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

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 "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 read and 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.
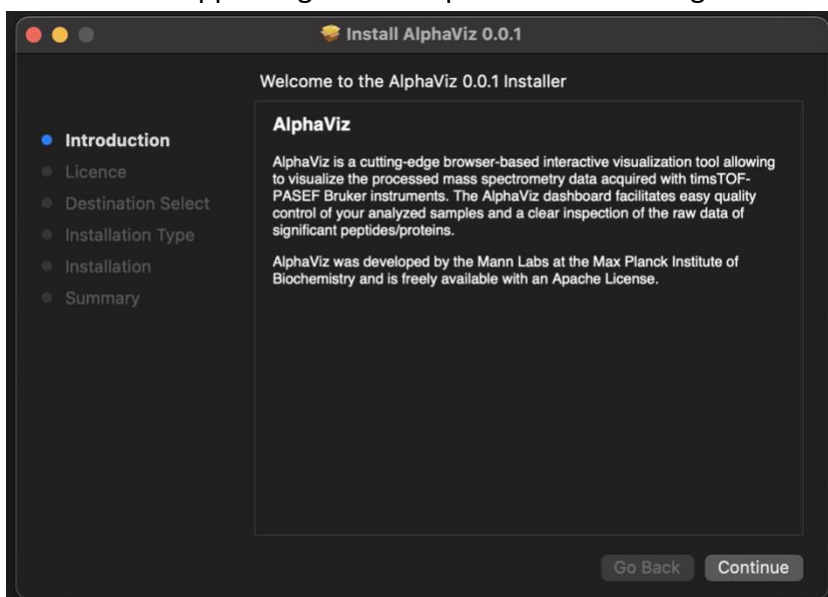
9. A new tab should be automatically opened in your default browser (Google Chrome or Mozilla Firefox is recommended for fast running and correct display of AlphaViz visualization) and start working with the tool.
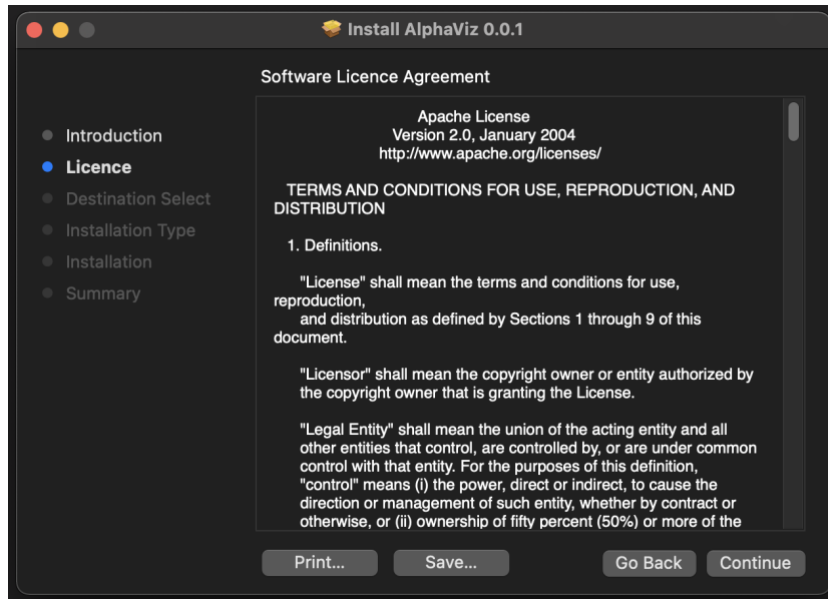
MacOS

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 GUI manual or in the AlphaTims CLI manual.
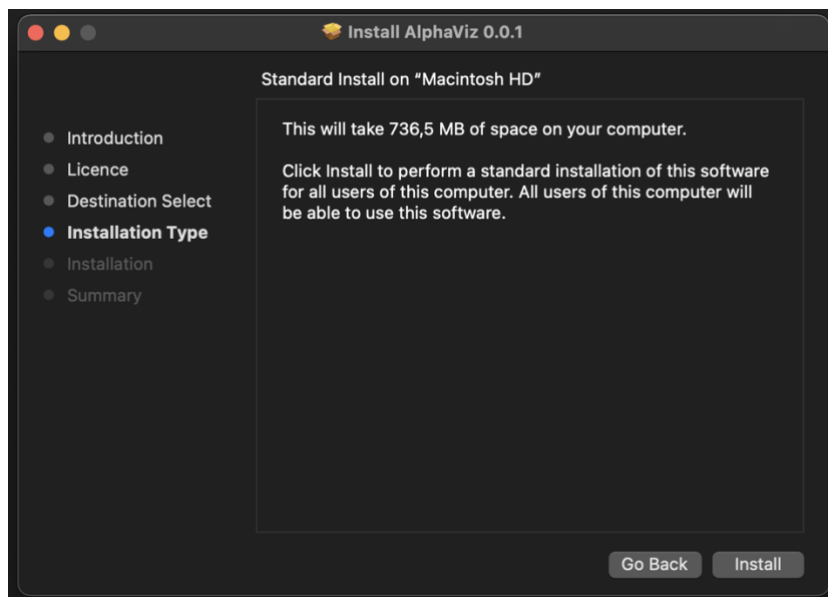
1. Download the latest release for macOS (alphaviz_gui_installer_macos.pkg) from the GitHub repository and open the .pkg file. If you receive a message that the file cannot be opened because it is from an unidentified developer, you can close the message by clicking "OK", go to the "Security & Privacy" 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.
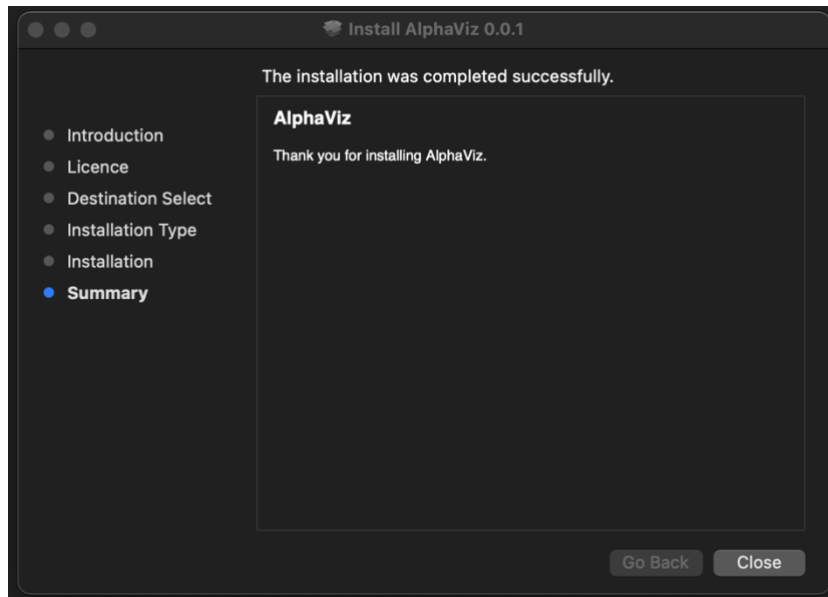

4. Press "Install" to start the installation (the size of the whole package is around 800 MB). This might take a few minutes.
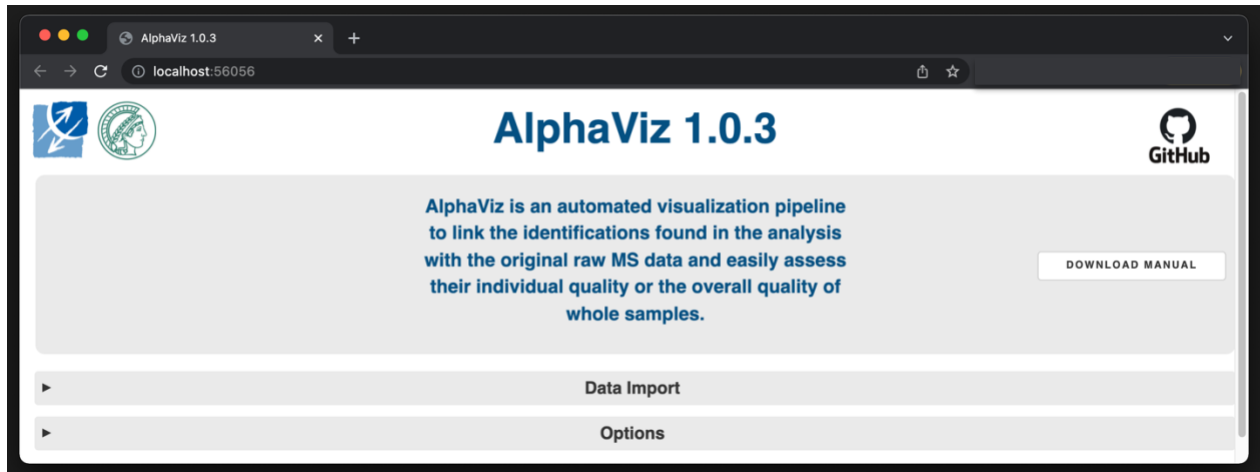

5. Click "Close" to quit the installation menu. AlphaViz is now available in the applications

folder on your MacOS.

How to use AlphaViz

After launching AlphaViz, firstly a terminal window will open 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 "localhost:XXXXX" URL. For optimal 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.

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). Hereafter, you can optionally select 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 or you can skip this step 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) 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.

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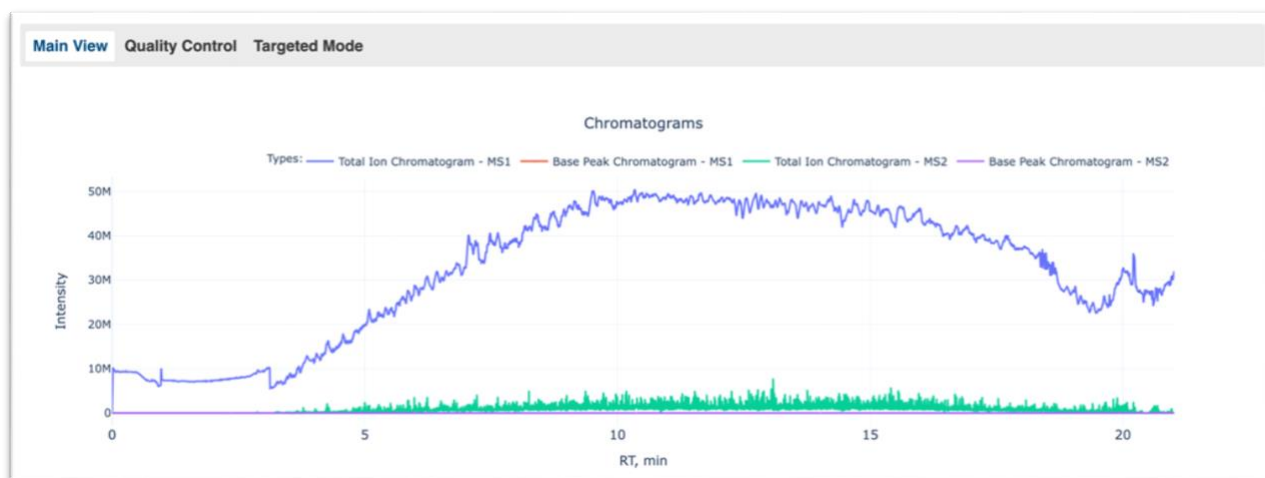
chromatograms (XICs) or mobilograms (in data-dependent acquisition, or DDA, mode) and 2D or 3D elution profile plots (in DIA mode).

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, 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.

-   **"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.

Search the protein by its gene name: (b)

Load a list of proteins: Choose file No file chosen (c)

Proteins table

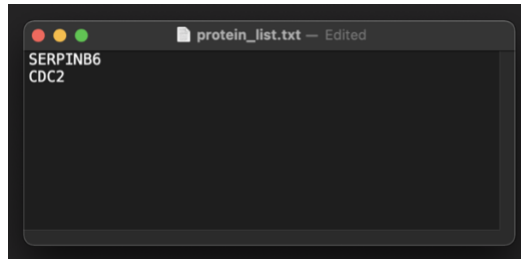| Protein IDs	Protein names	Gene names	# proteins	Mol weight, kDa	# MS/MS	Sequence lengths	(EXP) # peptides	(EX
A0A024QZ42;O75340-2;O75340;H...	Programmed cell death protein 6	PDCD6	5	14.45	4	121;189;191;104;123	3	
A0A024QZP7;P06493-2;P06493;A...	Cyclin-dependent kinase 1	CDC2;CDK1	5	34.081	5	297;240;297;224;189	5	
A0A2R8YD12;A0A024QZX5;A0A08...	Serpin B6	SERPINB6	5	38.124	12	332;380;395;376;171	10	
Q00341;A0A024R4E5;Q00341-2;H...	Vigilin	HDLBP	22	141.45	12	1268;1268;1235;973;...	11	
A0A024R4M0;P46781;B5MCT8;C9...	40S ribosomal protein S9	RPS9	4	22.591	2	194;194;139;156	2	

First Prev 1 2 3 4 5 Next Last (a2)

Peptides table

Sequence	Charge	m/z	Mass	1/K0	Retention time	Length	Modifications	Modified sequence	Oxidation (M) Probabilitie
ADFSGMSQTDLSLSK	2	793.872	1,585.7294	1.0724	11.565	15	Unmodified	_ADFSGMSQTDLSLSK_	NaN
ELNMIIMLPDETTDLR	2	968.4737	1,934.9329	1.1998	15.133	16	2 Oxidation (M)	_ELNM(Oxidation (M))IIM(Oxidation (M))LPDETTDLR_	ELNM(1)IIM(1)LPDET
FCADHPFLFFIQHSK	4	474.2331	1,892.9032	0.7917	15.627	15	Unmodified	_FCADHPFLFFIQHSK_	NaN
FYQAEMEELDFISAVEK	2	1,024.9797	2,047.9449	1.1931	18.812	17	Unmodified	_FYQAEMEELDFISAVEK_	NaN
HINTWVAEK	2	549.2905	1,096.5665	0.8901	5.8582	9	Unmodified	_HINTWVAEK_	NaN
IAELLSPGSVDPLTR	2	784.4381	1,566.8617	1.0724	14.563	15	Unmodified	_IAELLSPGSVDPLTR_	NaN
LEESYDMESVLR	2	735.8427	1,469.6708	1.0185	12.795	12	Unmodified	_LEESYDMESVLR_	NaN
LEESYDMESVLR	2	743.8401	1,485.6657	1.0252	9.1131	12	Oxidation (M)	_LEESYDM(Oxidation (M))ESVLR_	LEESYDM(1)ESVL

First Prev 1 2 Next Last

Protein coverage diagram (SERPINB6) (AA coverage 43.62%)

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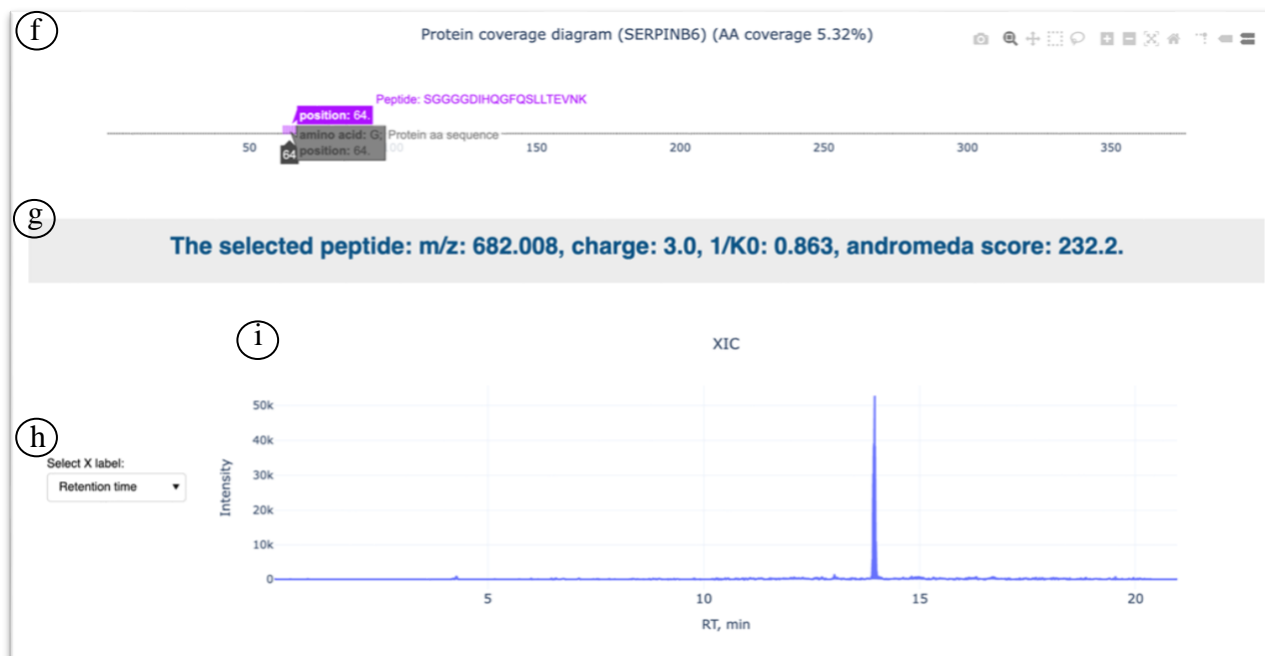
In all the ways described above, to undo any filtering press the "Undo" button (d). In this case the table content will be returned to the original proteins table after data loading.

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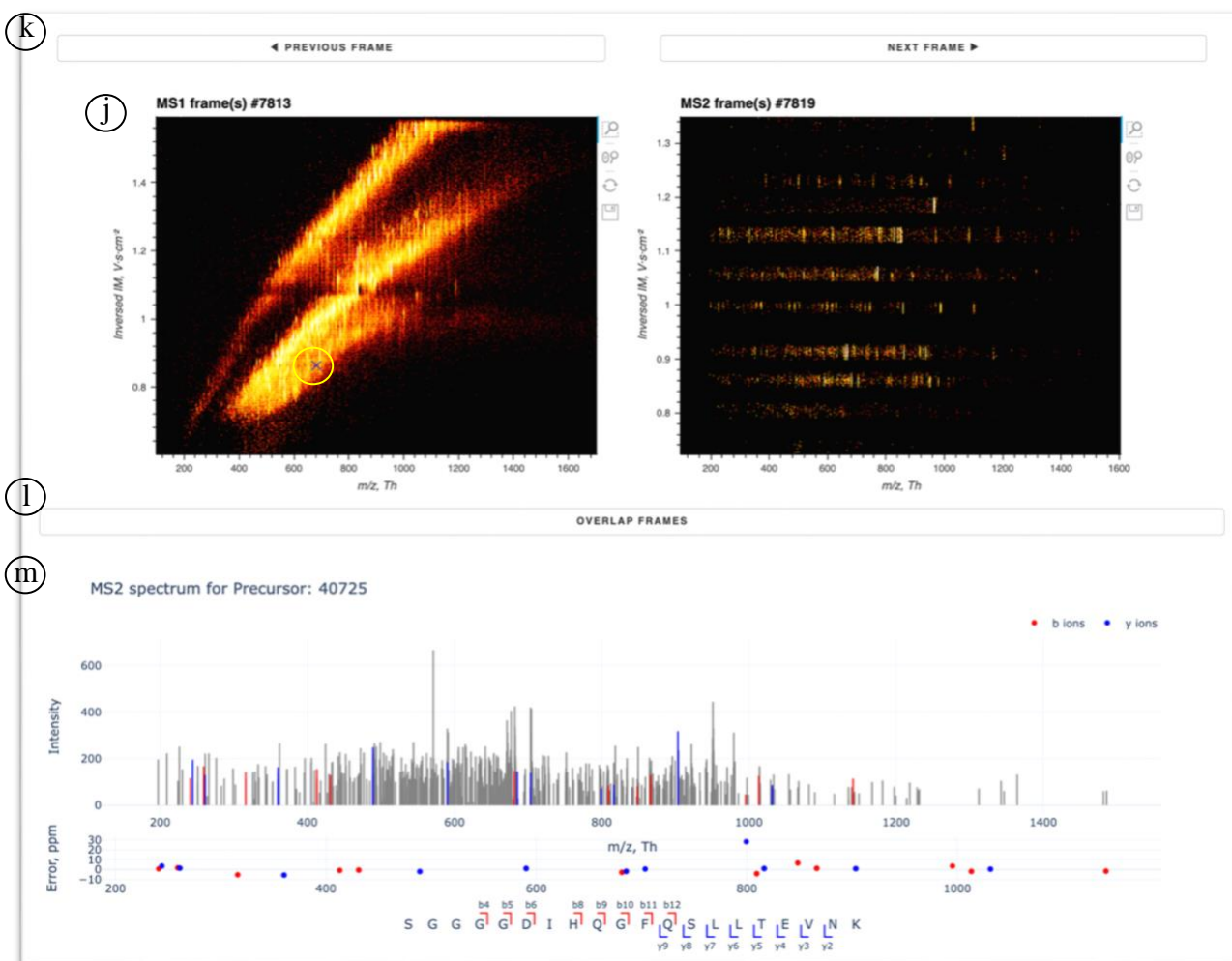
1. **DDA data analyzed with MaxQuant software:**

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

Protein coverage diagram (SERPINB6) (AA coverage 5.32%)

Peptide: SGGGGDIHQGFQSLLTEVNK

position: 64
amino acid: G, Protein aa sequence
64 position: 64.

50   100   150   200   250   300   350

The selected peptide: m/z: 682.008, charge: 3.0, 1/K0: 0.863, andromeda score: 232.2.

XIC

Intensity

50k
40k
30k
20k
10k
0

Select X label:
Retention time ▼

5   10   15   20

RT, min

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;
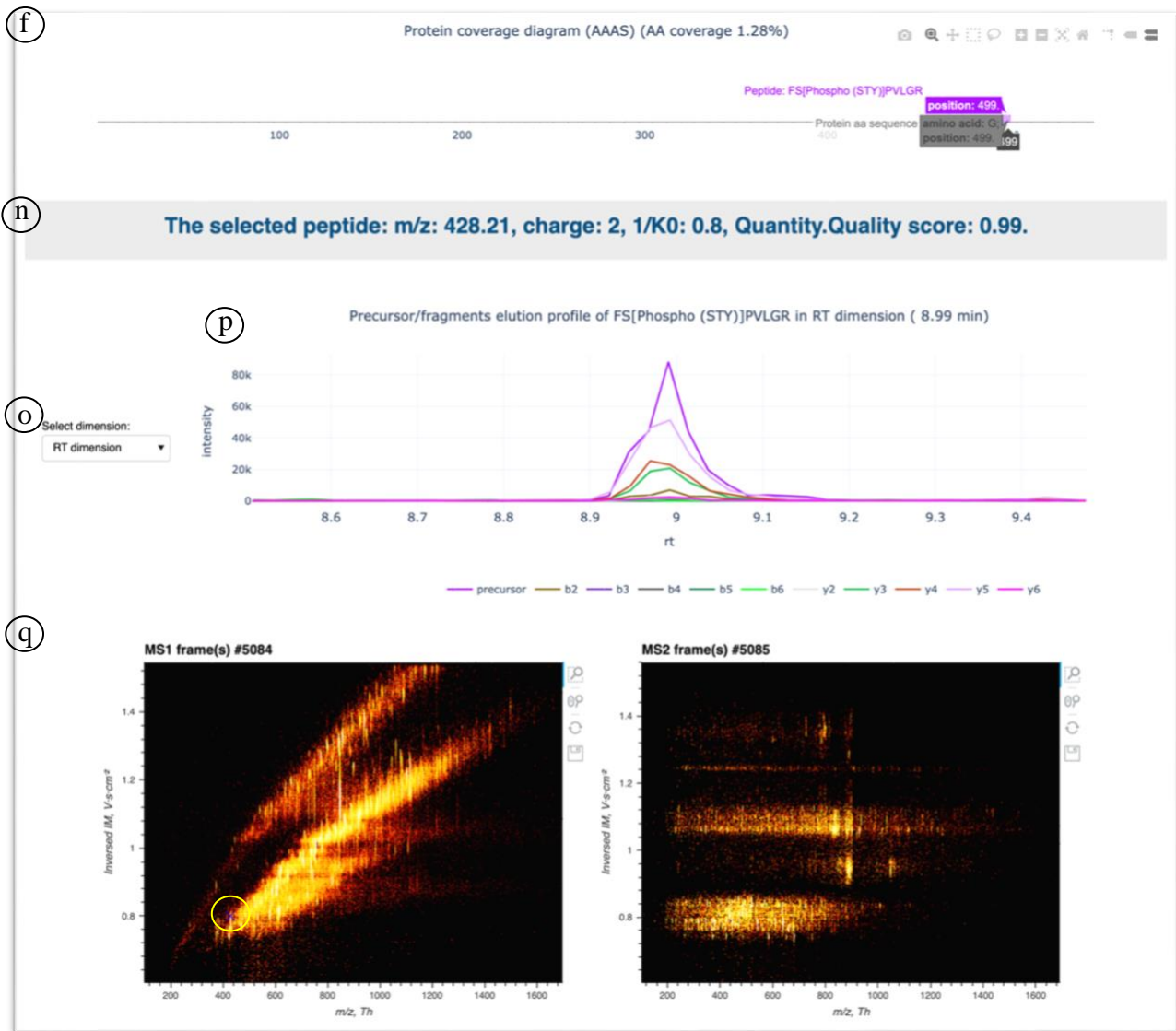
- two heatmaps (j) showing the intensity (colored) of the observed precursor (MS1) or fragment (MS2) masses (x-axis) against their ion mobility values (y-axis). The position of the peptide picked in the MS1 frame is marked as 'X'.
- "Previous frame" and "Next frame" buttons (k) allowing you to switch 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 compound graph (m) showing an MS2 spectrum highlighting the b- and y-ions identified by MaxQuant, an 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.

2. 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'.

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. For the 2D elution profile

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

- 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 each sample, how many MS1 and MS2 scans were processed and so on. Furthermore, you can 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.

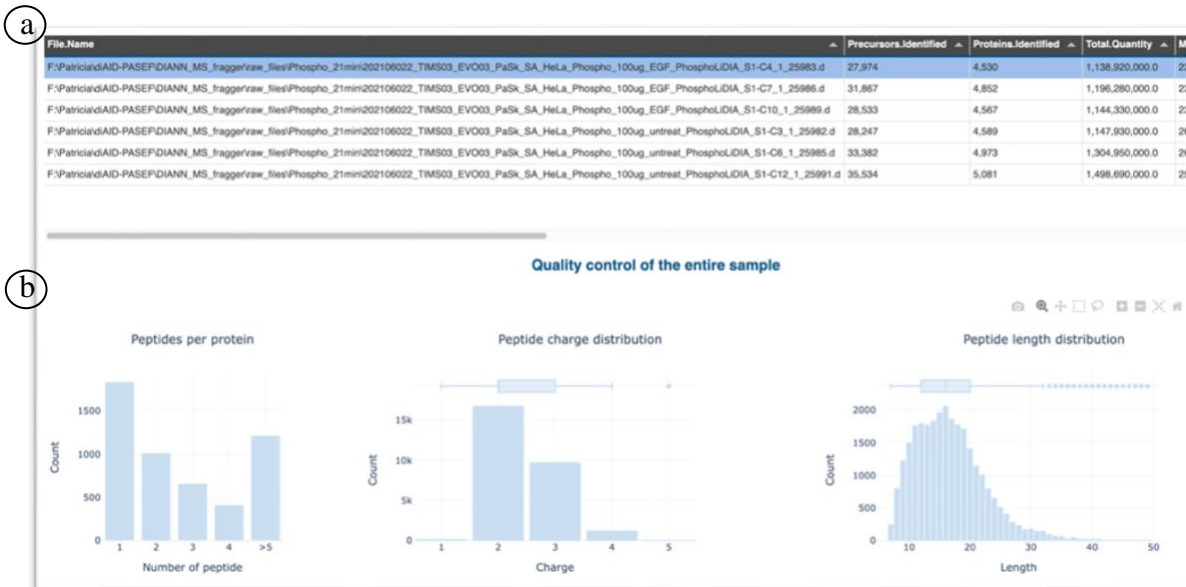*Figure 1A. The appearance of the "Quality Control" tab for the DDA data analyzed with MaxQuant.*

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 (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. 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:
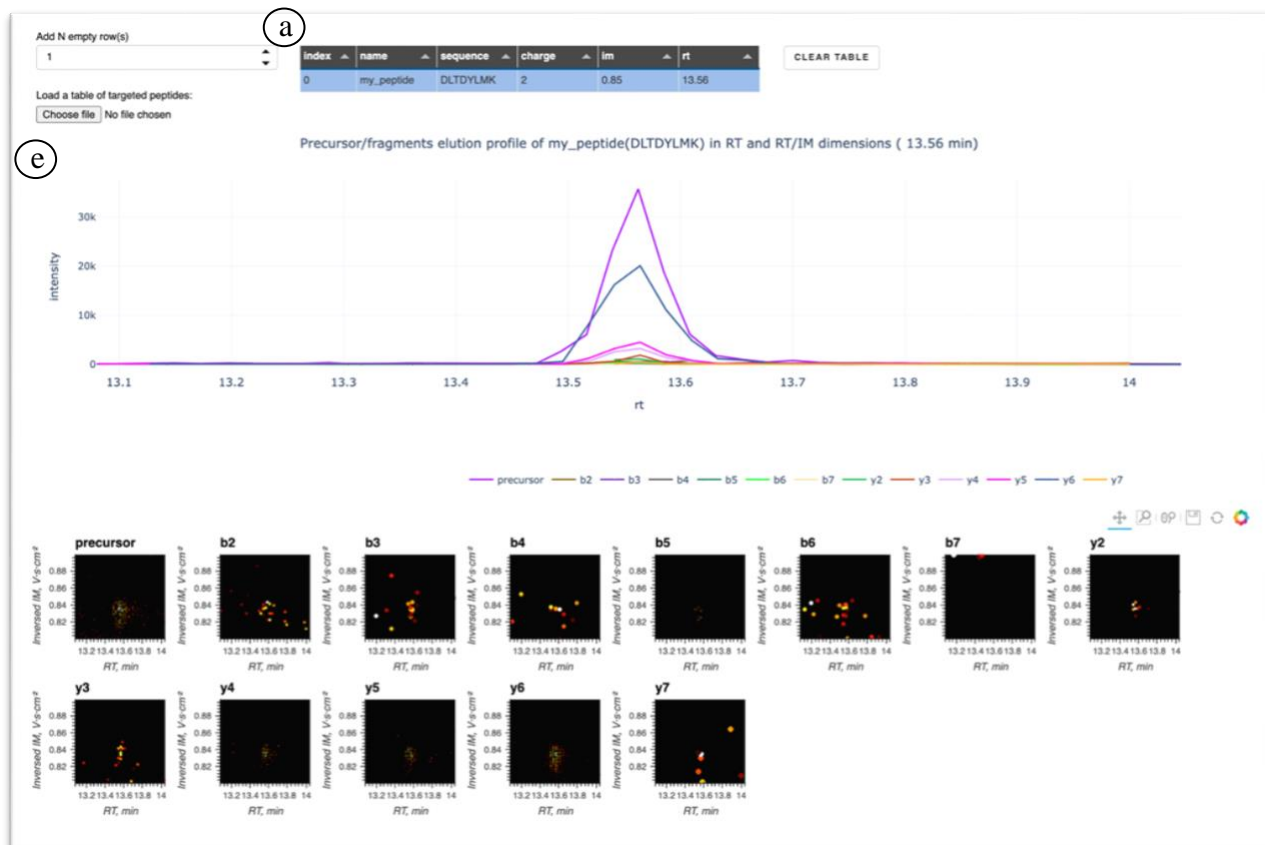
- its name that can contain a "unique identifier" or any other information;
- amino acid sequence;
- charge;
- ion mobility (im);
- retention time (rt).

The index column is automatically assigned in the table.

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 a file**: 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.