check your default browser (Google Chrome or Mozilla Firefox is recommended for fast running and correct display of AlphaViz visualization) and start working with the tool.

* 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").

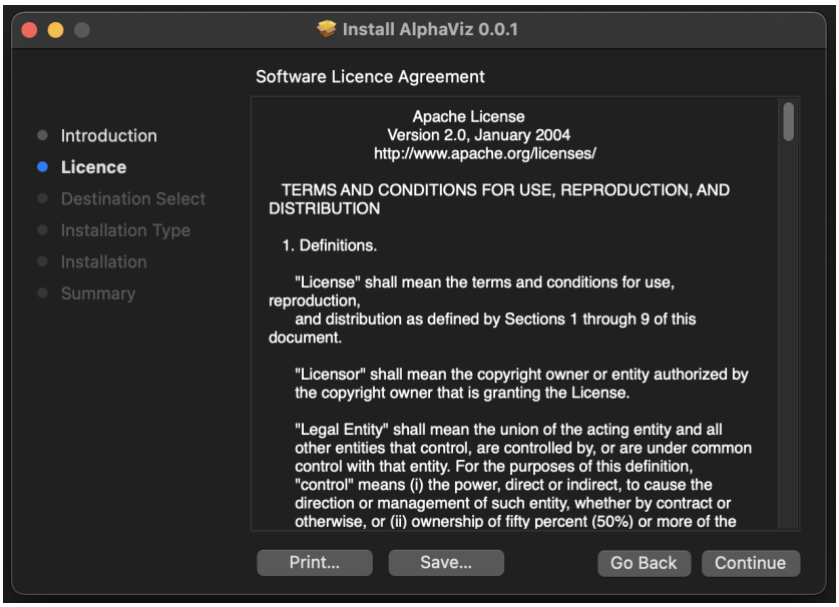
MacOS

Prerequisites: at least **macOS Catalina (10.15) or higher** (a system update might be necessary in case older versions do not work).

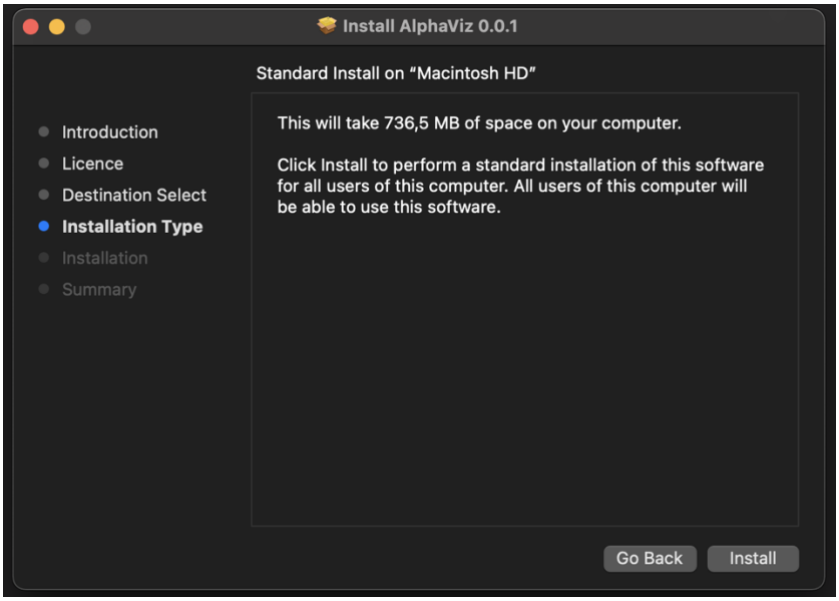
IMPORTANT WARNING: Since AlphaViz uses AlphaTims to read LC-TIMS-Q-TOF data from Bruker's timsTOF pro instrument (Bruker Daltonik), some calibration functions for it are provided by Bruker as libraries and are only available on Windows and Linux. Therefore, to avoid any problems with MS2 spectra quality 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 CLI manual.

1. Download the latest release for macOS (alphaviz_gui_installer_macos.pkg) from the GitHub repository and open the .pkg file.
2. Click "Continue" on the appearing "Install AlphaViz X.X.X" dialog window.

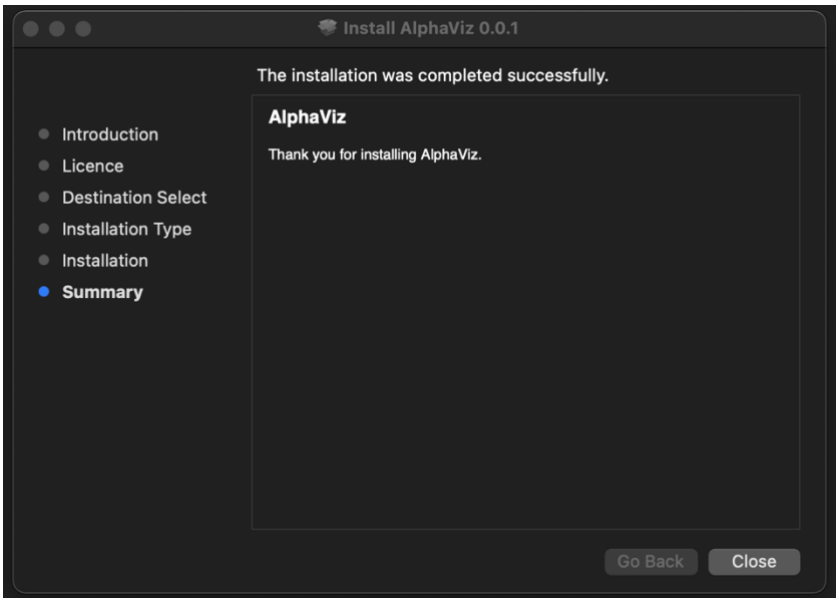
3. In the **License** section the Software License Agreement (Apache License) will be shown for you. To continue the installation, press "Continue" button and in the appeared pop-up window you need to agree with the regulations of the license.



4. Press “Install” to start the installation (the size of the whole package is around 800 MB). This might take a few minutes.

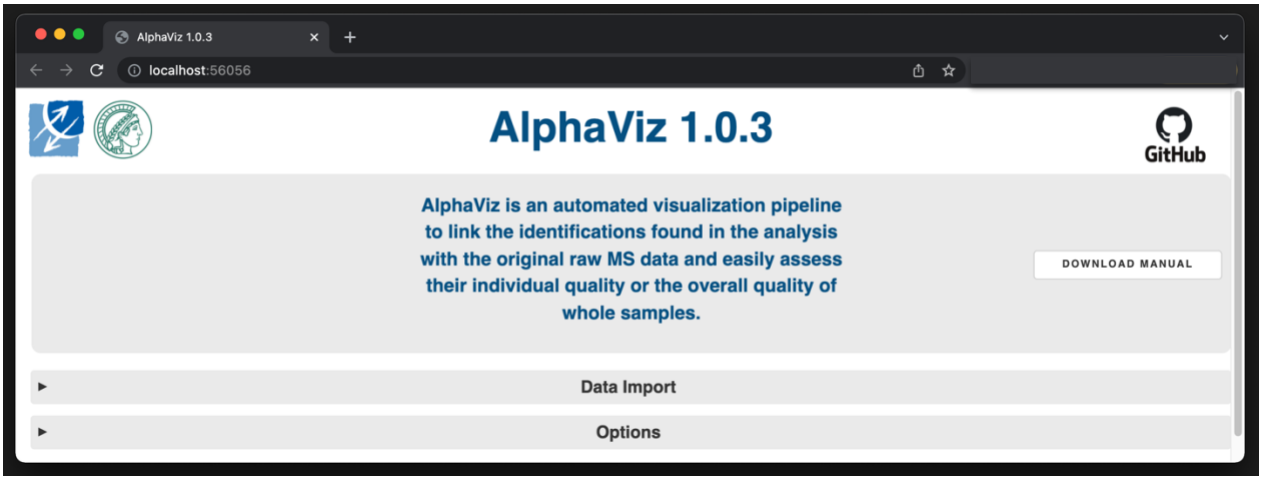


5. Click “Close” to quit the installation menu. AlphaViz is now available in the applications folder on your MacOS.



* In rare cases, you might get an error message that the “alphaviz_gui_install_macos.pkg” cannot be opened because it is coming from an unidentified developer. This error can be avoided by pressing the OPTION key or by going to ‘Security and Privacy’ under ‘System Preferences’ and enable AlphaViz installation.

How to use AlphaViz

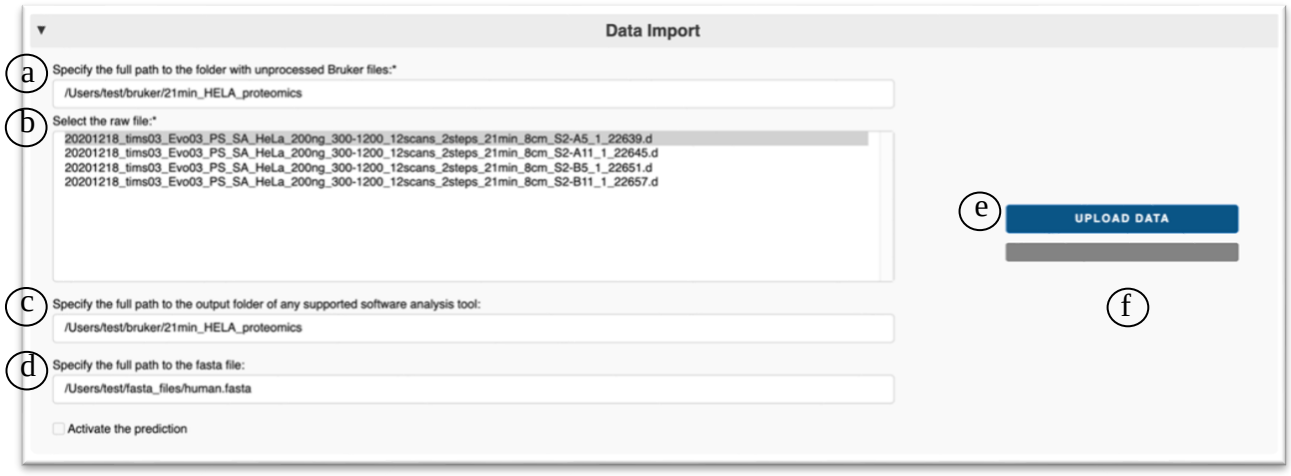


After launching AlphaViz, firstly a terminal window will open containing the tool's background information and after that you can access a new tab called "AlphaViz X.X.X" in the default browser with the "localhost:XXXXX" URL. For better performance, we recommend to use Google Chrome or Mozilla Firefox. To close the tool, just close the opened "AlphaViz X.X.X" tab or press "Ctrl + c" in the running terminal window.

Pressing "Download manual" button you can also download and check out this GUI manual.

Data import

To get started with AlphaViz in the "Data Import" panel you need to load the raw file itself or with the results of the analysis from any supported software, e.g. currently MaxQuant or DIA-NN, along with the fasta file that was used to analyse the data.



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).

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

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

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).

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

e) 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 (f) with detailed information.

Options

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

a) 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 3D elution profile plots for data-independent acquisition (DIA) data.

b) 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 generate extracted ion chromatograms (XICs) or mobilograms (in data-dependent acquisition, or DDA, mode) and 2D or 3D elution profile plots (in DIA mode).

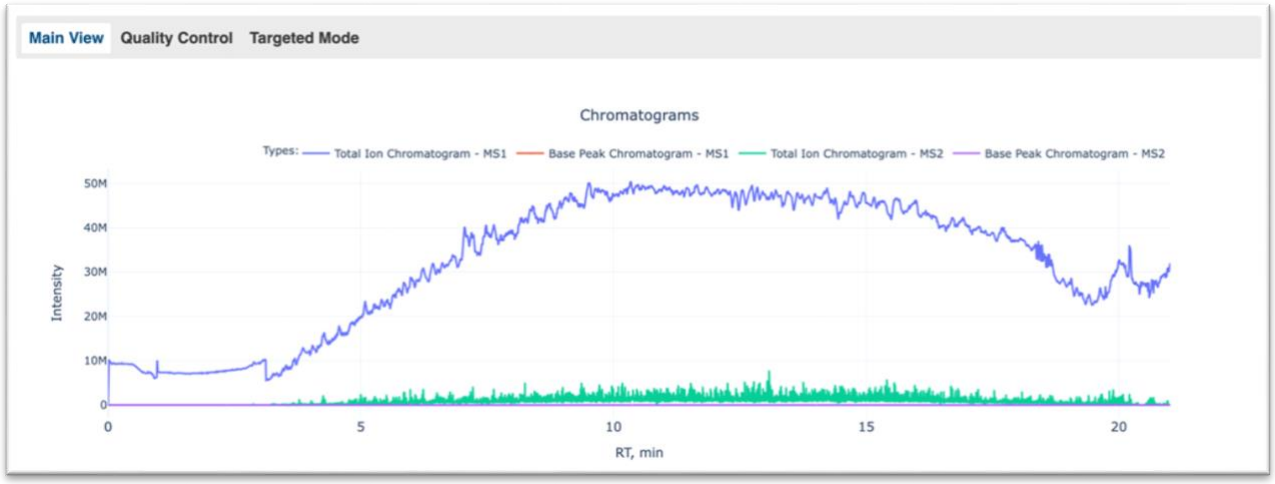
c) In the "Color scheme options" card the names of qualitative and sequential color scales can

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.

Data Visualization

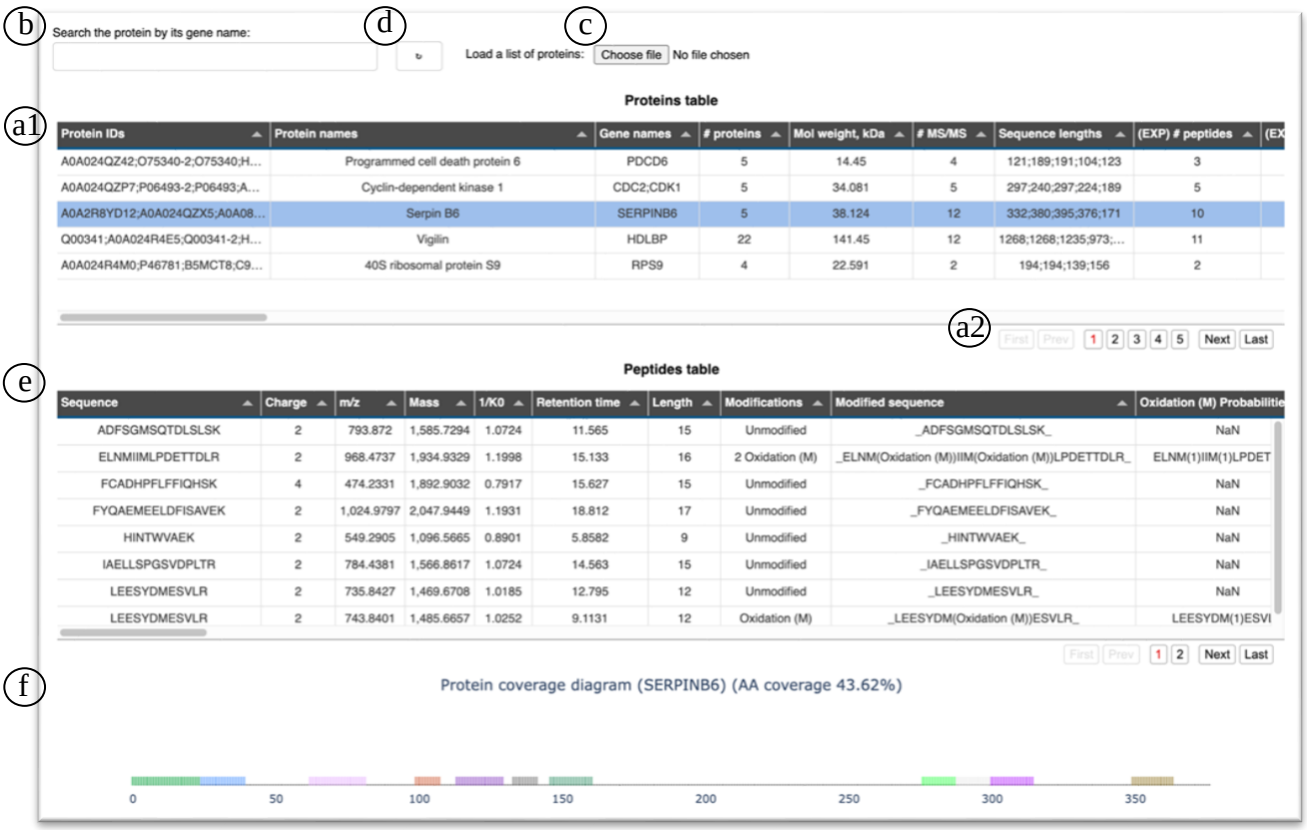
Once the data has been loaded into AlphaViz, several new tabs will appear, such as “Main View”, “Quality Control” and “Targeted Mode”. By default, the “Main View” tab is opened (highlighted in white), but you can switch between tabs by clicking on them.

- “Main View” tab



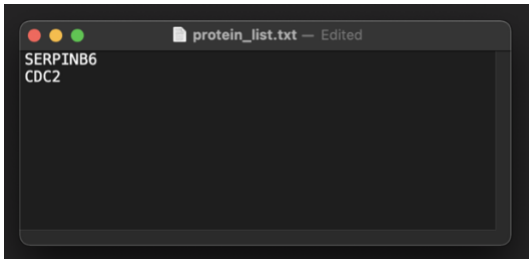
To assess the overall performance of the LC and MS instruments, various types of chromatograms including total ion chromatograms (TICs), showing the summed intensity of all detected precursor (MS1) / fragment (MS2) ions against the retention time, and base peak chromatograms (BPIs), showing the intensities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

IMPORTANT! If only raw data without any software output 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 analysis 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.



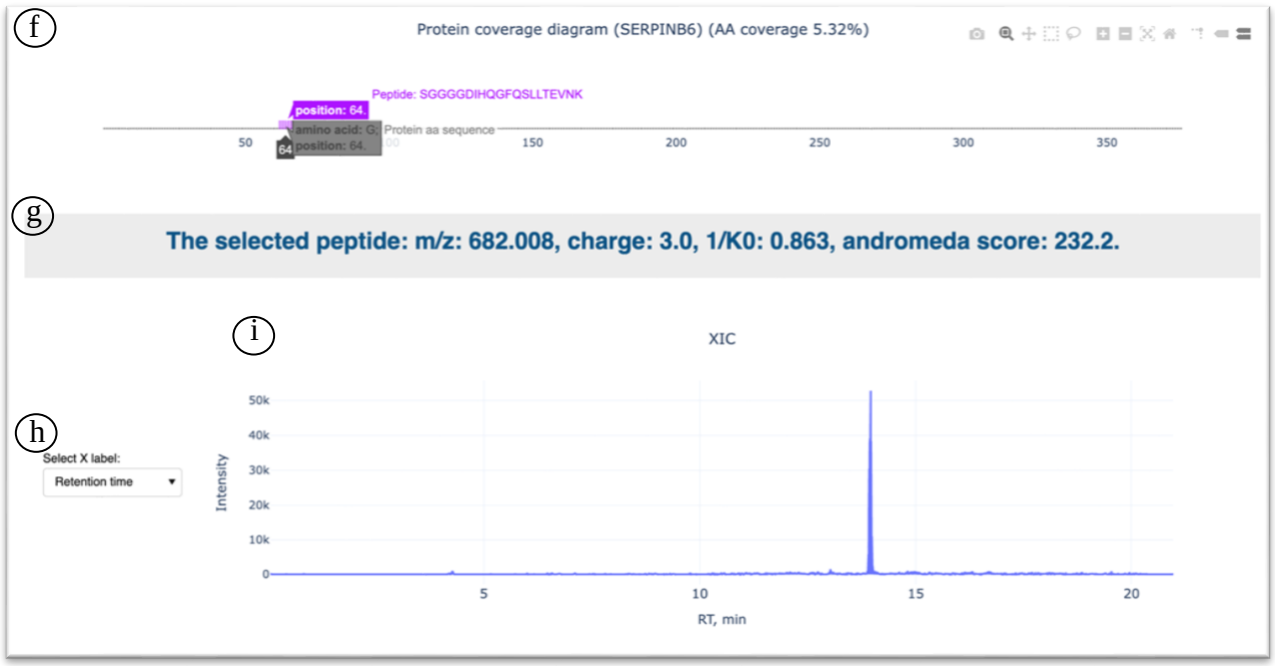
In all the ways described above, to undo any filtering press the “Undo” button (d). In this case the table content will be returned to the original proteins table after data loading.

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” (e). In addition, a protein coverage diagram (f) will be displayed below the “Peptides table” showing the position of the 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.

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

1. **DDA data analyzed with MaxQuant software**:

The following MaxQuant output files must be present in the output folder for AlphaViz to work correctly: 'allPeptides.txt', 'msms.txt', 'evidence.txt', 'proteinGroups.txt', 'summary.txt'.

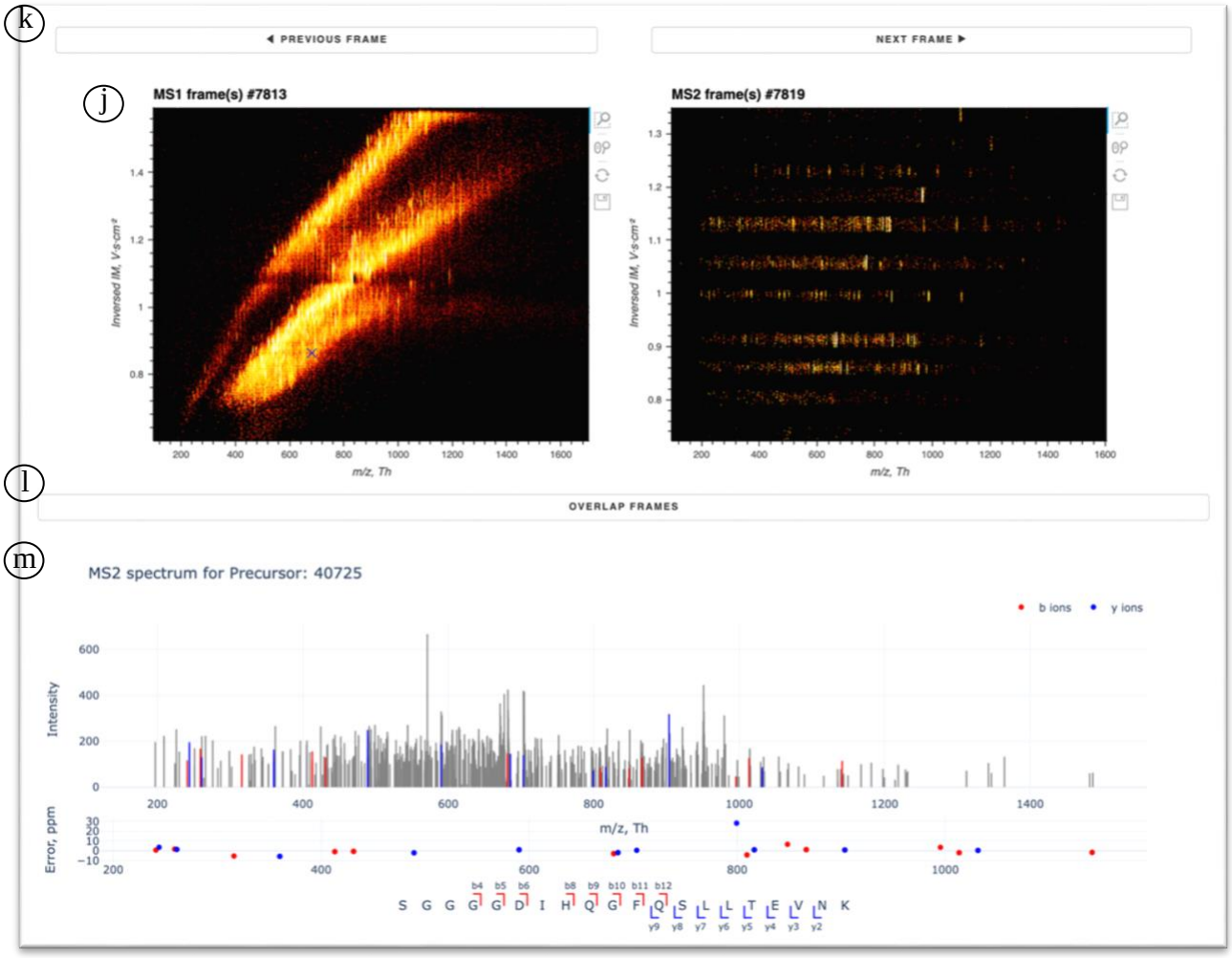


Below the “Peptides table”, the following information will be shown for the selected peptide:

- 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” and “Ion mobility” options, resulting in the type of line plot shown in (i), such as an extracted ion chromatogram (XIC) or mobilogram respectively;

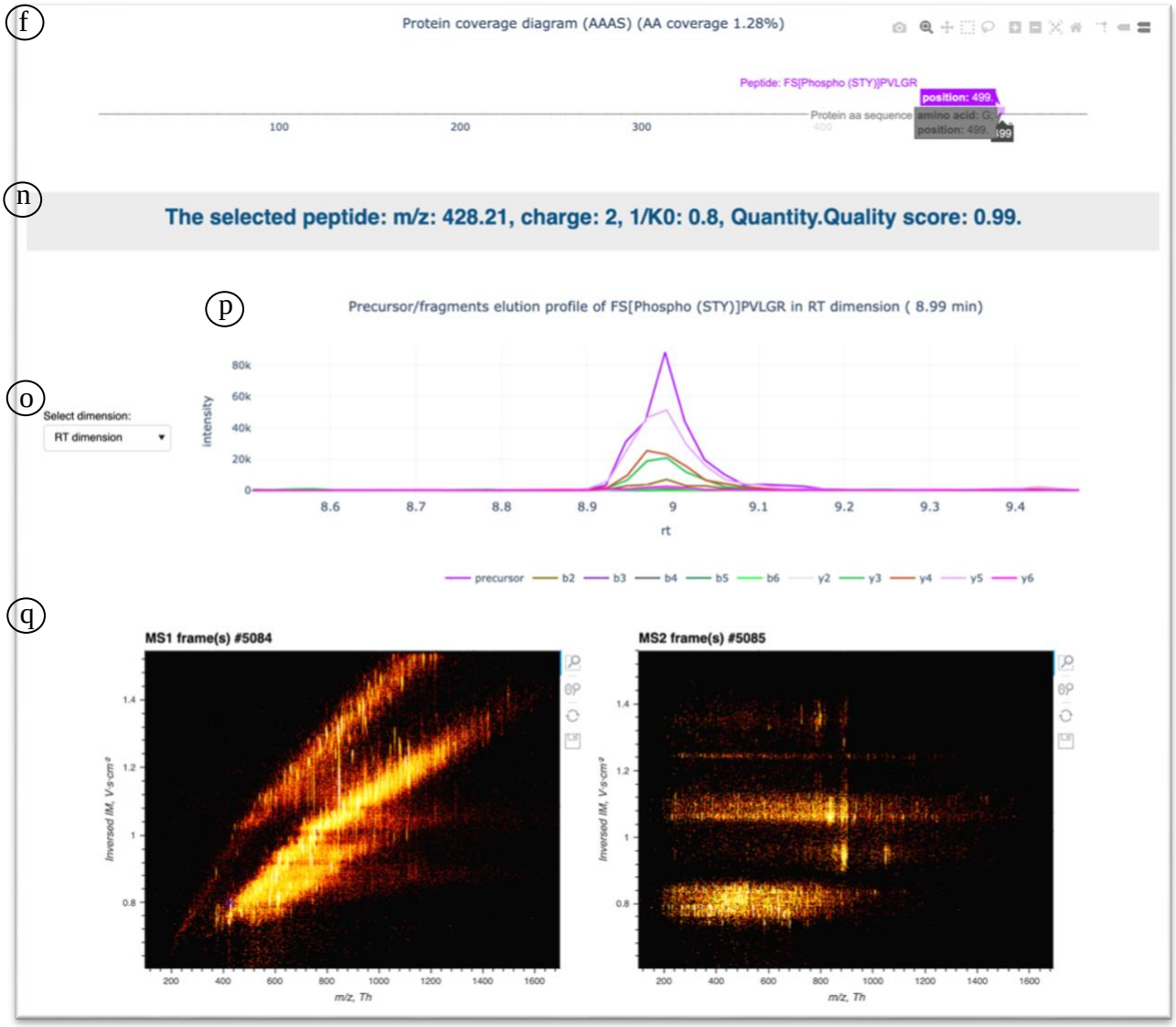


- two heatmaps (j) showing the intensity (colored) of the observed precursor (MS1) or fragment (MS2) masses (x-axis) against their ion mobility values (y-axis). The position of the peptide picked in the MS1 frame is marked as ‘X’.
- “Previous frame” and “Next frame” buttons (k) allowing you to switch between the frames in which the peptide was analyzed, if there are more than one such frame;
- an “Overlap frames” button (l) allowing overlapping of the frames in which the peptide was analyzed, if there are more than one such frame;
- a complex graph (m) showing an interactive MS2 spectrum highlighting the b- and y-ions 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.

2. **DIA data analyzed with DIA-NN software**:

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



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 (n) 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 (o) which allows to switch between the “RT dimension” and “RT/IM dimension” options, resulting in the type of the plot shown in (p), such as an 2D or 3D elution profile plots respectively;
- two heatmaps (q) 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’.

For the shown peptide, visualization of the 2D 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.

- “Quality Control” tab

Hereby you can view summary statistics of the whole experiment, including how many proteins and peptides were identified in the experiment, how many MS1 and MS2 scans were processed and so on, as well as interactively examine quality control plots of the whole 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 DIA data analyzed with DIA-NN.

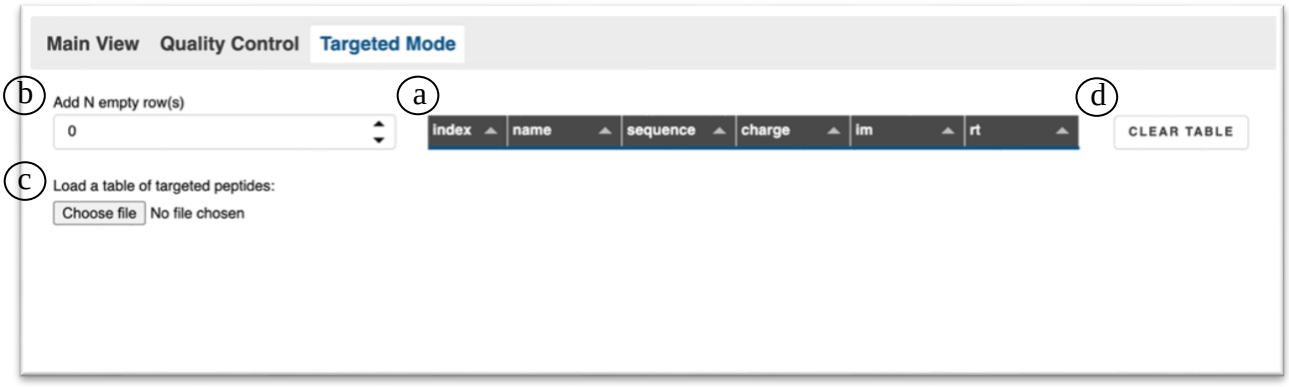
Figure 1B. The appearance of the "Quality Control" tab for the DDA data analyzed with MaxQuant.

For both DDA and DIA data the summary statistics table (a) will be shown at the top of the "Quality Control" tab with a pre-selected row (highlighted in blue) corresponding to the loaded experiment. In part (b) there are several distribution plots showing the number of identified peptides per protein, peptide charge/length/m/z distributions.

For DDA data analyzed with MaxQuant part (c) displays mass density plots before and after calibration.

- ## "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.



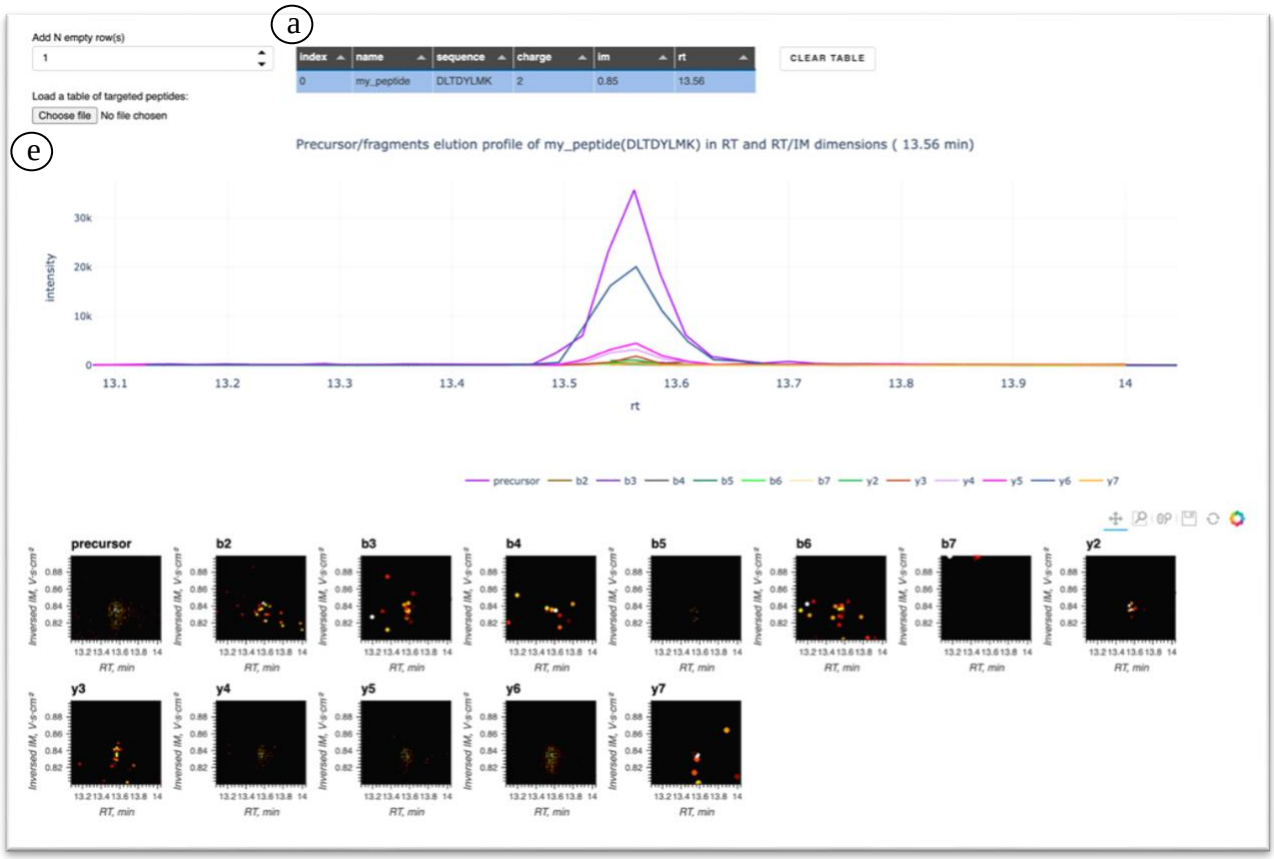
The obligatory information about the peptide of interest, which must be entered in table (a), is:

- its name;
- amino acid sequence;
- charge;
- ion mobility;
- retention time.

The update of the table (a) can be done in two ways:

- **Manually**: Set how many rows (peptides) to add to the table and using the integer input field (b) add some empty rows to the table. The table cells can then be modified manually.
- **By loading the table**: Load the table by pressing the “Choose file” button (c) containing the columns with the same names as for the table (a) on 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.

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), the elution profile plots appear in part (e).

For the shown peptide we see the elution profiles in both dimensions, confirming the presence of many ions from the y-series.