Protein Ligand Analyzer Guide

Basic Information

- Analysis is started by giving the necessary inputs then pressing “Analyze Data”
- Errors will pop up if the input has the wrong type. (rational instead of whole numbers, etc.) The user can close them and must change the corresponding input.
- Black messages in this text field explain the steps of the analysis as they happen.
- The analysis is going to take some time, increasing with the number of scans and different mass-over-charge values.

Standard Input

Most of the input fields already contain default values; these can of course be changed by entering different values.

Upload	
Upload Mass Spectrometry Data:	Select the .ms1 file that should be analyzed. See input information for more details.
Input	
Protein m/z	Specify the mass-over-charge value of the protein to use for the analysis.
Protein charge state	Define the charge state of the protein for accurate filtering and analysis.
Energy function protein	Choose the energy function to read from for the protein. Mostly “Low” = 1 and “High” = 2
Energy function ligands	Choose the energy function to read from for the ligands. Mostly “Low” = 1 and “High” = 2
Similarity settings	
DTW threshold	Set the similarity threshold for curve comparison using Dynamic Time Warping (DTW). The threshold is between 0 and positive infinity, the DTW value having to be below the threshold to count as similar.
Pearson threshold	Determine the minimum Pearson correlation coefficient for similarity checks. The threshold is between -1.0 and 1.0 (0 to 1 can be interpreted as a percentage correlation), the value having to be above the threshold to count as similar.
Different range Input	
Protein sampling range	Specify the range around the protein's mass-over-charge value to average these intensity curves together.
Ligand sampling range	Specify the range around the (possible) ligands’ mass-over-charge values to average these intensity curves together.
Analyzing range settings	
Analysis x-start	Set the starting scan number for analysis and visualization.
Analysis x-end	Set the final scan number for analysis and visualization. This value is automatically set to the maximum after the .ms1 file is selected. It can be changed afterwards.

Mode	
Untargeted:	Analyze all ligand (and protein) intensity curves in the mass spectrometry data without prior selection. The mode specific inputs are explained later.
Targeted:	Focus analysis on specific ligands (mass-over-charge values) selected by the user. The mode specific inputs are explained later.
Output settings	
Result Folder	Select the folder where the output PDF and the .csv files will be saved.
Normalization Mode	How graphs are normalized: Mode 1 – no normalization Mode 2 – ligand curves are individually normalized with their own maximum intensity value Mode 3 – ligand curves are normalized with one maximum intensity value across all ligands
Protein & Ligand graphs in:	One plot: both graphs are in a single plot with different colors. Two plots: each graph has one plot. Every page except the first has the protein graph and then a ligand graph underneath.
Creation of .csv files?	NO: the data of the matching ligands is not saved as .csv files. YES: the data of each matching ligand is saved in one .csv file.
Import/Export Settings	
Import settings	Choose text file to be treated as settings for the analysis. The easiest way is to take a text file produced by the program's "Export settings" button.
Export settings	Choose a directory to save the settings as a text file. This happens immediately after choosing a folder.

Mode: Targeted

In this mode only the selected ligands will be considered.

Select ligands	Choose ligands for targeted analysis by uploading a list. See input information for more details.

Mode: Untargeted

In this mode all ligands will be considered.

Charge exclusion range	Exclude specific charge states of the protein to eliminate false positives. Enter 0 if no other charge state of the protein should be considered. If the range is odd, the higher value is excluded Example: range = 5 and protein charge state 5 than 3,4,6,7, 8 are excluded.
Start m/z	Starting mass-over-charge value for the analysis. This value is automatically set after the .ms1 file is selected. It can be changed after this selection.
End m/z	Final mass-over-charge value for the analysis. This value is automatically set after the .ms1 file is selected. It can be changed after this selection.
Ligand m/z grouping range	Define the grouping range for ligands to group them by their mass-over-charge values. Out of each group the curve with the highest similarity is considered for further analysis. Example: range = 2 and mass-over-charge values = 499, 501, 502 Group 1 = [499, 501], Group 2 = [502]
Protein exclusion window	Specify a window to exclude mass-over-charge values near the protein to eliminate false positives.

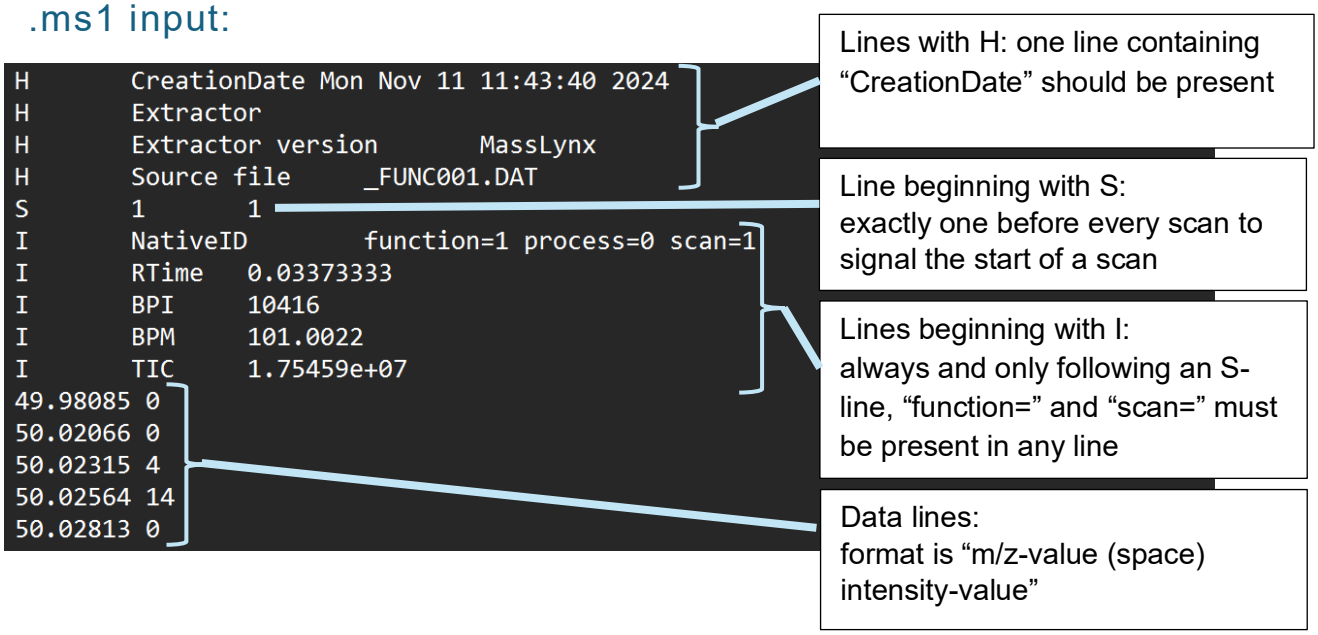
Advanced settings:

- These settings can be accessed from the main screen with “Advanced Settings”
- These settings should only be changed by experienced users.

Inputs	
Protein Charge state sum range	Through this mechanism the intensity values of the different charge states of the protein are summed. This new curve is used for analysis. Example: charge state 6 range 3 Sum of charge state= 5, 6, 7, 8
Savitzky-Golay filter window length	Specifies the number of data points used in smoothing to reduce noise while preserving signal shape.
Savitzky-Golay filter polyorder	Determines the degree of the polynomial used for fitting data within the specified window to smooth the signal.
Num of parse processes	Set the number of processes for parallel file parsing to improve performance. If greater than 1, it can get very RAM-intensive.
Num of analysis processes	Specify the number of processes for parallel data analysis tasks.
Max Cache Size	Define the maximum allowable cache size for storing intermediate data.
Cache Use	Enable (YES) or disable (NO) the use of a cache to optimize runtime efficiency on multiple analyses of the same scan set. The cache data and further runtime data are saved in “ProgramData/DAMS”.
Buttons	
Import cache	Choose zip archive to be treated as a cache for the program. The easiest way is to take a zip archive produced by the program’s “Export cache” button.
Export cache	Choose a directory to save the cache as a zip archive. This happens immediately after choosing a folder.
Delete cache	Empty the program’s cache.

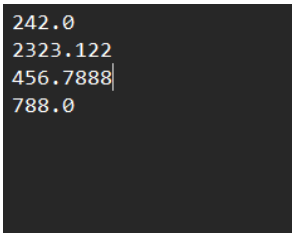
Input information:

.ms1 input:



.txt ligands input:

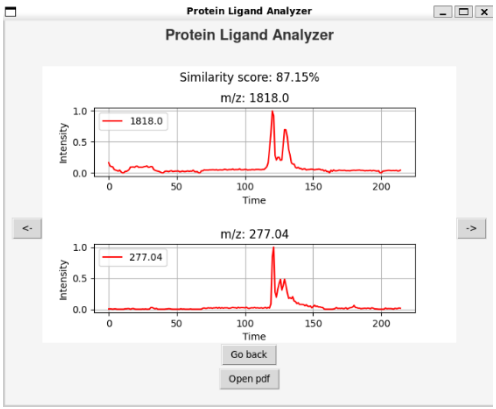
- Enter every ligand that should be considered in a new line.
- They need to be floats (with '.' as a delimiter)



Output information:

GUI output:

- The first plot is always the protein and the plot below one of the matching/selected ligands. Both come with their mass-over-charge value as identification.
- The similarity score displays the percentage similarity between the protein and the ligand.
- The buttons “->” and “<-“ iterate through the matching/selected ligands.
- The button “Go back” closes the result area and returns the app to the main interface.
- The button “Open pdf” opens the PDF result file.

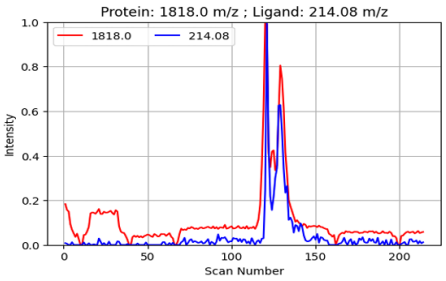


File output:

- The folder for these files is created in the selected folder (see above). If there is no folder selected, it is saved on the desktop.
- The name contains analysis date and time but also used scan date and time.
- There are two output types saved in the folder.

PDF:

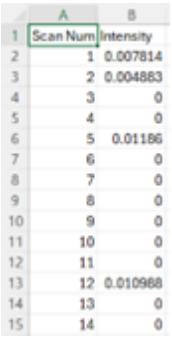
- The first page contains metadata about analysis and scan time.
- Contains plots like the GUI output with two modes (protein and ligand in a single plot or in two separate ones).



Intensity graph comparing protein and ligand 1

.csv (optional):

- The name of each csv file corresponds to the pdf ligand numbering and contains the mass-over-charge value.
- The files have two columns consisting of scan numbers and intensities, the values corresponding to the plots in the graphical output.



	A	B
1	Scan Num	Intensity
2	1	0.007814
3	2	0.004883
4	3	0
5	4	0
6	5	0.01186
7	6	0
8	7	0
9	8	0
10	9	0
11	10	0
12	11	0
13	12	0.010968
14	13	0
15	14	0

(Note: Certain machines use different separators in Excel for csv files, resulting in both columns being merged into one. Should this occur, research changing the “csv delimiter/separator” to ‘;’ for your application)